Subtracting the kinetic Sunyaev-Zeldovich effect from the cosmic microwave background with surveys of large-scale structure

Simon Foreman, Selim C. Hotinli, Mathew S. Madhavacheril, Alexander van Engelen, and Christina D. Kreisch

Subtracting the kinetic Sunyaev-Zeldovich effect from the cosmic microwave background with surveys of large-scale structure

Simon Foreman,[1, 2, 3] Selim C. Hotinli,[4] Mathew S. Madhavacheril,[1, 5]

Alexander van Engelen,[6] and Christina D. Kreisch[7]

[1]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada*

[2]*Dominion Radio Astrophysical Observatory, Herzberg Astronomy & Astrophysics Research Centre,*
National Research Council Canada, P.O. Box 248, Penticton, BC V2A 6J9, Canada

[3]*MIT Kavli Institute for Astrophysics and Space Research,*
Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA

[4]*William H. Miller III Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*

[5]*Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA*

[6]*School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287, USA*

[7]*Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA*

(Dated: February 16, 2023)

The kinetic Sunyaev-Zeldovich (kSZ) effect will be an important source of cosmological and astrophysical information in upcoming surveys of the cosmic microwave background (CMB). However, the kSZ effect will also act as the dominant source of noise for several other measurements that use small angular scales in CMB temperature maps, since its blackbody nature implies that standard component separation techniques cannot be used to remove it from observed maps. In this paper, we explore the idea of "de-kSZing": constructing a template for the late-time kSZ effect using external surveys of large-scale structure, and then subtracting this template from CMB temperature maps in order to remove some portion of the kSZ signal. After building intuition for general aspects of the de-kSZing procedure, we perform forecasts for the de-kSZing efficiency of several large-scale structure surveys, including BOSS, DESI, Roman, MegaMapper, and PUMA. We also highlight potential applications of de-kSZing to cosmological constraints from the CMB temperature power spectrum, CMB lensing reconstruction, and the moving-lens effect. While our forecasts predict achievable de-kSZing efficiencies of 10-20% at best, these results are specific to the de-kSZing formalism adopted in this work, and we expect that higher efficiencies are possible using improved versions of this formalism.

I. INTRODUCTION

The "primary" fluctuations of the cosmic microwave background (CMB), which encode the properties of the universe during the recombination era, have played a key role in the establishment of our current cosmological model. "Secondary" fluctuations, created by processes occurring after recombination, are now attracting significant theoretical and observational interest, as they can in principle be used to probe numerous aspects of astrophysics and fundamental physics.

In particular, much recent interest has been focused on the various incarnations of the Sunyaev-Zeldovich (SZ) effect ([1–4]; see Ref. [5] for a recent review), by which CMB photons scatter off of free electrons associated with massive dark matter halos. The thermal SZ (tSZ) effect, generated by thermal motions of electrons within a halo, distorts the CMB blackbody spectrum and probes the pressure of those electrons. The kinetic SZ (kSZ) effect, generated by bulk motion of the electrons, leads to additional small-scale CMB anisotropies and is sensitive to a combination of large-scale velocity flows and the spatial distribution of electrons. The polarized SZ (pSZ) effect, generated by a quadrupole in the radiation field observed by the scatterer (and therefore sensitive to the various effects that can contribute to that quadrupole), induces linear polarization in the scattered photons. Higher-order

corrections to these effects also have distinctive signatures and a variety of applications (e.g. [6–8]).

These effects are interesting signals in many contexts, but in other contexts, they represent sources of bias and/or noise that degrade our ability to make useful measurements. A significant portion of the tSZ effect can be removed from observed CMB maps thanks to its non-blackbody spectral shape, but the (blackbody) kSZ effect cannot be removed in this way, and is expected to dominate the observed temperature fluctuations at $\ell \gtrsim 4000$ after non-blackbody signals have been filtered out. For sufficiently high-resolution observations, reconstruction of gravitational lensing of the CMB will make use of temperature fluctuations at these scales, and the kSZ effect can significantly bias this reconstruction [9, 10], affecting its ability to constrain neutrino mass, dark energy, and dark matter (e.g. [11, 12]). Furthermore, kSZ fluctuations at small scales add noise to measurements of the Silk-damped regime of the primary CMB, which can be used to probe the effective number of free-streaming species $N_{\rm eff}$, the primordial helium abundance $Y_{\rm p}$, and the spectrum of primordial scalar perturbations. The kSZ effect also acts as noise for measurements of other CMB secondaries, such as the moving-lens effect [13–15].

This situation is analogous to gravitational lensing of the CMB, which is both a physically interesting signal and a source of bias and noise for other measurements (including primordial gravitational waves, acoustic os-

cillations in the primary CMB, and lensing itself). In the lensing case, methods to "de-lens" CMB temperature and polarization maps have been extensively investigated [16–20], applied to recent CMB observations [21–25], and integrated into the core analysis pipelines that are being developed for upcoming surveys [26]. "Internal" de-lensing uses lensing maps reconstructed from a CMB survey to remove lensing effects from temperature and polarization maps from the same survey, while "external" de-lensing uses an external dataset to estimate a lensing map that is used to de-lens the observed CMB.

In this work, we investigate the prospects for external "de-kSZing" of CMB temperature maps, whereby an external galaxy survey is used to construct a template for the kSZ effect, which is then subtracted from the observed maps. Specifically, we consider the portion of the kSZ effect sourced by the post-reionization universe (the "late-time" kSZ effect), since this is expected to either dominate over the kSZ signal from reionization or be comparable to it, depending on the models that are assumed [10]. Following Refs. [27, 28][1], we construct a template by using a galaxy survey to reconstruct the large-scale velocity field, and multiplying by a linear reconstruction of the electron density field; this forms a template for the electron momentum field, which can then be projected along the line of sight to yield an estimate of the kSZ-induced CMB temperature fluctuations. While Refs. [27, 28] explored the prospect of cross-correlating this template with the CMB in order to measure the kSZ effect for probing baryons in the warm-hot intergalactic medium, our motivation here is to use the template to remove some portion of the kSZ effect from CMB observations.

The bulk of our paper is concerned with assessing the performance of kSZ templates constructed from galaxy surveys with different properties, and performing an initial exploration of several possible applications of de-kSZing with external data: namely, constraining cosmological parameters using the CMB damping tail, improving the reconstruction of gravitational lensing, and measuring halo profiles using the moving-lens effect. While our overall conclusion is that it will be challenging to obtain significant benefits from de-kSZing within the formalism adopted in this work, it's likely that modifications to this formalism could yield substantial improvements, and we discuss this further in the final section of the paper.

Note that we will use the term "kSZ effect" to refer only to the *late-time* kSZ effect in the body of the paper; we will briefly discuss the implications of de-kSZing the late-time kSZ effect on the measurement of the kSZ signal from reionization at the end of Sec. V A.

This paper is organized as follows. In Sec. II, we review the main theoretical expressions used to describe the kSZ effect, as well as our formalism for constructing kSZ templates from galaxy surveys and how we model the relevant quantities. In Sec. III, we build intuition for the ability of different kSZ templates to capture a given fraction of the true kSZ signal, by enumerating the relevant redshifts and halo masses (III A), the separate impacts of the velocity and electron density templates (III B), and the influence of redshift uncertainties in the input galaxy survey (III D). In Sec. IV, we forecast the usefulness of specific galaxy surveys for de-kSZing, including BOSS, DESI, the Roman Space Telescope, the MegaMapper proposal, and the PUMA 21 cm intensity mapping proposal. In Sec. V, we discuss the potential of de-kSZing to improve measurements of cosmological parameters, CMB lensing, and the moving lens effect. Finally, we discuss plausible improved versions of de-kSZing and conclude in Sec. VI.

The appendices discuss the impact of shot noise in the galaxy survey used to construct the template (Appendix A), the details of our halo model approach to modelling the signal and templates (Appendix B), the halo occupation distribution models we assume in our computations (Appendix C), our approach to modelling 21 cm intensity mapping surveys (Appendix D), and further details of our CMB forecasts (Appendices E-F).

II. THEORY

A. Review of the kSZ effect

The kSZ effect adds the following contribution to the observed CMB temperature at sky location $\hat{\boldsymbol{n}}$ (e.g. [30]):

$$T_{\text{kSZ}}(\hat{\boldsymbol{n}}) = \int_0^{\chi_*} d\chi \, \tilde{K}(z[\chi]) \, q_{\text{r}}(\hat{\boldsymbol{n}}, \chi; z[\chi]) \,. \qquad (1)$$

In this expression, χ is comoving distance, χ_* corresponds to the earliest epoch we wish to include in our calculations of the kSZ effect, q_{r} is the radial (i.e. line-of-sight) component of the momentum of the free electrons at location $(\hat{\boldsymbol{n}}, \chi)$ and redshift $z[\chi]$, and $\tilde{K}(z)$ is a radial weight function that captures the scattering of CMB photons by these electrons. In this work, we are only concerned with the post-reionization kSZ effect, so we take χ_* to be the comoving distance to the end of reionization, at $z \approx 6$.

The electron momentum is well approximated by

$$q_{\text{r}}(\boldsymbol{x}; z) \approx v_{\text{r}}(\boldsymbol{x}; z)\delta_{\text{e}}(\boldsymbol{x}; z) \,, \qquad (2)$$

where v_{r} is the radial component of the velocity field and δ_{e} is the electron density contrast. At the large scales where the velocity field has the majority of its power, it can be treated as curl-free, such that

$$v_{\text{r}}(\boldsymbol{k}; z) = i\mu v(\boldsymbol{k}; z) \,, \qquad (3)$$

where μ is the cosine of the angle between $\boldsymbol{k}$ and the line

---

[1] See also Ref. [29] for a fully 3-dimensional formalism for kSZ templates.

of sight. Furthermore, at large scales, v is linearly related to the matter density contrast δ_m via

$$\delta_\mathrm{m}(\boldsymbol{k}) = \frac{k}{faH}v(\boldsymbol{k}) \ , \tag{4}$$

omitting the redshift arguments for brevity. Note that the electron momentum is given in full by $q_\mathrm{r} = v_\mathrm{r} + v_\mathrm{r}\delta_\mathrm{e}$, but the v_r term is far subdominant on nonlinear scales (where $\delta_\mathrm{e} \gg 1$) [30]. The radial weight function is

$$\tilde{K}(z) \equiv -T_\mathrm{CMB}\,\bar{n}_{\mathrm{e},0}\,\sigma_\mathrm{T}\,(1+z)^2 e^{-\tau(z)} \ , \tag{5}$$

where T_CMB is the mean CMB temperature, $\bar{n}_{\mathrm{e},0}$ is the mean electron number density at redshift zero, σ_T is the cross section for Thomson scattering, and $\tau(z)$ is the spatially-averaged optical depth to redshift z. We take $\tau(z) \approx 0$ after reionization.

In the Limber approximation [31, 32], the angular power spectrum of T_kSZ evaluates to

$$C_\ell^\mathrm{kSZ} = \int_0^{\chi_*} \frac{d\chi}{\chi^2}\tilde{K}(z[\chi])^2\,P_{q_\mathrm{r}q_\mathrm{r}}\left(\frac{\ell+1/2}{\chi};z[\chi]\right) \ , \tag{6}$$

where $P_{q_\mathrm{r}q_\mathrm{r}}$ is the 3d power spectrum of q_r, given by

$$P_{q_\mathrm{r}q_\mathrm{r}}(k_\mathrm{S};z) \approx \int \frac{d^3\boldsymbol{k}_\mathrm{L}}{(2\pi)^3}P_{v_\mathrm{r}v_\mathrm{r}}(k_\mathrm{L};z)P_\mathrm{ee}(k_\mathrm{S};z) \tag{7}$$

$$= \frac{1}{6\pi^2}\left(\int dk_\mathrm{L}\,k_\mathrm{L}^2 P_{vv}(k_\mathrm{L};z)\right)P_\mathrm{ee}(k_\mathrm{S};z) \ . \tag{8}$$

In writing Eq. (7), we have neglected the connected four-point function $\langle v_\mathrm{r}\delta_\mathrm{e}v_\mathrm{r}\delta_\mathrm{e}\rangle$, and only retained the dominant "squeezed" contribution to the power spectrum, arising from the large-scale (k_L) velocity power and small-scale (k_S) electron density power. The accuracy of these approximations has been argued both analytically [30, 33] and using simulations [30, 34, 35].[2]

---

[2] Often in the literature, the squeezed limit ($k_\mathrm{L} \ll k_\mathrm{S}$) is not taken when writing Eq. (7), but k_L is still assumed to be in the linear regime, in which case Eq. (7) becomes

$$P_{q_\mathrm{r}q_\mathrm{r}}(k_\mathrm{S}) \approx (faH)^2 \int \frac{d^3\boldsymbol{k}_\mathrm{L}}{(2\pi)^3}P_\mathrm{mm}(k_\mathrm{L})P_\mathrm{ee}(|\boldsymbol{k}_\mathrm{S}-\boldsymbol{k}_\mathrm{L}|)$$
$$\times \frac{k_\mathrm{S}(k_\mathrm{S}-2k_\mathrm{L}\mu_\mathrm{L})(1-\mu_\mathrm{L}^2)}{k_\mathrm{L}^2(k_\mathrm{S}^2+k_\mathrm{L}^2-2k_\mathrm{S}k_\mathrm{L}\mu_\mathrm{L})} \ . \tag{9}$$

If Eq. (9) is used when computing C_ℓ^kSZ instead of Eq. (7), the results differ by 30% at $\ell = 1000$ and by less than 5% at $\ell > 4000$. We use Eq. (7) in our computations, to enable us to compare the results with the templates described in Sec. II B, which also assume the squeezed limit.

B. Constructing templates from galaxy surveys

Given a galaxy density contrast δ_g^s observed (in redshift space, hence the superscript "s") by a galaxy survey, one can construct a template for the expected kSZ contribution to the observed CMB temperature.[3] First, one forms the following estimates for v_r and δ_e from δ_g^s:

$$\eta(\boldsymbol{k};z) \equiv i\mu\frac{P_{gv}(k,\mu;z)}{P_\mathrm{gg}^\mathrm{tot}(k,\mu;z)}\delta_\mathrm{g}^s(\boldsymbol{k};z) \ , \tag{10}$$

$$\epsilon(\boldsymbol{k};z) \equiv \frac{P_\mathrm{ge}(k,\mu;z)}{P_\mathrm{gg}^\mathrm{tot}(k,\mu;z)}\delta_\mathrm{g}^s(\boldsymbol{k};z) \ . \tag{11}$$

The redshift-space galaxy power spectrum $P_\mathrm{gg}^\mathrm{tot}(k,\mu;z)$ is defined to include the effect of shot noise (assumed to be Poissonian in this work),

$$P_\mathrm{gg}^\mathrm{tot}(k,\mu;z) \equiv P_\mathrm{gg}(k,\mu;z) + \frac{1}{\bar{n}_\mathrm{g}(z)} \ , \tag{12}$$

such that η and ϵ are Wiener-filtered quantities that downweight noise-dominated modes of δ_g^s. We will discuss our approach to modelling the power spectra needed for Eqs. (10)-(11) in Sec. II C.

One then forms an estimate for the line-of-sight electron momentum field,

$$\hat{q}_\mathrm{r}(\boldsymbol{x},z) = \eta(\boldsymbol{x},z)\epsilon(\boldsymbol{x},z) \ , \tag{13}$$

and computes a line-of-sight projection using the kSZ radial weight function, analogous with Eq. (1):

$$\hat{T}_\mathrm{kSZ}(\hat{\boldsymbol{n}}) = \int_0^{\chi_*} d\chi\,\tilde{K}(z[\chi])\,\hat{q}_\mathrm{r}(\hat{\boldsymbol{n}},\chi;z[\chi]) \ . \tag{14}$$

The cross-correlation between the template $\hat{T}_\mathrm{kSZ}$ and the true signal T_kSZ will depend on the 3d cross power spectrum between $\hat{q}_\mathrm{r}$ and q_r. Under the same assumptions applied to Eq. (7), this evaluates to

$$P_{\hat{q}_\mathrm{r}q_\mathrm{r}}(k_\mathrm{S},\mu_\mathrm{S};z)$$
$$\approx \int \frac{d^3\boldsymbol{k}_\mathrm{L}}{(2\pi)^3}P_{v_\mathrm{r}\eta}(k_\mathrm{L},\mu_\mathrm{L};z)P_{e\epsilon}(k_\mathrm{S},\mu_\mathrm{S};z) \tag{15}$$
$$= \frac{1}{4\pi^2}\left(\int dk_\mathrm{L}\,k_\mathrm{L}^2\int_{-1}^1 d\mu_\mathrm{L}\,\mu_\mathrm{L}^2\frac{P_{gv}(k_\mathrm{L},\mu_\mathrm{L};z)^2}{P_\mathrm{gg}^\mathrm{tot}(k_\mathrm{L},\mu_\mathrm{L};z)}\right)$$
$$\times \frac{P_\mathrm{ge}(k_\mathrm{S},\mu_\mathrm{S};z)^2}{P_\mathrm{gg}^\mathrm{tot}(k_\mathrm{S},\mu_\mathrm{S};z)} \ . \tag{16}$$

---

[3] Similar formalisms for kSZ templates were previously presented in Refs. [27, 28, 33]. Ref. [27] computed P_ge and P_gg using a linear bias model and Ref. [28] effectively assumed that $P_\mathrm{ge} = 1$, while we use a more detailed halo model approach based on Ref. [33], described in Sec. II C and Appendix B. Also, in contrast with Refs. [27, 33], we account for linear redshift-space distortions in the galaxy density when constructing our electron-density and velocity templates.

Redshift-space distortions in the observed galaxy density contrast δ_{g}^s imply that $P_{v_r\eta}$ and $P_{e\epsilon}$ both depend on the angle with respect to the line of sight, through μ_{L} and μ_{S} respectively. On the large scales that dominate the velocity contribution, redshift-distortions are approximately described by the Kaiser factor $\delta_{\mathrm{g}}^s(k,\mu) \propto (b_{\mathrm{g}} + f\mu^2)$ [36]. On the other hand, at the scales relevant for the electron density contribution, there will be "Finger of God" damping of large-μ_{S} power caused by velocity dispersions on small scales [37]. However, the angular power spectrum of the template is a line-of-sight projection weighted by the kSZ weight function from Eq. (5), and the width and smoothness of this weight function will largely suppress the influence of large-μ_{S} modes on the final result (e.g. [38, 39]). This, it will not be necessary to include the Finger of God effect in our computations. We may then write the angular power spectrum derived from the $\hat{q}_{\mathrm{r}}$-q_{r} cross-correlation as

$$C_\ell^{\hat{q}_{\mathrm{r}} q_{\mathrm{r}}} = \int_0^{\chi_*} \frac{d\chi}{\chi^2} \tilde{K}(z[\chi])^2 \, P_{\hat{q}_{\mathrm{r}} q_{\mathrm{r}}}\left(\frac{\ell+1/2}{\chi}, 0; z[\chi]\right) \,, \tag{17}$$

following the Limber approximation in setting $\mu_{\mathrm{S}} = 0$ [32].

A short calculation shows that the auto spectrum of the template $\hat{T}_{\mathrm{kSZ}}$ is also given by Eqs. (16)-(17). However, in this case we must revisit the approximation that the connected four-point function can be neglected: the power spectrum of $\hat{q}_{\mathrm{r}}$ is related to the trispectrum of δ_{g}^s, which has several contributions from the shot noise of the corresponding galaxy sample. We can write this as

$$C_\ell^{\hat{q}_{\mathrm{r}} \hat{q}_{\mathrm{r}}} = C_\ell^{\hat{q}_{\mathrm{r}} q_{\mathrm{r}}} + C_\ell^{\mathrm{shot}} \,, \tag{18}$$

with C_ℓ^{shot} including all four-point shot noise terms. In the approximation of Poissonian shot noise, these contributions depend on the mean number density of the sample, along with the galaxy power spectrum and bispectrum (e.g. [40, 41]). In Appendix A, we estimate the dominant such contributions, finding that the added power to a kSZ template can be as high as several tens of percents of the reconstructed kSZ power for a few of the surveys we consider. We do not include these terms in the computations in the body of the paper, but note that they should be accounted for in detailed de-kSZing analyses, and we refer the reader to Appendix A for more details.

C. Modelling

For modelling the various auto and cross power spectra defined in the previous two subsections, we follow the halo model approach of Ref. [33]. We briefly summarize this approach here, with further details contained in Appendix B. We use the public `hmvec` code for our computations.[4]

In the halo model, all matter is assumed to belong to bound, spherical halos. Correlation functions of matter may be computed once several ingredients are specified: the halo mass function, mass-dependent halo bias, and mass-dependent halo density profile. Correlation functions of galaxies additionally require models for how central and satellite galaxies occupy halos of a given mass, and the average spatial distribution of satellite galaxies with in a halo; similarly, predictions involving the electron number density require a model for the electron density profile within halos. Two-point correlations are composed of a "two-halo term," describing correlations between points located in different halos, and a "one-halo term," corresponding to two points located within the same halo.

In our calculations, we use the Sheth-Tormen halo mass function and halo bias [42], and a Navarro-Frenk-White (NFW) density profile [43] with the concentration-mass relation from Ref. [44]. Electron density power spectra are computed by replacing the NFW profile with an electron profile model based on hydrodynamical simulations from Ref. [45]; in particular, we use the model including AGN feedback as a baseline. Following Ref. [46], we truncate the electron profile at a maximum radius chosen so that the enclosed gas mass is the same as for a NFW profile truncated at $r_{\mathrm{cut}} = r_{\mathrm{200c}}$. To describe galaxy clustering, we use the halo occupation distribution (HOD) formalism (e.g. [47]), with specific HOD models introduced when they are used later in the paper.

Finally, we use linear expressions for power spectra involving velocities, since only the large-scale velocity power will enter our results:

$$P_{vv}(k) = \left(\frac{faH}{k}\right)^2 P_{\mathrm{m}}(k) \,, \tag{19}$$

$$P_{gv}(k,\mu) = \frac{faH}{k}(b_{\mathrm{g}} + f\mu^2)P_{\mathrm{m}}(k) \,, \tag{20}$$

where b_{g} is computed within the halo model and the z-dependence is left implicit.

III. INTUITION

Our ability to remove the late-time kSZ signal from observed CMB maps will clearly depend on how much of the true kSZ signal is captured by a given template. In this section, we build intuition for the relationship between a template's properties and its performance in a de-kSZing procedure.
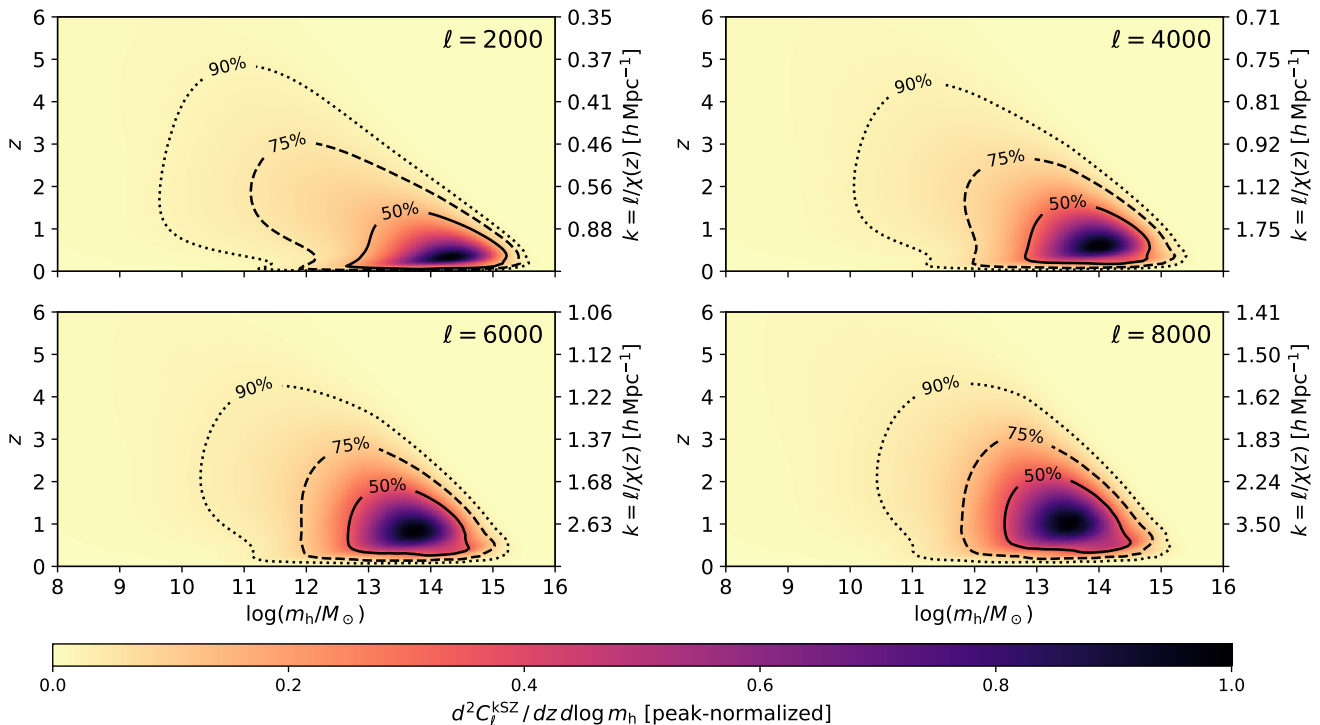
---

[4] https://github.com/simonsobs/hmvec

FIG. 1. Relative contribution to the kSZ angular power spectrum C_ℓ^{kSZ} from different redshifts and halo masses, with each panel corresponding to the indicated ℓ value. The right-hand axis of each plot denotes the comoving wavenumbers that are relevant at a given redshift in the Limber approximation ($k \approx \ell/\chi[z]$). The color scale of each panel is normalized to have a maximum value of unity. The contours enclose regions than contribute 50%, 75%, and 90% of the value of C_ℓ^{kSZ} at the given ℓ. We find that 50% the kSZ power generally comes from $z \lesssim 2$ and $12.5 \lesssim \log_{10}(m_{\mathrm{h}}/M_\odot) \lesssim 15$, implying that galaxy surveys focusing on these redshifts and halo masses are best suited for constructing kSZ templates.

A. Relevant redshifts and halo masses

We begin with Fig. 1, which visualizes the relative contribution to the kSZ auto power from different redshifts and halo masses. Specifically, we rewrite Eq. (6), Eq. (8), and the halo model expression for P_{ee} as

$$C_\ell^{\mathrm{kSZ}} = \int_0^{z_*} dz \int d\log m_{\mathrm{h}} \frac{d^2 C_\ell^{\mathrm{kSZ}}}{dz\, d\log m_{\mathrm{h}}} \qquad (21)$$

and plot the integrand of this expression, normalized to unity at its peak, for several representative ℓ values. (See Appendix B 3 for the details of how this is computed.) We show contours enclosing regions of the $z - m_{\mathrm{h}}$ plane that contribute 50%, 75%, and 90% of the total signal. We also translate the redshift axis into the corresponding wavenumbers probed in the Limber approximation ($k \approx \ell/\chi(z)$).

We find that 50% of the signal is generally localized to $z \lesssim 2$ and $12.5 \lesssim \log_{10}(m_{\mathrm{h}}/M_\odot) \lesssim 15$, while 75% of the signal arises from $z \lesssim 3$. The overall tilt of the contours stems from the fact that there are fewer higher-mass halos at higher redshift, so the contribution to C_ℓ^{kSZ} from higher redshifts is naturally concentrated at lower halo masses.

The ranges of redshift and halo mass that dominate the kSZ signal depend strongly on the halo mass function. The electron-momentum power spectrum $P_{q_{\mathrm{r}} q_{\mathrm{r}}}$ is proportional to the electron density power spectrum P_{ee} (recall Eq. 8); in the regime where this is dominated by the one-halo term, if we approximate $u_{\mathrm{gas}}(k, m_{\mathrm{h}}, z) \approx 1$, Eq. (B10) shows that $P_{\mathrm{ee}} \propto \int d\log m_{\mathrm{h}}\, m_{\mathrm{h}}^3 n(m_{\mathrm{h}}, z)$. The integrand of this expression peaks in the range $13.5 < \log_{10}(m_{\mathrm{h}}/M_\odot) < 14.5$ for $0.5 < z < 2$, with width $\Delta \log_{10}(m_{\mathrm{h}}/M_\odot)$ of order unity; furthermore, this integral declines sharply with redshift, with roughly a factor of 40 difference between $z = 1$ and $z = 3$. The assumed electron density profile also has an important effect: in the model from Ref. [45], the profile is strongly suppressed at $r \gtrsim 2R_{200}(m_{\mathrm{h}}, z)$, implying that the one-halo kSZ power from a halo with mass m_{h} will be suppressed at wavenumbers k satisfying $k \gtrsim 2/R_{200}(m_{\mathrm{h}}, z)$.[5] Thus, there is an effective (redshift-dependent) maximum

_____

[5] To see this mathematically, recall that the one-halo term depends on the Fourier transform of the density profile, $u_{\mathrm{gas}}(k) \propto \int dr\, r^2 \rho_{\mathrm{gas}}(k) \sin(kr)/kr$, and for a top-hat density profile with maximum radius r_{max}, this evaluates to $u_{\mathrm{gas}}(k) \propto k^{-3} j_1(kr_{\mathrm{max}})$, which has its first zero at $k \approx 4.5 r_{\mathrm{max}}^{-1}$ and is strongly suppressed at higher k.

k to which a mass-m_{h} halo will contribute significant kSZ power, and since R_{200} grows with halo mass, this maximum k will be smaller for more massive halos. We can see this in Fig. 1: lower ℓ values probe a higher maximum halo mass. In summary, the redshifts and halo masses that dominate the kSZ signal can mostly be understood in terms of (1) the abundance of halos of different masses at a given redshift, and (2) the spatial extent of the free electrons associated with each halo.

Overall, these results indicate that a galaxy survey (or combination of surveys) with broad coverage in redshift and halo mass will be necessary to construct a high-fidelity kSZ template. For example, a sample of emission-line galaxies (ELGs) would likely miss a significant part of the kSZ signal, since ELGs generally occupy halos with much lower mass (i.e. $m_{\mathrm{h}} \sim 10^{12} M_{\odot}$ [48–51]) than where the bulk of the kSZ contribution comes from. (We will explore specific galaxy samples, including ELGs, later in the paper.) Also, it will likely be challenging to recover more than roughly 80% of the kSZ signal unless one has access to a survey probing very low-mass ($m_{\mathrm{h}} \lesssim 10^{11.5} M_{\odot}$) halos at higher redshifts (such as a 21 cm intensity mapping survey [52, 53]), and is able to combine it with measurements of higher-mass halos.

B. Importance of electron density and velocity templates

1. Electron density template

The accuracies of the electron density template and velocity template both determine the fraction of true kSZ power captured by the kSZ template. We investigate the electron template in this subsection and the velocity template in Sec. III B 2.

To assess the fidelity of the electron template, we compute the squared correlation coefficient between δ_{g}^s and δ_{e}, for purely-transverse ($\mu_{\mathrm{S}} = 0$) modes:

$$r_{\mathrm{ge}}^2(k; z) = \frac{P_{\mathrm{ge}}(k_{\mathrm{S}}, 0; z)^2}{P_{\mathrm{gg}}^{\mathrm{tot}}(k_{\mathrm{S}}, 0; z) P_{\mathrm{ee}}(k_{\mathrm{S}}; z)} . \quad (22)$$

We compute this quantity for a family of fictional galaxy surveys which have the following HOD model:

$$\bar{N}_{\mathrm{c}}(m_{\mathrm{h}}) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\log m_{\mathrm{h}} - \log m_{\mathrm{cut}}}{\sigma_{\log m}}\right)\right] , \quad (23)$$

$$\bar{N}_{\mathrm{s}}(m_{\mathrm{h}}) = \left(\frac{m_{\mathrm{h}} - m_{\mathrm{cut}}}{10^{\beta} m_{\mathrm{cut}}}\right)^{\alpha} . \quad (24)$$

This is a simplified version of the HOD from Ref. [54], with the amplitude of the satellite occupation parameterized relative to the lower mass cutoff m_{cut}, to separate the effects of changing m_{cut} and this amplitude. For this exercise, we fix $\sigma_{\log m} = 0.3$ and $\alpha = 1$, and examine the effects of varying m_{cut} and β. The influence of these
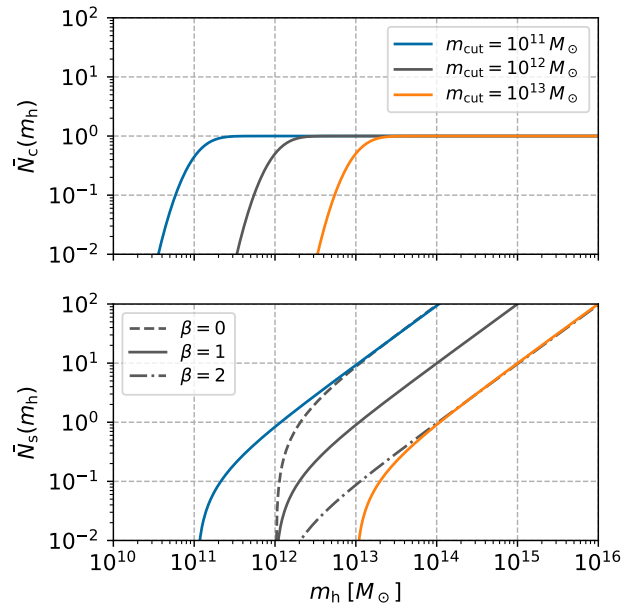

FIG. 2. Mean central occupation (*upper panel*) and satellite occupation (*lower panel*) for the HOD model used to investigate the electron density template in Sec. III B 1. In both panels, different colors denote different values of the lower mass cutoff m_{cut}, while in the lower panel, different linestyles denote different values for β, which controls the relative amplitude of $\bar{N}_{\mathrm{c}}$ and $\bar{N}_{\mathrm{s}}$. This figure can be used in tandem with Fig. 3 to gain intuition for how the properties of the HOD influence the fidelity of the electron density template that goes into a kSZ template.

parameters on the central and satellite occupations is visualized in Fig. 2.

In Fig. 3, we plot r_{ge}^2 as a function of z, at the k values relevant for the Limber approximation ($k \approx \ell/\chi[z]$) at two different ℓ values. We vary β between 0 and 2, noting that $\beta \approx 1$ is a typical value found in real or simulated galaxy samples (e.g. [51, 54–56]). We also vary the galaxy number density $\bar{n}_{\mathrm{g}}$, taken to be redshift-independent, and derive the corresponding (redshift-dependent) lower mass cutoff m_{cut} by solving Eq. (B4). The resulting m_{cut} values are shown in the lower panels of the figure.

The shapes of the curves in Fig. 3 can be interpreted in terms of whether the one-halo or two-halo term dominates. In particular, there are two noteworthy trends:

- At higher redshift, the g-e correlation improves with increasing $\bar{n}_{\mathrm{g}}$ (and decreasing m_{cut}). This is because, in the angular power spectrum, higher z corresponds to lower k, where the two-halo term is more important. At sufficiently low k (i.e. high z and low ℓ), the two-halo term completely dominates and perfect cross-correlation can be achieved, although the (angular) kSZ power will also have some sensitivity to lower z where the correlation is worse. The correlation is noticeably worse for $\beta = 0$ than for $\beta = 1$ or 2, because $\beta = 0$ implies a particularly
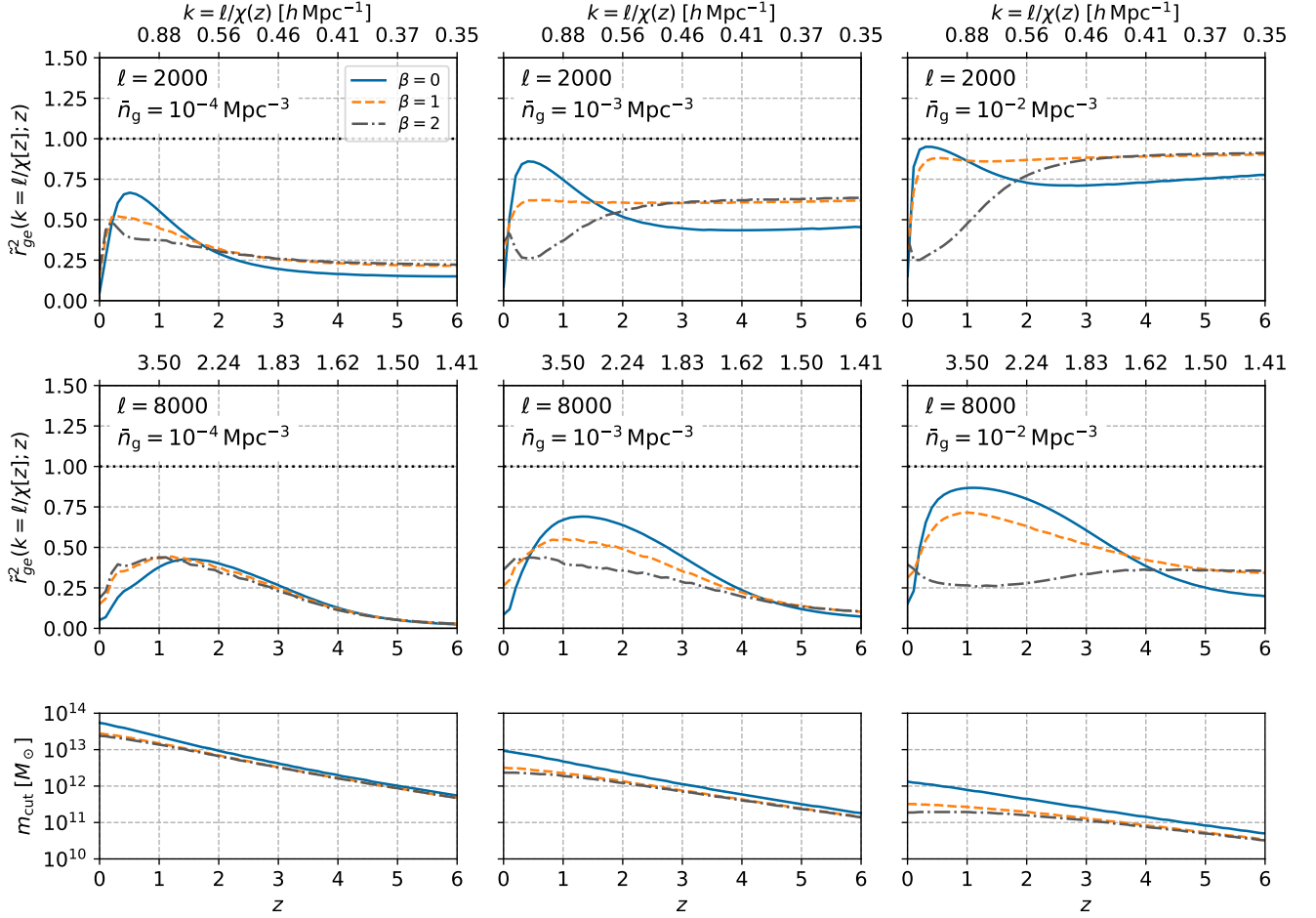
FIG. 3. An exploration of how accurate a template for the electron density δ_e can be constructed by Wiener-filtering the observed galaxy density δ_g^s. We plot the squared correlation coefficient between δ_g^s and δ_e, assuming different parameters in the simplified HOD model from Eqs. (23)-(24). Curves in the *upper* and *middle rows* are evaluated at $k = \ell/\chi[z]$ for $\ell = 2000$ or 8000, while the *lower row* shows the redshift-dependent minimum mass $m_{\rm cut}$ corresponding to the galaxy number density indicated in each column. In both the two-halo-dominated regime (at sufficiently high z and low k) and the one-halo-dominated regime (at low z and high k), a high-fidelity electron template is achievable if the input galaxy survey is sufficiently dense.

large amplitude for the one-halo term, moving the pure-two-halo regime to lower k.

- In the one-halo regime (low z, high k), a high-fidelity electron template is achievable for sufficiently high galaxy number density. In this regime, lower values of β result in a better template, because a higher satellite fraction implies a better measurement of the distribution of satellites within the halo, which will trace the distribution of electrons if the satellite profile and gas profile are similar.

2. Velocity template

Next, we assess the performance of the template for large-scale velocities. We do so by computing the ratio

of the velocity factors in Eq. (16) and (8):

$$R_{gv}(z) \equiv \frac{3}{2} \frac{\int dk_{\rm L}\, k_{\rm L}^2 \int_{-1}^1 d\mu_{\rm L}\, \mu_{\rm L}^2 \frac{P_{gv}(k_{\rm L},\mu_{\rm L};z)^2}{P_{gg}^{\rm tot}(k_{\rm L},\mu_{\rm L};z)}}{\int dk_{\rm L}\, k_{\rm L}^2 P_{vv}(k_{\rm L};z)} . \quad (25)$$

This is analogous to a correlation coefficient between δ_g^s and v, but is more relevant to the kSZ template than a standard correlation coefficient, since it directly reflects the redshift-dependent suppression of kSZ template power due to the properties of the velocity template.

Fig. 4 shows $R_{gv}(z)$ computed using the same HOD model as in Sec. III B 1, showing only the $\beta = 1$ case for simplicity. The inclusion of RSD slightly boosts the power in the velocity template thanks to the Kaiser factor increasing the signal-to-shot-noise ratio in $P_{gg}^{\rm tot}$. For $\bar{n}_g = 10^{-4}\,{\rm Mpc}^{-3}$, the accuracy of the velocity template can degrade the fraction of the kSZ power captured by the kSZ template by tens of percents, but for higher num-
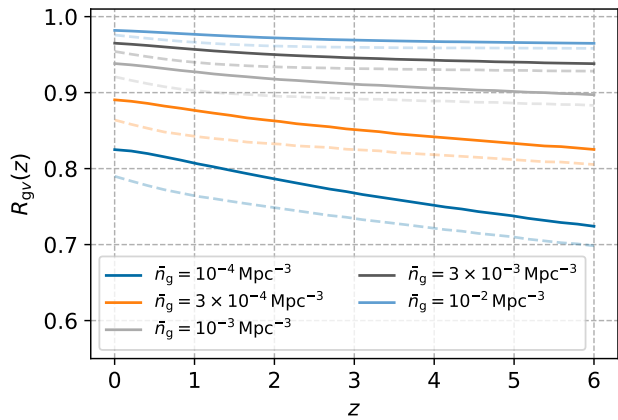
FIG. 4. Correlation-coefficient-like quantity, defined in Eq. (25), for the observed galaxy density and large-scale velocity field, reflecting the impact of the velocity template on the kSZ template. Different curves are computed as in Fig. 3 with $\beta = 1$, with faint lines ignoring the effects of RSD. For lower galaxy number density, the velocity template fidelity can degrade the kSZ template fidelity by a few tens of percents, while for higher number density, the electron template fidelity will mostly determine the fraction of true kSZ power that can be captured by the kSZ template.
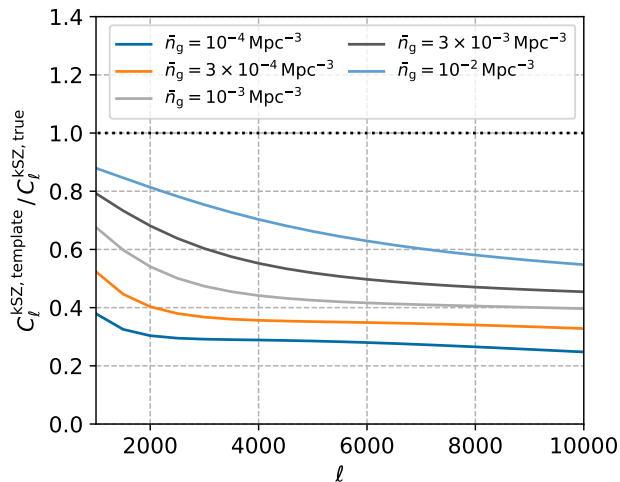
FIG. 5. Ratio of kSZ template power spectrum and true kSZ power spectrum, with the same line colors as Fig. 4. As expected, a template constructed from a denser galaxy survey captures a higher fraction of the true kSZ power, with better performance at lower multipoles that probe larger physical scales.

ber densities, a comparison with Fig. 3 reveals that the electron template will be the dominant factor in the performance of the kSZ template.

3. Combination

Fig. 5 shows the auto power spectrum of the kSZ template constructed using the same HOD as the curves in

Fig. 4, as a ratio to the true kSZ power. As expected, we find better performance at lower multipoles, because these multipoles probe larger physical scales and higher redshifts where the electron template is more accurate. The accuracy decreases at higher multipoles, but can still reach several tens of percents of the true kSZ power if a dense enough galaxy survey is used.

These general investigations provide context for survey-specific forecasts that we will present in Sec. IV.

C. Accuracy of model for galaxy-electron cross power spectrum

The kSZ template in Sec. II B requires input models for $P_{\mathrm{g}v}$, P_{ge}, and $P_{\mathrm{gg}}^{\mathrm{tot}}$. Since we only require $P_{\mathrm{g}v}$ on quasi-linear scales, it will be relatively straightforward to model, while $P_{\mathrm{gg}}^{\mathrm{tot}}$ is directly measurable from the associated galaxy survey. On the other hand, the modelling of P_{ge} is expected to be considerably more uncertain. This uncertainty leads to what is known as the "optical depth degeneracy" in kSZ studies (e.g., [57]) and is a limiting factor, for example, in using the reconstructed velocities from the kSZ effect as a probe of the growth rate of structure. Here, we consider what impact uncertainties in the assumed form of P_{ge} will have on a de-kSZing procedure.

Specifically, we are interested in the power spectrum of the kSZ effect that remains in a temperature map after a template has been subtracted off. This angular power spectrum, of $\hat{T}_{\mathrm{de\text{-}kSZ}} \equiv T_{\mathrm{kSZ}} - \hat{T}_{\mathrm{kSZ}}$, evaluates to

$$C_\ell^{\mathrm{de\text{-}kSZ}} = C_\ell^{\mathrm{kSZ}} - 2C_\ell^{\hat{q}_{\mathrm{r}} q_{\mathrm{r}}} + C_\ell^{\hat{q}_{\mathrm{r}} \hat{q}_{\mathrm{r}}} \ , \qquad (26)$$

where C_ℓ^{kSZ} is given by Eq. (6), $C_\ell^{\hat{q}_{\mathrm{r}} q_{\mathrm{r}}}$ is given by Eq. (17), and $C_\ell^{\hat{q}_{\mathrm{r}} \hat{q}_{\mathrm{r}}}$ is equal to $C_\ell^{\hat{q}_{\mathrm{r}} q_{\mathrm{r}}}$ if the assumed P_{ge} is exactly correct.

We can relax this assumption by writing the assumed spectrum as a sum of the true spectrum and an error term:

$$P_{\mathrm{ge}}^{\mathrm{model}}(k, \mu; z) = P_{\mathrm{ge}}(k, \mu; z) + \Delta P_{\mathrm{ge}}(k, \mu; z) \ . \qquad (27)$$

Repeating the derivations of $P_{\hat{q}_{\mathrm{r}} q_{\mathrm{r}}}$ (recall Eq. 16) and $P_{\hat{q}_{\mathrm{r}} \hat{q}_{\mathrm{r}}}$, we find that using $P_{\mathrm{ge}}^{\mathrm{model}}$ instead of P_{ge} in the construction of the electron density templates will change

each spectrum like so:

$$
\Delta P_{\hat{q}_r q_r}(k_S, \mu_S; z)
$$

$$
= \frac{1}{4\pi^2}\left(\int dk_L\, k_L^2 \int_{-1}^{1} d\mu_L\, \mu_L^2 \frac{P_{gv}(k_L, \mu_L; z)^2}{P_{gg}^{\mathrm{tot}}(k_L, \mu_L; z)}\right)
$$

$$
\times\ \frac{P_{ge}(k_S, \mu_S; z)\Delta P_{ge}(k_S, \mu_S; z)}{P_{gg}^{\mathrm{tot}}(k_S, \mu_S; z)}\ , \qquad (28)
$$

$$
\Delta P_{\hat{q}_r \hat{q}_r}(k_S, \mu_S; z)
$$

$$
= \frac{1}{4\pi^2}\left(\int dk_L\, k_L^2 \int_{-1}^{1} d\mu_L\, \mu_L^2 \frac{P_{gv}(k_L, \mu_L; z)^2}{P_{gg}^{\mathrm{tot}}(k_L, \mu_L; z)}\right)
$$

$$
\times\left[\frac{2P_{ge}(k_S, \mu_S; z)\Delta P_{ge}(k_S, \mu_S; z)}{P_{gg}^{\mathrm{tot}}(k_S, \mu_S; z)}\right.
$$

$$
\left. +\frac{\Delta P_{ge}(k_S, \mu_S; z)^2}{P_{gg}^{\mathrm{tot}}(k_S, \mu_S; z)}\right]\ . \qquad (29)
$$

If we compute $\Delta C_\ell^{\text{de-kSZ}}$ by combining these expressions, the definitions of $C_\ell^{\hat{q}_r q_r}$ and $C_\ell^{\hat{q}_r \hat{q}_r}$, and Eq. (26), we find that the term linear in ΔP_{ge} cancels out, such that the leading effect of ΔP_{ge} on $C_\ell^{\text{de-kSZ}}$ is of order ΔP_{ge}^2. As a concrete example, if $\Delta P_{ge} \approx \alpha P_{ge}$ where α is a constant, then the fractional impact on $C_\ell^{\text{de-kSZ}}$ is

$$
\frac{\Delta C_\ell^{\text{de-kSZ}}}{C_\ell^{\text{de-kSZ}}} \approx \frac{\alpha^2 C_\ell^{\hat{q}_r \hat{q}_r}}{C_\ell^{\text{kSZ}} - C_\ell^{\hat{q}_r \hat{q}_r}}\ . \qquad (30)
$$

This mild dependence on errors in P_{ge}, along with prospects for externally constraining it using e.g. dispersion measures of fast radio bursts [57], is a promising indication that modelling of the galaxy-electron cross power spectrum will not be a serious impediment to template-based de-kSZing.

D. Redshift precision

Thus far, we have assumed that the galaxy survey used for the template has negligible redshift uncertainty, but we can also consider surveys for which this uncertainty may be substantial. Specific examples include photometric surveys such as the Dark Energy Survey [58] and the Vera Rubin Observatory's Legacy Survey of Space and Time (LSST; [59]), or the SPHEREx satellite [60] which is expected to acquire redshifts with a wide range of uncertainties [60].

We will not attempt detailed forecasts for such surveys in this work, but instead present a simple computation that indicates what we might expect. Under the assumption that the uncertainty on each measured redshift is Gaussian, the galaxy power spectrum is multiplied by two powers of a Gaussian kernel

$$
W_{\Delta z}(k, \mu; z) = \exp\left[-\frac{\sigma_z^2 k^2 \mu^2 c^2}{2H(z)^2}\right]\ , \qquad (31)
$$

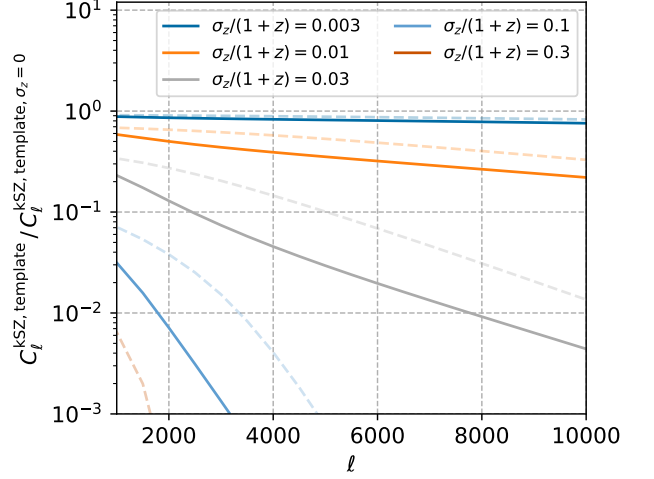while cross spectra between galaxies and another field are

FIG. 6. Impact of galaxy redshift uncertainties on a kSZ template constructed using those galaxies. Each curve is computed with the HOD from Sec. III B 1 with $\beta = 1$ and with the specified Gaussian redshift uncertainty, and has been divided by the corresponding template power spectrum assuming perfectly measured redshifts. *Solid curves* assume $\bar{n}_g = 10^{-4}\,\mathrm{Mpc}^{-3}$, while *dashed curves* assume $\bar{n}_g = 10^{-2}\,\mathrm{Mpc}^{-3}$. For $\sigma_z/(1+z) = 0.03$, which is the target for the Vera Rubin Observatory's LSST, and for which SPHEREx will obtain a galaxy sample with $\bar{n}_g \sim \mathcal{O}(10^{-4})\,\mathrm{Mpc}^{-3}$ at $z \sim 1$, significant suppression of the kSZ template power is expected. This motivates our focus on spectroscopic and 21 cm surveys in this work.

multiplied by a single power (e.g. [33]):

$$
P_{gg}(k, \mu; z) \rightarrow P_{gg}(k, \mu; z)W_{\Delta z}(k, \mu; z)^2\ ,
$$

$$
P_{ge}(k, \mu; z) \rightarrow P_{ge}(k, \mu; z)W_{\Delta z}(k, \mu; z)\ ,
$$

$$
P_{gv}(k, \mu; z) \rightarrow P_{gv}(k, \mu; z)W_{\Delta z}(k, \mu; z)\ . \qquad (32)
$$

Since the redshift kernel for the kSZ effect is quite broad, one might expect that redshift errors would not have a strong effect on a kSZ template, as is the case when constructing templates for the integrated Sachs-Wolfe effect (e.g. [61]), for example. However, this intuition does not hold in this case, because our kSZ template involves a product of separate 3d templates for velocity and electron density, and low-μ modes of this product are affected by higher-μ modes of each component template. In particular, in the squeezed limit we have assumed for the kSZ effect in Eq. (16), the effect of redshift uncertainties on higher-μ modes of the velocity template propagates into the final projected kSZ template.

In Fig. 6, we show how different levels of redshift uncertainty affect $C_\ell^{\text{kSZ,template}}$ computed using the HOD from Sec. III B 1 with $\beta = 1$ and $\bar{n}_g = 10^{-4}\,\mathrm{Mpc}^{-3}$ (solid lines) or $10^{-2}\,\mathrm{Mpc}^{-3}$ (dashed lines). For $\sigma_z/(1+z) = 0.03$, which is the target for LSST [62], the kSZ template power is suppressed at $\ell = 2000$ by 85% (70%) for $\bar{n}_g = 10^{-4}\,\mathrm{Mpc}^{-3}$ ($10^{-2}\,\mathrm{Mpc}^{-3}$), and more than 90% for $\ell \gtrsim 2500$ (5000). SPHEREx is expected to achieve a

galaxy number density of $\mathcal{O}(10^{-4})\,\mathrm{Mpc}^{-3}$ at $z \sim 1$ with $\sigma_z/(1+z) \sim 0.03$ [60], implying a similar suppression of kSZ template power for a SPHEREx-based template. For this reason, we will focus on spectroscopic surveys, which have $\sigma_z/(1+z) \sim \mathcal{O}(0.001)$ (e.g. [63]), or 21 cm surveys with equivalent redshift precision, for the remainder of this work.

IV. FORECASTS

In this section, we investigate the amount of kSZ power that can be recovered when different large-scale structure surveys are used to construct kSZ templates using the procedure in Sec. II B. As discussed in Sec. III, we will find that the ranges of redshift and halo mass probed, along with the number density of the galaxy sample, play an important role in the performance of the templates.

A. Spectroscopic surveys

First, we consider a representative set of spectroscopic galaxy surveys: BOSS, DESI, the High Latitude Spectroscopic Survey of the Roman Space Telescope, and the MegaMapper proposal. We state our assumptions about each survey, and their implications for kSZ reconstruction, in the following subsections. As a visual aid for these discussions, in Fig. 7 we plot the HOD model assumed for each survey at a representative redshift. Further details about HOD modelling are included in Appendix C. Our main results are summarized in Figs. 8 and 9, where we show the angular power spectrum of each kSZ template normalized to the true kSZ power spectrum, along with the linear galaxy bias and number density assumed for each survey.

1. BOSS

BOSS (the Baryon Oscillation Spectroscopic Survey [64]) measured the redshifts of 1.5 million luminous red galaxies (LRGs) over $0.4 \lesssim z \lesssim 0.7$ and $10000\,\mathrm{deg}^2$. We use the HOD from Ref. [56], which was jointly fit to the projected correlation function and void probability function measured from a subset of the BOSS CMASS sample spanning $0.46 \lesssim z \lesssim 0.57$. We use this HOD over the full BOSS redshift range, where it predicts a roughly constant number density of $\bar{n}_g \approx 10^{-4}\,\mathrm{Mpc}^{-3}$ and a linear bias ranging from 1.9 at $z = 0.4$ to 2.2 at $z = 0.7$.

Fig. 8 shows that the corresponding kSZ template can recover at most 7% of the total kSZ power at $\ell \approx 3000$, and even less at lower and higher multipoles. This is consistent with our intuition from Sec. III: Fig. 5 showed that a survey with $\bar{n}_g \approx 10^{-4}\,\mathrm{Mpc}^{-3}$ over $0 < z < 6$ and a BOSS-like HOD can reproduce at most 30% of the true kSZ power for $\ell \gtrsim 2000$, but BOSS covers a considerably smaller redshift range than what is significant for
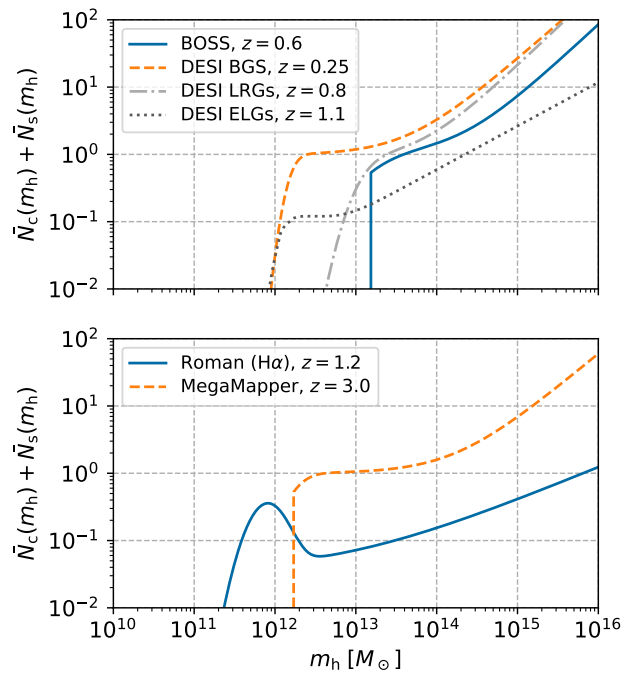

FIG. 7. HOD models for the spectroscopic surveys we forecast for in Sec. IV A. Specifically, we show the sum of central and satellite occupations, evaluated at a representative redshift for each survey. For visual clarity, we plot the models in two separate panels. These models are described in detail in Sec. IV A and Appendix C.

kSZ (recall Fig. 1), so we would expect a correspondingly smaller fraction of the kSZ power to be captured by the template.

2. DESI

DESI (the Dark Energy Spectroscopic Instrument [63]) is an ongoing survey project with the goal of measuring roughly 30 million spectroscopic galaxy and quasar redshifts over $14000\,\mathrm{deg}^2$. We separately consider the planned Bright Galaxy Sample (BGS), which will have number density greater than BOSS for $z \lesssim 0.4$; the LRG sample over $0.6 \lesssim z \lesssim 1$; the emission-line galaxy (ELG) sample over $0.6 \lesssim z \lesssim 1.6$; and the quasar (QSO) clustering sample over $0.6 \lesssim z \lesssim 1.8$. In some cases, we adjust a given HOD model or two-halo contribution to agree with the expected galaxy number density or linear bias, and we describe these adjustments below.

BGS: We use the HOD from Ref. [55], which is fit to a mock galaxy catalog with the same apparent magnitude threshold as the BGS sample. Specifically, the authors fit their 5 free HOD parameters to the clustering of galaxy subsamples corresponding to different absolute magnitudes. We use the parameter fits corresponding to $^{0.1}M_r - 5\log h = -20.5$, motivated by evidence from Ref. [65] that the distribution of absolute magnitudes of
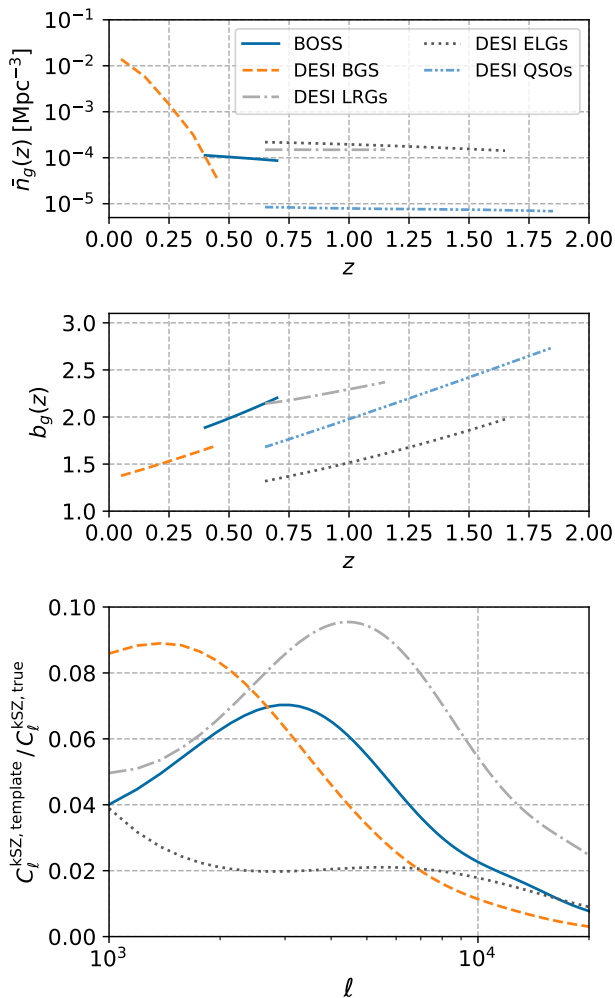
FIG. 8. Assumed number density (*upper panel*) and linear bias (*middle panel*) for BOSS and various clustering samples from DESI, along with forecasts for the ratio of kSZ template power spectrum and true kSZ power spectrum for templates constructed from each sample (*lower panel*). The DESI QSO template recovers less than 1% of the true kSZ power, so we do not show it in the lower panel. Overall, the amount of kSZ power that can be recovered by these templates is limited by the halo masses and redshift range probed by each galaxy sample, along with galaxy shot noise that restricts access to smaller-scale modes.

BGS targets peaks at this value. However, we rescale the minimum halo mass and normalization of the satellite occupation by a redshift-dependent prefactor such that the predicted mean galaxy density matches the forecast values from Ref. [63] (see Appendix C 2). We also rescale the predicted two-halo term of the power spectrum to agree with the linear bias assumed for BGS in Ref. [63] in the low-k limit.

LRGs: We use the LRG HOD from Ref. [51], which was fit to galaxies selected from the IllustrisTNG simulation according to color cuts corresponding to the LRG sample. As with our BGS model, we rescale the HOD

such that the mean galaxy density matches the LRG target density of $1.5 \times 10^{-4}\,\mathrm{Mpc}^{-3}$ quoted in Ref. [51]. We use the linear bias predicted by this model, which ranges from 2.2 at $z = 0.6$ to 2.4 at $z = 1$. If we instead rescale the HOD to match the galaxy number density used for forecasting in Ref. [63], or rescale the two-halo term to match the linear bias used in that work, the end result for $C_\ell^{\mathrm{kSZ,template}}$ can change by up to 75% at lower multipoles, but never exceeds 6% of the true kSZ power spectrum. We have also compared with results from using a different HOD fit to imaging of LRG targets in Ref. [66], and found that $C_\ell^{\mathrm{kSZ,template}}$ is around 20% higher, which is not sufficient to qualitatively change our conclusions.

ELGs: We use the ELG HOD from Ref. [51], which is based on the "high-mass quenched" ELG model from Ref. [49] and which was also fit to galaxies selected from the IllustrisTNG simulation. We use the galaxy density $(\sim 1.5 \times 10^{-4}\mathrm{Mpc}^{-3})$ and linear bias (ranging from 1.3 at $z = 0.65$ to 2.0 at $z = 1.65$) predicted by this model. In contrast to other HODs described above, the expected number of centrals in the ELG model peaks in lower-mass halos $(m_{\mathrm{h}} \sim 10^{12} M_\odot)$ and declines at higher halo masses[6], reflecting the expectation that star formation is quenched in higher-mass halos and therefore these halos are less likely to host ELGs as their central galaxy (e.g. [48–51]). Using the galaxy density and linear bias from Ref. [63] decreases $C_\ell^{\mathrm{kSZ,template}}$ by roughly 30%, while using a parameterized version of the HOD from Ref. [67] changes it by around a factor of 2, but as with the LRGs, this does not alter the nature of our conclusions.

QSOs: We use the error-function–based HOD from Ref. [49], with the best-fit parameters fit to eBOSS quasar clustering but with the lower halo mass threshold and satellite fraction normalization adjusted to match the QSO number density given in Ref. [63], and with the 2-halo term rescaled to match the linear bias from Ref. [63] in the low-k limit. Due to the low expected QSO number density $(\lesssim 10^{-5}\,\mathrm{Mpc}^{-3})$, the QSO sample contributes negligibly to a reconstruction of the late-time kSZ signal, recovering less than 1% of the true kSZ power, so we do not show the corresponding template in Fig. 8 or discuss the QSO sample further.

Amongst the different DESI samples, Fig. 8 shows that a template constructed from the LRGs captures the largest fraction of kSZ power, because its combination of halo mass range and redshift range is best matched to the ranges relevant for kSZ. However, for the same reasons as BOSS, DESI LRGs can still only recover a small fraction of kSZ power. The ranges of halo mass and/or redshift

---

[6] This feature of the central galaxy occupation for DESI ELGs is not visible in Fig. 7 because we have plotted the sum of the central and satellite occupations, but a peak at $m_{\mathrm{h}} \sim 10^{12} M_\odot$ is visible in the Roman (Hα) HOD plotted in the lower panel, for which the satellite occupation has lower amplitude.

are less optimal for the other samples, which fare even worse. Finally, Fig. 9 shows that a combination of all BOSS and DESI samples could recover around 10-20% of the true kSZ power over a wide range of multipoles.

3. Roman

The Roman Space Telescope [68] is a planned multi-purpose satellite with cosmology among its scientific drivers. In particular, the Roman High Latitude Spectroscopic Survey will obtain redshifts for 12 million ELGs over $2000 \deg^2$, using Hα emission over $1 < z < 2$ and [OIII] emission over $1 < z < 3$ [69]. We only consider the Hα sample in this work, due to its higher number density and more relevant redshift range for construction of a kSZ template.

For an HOD, we use a parameterized fit to the HOD measured from a mock ELG catalog generated from the GALACTICUS semi-analytical model applied to the UNIT N-body simulation [70]. This model has a similar form to the DESI ELG used above, but with a central occupation that peaks at slightly lower halo masses and with a lower satellite fraction. We rescale the amplitudes of both the central and satellite occupations to match the Hα number densities from Ref. [69] corresponding to fluxes $> 10^{16} \, \mathrm{erg \, s^{-1} cm^{-2}}$ and dust attenuation parameter $A_V = 1.65$, matching the catalog from Ref. [70]. We also rescale the 2-halo term such that the linear bias agrees with that in Ref. [69] ($b(z) = 1 + 0.5z$) in the low-k limit.

Fig. 9 shows that the corresponding kSZ template can recover 7% of the kSZ power at $\ell \approx 1000$, declining to less than 1% at $\ell \gtrsim 3500$. Despite the relatively high ELG number density projected for the Roman Hα sample, the corresponding halo masses are too low and the redshift range is too high to capture a significant part of the kSZ signal. The de-kSZing efficiency exhibits a minimum at $\ell \approx 6000$, which corresponds to spatial scales in the regime where neither the two-halo nor one-halo terms are dominant at the relevant redshifts.

In principle, one may also consider the Euclid satellite, which will perform its own spectroscopic survey of Hα emitters. However, Table 3 of Ref. [71] indicates that this sample will have a lower number density than Roman and similar linear bias, so a kSZ template built from Euclid's Hα sample is not expected to fare any better than what we have forecast for Roman.

4. MegaMapper

The MegaMapper [72] is a proposed ground-based telescope outfitted with DESI fiber-positioning technology that would target Lyman-break galaxies (LBGs) and Lyman-alpha emitters over $2 < z < 5$ using imaging from the Vera Rubin Observatory's Legacy Survey of Space and Time. We perform a forecast using the HOD for
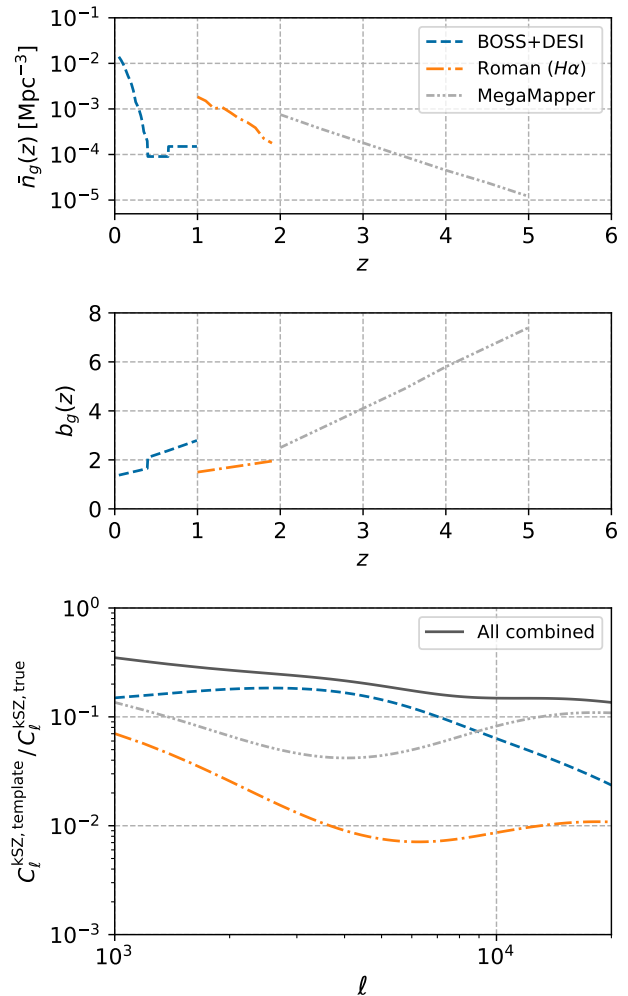

FIG. 9. Same as Fig. 8, but for kSZ templates constructed from the combination of BOSS and DESI (*blue dashed*), the Hα emitter sample from the Roman Space Telescope (*orange dot-dashed*), the Lyman-break galaxy sample from MegaMapper (*grey dot-dot-dashed*), and an idealistic combination of all of the above (*solid black*) that assumes identical sky coverage for each survey. The first three samples are mostly limited to recovering 20% or less of the true kSZ power, while the combination of all of them can recover 35% at $\ell \approx 1000$ and between 15 and 25% for $\ell \gtrsim 2000$.

LBGs obtained from early observations of roughly 6×10^5 objects by the Hyper Suprime-Cam [73]. In particular, we use their "linear HOD model" fit at $z \approx 3.8$ using galaxies with threshold apparent magnitude $m_{\mathrm{UV}}^{\mathrm{th}} = 24.5$, which is the same limiting magnitude assumed for the "idealised sample" from Ref. [74]. The galaxy number density and linear bias predicted by this model are very close to the values used for forecasting in Refs. [72, 74] (the number density ranges from $5 \times 10^{-3} \, \mathrm{Mpc^{-3}}$ at $z = 2$ to $10^{-5} \, \mathrm{Mpc^{-3}}$ at $z = 5$, and the bias ranges from 2.5 at $z = 2$ to 7 at $z = 5$), so we use them as is.

Fig. 9 shows that the corresponding kSZ template can recover roughly 15% of kSZ power at $\ell \approx 1000$ and less

than 10% at $1500 \lesssim \ell \lesssim 14000$. As with the Roman forecast discussed above, the minimum at $\ell \lesssim 4000$ is due to the transition between the two-halo and one-halo regimes, where the performance of the halo model framework is known to be particularly poor in general [75, 76]. The high linear bias of LBGs enhances the template's performance at lower ℓ (where the two-halo term is more important) compared to other samples, while the performance at higher ℓ (into the pure one-halo regime) also shows promise.

5. Combination: BOSS + DESI + Roman + MegaMapper

The solid black line in the lower panel of Fig. 9 shows the kSZ template that results in combining the templates from BOSS, DESI, Roman, and MegaMapper, using only the dominant survey at a given redshift rather than optimally combining templates at redshifts where surveys overlap. (For these forecasts, only a small loss of information is incurred by using this sub-optimal combination.) Note that we assume identical sky coverage of each survey when forming this combination, which will not be true in practice. With this caveat, we find that even this combination can recover roughly 35% of the late-time kSZ power at $\ell \sim 1000$, between 20 and 25% for $2000 \lesssim \ell \lesssim 6000$, and between 15 and 20% at higher multipoles.

B. 21 cm intensity mapping

Post-reionization large-scale structure can also be measured via 21 cm emission from neutral hydrogen (e.g. [77–79]). In this subsection, we consider the PUMA (Packed Ultra-wideband Mapping Array [80, 81]) proposal for a next-generation 21 cm intensity mapping project, intended to map LSS over $0.3 < z < 6$. We also comment on other 21 cm projects below.

Our forecasts follow a similar procedure to those for spectroscopic surveys, but with several distinctions (see Appendix D for more details). Our chosen "HOD" model does not parameterize the occupation statistics of galaxies in halos, but instead makes use of a model for the relation between neutral hydrogen (HI) mass and halo mass in a given halo, following Refs. [52, 82]. In particular, we perform forecasts that assume either the $m_{HI}(m_h)$ relation fit to the IllustrisTNG simulations in Ref. [52], or the parameterized function from Ref. [82] that was fit to observations of HI in resolved galaxies at $z \sim 0$, constraints on the mean HI density at $z \sim 1$, and properties of damped Lyman-α absorbers at $z \gtrsim 2$. We show these two options in Fig. 10. We also assume the HI density profile within a halo is given by the fitting formula from Ref. [52], again determined from IllustrisTNG; we show in Appendix D 4 that using the exponential HI profile from Ref. [82] changes our forecasts negligibly at low ℓ
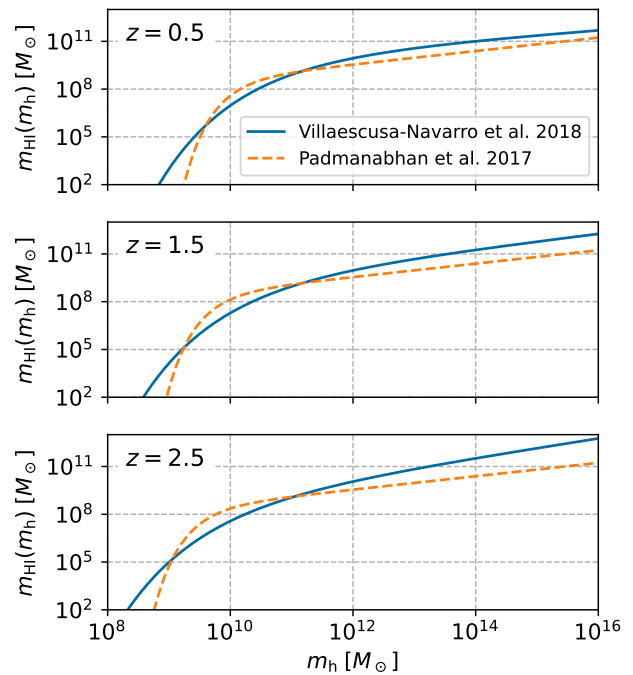

FIG. 10. The two HI mass-halo mass relations that we use in our forecasts for 21 cm intensity mapping surveys: the relation from Ref. [52] (*blue solid lines*), which was fit to the IllustrisTNG simulation over $0 < z < 5$, and that from Ref. [82] (*orange dashed lines*), which was jointly fit to a variety of low- and high-redshift observations of neutral hydrogen. Halos with masses relevant to the kSZ effect ($m_h \sim 10^{13.5} M_\odot$) contain more HI if the former relation is true, such that a kSZ template constructed from a 21 cm survey would have higher fidelity than if the latter relation is true.

and by tens of percents at higher ℓ, which is subdominant to the impact of the $m_{HI}(m_h)$ relation.

For the noise power of the HI auto spectrum, we add the (scale-dependent) instrumental noise power spectrum from Ref. [78] to the intrinsic shot noise of 21 cm emitters, computed using an updated version of the model from Ref. [83]. To ease comparisons with the surveys from Sec. IV A, we show an "effective number density" $\bar{n}_{eff}$, equal to the inverse of the 21 cm instrumental+shot noise power spectrum, in Fig. 11. We consider configurations of PUMA with 32000 or 5000 dishes ("PUMA-32k" and "PUMA-5k" respectively, as considered in Refs. [78, 81]). The effective number density decreases at higher $k_\perp$ due to the (redshift-dependent) finite angular resolution of the instrument. For PUMA-5k, the instrumental noise dominates over the shot noise, and the former increases with redshift due to the higher system temperature at lower frequencies, such that $\bar{n}_{eff}$ decreases at higher redshift. For PUMA-32k, the instrumental and shot noise both contribute to the total noise, with the former increasing and the latter decreasing at higher redshifts, leading to the observed non-monotonic behavior of $\bar{n}_{eff}$ with redshift.

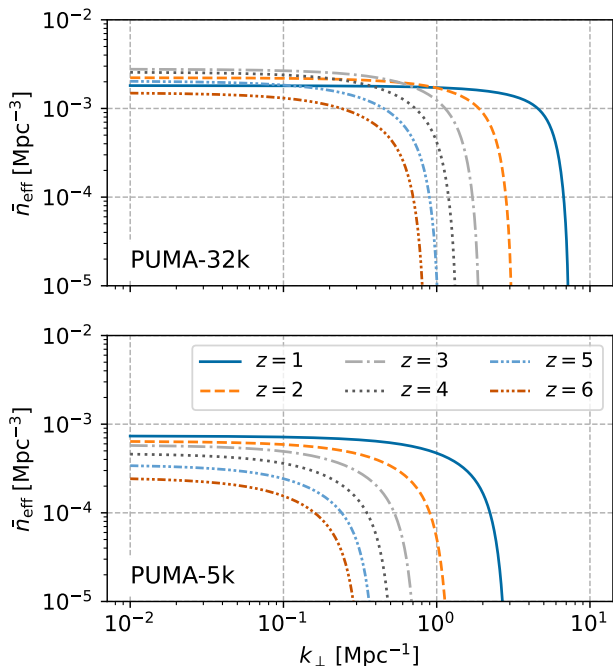Finally, Galactic and extragalactic foregrounds are ex-

FIG. 11. For ease of comparison with our forecasts for spectroscopic galaxy surveys, we translate the expected noise in a 21 cm survey, which has contributions from both the intrinsic shot noise of 21 cm emitters and instrumental noise, into an "effective shot noise" $\bar{n}_{\rm eff}$ which varies with redshift and transverse wavenumber $k_\perp$. This decreases at high $k_\perp$ due to the telescope's finite angular resolution. For PUMA-32k, the non-monotonicity with redshift at low $k_\perp$ is due to the competing influence of intrinsic shot noise and instrumental noise, which scale oppositely with redshift, while for PUMA-5k, instrumental noise is dominant.

pected to prevent certain Fourier modes of the HI distribution from being observed, and we account for this by assuming that modes with $k_\parallel < 0.03 \, \mathrm{Mpc}^{-1}$ or within a "foreground wedge" (see e.g. [84–86]) defined by 3 times the width of a PUMA dish's primary beam (following Ref. [78]) will be inaccessible. (We will comment on the impact of these assumptions below.) Note that the loss of low-$k_\parallel$ modes in the 21 cm maps does *not* imply that the kSZ template's angular power spectrum has negligible amplitude, because the template is constructed from a quadratic combination of 21 cm maps; see Appendix D 3 for more discussion on this point.

We show the resulting kSZ template forecasts in Fig. 12. In the complete absence of foregrounds, a template from PUMA-32k would be able to recover $\sim$45% of the true kSZ power at $\ell \approx 1000$, $\sim$20% at $\ell \approx 2500$, and less at higher multipoles. This strong performance compared to the spectroscopic surveys from Sec. IV A is due to the high sensitivity of PUMA-32k (equivalent to very low shot noise) shown in Fig. 11. The results at higher multipoles are strongly dependent on the form of the $m_{\rm HI}(m_{\rm h})$ relation: the relation from Ref. [52] ascribes more HI to halo masses relevant for kSZ, and therefore
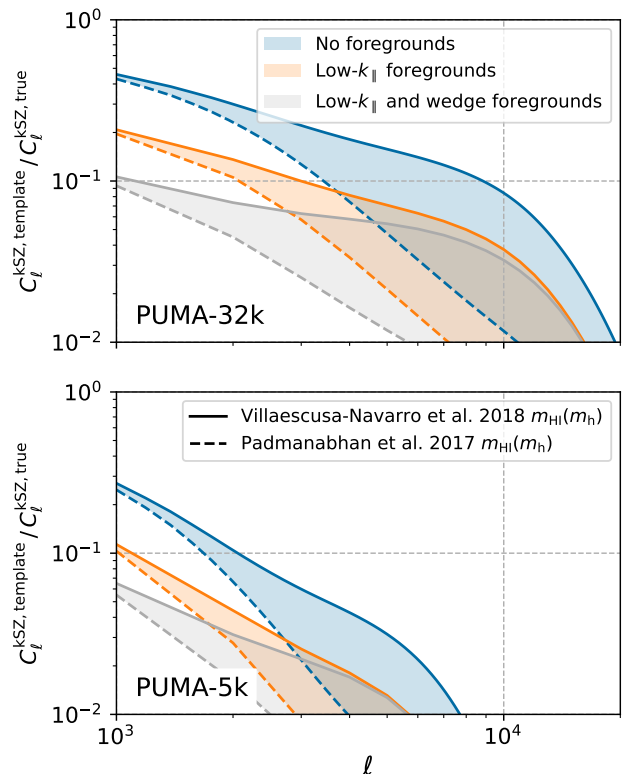

FIG. 12. Ratio of kSZ template power spectrum and true kSZ power spectrum for templates constructed from 32000-dish (*upper panel*) and 5000-dish (*lower panel*) versions of the proposed PUMA 21 cm intensity mapping survey. Shaded bands denote uncertainty related to the choice of HI mass-halo mass relation in the forecasts, computed using either of the two options from Fig. 10. In the absence of foregrounds (*blue*), the template performance at lower multipoles exceeds our forecasts for spectroscopic galaxy surveys, but when we include the loss of low-$k_\parallel$ modes due to spectrally-smooth foregrounds (*orange*) and leakage of these foreground into the "foreground wedge" (*grey*), the template's performance is significantly degraded. Furthermore, as explained in the main text and Appendix D 4, the amplitude of the full-foreground forecasts is sensitive to the precise assumptions made about foregrounds, so the grey bands should only be taken as a rough indication.

results in a better kSZ template, than the relation from Ref. [82].

Accounting for the loss of low-$k_\parallel$ modes due to foregrounds, the fraction of recovered kSZ power drops to roughly 20%, and also accounting for 21 cm signal loss within the foreground wedge reduces the recovered kSZ power to less than 10%. Note that the detrimental effect of the foreground wedge can in principle be completely removed with sufficiently accurate characterization of the 21 cm telescope and careful analysis (e.g. [87]), or via machine learning methods [88]. Also, low-$k_\parallel$ modes can in principle be recovered with various reconstruction techniques (e.g. [41, 89, 90]). However, modes recovered in this way will have different noise properties than those

assumed in our forecasts, so we leave it to future work to assess their impact on a 21 cm-based kSZ template.

As expected, PUMA-5k fares worse than PUMA-32k, due to its lower sensitivity and worse angular resolution (which prevents it from accessing higher-ℓ modes). At $\ell \approx 1000$, PUMA-5k could recover 25% of the kSZ power in the absence of foregrounds, 10% if low-$k_\parallel$ modes are lost, and 6% if wedge modes are also lost. The performance of the kSZ template also degrades more quickly with ℓ than for PUMA-32k.

We have carried out forecasts for the HIRAX [91] and CHORD [92] instruments, but found that neither could recover more than 5% of the true kSZ power even in the absence of foregrounds, so we have not shown them in Fig. 12.

In these forecasts, we have followed Ref. [78] in our assumptions about 21 cm foreground cleaning, but we note that different assumptions can lead to significant variations in the grey curves in Fig. 12. In Appendix D 4, we explore different choices for the minimum accessible $k_\parallel$ and the extent of the foreground wedge, and we find that the performance of the ensuing kSZ template can vary by as much as a factor of 10 (at low ℓ) or 5 (at higher ℓ), mostly in the downward direction (in no case do the analogs of the grey curves exceed the orange curves plotted in Fig. 12). The reader should thus bear in mind that the grey band shown in Fig. 12 is only roughly indicative of the template performance with a full treatment of foregrounds.

A further source of uncertainty in these forecasts is the stochasticity of the distribution of neutral hydrogen, which we have assumed is solely attributable to the discreteness of 21 cm-emitting objects. Ref. [93] has recently used hydrodynamical simulations to show that the stochasticity of the HI density can far exceed the Poissonian approximation at $z = 0$ but is of the same order as this approximation at $z = 1$. We leave it to future work to incorporate more accurate estimates of HI stochasticity into our forecasts.

V. APPLICATIONS

The kSZ effect constitutes the dominant contribution to the CMB black-body anisotropies on small scales ($\ell \gtrsim 4000$) and, for the next-generation Stage-3 and Stage-4 CMB experiments [95–98], is expected to be only within a factor ~ 2 smaller than the frequency-dependent CMB foregrounds, the cosmic infrared background (CIB) and the thermal SZ (tSZ) effects, after standard harmonic-space internal linear combination (ILC) cleaning [99]. For futuristic lower-noise CMB experiments like CMB-HD, the kSZ effect will likely dominate the CMB signal due to better cleaning of frequency-dependent foregrounds [100]. The removal of the kSZ effect via de-kSZing discussed in this paper hence may allow a significant reduction of the observed small-scale CMB variance.

In Fig. 13, we demonstrate the reduction of the total ILC-cleaned CMB variance by removing the (late-time) kSZ contribution. We describe our implementation of the ILC and the CMB forecasts in Appendix E. The left panel demonstrates the fractional reduction of the CMB variance at $\ell = 5000$ for the CMB-S4 and CMB-HD surveys as a function of the fractional kSZ power spectrum amplitude that remains after de-kSZing (assuming that a scale-independent fraction of the amplitude is removed). The right panel demonstrates the same reduction for a range of CMB multipoles satisfying $\ell \in [1, 20000]$. In particular, for futuristic surveys such as CMB-HD, de-kSZing can lead to a $\sim 20\%$ reduction of the CMB variance if $\sim 60\%$ of the kSZ signal can be removed. We also show forecasts for CMB-HD that omit frequency-dependent foregrounds, as a proxy for a case where these foregrounds can be cleaned extremely efficiently; in this case, de-kSZing allows for a substantial reduction in temperature variance, particularly at higher multipoles.

In this section, we briefly highlight various ways in which de-kSZing may improve cosmological inference, leaving a more detailed analysis to future work. We envision a de-kSZing procedure in which a kSZ template $\hat{T}_{\mathrm{kSZ}}$ is subtracted from an observed temperature map. Based on the results in Sec. IV, we approximate this procedure as removing a scale-independent fraction of kSZ power, and examine the consequences of such a reduction. We find that significant gains in cosmological inference generally require de-kSZing to perform better than the detailed forecasts from Sec. IV, and we take this as motivation to explore improvements to the de-kSZing formalism used in this paper.

A. Improving parameter constraints from the CMB power spectrum

The small-scale primary CMB anisotropies are exponentially suppressed due to Silk damping. The angular scale of this damping is determined by the expansion rate of the Universe and the free electron density before recombination, which are in turn sensitive to cosmological parameters such as the number of free-streaming species N_{eff}, the baryon density Ω_b, and the Helium abundance Y_{p}. De-kSZing can improve the measurement precision of the CMB damping tail, in principle leading to improved parameter constraints. Unfortunately, for the cases we have examined, including polarization information, these improvements turn out to be quite mild: even assuming 90% de-kSZing efficiency, we find that the uncertainty on determinations of N_{eff} and Ω_b decreases by no more than 10% (see Fig. 14).

For our forecasts in Fig. 14, we used the publicly available forecasting software FisherLens[7] [20], and considered a cosmological model with parameters

---

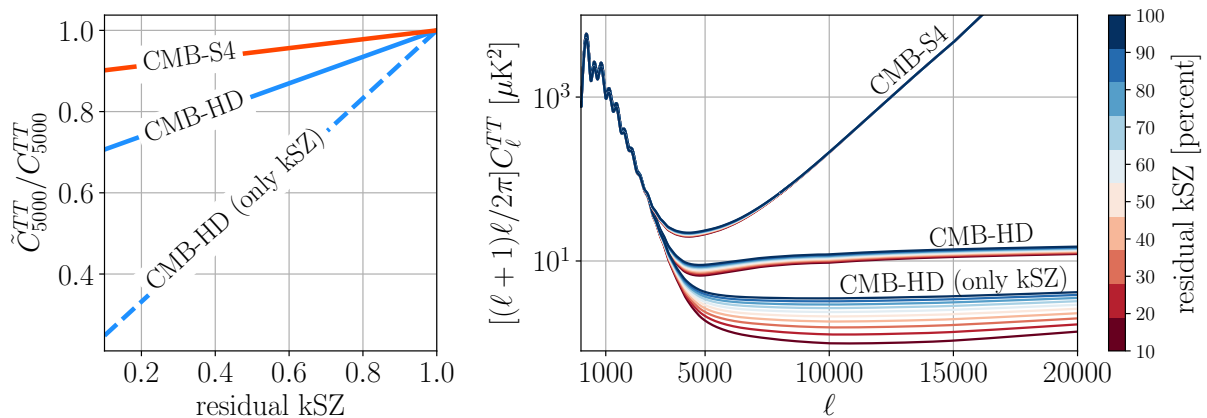[7] https://github.com/ctrendafilova/FisherLens

FIG. 13. Reducing the CMB variance via de-kSZing. *Left panel*: Reduction of the CMB temperature variance at multipole $\ell = 5000$ for the CMB-S4 and CMB-HD surveys as a function of fractional residual kSZ power spectrum amplitude after de-kSZing. The CMB variance is reduced by 5% (10%) for 50% (90%) kSZ removal for CMB-S4 and by 20% (60%) for 60% (90%) kSZ removal for CMB-HD. *Right panel*: The CMB variance as a function of ℓ for varying levels of residual kSZ power spectrum amplitude from 10% to 100%. The reduction of the CMB variance is marginal for CMB-S4, while future experiments such as CMB-HD can potentially significantly benefit from de-kSZing. Our forecasts for lines labelled CMB-S4 and CMB-HD include CMB foregrounds from CIB and tSZ after ILC-cleaning, as well as the reionization kSZ [e.g. 94], which is not removed. Forecasts for lines labelled "CMB-HD (only kSZ)" include only the late-time and reionization kSZ as foregrounds for CMB-HD, and indicate that the benefit of de-kSZing would be much greater if frequency-dependent foregrounds could be cleaned much better than currently available ILC techniques.
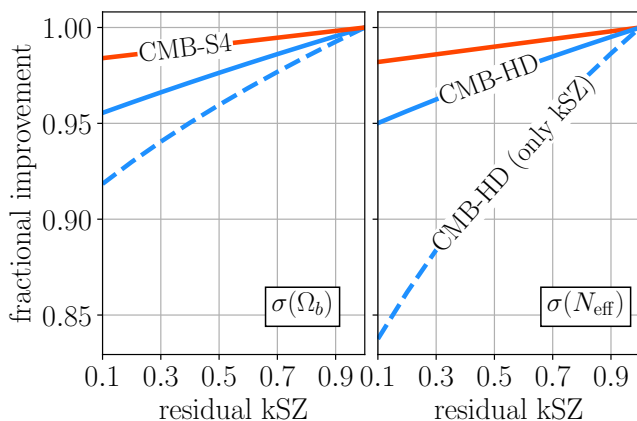

FIG. 14. Improvement from de-kSZing on measurement errors on two example cosmological parameters: baryon density Ω_b and number of free-streaming species $N_{\rm eff}$. We find 50% de-kSZing with CMB-S4 (*solid orange*) may improve errors on these parameters by around a percent, while CMB-HD (*solid blue*) can obtain 3 to 5% improvement for similar de-kSZing efficiency. We also demonstrate the power de-kSZing in the absence of frequency-dependent foregrounds with the dashed blue curves.

$\{\Omega_c h^2, \Omega_b h^2, \theta_s, \tau, A_s, n_s, N_{\rm eff}\}$ with fiducial values set to match the parameters determined by *Planck* [23]. Here, $\Omega_c h^2$ is the physical cold dark matter density, $\Omega_b h^2$ is the physical baryon density, θ_s is the angle subtended by the acoustic scale, τ is the Thomson optical depth to recombination, A_s is the primordial scalar fluctuation amplitude, and n_s is the primordial scalar fluctuation slope. We define the information matrix with elements

given by

$$F_{ij} = \sum_{\ell_1, \ell_2} \sum_{WXYZ} \frac{\partial C_{\ell_1}^{XY}}{\partial \lambda^i} \left[\text{Cov}_{\ell_1 \ell_2}^{XY,WZ} \right]^{-1} \frac{\partial C_{\ell_2}^{WZ}}{\partial \lambda^j} . \quad (33)$$

where λ_i are the cosmological parameters. We include lensed TT, TE, EE, and dd spectra, where $C_\ell^{dd} = \ell(\ell + 1)C_\ell^{\phi\phi}$ is the lensing deflection spectrum. We set the range of multipoles considered in our analysis to $\ell \in [30, 10^4]$ and describe our modelling of the foregrounds and noise in Appendix E. The covariances include the lensing-induced non-Gaussian contributions as introduced in Ref. [20], which we calculate using `FisherLens`. We set the sky fraction as $f_{\rm sky} = 0.5$. We also included a prior on τ with $\sigma_\tau = 0.007$, similar to what has been achieved by *Planck* [101].

Another observable of the CMB power spectrum is the kSZ effect from reionization. In this paper we focussed our attention on the portion of the kSZ effect that is caused by the ionized gas in moving massive halos in the relatively recent universe, at $z \lesssim 2$, but CMB maps also contain the signatures of the kSZ effect during reionization at $6 \lesssim z \lesssim 12$, originating mainly from the motions of ionized bubbles of the intergalactic medium as reionization progresses. The power spectrum of this effect has both the same spectral dependence (as a function of electromagnetic frequency) and a very similar power spectrum shape (as a function of ℓ) as that of the late-time kSZ. Its amplitude is also expected to be comparable. If a significant fraction of the late-time kSZ could be removed using the methods that we have explored, this could in principle help isolate the earlier effect and

thereby reduce the resulting uncertainty on the duration of reionization, which is the property to which the kSZ power spectrum is most sensitive [e.g., 102–104]. This would be achieved both by removing part of the highly degenerate signal power spectrum from low redshift, and, less significantly, by reducing the power spectrum variance. Finally, removing some of the late-time kSZ could help the search for signatures of reionization kSZ by using the four-point function [105]. Although one of the main advantages of this method is in helping to distinguish the late-time and reionization signals, reducing the late-time four-point signal by de-kSZing would only help in the use of this four-point method to study reionization in a new way.

B. Measurement of other CMB secondaries

1. CMB lensing

Reducing the CMB temperature variance induced by kSZ also improves lensing reconstruction. In Figure 15, we demonstrate the reduction of the TT-TT lensing quadratic-estimator noise (*upper panels, dotted curves*) together with the improvement on the minimum-variance noise from combining all lensing quadratic estimators (*solid curves*). Here, we use the class_delens software [14] and assume a CMB-HD-like survey including CIB and tSZ foregrounds after ILC-cleaning, as well as the reionization kSZ. The TT-TT lensing quadratic-estimator reconstruction noise is given by [106]

$$N_\ell^{TT-TT} = (2\ell+1)\left[\sum_{\ell_1\ell_2}\frac{|f_{\ell_1\ell\ell_2}^{TT}|^2}{2\tilde{C}_{\ell_1}^{TT}\tilde{C}_{\ell_2}^{TT}}\right]^{-1}, \quad (34)$$

where $f_{\ell_1\ell\ell_2}^{TT}$ is the optimal filter for the full-sky CMB lensing TT-TT quadratic estimator [106] and $\tilde{C}_\ell^{TT}$ is the observed CMB temperature spectrum. Performing lensing reconstruction on CMB maps after de-kSZing the CMB temperature would hence reduce the reconstruction noise, increasing the fidelity of lensing measurements.

Temperature will be the dominant lensing channel for pre-S4 experiments, such as the Simons Observatory [95, 96]. At CMB-S4 noise levels, however, the benefit of de-kSZing for reconstructing large-scale lensing modes will be much less, due to the sub-dominant contribution of the TT-TT estimator to the minimum variance lensing estimator,

$$N_\ell^{\mathrm{mv}} = \frac{1}{\sum_{\alpha\beta}(\boldsymbol{N}_\ell^{-1})^{\alpha\beta}}, \quad (35)$$

where $\boldsymbol{N}$ is the covariance of all quadratic estimators from temperature and polarization maps. In particular, on scales $\ell \lesssim 10^4$, the minimum-variance lensing reconstruction is dominated by the EB-EB quadratic estimator [106], which is not improved by de-kSZing. On the other hand, when reconstructing the smaller-scale ($\ell \gtrsim 10^4$) lensing modes, the TT-TT estimator plays a more significant role, suggesting that de-kSZing can be important for improving small-scale CMB lensing reconstruction in the future.[8] We find that for a CMB-HD-like survey, the TT-TT lensing quadratic-estimator noise decreases by $\sim$15% on small scales if $\sim$30% of the kSZ signal can be removed.

Small-scale CMB lensing reconstruction has applications including cosmological parameter inference, distinguishing between different dark-matter models [11], constraining high-redshift astrophysics, and validating galaxy weak-lensing shear measurements [110].

2. Moving-lens effect

Reducing the CMB variance on small scales can improve the measurement precision of cluster and halo profiles from several CMB secondaries. In Fig. 16, we consider the moving-lens effect [13–15] as a concrete example. In the moving-lens effect, a halo with peculiar motion transverse to the line of sight creates a small-scale dipolar temperature anisotropy centered on the halo's location on the sky. If we use $\boldsymbol{r}_\perp$ to denote the transverse proper distance from the halo center, and define $\boldsymbol{x}_\perp \equiv \boldsymbol{r}_\perp/r_s$ where r_s is the halo scale radius, the moving lens signal in the dimensionless CMB temperature $\Theta \equiv \Delta T/T_{\mathrm{CMB}}$ is

$$\Theta_{\mathrm{ML}}(\boldsymbol{x}_\perp) = -a_0\boldsymbol{v}_{\mathrm{b},\perp}\cdot\boldsymbol{\mathcal{M}}(\boldsymbol{x}_\perp), \quad (36)$$

where $\boldsymbol{v}_{\mathrm{b},\perp}$ is the halo's transverse velocity. The prefactor a_0 and moving-lens profile $\boldsymbol{\mathcal{M}}(\boldsymbol{x}_\perp)$ depend on the halo density profile; assuming an NFW profile, they evaluate to [14]

$$a_0 \equiv \frac{16\pi G\rho_s r_s^2}{c^3}, \quad (37)$$

$$\boldsymbol{\mathcal{M}}(\boldsymbol{x}_\perp) \equiv \frac{\boldsymbol{x}_\perp}{2x_\perp^2}\left[\left|\frac{2\sec^{-1}(x_\perp)}{\sqrt{x_\perp^2-1}}\right| + \ln\left(\frac{x_\perp^2}{4}\right)\right], \quad (38)$$

where

$$\rho_s = \frac{M}{4\pi r_s^3}\left[-\frac{r_{\mathrm{vir}}}{r_s+r_{\mathrm{vir}}} - \ln\left(\frac{r_s+r_{\mathrm{vir}}}{r_s}\right)\right]. \quad (39)$$

Here, we forecast the measurement precision on the angle-averaged moving-lens signal of a 10^{14} solar-mass halo, in $x_\perp$ bins with width $\Delta x_\perp = 0.5$ and taking the fiducial density profile to be NFW. For this forecast, we

---

[8] Note also that de-kSZing can potentially reduce the kSZ-induced biases on the lensing reconstruction calculated in Refs [9, 107, 108], and the corresponding bias on de-lensed B-modes discussed in Ref. [109].
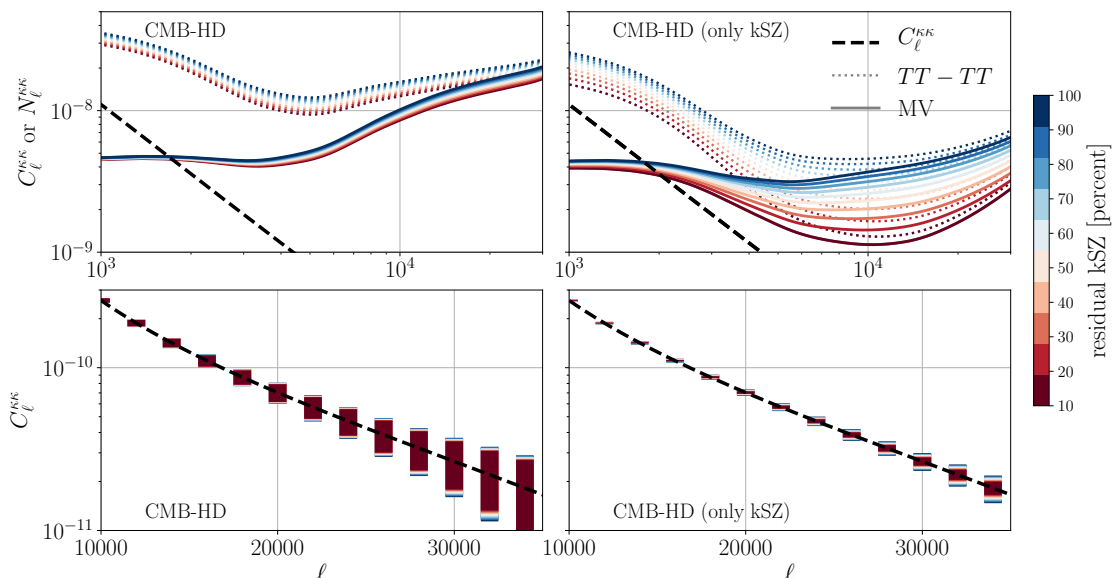
FIG. 15. *Upper panels:* Improvement of the lensing-reconstruction TT-TT quadratic-estimator noise (*dotted curves*) together with the improvement on the minimum-variance noise from combining all temperature and polarization quadratic-estimators (*solid curves*). *Lower panels:* Lensing convergence power spectra (*solid black*). The error bars shown are from the diagonal terms of the minimum-variance lensing reconstruction noise. We use the `class_delens` software [14] and consider a CMB-HD-like survey throughout. *Left panels:* We include all significant foregrounds (CIB and tSZ, after ILC-cleaning, as well as the reionization kSZ) to demonstrate the power of de-kSZing in the presence of frequency dependent foregrounds. *Right panels:* We include only the kSZ foreground (both late-time and reionization). De-kSZing improves the TT-TT lensing reconstruction noise at all scales, while the minimum-variance noise is only improved at smaller scales where the TT-TT estimator contributes significantly to the minimum-variance lensing reconstruction.

use the optimal matched filter derived in Ref. [14], applied separately to each $x_\perp$ bin to enable a measurement of the radially-binned moving lens signal (see Appendix F for details).

De-kSZing reduces the observed CMB variance $\tilde{C}_\ell^{TT}$ on scales where the moving-lens reconstruction noise gets its dominant contribution. The uncertainty bands in Fig. 16 are scaled to match the net uncertainty anticipated from measurements of 5×10^4 halos of the same mass at redshift $z = 1$, roughly representative of the number of cluster-size halos probed by recent and upcoming imaging surveys [111]. The improvement in the uncertainty in a given bin can be seen as a function of the residual kSZ power remaining in the CMB maps after de-kSZing. We find that $\sim 50\%$ removal of the kSZ effect can lead up to $\sim 30\%$ improvement of the measurement of the moving-lens signal at a given distance from the halo center, depending on the survey systematics and other foregrounds.

Unlike kSZ or other Sunyaev Zel'dovich effects, the moving-lens effect is purely gravitational and can be used to probe quantities such as the growth rate of cosmological structure [15], for example, without depending on the modelling of the electron gas. De-kSZing would not only boost the detection significance of the moving-lens effect but also potentially increase the prospects to perform cosmological and astrophysical inference with this promising observable.

VI. CONCLUSION

In this paper, we have explored the prospects for "de-kSZing" the cosmic microwave background: constructing a template for the late-time kinetic Sunyaev-Zeldovich effect using a galaxy or 21 cm survey, and subtracting this template from observed CMB temperature maps. The template is constructed following the procedure from Refs. [27, 28, 33]: separate templates for the small-scale electron density and large-scale velocity field are formed by Wiener-filtering the observed galaxy density; these templates are combined into an estimate of the electron momentum field; and this combination is projected along the line of sight, using the theoretical kSZ redshift kernel, to form an estimate of kSZ-induced temperature fluctuations in the CMB. We have identified the properties of the input galaxy survey that most strongly determine the fidelity of the associated kSZ template: redshift and halo mass coverage (Sec. III A) and galaxy number density (Sec. III B).

We have also carried out detailed forecasts for a number of recent, upcoming, or proposed surveys, assessing the ability of each survey to provide a template that could be used for de-kSZing. Since redshift uncertainties in photometric surveys will significantly degrade the usefulness of such a template (Sec. III D), we have focused on spectroscopic and 21 cm surveys in our forecasts.

The combination of BOSS and several of DESI's planned samples could in principle yield a template that
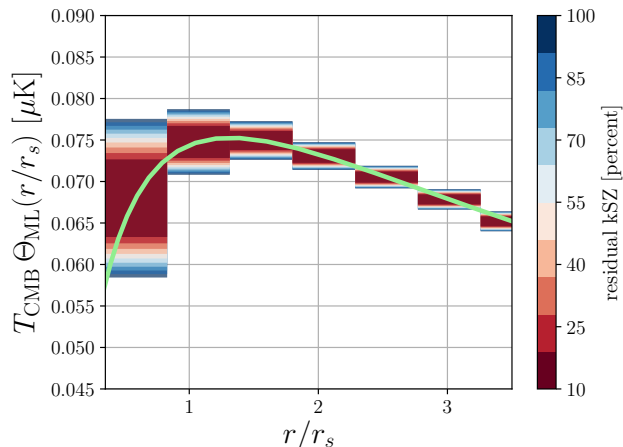
FIG. 16. Improvement on the measurement accuracy of the moving-lens profile. The error bars are scaled to match the net uncertainty anticipated from measurements of 5×10^4 halos of the same mass ($M = 10^{14} M_\odot$) at redshift $z = 1$ using CMB-HD (considering only the kSZ foreground). Different bins in $x = r/r_s$ are calculated by filtering the CMB around halos in radial bins with radial distance x from the halo center with width $\Delta x \simeq 0.5$. The improvement in the measurement error of the halo profile at a given x-bin can be seen as a function of the residual kSZ (in percentage) in the CMB maps after de-kSZing. The degrading effect of different foregrounds on the detection and characterisation of the moving-lens effect could depend on the statistics used in the analysis, which may determine the benefit of de-kSZing in practice.

captures 10-20% of the total kSZ power, while templates derived from the Hα emission-line galaxy sample from the Roman Space Telescope or the Lyman-break galaxy sample from the proposed MegaMapper telescope can capture no more than 10% of the kSZ power at $\ell \gtrsim 1500$ (Fig. 9). Once foreground contamination is accounted for, our forecasts indicate that the proposed PUMA 21 cm intensity mapping survey can do no better than the combination of BOSS and DESI, although uncertainty in modelling of the 21 cm signal on small scales, along with the achievable level of foreground cleaning, makes this conclusion rather uncertain (Fig. 12). Table I provides representative numbers for the surveys we have considered.

However, we emphasize that we have only performed an initial exploration of a specific de-kSZing procedure in this work, and we expect that alternative procedures can yield significant improvements to the results we have presented here. With this in mind, we have highlighted several applications of the idea of de-kSZing, all of which rely on the associated reduction of the small-scale variance in CMB temperature maps: better recovery of cosmological information from the CMB damping tail (Sec. V A); lower noise in measurements of CMB lensing at small scales, which can be used to test models for dark matter (Sec. V B 1); and more precise measurements of the moving-lens effect, which can be used to probe the cosmic growth rate and the properties of dark matter halos

Survey	De-kSZing Efficiency (%)		
	$\ell = 1000$	$\ell = 4000$	$\ell = 8000$
BOSS + DESI	15	17	8
Roman (Hα)	7	<1	<1
MegaMapper	13	4	7
All of the above	35	21	15
PUMA-32k	10	4	<1

TABLE I. Representative de-kSZing efficiencies for kSZ templates constructed from the surveys considered in this work. See Sec. IV for discussions of important caveats associated with these numbers; in particular, the "all of the above" line only applies to the overlapping sky footprints of the surveys it combines, and the PUMA-32k numbers are highly uncertain due to modelling uncertainties in the HI distribution at small scales and the achievable level of foreground cleaning in 21 cm surveys. De-kSZing procedures that improve upon that presented in this work could potentially improve these efficiencies substantially.

(Sec. V B 2).

Motivated by these applications, there are several pathways towards an improved de-kSZing procedure that would be worth pursuing:

- Different surveys (or combinations of surveys) could be used for the velocity and electron density templates, which would make better use of the properties of each survey. In particular, photometric redshift errors would be tolerable if a photometric survey was only used for the electron density template. As a concrete example, the LSST Y10 Gold lens sample is estimated[9] to have a number density of $4 \times 10^{-3} \, \mathrm{Mpc}^{-3}$ for $0.5 < z < 2$. Assuming that this sample is used for the electron template but that the velocity template comes from a sample with $\bar{n}_g \sim 10^{-4} \, \mathrm{Mpc}^{-3}$, we infer from Figs. 4-5 that a de-kSZing efficiency of roughly 40% at $\ell > 3000$ may be achievable. More detailed forecasts will be needed to confirm this conclusion, however.

- Extra weights could be applied to the input galaxy survey to optimize the correlation of the kSZ template with the true kSZ fluctuations (in a statistical sense). For example, if halo mass estimates for each galaxy are available, different mass bins could be weighted differently (e.g. [112, 113]) in order to better account for the mass- and redshift-dependence of the kSZ signal (recall Fig. 1). The expected improvement would depend on the accuracy of the halo masses, which would determine the number of mass bins that could be used.

---

[9] This estimate is based on a total angular number density of $48 \, \mathrm{arcmin}^{-2}$ over $18000 \, \mathrm{deg}^2$, and a galaxy redshift distribution of $dN/dz \propto z^2 \exp[-(z/0.28)^{0.9}]$ [62].

- Tomographic reconstruction of large-scale velocities using the kSZ effect [33, 35, 114–121] could possibly be integrated into a de-kSZing procedure as a way to improve upon an external velocity template. This would have the advantage of constructing the velocity template from small-scale modes of the galaxy density (along with small-scale CMB temperature modes), but will introduce higher-point biases and noise contributions that would need to be accounted for, since the velocity template will then be quadratic ($\sim \delta_g^s T$) instead of linear ($\sim \delta_g^s$). This procedure would be analogous to iterative de-lensing of CMB polarization [122].

- Rather than viewing de-kSZing as a subtraction procedure, it may be possible to use Bayesian (e.g. [123]) or machine learning (e.g. [124, 125]) techniques to jointly estimate the statistics of the kSZ effect, the primary CMB, and/or other CMB secondaries. The former approach would allow for uncertainties in the procedure and data to be accounted for more systematically, while the latter may be able to exploit features of the kSZ signal (such as non-Gaussian information) that are not incorporated in our template construction. We note, however, that either approach would likely require significant computing power.

We leave these possibilities to future work.

Appendix A: Shot noise contributions to kSZ template auto spectrum

The kSZ template described in Sec. II B is related to a product of filtered copies of an observed galaxy density contrast δ_g, and therefore the autocorrelation of the template is sensitive to the four-point function of δ_g. (For simplicity, we neglect the impact of redshift-space distortions in this appendix, and therefore use δ_g instead of δ_g^s.) This four-point function contains several contributions from the shot noise in δ_g; in Fourier space and for Poissonian shot noise, these contributions are given by (e.g. [40, 41])

$$\langle \delta_g(\boldsymbol{k}_1)\delta_g(\boldsymbol{k}_2)\delta_g(\boldsymbol{k}_3)\delta_g(\boldsymbol{k}_4)\rangle_{\text{shot}} = (2\pi)^3 \delta_D(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 + \boldsymbol{k}_4)\left[S_4 + S_{3,1} + S_{2,2} + S_{2,1,1}\right] \tag{A1}$$

where

$$S_4 = \frac{1}{\bar{n}_g^3} \,,$$

$$S_{3,1} = \frac{1}{\bar{n}_g^2}\left[P_{gg}(k_1) + P_{gg}(k_2) + P_{gg}(k_3) + P_{gg}(k_4)\right] \,,$$

$$S_{2,2} = \frac{1}{\bar{n}_g^2}\left[P_{gg}(|\boldsymbol{k}_1 + \boldsymbol{k}_2|) + P_{gg}(|\boldsymbol{k}_1 + \boldsymbol{k}_3|) + P_{gg}(|\boldsymbol{k}_1 + \boldsymbol{k}_4|)\right] \,,$$

$$S_{2,1,1} = \frac{1}{\bar{n}_g}\left[B_{ggg}(\boldsymbol{k}_1 + \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4) + B_{ggg}(\boldsymbol{k}_1 + \boldsymbol{k}_3, \boldsymbol{k}_2, \boldsymbol{k}_4) + B_{ggg}(\boldsymbol{k}_1 + \boldsymbol{k}_4, \boldsymbol{k}_2, \boldsymbol{k}_3)\right.$$
$$\left. + B_{ggg}(\boldsymbol{k}_2 + \boldsymbol{k}_3, \boldsymbol{k}_1, \boldsymbol{k}_4) + B_{ggg}(\boldsymbol{k}_2 + \boldsymbol{k}_4, \boldsymbol{k}_1, \boldsymbol{k}_3) + B_{ggg}(\boldsymbol{k}_3 + \boldsymbol{k}_4, \boldsymbol{k}_1, \boldsymbol{k}_2)\right] \tag{A2}$$

and we have omitted the redshift-dependence and the arguments of the S functions for brevity. Each contribution arises from a different subset of points taken to be at zero lag in the position-space correlation function: S_4 is when all four points are at the same spatial location; $S_{3,1}$ is when three points are at the same location and one is elsewhere; $S_{2,2}$ is when two pairs of points are each at the same location, but the pairs are not co-located; and $S_{2,1,1}$ is when two points are at the same location, while the other two points are each at different locations.

The power spectrum of the electron momentum template in Eq. (13) evaluates to

$$\langle \hat{q}_{\rm r}(\boldsymbol{k}_1)\hat{q}_{\rm r}(\boldsymbol{k}_2)\rangle = \left\langle \int \frac{d^3\boldsymbol{k}_1'}{(2\pi)^3}\eta(\boldsymbol{k}_1')\epsilon(\boldsymbol{k}_1-\boldsymbol{k}_1')\int \frac{d^3\boldsymbol{k}_2'}{(2\pi)^3}\eta(\boldsymbol{k}_2')\epsilon(\boldsymbol{k}_2-\boldsymbol{k}_2')\right\rangle$$

$$= (-1)\int \frac{d^3\boldsymbol{k}_1'}{(2\pi)^3}\int \frac{d^3\boldsymbol{k}_2'}{(2\pi)^3}\mu_1'\mu_2'\frac{P_{gv}(k_1')}{P_{\rm gg}^{\rm tot}(k_1')}\frac{P_{gv}(k_2')}{P_{\rm gg}^{\rm tot}(k_2')}\frac{P_{\rm ge}(|\boldsymbol{k}_1-\boldsymbol{k}_1'|)}{P_{\rm gg}^{\rm tot}(|\boldsymbol{k}_1-\boldsymbol{k}_1'|)}\frac{P_{\rm ge}(|\boldsymbol{k}_2-\boldsymbol{k}_2'|)}{P_{\rm gg}^{\rm tot}(|\boldsymbol{k}_2-\boldsymbol{k}_2'|)}$$

$$\times \left\langle \delta_{\rm g}(\boldsymbol{k}_1')\delta_{\rm g}(\boldsymbol{k}_1-\boldsymbol{k}_1')\delta_{\rm g}(\boldsymbol{k}_2')\delta_{\rm g}(\boldsymbol{k}_2-\boldsymbol{k}_2')\right\rangle, \tag{A3}$$

and combined with Eqs. (A1) and (A2), this yields

$$P_{\hat{q}_{\rm r}\hat{q}_{\rm r}}(\boldsymbol{k}_1) = (-1)\int \frac{d^3\boldsymbol{k}_1'}{(2\pi)^3}\int \frac{d^3\boldsymbol{k}_2'}{(2\pi)^3}\mu_1'\mu_2'\frac{P_{gv}(k_1')}{P_{\rm gg}^{\rm tot}(k_1')}\frac{P_{gv}(k_2')}{P_{\rm gg}^{\rm tot}(k_2')}\frac{P_{\rm ge}(|\boldsymbol{k}_1-\boldsymbol{k}_1'|)}{P_{\rm gg}^{\rm tot}(|\boldsymbol{k}_1-\boldsymbol{k}_1'|)}\frac{P_{\rm ge}(|\boldsymbol{k}_1+\boldsymbol{k}_2'|)}{P_{\rm gg}^{\rm tot}(|\boldsymbol{k}_1+\boldsymbol{k}_2'|)}\left[\hat{S}_4 + \hat{S}_{3,1} + \hat{S}_{2,2} + \hat{S}_{2,1,1}\right], \tag{A4}$$

where

$$\hat{S}_4 = \frac{1}{\bar{n}_{\rm g}^3},$$

$$\hat{S}_{3,1} = \frac{1}{\bar{n}_{\rm g}^2}\left[P_{\rm gg}(k_1') + P_{\rm gg}(|\boldsymbol{k}_1-\boldsymbol{k}_1'|) + P_{\rm gg}(k_2') + P_{\rm gg}(|\boldsymbol{k}_1+\boldsymbol{k}_2'|)\right],$$

$$\hat{S}_{2,2} = \frac{1}{\bar{n}_{\rm g}^2}\left[P_{\rm gg}(k_1) + P_{\rm gg}(|\boldsymbol{k}_1'+\boldsymbol{k}_2'|) + P_{\rm gg}(|\boldsymbol{k}_1'-\boldsymbol{k}_1-\boldsymbol{k}_2'|)\right],$$

$$\hat{S}_{2,1,1} = \frac{1}{\bar{n}_{\rm g}}\left[B_{\rm ggg}(\boldsymbol{k}_1,\boldsymbol{k}_2',-\boldsymbol{k}_1-\boldsymbol{k}_2') + B_{\rm ggg}(\boldsymbol{k}_1'+\boldsymbol{k}_2',\boldsymbol{k}_1-\boldsymbol{k}_1',-\boldsymbol{k}_1-\boldsymbol{k}_2') + B_{\rm ggg}(\boldsymbol{k}_1-\boldsymbol{k}_1-\boldsymbol{k}_2',\boldsymbol{k}_1-\boldsymbol{k}_1',\boldsymbol{k}_2')\right.$$

$$\left.+ B_{\rm ggg}(\boldsymbol{k}_1-\boldsymbol{k}_1'+\boldsymbol{k}_2',\boldsymbol{k}_1',-\boldsymbol{k}_1-\boldsymbol{k}_2') + B_{\rm ggg}(-\boldsymbol{k}_1'-\boldsymbol{k}_2',\boldsymbol{k}_1',\boldsymbol{k}_2') + B_{\rm ggg}(-\boldsymbol{k}_1,\boldsymbol{k}_1',\boldsymbol{k}_1-\boldsymbol{k}_1')\right]. \tag{A5}$$

In the squeezed limit $(k_1', k_2' \ll k_1)$, many of these terms will integrate to zero in Eq. (A4), because they have no dependence on either μ_1' or μ_2' but are integrated against $\mu_1'\mu_2'$. (Physically, this reflects the fact that the kSZ effect is sensitive to the line-of-sight component of the large-scale velocity, and the angular average of this is zero.) However, terms that involve the sum $\boldsymbol{k}_1' + \boldsymbol{k}_2'$ may contribute non-negligibly. In particular, such terms involve power at larger scales than one might naively expect, because it is possible to have $|\boldsymbol{k}_1' + \boldsymbol{k}_2'| \ll k_1', k_2'$ if $\boldsymbol{k}_1'$ and $\boldsymbol{k}_2'$ are similar in magnitude and anti-aligned.[10]

In Fig. 17, we evaluate the two dominant terms of this type in the squeezed limit, and compare them to the Gaussian contribution to $P_{\hat{q}_{\rm r}\hat{q}_{\rm r}}$. Specifically, we compute

$$P_{\hat{q}_{\rm r}\hat{q}_{\rm r}}^{\rm shot,P}(k) = (-1)\int \frac{d^3\boldsymbol{k}_1'}{(2\pi)^3}\int \frac{d^3\boldsymbol{k}_2'}{(2\pi)^3}\mu_1'\mu_2'\frac{P_{gv}(k_1')}{P_{\rm gg}^{\rm tot}(k_1')}\frac{P_{gv}(k_2')}{P_{\rm gg}^{\rm tot}(k_2')}\left(\frac{P_{\rm ge}(k)}{P_{\rm gg}^{\rm tot}(k)}\right)^2\frac{1}{\bar{n}_{\rm g}^2}P_{\rm gg}(|\boldsymbol{k}_1'+\boldsymbol{k}_2'|) \tag{A6}$$

and

$$P_{\hat{q}_{\rm r}\hat{q}_{\rm r}}^{\rm shot,B}(k) = (-1)\int \frac{d^3\boldsymbol{k}_1'}{(2\pi)^3}\int \frac{d^3\boldsymbol{k}_2'}{(2\pi)^3}\mu_1'\mu_2'\frac{P_{gv}(k_1')}{P_{\rm gg}^{\rm tot}(k_1')}\frac{P_{gv}(k_2')}{P_{\rm gg}^{\rm tot}(k_2')}\left(\frac{P_{\rm ge}(k)}{P_{\rm gg}^{\rm tot}(k)}\right)^2\frac{1}{\bar{n}_{\rm g}}B_{\rm ggg}(-\boldsymbol{k}_1'-\boldsymbol{k}_2',\boldsymbol{k}_1',\boldsymbol{k}_2'), \tag{A7}$$

and compare with Eq. (16), ignoring redshift-space distortions:

$$P_{\hat{q}_{\rm r}q_{\rm r}}^{\rm Gaus}(k) = \frac{1}{6\pi^2}\left(\int dk_{\rm L}\, k_{\rm L}^2\frac{P_{gv}(k_{\rm L})^2}{P_{\rm gg}^{\rm tot}(k_{\rm L})}\right)\frac{P_{\rm ge}(k)^2}{P_{\rm gg}^{\rm tot}(k)}. \tag{A8}$$

We evaluate the various power spectra using the halo model framework described in the main text and Appendix B, for each of the spectroscopic surveys considered in Sec. IV A. We numerically compute Eqs. (A6) and (A7) using the Monte Carlo integration algorithm described in Ref. [126], as implemented in the `vegas` Python package [127].

For most of the surveys we consider, these shot noise contributions are below 25% of the Gaussian contribution at $\ell = 2000$ and below 40% at $\ell = 8000$. (A noteworthy exception is the "shot,B" contribution for the DESI BGS

---

[10] In Ref. [41], these terms were found to significantly affect the precision with which long-wavelength modes of $\delta_{\rm g}$ can be reconstructed with a quadratic estimator.
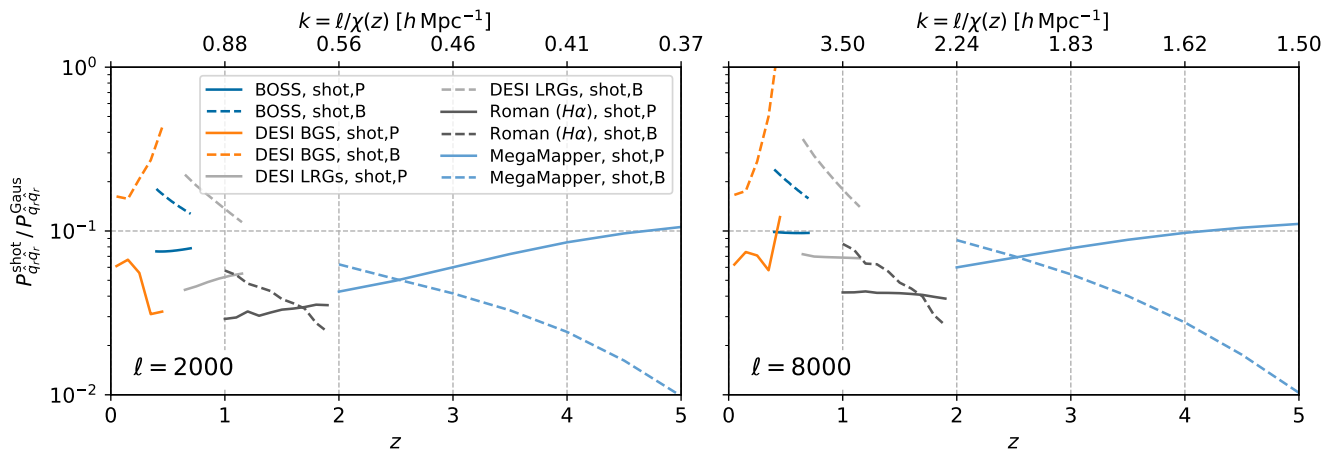
FIG. 17. The two dominant shot-noise contributions to the auto power spectrum of the electron momentum template $\hat{q}_r$, plotted as ratios to the Gaussian (i.e. reconstructed kSZ signal) contribution to $P_{\hat{q}_r\hat{q}_r}$. *Solid lines* denote the contribution that involves the galaxy power spectrum (Eq. A6), while *dashed lines* correspond to the contribution related to the galaxy bispectrum (Eq. A7). Different colors denote different spectroscopic surveys, described in Sec. IV A. With the exception of DESI BGS, these shot-noise contributions are less than $\sim 25\%$ of the template power at $\ell = 2000$ and less than $\sim 40\%$ at $\ell = 8000$. While subdominant to the desired signal in the template, this shot-noise-induced power should be accounted for in future de-kSZing analyses.

sample, which strongly increases at the higher-redshift edge of the BGS sample, $z \approx 0.4$, at $\ell = 8000$; this is caused by the decreasing number density of the BGS sample at this redshift, combined with the overall higher amplitude of the galaxy bispectrum in the BGS redshift range compared to the higher redshifts covered by the other surveys.) These numbers can be interpreted as upper bounds on the amount of power in a kSZ template that arises from shot noise in the galaxy sample instead of reconstructed kSZ signal. Future modelling related to applications of de-kSZing will need to account for this.

Appendix B: Details of halo model

In this appendix, we provide more details of the halo model approach we use to model quantities related to spectroscopic surveys. The galaxy power spectrum, electron power spectra, integrand of the kSZ angular power spectrum (see Eq. 21), and galaxy bispectrum are discussed in Sec. B 1, B 2, B 3, and B 4 respectively. For brevity, we have omitted redshift arguments in Sec. B 1, B 2, and B 4.

1. Galaxy power spectrum

Our formalism is based on Ref. [33]. The galaxy power spectrum is a sum of two-halo, one-halo, and shot noise terms,

$$P_{\mathrm{gg}}(k,\mu) = P_{\mathrm{gg}}^{2\mathrm{h}}(k,\mu) + P_{\mathrm{gg}}^{1\mathrm{h}}(k) + P_{\mathrm{gg}}^{\mathrm{shot}} . \tag{B1}$$

As stated in the main text, we do not include Finger of God damping in our computations, and therefore the one-halo term has no μ-dependence.[11] The two-halo term is given by

$$P_{\mathrm{gg}}^{2\mathrm{h}}(k,\mu) = \left[b_{\mathrm{g}}(k) + f\mu^2 \right]^2 P_{\mathrm{lin}}(k) , \tag{B2}$$

---

[11] See Ref. [128] for a formalism that consistently includes the Finger of God effect in a halo model framework.

where P_{lin} is the linear matter power spectrum and the prefactor incorporates the leading effect of redshift space distortions at large scales. The linear bias b_{g} is given by

$$b_{\text{g}}(k) = \frac{1}{\bar{n}_{\text{g}}} \int dm_{\text{h}}\, n(m_{\text{h}}) b_{\text{h}}(m_{\text{h}}) \left[\bar{N}_{\text{c}}(m_{\text{h}}) + \bar{N}_{\text{s}}(m_{\text{h}}) u_{\text{s}}(k, m_{\text{h}}) \right] \tag{B3}$$

where $n(m_{\text{h}})$ is the halo mass function, $b_{\text{h}}(m_{\text{h}})$ is the halo bias, $\bar{N}_{\text{c}}$ and $\bar{N}_{\text{s}}$ are the mean central and satellite occupation numbers, u_{s} is the Fourier transform of the assumed profile of satellite galaxies, and the mean galaxy number density $\bar{n}_{\text{g}}$ is

$$\bar{n}_{\text{g}} = \int dm_{\text{h}}\, n(m_{\text{h}}) \left[\bar{N}_{\text{c}}(m_{\text{h}}) + \bar{N}_{\text{s}}(m_{\text{h}}) \right] \ . \tag{B4}$$

The one-halo term is

$$P_{\text{gg}}^{\text{1h}}(k) = \frac{1}{\bar{n}_{\text{g}}^2} \int dm_{\text{h}}\, n(m_{\text{h}}) \left[2 \left\langle N_{\text{c}}(m_{\text{h}}) N_{\text{s}}(m_{\text{h}}) \right\rangle u_{\text{s}}(k, m_{\text{h}}) + \left\langle N_{\text{s}}(m_{\text{h}})[N_{\text{s}}(m_{\text{h}}) - 1] \right\rangle u_{\text{s}}(k, m_{\text{h}})^2 \right] \ . \tag{B5}$$

The expectation values of halo occupation numbers depend on what is assumed about correlations between numbers of central and satellite galaxies within a given halo. We use the "minimally correlated" assumption from Ref. [33], in which the central and satellite occupation numbers are independent Poisson random variables; this yields

$$\left\langle N_{\text{c}}(m_{\text{h}}) N_{\text{s}}(m_{\text{h}}) \right\rangle = \bar{N}_{\text{c}}(m_{\text{h}}) \bar{N}_{\text{s}}(m_{\text{h}}) \ , \qquad \left\langle N_{\text{s}}(m_{\text{h}})[N_{\text{s}}(m_{\text{h}}) - 1] \right\rangle = \bar{N}_{\text{s}}(m_{\text{h}})^2 \ . \tag{B6}$$

In specific HOD models where a central-satellite correlation is assumed, we implement this correlation by including $\bar{N}_{\text{c}}(m_{\text{h}})$ as a prefactor in the model for $\bar{N}_{\text{s}}(m_{\text{h}})$, ensuring that $\bar{N}_{\text{s}} = 0$ if $\bar{N}_{\text{c}} = 0$ at a given halo mass.

Finally, the shot noise term is given by

$$P_{\text{gg}}^{\text{shot}} = \frac{1}{\bar{n}_{\text{g}}} \ . \tag{B7}$$

2. Electron power spectra

In this paper, as in Ref. [33], we assume that free electrons trace the gas within halos, such that we can describe electron distributions using the standard halo model for matter, but with the matter density profile replaced with the gas profile. The electron density power spectrum is then given by

$$P_{\text{ee}}(k) = P_{\text{ee}}^{\text{2h}}(k) + P_{\text{ee}}^{\text{1h}}(k) \ , \tag{B8}$$

with

$$P_{\text{ee}}^{\text{2h}}(k) = \left[\int dm_{\text{h}}\, n(m_{\text{h}}) b_{\text{h}}(m_{\text{h}}) \frac{m_{\text{h}}}{\bar{\rho}_{\text{m}}} u_{\text{gas}}(k, m_{\text{h}}) \right]^2 P_{\text{lin}}(k) \tag{B9}$$

and

$$P_{\text{ee}}^{\text{1h}}(k) = \int dm_{\text{h}}\, n(m_{\text{h}}) \left(\frac{m_{\text{h}}}{\bar{\rho}_{\text{m}}} \right)^2 |u_{\text{gas}}(k, m_{\text{h}})|^2 \ . \tag{B10}$$

Similarly, the galaxy-electron cross power spectrum is

$$P_{\text{ge}}(k, \mu) = P_{\text{ge}}^{\text{2h}}(k, \mu) + P_{\text{ge}}^{\text{1h}}(k) \ , \tag{B11}$$

with

$$P_{\text{ge}}^{\text{2h}}(k, \mu) = \left[b_{\text{g}}(k) + f\mu^2 \right] \left[\int dm_{\text{h}}\, n(m_{\text{h}}) b_{\text{h}}(m_{\text{h}}) \frac{m_{\text{h}}}{\bar{\rho}_{\text{m}}} u_{\text{gas}}(k, m_{\text{h}}) \right] P_{\text{lin}}(k) \tag{B12}$$

and

$$P_{\text{ge}}^{\text{1h}}(k) = \int dm_{\text{h}}\, n(m_{\text{h}}) \frac{m_{\text{h}}}{\bar{\rho}_{\text{m}}} u_{\text{gas}}(k, m_{\text{h}}) \frac{\bar{N}_{\text{c}}(m_{\text{h}}) + \bar{N}_{\text{s}}(m_{\text{h}}) u_{\text{s}}(k, m_{\text{h}})}{\bar{n}_{\text{g}}} \ . \tag{B13}$$

Note that these quantities refer to the correlation between the observed galaxy density (which is affected by redshift space distortions) and the "true" electron density (which has no RSD contribution).

3. Integrand of kSZ angular power spectrum

It's helpful to rewrite C_{ℓ}^{kSZ} as an integral over redshift and halo mass, so that we can compare the relative contributions from different ranges of these two quantities (see Fig. 1). Combining Eqs. (6) and (8), and changing variables from χ to z, we can write

$$C_{\ell}^{\text{kSZ}} = \int_0^{z_*} dz \frac{d\chi}{dz} \frac{1}{\chi[z]^2} \tilde{K}(z)^2 \frac{1}{6\pi^2} \left[\int dk_{\text{L}}\, k_{\text{L}}^2 P_{vv}(k_{\text{L}}; z) \right] P_{\text{ee}}\left(\frac{\ell + 1/2}{\chi[z]}; z \right) \ . \tag{B14}$$

We can further rewrite P_{ee} as

$$P_{\text{ee}}(k; z) = \int d\log m_{\text{h}}\, m_{\text{h}} \left(\frac{dP_{\text{ee}}^{\text{2h}}(k; z)}{dm_{\text{h}}} + \frac{dP_{\text{ee}}^{\text{1h}}(k; z)}{dm_{\text{h}}} \right) \tag{B15}$$

with

$$\frac{dP_{\text{ee}}^{\text{2h}}(k; z)}{dm_{\text{h}}} = 2n(m_{\text{h}}) b_{\text{h}}(m_{\text{h}}) \frac{m_{\text{h}}}{\bar{\rho}_{\text{m}}} u_{\text{gas}}(k, m_{\text{h}}) \left[\int_{m_{\text{h,min}}}^{m_{\text{h}}} dm_{\text{h}}'\, n(m_{\text{h}}') b_{\text{h}}(m_{\text{h}}') \frac{m_{\text{h}}'}{\bar{\rho}_{\text{m}}} u_{\text{gas}}(k, m_{\text{h}}') \right] P_{\text{lin}}(k) \ , \tag{B16}$$

$$\frac{dP_{\text{ee}}^{\text{1h}}(k; z)}{dm_{\text{h}}} = n(m_{\text{h}}) \left(\frac{m_{\text{h}}}{\bar{\rho}_{\text{m}}} \right)^2 |u_{\text{gas}}(k, m_{\text{h}})|^2 \ . \tag{B17}$$

Thus, we can finally write

$$C_{\ell}^{\text{kSZ}} = \int_0^{z_*} dz \int d\log m_{\text{h}} \frac{d^2 C_{\ell}^{\text{kSZ}}}{dz\, d\log m_{\text{h}}} \tag{B18}$$

with

$$\frac{d^2 C_{\ell}^{\text{kSZ}}}{dz\, d\log m_{\text{h}}} = \frac{d\chi}{dz} \frac{1}{\chi[z]^2} \tilde{K}(z)^2 \frac{1}{6\pi^2} \left[\int dk_{\text{L}}\, k_{\text{L}}^2 P_{vv}(k_{\text{L}}; z) \right] m_{\text{h}} \left(\frac{dP_{\text{ee}}^{\text{2h}}(k; z)}{dm_{\text{h}}} + \frac{dP_{\text{ee}}^{\text{1h}}(k; z)}{dm_{\text{h}}} \right) \ , \tag{B19}$$

and the terms in parentheses given by Eqs. (B16)-(B17).

4. Galaxy bispectrum

To compute the shot noise contribution to the power spectrum of a kSZ template, we also need to compute the galaxy bispectrum (see Appendix A). This is a function of three 3d wavevectors $(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ constrained such that $\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 = 0$; in the absence of redshift space distortions (which we ignore for simplicity), the resulting triangle can be specified by 3 numbers, which we choose to be the magnitudes of each wavevector. The halo model description of the galaxy bispectrum can then be written as (e.g. [129])

$$B_{\text{ggg}}(k_1, k_2, k_3) = B_{\text{ggg}}^{\text{3h}}(k_1, k_2, k_3) + B_{\text{ggg}}^{\text{2h}}(k_1, k_2, k_3) + B_{\text{ggg}}^{\text{1h}}(k_1, k_2, k_3) \ . \tag{B20}$$

There is also a shot noise contribution, but we do not require this for our calculations in Appendix A, so we have not written it above.

The three-halo term is

$$B_{\text{ggg}}^{\text{3h}}(k_1, k_2, k_3) = b_{\text{eff}}(k_1) b_{\text{eff}}(k_2) b_{\text{eff}}(k_3) B_{\text{tree}}(k_1, k_2, k_3) \ , \tag{B21}$$

where the matter bispectrum at leading order ("tree-level") in perturbation theory is

$$B_{\text{tree}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = 2F_2(\boldsymbol{k}_1, \boldsymbol{k}_2)P_{\text{lin}}(k_1)P_{\text{lin}}(k_2) + 2 \text{ perms} \tag{B22}$$

and the symmetrized F_2 kernel is

$$F_2(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{5}{7} + \frac{1}{2}\hat{\boldsymbol{k}}_1 \cdot \hat{\boldsymbol{k}}_2 \left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{2}{7}\left(\hat{\boldsymbol{k}}_1 \cdot \hat{\boldsymbol{k}}_2\right)^2 . \tag{B23}$$

The two-halo term is

$$B_{\text{ggg}}^{2\text{h}}(k_1, k_2, k_3) = \left[\frac{1}{\bar{n}_{\text{g}}^2}\int dm_{\text{h}}\, n(m_{\text{h}})b_{\text{h}}(m_{\text{h}})\bigg\{ \langle N_{\text{c}}(m_{\text{h}})N_{\text{s}}(m_{\text{h}})\rangle \left(u_{\text{s}}(k_1, m_{\text{h}}) + u_{\text{s}}(k_2, m_{\text{h}})\right)\right.$$

$$\left. + \langle N_{\text{s}}(m_{\text{h}})\left[N_{\text{s}}(m_{\text{h}}) - 1\right]\rangle\, u_{\text{s}}(k_1, m_{\text{h}})u_{\text{s}}(k_2, m_{\text{h}})\bigg\}\right]b_{\text{eff}}(k_3)P_{\text{lin}}(k_3) + 2 \text{ perms} , \quad (B24)$$

while the one-halo term is

$$B_{\text{ggg}}^{1\text{h}}(k_1, k_2, k_3) = \frac{1}{\bar{n}_{\text{g}}^3}\int dm_{\text{h}}\, n(m_{\text{h}})\bigg\{ \langle N_{\text{c}}(m_{\text{h}})N_{\text{s}}(m_{\text{h}})(N_{\text{s}}(m_{\text{h}}) - 1)\rangle \left(u_{\text{s}}(k_1, m_{\text{h}})u_{\text{s}}(k_2, m_{\text{h}}) + 2 \text{ perms}\right)$$

$$+ \langle N_{\text{s}}(m_{\text{h}})(N_{\text{s}}(m_{\text{h}}) - 1)(N_{\text{s}}(m_{\text{h}}) - 2)\rangle\, u_{\text{s}}(k_1, m_{\text{h}})u_{\text{s}}(k_2, m_{\text{h}})u_{\text{s}}(k_3, m_{\text{h}})\bigg\} . \quad (B25)$$

For independent Poisson-distributed numbers of centrals and satellites, the expectation values in Eq. (B25) are given by

$$\langle N_{\text{c}}(m_{\text{h}})N_{\text{s}}(m_{\text{h}})(N_{\text{s}}(m_{\text{h}}) - 1)\rangle = \bar{N}_{\text{c}}(m_{\text{h}})\bar{N}_{\text{s}}(m_{\text{h}})^2 , \qquad \langle N_{\text{s}}(m_{\text{h}})(N_{\text{s}}(m_{\text{h}}) - 1)(N_{\text{s}}(m_{\text{h}}) - 2)\rangle = \bar{N}_{\text{s}}(m_{\text{h}})^3 . \tag{B26}$$

Appendix C: Halo occupation modelling

In this section, we record the details of the HOD models used for the survey-specific forecasts in Sec. IV. Note that we write parameters with dimension of mass either in $h^{-1}M_\odot$ or $M_\odot$ units, depending on what was used in the original reference, but we translate all parameters to $M_\odot$ units in our numerical computations in this work.

1. BOSS

For BOSS, we use the following 5-parameter HOD from Ref. [56]:

$$\bar{N}_{\text{c}}(m_{\text{h}}) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\log m_{\text{h}} - \log M_{\text{min}}}{\sigma_{\log M}}\right)\right] , \tag{C1}$$

$$\bar{N}_{\text{s}}(m_{\text{h}}) = \bar{N}_{\text{c}}(m_{\text{h}})\left(\frac{M}{M_1}\right)^\alpha e^{-M_{\text{cut}}/m_{\text{h}}} , \tag{C2}$$

with the best-fit parameter values from the joint fit to the projected correlation function (w_p) and void probability function (P_0):

$$M_{\text{min}} = 10^{13.18}\, h^{-1}M_\odot , \quad \sigma_{\log M} = 0.55 , \quad M_1 = 10^{14.28}\, h^{-1}M_\odot , \quad \alpha = 1.12 , \quad M_{\text{cut}} = 10^{4.87}\, h^{-1}M_\odot . \tag{C3}$$

2. DESI BGS

For the DESI BGS sample, we use the HOD from Ref. [55] with the assumption that the scatter set by the $\sigma_{\log M}$ parameter is Gaussian:

$$\bar{N}_{\rm c}(m_{\rm h}) = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log m_{\rm h} - R\log M_{\rm min}}{\sigma_{\log M}}\right)\right] , \tag{C4}$$

$$\bar{N}_{\rm s}(m_{\rm h}) = \bar{N}_{\rm c}(m_{\rm h})\left(\frac{m_{\rm h} - M_0}{(M_1')^R}\right)^\alpha , \tag{C5}$$

with fitted parameter values corresponding to absolute r-band magnitude $^{0.1}M_r - 5\log h = -20.5$:

$$M_{\rm min} = 10^{12.2}\,h^{-1}M_\odot , \quad \sigma_{\log M} = 0.15 , \quad M_1' = 10^{13.5}\,h^{-1}M_\odot , \quad \alpha = 1.05 , \quad M_0 = 10^{12}\,h^{-1}M_\odot . \tag{C6}$$

We fix R by demanding that the $\bar{n}_{\rm g}$ expression in Eq. (B4) evaluates to the value computed from Table 2.5 of Ref. [63] over $0.05 \leq z \leq 0.45$. The resulting R values are well fit by $R(z) = 0.88 + 0.23z + 0.76z^2$.

3. DESI LRG

For the DESI LRG sample, we use the HOD from Ref. [51]:

$$\bar{N}_{\rm c}(m_{\rm h}) = \frac{f_{\rm ic}}{2}{\rm erfc}\left(\frac{R\log M_{\rm cut} - \log m_{\rm h}}{\sqrt{2}\sigma_{\log M}}\right) , \tag{C7}$$

$$\bar{N}_{\rm s}(m_{\rm h}) = \bar{N}_{\rm c}(m_{\rm h})\left(\frac{m_{\rm h} - \kappa M_{\rm cut}}{M_1^R}\right)^\alpha , \tag{C8}$$

with

$$M_{\rm cut} = 10^{12.7}\,h^{-1}M_\odot , \quad \sigma_{\log M} = 0.2 , \quad M_1 = 10^{13.6}\,h^{-1}M_\odot , \quad \alpha = 1.15 , \quad \kappa = 0.08 , \quad f_{\rm ic} = 0.8 . \tag{C9}$$

We fix R such that $\bar{n}_{\rm g}$ evaluates to the target density of $5\times10^{-4}\,{\rm Mpc}^{-3}$ quoted in Ref. [51] over $0.6 \leq z \leq 1.1$. The resulting R values are well fit by $R(z) = 1.05 - 0.033z$.

4. DESI ELG

For the DESI ELG sample, we use the HOD from Ref. [51]:

$$\bar{N}_{\rm c}(m_{\rm h}) = R\left\{2A\phi(\log m_{\rm h})\Phi(\log m_{\rm h}) + \frac{1}{2Q}\left[1 + {\rm erf}\left(\frac{\log m_{\rm h} - \log M_{\rm cut}}{0.01}\right)\right]\right\} , \tag{C10}$$

$$\bar{N}_{\rm s}(m_{\rm h}) = R\left(\frac{m_{\rm h} - \kappa M_{\rm cut}}{M_1}\right)^\alpha , \tag{C11}$$

where

$$\phi(\log m_{\rm h}) = \frac{1}{\sqrt{2\pi\sigma_{\rm M}^2}}\exp\left[-\frac{(\log m_{\rm h} - \log M_{\rm cut})^2}{2\sigma_{\rm M}^2}\right] , \tag{C12}$$

$$\Phi(\log m_{\rm h}) = \frac{1}{2}\left[1 + {\rm erf}\left(\gamma\frac{\log m_{\rm h} - \log M_{\rm cut}}{\sqrt{2}\sigma_{\rm M}}\right)\right] , \tag{C13}$$

$$A = p_{\rm max} - \frac{1}{Q} \tag{C14}$$

and

$$p_{\text{max}} = 0.075 \ , \ Q = 95 \ , \ M_{\text{cut}} = 10^{11.9} \, h^{-1} M_\odot \ , \ \sigma_M = 0.5 \ , \ \gamma = 5 \ , \ M_1 = 10^{14.2} \, h^{-1} M_\odot \ , \ \alpha = 0.65 \ , \ \kappa = 1.35 \ . \tag{C15}$$

We use $R = 1$ for our baseline modelling, but in cases where we wish to ensure that the predicted number density matches a desired input, we fix R appropriately at each redshift.

5. DESI QSO

For the DESI QSO sample, we use the error-function–based HOD from Ref. [49]:

$$\bar{N}_c(m_h) = \frac{1}{2} p_{\text{max}} \text{erfc} \left(\frac{R \log M_c - \log m_h}{\sqrt{2} \log(e) \sigma_M} \right) \ , \tag{C16}$$

$$\bar{N}_s(m_h) = \left(\frac{m_h - \kappa M_c}{(M_1')^R} \right)^\alpha \ , \tag{C17}$$

with

$$M_c = 10^{12.21} \, h^{-1} M_\odot \ , \quad \sigma_M = 0.6 \ , \quad M_1 = 10^{14.09} \, h^{-1} M_\odot \ , \quad \kappa = 1.0 \ , \quad \alpha = 0.39 \ , \quad p_{\text{max}} = 0.033 \ . \tag{C18}$$

We fix R such that $\bar{n}_g$ evaluates to the QSO density given in Ref. [63] over $0.65 \leq z \leq 1.85$. The resulting R values are well fit by $R(z) = 1.16 - 0.040z$.

6. Roman

For the Hα sample of ELGs from the Roman High Latitude Spectroscopic Survey, we fit the following parameterized model to measurements from the mock catalog in Ref. [70]:

$$\bar{N}_c(m_h) = R \left\{ A_1 \exp \left[-\frac{(\log m_h - \log M_c)^2}{2\sigma_{\log M}^2} \right] + A_2 \left[1 + \text{erf} \left(\frac{\log m_h - \log M_c}{\sigma_{\log M}} \right) \right] \right\} \ ,$$

$$\bar{N}_s(m_h) = \begin{cases} R A_2 \left(\frac{m_h}{M_1} \right)^{\alpha_1} \ , & m_h < M_1 \ , \\ R A_2 \left(\frac{m_h}{M_1} \right)^{\alpha_2} \ , & m_h \geq M_1 \ . \end{cases} \tag{C19}$$

The central occupation is composed of a peak with amplitude A_1 at $m_h = M_c$ and a smooth transition to a plateau with amplitude $2A_2$ at higher masses, while the satellite occupation is a broken power law with tilt α_1 for $m_h < M_1$ and α_2 for $m_h \geq M_1$. Note that the measured HODs from Ref. [70] typically have a double-peaked structure as a function of halo mass, but the higher-mass peak is typically weaker and its physical origin is somewhat unclear, so our parameterization only incorporates the lower-mass peak. We fix the free parameters to approximately match the measurements shown in Figure 8 of Ref. [70] for $1.0 < z < 1.1$, $A_V = 1.65$, $f_{\text{lin},1} = 10^{-16} \, \text{erg s}^{-1} \, \text{cm}^{-2}$, and $f_{\text{lin},2R} = 0.5$, resulting in

$$A_1 = 0.2 \ , \quad A_2 = 0.01 \ , \quad M_c = 11.9 M_\odot \ , \quad \sigma_{\log M} = 0.2 \ , \quad M_1 = 12.3 M_\odot \ , \quad \alpha_1 = 2.5 \ , \quad \alpha_2 = 0.5 \ . \tag{C20}$$

We fix R by such that the predicted galaxy number density matches that from Ref. [69] corresponding to fluxes $> 10^{16} \, \text{erg s}^{-1} \text{cm}^{-2}$ and dust attenuation parameter $A_V = 1.65$. The resulting R values are well fit by $R(z) = 5.41 - 2.77z$.

7.   MegaMapper

For the MegaMapper LBG sample, we use the "linear HOD model" from Ref. [73]:

$$\bar{N}_{\rm c}(m_{\rm h}) = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log m_{\rm h} - \log M_{\rm min}}{\sqrt{2}\sigma_{\log M}}\right)\right] \, , \tag{C21}$$

$$\bar{N}_{\rm s}(m_{\rm h}) = \bar{N}_{\rm c}(m_{\rm h})\left(\frac{m_{\rm h} - M_{\rm cut}}{M_{\rm sat}}\right)^{\alpha} \, . \tag{C22}$$

We use the best-fit parameters corresponding to $z \approx 3.8$ with $m_{\rm UV}^{\rm th} = 24.5$:

$$M_{\rm min} = M_{\rm cut} = 10^{12.22} M_{\odot} \, , \quad \sigma_{\log M} = 0.2 \, , \quad M_{\rm sat} = 10^{14.23} M_{\odot} \, , \quad \alpha = 1.0 \, . \tag{C23}$$

Note that we have chosen to set $M_{\rm cut}$ equal to $M_{\rm min}$, which differs from the choice made in Ref. [73], but the precise value of $M_{\rm cut}$ has a minimal impact on our results.

Appendix D: Modelling for 21 cm intensity mapping

In this appendix, we provide the details of our approach to modelling 21 cm observations of large-scale structure.

1.   Halo model

a.   Formalism

The halo model framework we use for modelling the distribution of HI generally follows that in Appendix B, with a few differences adapted from the formalism in Ref. [128]. HI halo models are often written as predictions for the statistics of the observed brightness temperature, but for consistency with our other forecasts, we write predictions for the HI overdensity instead of the brightness temperature fluctuations. In this approach, the halo occupation functions $\bar{N}_{\rm c}(m_{\rm h})$ and $\bar{N}_{\rm s}(m_{\rm h})$ are replaced by a HI mass-halo mass relation $m_{\rm HI}(m_{\rm h})$, and a HI-specific halo density profile $u_{\rm HI}$ is used, such that the linear bias from Eq. (B3) transforms into

$$b_{\rm HI}(k) = \frac{1}{\bar{\rho}_{\rm HI}}\int dm_{\rm h}\, n(m_{\rm h})b_{\rm h}(m_{\rm h})m_{\rm HI}(m_{\rm h})u_{\rm HI}(k, m_{\rm h}) \tag{D1}$$

and the one-halo term from Eq. (B5) becomes

$$P_{\rm HI}^{\rm 1h}(k) = \frac{1}{\bar{\rho}_{\rm HI}^2}\int dm_{\rm h}\, n(m_{\rm h})m_{\rm HI}(m_{\rm h})^2 u_{\rm HI}(k, m_{\rm h})^2 \, , \tag{D2}$$

where

$$\bar{\rho}_{\rm HI} = \int dm_{\rm h}\, n(m_{\rm h})m_{\rm HI}(m_{\rm h}) \, . \tag{D3}$$

Note that the two-halo term in Eq. (B2) retains the same form.

b.   HI mass-halo mass relation

For the HI mass-halo mass relation $m_{\rm HI}(m_{\rm h}, z)$, we consider two fitting functions from the literature. The first is from Ref. [52], and has been fit to measurements of HI from the IllustrisTNG simulations:

$$m_{\rm HI}(m_{\rm h}) = m_0 \left(\frac{m_{\rm h}}{m_{\rm min}}\right)^{\alpha} \exp\left[-\left(\frac{m_{\rm min}}{m_{\rm h}}\right)^{0.35}\right] \, , \tag{D4}$$

with the best-fit "FoF-SO" values for α, m_0, and $m_{\min}$ listed for $z = 0$ to 5 in their Table 1.

The second is from Ref. [82], and has been fit to a variety of HI observations (resolved low-redshift galaxies, intensity-mapping determinations of the mean HI density, and higher-redshift damped Lyman-α absorbers):

$$m_{\mathrm{HI}}(m_{\mathrm{h}}, z) = \alpha f_{\mathrm{H},c} m_{\mathrm{h}} \left(\frac{m_{\mathrm{h}}}{10^{11} h^{-1} M_\odot} \right)^{\beta} \exp\left[- \left(\frac{v_{c,0}}{v_c(m_{\mathrm{h}}, z)} \right)^{3} \right] , \tag{D5}$$

with $f_{\mathrm{H},c} = \Omega_{\mathrm{b}}(1 - Y_{\mathrm{p}})/\Omega_{\mathrm{m}}$, halo virial velocity $v_c(m_{\mathrm{h}})$ given by

$$v_c(m_{\mathrm{h}}, z) = \sqrt{\frac{G m_{\mathrm{h}}}{r_{\mathrm{vir}}(m_{\mathrm{h}}, z)}} , \tag{D6}$$

and best-fit parameters

$$\alpha = 0.09 , \quad \log\left(\frac{v_{c,0}}{\mathrm{km\,s^{-1}}}\right) = 1.56 , \quad \beta = -0.58 . \tag{D7}$$

c. HI density profile

Likewise, we consider two forms of the HI density profile. The first, which we use for our main computations, is from Ref. [52]:

$$\rho_{\mathrm{HI}}(r, m_{\mathrm{h}}, z) = \frac{\rho_0}{r^{\alpha_*}} \exp\left(-\frac{r_0}{r}\right) , \tag{D8}$$

using the best-fit parameters for $z = 0$ to 5 and $m = 10^9 h^{-1} M_\odot$ to $10^{15} h^{-1} M_\odot$ listed in their Table 2. We numerically evaluate a fast Fourier transform of this profile to obtain $u_{\mathrm{HI}}(k, m_{\mathrm{h}}, z)$.

The second profile is from Ref. [82]:

$$\rho_{\mathrm{HI}}(r, m_{\mathrm{h}}, z) = \rho_0(m_{\mathrm{h}}, z) \exp\left(-\frac{r}{r_{\mathrm{s}}(m_{\mathrm{h}}, z)}\right) , \tag{D9}$$

with halo scale radius $r_{\mathrm{s}}(m_{\mathrm{h}}, z) = r_{\mathrm{vir}}(m_{\mathrm{h}}, z)/c_{\mathrm{HI}}(m_{\mathrm{h}}, z)$ using an HI-specific concentration-mass relation:

$$c_{\mathrm{HI}}(m_{\mathrm{h}}, z) = c_{\mathrm{HI},0} \left(\frac{m_{\mathrm{h}}}{10^{11} M_\odot} \right)^{-0.109} \frac{4}{(1 + z)^\gamma} , \tag{D10}$$

with

$$c_{\mathrm{HI},0} = 28.65 , \quad \gamma = 1.45 . \tag{D11}$$

The normalization $\rho_0(m_{\mathrm{h}}, z)$ is fixed so that the HI mass enclosed within the virial radius is equal to $m_{\mathrm{HI}}(m_{\mathrm{h}}, z)$. In Fourier space, the normalized profile u_{HI} is then given by

$$u_{\mathrm{HI}}(k, m_{\mathrm{h}}, z) = \frac{2}{[1 + k^2 r_{\mathrm{s}}(m_{\mathrm{h}}, z)^2]^2} . \tag{D12}$$

In Appendix D 4, we compare the results if either of these profiles is used.

2. Instrumental noise

In 21 cm intensity mapping, the dominant noise in measurements of the power spectrum is typically the instrumental noise associated with finite observing time and properties of the instrument, rather than the intrinsic shot noise of 21 cm-emitting objects. Thus, we replace the shot power term in Eq. (12) with an "effective" shot noise that includes

both instrumental and Poisson contributions:

$$P_{\mathrm{HI}}^{\mathrm{shot,eff}}(k_\perp, z) = \frac{P_{\mathrm{N}}(k_\perp, z)}{\bar{T}_{\mathrm{HI}}(z)^2} + \frac{1}{n_{\mathrm{HI}}(z)} \ . \tag{D13}$$

The instrumental noise P_{N} is associated with 21 cm brightness temperature fluctuations, such that dividing it by $\bar{T}_{\mathrm{HI}}(z)^2$ converts into the noise power spectrum associated with the HI overdensity. We compute the noise power spectrum for PUMA, mean brightness temperature, and intrinsic shot noise as described in the appendices of Ref. [78] and implemented in the PUMANoise code[12]. From Eq. D13, we can define an effective number density $\bar{n}_{\mathrm{eff}}(k_\perp, z) \equiv P_{\mathrm{HI}}^{\mathrm{shot,eff}}(k_\perp, z)^{-1}$, which we plot in Fig. 11.

3. Foregrounds

The need to remove bright foregrounds, dominated by Galactic synchrotron emission and extragalactic radio sources, from 21 cm observations is expected to impose two restrictions on the modes that are available for cosmological analysis[13]:

1. Modes with $|k_\parallel| < k_{\parallel \mathrm{min}}$ will be indistinguishable from smooth-spectrum foregrounds, and will therefore be filtered out. We use $k_{\parallel \mathrm{min}} = 0.03\,\mathrm{Mpc}$, following Ref. [131], who found that an optimal foreground cleaning approach imposed $k_{\parallel \mathrm{min}} \approx 0.02 h^{-1}\,\mathrm{Mpc}$ when applied to simulations.

2. In interferometric observations, spectrally-smooth foreground power will leak beyond pure low-$k_\parallel$ modes into higher-$k_\parallel$ modes in a baseline-dependent way, creating a so-called "foreground wedge" of contamination (e.g. [84–86]). This is described by by $|k_\parallel| < \beta(z) k_\perp$ where

$$\beta(z) \equiv \frac{\chi(z) H(z)}{c(1+z)} \sin \theta_{\mathrm{w}}(z) \ . \tag{D14}$$

The extent of the contamination is set by $\theta_{\mathrm{w}}(z)$, the maximum angle away from the receiver's phase center at which a spectrally smooth sky signal can contaminate higher-$k_\parallel$ modes. Following Ref. [78], we set this to 3 times the width of the PUMA primary beam, i.e. $\theta_{\mathrm{w}}(z) = 3 \times 1.2\lambda(z)/D_{\mathrm{eff}}$, with $\lambda(z) \equiv 21(1+z)\,\mathrm{cm}$ and $D_{\mathrm{eff}} = 5\,\mathrm{m}$. We explore our sensitivity to this choice, and the choice of $k_{\parallel \mathrm{min}}$, in Appendix D 4.

These restrictions will apply separately to templates for velocity and electron density fields constructed from 21 cm observations. However, the total kSZ template is constructed from the product of these templates (recall Eqs. 10-11 and 13), which becomes a convolution in Fourier space:

$$\hat{q}_r(\boldsymbol{k}_{\mathrm{S}}, z) = \int_{\boldsymbol{k}_{\mathrm{L}}} \eta(\boldsymbol{k}_{\mathrm{L}}, z) \epsilon(\boldsymbol{k}_{\mathrm{S}} - \boldsymbol{k}_{\mathrm{L}}, z) \ . \tag{D15}$$

Thus, low-$k_\parallel$ modes of $\hat{q}_r(\boldsymbol{k}_{\mathrm{S}}, z)$ are not obscured by foregrounds, since they can be sourced by pairs of η and ϵ modes whose line-of-sight wavenumbers are individually much larger (i.e. it is possible that $|k_{\mathrm{S}\parallel}| < k_{\parallel \mathrm{min}}$ while having $|k_{\mathrm{L}\parallel}|$ and $|k_{\mathrm{S}\parallel} - k_{\mathrm{L}\parallel}|$ both greater than $k_{\parallel \mathrm{min}}$). We will omit redshift arguments in what follows.

To implement these foreground restrictions in computations, we start by recalling the cross power spectrum of $\hat{q}_r$ and q_r from Eq. (15), prior to taking the squeezed limit of $P_{e\epsilon}$ (the discussion below also applies to the auto spectrum of $\hat{q}_r$):

$$P_{\hat{q}_r q_r}(\boldsymbol{k}_{\mathrm{S}}) \approx \int \frac{d^3 \boldsymbol{k}_{\mathrm{L}}}{(2\pi)^3} P_{v_r \eta}(\boldsymbol{k}_{\mathrm{L}}) P_{e\epsilon}(\boldsymbol{k}_{\mathrm{S}} - \boldsymbol{k}_{\mathrm{L}}) \ . \tag{D16}$$

---

[12] https://github.com/slosar/PUMANoise
[13] We note that observational and methodological work is underway to understand in more detail which modes will be eliminated by 21 cm foreground cleaning and how to incorporate this into forecasts (e.g. [130]), but in this work we adopt assumptions which have been motivated by previous studies [79]. We also note that reconstruction techniques are under development which could allow for the recovery of foreground-obscured modes (e.g. [41, 89, 90]), but such reconstructed modes will have different noise than directly-observed modes, so we only consider directly-observed modes in this work.

We must set the integrand to zero *outside* of the region defined by the union of the following four conditions:

$$|k_{\mathrm{L}\parallel}| > k_{\parallel\mathrm{min}} , \quad |k_{\mathrm{L}\parallel} - k_{\mathrm{S}\parallel}| > k_{\parallel\mathrm{min}} , \quad |k_{\mathrm{L}\parallel}| > \beta k_{\mathrm{L}\perp} , \quad |k_{\mathrm{S}\parallel} - k_{\mathrm{L}\parallel}| > \beta |\boldsymbol{k}_{\mathrm{S}\perp} - \boldsymbol{k}_{\mathrm{L}\perp}| . \tag{D17}$$

With $k_{\mathrm{L}\parallel} = k_{\mathrm{L}}\mu_{\mathrm{L}}$ and $k_{\mathrm{L}\perp} = k_{\mathrm{L}}\sqrt{1 - \mu_{\mathrm{L}}^2}$ and analogously for $\boldsymbol{k}_{\mathrm{S}}$, and also defining $\cos\phi_{\perp} \equiv \hat{\boldsymbol{k}}_{\mathrm{S}\perp} \cdot \hat{\boldsymbol{k}}_{\mathrm{L}\perp}$, the fourth condition can be written as $\cos\phi_{\perp} > \gamma(k_{\mathrm{S}}, \mu_{\mathrm{S}}; k_{\mathrm{L}}, \mu_{\mathrm{L}})$, where

$$\gamma(k_{\mathrm{S}}, \mu_{\mathrm{S}}; k_{\mathrm{L}}, \mu_{\mathrm{L}}) \equiv \frac{\beta^2\left[(1 - \mu_{\mathrm{S}}^2)k_{\mathrm{S}}^2 + (1 - \mu_{\mathrm{L}}^2)k_{\mathrm{L}}^2\right] - (\mu_{\mathrm{S}}k_{\mathrm{S}} - \mu_{\mathrm{L}}k_{\mathrm{L}})^2}{2\beta^2\sqrt{1 - \mu_{\mathrm{S}}^2}\sqrt{1 - \mu_{\mathrm{L}}^2}k_{\mathrm{S}}k_{\mathrm{L}}} . \tag{D18}$$

We can integrate this constraint out of Eq. (D16) by approximating $P_{e\epsilon}(\boldsymbol{k}_{\mathrm{S}} - \boldsymbol{k}_{\mathrm{L}}) \approx P_{e\epsilon}(\boldsymbol{k}_{\mathrm{S}})$ and using

$$\mathcal{C}_{\perp}(k_{\mathrm{S}}, \mu_{\mathrm{S}}; k_{\mathrm{L}}, \mu_{\mathrm{L}}) \equiv \int_0^{2\pi} d\phi_{\perp} \, \Theta(\cos\phi_{\perp} - \gamma) = \begin{cases} 2\pi , & \gamma \leq -1 , \\ 2\cos^{-1}\gamma , & -1 < \gamma < 1 , \\ 0 , & \gamma \geq 1 , \end{cases} \tag{D19}$$

where $\Theta(\cdots)$ is a step function. Meanwhile, the first three conditions in Eq. (D17) are equivalent to setting the integrand to zero if any of the following conditions *are* satisfied:

$$|\mu_{\mathrm{L}}| < \frac{k_{\parallel\mathrm{min}}}{k_{\mathrm{L}}} , \quad |k_{\mathrm{L}}\mu_{\mathrm{L}} - k_{\mathrm{S}}\mu_{\mathrm{S}}| < k_{\parallel\mathrm{min}} , \quad |\mu_{\mathrm{L}}| < \frac{\beta}{\sqrt{1 + \beta^2}} , \tag{D20}$$

and we enforce these numerically when evaluating Eq. (D16).

Finally, we compute $C_\ell^{\hat{q}_r q_r}$ in the Limber approximation, which requires taking the $\mu_{\mathrm{S}} \to 0$ limit of Eq. (D16):

$$P_{\hat{q}_r q_r}^{21\mathrm{cm}}(k_{\mathrm{S}}, 0) = \frac{1}{8\pi^3} \int dk_{\mathrm{L}} \, k_{\mathrm{L}}^2 \int_{-1}^1 d\mu_{\mathrm{L}} \, \mu_{\mathrm{L}}^2 \frac{P_{gv}(k_{\mathrm{L}}, \mu_{\mathrm{L}})^2}{P_{gg}^{\mathrm{tot}}(k_{\mathrm{L}}, \mu_{\mathrm{L}})} \frac{P_{ge}(k_{\mathrm{S}}, 0)^2}{P_{gg}^{\mathrm{tot}}(k_{\mathrm{S}}, 0)} \mathcal{C}_{\perp}(k_{\mathrm{S}}, 0; k_{\mathrm{L}}, \mu_{\mathrm{L}}) \, \mathcal{C}_{\parallel}(k_{\mathrm{S}}, 0; k_{\mathrm{L}}, \mu_{\mathrm{L}}) , \tag{D21}$$

where $\mathcal{C}_{\parallel}(k_{\mathrm{S}}, \mu_{\mathrm{S}}; k_{\mathrm{L}}, \mu_{\mathrm{L}})$ implements the conditions in Eq. (D20). Note that we take the $\mu_{\mathrm{S}} \to 0$ limit *after* accounting for foreground mode cuts as described above, and therefore we only consider modes of $\hat{q}_r$ that survive these cuts.

4. Dependence of forecasts on HI profile and foreground cuts

In Fig. 18, we compare kSZ template forecasts assuming the simulation-based HI profile from Ref. [52] with those assuming the exponential profile from Ref. [82]. We find that this choice does not qualitatively change the results: in particular, we see a negligible change at $\ell \lesssim 2000$ that becomes more important at larger multipoles, with as much as a 10% (15%) change in the results for PUMA-32k (PUMA-5k) at $\ell \sim 4000$, and a 30% (70%) change at $\ell \sim 10000$. This uncertainty in our forecasts is subdominant to that from the uncertain form of the $m_{\mathrm{HI}}(m_{\mathrm{h}})$ relation (this uncertainty is also lower at lower multipoles).

In Fig. 19, we explore how our forecasts depend on assumptions about 21 cm foregrounds. Each column corresponds to a different choice of $k_{\parallel\mathrm{min}}$, while the different linestyles for the grey curves denote different assumptions about the foreground wedge: an angular extent of 3 times the primary beam width of a single 6 m dish, 1 times the primary beam width (a more optimistic case), or the full sky visible above the horizon (a more pessimistic case). Between the most optimistic and pessimistic assumptions for the grey curves, we find roughly an order of magnitude variation at $\ell \sim 1000$ and a factor of 5 at $\ell \sim 10000$, indicating that the performance of a 21 cm-based kSZ template will be very sensitive to the level of foreground cleaning.

Appendix E: CMB Forecasts

We model the CMB noise including the pink and white noise components as

$$N_\ell = \Delta_T^2 \exp\left[\ell(\ell+1)\frac{\theta_{\mathrm{FWHM}}^2}{8\log(2)}\right]\left[1 + (\ell_{\mathrm{knee}}/\ell)^{\alpha_{\mathrm{knee}}}\right] \tag{E1}$$

where we set the beam full-width half-maximum (FWHM), the temperature noise RMS Δ_T, and the $\{\alpha_{\mathrm{knee}}, \ell_{\mathrm{knee}}\}$ parameters which model the effect of the Earth's atmosphere to match upcoming experiments as given in Table II.
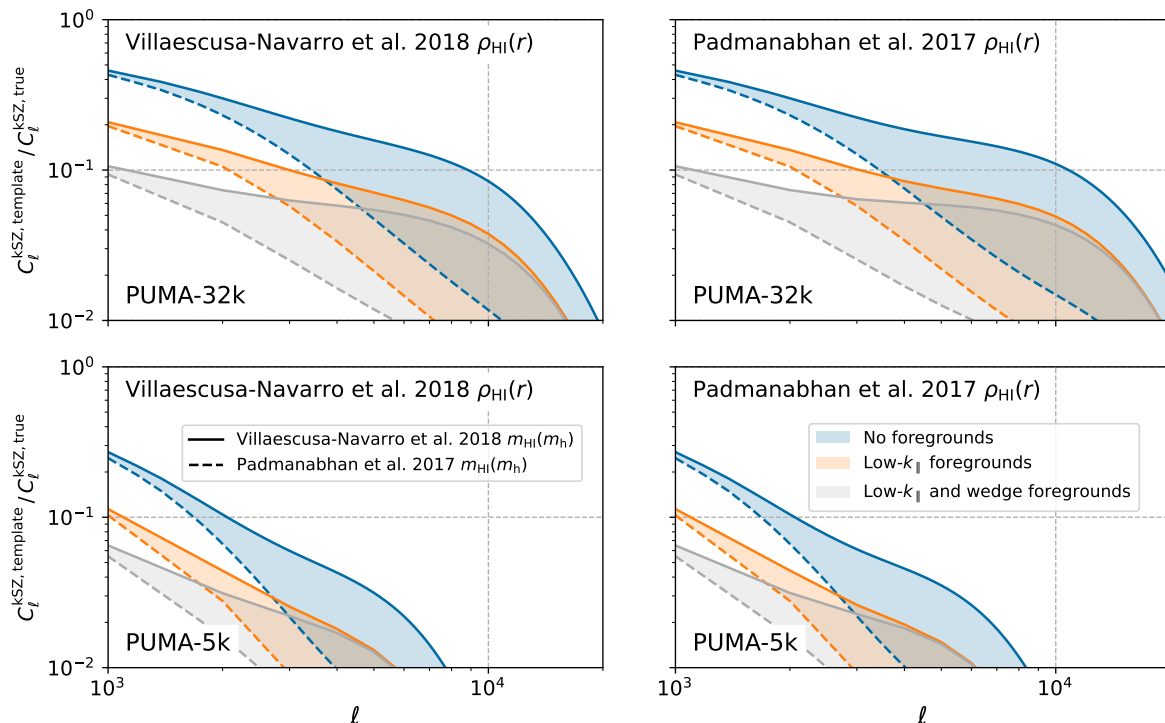
FIG. 18. Forecasts for kSZ templates constructed from 21 cm intensity mapping measurements from PUMA, comparing our baseline forecasts that assume the simulation-based HI density profile from Ref. [52] (*left panels*) with forecasts that use the the exponential profile from Ref. [82] (*right panels*). The choice of HI profile has negligible effect at lower multipoles, while it can affect the results by as much as 30% (70%) for PUMA-32k (PUMA-5k) at $\ell \sim 10000$.

	Beam FWHM		Noise RMS (μK-arcmin)	
	S4	HD	S4	HD
39 GHz	5.1′	0.94′	12.4	3.4
93 GHz	2.2′	0.42′	2.0	0.7
145 GHz	1.4′	0.25′	2.0	0.8
225 GHz	1.0′	0.17′	6.9	2.0
280 GHz	0.9′	0.13′	16.7	2.7

TABLE II. Inputs to ILC noise for the CMB configurations. We have chosen the temperature noise RMS and the beam parameters to approximately match CMB-S4 and CMB-HD. We account for the effect of Earth's atmosphere by setting $\ell_{\rm knee} = 100$ and $\alpha_{\rm knee} = -3$. We define the polarization noise as $\Delta_E = \Delta_B = \sqrt{2}\Delta_T$.

The contributions to the signal in the millimeter wavelength include clustered CIB, Poisson CIB and tSZ foregrounds, and the black-body late-time kSZ following Ref. [132] based on fits to data from Ref. [133]. We do not include the correlation between tSZ and CIB. We also include radio sources in the 39, 93 and 145 GHz channels, using the flux-limit-dependent radio-source power model from Ref. [134]. For CMB-S4, we assume flux limits of 10, 7 and 10 mJy, respectively in those channels. For CMB-HD, we assume flux limits of 2, 1 and 1 mJy respectively. In addition, we include the lensed CMB black-body contribution which we calculate using CAMB [135], and the kSZ signal from reionization following Ref. [94].

We calculate the total covariance between two frequency channels at each multipole ℓ, $\boldsymbol{C}_\ell^{ij}$, taking into account the signal components correlated across frequencies, as well as the uncorrelated beam-deconvolved noise, indexed by i. The resulting minimum-variance standard ILC noise for the black-body signal (lensed CMB+kSZ) including all other contributions is then given by

$$N_\ell = \left[\sum_{ij} \left(\boldsymbol{C}^{-1} \right)_\ell^{ij} \right]^{-1}. \tag{E2}$$
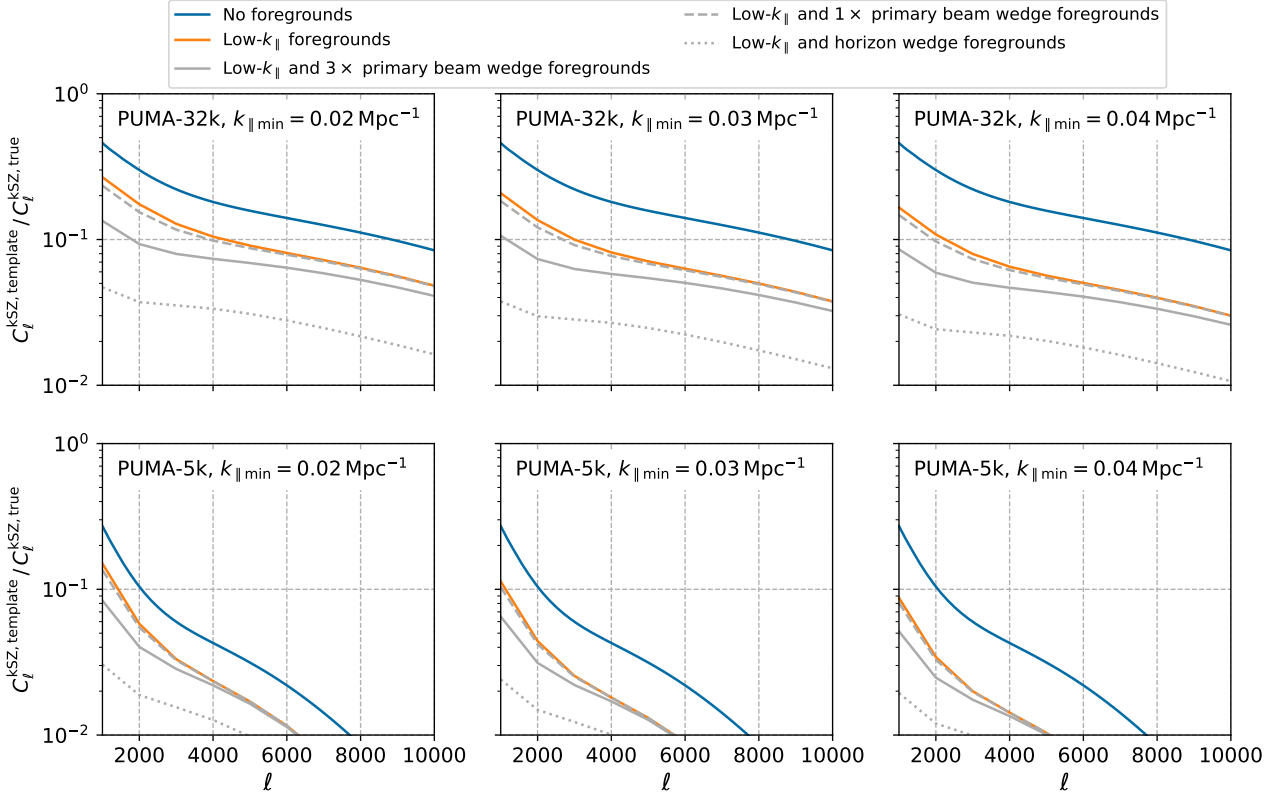
FIG. 19. Dependence of a 21 cm-based kSZ template on assumptions about foreground cleaning. For simplicity, we only show results assuming the HI mass-halo mass relation and HI density profile from Ref. [52]. Different columns show different assumptions about the minimum $k_{\parallel}$ that can be cleaned of foregrounds, while different linestyles of the grey curves show different assumptions about the severity of the foreground wedge. We find a factor of 5 (at higher ℓ) to 10 (at lower ℓ) difference between the most optimistic and pessimistic cases that include the wedge.

Appendix F: Forecasts for moving-lens signal

Ref. [14] derives an optimal matched filter for measuring transverse velocities with the moving-lens effect, assuming a known form for the halo density profile. In Sec. VB2, we present a forecast that uses a modified form of this procedure to obtain binned information about the moving-lens profile (and therefore the halo density profile) itself. The impact of the moving-lens effect on the CMB has been written in Eqs. (36) to (39).

In the form derived in Ref. [14], an optimal filter $\tilde{\Psi}_i$ for the moving-lens signal along transverse direction i is given by

$$\tilde{\Psi}_i(\vec{\ell}) = \sigma_i \frac{\tilde{\mathcal{M}}_i(\vec{\ell})}{\tilde{C}_{\ell}^{TT}} \, , \tag{F1}$$

where i is either of two unit vectors, which we call $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$, which span the $\boldsymbol{x}_{\perp}$ plane. We have written $\mathcal{M}_x = \hat{\boldsymbol{x}} \cdot \boldsymbol{\mathcal{M}}$ and $\mathcal{M}_y = \hat{\boldsymbol{y}} \cdot \boldsymbol{\mathcal{M}}$. The measurement noise σ_i is given by

$$\sigma_i = a_0^{-2} \int \frac{\mathrm{d}^2 \boldsymbol{\ell}}{(2\pi)^2} |\tilde{\Psi}_i(\boldsymbol{\ell})|^2 \tilde{C}_{\ell}^{TT} \, . \tag{F2}$$

In our binned forecasts, we use the same approach, but for a bin centered at $x_{\perp j}$, we suppose that we have applied a circular ring mask $\mathcal{R}_j(x_{\perp})$ with outer (inner) radius $x_{\perp j} + \Delta x_{\perp}/2$ $(x_{\perp j} - \Delta x_{\perp}/2)$ around the halo center, such that $\tilde{\mathcal{M}}_i(\boldsymbol{\ell})$ in the numerator of Eq. (F1) is the harmonic transform of $\mathcal{R}_j(x_{\perp})\mathcal{M}_i(\boldsymbol{x}_{\perp})$ instead of $\mathcal{M}_i(\boldsymbol{x}_{\perp})$. We then compute the modified measurement noise using Eq. (F2), taking $\sigma \equiv \sigma_x = \sigma_y$. Our results are demonstrated in

Fig. 16.

[1] Y. B. Zeldovich and R. A. Sunyaev, Astrophys. Space Sci. **4**, 301 (1969).
[2] R. A. Sunyaev and Y. B. Zeldovich, Comments Astrophys. Space Phys. **4**, 173 (1972).
[3] R. A. Sunyaev and Y. B. Zeldovich, Ann. Rev. Astron. Astrophys. **18**, 537 (1980).
[4] S. Y. Sazonov and R. A. Sunyaev, Mon. Not. Roy. Astron. Soc. **310**, 765 (1999), arXiv:astro-ph/9903287.
[5] T. Mroczkowski *et al.*, Space Sci. Rev. **215**, 17 (2019), arXiv:1811.02310 [astro-ph.CO].
[6] N. Itoh, Y. Kohyama, and S. Nozawa, Astrophys. J. **502**, 7 (1998), arXiv:astro-ph/9712289.
[7] W. R. Coulton, A. Ota, and A. van Engelen, Phys. Rev. Lett. **125**, 111301 (2020), arXiv:1910.10152 [astro-ph.CO].
[8] S. C. Hotinli, G. P. Holder, M. C. Johnson, and M. Kamionkowski, (2022), arXiv:2204.12503 [astro-ph.CO].
[9] S. Ferraro and J. C. Hill, Phys. Rev. D **97**, 023512 (2018), arXiv:1705.06751 [astro-ph.CO].
[10] H. Cai, M. S. Madhavacheril, J. C. Hill, and A. Kosowsky, Phys. Rev. D **105**, 043516 (2022), arXiv:2111.01944 [astro-ph.CO].
[11] H. N. Nguyen, N. Sehgal, and M. S. Madhavacheril, Phys. Rev. D **99**, 023502 (2019), arXiv:1710.03747 [astro-ph.CO].
[12] D. Han and N. Sehgal, Phys. Rev. D **105**, 083516 (2022), arXiv:2112.02109 [astro-ph.CO].
[13] S. C. Hotinli, J. Meyers, N. Dalal, A. H. Jaffe, M. C. Johnson, J. B. Mertens, M. Munchmeyer, K. M. Smith, and A. van Engelen, Phys. Rev. Lett. **123**, 061301 (2019), arXiv:1812.03167 [astro-ph.CO].
[14] S. C. Hotinli, M. C. Johnson, and J. Meyers, Phys. Rev. D **103**, 043536 (2021), arXiv:2006.03060 [astro-ph.CO].
[15] S. C. Hotinli, K. M. Smith, M. S. Madhavacheril, and M. Kamionkowski, Phys. Rev. D **104**, 083529 (2021), arXiv:2108.02207 [astro-ph.CO].
[16] K. M. Smith, D. Hanson, M. LoVerde, C. M. Hirata, and O. Zahn, JCAP **2012** (6), 014, arXiv:1010.0048 [astro-ph.CO].
[17] D. Green, J. Meyers, and A. van Engelen, JCAP **12**, 005, arXiv:1609.08143 [astro-ph.CO].
[18] N. Sehgal, M. S. Madhavacheril, B. Sherwin, and A. van Engelen, Phys. Rev. D **95**, 103512 (2017), arXiv:1612.03898 [astro-ph.CO].
[19] A. Manzotti, Phys. Rev. D **97**, 043527 (2018), arXiv:1710.11038 [astro-ph.CO].
[20] S. C. Hotinli, J. Meyers, C. Trendafilova, D. Green, and A. van Engelen, JCAP **04** (04), 020, arXiv:2111.15036 [astro-ph.CO].
[21] P. Larsen, A. Challinor, B. D. Sherwin, and D. Mak, Phys. Rev. Lett. **117**, 151102 (2016), arXiv:1607.05733 [astro-ph.CO].
[22] J. Carron, A. Lewis, and A. Challinor, JCAP **05**, 035, arXiv:1701.01712 [astro-ph.CO].
[23] N. Aghanim *et al.* (Planck), Astron. Astrophys. **641**, A8 (2020), arXiv:1807.06210 [astro-ph.CO].
[24] P. A. R. Ade *et al.* (SPTpol, BICEP, Keck), Phys. Rev. D **103**, 022004 (2021), arXiv:2011.08163 [astro-ph.CO].
[25] D. Han *et al.* (ACT), JCAP **01**, 031, arXiv:2007.14405 [astro-ph.CO].
[26] T. Namikawa *et al.*, Phys. Rev. D **105**, 023511 (2022), arXiv:2110.09730 [astro-ph.CO].
[27] S. Ho, S. Dedeo, and D. Spergel, (2009), arXiv:0903.2845 [astro-ph.CO].
[28] J. Shao, P. Zhang, W. Lin, Y. Jing, and J. Pan, Mon. Not. Roy. Astron. Soc. **413**, 628 (2011), arXiv:1004.1301 [astro-ph.CO].
[29] D. Munshi, I. T. Iliev, K. L. Dixon, and P. Coles, Mon. Not. Roy. Astron. Soc. **463**, 2425 (2016), arXiv:1511.03449 [astro-ph.CO].
[30] C.-P. Ma and J. N. Fry, Phys. Rev. Lett. **88**, 211301 (2002), arXiv:astro-ph/0106342 [astro-ph].
[31] D. N. Limber, Astrophys. J. **119**, 655 (1954).
[32] M. LoVerde and N. Afshordi, Phys. Rev. **D78**, 123506 (2008), arXiv:0809.5112 [astro-ph].
[33] K. M. Smith, M. S. Madhavacheril, M. Münchmeyer, S. Ferraro, U. Giri, and M. C. Johnson, (2018), arXiv:1810.13423 [astro-ph.CO].
[34] L. D. Shaw, D. H. Rudd, and D. Nagai, Astrophys. J. **756**, 15 (2012), arXiv:1109.0553 [astro-ph.CO].
[35] U. Giri and K. M. Smith, (2020), arXiv:2010.07193 [astro-ph.CO].
[36] N. Kaiser, Mon. Not. Roy. Astron. Soc. **227**, 1 (1987).
[37] J. C. Jackson, Mon. Not. Roy. Astron. Soc. **156**, 1P (1972), arXiv:0810.3908 [astro-ph].
[38] M. Jalilvand, B. Ghosh, E. Majerotto, B. Bose, R. Durrer, and M. Kunz, Phys. Rev. D **101**, 043530 (2020), arXiv:1907.13109 [astro-ph.CO].
[39] H. S. Grasshorn Gebhardt and D. Jeong, Phys. Rev. D **102**, 083521 (2020), arXiv:2008.08706 [astro-ph.CO].
[40] N. S. Sugiyama, S. Saito, F. Beutler, and H.-J. Seo, Mon. Not. Roy. Astron. Soc. **497**, 1684 (2020), arXiv:1908.06234 [astro-ph.CO].
[41] O. Darwish, S. Foreman, M. M. Abidi, T. Baldauf, B. D. Sherwin, and P. D. Meerburg, Phys. Rev. D **104**, 123520 (2021), arXiv:2007.08472 [astro-ph.CO].
[42] R. K. Sheth, H. J. Mo, and G. Tormen, Mon. Not. Roy. Astron. Soc. **323**, 1 (2001), arXiv:astro-ph/9907024.
[43] J. F. Navarro, C. S. Frenk, and S. D. M. White, Astrophys. J. **462**, 563 (1996), arXiv:astro-ph/9508025.
[44] A. R. Duffy, J. Schaye, S. T. Kay, and C. Dalla Vecchia, Mon. Not. Roy. Astron. Soc. **390**, L64 (2008), [Erratum: Mon.Not.Roy.Astron.Soc. 415, L85 (2011)], arXiv:0804.2486 [astro-ph].
[45] N. Battaglia, JCAP **08**, 058, arXiv:1607.02442 [astro-ph.CO].
[46] B. Bolliet, J. C. Hill, S. Ferraro, A. Kusiak, and A. Krolewski, (2022), arXiv:2208.07847 [astro-ph.CO].

[47] A. A. Berlind and D. H. Weinberg, Astrophys. J. **575**, 587 (2002), arXiv:astro-ph/0109001.
[48] V. Gonzalez-Perez, J. Comparat, P. Norberg, C. M. Baugh, S. Contreras, C. Lacey, N. McCullagh, A. Orsi, J. Helly, and J. Humphries, Mon. Not. Roy. Astron. Soc. **474**, 4024 (2018), [Erratum: Mon.Not.Roy.Astron.Soc. 503, 28–30 (2021)], arXiv:1708.07628 [astro-ph.GA].
[49] S. Alam, J. A. Peacock, K. Kraljic, A. J. Ross, and J. Comparat, Mon. Not. Roy. Astron. Soc. **497**, 581 (2020), arXiv:1910.05095 [astro-ph.CO].
[50] S. Avila *et al.*, Mon. Not. Roy. Astron. Soc. **499**, 5486 (2020), arXiv:2007.09012 [astro-ph.CO].
[51] S. Yuan, B. Hadzhiyska, S. Bose, and D. J. Eisenstein, (2022), arXiv:2202.12911 [astro-ph.CO].
[52] F. Villaescusa-Navarro *et al.*, Astrophys. J. **866**, 135 (2018), arXiv:1804.09180 [astro-ph.CO].
[53] C. Modi, E. Castorina, Y. Feng, and M. White, JCAP **09**, 024, arXiv:1904.11923 [astro-ph.CO].
[54] Z. Zheng, A. A. Berlind, D. H. Weinberg, A. J. Benson, C. M. Baugh, S. Cole, R. Dave, C. S. Frenk, N. Katz, and C. G. Lacey, Astrophys. J. **633**, 791 (2005), arXiv:astro-ph/0408564.
[55] A. Smith, S. Cole, C. Baugh, Z. Zheng, R. Angulo, P. Norberg, and I. Zehavi, Mon. Not. Roy. Astron. Soc. **470**, 4646 (2017), arXiv:1701.06581 [astro-ph.CO].
[56] K. Walsh and J. Tinker, Mon. Not. Roy. Astron. Soc. **488**, 470 (2019), arXiv:1905.07024 [astro-ph.CO].
[57] M. S. Madhavacheril, N. Battaglia, K. M. Smith, and J. L. Sievers, Phys. Rev. D **100**, 103532 (2019), arXiv:1901.02418 [astro-ph.CO].
[58] T. Abbott *et al.* (DES), (2005), arXiv:astro-ph/0510346.
[59] P. A. Abell *et al.* (LSST Science, LSST Project), (2009), arXiv:0912.0201 [astro-ph.IM].
[60] O. Doré *et al.*, (2014), arXiv:1412.4872 [astro-ph.CO].
[61] S. Ferraro, E. Schaan, and E. Pierpaoli, (2022), arXiv:2205.10332 [astro-ph.CO].
[62] R. Mandelbaum *et al.* (LSST Dark Energy Science), (2018), arXiv:1809.01669 [astro-ph.CO].
[63] A. Aghamousa *et al.* (DESI), (2016), arXiv:1611.00036 [astro-ph.IM].
[64] K. S. Dawson *et al.* (BOSS), Astron. J. **145**, 10 (2013), arXiv:1208.0022 [astro-ph.CO].
[65] P. Zarrouk *et al.*, Mon. Not. Roy. Astron. Soc. **509**, 1478 (2021), arXiv:2106.13120 [astro-ph.GA].
[66] R. Zhou *et al.*, Mon. Not. Roy. Astron. Soc. **501**, 3309 (2021), arXiv:2001.06018 [astro-ph.CO].
[67] B. Hadzhiyska, S. Tacchella, S. Bose, and D. J. Eisenstein, Mon. Not. Roy. Astron. Soc. **502**, 3599 (2021), arXiv:2011.05331 [astro-ph.GA].
[68] D. Spergel *et al.*, (2015), arXiv:1503.03757 [astro-ph.IM].
[69] Y. Wang *et al.*, (2021), arXiv:2110.01829 [astro-ph.CO].
[70] Z. Zhai, Y. Wang, A. Benson, C.-H. Chuang, and G. Yepes, Mon. Not. Roy. Astron. Soc. **505**, 2784 (2021), arXiv:2103.11063 [astro-ph.CO].
[71] A. Blanchard *et al.* (Euclid), Astron. Astrophys. **642**, A191 (2020), arXiv:1910.09273 [astro-ph.CO].
[72] D. J. Schlegel *et al.*, (2019), arXiv:1907.11171 [astro-ph.IM].
[73] Y. Harikane *et al.* 10.1093/pasj/psx097 (2017), arXiv:1704.06535 [astro-ph.GA].
[74] S. Ferraro *et al.*, (2019), arXiv:1903.09208 [astro-ph.CO].
[75] A. Mead, J. Peacock, C. Heymans, S. Joudaki, and A. Heavens, Mon. Not. Roy. Astron. Soc. **454**, 1958 (2015), arXiv:1505.07833 [astro-ph.CO].
[76] A. J. Mead and L. Verde, Mon. Not. Roy. Astron. Soc. **503**, 3095 (2021), arXiv:2011.08858 [astro-ph.CO].
[77] P. Bull, P. G. Ferreira, P. Patel, and M. G. Santos, Astrophys. J. **803**, 21 (2015), arXiv:1405.1452 [astro-ph.CO].
[78] R. Ansari *et al.* (Cosmic Visions 21 cm), (2018), arXiv:1810.09572 [astro-ph.CO].
[79] A. Liu and J. R. Shaw, Publ. Astron. Soc. Pac. **132**, 062001 (2020), arXiv:1907.08211 [astro-ph.IM].
[80] A. Slosar *et al.* (PUMA), Bull. Am. Astron. Soc. **51**, 53 (2019), arXiv:1907.12559 [astro-ph.IM].
[81] E. Castorina *et al.*, (2020), arXiv:2002.05072 [astro-ph.IM].
[82] H. Padmanabhan, A. Refregier, and A. Amara, Mon. Not. Roy. Astron. Soc. **469**, 2323 (2017), arXiv:1611.06235 [astro-ph.CO].
[83] E. Castorina and F. Villaescusa-Navarro, Mon. Not. Roy. Astron. Soc. **471**, 1788 (2017), arXiv:1609.05157 [astro-ph.CO].
[84] M. F. Morales, B. Hazelton, I. Sullivan, and A. Beardsley, Astrophys. J. **752**, 137 (2012), arXiv:1202.3830 [astro-ph.IM].
[85] A. R. Parsons, J. C. Pober, J. E. Aguirre, C. L. Carilli, D. C. Jacobs, and D. F. Moore, Astrophys. J. **756**, 165 (2012), arXiv:1204.4749 [astro-ph.IM].
[86] A. Liu, A. R. Parsons, and C. M. Trott, Phys. Rev. D **90**, 023018 (2014), arXiv:1404.2596 [astro-ph.CO].
[87] A. Ghosh, F. Mertens, and L. V. E. Koopmans, Mon. Not. Roy. Astron. Soc. **474**, 4552 (2018), arXiv:1709.06752 [astro-ph.CO].
[88] S. Gagnon-Hartman, Y. Cui, A. Liu, and S. Ravanbakhsh, Mon. Not. Roy. Astron. Soc. **504**, 4716 (2021), arXiv:2102.08382 [astro-ph.CO].
[89] H.-M. Zhu, U.-L. Pen, Y. Yu, and X. Chen, Phys. Rev. D **98**, 043511 (2018), arXiv:1610.07062 [astro-ph.CO].
[90] C. Modi, M. White, A. Slosar, and E. Castorina, JCAP **11**, 023, arXiv:1907.02330 [astro-ph.CO].
[91] D. Crichton *et al.*, J. Astron. Telesc. Instrum. Syst. **8**, 011019 (2022), arXiv:2109.13755 [astro-ph.IM].
[92] K. Vanderlinde *et al.* 10.5281/zenodo.3765414 (2019), arXiv:1911.01777 [astro-ph.IM].
[93] A. Obuljen, M. Simonović, A. Schneider, and R. Feldmann, (2022), arXiv:2207.12398 [astro-ph.CO].
[94] H. Park, P. R. Shapiro, E. Komatsu, I. T. Iliev, K. Ahn, and G. Mellema, Astrophys. J. **769**, 93 (2013), arXiv:1301.3607 [astro-ph.CO].
[95] P. Ade *et al.* (Simons Observatory), JCAP **02**, 056, arXiv:1808.07445 [astro-ph.CO].
[96] M. H. Abitbol *et al.* (Simons Observatory), Bull. Am. Astron. Soc. **51**, 147 (2019), arXiv:1907.08284 [astro-ph.IM].

[97] K. N. Abazajian *et al.* (CMB-S4), (2016), arXiv:1610.02743 [astro-ph.CO].

[98] K. Abazajian *et al.*, (2019), arXiv:1907.04473 [astro-ph.IM].

[99] M. Tegmark, A. de Oliveira-Costa, and A. J. S. Hamilton, Phys. Rev. D **68**, 123523 (2003), arXiv:astro-ph/0302496.

[100] S. Aiola *et al.* (CMB-HD), (2022), arXiv:2203.05728 [astro-ph.CO].

[101] L. Pagano, J. M. Delouis, S. Mottet, J. L. Puget, and L. Vibert, Astron. Astrophys. **635**, A99 (2020), arXiv:1908.09856 [astro-ph.CO].

[102] A. Gruzinov and W. Hu, Astrophys. J. **508**, 435 (1998), arXiv:astro-ph/9803188.

[103] A. Mesinger, M. McQuinn, and D. N. Spergel, Mon. Not. Roy. Astron. Soc. **422**, 1403 (2012), arXiv:1112.1820 [astro-ph.CO].

[104] O. Zahn *et al.*, Astrophys. J. **756**, 65 (2012), arXiv:1111.6386 [astro-ph.CO].

[105] K. M. Smith and S. Ferraro, Phys. Rev. Lett. **119**, 021301 (2017), arXiv:1607.01769 [astro-ph.CO].

[106] T. Okamoto and W. Hu, Phys. Rev. D **67**, 083002 (2003), arXiv:astro-ph/0301031.

[107] O. Darwish, B. D. Sherwin, N. Sailer, E. Schaan, and S. Ferraro, (2021), arXiv:2111.00462 [astro-ph.CO].

[108] N. Sailer, E. Schaan, and S. Ferraro, Phys. Rev. D **102**, 063517 (2020), arXiv:2007.04325 [astro-ph.CO].

[109] A. Baleato Lizancos and S. Ferraro, (2022), arXiv:2205.09000 [astro-ph.CO].

[110] B. Hadzhiyska, B. D. Sherwin, M. Madhavacheril, and S. Ferraro, Phys. Rev. D **100**, 023547 (2019), arXiv:1905.04217 [astro-ph.CO].

[111] J. Gao, X. Xu, X. Zhou, J. Ma, Z. Zhou, T. Zhang, J. Nie, J. Wang, S. Xue, and H. Zou, Astrophys. J. Suppl. **253**, 56 (2021), arXiv:2101.12340 [astro-ph.GA].

[112] N. Hamaus, U. Seljak, and V. Desjacques, Phys. Rev. D **84**, 083509 (2011), arXiv:1104.2321 [astro-ph.CO].

[113] M. Münchmeyer, M. S. Madhavacheril, S. Ferraro, M. C. Johnson, and K. M. Smith, Phys. Rev. D **100**, 083508 (2019), arXiv:1810.13424 [astro-ph.CO].

[114] A. Terrana, M.-J. Harris, and M. C. Johnson, JCAP **1702** (02), 040, arXiv:1610.06919 [astro-ph.CO].

[115] A.-S. Deutsch, E. Dimastrogiovanni, M. C. Johnson, M. Munchmeyer, and A. Terrana, Phys. Rev. D **98**, 123501 (2018), arXiv:1707.08129 [astro-ph.CO].

[116] J. Cayuso, R. Bloch, S. C. Hotinli, M. C. Johnson, and F. McCarthy, (2021), arXiv:2111.11526 [astro-ph.CO].

[117] S. C. Hotinli and M. C. Johnson, Phys. Rev. D **105**, 063522 (2022), arXiv:2012.09851 [astro-ph.CO].

[118] S. C. Hotinli, J. B. Mertens, M. C. Johnson, and M. Kamionkowski, Phys. Rev. D **100**, 103528 (2019), arXiv:1908.08953 [astro-ph.CO].

[119] N. Anil Kumar, G. Sato-Polito, M. Kamionkowski, and S. C. Hotinli, (2022), arXiv:2205.03423 [astro-ph.CO].

[120] S. C. Hotinli, S. Ferraro, G. P. Holder, M. C. Johnson, M. Kamionkowski, and P. La Plante, (2022), arXiv:2207.07660 [astro-ph.CO].

[121] N. A. Kumar, S. C. Hotinli, and M. Kamionkowski, (2022), arXiv:2208.02829 [astro-ph.CO].

[122] K. M. Smith, D. Hanson, M. LoVerde, C. M. Hirata, and O. Zahn, JCAP **06**, 014, arXiv:1010.0048 [astro-ph.CO].

[123] M. Millea, E. Anderes, and B. D. Wandelt, Phys. Rev. D **102**, 123542 (2020), arXiv:2002.00965 [astro-ph.CO].

[124] E. Guzman and J. Meyers, Phys. Rev. D **104**, 043529 (2021), arXiv:2101.01214 [astro-ph.CO].

[125] H. Tanimura, N. Aghanim, V. Bonjean, and S. Zaroubi, Astron. Astrophys. **662**, A48 (2022), arXiv:2201.01643 [astro-ph.CO].

[126] G. P. Lepage, J. Comput. Phys. **439**, 110386 (2021), arXiv:2009.05112 [physics.comp-ph].

[127] P. Lepage, gplepage/vegas: vegas version 5.1.1 (2022).

[128] E. Schaan and M. White, JCAP **05**, 068, arXiv:2103.01964 [astro-ph.CO].

[129] K. Yamamoto, Y. Nan, and C. Hikage, Phys. Rev. D **95**, 043528 (2017), arXiv:1610.03665 [astro-ph.CO].

[130] S. Cunnington, S. Camera, and A. Pourtsidou, Mon. Not. Roy. Astron. Soc. **499**, 4054 (2020), arXiv:2007.12126 [astro-ph.CO].

[131] J. R. Shaw, K. Sigurdson, M. Sitwell, A. Stebbins, and U.-L. Pen, Phys. Rev. D **91**, 083514 (2015), arXiv:1401.2095 [astro-ph.CO].

[132] M. S. Madhavacheril, N. Battaglia, and H. Miyatake, Phys. Rev. D **96**, 103525 (2017), arXiv:1708.07502 [astro-ph.CO].

[133] J. Dunkley *et al.*, JCAP **07**, 025, arXiv:1301.0776 [astro-ph.CO].

[134] G. Lagache, M. Béthermin, L. Montier, P. Serra, and M. Tucci, Astron. Astrophys. **642**, A232 (2020), arXiv:1911.09466 [astro-ph.CO].

[135] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000), arXiv:astro-ph/9911177 [astro-ph].