This is the accepted manuscript made available via CHORUS. The article has been published as:

Primordial trispectrum from kinetic Sunyaev-Zel'dovich tomography

Neha Anil Kumar, Gabriela Sato-Polito, Marc Kamionkowski, and Selim C. Hotinli

Phys. Rev. D **106**, 063533 — Published 27 September 2022

DOI: 10.1103/PhysRevD.106.063533

Primordial trispectrum from kSZ tomography

Neha Anil Kumar,* Gabriela Sato-Polito, Marc Kamionkowski, and Selim C. Hotinli
William H. Miller III Department of Physics and Astronomy,
Johns Hopkins University Baltimore, MD 21218, USA

The kinetic Sunyaev Zel'dovich effect is a secondary CMB temperature anisotropy that provides a powerful probe of the radial-velocity field of matter distributed across the Universe. This velocity field is reconstructed by combining high-resolution CMB measurements with galaxy survey data, and it provides an unbiased tracer of matter perturbations in the linear regime. In this paper, we show how this measurement can be used to probe primordial non-Gaussianity of the local type, particularly focusing on the trispectrum amplitude $\tau_{\rm NL}$, as may arise in a simple two-field inflation model that we provide by way of illustration. Cross-correlating the velocity-field-derived matter distribution with the biased large-scale galaxy density field allows one to measure the scale-dependent bias factor with sample variance cancellation. We forecast that a configuration corresponding to CMB-S4 and VRO results in a sensitivity of $\sigma_{f_{\rm NL}} \approx 0.59$ and $\sigma_{\tau_{\rm NL}} \approx 1.5$. These forecasts predict improvement factors of 10 and 195 for $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$, respectively, over the sensitivity using VRO data alone, without internal sample variance cancellation. Similarly, for a configuration corresponding to DESI and SO, we forecast a sensitivity of $\sigma_{f_{\rm NL}} \approx 3.1$ and $\sigma_{\tau_{\rm NL}} \approx 69$, with improvement factors of 2 and 5, respectively, over the use of the DESI data-set in isolation. We find that a high galaxy number density and large survey volume considerably improve our ability to probe the amplitude of the primordial trispectrum for the multi-field model considered.

I. INTRODUCTION

Detecting and constraining characteristics of the primordial Universe to understand the origin of structure is one of the primary goals of many upcoming large-scale structure surveys and CMB experiments [1–8]. The most widely accepted paradigm is that of inflation [9–11], which addresses most of the problems of the original Big-Bang scenario and has a set of predictions that are compatible with many current observations [12–15]. Although the general predictions of the inflationary model, such as a flat universe and largely scale-invariant set of initial fluctuations, have been confirmed by recent observations, the specific physical processes that govern this epoch are yet to be understood.

Comparing the predictions of various inflationary models to astrophysical observations allows one to probe the physics of the ultra-high energy scales that are otherwise not directly accessible to experiments. Searching for imprints of primordial non-Gaussianity in the CMB spectrum or on the large-scale matter distribution are possible methods to effectively distinguish between various models of inflation and the number of degrees of freedom governing the epoch [e.g., 16–33].

One simple, and widely studied class of such non-Gaussianity is the local-type or $f_{\rm NL}$-parametrization, in which one includes a quadratic term in the primordial potential $\Phi = \phi + f_{\rm NL}\phi^2$. In this model, both linear and quadratic terms in the potential originate from the same Gaussian field ϕ, called the inflaton. The current best bound is $f_{\rm NL} = -0.9 \pm 5.1$, coming from the latest Planck satellite CMB analysis [34] with the growth factor normalized to one during matter domination.

Non-Gaussianity also naturally arises in models of inflation that involve more than one field. This could be due to a coupling term across the two fields [35] or the addition of a field with its own quadratic term [36]. This can enhance the inflaton four-point function (or trispectrum) while not affecting the more widely considered three-point function (or bispectrum) [37, 38], making the primordial trispectrum a valuable signature of extra degrees of freedom in the early Universe. The amplitude of the primordial trispectrum has also been constrained by the Planck CMB data, with the most recent estimate being $\tau_{\rm NL} = (-5.8 \pm 6.5) \times 10^4$ [34].

Given Silk damping of the temperature fluctuations, there is not much room to significantly improve upon the $f_{\rm NL}$ measurements with CMB measurements alone. However, there are a few proposed methods to further constrain $\tau_{\rm NL}$ using its signature in the trispectrum of the 21-cm brightness temperature [39], the halo bias [40, 41], and 3-point correlations between two-CMB-temperature and one-μ-spectral-distortion fluctuations [42]. The expected sensitivities for these proposals are $\tau_{\rm NL} \approx 50-100$.

In the case of local non-Gaussianity, we also expect to obtain constraints from the distribution of galaxies on large scales, relying only on the measurement of the galaxy power spectrum in the linear regime [43, 44]. This constraint can be obtained using the fact that a non-zero $f_{\rm NL}$ induces a scale-dependent bias factor [36, 45], providing a unique signal that is not mimicked by changes in the other standard cosmological parameters. However, reaching the predicted multi-field threshold of $f_{\rm NL} \gtrsim 1$ remains difficult because of sample variance.

In this paper, we forecast the sensitivity of kSZ tomography to both $f_{\rm NL}$ and $\tau_{\rm NL}$, assuming that the primordial non-Gaussianity is induced by two different fields [36]. The kinetic Sunyaev Zel'dovich (kSZ) effect is the secondary CMB temperature anisotropy induced by the

* nanilku1@jhu.edu

peculiar velocity of interspersed free-electrons that scatter the CMB photons [46–50]. Cross-correlating high-resolution CMB maps with large-scale structure surveys will allow for the measurement of this kSZ contribution as a function of redshift, a technique termed kSZ tomography [51–55]. This cross-correlation can be used to reconstruct the radial velocity field of free electrons in a 3-dimensional volume, from which the large scale matter distribution can be inferred.

By comparing this kSZ-tomography-based matter perturbation amplitude with the amplitude of the galaxy power spectrum, one can obtain excellent constraints on the scale-dependent bias. Since the matter and galaxy distributions are determined independently, the bias can be measured on a mode-by-mode basis, thus circumventing the cosmic-variance limit that usually arises when inferring $f_{\rm NL}$ and $\tau_{\rm NL}$ from the galaxy distribution data-set in isolation [56].

We forecast the precision with which our model of primordial non-Gaussianity can be measured in two distinct scenarios: one in which only a single tracer from a galaxy survey is considered and another in which both the galaxy survey and the kSZ-reconstructed velocity field are jointly measured. The forecasts that include both the galaxy distribution and velocity field are based on the kSZ bispectrum formalism developed in Ref. [57], which accounts for photo-z errors and the optical-depth degeneracy. We consider, in our forecasts, two baseline experimental configurations: 'baseline 1' corresponding to the combination of Vera Rubin Observatory (VRO) [58] and CMB-S4 [59] and 'baseline 2' corresponding to Dark Energy Spectroscopic Instrument (DESI) [60] and Simons Observatory (SO) [1, 2]. For these forecasts, we closely follow the method and experimental parameter values used in Ref. [43], in which the same experimental configurations were used to forecast survey sensitivity to the single field $f_{\rm NL}$ model of inflation.

Our forecasts find that for the configuration of VRO and CMB-S4, $\sigma_{f_{\rm NL}} \approx 0.59$ and $\sigma_{\tau_{\rm NL}} \approx 1.5$, which corresponds to improvement factors of 10 and 195, respectively, over the use of VRO data alone. Similarly, for the configuration of the DESI and SO, we find that $\sigma_{f_{\rm NL}} \approx 3.1$ and $\sigma_{\tau_{\rm NL}} \approx 69$, with improvement factors of 2 and 5, respectively, compared to the forecasts made using DESI data in isolation. We find that our forecasts on the galaxy distribution data sets alone are compatible with the single-tracer results of Ref. [44] when differences in our survey parameters are taken into account.

Through the variation of experimental parameters we also determine that changes in redshift dispersion arising from photo-z errors, as well as increases in CMB sensitivity and CMB resolution have a relatively minimal effect on our ability to measure the non-Gaussianity. In contrast, we find that a large survey volume, with well measured large-scale modes, and a high galaxy number density most prominently decrease the error with which both $f_{\rm NL}$ and $\tau_{\rm NL}$ can be measured.

Throughout this paper, we adopt the ΛCDM Cosmology as our fiducial model with the following parameters from *Planck* 2018 [61]: reduced Hubble constant $h = 0.674$, baryon and cold-dark-matter density parameters today $\Omega_b = 0.049$, and $\Omega_{\rm cdm} = 0.264$ respectively, spectral index $n_s = 0.965$ and amplitude of the primordial scalar power spectrum $A_s = 2.2 \times 10^{-9}$. In all our equations, we work under the convention $c = 1$.

This paper is organised as follows. In Sec. II, we introduce our scale-dependent biasing model, derived using the peak-background-split methodology for the multi-field model of inflation presented in Ref. [36]. We also explain how kSZ tomography can be used for sample variance cancellation. In Sec. III, we describe the experimental parameters in our forecast. The forecast set up is described in Sec. IV. Finally, the results of our analysis and our final set of forecasts are detailed in Sec. V.

II. THEORY

Before explaining in detail how kSZ tomography can be used in tandem with large-scale galaxy survey data, we first introduce the $\tau_{\rm NL}$ model of primordial non-Gaussianity and derive the relevant power spectra relations via the peak background split formalism. Further details on this model and its derivations can be found in Ref. [36]. Furthermore, we address the possible extension of the results in this paper to another commonly considered, non-Gaussian model of inflation. Finally, we then also briefly address how the velocity field is reconstructed, given temperature-anisotropy and galaxy-distribution data.

A. Local non-Gaussianity in Peak Background Split Formalism

In this paper, we consider the curvaton model for the primordial gravitational potential in which two different fields, the inflaton and the curvaton, contribute to the curvature perturbation. The contribution from the inflaton is purely Gaussian while the perturbations of the curvaton field generate the non-Gaussianity. In this model, therefore, the primordial potential takes the following form:

$$\Phi(\boldsymbol{x}) = \phi(\boldsymbol{x}) + \psi(\boldsymbol{x}) + f_{\rm NL}(1+\Pi)^2(\psi^2(\boldsymbol{x}) - \langle\psi^2\rangle). \quad (1)$$

Here, ϕ and ψ are uncorrelated Gaussian random fields with power spectra that are proportional to each other, with proportionality constant $\Pi \equiv P_\phi/P_\psi$. For this model, the three and four point functions take the local form

$$\xi^{(3)}_\phi(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = f_{\rm NL}[P_1P_2 + 5 \text{ perms.}] + \mathcal{O}(f^3_{\rm NL}), \quad (2)$$

$$\xi^{(4)}_\phi(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) = 2\Big(\frac{5}{6}\Big)^2\tau_{\rm NL}[P_1P_2P_{13} + 23 \text{ perms.}] + \mathcal{O}(\tau^2_{\rm NL}), \quad (3)$$

where we have defined $P_i \equiv P_\Phi(k_i)$, $P_{ij} = P_\Phi(|\boldsymbol{k_i} + \boldsymbol{k_j}|)$, and $\tau_{\rm NL} \equiv (6f_{\rm NL}/5)^2(1+\Pi)$. Therefore, for $\Pi = 0$, this model reduces to the singly-parametrized $f_{\rm NL}$ model.

Large-scale halo bias is usually treated under the context of the peak background split formalism [62], where one can split the density field into a long-wavelength piece δ_ℓ and a short-wavelength piece δ_s as in

$$\rho(\boldsymbol{x}) = \bar{\rho}(1 + \delta_\ell + \delta_s). \tag{4}$$

The local Lagrangian number density of halos $n(\boldsymbol{x})$, at position $\boldsymbol{x}$ is dependent on the local value of the long-wavelength perturbation δ_ℓ as well as the local small scale power $\sigma_8^{\rm local}(\boldsymbol{x})$. In the Gaussian case, since the small scale power is a constant, the Lagrangian bias is solely dependent on the variation in the halo number density as a function of the large-scale matter overdensity field [45, 63].

When non-Gaussianity is present, the analysis under this formalism is complicated by the fact that the large- and small-scale density fluctuations are no longer independent. This becomes evident in the $\tau_{\rm NL}$ model when the long- and short-wavelength pieces of the Gaussian potential fluctuations are separated as follows:

$$\phi = \phi_\ell + \phi_s, \qquad \psi = \psi_\ell + \psi_s. \tag{5}$$

Plugging this into Eq. (1) will show that a few of the terms contain both short- and long-wavelength pieces. Therefore, this scenario needs more careful handling.

We start by establishing the Fourier-space relation between the primordial potential and matter overdensity field $\delta(\boldsymbol{k}, z) = \alpha(k, z)\Phi(\boldsymbol{k})$, where the form of the Poisson equation based operator $\alpha(k)$ is given by [45]:

$$\alpha(k, z) = \frac{2k^2 T(k) G(z)}{3\Omega_m H_0^2}. \tag{6}$$

Here, $G(z)$ is the linear growth rate normalized such that $G(z) = 1/(1+z)$ during matter domination and $T(k)$ is the transfer function normalized to 1 at low k. Since this operator is usually defined in terms of its action in Fourier space, when applied to a real-space function such as $\phi(\boldsymbol{x})$, we use the convention

$$\alpha\phi(\boldsymbol{x}) \equiv \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\alpha(k)e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \int d^3\boldsymbol{y}\ \phi(\boldsymbol{y})e^{-i\boldsymbol{k}\cdot\boldsymbol{y}}. \tag{7}$$

With this relation in hand, the contributions from both the inflaton and the curvaton field, to the long-wavelength piece of the matter overdensity fluctuation, can be written as:

$$\delta_\ell(\boldsymbol{x}) = \alpha[\phi_\ell(\boldsymbol{x}) + \psi_\ell(\boldsymbol{x})], \tag{8}$$

where the remaining terms are either much smaller ($f_{\rm NL}\psi_\ell^2$) or contain short-wavelength pieces. Similarly, within a region of given large-scale overdensity δ_ℓ and potential $[\phi_\ell + \psi_\ell]$, the short-wavelength modes of the matter overdensity field are:

$$\delta_s = \alpha[\phi_s + \psi_s(1 + 2f_{\rm NL}(1+\Pi)^2\psi_\ell)], \tag{9}$$

where the white-noise term $[f_{\rm NL}(1+\Pi)^2\psi_s^2]$, that is spatially invariant when averaged over, has been disregarded, and the explicit $\boldsymbol{x}$-dependence of the terms has been dropped for ease of notation.

Given the above split, it is evident that the mixing of the short- and long-wavelength pieces induces a scale dependence on the local small-scale power of the matter overdensity field:

$$\begin{aligned}\sigma^2 &= \alpha^2\{\langle\phi_s^2\rangle + \langle\psi_s^2\rangle[1 + 2f_{\rm NL}(1+\Pi)^2\psi_\ell]^2\} \\ &= \bar{\sigma}^2[1 + 4f_{\rm NL}(1+\Pi)\psi_\ell],\end{aligned} \tag{10}$$

where $\bar{\sigma}^2 = \alpha^2\langle\psi_s^2\rangle(1+\Pi)$ and we have, once again, dropped any terms quadratic in ψ_ℓ. The above expression indicates that when there exists primordial non-Gaussianity, the number density of halos varies not only with the large-scale matter overdensity modes but also with the local small-scale power. This can be accounted for in the derivation of the Lagrangian halo bias as

$$\delta_h \equiv \frac{\delta n_h}{n_h} = b_h\delta_\ell + \beta_f(1+\Pi)f_{\rm NL}\psi_\ell, \tag{11}$$

where

$$b_h \equiv \frac{\partial \ln n_h}{\partial \delta_\ell} \quad \text{and} \quad \beta_f \equiv 2\frac{\partial \ln n_h}{\partial \ln \sigma} = 2\delta_c(b_h - 1). \tag{12}$$

Given the above form, it is straightforward to calculate the matter-halo (P_{mh}) and halo-halo (P_{hh}) power spectra. Using the fact that the inflaton and curvaton fields are uncorrelated, and that their power spectra are proportional to each other, one can derive

$$P_{mh}(k, z) = \left[b_h + \beta_f\frac{f_{\rm NL}}{\alpha(k, z)}\right]P_{mm}(k, z), \tag{13}$$

and

$$P_{hh}(k, z) = \left[b_h^2 + 2b_h\beta_f\frac{f_{\rm NL}}{\alpha(k, z)} + \beta_f^2\frac{\left(\frac{5}{6}\right)^2\tau_{\rm NL}}{\alpha^2(k, z)}\right]P_{mm}(k, z), \tag{14}$$

where $P_{mm}(k, z)$ refers to the large-scale matter power spectrum, i.e., the Fourier space variance in our large-scale overdensity δ_ℓ. From this point on, since we will primarily be dealing with matter overdensities on linear scales, we will label our large-scale overdensity with δ_m. On these linear scales, we will continue to use a subscript of h when referring to the halo power spectra.

The forecasts in this paper will, therefore, focus on calculating survey sensitivity to both $f_{\rm NL}$ and $\tau_{\rm NL}$ under the null hypothesis ($f_{\rm NL} = \tau_{\rm NL} = 0$), using both P_{mh} and P_{hh} as parametrized above. Although the parameter Π more directly provides information on whether the primordial potential is defined by two different fields, we choose not to explore the parameter space in terms of $[f_{\rm NL}, \Pi]$. This is because Π is defined as the ratio between two power spectra and can realistically be infinite in the absence of the curvaton field ($P_\psi = 0$).

Neverthless, we can still attempt to probe the degrees of freedom during inflation using the fact that the trispectrum amplitude satisfies $\tau_{\rm NL} \geq (6/5)^2 f_{\rm NL}^2$ for multi-field models. This relation can be used to define the parameter

$$r_{\rm NL} = (5/6)^2 \tau_{\rm NL} - f_{\rm NL}^2\,, \tag{15}$$

which can deviate away from zero in the presence of additional degrees of freedom in the early Universe, depending on the value of Π. Hence, $\tau_{\rm NL}$ can be thought of as a probe of the extra degrees of freedom during inflation.

B. Redshift Space Distortions

Redshift maps of galaxies distributed in a given survey volume are distorted by their peculiar velocities along the line of sight. When the bias relation is linear, the redshift-distorted halo overdensity is the sum of the biased matter overdensity in real space and a correction from the peculiar velocity of galaxies [64]

$$\delta_{h,\rm RSD}(\boldsymbol{x}) = b_1 \delta_m(\boldsymbol{x}) + \frac{\partial}{\partial \boldsymbol{x}}\Big[\frac{\boldsymbol{u}(\boldsymbol{x})\cdot\hat{\boldsymbol{x}}}{aH}\Big], \tag{16}$$

where $\boldsymbol{u}$ refers to the peculiar velocity of the galaxies and $\boldsymbol{x}$ refers to the position of the observed galaxy. To simplify the conversion to Fourier space, we use the late-time, linearized, continuity-equation-based relation between the peculiar-velocity field and matter-overdensity field,

$$\boldsymbol{u}(\boldsymbol{k}, z) = aHf\frac{i\boldsymbol{k}}{k^2}\delta_m(\boldsymbol{k}, z). \tag{17}$$

Here, f refers to the linear growth rate $d\ln G/d\ln a$. With the above relation in hand, the Fourier transform of the redshift space linear bias relation simplifies to

$$\delta_{h,\rm RSD}(\boldsymbol{k}) = [b_h + f\mu_k^2]\delta_m(\boldsymbol{k}) \tag{18}$$

where μ_k is defined to be $\hat{\boldsymbol{e}}_z \cdot \hat{\boldsymbol{k}}$, the cosine of the angle between the line of sight and the wavevector $\hat{\boldsymbol{k}}$.

It is straightforward to extend this derivation to the bias relation in Fourier space, for the $\tau_{\rm NL}$ model of primordial non-Gaussianity. The updated form of the halo overdensity is simply

$$\delta_{h,\rm RSD}(\boldsymbol{k}) = [b_h + f\mu_k^2]\delta_m(\boldsymbol{k}) + \beta_f(1+\Pi)f_{\rm NL}\psi_\ell(\boldsymbol{k}), \tag{19}$$

where the same correction term is added to the original form introduced in Eq. (11). The power spectra can therefore be updated, under the effects of RSD, by replacing every instance of b_h in our previously derived halo power spectra models with $b_{h,\rm RSD} = [b_h + f\mu_k^2]$.

C. Extension to the $g_{\rm NL}$ Model

In this work, we primarily focus on $\tau_{\rm NL}$- and $f_{\rm NL}$- type non-Gaussianities that have a clear correspondence predicted in the case of single-field slow-roll inflation. However, the forecasts in this paper can be extended to another possible model, parametrized by $g_{\rm NL}$, in which the primordial potential takes the following form:

$$\Phi(\boldsymbol{x}) = \phi(\boldsymbol{x}) + g_{\rm NL}[\phi^3(\boldsymbol{x}) - 3\langle\phi^2\rangle\phi(\boldsymbol{x})]. \tag{20}$$

For this single-field model, using the peak-background-split formalism from above, one can show that:

$$P_{mh}(k,z) = \left[b_{h,\rm RSD} + \beta_g \frac{g_{\rm NL}}{\alpha(k,z)}\right] P_{mm}(k,z), \tag{21}$$

$$P_{hh}(k,z) = \left[b_{h,\rm RSD} + \beta_g \frac{g_{\rm NL}}{\alpha(k,z)}\right]^2 P_{mm}(k,z) \tag{22}$$

where $\beta_g = 3\partial \ln n_h/\partial f_{\rm NL}$. Since, the barrier crossing prediction for β_g does not agree well with N-body simulations, previous forecasts on this model have used simulation-based fit-functions for β_g that are independent of f_{NL} under the null hypothesis (see, for example, Ref. [44]).

When compared with the single-field $f_{\rm NL}$-cosmology [where $\tau_{\rm NL} = (6/5 f_{\rm NL})^2$], one can see that the contribution of $g_{\rm NL}$ to the halo bias is the same as the contribution from $f_{\rm NL} = (\beta_f/\beta_g)g_{\rm NL}$. Therefore, when only a single population of tracers is being considered to calculate $P_{hh}(k,z)$, the two parameters are indistinguishable. Since the forecasts in this paper are calculated using the cross-correlation of a single galaxy tracer with kSZ tomography, under the null hypothesis one can use these forecasts to obtain constraints on $g_{\rm NL}$ by expressing $\sigma_{g_{\rm NL}} = (\beta_f/\beta_g)\sigma_{f_{\rm NL}}$ [44].

D. The kSZ Effect

The temperature fluctuation attributed to the kSZ in the $\hat{\boldsymbol{n}}$ direction in the sky is given by the integral [57],

$$T(\hat{\boldsymbol{n}}) = -T_{\rm CMB}\sigma_T \int \frac{d\chi}{(1+z)} e^{-\tau(\chi)} n_e(\hat{\boldsymbol{n}}, \chi)\hat{\boldsymbol{n}}\cdot\boldsymbol{v}, \tag{23}$$

where $T_{\rm CMB}$ is the average temperature of the CMB today, σ_T is the Thomson Scattering cross-section, and τ is the optical depth to the scattering electron with velocity $\boldsymbol{v}$ at comoving distance χ, and redshift z. The fluctuation is also dependent on the electron number density $n_e(\hat{\boldsymbol{n}}, \chi) = \bar{n}_e(\chi)(1+\delta_e)$.

To use this anisotropy data and derive redshift dependent information one must cross-correlate the kSZ data set with a tracer of large-scale structure. Ref. [57] shows that most of the varied approaches to this technique are equivalent to using a bispectrum of the form $\langle\delta\delta T\rangle$ to reconstruct the radial-velocity field. In the next few sections we summarize how this bispectrum is used to derive the expected form of the signal and noise. For a more detailed derivation of the results, see Ref. [57].

1. Bispectrum Based Estimator

According to Ref. [57], the statistic that carries the kSZ tomography signal is a 3-point function defined as,

$$\langle \delta_X(\boldsymbol{k})\delta_X(\boldsymbol{k'})T(\ell)\rangle = B(\boldsymbol{k},\boldsymbol{k'},\ell)(2\pi)^3\delta^3\left(\boldsymbol{k}+\boldsymbol{k'}+\frac{\ell}{\chi_*}\right), \tag{24}$$

where δ_X refers to the overdensity of the tracer in consideration, and all terms marked with a $*$ refer to quantities evaluated at redshift z_*. It can be shown that the kSZ is dominant in the squeezed limit [57], in which the bispectrum takes the form:

$$B(k_L,k_S,\ell,k_{Lr}) = -\frac{K_* k_{Lr}}{\chi_*^2}\frac{P_{Xv}(k_L)}{k_L}P_{Xe}(k_S), \tag{25}$$

where k_L refers to the long-wavelength mode, k_{Lr} is its component along the line of sight, k_S refers to the short-wavelength mode, and P_{Xv} and P_{Xe} refer to the cross-spectra of the tracer overdensity field with the velocity field, and the electron density perturbations, respectively. In the above equation we have also defined,

$$K_* \equiv -T_{\rm CMB}\sigma_T\bar{n}_{e,0}e^{-\tau(\chi_*)}(1+z_*)^2, \tag{26}$$

where $\bar{n}_{e,0}$ is the mean electron density today.

2. Velocity Reconstruction

As shown in Ref. [57], a quadratic estimator for the long-wavelength velocity modes can be constructed by summing over the pairs $[\delta_X(\boldsymbol{k}_S)T(\boldsymbol{\ell})]$ of short-wavelength modes in the galaxy and CMB maps. This method is equivalent to the optimal kSZ bispectrum estimator [57].

Given the form of $B(k_L,k_S,\ell,k_{Lr})$ in Eq. (25), the signal to noise ratio S/N of the kSZ bispectrum in the squeezed limit is

$$\frac{S}{N} = V\int\frac{d^3\boldsymbol{k}_L}{(2\pi)^3}\frac{k_{Lr}^2}{k_L^2}\frac{P_{gv}(k_L)^2}{P_{gg}^{\rm tot}(k_L)N_{v_r}(k_L)}, \tag{27}$$

where N_{v_r} is the noise associated with radial velocity reconstruction. This noise is modelled as

$$N_{v_r}(k_L,\mu_L) = \frac{2\pi\chi_*^2}{K_*^2}\left[\int dk_S\frac{k_S P_{ge}(k_S)^2}{P_{gg}^{\rm tot}(k_S)\ C_{\ell=k_S\chi_*}^{\rm tot}}\right]^{-1}. \tag{28}$$

In the above two equations, we have explicitly used a subscript of g to label our tracer X. Therefore, $P_{gg}(k_S,\mu_S)$ refers to the small-scale galaxy-galaxy auto-power spectrum and $P_{ge}(k_S,\mu_S)$ is the small-scale galaxy-electron power spectrum. Finally, μ_L refers to the angle of the large-scale mode with respect to the line of sight, i.e. $\mu_L = \hat{\boldsymbol{k}}_L\cdot\hat{\boldsymbol{n}}$. However, it is important to note that μ_S and μ_L are not independent of each other. The value of μ_S is completely determined by k_L,μ_L and k_S since the line of sight components of the Fourier modes k_L and k_S, are equal to each other. The total noise in our velocity reconstruction N_{vv} is then $N_{vv} = \mu_L^{-2}N_{v_r}$.

Here and below, a subscript of 'g' will be used to denote small-scale galaxy power spectra that appear in kSZ tomography, in contrast to the subscript h that has so far been used to label large-scale halo power spectra. While on large scales, we will assume that a single galaxy occupies each halo, small scale galaxy power spectra will be calculated within the halo model including the halo occupation distribution (HOD) [65, 66]. The modelling assumptions and parameter values used to construct the small-scale spectra under this model can be found in Appendix A.

In our model of the velocity reconstruction noise, we also include the effect of photo-z errors via a Gaussian kernel of the form

$$W^2(k,\mu) = e^{-k^2\mu^2\sigma^2(z)/H^2(z)}, \tag{29}$$

where $\sigma(z)$ is the redshift scattering of the galaxy survey in consideration. This induces a $\mu_S = \hat{\boldsymbol{k}}_S\cdot\hat{\boldsymbol{n}}$ dependence on the small-scale galaxy-galaxy and galaxy-electron spectra. Further details on the noise in the kSZ velocity reconstruction due to photo-z errors can be found in Ref. [57].

Finally, based on the linear relation between matter overdensities and peculiar velocities [Eq. (17)], the noise in the reconstructed density perturbation field is

$$N_{mm}^{\rm rec}(k_L,\mu) = \frac{k_L^2}{(faH)_*^2}N_{vv}(k_L,\mu). \tag{30}$$

It is important to note that the noise is proportional to the magnitude k_L^2. This implies that the reconstruction noise is lowest on largest scales, which corresponds to the regime where cosmic variance is a dominant noise source. Therefore, it is on these scales that we would expect this independent probe of large-scale structure to significantly contribute to sample variance cancellation.

III. EXPERIMENT SPECIFICATIONS

The primary set of forecasts presented in this paper consider two next generation large-scale structure experiments, DESI and VRO. VRO is an example of a high number density galaxy survey with photometric redshifts. In contrast, DESI is a low number density survey with precise, spectroscopic redshifts. Our forecasts assume a cross-correlation of these data-sets with kSZ data from CMB-S4, as well as data from a configuration similar to that of SO, to test and display the effects of sample variance cancellation.

A. Large-Scale Structure Experiments

For our forecasts on VRO, we use the specifications for the *LSST Gold Sample* as prescribed in the LSST science

book [58]. The galaxy number density for this data-set, per arcmin^2, is given by

$$n(z) = n_0 \frac{1}{2z_0}\left(\frac{z}{z_0}\right)^2 \exp(-z/z_0), \tag{31}$$

with $z_0 = 0.3$ and $n_0 = 40$ arcmin^2. At $z = 1$, this corresponds to a galaxy number density of approximately $n_{\rm gal} = 10^{-2}$ Mpc^{-3}. The photometric redshift error for this survey is

$$\sigma_z = 0.03(1+z), \tag{32}$$

which corresponds to a redshift dispersion of 0.06 at $z = 1$. Finally, the bias for this sample is also specified to be

$$b(z) = 0.95/G(z), \tag{33}$$

with their growth factor normalized such that $G(z = 0) = 1$. This corresponds to a bias of 1.6 at a redshift of 1. For DESI, we make a single forecast assuming a galaxy number density of $n_{\rm gal} = 10^{-4}$ Mpc^{-3} with a Gaussian halo bias of 1.6 at redshift 1, in accordance with the specifications provided in the DESI white-paper [60].

To ensure that the small-scale power spectra generated based on the HOD model are consistent with the specifications of the experiments in consideration, we use the following prescription. In the HOD model, the galaxy sample is specified by imposing a particular threshold stellar mass $m_\star^{\rm thresh}$ of observable galaxies. Since the remaining parameters defining the galaxy distribution for P_{gg} and P_{ge} are dependent on this parameter, we match the value of $m_\star^{\rm thresh}$ so that the total predicted galaxy number density matches number density expected for a given experiment. The details on these power spectra's dependencies on $n_{\rm gal}$ and $m_\star^{\rm thresh}$ can be found in Appendix A.

B. CMB Experiments

Most of our forecasts are based on the planned CMB-S4 experiment specifications. Although the exact instrument specifications are still pending, we assume an effective beam with full-width-half-maximum (FWHM) of 1.5 arcminutes and a sensitivity of 1.0 μK$-$arcmin, which is one of many possible configurations. The effects of atmospheric noise are not included since they are expected to be sub-dominant to the instrument and kSZ contributions at the relevant high multipoles of $\ell > 3000$. The final set of contributions to the CMB spectrum that enters Eq. (28) can be written as

$$C_\ell^{\rm tot} = C_\ell^{TT} + C_\ell^{\rm kSZ\text{-}late\text{-}time} + N_\ell. \tag{34}$$

Here, C_ℓ^{TT} is the lensed CMB temperature power spectrum, $C_\ell^{\rm kSZ\text{-}late\text{-}time}$ is the low redshift contribution to kSZ and finally N_ℓ is the instrumental noise power spectrum of the CMB map, which is modelled as

$$N(\ell) = s^2 \exp\left[\frac{\ell(\ell+1)\theta^2_{\rm FWHM}}{8\ \ln 2}\right], \tag{35}$$

where s labels the sensitivity of the instrument and $\theta_{\rm FWHM}$ is the resolution. We also make a forecast for a configuration with noise and beam comparable to SO. To make this estimate we use a beam with a resulotion of 1.5 armcmin and an effective white noise level of 5.0 μK$-$arcmin, matching the set up in Ref. [43].

IV. FORECAST SETUP

In this section, we briefly describe the construction of the information matrix and the relevant systematics, focusing on the methodology used for forecasts on the cross-correlated data-sets. We then establish the models and parameter space over which the information matrix is constructed.

For our forecast, the measured modes are the large-scale modes of $[v(\boldsymbol{k}), \delta_h(\boldsymbol{k})]$ where $\{v(\boldsymbol{k})\}$ are the kSZ velocity reconstruction modes and $\{\delta_h(\boldsymbol{k})\}$ are large-scale halo overdensity modes. The halo-overdensity modes are obtained from the survey data set assuming that each halo is occupied by exactly one galaxy. Therefore, the signal and noise matrices are

$$\mathbf{S}(k,\mu,z) = \begin{pmatrix} P_{vv} & P_{vh} \\ P_{vh} & P_{hh} \end{pmatrix}, \tag{36}$$

$$\mathbf{N}(k,\mu,z) = \begin{pmatrix} N_{vv} & 0 \\ 0 & N_{hh} \end{pmatrix}. \tag{37}$$

The covariance matrix of our measured signal is the sum of the above two matrices,

$$\mathbf{C}(k,\mu,z) = \mathbf{S}(k,\mu,z) + \mathbf{N}(k,\mu,z). \tag{38}$$

The information matrix at redshift bin z_* is, therefore,

$$\begin{aligned} F_{ab} &= \frac{V}{2}\int \frac{d^3k}{(2\pi)^3} {\rm Tr}[\mathbf{C}(\boldsymbol{k})_{,a}\mathbf{C}(\boldsymbol{k})^{-1}\mathbf{C}(\boldsymbol{k})_{,b}\mathbf{C}(\boldsymbol{k})^{-1}] \\ &= \frac{V}{2}\int \frac{k^2 dk\ d\mu}{(2\pi)^2} {\rm Tr}[\mathbf{C}(\boldsymbol{k})_{,a}\mathbf{C}(\boldsymbol{k})^{-1}\mathbf{C}(\boldsymbol{k})_{,b}\mathbf{C}(\boldsymbol{k})^{-1}], \end{aligned} \tag{39}$$

where we have accounted for the fact that the covariance-matrix elements are only dependent on k and μ, with the latter being induced by the kSZ based velocity reconstruction and the inclusion of photo-z errors. We assume that the integral can be performed from a lower limit $k_{\rm min} \equiv \pi/V^{1/3}$, restricted by the survey volume V, to an upper limit $k_{\rm max} \approx 10^{-1}$ Mpc^{-1}.

The final models for $P_{hh}(k,\mu,z)$ and $P_{vh}(k,\mu,z)$, as they appear in the covariance matrix, are:

$$P_{vh}(k) = \frac{b_v f a H}{k}\left[b_{h,\rm RSD} + \beta_f \frac{f_{\rm NL}}{\alpha(k)}\right] P_{mm}(k) \tag{40}$$

$$P_{hh}(k) = \left[b^2_{h,\rm RSD} + 2b_{h,\rm RSD}\beta_f \frac{f_{\rm NL}}{\alpha(k)} + \beta_f^2 \frac{\left(\frac{5}{6}\right)^2 \tau_{\rm NL}}{\alpha^2(k)}\right] P_{mm}(k), \tag{41}$$

where $b_{h,\mathrm{RSD}} = [b_h + f\mu^2]$, and the explicit dependence of some terms on z and μ have been dropped for ease of notation. These models are a direct result of the derivations in Sec. II.

To model the final signal term $P_{vv}(k,z)$, we use the relation between the velocity and matter power spectra introduced in Eq. (17),

$$P_{vv}(k,z) = \left(\frac{b_v f a H}{k}\right)^2 P_{mm}(k,z). \tag{42}$$

Here, we have introduced a the optical-depth degeneracy parameter b_v. This parameter, with an expected value of 1.0, is introduced to account for the fact that kSZ data allows for the measurement of the product of the P_{ge} and P_{gv}, which means that a constant factor of scale could be exchanged between the two while keeping the signal unchanged.

In summary, the measurement covariance matrix $\mathbf{C}(k,\mu,z)$ is constructed based on the above three models for $P_{hh}(k,\mu,z)$, $P_{vh}(k,\mu,z)$ and $P_{vv}(k,z)$. This is used to construct a 4×4 information matrix over the parameter space spanned by $[b_h,\ b_v,\ f_{\mathrm{NL}},\ \tau_{\mathrm{NL}}]$ with the fiducial values set to $[1.6,\ 1.0,\ 0.0,\ 0.0]$, respectively. We invert this matrix and marginalise over the parameters b_h and b_v to obtain error estimates for f_{NL} and τ_{NL}. We also experimented with marginalising over cosmological parameters, but found that these do not significantly change our error estimates.

It is important to note that the value of Π, under the curvaton model constraint $\tau_{\mathrm{NL}} = (6f_{\mathrm{NL}}/5)^2(1+\Pi)$, is not well-defined for the assumed null hypothesis, $f_{\mathrm{NL}} = \tau_{\mathrm{NL}} = 0$. Therefore, in our forecasts, we assume that the models presented in Eq. (42) represent one possible parametrization of non-Gaussian power spectra under a multi-field model of inflation. That is, we vary the parameters τ_{NL} and f_{NL} independently, around their fiducial values, to construct our information matrix and make our forecasts. These estimates are propagated to quote a constraint for the parameter r_{NL} introduced in Eq. (15).

The noise spectrum for the velocity reconstruction term N_{vv} is given by $\mu_L^{-2} N_{v_r}$, where the form of N_{v_r} was introduced in Eq. (30). For halos, we assume that the noise is given primarily by the galaxy shot noise along with photo-z errors. Photo-z errors can be implemented for halos by a convolution of the halo density field with a Gaussian kernel in the radial direction, the form of which was introduced in Eq. (29). The halo noise power spectrum is then

$$N_{hh}(k,\mu) = \frac{1}{W^2(k,\mu) n_{\mathrm{gal}}}, \tag{43}$$

where we have directly used the galaxy number density n_{gal} based on our assumption that the galaxy distribution has a one-to-one correspondence with the distribution of halos, on large scales.

TABLE I. Baseline configurations for the cross-correlated CMB and LSS experiments. Values for baseline 1 match the specifications of the VRO survey and CMB-S4. The values for baseline 2 are similar to those expected for DESI and SO. The survey volumes were kept the same across the two forecasts to emphasize the dependence of the forecasts on galaxy density and photo-z errors.

		baseline 1	baseline 2
redshift	z	1.0	1.0
survey volume	V	100 Gpc^3	100 Gpc^3
halo bias	b_h	1.6	1.6
galaxy density	n_{gal}	10^{-2} Mpc^{-3}	2×10^{-4} Mpc^{-3}
photo-z error	σ_z	0.06	-
CMB resolution	θ_{FWHM}	1.5 arcmin	1.5 arcmin
CMB sensitivity	s	1 μK−arcmin	5 μK−arcmin

V. FORECAST RESULTS

In this section we provide forecasts for different experimental configurations. In the first part we analyse two different baseline configurations and provide estimates based on our assumed specifications on the galaxy survey and CMB measurement instrumentation. We then consider one of these baselines and vary each of the esperimental parameters, in isolation, to display the effects of these variations on our ability to constrain f_{NL} and τ_{NL}.

A. Baseline Forecasts

To establish the parameter dependencies of our forecast, we first display the results from our information matrix analysis for the two sets of baseline experiments described in Sec. III. Their instrumental specifications have been summarized in Table I. Baseline 1 specifications were chosen to resemble the experimental configuration of VRO and CMB-S4. Similarly, baseline 2 corresponds to the combination of DESI and an SO-like CMB experiment.

Our forecasts on the aforementioned baseline configurations have been summarized in Table II. This table also includes constraints on the parameter r_{NL}, around a fiducial value of 0, corresponding to the assumed null hypothesis. To simplify our calculations, we assume a cubic geometry for the survey volume and for our kSZ formalism. Therefore, these forecasts do not include the effects of the time evolution of power spectra and biases on the light cone. For a complementary kSZ formalism using maps on the light-cone, see e.g., Refs. [67–69].

For the baseline 1 experiments, the improvement factor in our ability to measure f_{NL} and τ_{NL}, arising from the cross-correlation with kSZ data, is 10 and 195, respectively. For baseline two, the improvement is 2 and 5 for the two parameters, respectively. In both cases, the improvement factor in the correlation-coefficient between

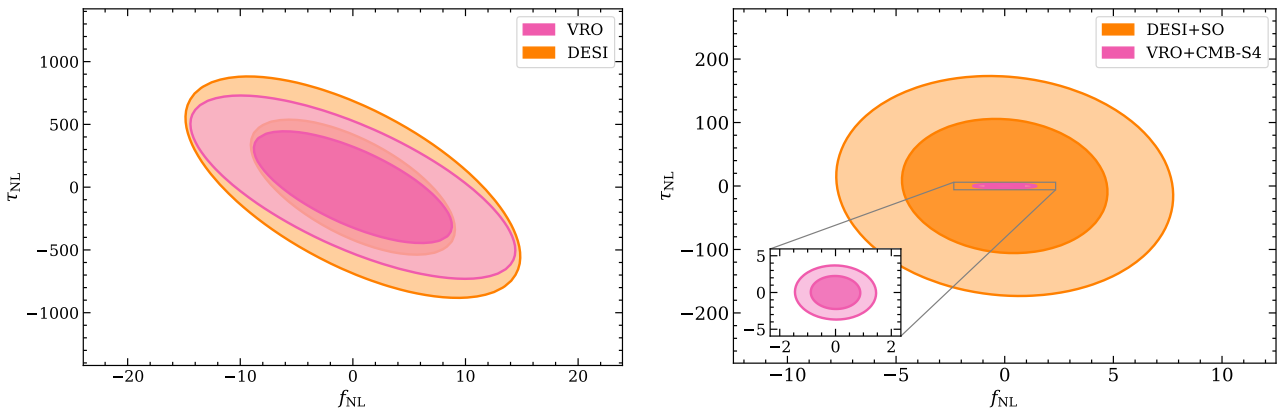

FIG. 1. Forecasted error ellipses on $f_{\rm NL}$ and $\tau_{\rm NL}$ at 68% and 95% confidence intervals, after marginalizing over b_h and b_v. *Left:* results when only galaxy survey data is considered. *Right:* results when velocity reconstruction data is added to the analysis. Each color corresponds to one of the baselines defined in Table I.

the two parameters is much higher. This is explained by the fact that the cross-correlation of the two data sets allows for the inclusion of the $P_{vh}(k)$ signal, which offers an independent constraint for $f_{\rm NL}$.

B. Experiment Parameter Variations

In order to assess which experimental limitations have the greatest impact on measurements of primordial non-Gaussianity, we isolate the effects of certain experimental parameters on our ability to constrain $f_{\rm NL}$ and $\tau_{\rm NL}$ by varying each parameter in isolation. For the following forecasts, we assume the baseline 1 configuration, the specifics of which are provided in Table II.

To highlight the scales that contribute most to the signal, we plot both $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$ as a function of the smallest measurable Fourier mode for our galaxy survey. This corresponds to varying the largest recoverable k-mode from the survey volume V, directly impacting the value of $k_{\rm min}$ as it appears in Eq. (39). These plots are displayed in Fig. 2. In both cases, the effects of sample variance cancellation become evident below $k \approx 2\times10^{-2}$ Mpc^{-1}. This behaviour can be explained by comparing the model for $P_{vv}(k)$ [Eq. (42)] to the assumed model for N_{vv}[derived from Eq. (28)]. These models indicate that the signal-to-noise ratio (SNR) of our velocity reconstruction is inversely proportional to k^2. Therefore, on small scales, we are only dependent the signal from $P_{hh}(k)$ to constrain both $f_{\rm NL}$ and $\tau_{\rm NL}$, causing the errors to coincide across the estimates from the single (galaxy) data-set and the cross-correlated (galaxy-kSZ) data-sets. In contrast, the higher SNR on larger scales allows us to constrain the non-Gaussian parameters using the models for both $P_{mh}(k)$ and $P_{hh}(k)$. Therefore, on these larger scales, the effects of sample variance cancellation are on full display.

To explore, more carefully, the information contained in the signal across both cases on large scales, we also display the dependence of $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$ on $k_{\rm min}$ when there is no contamination in our signal coming from shot noise, photo-z errors, and CMB instrument noise. These plots are displayed in Fig. 3.

At large values of k, we no longer see the 'Galaxy' curve coincide with the 'Galaxy+kSZ' results. This is a consequence of the much lower velocity reconstruction noise in the absence of shot noise. However, it is clear that an extension of the curves to smaller scales would reveal behaviour similar to the curves in Fig. 2, as a result of cosmic variance. The behaviour on larger scales is a lot more noteworthy for this set-up. In the case of $\sigma_{f_{\rm NL}}$, we see an inflection point at $k \approx 3\times10^{-3}$ Mpc^{-1}, after which the slope of the curve gets closer to zero. In contrast, when the galaxy survey data is combined with velocity reconstruction data, the forecast on $\sigma_{f_{\rm NL}}$ decreases steadily below $k \approx 10^{-3}$ Mpc^{-1}. This indicates that the large-scale modes contain a significant amount of data

TABLE II. Information matrix based estimates on the errors across the two parameters, $f_{\rm NL}$ and $\tau_{\rm NL}$. The results of error propagation to $r_{\rm NL}$ are also included. To arrive at these estimates, a mean value of 0 was chosen for both $f_{\rm NL}$ and $\tau_{\rm NL}$, while the halo bias was defined for each experiment as shown in Table I.

		baseline 1	baseline 2
$f_{\rm NL}$ error	$\sigma^{\rm gal}_{f_{\rm NL}}$	5.8	6.0
	$\sigma^{\rm kSZ+gal}_{f_{\rm NL}}$	5.9×10^{-1}	3.1
$\tau_{\rm NL}$ error	$\sigma^{\rm gal}_{\tau_{\rm NL}}$	2.9×10^{2}	3.6×10^{2}
	$\sigma^{\rm kSZ+gal}_{\tau_{\rm NL}}$	1.5	6.9×10^{1}
$r_{\rm NL}$ error	$\sigma^{\rm gal}_{r_{\rm NL}}$	2.0×10^{2}	2.5×10^{2}
	$\sigma^{\rm kSZ+gal}_{r_{\rm NL}}$	1.0	4.8×10^{1}

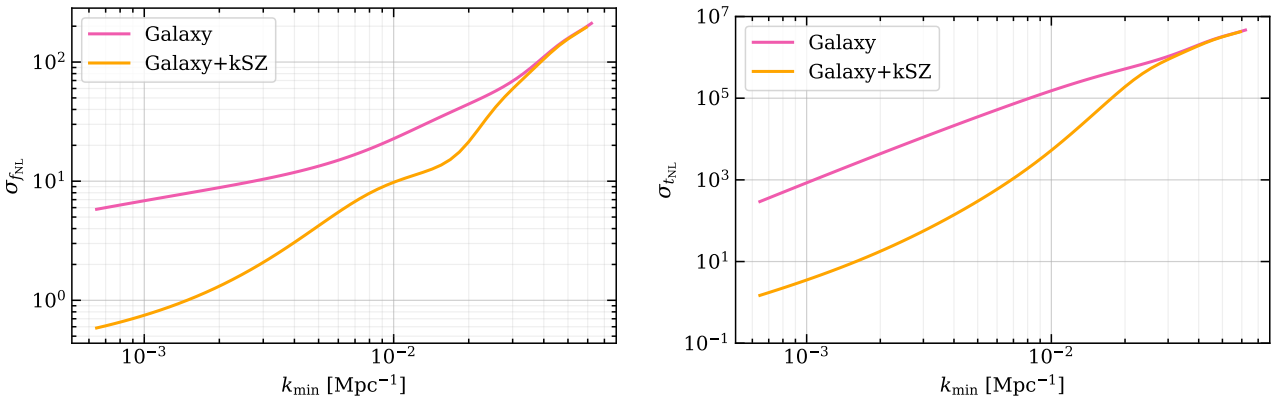

FIG. 2. *Left:* $\sigma_{f_{\rm NL}}$ as a function of $k_{\rm min}$ for baseline 1. *Right:* $\sigma_{\tau_{\rm NL}}$ as a function of $k_{\rm min}$ for baseline 1. The lower bound for k, on the left of the plots, is defined by the survey volume V. When the two data sets are cross-correlated, the error in $f_{\rm NL}$ and $\tau_{\rm NL}$ drastically decreases below $k \approx 2 \times 10^{-2}$ Mpc^{-1} (in comparison to the 'Galaxy' case), owing to the low noise in velocity reconstruction on large scales.

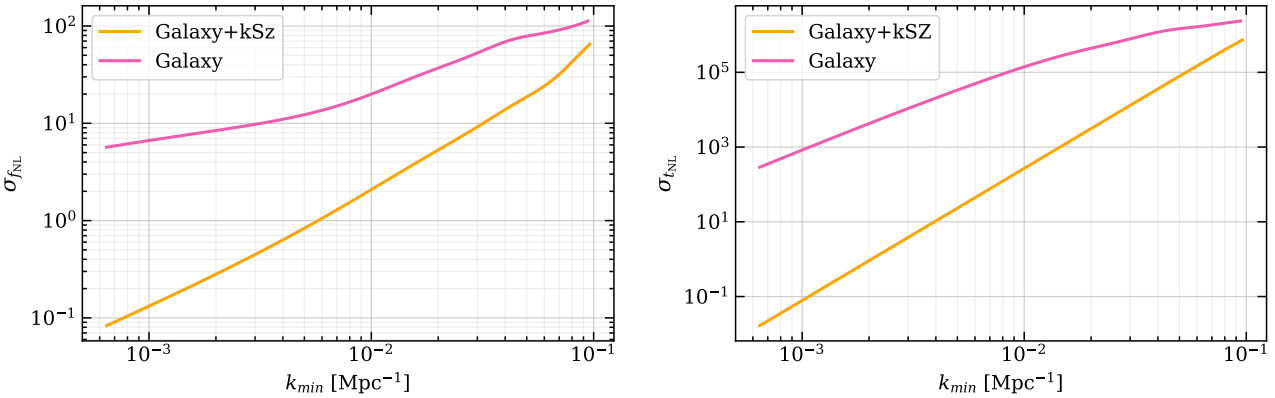

FIG. 3. *Left:* $\sigma_{f_{\rm NL}}$ as a function of $k_{\rm min}$ for baseline 1. *Right:* $\sigma_{\tau_{\rm NL}}$ as a function of $k_{\rm min}$ for baseline 1. For both the above sets of data we assume that the halo shot noise is zero and there are no photo-z errors. The lower bound for k, at the left end of the plots, is defined by the survey volume V. The behaviour in this unrealistic case matches the results in Fig. 2, where once again we see a sharp decrease in the error in $f_{\rm NL}$ and $\tau_{\rm NL}$ when the two data sets are cross-correlated due to high SNR in the velocity reconstruction on larger scales.

that allows us to better constrain $f_{\rm NL}$ with the cross-correlated data-sets. In contrast, the 'Galaxy' estimate of $\sigma_{\tau_{\rm NL}}$ never experiences the inflection point seen in the corresponding $\sigma_{f_{\rm NL}}$ curve, within the range of $k_{\rm min}$ plotted. However, the benefit of cross-correlating the data-sets is still evident in the relatively steeper decrease in $\sigma_{\tau_{\rm NL}}$ on larger scales. It is also important to note that when the two data sets are combined, the error in $\tau_{\rm NL}$ decreases more steeply than the error in $f_{\rm NL}$ i.e.; on the largest scales $\sigma_{\tau_{\rm NL}}$ reaches a minimum of $\sim 10^{-3}$ whereas $\sigma_{f_{\rm NL}}$ is only improved to $\sim 10^{-2}$.

The difference in the behaviour of $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$ as a function of scale $k_{\rm min}$ can be understood by analysing the non-Gaussian model for $P_{hh}(k)$ [Eq. (42)]. The contribution of $\tau_{\rm NL}$ to this signal comes from a term that is dependent on $\alpha(k)^{-2}$, or equivalently on k^{-4} [where $\alpha(k)$ is defined in Eq. (6)]. In contrast, the contribution of $f_{\rm NL}$ to this signal comes from a term that scales as k^{-2}. We conclude that, because the $P_{hh}(k)$ signal is dominated by the $\tau_{\rm NL}$ contribution, $\tau_{\rm NL}$ is constrained much better than $f_{\rm NL}$ on larger scales. Furthermore, this difference is most evident in the 'Galaxy+kSZ' case because an improved constraint on $f_{\rm NL}$ more directly translates to an improved constraint on $\tau_{\rm NL}$ due to the inclusion of the P_{mh} signal.

It is precisely this dependence on the large scales that explains the behaviour of the forecasts under varying galaxy number density $n_{\rm gal}$. The dependence of $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$ on varying values of $n_{\rm gal}$ has been plotted in Fig. 4. The solid lines correspond to results derived based on the experimental configuration corresponding to baseline 1 in Table I, including shot noise and photo-z errors. The

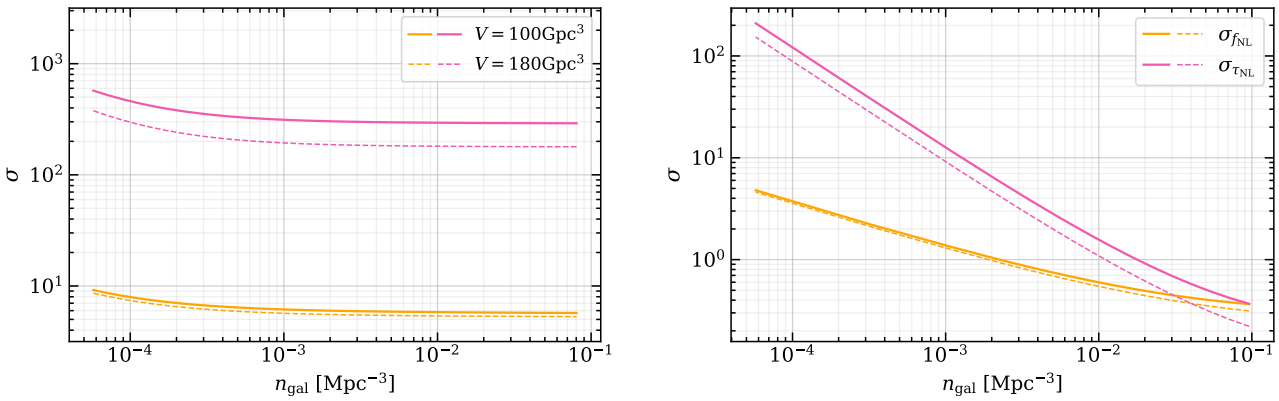

FIG. 4. *Left:* $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$ as a function of $n_{\rm gal}$, assuming that the galaxy data is used in isolation. *Right:* $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$ as a function of $n_{\rm gal}$, assuming that the galaxy data is cross-correlated with the kSZ data set. The solid (dashed) lines correspond to baseline 1 with a survey volume of 100 Gpc^3 (180 Gpc^3). Galaxy density not only defines the shape of the small scale power spectra but also defines the amount of shot noise in galaxy survey data. The sharp decrease in both $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$ for the cross-correlated data-set (right) is due to the fact that a high galaxy number density allows one to probe larger scales due to the lowered galaxy shot noise.

dashed lines represent results derived from the same set-up, with only the survey volume V updated to 180 Gpc^3. For clarity, the fiducial value of b_h was held constant.

When considering only the galaxy survey data, the results displayed on the left in Fig. 4 show that although the estimates improve slightly with decreasing shot noise, the cosmic variance limit is quickly reached for both our estimates of $f_{\rm NL}$ and $\tau_{\rm NL}$, irrespective of the assumed survey volume. This is because, using galaxy survey data alone forces us to constrain both $f_{\rm NL}$ and $\tau_{\rm NL}$ using the model for $P_{hh}(k)$ in isolation. Furthermore, as seen in Fig. 3, the slope in the $\sigma_{f_{\rm NL}}$ error is closer to zero with inclusion of low k-modes. Therefore, although the lower shot noise (higher $n_{\rm gal}$) allows us to probe larger and larger scales, the ability to constrain the non-Gaussian parameters eventually plateaus, as is seen in our results.

However, when the galaxy survey data is cross-correlated with the kSZ data, the lowered shot noise has a much more pronounced impact on both $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$ (right of Fig. 4). In fact, the effect of the higher $n_{\rm gal}$ on $\sigma_{\tau_{\rm NL}}$ is much steeper, with the results indicating that at high enough galaxy number density one can constrain $\tau_{\rm NL}$ better than $f_{\rm NL}$. This is because a higher number density allows for the use of more signal from large-scale k-modes. When cross-correlating the two data sets, the inclusion of these modes allows for a steady improvement in the ability to measure $f_{\rm NL}$ in the absence of noise, as shown in Fig. 3. This improved constraint on $f_{\rm NL}$ translates into a better measurement of $\tau_{\rm NL}$. This, combined with the difference in the contribution of each of these terms to the P_{hh} signal, allows for a tighter constraint on $\tau_{\rm NL}$ than $f_{\rm NL}$ when larger scales can be included as a result of lower shot noise.

The threshold value of $n_{\rm gal}$ required to measure $\tau_{\rm NL}$ with a higher sensitivity than $f_{\rm NL}$ is dependent on the survey volume. The results from the two different survey volumes, presented on the right of Fig. 4, indicate that the two uncertainty curves for $f_{\rm NL}$ and $\tau_{\rm NL}$ intersect at lower values of $n_{\rm gal}$ for larger survey volumes, as expected. While a survey with $V = 100$ Gpc^3 (under the baseline 1 configuration) can only achieve $\sigma_{\tau_{\rm NL}} \lesssim \sigma_{f_{\rm NL}}$ with $n_{\rm gal} \approx 10^{-1}$ Mpc^{-3}, increasing the survey volume to $V = 180$ Gpc^3 allows for the sensitivities to intersect at a more achievable $n_{\rm gal} \approx 3 \times 10^{-2}$ Mpc^{-3}.

To establish the dependence of these errors on CMB instrumental noise [Eq. (35)], we also calculated the values of $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$ under varying values of sensitivity s and resolution $\theta_{\rm FWHM}$, independently. In both these calculations we assumed the baseline 1 configuration for all other parameters. In our results we find that varying the sensitivity from 0.25 μK-arcmin to 14 μK-arcmin approximately increases our error in $f_{\rm NL}$ by 3.1x and $\tau_{\rm NL}$ by 6.3x. In contrast, the dependence of the errors on the CMB telescope resolution is a more pronounced. When the resolution is varied from 0.5-10 arcmin, $\sigma_{f_{\rm NL}}$ steadily increases by a factor of 9.8. Similarly, the forecasted error in the estimation of $\tau_{\rm NL}$ increases steeply by a factor of 100 for the same variation in CMB telescope resolution.

Finally, the values of $\sigma_{f_{\rm NL}}$ and $\sigma_{\tau_{\rm NL}}$ were also calculated for varying values of photo-z error σ_z. The value of σ_z was varied from 0.0 to 1.0, which approximately resulted an increase in $\sigma_{f_{\rm NL}}$ by a factor of 3.2 and an increase in $\sigma_{\tau_{\rm NL}}$ by a factor of 2.5. This minimal effect of varying σ_z is explained by the scale dependence of the $f_{\rm NL}$ and $\tau_{\rm NL}$ terms that makes the constraints most dependent on the largest scales measured.

VI. CONCLUSIONS

kSZ tomography is a powerful probe of the large-scale matter distribution that will be accessible with the next generation CMB and large-scale structure surveys. Cross-correlating this data set, with galaxy distribution data from upcoming large-scale structure surveys, such as the VRO survey and DESI, leads to sample variance cancellation in the measurement of galaxy bias and other quantities. In this paper, we have calculated the sensitivity with which both $f_{\rm NL}$ and $\tau_{\rm NL}$ can be measured using this method of cross-correlation. We also display the improvement coming from the addition of the kSZ data set and identify the experimental factors that most prominently contribute to better sensitivity in our measurements.

The statistical power of this method is most evident at large scales ($k < 10^{-2}$ Mpc^{-1}), arising from the low noise in the velocity reconstruction from kSZ data. For a cross-correlation between VRO survey data and CMB-S4, we find that one can reach $\sigma_{f_{\rm NL}} \approx 0.59$ and $\sigma_{\tau_{\rm NL}} \approx 1.5$ and improvement factors of 10 and 195, respectively, in comparison to estimates that use VRO data alone, without internal sample variance cancellation. Similarly, for the combination of DESI and a SO like survey we calculate $\sigma_{f_{\rm NL}} \approx 3.1$ and $\sigma_{\tau_{\rm NL}} \approx 69$, with a corresponding improvement factors of 2 and 5, respectively. This forecast includes marginalization over all relevant parameters and realistic photo-z errors as well as redshift space distortions. In our analysis of the experimental parameters which most heavily influence our sensitivity to measuring the scale-dependent bias, we find that the best results are achieved when the galaxy survey data is obtained from a large survey volume, with well measured large-scale modes, in combination with a high galaxy number density count. Furthermore, we expect that binning the galaxy survey data by mass, population, or redshift, to achieve internal sample variance cancellation, will further improve sensitivity to both $\tau_{\rm NL}$ and $f_{\rm NL}$, following the analysis in [43, 44].

In our work we have used a simplified 3-dimensional box geometry to illustrate the properties of the method and highlight the potential to measure signatures of non-Gaussianity using kSZ tomography data. We assume a fixed, functional form for β_f [Eq. (12)] to explicitly display forecasts on $f_{\rm NL}$ and $\tau_{\rm NL}$ alone. However, the dependence of β_f on b_h may require further simulation-based analysis for a non-Gaussian universe (see Ref. [70]). Although there are other modelling assumptions intrinsic to our calculation of velocity reconstruction noise, such as the assumed distribution of electron gas within halos, we expect a marginalisation over these model parameters to have a minimal impact on our sensitivity to $f_{\rm NL}$ and $\tau_{\rm NL}$. In these forecasts we also account for optical depth degeneracy via an added parameter in our information matrix; however, we expect that the measurements of electron profiles from fast-radio burst searches [71], as well as cross-correlation between radial and transverse velocities (latter reconstructed from so called 'moving-lens' tomography [72–74]) can potentially mitigate this bias in the near future. Although the inclusion of GR effects could lead to degeneracies with the existing $f_{\rm NL}$ and $\tau_{\rm NL}$ parameters (explored for the $f_{\rm NL}$ case in Refs. [75, 76]), we expect the effects of these degeneracies to be minimised by redshift binning or the consideration of multiple populations of halos.

We find that our forecasts compare well with other attempts at constraining local non-Gaussianity under sample variance cancellation, using different tools for cross-correlation. Our constraints on both $f_{\rm NL}$ and $\tau_{\rm NL}$ are slightly better than (within a factor of ~ 2 and ~ 3, respectively) the forecasts presented in Ref. [44], where sample variance cancellation was achieved by considering multiple populations of halos, assuming an LSST-type survey. Moreover, cross-correlation of reconstructed CMB lensing potential and galaxy clustering can also probe local type of non-Gaussianities, as showin in Ref. [77]. Their forecast on $f_{\rm NL}$, considering the survey combination of LSST and CMB-S4, with redshift binning, is comparable to the value presented in this paper. However, such forecasts would be sensitive to lensing reconstruction biases, which are likely more detrimental [78] than similar biases in kSZ tomography [69]. Including the CMB-lensing or moving-lens tomography [72–74] for additional sample variance cancellation from transverse modes could lead to some improvement, the analysis of which is left to future work. Ultimately, our forecasts in this paper indicate that kSZ tomography is a prominent tool for cross-correlation physics, allowing for impressive constraints on the PNG parameters in the curvaton scenario.

ACKNOWLEDGMENTS

We are thankful to Mesut Çalışkan for valuable discussions and help with editing the presentation of this work. SCH thanks Kendrick Smith and Mat Madhavacheril for useful conversations. GSP was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE1746891. SCH was supported by a Johns Hopkins Horizons Fellowship. This work was also supported by NSF Grant No. 1818899 and the Simons Foundation.

Appendix A: Halo Model

For the forecasts in this paper, we use the halo model to calculate the non-linear power spectra involving electron and galaxy fields. These power spectra are used to calculate the noise in our velocity reconstruction from the measured kSZ anisotropies. In this section, we present a short overview of this modelling methodology and present our modelling assumptions.

The halo model is dependent on the fundamental assumption that all the dark and baryonic matter is bound in halos of varying masses and density profiles. The correlation function for the matter-density or galaxy-density fluctuations then receives two contributions, one which accounts for the clustering of distinct halos ("two-halo" term) and another which accounts for the clustering within each individual halo ("one-halo" term). A review of this model can be found in Ref. [79].

1. Dark Matter

Although the non-linear power spectrum of dark matter clustering is not directly used in the velocity reconstruction estimates, we make assumptions on the clustering of these halos that define the form of the electron and galaxy power spectra. These specifics are described below.

Given the linear matter power spectrum $P_{mm}(k)$, and the cosmological matter density ρ_m (at the redshift of consideration), the rms variance of mass within a sphere of radius R that contains mass $m = 4\pi\rho_m R^3/3$ is defined as:

$$\sigma^2(m,z) = \frac{1}{2\pi^2}\int_0^\infty dk\ k^2\ P_{mm}(k,z)W^2(kR). \quad \text{(A1)}$$

Here, $R = R(m)$ and $W(kR)$ is the window function in Fourier space:

$$W(kR) = \frac{3[\sin(kR) - kR\cos(kR)]}{(kR)^3}. \quad \text{(A2)}$$

This is then used to define the halo mass function,

$$n(m,z) = f(\sigma,z)\frac{\rho_m}{m^2}\frac{d\ln[\sigma(m,z)^{-1}]}{d\ln(m)}, \quad \text{(A3)}$$

where m is the halo mass. This quantity denotes the number density of halos per mass interval, at a specific redshift z. For our calculations, we assume the Tinker collapse fraction [80]:

$$f(\sigma,z) = A\left[\left(\frac{\sigma}{b}\right)^{-a} + 1\right]e^{-c/\sigma^2}, \quad \text{(A4)}$$

with $A = 0.186$, $a = 1.47$, $b = 2.57$, and $c = 1.19$. The linear halo bias, consistent with the above collapse fraction, is assumed to be [81]:

$$b_h(\nu) = 1 + \frac{1}{\sqrt{a}\delta_c}\Big[\sqrt{a}(a\nu^2) + \sqrt{a}b(a\nu^2)^{1-c} - \frac{(a\nu^2)^c}{(a\nu^2)^c + b(1-c)(1-c/2)}\Big], \quad \text{(A5)}$$

where, in this model, $a = 0.707$, $b = 0.5$, $c = 0.6$, and we have defined $\nu(m,z) = \delta_c/\sigma(m,z)$. Note that these set of equations satisfy the consistency relation:

$$\int_{-\infty}^{\infty} d\ln m\ m^2 n(m,z)\Big(\frac{m}{\rho_m}\Big)b_h(m,z) = 1. \quad \text{(A6)}$$

2. Galaxies

The distribution of galaxies inside each halo is modelled according to the Halo Occupation Distribution (HOD) [82]. Under this model, we assume separate distributions for central and satellite galaxies, the forms of which are determined in [65].

The number of central galaxies in a halo is either 0 or 1. They are always located exactly at the halo's center. The mean number of centrals in a halo of mass m is fixed by the amount of stellar mass in each halo and is given by:

$$\bar{N}_c(m) = \frac{1}{2} - \frac{1}{2}\mathrm{erf}\Big[\frac{\log_{10}(m_\star^{\mathrm{thresh}})\ -\ \log_{10}(m_*(m))}{\sqrt{2}\sigma_{\log m_*}}\Big]. \quad \text{(A7)}$$

Here, $m_*(m)$ is the stellar mass in each halo of mass m and is modelled according to the form provided in equation 13 of Ref. [65]. The galaxy sample is defined by imposing a threshold stellar mass $m_\star^{\mathrm{thresh}}$ of observable galaxies. This model assumes a log normal distribution for stellar mass in a fixed halo of mass m, with a constant redshift independent scatter $\sigma_{\log m_*}$. For our calculations we set the value of this scatter to 0.2.

The mean number of satellite galaxies in a halo of mass m is given by:

$$\bar{N}_s(m) = \bar{N}_c(m)\Big(\frac{m}{m_{\mathrm{sat}}}\Big)^{\alpha_{\mathrm{sat}}}\exp\Big(\frac{-m_{\mathrm{cut}}}{m}\Big). \quad \text{(A8)}$$

The free parameters in this model, m_{sat}, α_{sat} and m_{cut}, depend on the choice of $m_\star^{\mathrm{thresh}}$. Their dependence on the threshold stellar mass is consistent with the 'SIGMOD1' model in Ref. [65] at redshift $z = 1$.

The total galaxy-galaxy power spectrum is the sum of the one halo and two halo contributions, which are defined as:

$$P_{gg}^{1h}(k,z) = \int_{-\infty}^{\infty} d\ln m \; \frac{mn(m,z)}{n_{\rm gal}^2}\Big[2\langle N_c(m)N_s(m)\rangle u_c(k)u_s(k|m,z) + \langle N_s(m)\;(N_s(m)-1)\rangle u_s(k|m,z)^2\Big],$$
$$P_{gg}^{2h}(k,z) = P_{mm}(k,z)\left[\int_{-\infty}^{\infty} d\ln m \; mn(m,z)b_h(m,z)\frac{\bar{N}_c(m)u_c(k)+\bar{N}_s(m)u_s(k|m,z)}{n_{\rm gal}}\right]^2. \tag{A9}$$

where $n_{\rm gal}$ is the mean number of galaxies in the simulated survey. It is dependent on the chosen value of $m_\star^{\rm thresh}$, and is defined as:

$$n_{\rm gal} = \int_{-\infty}^{\infty} d\;\ln m \; mn(m,z)[\bar{N}_s(m)+\bar{N}_c(m)]. \tag{A10}$$

Furthermore, $u_c(k)$ and $u_s(k|m,z)$ represent the Fourier space distribution profiles of centrals and satellite galaxies, respectively. Since we assume that the centrals are at exact halo centers, we set $u_c = 1$. We assume that the satellite galaxies follow an NFW profile:

$$\rho(r|m,z) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2}, \tag{A11}$$

where the scale radius r_s is related to the virial radius $r_{\rm vir}$ via the concentration parameter $c = r_{\rm vir}/r_s$. The mass and redshift dependence in this distribution arises from the assumed model for the concentration parameter:

$$c(m,z) = A\Big(\frac{m}{2\times 10^{12}\;h^{-1}M_\odot}\Big)^{\alpha}(1+z)^{\beta}, \tag{A12}$$

where $A = 7.85$, $\alpha = -0.081$, and $\beta = -0.71$ [57].

Finally, the expectation values $\langle N_c(m)N_s(m)\rangle$ and $\langle N_s(m)\;(N_s(m)-1)\rangle$, appearing in Eq. (A9), are defined to be $\bar{N}_s(m)$ and $\bar{N}_s(m)^2/\bar{N}_c(m)$, respectively, assuming $N_c(m)$ and $N_s(m)$ are maximally correlated.

3. Electrons

The electron distribution is modelled under the assumption that all the electron gas is bound within dark matter halos. Given this assumption, the auto-power spectrum of the electron gas is a sum of a one-halo and two-halo contribution, each of which is defined as:

$$P_{ee}^{1h}(k,z) = \int_{-\infty}^{\infty} d\ln m \; mn(m,z)\Big(\frac{m}{\rho_m}\Big)^2|u_e(k|m,z)|^2,$$
$$P_{ee}^{2h}(k,z) = P_{mm}(k,z)\left[\int_{-\infty}^{\infty} d\ln m \; mn(m,z)\Big(\frac{m}{\rho_m}\Big)b_h(m,z)u_e(k|m,z)\right]^2. \tag{A13}$$

Here, $u_e(k|m,z)$ refers to the Fourier-space distribution profile of the electron gas, which we assume to be a function of halo mass m and redshift z only. We use the AGN model-based fit function for the real-space mass distribution of the electron gas [83],

$$\rho_{\rm gas} = \frac{\Omega_b}{\Omega_m}\rho_c(z)\bar{\rho}_0\Big(\frac{x}{x_c}\Big)^{\gamma}\left[1+\Big(\frac{x}{x_c}\Big)^{\alpha}\right]^{-\frac{\beta-\gamma}{\alpha}}, \tag{A14}$$

where we have dropped the explicit dependence of some of the above parameters on mass and redshift for ease of notation. In the above model from Ref. [83], $x = r/R_{200}(m,z)$ where R_{200} is the radius at which the dark matter halo reaches a density of $200\rho_c(z)$. Furthermore, we have $\gamma = -0.2$ and $x_c = 0.5$. The remaining parameters $\bar{\rho}_0(m,z), \alpha(m,z)$, and $\beta(m,z)$ are fitted with a power law in halo mass and redshift:

$$A = A_0^x\left(\frac{m}{10^{14}M_\odot}\right)^{\alpha_m^x}(1+z)^{\alpha_z^x}, \tag{A15}$$

where the parameters for the AGN model used in this paper have been lifted from Table 2 of Ref. [83].

Given the auto - power spectra defined in Eq. (A9) and (A13) the cross spectra can be calculated as defined in Appendix B of Ref. [57]. One example set of spectra, constructed based on an $m_\star^{\rm thresh}$ value that generates a galaxy number density similar to that of VRO [58], has been shown in Fig. 5.

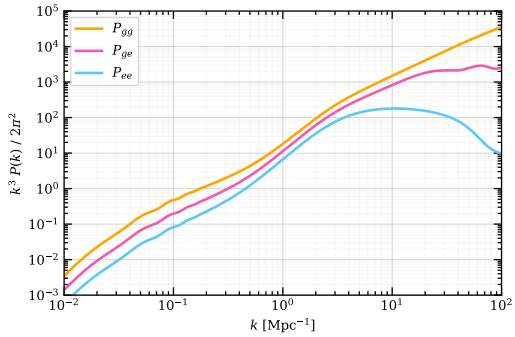

FIG. 5. Auto and cross power spectra in our halo model assuming the 'AGN' model of electron gas profile plotted at z=0. The shown galaxy auto-power spectrum was constructed to match the predicted galaxy number density for VRO.

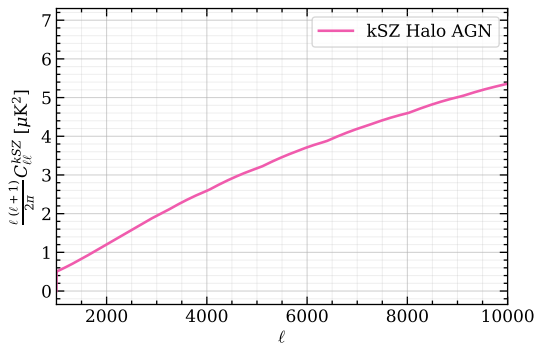

FIG. 6. CMB Power spectrum for kSZ for redshifts $0 < z < 6.5$ calculated in the halo model, assuming the AGN model based electron gas density profile.

4. kSZ Model

The late-time kSZ contribution to the CMB power spectrum [labelled $C_\ell^{\rm kSZ\text{-}late\text{-}time}$ in Eq. (34)] is also modelled based on the above power spectra. The kSZ angular power spectrum at large values of ℓ, where its contribution to the CMB spectrum is the largest, is dominated by the power spectrum of the transverse momentum field $P_{q\perp}(k)$ and is given by [84]

$$C_\ell^{\rm kSZ} = \frac{(\sigma_T \bar{n}_{e,0})^2}{2} \int \frac{d\chi}{\chi^2 a^4} e^{-2\tau} P_{q\perp}\Big(k = \frac{\ell}{\chi}, \chi\Big). \quad \text{(A16)}$$

We calculate the power spectrum of the transverse momentum field based on the form provided in [85]

$$P_{q\perp}(k,z) = (fHa)^2 \int_{-\infty}^{\infty} \frac{d^3k'}{(2\pi)^3} P_{ee}(|\boldsymbol{k} - \boldsymbol{k'}|, z) \frac{k(k-2k'\mu')(1-\mu'^2)}{k'^2(k^2+k'^2-2kk'\mu')}. \quad \text{(A17)}$$

A plot of the computed $C_\ell^{\rm kSZ}$ used in the forecasts has been presented in Fig. 6.

[1] J. Aguirre et al. (Simons Observatory), The Simons Observatory: Science goals and forecasts (2018), arXiv:1808.07445 [astro-ph.CO].

[2] A. Lee, M. H. Abitbol, S. Adachi, P. Ade, J. Aguirre, Z. Ahmed, S. Aiola, A. Ali, D. Alonso, M. A. Alvarez, and et al., The Simons Observatory, in Bulletin of the American Astronomical Society, Vol. 51 (2019) p. 147, arXiv:1907.08284 [astro-ph.IM].

[3] K. N. Abazajian et al. (CMB-S4), CMB-S4 Science Book, First Edition (2016), arXiv:1610.02743 [astro-ph.CO].

[4] K. Abazajian et al., Snowmass 2021 CMB-S4 White Paper, in 2022 Snowmass Summer Study (2022) arXiv:2203.08024 [astro-ph.CO].

[5] N. Sehgal et al., CMB-HD: An Ultra-Deep, High-Resolution Millimeter-Wave Survey Over Half the Sky (2019), arXiv:1906.10134 [astro-ph.CO].

[6] S. Aiola et al. (CMB-HD), Snowmass2021 CMB-HD White Paper (2022), arXiv:2203.05728 [astro-ph.CO].

[7] A. Aghamousa et al. (DESI), The DESI Experiment Part I: Science,Targeting, and Survey Design (2016), arXiv:1611.00036 [astro-ph.IM].

[8] D. J. Schlegel et al., Astro2020 APC White Paper: The MegaMapper: a z > 2 Spectroscopic Instrument for the Study of Inflation and Dark Energy (2019), arXiv:1907.11171 [astro-ph.IM].

[9] A. H. Guth, The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems, Phys. Rev. D **23**, 347 (1981).

[10] A. D. Linde, The new inflationary universe scenario. 1. High temperature effects (1982).

[11] A. Albrecht and P. J. Steinhardt, Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking, Phys. Rev. Lett. **48**, 1220 (1982).

[12] S. W. Hawking, The Development of Irregularities in a Single Bubble Inflationary Universe, Phys. Lett. B **115**, 295 (1982).

[13] A. H. Guth and S. Y. Pi, Fluctuations in the New Inflationary Universe, Phys. Rev. Lett. **49**, 1110 (1982).

[14] A. A. Starobinsky, Dynamics of Phase Transition in the New Inflationary Universe Scenario and Generation of Perturbations, Phys. Lett. B **117**, 175 (1982).

[15] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, Spontaneous Creation of Almost Scale - Free Density Pertur-

bations in an Inflationary Universe, Phys. Rev. D **28**, 679 (1983).

[16] D. Baumann and D. Green, Signatures of Supersymmetry from the Early Universe, Phys. Rev. D **85**, 103520 (2012), arXiv:1109.0292 [hep-th].

[17] V. Assassi, D. Baumann, and D. Green, On Soft Limits of Inflationary Correlation Functions, JCAP **11**, 047, arXiv:1204.4207 [hep-th].

[18] X. Chen and Y. Wang, Quasi-Single Field Inflation with Large Mass, JCAP **09**, 021, arXiv:1205.0160 [hep-th].

[19] S. Pi and M. Sasaki, Curvature Perturbation Spectrum in Two-field Inflation with a Turning Trajectory, JCAP **10**, 051, arXiv:1205.0161 [hep-th].

[20] T. Noumi, M. Yamaguchi, and D. Yokoyama, Effective field theory approach to quasi-single field inflation and effects of heavy fields, JHEP **06**, 051, arXiv:1211.1624 [hep-th].

[21] N. Arkani-Hamed and J. Maldacena, Cosmological Collider Physics (2015), arXiv:1503.08043 [hep-th].

[22] J.-O. Gong, S. Pi, and M. Sasaki, Equilateral non-Gaussianity from heavy fields, JCAP **11**, 043, arXiv:1306.3691 [hep-th].

[23] H. Lee, D. Baumann, and G. L. Pimentel, Non-Gaussianity as a Particle Detector, JHEP **12**, 040, arXiv:1607.03735 [hep-th].

[24] A. Kehagias and A. Riotto, On the Inflationary Perturbations of Massive Higher-Spin Fields, JCAP **07**, 046, arXiv:1705.05834 [hep-th].

[25] S. Kumar and R. Sundrum, Heavy-Lifting of Gauge Theories By Cosmic Inflation, JHEP **05**, 011, arXiv:1711.03988 [hep-ph].

[26] H. An, M. McAneny, A. K. Ridgway, and M. B. Wise, Quasi Single Field Inflation in the non-perturbative regime, JHEP **06**, 105, arXiv:1706.09971 [hep-ph].

[27] H. An, M. McAneny, A. K. Ridgway, and M. B. Wise, Non-Gaussian Enhancements of Galactic Halo Correlations in Quasi-Single Field Inflation, Phys. Rev. D **97**, 123528 (2018), arXiv:1711.02667 [hep-ph].

[28] D. Baumann, G. Goon, H. Lee, and G. L. Pimentel, Partially Massless Fields During Inflation, JHEP **04**, 140, arXiv:1712.06624 [hep-th].

[29] S. Kumar and R. Sundrum, Seeing Higher-Dimensional Grand Unification In Primordial Non-Gaussianities, JHEP **04**, 120, arXiv:1811.11200 [hep-ph].

[30] G. Goon, K. Hinterbichler, A. Joyce, and M. Trodden, Shapes of gravity: Tensor non-Gaussianity and massive spin-2 fields, JHEP **10**, 182, arXiv:1812.07571 [hep-th].

[31] D. Anninos, V. De Luca, G. Franciolini, A. Kehagias, and A. Riotto, Cosmological Shapes of Higher-Spin Gravity, JCAP **04**, 045, arXiv:1902.01251 [hep-th].

[32] S. Kumar and R. Sundrum, Cosmological Collider Physics and the Curvaton, JHEP **04**, 077, arXiv:1908.11378 [hep-ph].

[33] A. Hook, J. Huang, and D. Racco, Searches for other vacua. Part II. A new Higgstory at the cosmological collider, JHEP **01**, 105, arXiv:1907.10624 [hep-ph].

[34] Y. Akrami et al. (Planck), Planck 2018 results. IX. Constraints on primordial non-Gaussianity, Astron. Astrophys. **641**, A9 (2020), arXiv:1905.05697 [astro-ph.CO].

[35] D. Jeong and M. Kamionkowski, Clustering Fossils from the Early Universe, Phys. Rev. Lett. **108**, 251301 (2012), arXiv:1203.0302 [astro-ph.CO].

[36] D. Tseliakhovich, C. Hirata, and A. Slosar, Non-Gaussianity and large-scale structure in a two-field inflationary model, Phys. Rev. D **82**, 043531 (2010), arXiv:1004.3302 [astro-ph.CO].

[37] X. Chen and Y. Wang, Quasi-Single Field Inflation and Non-Gaussianities, JCAP **04**, 027, arXiv:0911.3380 [hep-th].

[38] T. Suyama and S. Yokoyama, Extension of local-type inequality for the higher order correlation functions, JCAP **07**, 033, arXiv:1105.5851 [astro-ph.CO].

[39] A. Cooray, C. Li, and A. Melchiorri, The trispectrum of 21-cm background anisotropies as a probe of primordial non-Gaussianity, Phys. Rev. D **77**, 103506 (2008), arXiv:0801.3463 [astro-ph].

[40] D. Yamauchi and K. Takahashi, Probing higher-order primordial non-Gaussianity with galaxy surveys, Phys. Rev. D **93**, 123506 (2016), arXiv:1509.07585 [astro-ph.CO].

[41] T. Sekiguchi, T. Takahashi, H. Tashiro, and S. Yokoyama, Probing primordial non-Gaussianity with 21 cm fluctuations from minihalos, JCAP **02**, 033, arXiv:1807.02008 [astro-ph.CO].

[42] N. Bartolo, M. Liguori, and M. Shiraishi, Primordial trispectra and CMB spectral distortions, JCAP **03**, 029, arXiv:1511.01474 [astro-ph.CO].

[43] M. Münchmeyer, M. S. Madhavacheril, S. Ferraro, M. C. Johnson, and K. M. Smith, Constraining local non-Gaussianities with kinetic Sunyaev-Zel'dovich tomography, Phys. Rev. D **100**, 083508 (2019), arXiv:1810.13424 [astro-ph.CO].

[44] S. Ferraro and K. M. Smith, Using large scale structure to measure f_{NL}, g_{NL} and τ_{NL}, Phys. Rev. D **91**, 043506 (2015), arXiv:1408.3126 [astro-ph.CO].

[45] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, Constraints on local primordial non-Gaussianity from large scale structure (2008), arXiv:0805.3580 [astro-ph].

[46] R. A. Sunyaev and Y. B. Zeldovich, The Velocity of clusters of galaxies relative to the microwave background. The Possibility of its measurement, Mon. Not. Roy. Astron. Soc. **190**, 413 (1980).

[47] Y. B. Zeldovich and R. A. Sunyaev, The Interaction of Matter and Radiation in a Hot-Model Universe, Astrophys. Space Sci. **4**, 301 (1969).

[48] Y. B. Zeldovich, Gravitational instability: An Approximate theory for large density perturbations, Astron. Astrophys. **5**, 84 (1970).

[49] R. A. Sunyaev and Y. B. Zeldovich, Microwave background radiation as a probe of the contemporary structure and history of the universe, Ann. Rev. Astron. Astrophys. **18**, 537 (1980).

[50] S. Y. Sazonov and R. A. Sunyaev, Microwave polarization in the direction of galaxy clusters induced by the CMB quadrupole anisotropy, Mon. Not. Roy. Astron. Soc. **310**, 765 (1999), arXiv:astro-ph/9903287.

[51] P.-J. Zhang and U.-L. Pen, Deprojecting Sunyaev-Zeldovich statistics, Astrophys. J. **549**, 18 (2001), arXiv:astro-ph/0007462.

[52] S. Ho, S. Dedeo, and D. Spergel, Finding the Missing Baryons Using CMB as a Backlight (2009), arXiv:0903.2845 [astro-ph.CO].

[53] J. Shao and T. Fang, The kinetic Sunyaev-Zel'dovich tomography II: probing the circumgalactic medium, Mon. Not. Roy. Astron. Soc. **458**, 3773 (2016), arXiv:1602.08932 [astro-ph.CO].

[54] P. Zhang and A. Stebbins, Confirmation of the Copernican Principle at Gpc Radial Scale and above from the Ki-

netic Sunyaev Zel'dovich Effect Power Spectrum, Phys. Rev. Lett. **107**, 041301 (2011), arXiv:1009.3967 [astro-ph.CO].

[55] D. Munshi, I. T. Iliev, K. L. Dixon, and P. Coles, Extracting the late-time kinetic Sunyaev–Zel'dovich effect, Mon. Not. Roy. Astron. Soc. **463**, 2425 (2016), arXiv:1511.03449 [astro-ph.CO].

[56] U. Seljak, Extracting primordial non-gaussianity without cosmic variance, Phys. Rev. Lett. **102**, 021302 (2009), arXiv:0807.1770 [astro-ph].

[57] K. M. Smith, M. S. Madhavacheril, M. Münchmeyer, S. Ferraro, U. Giri, and M. C. Johnson, KSZ tomography and the bispectrum (2018), arXiv:1810.13423 [astro-ph.CO].

[58] P. A. Abell et al. (LSST Science, LSST Project), LSST Science Book, Version 2.0 (2009), arXiv:0912.0201 [astro-ph.IM].

[59] K. N. Abazajian et al. (CMB-S4), CMB-S4 Science Book, First Edition (2016), arXiv:1610.02743 [astro-ph.CO].

[60] A. Aghamousa et al. (DESI), The DESI Experiment Part I: Science,Targeting, and Survey Design (2016), arXiv:1611.00036 [astro-ph.IM].

[61] N. Aghanim et al. (Planck), Planck 2018 results. VI. Cosmological parameters, Astron. Astrophys. **641**, A6 (2020), [Erratum: Astron.Astrophys. 652, C4 (2021)], arXiv:1807.06209 [astro-ph.CO].

[62] S. Cole and N. Kaiser, Biased clustering in the cold dark matter cosmogony (1989).

[63] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, The imprints of primordial non-gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects, Phys. Rev. D **77**, 123514 (2008), arXiv:0710.4560 [astro-ph].

[64] N. Kaiser, Clustering in real space and in redshift space (1987).

[65] A. Leauthaud et al., New constraints on the evolution of the stellar-to-dark matter connection: a combined analysis of galaxy-galaxy lensing, clustering, and stellar mass functions from z=0.2 to z=1, Astrophys. J. **744**, 159 (2012), arXiv:1104.0928 [astro-ph.CO].

[66] A. Leauthaud, J. Tinker, P. S. Behroozi, M. T. Busha, and R. Wechsler, A theoretical framework for combining techniques that probe the link between galaxies and dark matter, Astrophys. J. **738**, 45 (2011), arXiv:1103.2077 [astro-ph.CO].

[67] A. Terrana, M.-J. Harris, and M. C. Johnson, Analyzing the cosmic variance limit of remote dipole measurements of the cosmic microwave background using the large-scale kinetic Sunyaev Zel'dovich effect, JCAP **02**, 040, arXiv:1610.06919 [astro-ph.CO].

[68] A.-S. Deutsch, E. Dimastrogiovanni, M. C. Johnson, M. Münchmeyer, and A. Terrana, Reconstruction of the remote dipole and quadrupole fields from the kinetic Sunyaev Zel'dovich and polarized Sunyaev Zel'dovich effects, Phys. Rev. D **98**, 123501 (2018), arXiv:1707.08129 [astro-ph.CO].

[69] J. Cayuso, R. Bloch, S. C. Hotinli, M. C. Johnson, and F. McCarthy, Velocity reconstruction with the cosmic microwave background and galaxy surveys, arXiv:2111.11526 [astro-ph.CO].

[70] A. Barreira, Can we actually constrain $f_{\rm NL}$ using the scale-dependent bias effect? An illustration of the impact of galaxy bias uncertainties using the BOSS DR12 galaxy power spectrum (2022), arXiv:2205.05673 [astro-ph.CO].

[71] M. S. Madhavacheril, N. Battaglia, K. M. Smith, and J. L. Sievers, Cosmology with the kinematic Sunyaev-Zeldovich effect: Breaking the optical depth degeneracy with fast radio bursts, Phys. Rev. D **100**, 103532 (2019), arXiv:1901.02418 [astro-ph.CO].

[72] S. C. Hotinli, J. Meyers, N. Dalal, A. H. Jaffe, M. C. Johnson, J. B. Mertens, M. Münchmeyer, K. M. Smith, and A. van Engelen, Transverse Velocities with the Moving Lens Effect, Phys. Rev. Lett. **123**, 061301 (2019), arXiv:1812.03167 [astro-ph.CO].

[73] S. C. Hotinli, M. C. Johnson, and J. Meyers, Optimal filters for the moving lens effect, Phys. Rev. D **103**, 043536 (2021), arXiv:2006.03060 [astro-ph.CO].

[74] S. C. Hotinli, K. M. Smith, M. S. Madhavacheril, and M. Kamionkowski, Cosmology with the moving lens effect, Phys. Rev. D **104**, 083529 (2021), arXiv:2108.02207 [astro-ph.CO].

[75] R. Maartens, S. Jolicoeur, O. Umeh, E. M. De Weerd, and C. Clarkson, Local primordial non-Gaussianity in the relativistic galaxy bispectrum, JCAP **04**, 013, arXiv:2011.13660 [astro-ph.CO].

[76] A. Raccanelli, F. Montanari, D. Bertacca, O. Doré, and R. Durrer, Cosmological Measurements with General Relativistic Galaxy Correlations, JCAP **05**, 009, arXiv:1505.06179 [astro-ph.CO].

[77] M. Schmittfull and U. Seljak, Parameter constraints from cross-correlation of CMB lensing with galaxy clustering, Phys. Rev. D **97**, 123540 (2018), arXiv:1710.09465 [astro-ph.CO].

[78] G. Fabbian, A. Lewis, and D. Beck, CMB lensing reconstruction biases in cross-correlation with large-scale structure probes, JCAP **10**, 057, arXiv:1906.08760 [astro-ph.CO].

[79] A. Cooray and R. K. Sheth, Halo Models of Large Scale Structure, Phys. Rept. **372**, 1 (2002), arXiv:astro-ph/0206508.

[80] J. L. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. S. Warren, G. Yepes, S. Gottlober, and D. E. Holz, Toward a halo mass function for precision cosmology: The Limits of universality, Astrophys. J. **688**, 709 (2008), arXiv:0803.2706 [astro-ph].

[81] J. L. Tinker, B. E. Robertson, A. V. Kravtsov, A. Klypin, M. S. Warren, G. Yepes, and S. Gottlober, The Large Scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests, Astrophys. J. **724**, 878 (2010), arXiv:1001.3162 [astro-ph.CO].

[82] A. A. Berlind and D. H. Weinberg, The Halo occupation distribution: Towards an empirical determination of the relation between galaxies and mass, Astrophys. J. **575**, 587 (2002), arXiv:astro-ph/0109001.

[83] N. Battaglia, The Tau of Galaxy Clusters (2016), arXiv:1607.02442 [astro-ph.CO].

[84] E. T. Vishniac, Reionization and small-scale fluctuations in the microwave background, Astrophys. J. **322**, 597 (1987).

[85] W. Hu, Reionization revisited: secondary cmb anisotropies and polarization, Astrophys. J. **529**, 12 (2000), arXiv:astro-ph/9907103.