



HAL
open science

Super sample covariance of the thermal Sunyaev-Zel'dovich effect

Ken Osato, Masahiro Takada

► **To cite this version:**

Ken Osato, Masahiro Takada. Super sample covariance of the thermal Sunyaev-Zel'dovich effect. Physical Review D, 2021, 103 (6), pp.063501. 10.1103/PhysRevD.103.063501 . hal-02973148

HAL Id: hal-02973148

<https://hal.science/hal-02973148>

Submitted on 12 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Super sample covariance of the thermal Sunyaev-Zel'dovich effect

Ken Osato^{1,*} and Masahiro Takada²

¹*Institut d'Astrophysique de Paris, Sorbonne Université,
CNRS, UMR 7095, 98bis boulevard Arago, 75014 Paris, France*

²*Kavli Institute for the Physics and Mathematics of the Universe (WPI),
The University of Tokyo Institutes for Advanced Study, The University of Tokyo,
5-1-5 Kashiwanoha, Kashiwa, 277-8583 Chiba, Japan*



(Received 8 October 2020; accepted 1 February 2021; published 2 March 2021)

The thermal Sunyaev-Zel'dovich (tSZ) effect is a powerful probe of cosmology. The statistical errors in the tSZ power spectrum measurements are dominated by the presence of massive clusters in a survey volume that are easy to identify on an individual cluster basis. First, we study the impact of super sample covariance (SSC) on the tSZ power spectrum measurements, and find that the sample variance is dominated by the connected non-Gaussian (cNG) covariance arising mainly from Poisson number fluctuations of massive clusters in the survey volume. Second, we find that removing such individually detected, massive clusters from the analysis significantly reduces the cNG contribution, thereby causing the SSC to be a leading source of the sample variance. We then show, based on a Fisher analysis, that the power spectrum measured from the remaining diffuse tSZ effects can be used to obtain tight constraints on cosmological parameters as well as the hydrostatic mass bias parameter. Our method allows the use of both individual tSZ cluster counts and the power spectrum measurements of diffuse tSZ signals for cosmology and intracluster gas physics.

DOI: [10.1103/PhysRevD.103.063501](https://doi.org/10.1103/PhysRevD.103.063501)

I. INTRODUCTION

The anisotropy of the cosmic microwave background (CMB) is one of the most important probes of physical states in the Universe. Among different components of the CMB anisotropy, the Sunyaev-Zel'dovich (SZ) effect [1–3] provides us with rich information of cosmic structures of the late-time Universe (for reviews, see Refs. [4,5]). The SZ effect is divided into two classes according to the physical processes giving rise to the anisotropy. One is the thermal SZ effect (tSZ), where energy is transferred from hot electrons to CMB photons through inverse Compton scattering. The other one is the kinetic SZ effect (kSZ), which is caused by the Doppler effect due to the bulk motion of electrons with respect to the rest frame of CMB photons. The tSZ effect arises mainly from thermal electrons, most of which reside in galaxy clusters, and enable one to detect distant galaxy clusters up to high redshifts because the cosmological surface brightness dimming is compensated by the increase of the CMB photon number density at higher redshifts. Since the abundance and spatial distribution of galaxy clusters reflect the degree of structure formation of the Universe, we can take advantage of SZ effects to constrain cosmological parameters [6–10]. Since the amplitude of the tSZ effect is larger than that of the kSZ effect, and the tSZ effect is easier

to distinguish from other anisotropies due to specific frequency dependence, the tSZ effect has now been intensively studied.

The tSZ effect has already been measured by several survey programs. The *Planck* satellite has conducted measurements for all-sky coverage [11,12], and ground-based telescopes aim to measure the tSZ effect with high angular resolution (~ 1 arcmin), which includes the Atacama Cosmology Telescope (ACT) [13,14], South Pole Telescope [15,16], Simons Observatory [17], and CMB-S4 [18].

As a summary statistics of the tSZ effect, power spectrum is commonly employed because the theoretical approach is readily available [19]. Among them, the halo model prescription [20–23] can reproduce the results of hydrodynamical simulations [24,25] and is widely employed in practical analyses [11,26]. In addition, by cross correlating the tSZ map with other observables of large-scale structure, e.g., cosmic shear or nearby galaxies, the significance of the measurements can be enhanced [27–34]. Since the tSZ effect is sourced by hot free electrons and most of them are found in galaxy clusters, i.e., massive halos, the halo model picture is justified. Similarly, we can also estimate the covariance matrix based on the halo model prescription. However, the covariance matrix suffers from large-scale mass fluctuations with wavelengths comparable to or greater than the size of the survey volume, which is not a direct observable. This additional contribution to the covariance is referred to as the super sample covariance (SSC)

*ken.osato@iap.fr

[35]. If the SSC contribution is not taken into account, the covariance matrix is underestimated and the statistical significance is overestimated. In order to carry out the analysis with the power spectrum in an unbiased way, we need to incorporate such additional contributions.

In Refs. [26,36] it was shown that the connected non-Gaussian term is a dominant contribution to the covariance of the power spectrum. Furthermore, this term is sourced mainly by massive clusters at low redshifts. Therefore, removing such clusters from the measurement by masking the regions leads to a suppression of the covariance matrix and enhancement of the significance, though some fraction of signals can be missed [37]. In this paper, we estimate the SSC contribution for the tSZ effect based on the halo model prescription and quantify the effects on cosmological parameter estimation with the tSZ power spectrum. Furthermore, we also investigate the effect when massive clusters are masked.

This paper is organized as follows. First, we review the halo model prescription to compute the tSZ power spectrum in Sec. II and the covariance matrix in Sec. III. In Sec. IV, we give specific details on experimental conditions. In Sec. V, we present results of statistical significance and forecasts of constraints on parameters. We make concluding remarks in Sec. VI.

Throughout this paper, we assume the flat Λ cold dark matter (CDM) Universe. Unless otherwise stated, we adopt the best-fit cosmological parameters inferred with observations of temperature and polarization anisotropies and gravitational lensing of CMB (TT, TE, EE + lowE + lensing) measured by *Planck* [38,39]: the physical CDM density $\Omega_c h^2 = 0.12011$, the physical baryon density $\Omega_b h^2 = 0.022383$, the acoustic angular scale $100\theta_* = 1.041085$, the amplitude and slope of the scalar perturbation $\ln(10^{10} A_s) = 3.0448$, $n_s = 0.96605$ at the pivot scale $k_{\text{pivot}} = 0.05 \text{ Mpc}^{-1}$, and the optical depth $\tau_{\text{reio}} = 0.0543$. The neutrino component is composed of two massless neutrinos and one massive neutrino with mass $m_\nu = 0.06 \text{ eV}$, and the massive neutrino density parameter is $\Omega_\nu h^2 = m_\nu / (93.14 \text{ eV}) = 0.000644$. Thus, the matter component is composed of CDM, baryons, and massive neutrinos, and the matter density parameter is $\Omega_m = \Omega_c + \Omega_b + \Omega_\nu = 0.31816$. Then, the energy density of the cosmological constant is given as $\Omega_\Lambda = 1 - \Omega_m = 0.68184$. The amplitude of matter fluctuations at the scale $8 h^{-1} \text{ Mpc}$ is $\sigma_8 = 0.811$. For notation, the symbol “log” denotes the common logarithm “ \log_{10} ” and the symbol “ln” denotes the natural logarithm “ \log_e .”

II. THE tSZ POWER SPECTRUM

A. Halo model of the tSZ power spectrum

Here we briefly review the halo model calculations of the tSZ power spectrum. The temperature variation ΔT due to the tSZ effect [1–3] is given by

$$\frac{\Delta T}{T_{\text{CMB}}} = g_\nu(x)y = g_\nu(x) \frac{\sigma_T}{m_e c^2} \int P_e dl, \quad (1)$$

where y is the Compton y parameter, $T_{\text{CMB}} = 2.7255 \text{ K}$ is the CMB temperature, σ_T is the Thomson scattering cross section, m_e is the electron mass, c is the speed of light, P_e is the free electron pressure, and in the last term the line-of-sight integration is carried out with respect to the physical length l . The frequency-dependent part $g_\nu(x)$ is given by

$$g_\nu(x) = x \frac{e^x - 1}{e^x + 1} - 4, \quad x \equiv \frac{h\nu}{k_B T_{\text{CMB}}}, \quad (2)$$

where ν is the frequency, h is the Planck constant, and k_B is the Boltzmann constant.

First, we give formulations for the halo model calculation of the tSZ power spectrum [20,21]:

$$C(\ell) = C^{\text{1h}}(\ell) + C^{\text{2h}}(\ell), \quad (3)$$

$$C^{\text{1h}}(\ell) = \int_0^{z_{\text{reio}}} d\chi \frac{d^2 V}{d\chi d\Omega} \times \int_{M_{\text{min}}}^{M_{\text{max}}} dM \frac{dn_h}{dM}(M, z) |\tilde{y}(\ell; M, \chi)|^2, \quad (4)$$

$$C^{\text{2h}}(\ell) = \int_0^{z_{\text{reio}}} d\chi \frac{d^2 V}{d\chi d\Omega} P_L \left(k = \frac{\ell + 1/2}{\chi}; \chi \right) \times \left[\int_{M_{\text{min}}}^{M_{\text{max}}} dM \frac{dn_h}{dM}(M, z) b_h(M, z) \tilde{y}(\ell; M, \chi) \right]^2, \quad (5)$$

where χ is the comoving distance [40], $d^2 V / d\chi d\Omega = \chi^2$ is the comoving volume per unit comoving distance and unit solid angle, $dn_h / dM(M, z)$ is the halo mass function, $P_L(k; \chi)$ is the linear matter power spectrum at redshift z corresponding to the comoving radial distance $\chi(z)$, and $b_h(M, z)$ is the halo bias. The quantity $\tilde{y}(\ell; M, \chi)$ is the Fourier transform of the electron pressure profile. The modeling of the pressure profile of halos will be discussed later in this section. The one-halo term (4) corresponds to the correlation between the same halo and the two-halo term (5) denotes the correlation between different halos. We assume that reionization occurs instantaneously at $z_{\text{reio}} = 7$, and at this redshift all hydrogen and helium is fully ionized. We use the linear Boltzmann code CLASS [41] to compute the linear matter power spectrum $P_L(k, z)$. In the mass integration, we adopt the virial mass M_{vir} as the halo mass definition and employ $M_{\text{min}} = 10^{11} h^{-1} M_\odot$, $M_{\text{max}} = 10^{16} h^{-1} M_\odot$ for the lower and upper limits of the halo mass integral, respectively. The virial radius R_{vir} for a halo with virial mass M_{vir} is obtained from the expression

$$M_{\text{vir}} = \frac{4\pi}{3} \Delta_{\text{vir}}(z) \rho_{\text{cr}}(z) R_{\text{vir}}^3, \quad (6)$$

where $\rho_{\text{cr}}(z)$ is the critical density. The virial overdensity Δ_{vir} is given as [42]

$$\Delta_{\text{vir}} = 18\pi^2 + 82(\Omega_{\text{m}}(z) - 1) - 39(\Omega_{\text{m}}(z) - 1)^2, \quad (7)$$

where

$$\Omega_{\text{m}}(z) = \Omega_{\text{m}}(1+z)^3 E^{-2}(z), \quad (8)$$

and the expansion factor is defined as

$$E(z) = \frac{H(z)}{H_0} = [\Omega_{\text{m}}(1+z)^3 + \Omega_{\Lambda}]^{\frac{1}{2}}. \quad (9)$$

Since fitting formulas of the halo mass function, halo bias, and electron pressure profile adopt halo mass definitions that are different from the virial mass, we also use the alternative halo mass definitions M_{500} and M_{200b} :

$$M_{500} = \frac{4\pi}{3} 500 \rho_{\text{cr}}(z) R_{500}^3, \quad (10)$$

$$M_{200b} = \frac{4\pi}{3} 200 \rho_{\text{m}}(z) R_{200b}^3, \quad (11)$$

where $\rho_{\text{m}}(z) = \rho_{\text{cr}}(z) \Omega_{\text{m}}(z)$. These halo masses can be converted from the virial mass if we assume that the density profile of halos follows the Navarro-Frenk-White profile [43,44]:

$$\rho(r) = \frac{\rho_{\text{s}}}{(r/r_{\text{s}})(1+r/r_{\text{s}})^2}. \quad (12)$$

The scale density ρ_{s} is determined by the relation

$$M_{\text{vir}} = \int_0^{R_{\text{vir}}} dr 4\pi r^2 \rho(r), \quad (13)$$

and the scale radius r_{s} is determined from the mass-concentration relation

$$c_{\text{vir}}(M_{\text{vir}}, z) = \frac{R_{\text{vir}}}{r_{\text{s}}}, \quad (14)$$

where we adopt the fitting formula calibrated by N -body simulations in Ref. [45]. In addition, for the halo mass function and halo bias, we adopt the fitting formulas calibrated by N -body simulations: the fitted halo mass function with respect to M_{200b} in Ref. [46] and the fitted halo bias in Ref. [47].

Next, we derive the expression for the Fourier transform of the Compton y from a single halo $\tilde{y}(\ell; M, z)$:

$$\tilde{y}(\ell; M, z) = \frac{4\pi R_{\text{s}}}{\ell_s^2} \frac{\sigma_{\text{T}}}{m_{\text{e}} c^2} \times \int dx x^2 P_{\text{e}}(x) \frac{\sin((\ell + 1/2)x/\ell_s)}{(\ell + 1/2)x/\ell_s}, \quad (15)$$

where $x = r/R_{\text{s}}$, $\ell_s = D_{\text{A}}(\chi)/R_{\text{s}}$, R_{s} is the arbitrary scale radius, and $D_{\text{A}}(\chi)$ is the angular diameter distance. For the pressure profile of free electrons, we make use of the universal pressure profile [48] calibrated by *Planck* observations [49]:

$$\frac{P_{\text{e}}(r)}{P_{500}} = p(x) \left[\frac{M_{500}^{\text{HSE}}}{3 \times 10^{14} h_{70}^{-1} M_{\odot}} \right]^{0.12}, \quad (16)$$

$$p(x) \equiv \frac{P_0}{(c_{500} x)^{\gamma} [1 + (c_{500} x)^{\alpha}]^{(\beta-\gamma)/\alpha}}, \quad (17)$$

$$P_{500} = 1.65 \times 10^{-3} E(z)^{\frac{8}{3}} \times \left[\frac{M_{500}^{\text{HSE}}}{3 \times 10^{14} h_{70}^{-1} M_{\odot}} \right]^{\frac{2}{3}} h_{70}^2 \text{ keV cm}^{-3}, \quad (18)$$

where $(P_0, c_{500}, \gamma, \alpha, \beta) = (6.41, 1.81, 0.31, 1.33, 4.13)$, $x = r/R_{500}^{\text{HSE}}$, and $h_{70} \equiv h/0.7$. In the calibration of the pressure profile, the mass is determined assuming hydrostatic equilibrium, where only thermal pressure is balanced with the self-gravity of the halo. However, in addition to thermal pressure, nonthermal processes, e.g., turbulence or magnetic fields, could also contribute to the total pressure supporting the mass of galaxy clusters [50–53]. Thus, the hydrostatic mass M_{500}^{HSE} is generally lower than the true mass M_{500} . In order to take this effect into account, we parametrize the mass and radius with the hydrostatic mass bias parameter b_{HSE} :

$$M_{500}^{\text{HSE}} = (1 - b_{\text{HSE}}) M_{500}, \quad (19)$$

$$R_{500}^{\text{HSE}} = (1 - b_{\text{HSE}})^{\frac{1}{3}} R_{500}. \quad (20)$$

We adopt the fiducial value $b_{\text{HSE}} = 0.2$, which is suggested by mass calibration measurements [54,55] and hydrodynamical simulations [56]. The amplitude of the tSZ power spectrum is sensitive to the hydrostatic bias parameter, and thus the parameter can be constrained through the power spectrum in a way complementary to the mass-calibration measurements.

B. Selection function of massive clusters

Here we introduce an observable in the cluster survey that is relevant for the tSZ effect. For each cluster, the integrated flux of the tSZ effect corresponds to the thermal energy stored in the galaxy cluster. We define the three-dimensional integrated Compton y parameter, which is denoted as Y_{500} :

$$Y_{500}(M, z) = \frac{\sigma_T}{m_e c^2} \int_0^{R_{500}} dr 4\pi r^2 P_e(r; M, z). \quad (21)$$

This quantity is proportional to the thermal energy of gas in the galaxy cluster and can be measured in the tSZ survey if the redshift of the cluster is known [57]. Furthermore, this quantity exhibits a tight scaling relation with the halo mass [58]. As we will show later, since the dominant source of the tSZ covariance is caused by massive clusters, we study how the masking of such massive clusters, based on the integrated Compton y parameter, helps to reduce the sample covariance. We introduce the selection function $S(M, z)$ based on the integrated Compton y parameter:

$$S(M, z) = \begin{cases} 1 & (Y_{500}(M, z) \leq Y_{\text{thres}}), \\ 0 & (Y_{500}(M, z) > Y_{\text{thres}}), \end{cases} \quad (22)$$

where Y_{thres} is the threshold value. When cluster masking is applied, this selection function is inserted into the mass and redshift integrations in the halo model expressions (see Sec. III C). In Fig. 1 we show how the integrated Compton y parameter varies with the virial mass and redshift, and the different lines denote three representative values of $\log(Y_{500}/\text{Mpc}^2) = -6, -5, -4$. In Fig. 2 we show the tSZ power spectrum along with the CMB primary spectrum and noise power spectrum, which will be discussed in Sec. IV. The tSZ power spectra with and without cluster masking are shown. When cluster masking is applied, the amplitude is suppressed by less than half but the covariance matrix is also reduced, as shown in Sec. III. Both the one-halo (dashed lines) and two-halo (dot-dashed lines) terms are reduced by cluster masking, but the one-halo term is more suppressed because the main source of the term is

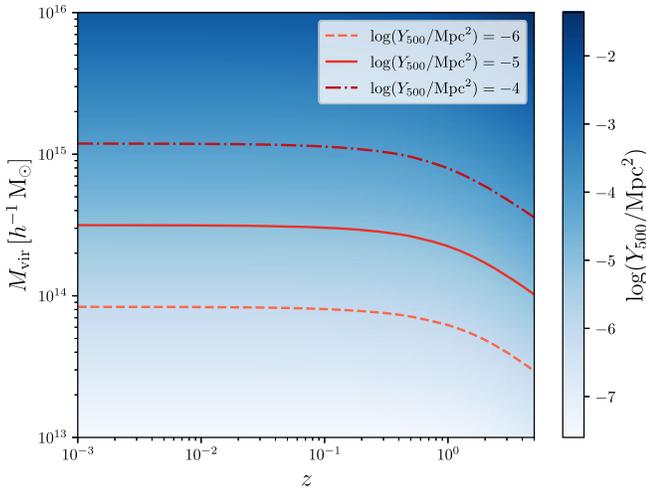


FIG. 1. Integrated Compton y parameter for clusters of given redshift z and virial mass M_{vir} . The dashed, solid, and dot-dashed lines correspond to $\log(Y_{500}/\text{Mpc}^2) = -6, -5, -4$, respectively. For each threshold, clusters located in the upper regions divided by the lines are masked.

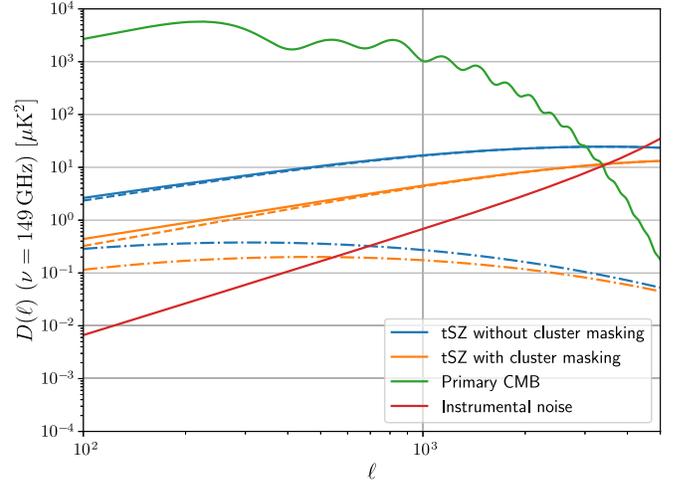


FIG. 2. Power spectra of temperature anisotropy due to the tSZ effect with and without cluster masking. For comparison, the primary CMB anisotropy power spectrum and instrumental noise power spectrum (see Sec. IV) are also shown. For halo model calculations, the dashed (dot-dashed) lines correspond to the one-halo (two-halo) term contribution. Instead of the raw power spectrum $C(\ell)$, we show $D(\ell) = \ell(\ell + 1)/(2\pi)C(\ell)g_V^2 T_{\text{CMB}}^2$ at the frequency $\nu = 149$ GHz. For cluster masking, we adopt the threshold $\log(Y_{\text{thres}}/\text{Mpc}^2) = -5$.

nearby massive clusters. As a result, the signal-to-noise ratio is enhanced by masking clusters. We will discuss the statistical significance in Sec. VA.

III. COVARIANCE OF THE tSZ POWER SPECTRUM

A. Compton y with the survey mask

Here we derive formulas for the covariance matrix of the Compton y power spectrum, taking into account the survey mask. First, we define the survey window function $W(\boldsymbol{\theta})$ and the observed Compton y field as

$$y_W(\boldsymbol{\theta}) = W(\boldsymbol{\theta})y(\boldsymbol{\theta}), \quad (23)$$

where $W(\boldsymbol{\theta}) = 1$ for observed regions and $W(\boldsymbol{\theta}) = 0$ for masked regions. Then, the Fourier transform of the field under the flat-sky approximation is given by

$$\tilde{y}_W(\ell) = \int \frac{d^2\ell'}{(2\pi)^2} \tilde{W}(\ell - \ell') \tilde{y}(\ell'), \quad (24)$$

where the tilde symbol denotes the Fourier transform of the quantity. The estimator of the power spectrum for the i th multipole bin is given as

$$\hat{C}(\ell_i) \equiv \frac{1}{\Omega_W} \int_{\ell \in \ell_i} \frac{d^2\ell}{\Omega_{\ell_i}} \tilde{y}_W(\ell) \tilde{y}_W(-\ell), \quad (25)$$

where Ω_W is the effective survey area $\Omega_W = \int d^2\theta W(\boldsymbol{\theta})$ and Ω_{ℓ_i} is the number of modes $\Omega_{\ell_i} = \int_{\ell \in \ell_i} d^2\ell$. When the

bin width is sufficiently smaller than the multipole, the number of modes can be approximated as

$$\Omega_{\ell_i} \approx 2\pi\ell_i^2 \Delta \ln \ell, \quad (26)$$

where we consider the logarithmically evenly spaced bins for multipoles and $\Delta \ln \ell$ is the bin width in logarithmic space. Then, the ensemble average of the estimator becomes the sum of the power spectrum $C(\ell)$ and noise power spectrum $N(\ell)$:

$$\langle \hat{C}(\ell_i) \rangle = C(\ell_i) + N(\ell_i). \quad (27)$$

Hereafter, the parentheses denote the ensemble average.

B. Covariance matrix of the tSZ power spectrum

Next, we derive the expression for the covariance matrix \mathcal{C} as

$$\begin{aligned} \mathcal{C}_{ij} &\equiv \langle \hat{C}(\ell_i) \hat{C}(\ell_j) \rangle - \langle \hat{C}(\ell_i) \rangle \langle \hat{C}(\ell_j) \rangle \\ &= \frac{1}{\Omega_W} \left[\frac{(2\pi)^2}{\Omega_{\ell_i}} 2[C(\ell_i) + N(\ell_i)]^2 \delta_{ij} + \bar{T}_W(\ell_i, \ell_j) \right], \end{aligned} \quad (28)$$

where δ_{ij} is the Kronecker delta. The first term is referred to as the Gaussian term (labeled as G):

$$\mathcal{C}_{ij}^G = \frac{1}{\Omega_W} \frac{(2\pi)^2}{\Omega_{\ell_i}} 2[C(\ell_i) + N(\ell_i)]^2 \delta_{ij}. \quad (29)$$

The remaining term is the non-Gaussian term, which does not depend on the binning of the multipoles, in contrast to the Gaussian term. The trispectrum convoluted with the window function \bar{T}_W is given as

$$\begin{aligned} \bar{T}_W(\ell_i, \ell_j) &= \frac{1}{\Omega_W} \int_{l \in \ell_i} \frac{d^2 \ell}{\Omega_{\ell_i}} \int_{l \in \ell_j} \frac{d^2 \ell'}{\Omega_{\ell_j}} \int \left[\prod_{a=1}^4 \frac{d^2 q_a}{(2\pi)^2} \tilde{W}(\mathbf{q}_a) \right] \\ &\quad \times (2\pi)^2 \delta_{\mathbb{D}}(\mathbf{q}_{1234}) T(\ell + \mathbf{q}_1, \\ &\quad -\ell + \mathbf{q}_2, \ell' + \mathbf{q}_3, -\ell' + \mathbf{q}_4), \end{aligned} \quad (30)$$

where $\delta_{\mathbb{D}}$ is the Dirac delta function and $\mathbf{q}_{1234} \equiv \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_4$.

Then, we derive the expressions for the covariance matrix based on the halo model [59]. First, we introduce the following notation:

$$I_{\mu}^{\beta}(\ell_1, \dots, \ell_{\mu}; \chi) \equiv \int dM \frac{dn_h}{dM} b_{\beta} \tilde{y}(\ell_1) \cdots \tilde{y}(\ell_{\mu}), \quad (31)$$

where $b_0 = 1$ and $b_1 = b_h(M, z)$. The halo model expression for the tSZ power spectrum becomes

$$\begin{aligned} C(\ell) &= \int d\chi \frac{d^2 V}{d\chi d\Omega} \left[I_2^0(\ell, \ell; \chi) + [I_1^1(\ell; \chi)]^2 \right. \\ &\quad \left. \times P_L\left(k = \frac{\ell + 1/2}{\chi}, \chi\right) \right]. \end{aligned} \quad (32)$$

For projected fields such as Compton y , we can compute the non-Gaussian terms for the covariance as follows [8,60,61]:

$$\mathcal{C}_{ij}^{\text{NG}} = \mathcal{C}_{ij}^{\text{cNG}} + \mathcal{C}_{ij}^{\text{SSC}}, \quad (33)$$

$$\mathcal{C}_{ij}^{\text{cNG}} = \frac{1}{\Omega_W} \int_{l \in \ell_i} \frac{d^2 \ell}{\Omega_{\ell_i}} \int_{l \in \ell_j} \frac{d^2 \ell'}{\Omega_{\ell_j}} T(\ell, -\ell, \ell', -\ell'), \quad (34)$$

$$\mathcal{C}_{ij}^{\text{SSC}} = \mathcal{C}_{ij}^{\text{HSV}} + \mathcal{C}_{ij}^{\text{HSV-BC}} + \mathcal{C}_{ij}^{\text{BC}}, \quad (35)$$

$$\begin{aligned} \mathcal{C}_{ij}^{\text{HSV}} &= \int d\chi \frac{d^2 V}{d\chi d\Omega} I_2^1(\ell_i, \ell_i; \chi) I_2^1(\ell_j, \ell_j; \chi) \\ &\quad \times \frac{1}{\Omega_W^2} \int \frac{d^2 \ell}{(2\pi)^2} |\tilde{W}(\ell)|^2 P_L(k; \chi) \\ &= \int d\chi \frac{d^2 V}{d\chi d\Omega} I_2^1(\ell_i, \ell_i; \chi) I_2^1(\ell_j, \ell_j; \chi) [\sigma_W^L(\chi)]^2, \end{aligned} \quad (36)$$

$$\begin{aligned} \mathcal{C}_{ij}^{\text{BC}} &= \int d\chi \frac{d^2 V}{d\chi d\Omega} \left(\frac{68}{21} \right)^2 [I_1^1(\ell_i; \chi) I_1^1(\ell_j; \chi)]^2 \\ &\quad \times P_L(k_i; \chi) P_L(k_j; \chi) [\sigma_W^L(\chi)]^2, \end{aligned} \quad (37)$$

$$\begin{aligned} \mathcal{C}_{ij}^{\text{HSV-BC}} &= \int d\chi \frac{d^2 V}{d\chi d\Omega} \frac{68}{21} \{ [I_1^1(\ell_i; \chi)]^2 I_2^1(\ell_j, \ell_j; \chi) P_L(k_i; \chi) \\ &\quad + [I_1^1(\ell_j; \chi)]^2 I_2^1(\ell_i, \ell_i; \chi) P_L(k_j; \chi) \} [\sigma_W^L(\chi)]^2, \end{aligned} \quad (38)$$

where we define the variance of mass density fluctuations of super-survey modes as

$$[\sigma_W^L(\chi)]^2 \equiv \frac{1}{\Omega_W^2} \int \frac{d^2 \ell}{(2\pi)^2} |\tilde{W}(\ell)|^2 P_L(k; \chi), \quad (39)$$

and $k \equiv (\ell + 1/2)/\chi$. The term that is sourced from the trispectrum with a parallelogram configuration is referred to as the connected non-Gaussian term (labeled as cNG). The latter three terms are referred to as the halo sample variance (HSV) [61] the beat coupling (BC) [60,62], and their cross correlation (HSV-BC) terms, respectively. The sum of the three terms is referred to as super sample variance (SSC) [35]. Since the trispectrum is dominated by the one-halo term at all scales [59], we ignore the two-, three-, and four-halo terms of the trispectrum for simplicity:

$$\begin{aligned} T(\ell, -\ell, \ell', -\ell') &\approx T^{\text{1h}}(\ell, \ell') \\ &= \int d\chi \frac{d^2 V}{d\chi d\Omega} \int dM \frac{dn_h}{dM} \\ &\quad \times |\tilde{y}(\ell; M, \chi)|^2 |\tilde{y}(\ell'; M, \chi)|^2. \end{aligned} \quad (40)$$

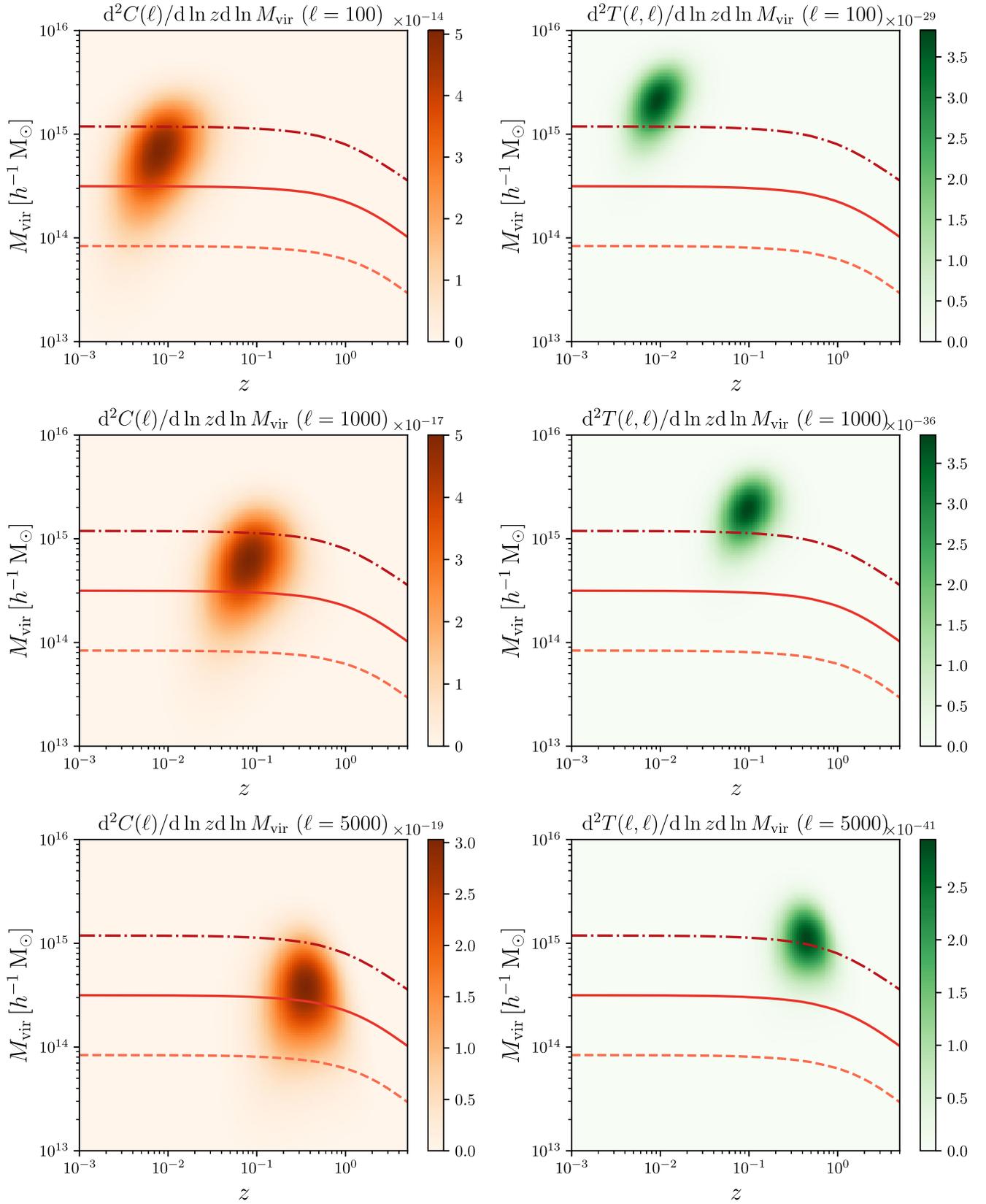


FIG. 3. Differential contributions of clusters in infinitesimal intervals of redshift and halo mass around a given redshift (x axis) and mass (y axis) to the power spectrum and the diagonal term of the trispectrum for $\ell = 100$ (upper panels), 1000 (medium panels), and 5000 (lower panels), respectively. The dashed, solid, and dot-dashed lines correspond to $\log(Y_{500}/\text{Mpc}^2) = -6, -5,$ and $-4,$ respectively.

C. Implementation of cluster masking in the halo model

Here we describe how to incorporate the cluster masking in the halo model expressions of the tSZ power spectrum and covariance matrix. The selection function $S(M, \chi)$ [Eq. (22)] is inserted into the mass integration in I_μ^β and T^{1h} :

$$\begin{aligned} I_\mu^\beta(\ell_1, \dots, \ell_\mu; \chi) &\rightarrow \hat{I}_\mu^\beta(\ell_1, \dots, \ell_\mu; \chi) \\ &\equiv \int dM \frac{dn_h}{dM} S(M, \chi) b_\beta \tilde{y}(\ell_1) \cdots \tilde{y}(\ell_\mu), \end{aligned} \quad (41)$$

and

$$\begin{aligned} T^{1h}(\ell, \ell') &\rightarrow \hat{T}^{1h}(\ell, \ell') \\ &\equiv \int d\chi \frac{d^2V}{d\chi d\Omega} \int dM \frac{dn_h}{dM} S(M, \chi) \\ &\quad \times |\tilde{y}(\ell; M, \chi)|^2 |\tilde{y}(\ell'; M, \chi)|^2. \end{aligned} \quad (42)$$

Then, in order to apply cluster masking, I_μ^β and T^{1h} in the halo model expressions (32), (33), (36), (37), and (37) are replaced with \hat{I}_μ^β and \hat{T}^{1h} , respectively.

D. Weight function

In order to investigate which halos contribute the signal and covariance, we compute the weight function with respect to mass and redshift. For the power spectrum, the weight function is given as

$$\begin{aligned} \frac{d^2}{dz dM} C(\ell) &= \frac{d^2V}{dz d\Omega} \left[\frac{dn_h}{dM} |\tilde{y}(\ell)|^2 + 2 \frac{dn_h}{dM} b_h \tilde{y}(\ell) \right. \\ &\quad \left. \times P_L \left(k = \frac{\ell + 1/2}{\chi}, z \right) \int dM \frac{dn_h}{dM} b_h \tilde{y}(\ell) \right]. \end{aligned} \quad (43)$$

Similarly, the weight function of the trispectrum is given as

$$\begin{aligned} \frac{d^2}{dz dM} T(\ell, -\ell, \ell', -\ell') &\approx \frac{d^2}{dz dM} T^{1h}(\ell, \ell') \\ &= \frac{d^2V}{dz d\Omega} \frac{dn_h}{dM} |\tilde{y}(\ell)|^2 |\tilde{y}(\ell')|^2. \end{aligned} \quad (44)$$

Note that these expressions are weight functions with respect to the redshift z , instead of the comoving distance χ , and the comoving volume with respect to redshift is $d^2V/dz d\Omega = \chi^2(z)H(z)/c$.

Figure 3 shows the weight functions for diagonal components, i.e., $\ell = \ell'$ of the power spectrum and trispectrum for three representative multipoles, $\ell = 100, 1000, 5000$. Obviously, the trispectrum is more sensitive to massive clusters than the power spectrum at all scales. The integrand of the mass integration of the trispectrum contains \tilde{y}^4 in contrast to \tilde{y}^2 at highest for power spectrum, and thus the

contribution from massive clusters becomes prominent in the trispectrum. Similarly, we can expect that the super sample covariance involves the mass integration of \tilde{y}^2 at highest and is less sensitive to massive clusters compared to the cNG term. We also show the three critical lines of $\log(Y_{500}/\text{Mpc}^2) = -6$ (dashed), -5 (solid), and -4 (dot-dashed). In order to enhance the significance of the detection, we need to keep the power spectrum but remove the trispectrum contribution as much as possible, which is the dominant source of the covariance. We adopt the fiducial threshold $\log(Y_{\text{thres}}/\text{Mpc}^2) = -5$ when masking massive clusters. Apparently, the threshold $\log(Y_{\text{thres}}/\text{Mpc}^2) = -4$ keeps the large fraction of the signal but our fiducial threshold gives a higher significance. Since the contribution of the cNG term is quite larger than the other terms, rather than maintaining the signal, completely excluding the trispectrum is a more effective strategy.

IV. EXPERIMENTAL CONDITIONS

In this section we define the experimental conditions for statistical analysis. We consider a practical case which is similar to the Advanced ACT measurement [63]. We assume that the sky coverage is $\Omega_w = 2100 \text{ deg}^2$ and the survey window is circularly symmetric for simplicity. The power spectrum of the mask is given as

$$|\tilde{W}(\ell)|^2 = \Omega_w^2 \left[2 \frac{J_1(\ell \Theta_w)}{\ell \Theta_w} \right]^2, \quad (45)$$

where $J_1(x)$ is the first-order Bessel function and we employ $\Theta_w \equiv \sqrt{\Omega_w/\pi}$. In general, the survey window function has an irregular shape or the survey regions are divided into multiple separate patches. The relative strength of SSC depends on the geometry of the survey region or the

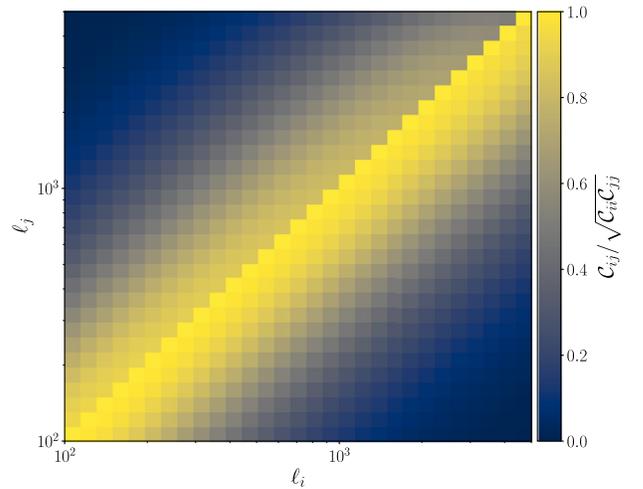


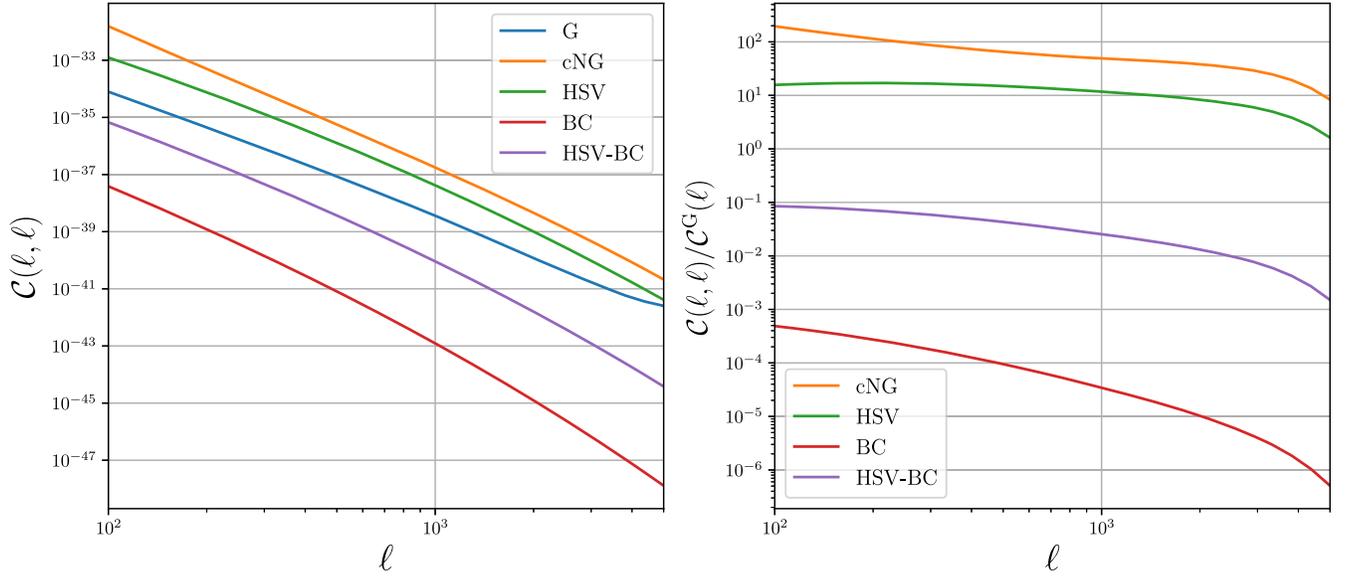
FIG. 4. Total covariance matrix of the tSZ power spectrum. The upper left (lower right) part corresponds to the covariance matrix with (without) cluster masking.

degree of discontinuity of survey regions [64]. A compact survey geometry such as the circular geometry considered here has the largest SSC contribution. Hence, the following estimate can be considered as the worst-case scenario for the impact of the SSC contribution. The binning is logarithmically evenly spaced with the minimum $\ell_{\min} = 100$, maximum $\ell_{\max} = 5000$, and number of bins $n_{\ell} = 30$. For the noise power spectrum, we assume that the instrumental noise is Gaussian and its variance is $\sigma_{\text{inst}} = 7 \mu\text{K arcmin}$ with a single band at $\nu_0 = 149 \text{ GHz}$. Then, the noise power spectrum is given as

$$N(\ell) = \left[\frac{\sigma_{\text{inst}}}{g(\nu_0) T_{\text{CMB}}} \right]^2 e^{\ell^2 \theta_{\text{FWHM}}^2 / (8 \ln 2)}, \quad (46)$$

where $\theta_{\text{FWHM}} = 1.4 \text{ arcmin}$ is the full width at half maximum of the beam size [63]. In addition to the instrumental noise, radio and infrared point sources, cosmic infrared background, and primary CMB leak to Compton y estimates due to incomplete separation and can be a source of noise. However, the instrumental noise dominates at small scales [19,26], and we ignore the contributions from other sources in the subsequent analyses for simplicity.

Without cluster masking



With cluster masking

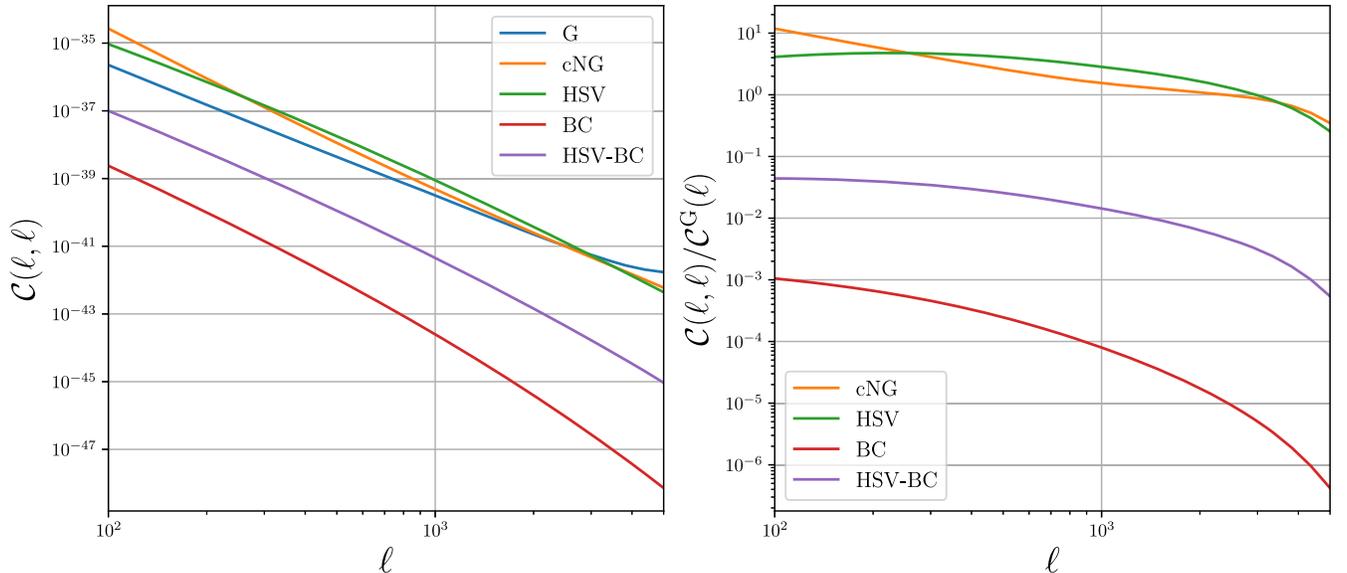


FIG. 5. Diagonal terms of the covariance matrix of the tSZ power spectrum (left panels) and each term divided by the Gaussian term (right panels) without (upper panels) and with (lower panels) cluster masking. The HSV, BC, and HSV-BC terms correspond to the SSC.

The total covariance matrix is shown in Fig. 4. The diagonal terms of the covariance matrix for each component are shown in Fig. 5. In the case of no cluster masking, the cNG term dominates at all scales. On the other hand, cluster masking removes clusters that contribute to the cNG term, and Gaussian and SSC terms become important when clusters are masked.

V. RESULTS

A. Statistical significance

In this section we discuss the statistical significance of the detection of the tSZ power spectrum and the effects of cluster masking and SSC on it. The signal-to-noise ratio (SNR) S/N is computed as

$$\frac{S}{N} = \left[\sum_{\ell_{\min} \leq \ell_i, \ell_j \leq \ell_{\max}} C(\ell_i) C_{ij}^{-1} C(\ell_j) \right]^{\frac{1}{2}}. \quad (47)$$

Figure 6 shows the SNR as a function of the maximum multipole ℓ_{\max} . The SSC contributes only 5% of the SNR without cluster masking because the cNG term dominates. However, by masking massive clusters, the overall SNR is enhanced due to the suppression of the cNG term but the relative contribution from SSC also increases to 30% because the SSC is less sensitive to the cluster masking.

Next, we discuss how SSC affects the chi-squared statistic that an observer could obtain for a given realization of data. The chi-squared statistic is defined as

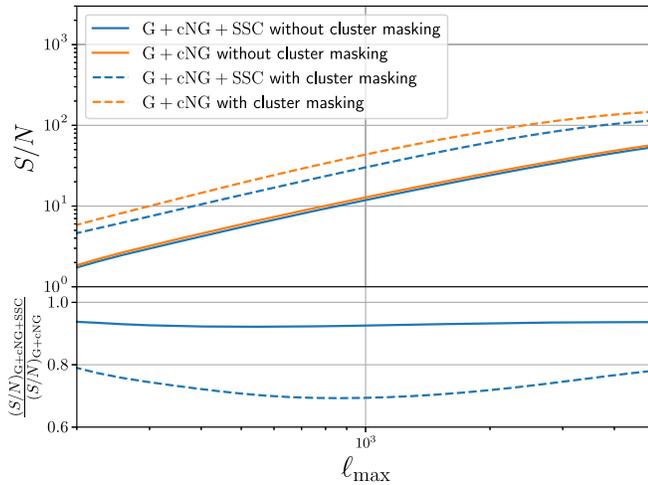


FIG. 6. SNR as a function of maximum multipole ℓ_{\max} with the total covariance including SSC or the covariance including only the Gaussian and trispectrum (cNG) contributions (i.e., without SSC), as denoted in the legend. The lowest multipole is fixed as $\ell_{\min} = 100$ and the highest multipole is varied in the range $200 \leq \ell_{\max} \leq 5000$. The bin width is also fixed. The lower panel shows the ratio of S/N values with and without the SSC contribution, for the two cases where we remove massive clusters with $Y_{500} > 10^{-5} \text{ Mpc}^{-2}$ or not (with and without “cluster masking”).

$$\chi^2 \equiv (\mathbf{d} - \bar{\mathbf{d}})^T \mathbf{C}^{-1} (\mathbf{d} - \bar{\mathbf{d}}), \quad (48)$$

where \mathbf{d} is the data vector defined as

$$\mathbf{d} = (C(\ell_1), \dots, C(\ell_{n_\ell})), \quad (49)$$

and $\bar{\mathbf{d}} \equiv \langle \mathbf{d} \rangle$. The expectation value of the χ^2 value is estimated as

$$\begin{aligned} \langle \chi^2 \rangle &= \langle (\mathbf{d} - \bar{\mathbf{d}})^T \mathbf{C}^{-1} (\mathbf{d} - \bar{\mathbf{d}}) \rangle \\ &= \text{Tr}(\langle (\mathbf{d} - \bar{\mathbf{d}})(\mathbf{d} - \bar{\mathbf{d}})^T \rangle \mathbf{C}^{-1}) \\ &= \text{Tr}(\mathbf{C} \mathbf{C}^{-1}) = n_\ell. \end{aligned} \quad (50)$$

However, if the wrong covariance is employed in the analysis, the expected chi-squared deviates from the number of bins. Let us consider the case where the true covariance is composed of Gaussian, connected non-Gaussian, and super sample covariance, and the wrong covariance is that without the super sample covariance:

$$\mathbf{C}_{\text{true}} \equiv \mathbf{C}^G + \mathbf{C}^{\text{cNG}} + \mathbf{C}^{\text{SSC}}, \quad (51)$$

$$\mathbf{C}_{\text{wrong}} \equiv \mathbf{C}^G + \mathbf{C}^{\text{cNG}}. \quad (52)$$

Then, the expectation value of chi-squared with the wrong covariance is given as

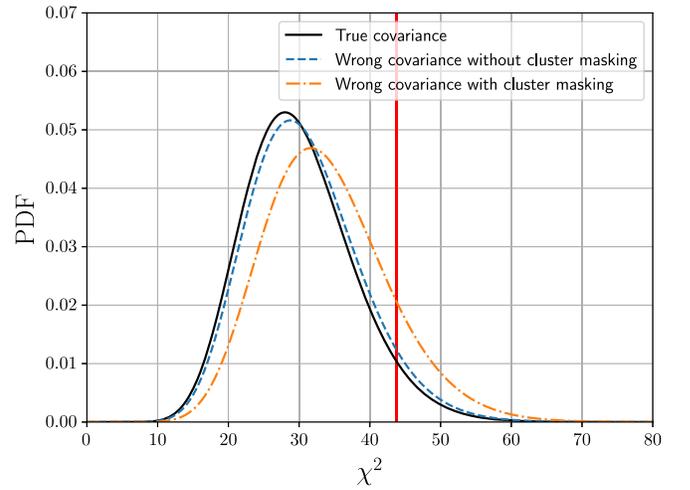


FIG. 7. PDFs of chi-squared with true and wrong covariances and with and without cluster masking. The black solid line shows the chi-squared distribution with degrees of freedom $n = 30$, and for the other lines the argument is scaled to match the expectation values. The blue dashed line shows the PDF with the wrong covariance, i.e., without the SSC term, and furthermore the orange dot-dashed line shows the one when massive clusters are masked. The red vertical line corresponds to $\chi^2 = 43.8$ where the p -value is 0.05 with the true covariance.

TABLE I. p -values with wrong covariances with and without cluster masking. The lower limit of chi-squared is determined as the p -value with the true covariance is 0.05.

Covariance	$\langle \chi^2 \rangle$	p -value
True covariance	30	0.05
Wrong covariance without cluster masking	30.80	0.061
Wrong covariance with cluster masking	33.91	0.120

$$\begin{aligned} \langle \chi_{\text{wrong}}^2 \rangle &= \langle (\mathbf{d} - \bar{\mathbf{d}})^T (\mathbf{C}_{\text{wrong}})^{-1} (\mathbf{d} - \bar{\mathbf{d}}) \rangle \\ &= \text{Tr}(\mathbf{C}_{\text{true}} (\mathbf{C}_{\text{wrong}})^{-1}) > n_\ell. \end{aligned} \quad (53)$$

The chi-squared with the wrong covariance no longer follows the chi-squared distribution, but if the wrong covariance is close to the true covariance we can approximate the probability distribution function (PDF) of χ_{wrong}^2 as the chi-squared distribution with scaling to match the expectation value:

$$P(\chi_{\text{wrong}}^2) d\chi_{\text{wrong}}^2 \approx P_{\chi^2} \left(\frac{n_\ell}{\alpha} \chi_{\text{wrong}}^2; n_\ell \right) \frac{n_\ell}{\alpha} d\chi_{\text{wrong}}^2, \quad (54)$$

where $P_{\chi^2}(\chi^2; n)$ is the PDF of the chi-squared distribution with degrees of freedom n , and $\alpha \equiv \text{Tr}(\mathbf{C}_{\text{true}} (\mathbf{C}_{\text{wrong}})^{-1})$.

Figure 7 shows the PDFs with the true covariance and the wrong ones, i.e., without SSC, for the cases with and without cluster masking. The PDFs with wrong covariances are skewed rightward compared to the true one because the wrong covariance underestimates the true covariance or, more exactly, the amplitude of statistical scatters, and it apparently gives a higher significance. In order to see how the significance is overestimated with the wrong covariance, we address how the p -value changes. First, we compute $\chi_{p=0.05}^2$, which gives the p -value as 0.05 with the true covariance. When the true covariance is adopted, the resultant chi-squared follows the chi-squared distribution, and from the cumulative distribution function, $\chi_{p=0.05}^2 = 43.8$ is calculated. Then, we compute the p -values with wrong covariances, i.e.,

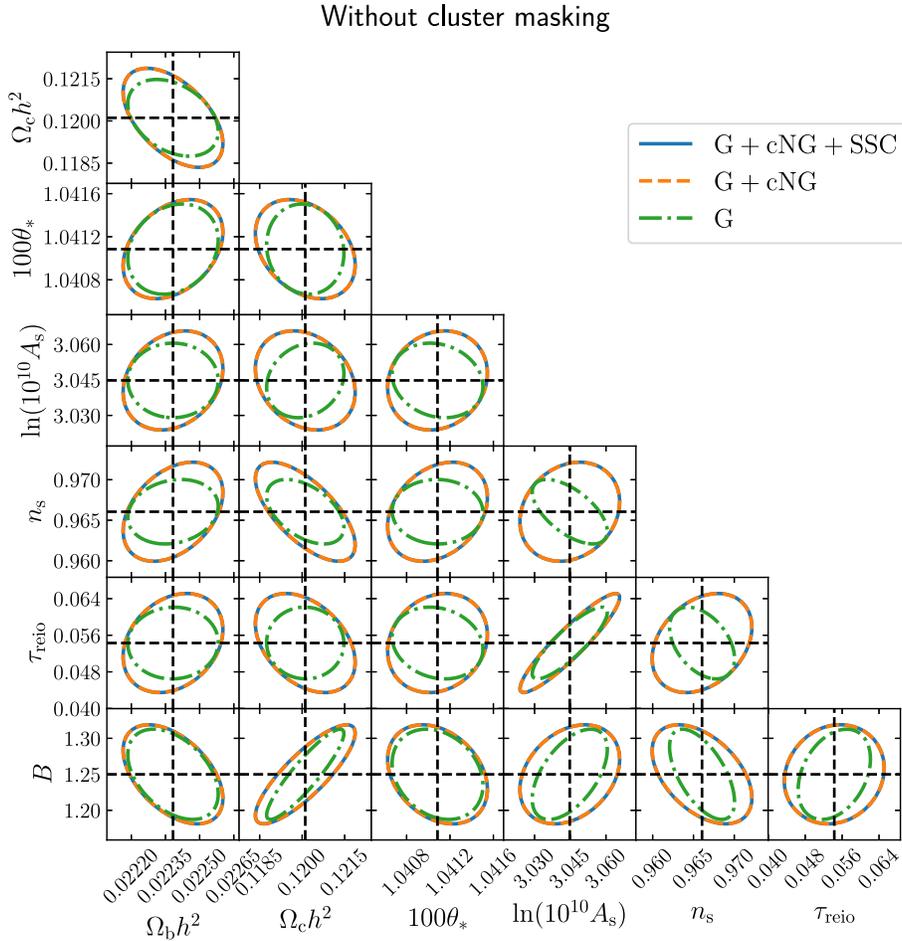


FIG. 8. Expected 1σ constraints on parameters including marginalization over other parameters, obtained based on the Fisher matrix (see text for details). The cluster masking is not applied. The blue solid, orange dashed, and green dot-dashed lines correspond to the results with total, Gaussian and connected non-Gaussian, Gaussian only covariances, respectively.

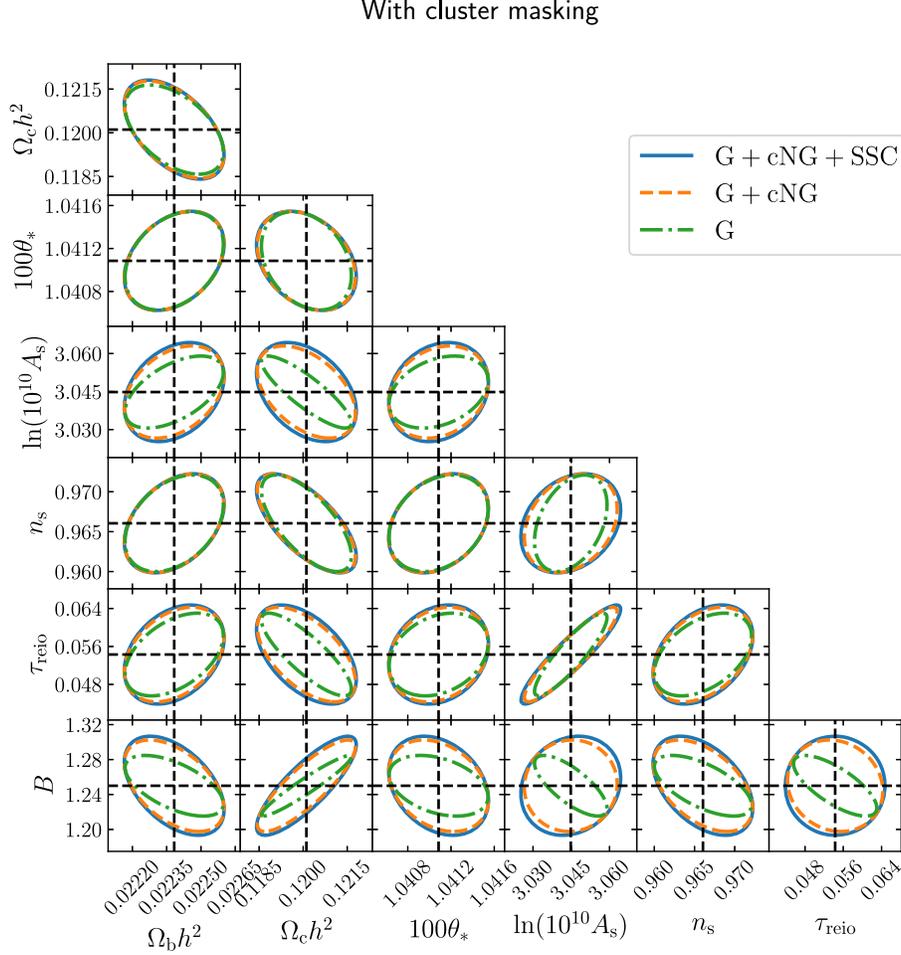


FIG. 9. Similar to the previous figure, but for the expected constraints when the cluster masking is applied. The blue solid, orange dashed, and green dot-dashed lines correspond to the results with total, Gaussian and connected non-Gaussian, Gaussian only covariances, respectively.

$$p = \int_{\chi^2_{p=0.05}}^{\infty} P(\chi^2_{\text{wrong}}) d\chi^2_{\text{wrong}}. \quad (55)$$

In Table I we show the p -values for the two cases with and without cluster masking. When cluster masking is not applied, the effect of the super sample covariance is subdominant and the p -value increases only by 1.1%. However, when massive clusters are masked, the SSC becomes relatively important and the p -value is 12.0%. This result implies that it is 7.0% more likely to derive optimistic significance of the detection. For the case of parameter inference, however, even a true model gives an apparent bad fit to the data because there might be a higher chance to have a relatively large chi-squared value for the best-fit model due to the underestimation in the covariance amplitude.

B. Fisher analysis

In this section we quantify the effects of SSC on the parameter constraints with the tSZ power spectrum based

on a Fisher forecast [65,66]. The tSZ power spectrum is sensitive to the matter fluctuation and has the potential to place tight constraints on the amplitude A_s . Furthermore, the hydrostatic mass bias parameter can also be constrained, which information is not accessible solely from CMB temperature and polarization analysis.

When the likelihood is assumed to be a multivariate Gaussian, the Fisher matrix is given as

$$F_{ij}^{\text{tSZ}} = \frac{\partial \mathbf{d}^T}{\partial \theta_i} \mathbf{C}^{-1} \frac{\partial \mathbf{d}}{\partial \theta_j} + \frac{1}{2} \text{Tr} \left(\mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_i} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_j} \right). \quad (56)$$

We consider the parameter space $\boldsymbol{\theta} = (\Omega_b h^2, \Omega_c h^2, 100\theta_*, \ln(10^{10} A_s), n_s, \tau_{\text{reio}}, B)$, where $B = (1 - b_{\text{HSE}})^{-1}$. In particular, the hydrostatic bias parameter B characterizes the nonthermal pressure support of galaxy clusters and groups, and it is not well constrained compared to cosmological parameters. Since severe degeneracy between parameters only occurs with the tSZ power spectrum, we add information from the *Planck* 2018 TT, TE,

TABLE II. Expected 1σ error on each parameter with CMB anisotropies and the tSZ power spectrum. We show the values with three different covariances: G, G + cNG, and G + cNG + SSC.

	$\Omega_b h^2$	$\Omega_c h^2$	$100\theta_*$	$\ln(10^{10}A_s)$	n_s	τ_{reio}	B
CMB	0.000146	0.00118	0.000307	0.0141	0.00415	0.00730	—
Without cluster masking							
CMB + tSZ (G)	0.000132	0.00090	0.000278	0.0105	0.00264	0.00520	0.0416
CMB + tSZ (G + cNG)	0.000145	0.00117	0.000305	0.0139	0.00404	0.00716	0.0455
CMB + tSZ (G + cNG + SSC)	0.000145	0.00117	0.000305	0.0139	0.00405	0.00717	0.0456
With cluster masking							
CMB + tSZ (G)	0.000142	0.00102	0.000304	0.0094	0.00401	0.00577	0.0230
CMB + tSZ (G + cNG)	0.000144	0.00111	0.000305	0.0121	0.00409	0.00661	0.0347
CMB + tSZ (G + cNG + SSC)	0.000144	0.00113	0.000306	0.0130	0.00411	0.00691	0.0376

EE + lowE + lensing result [39]. Thus, the resultant Fisher matrix F is given by

$$F = F^{\text{tSZ}} + F^{\text{CMB}}, \quad (57)$$

where the Fisher matrix for CMB anisotropies is computed from chains obtained in the *Planck* analysis. In Figs. 8 and 9 we show forecasts of the constraints from the tSZ power spectrum and CMB anisotropies with and without cluster masking. Table II shows the forecast of 1σ errors. When adding the information about the tSZ power spectrum we can obtain tighter constraints on parameters, especially the amplitude of primordial curvature perturbations: $\ln(10^{10}A_s)$. Note that the hydrostatic bias parameter B can also be tightly constrained. When clusters are not masked, the effect of SSC is subdominant since the cNG term is much larger. On the other hand, with cluster masking the SSC comes into effect because the SSC is not sensitive to cluster masking compared to the cNG term. As the SNR improves from cluster masking, the constraints become tighter.

VI. CONCLUSIONS

The tSZ effect is one of the most important probes in cosmology. With the power spectrum, we can constrain the cosmological parameters and investigate the astrophysical effects, e.g., hydrostatic mass bias. Since any survey is done for a finite volume, even for a full-sky survey, it is important to realize the impact of the super survey modes on the statistical power of the tSZ power spectrum. We quantified the contribution based on the halo model approach and addressed biases in parameter estimation. However, it was found that the super survey covariance is subdominant compared with the cNG term, which is sourced from the trispectrum with a parallelogram configuration.

In order to enhance the statistical significance, we proposed that massive nearby galaxy clusters, which are the dominant source of the trispectrum, should be masked to suppress the contribution of the cNG term [37].

We proposed the cluster masking based on the integrated Compton y parameter, which corresponds to the thermal energy content in galaxy clusters, and the method enables one to significantly reduce the cNG term. On the other hand, even after masking clusters, the SSC remains because it originates from relatively less massive clusters, similarly to the power spectrum signal. Though the overall statistical significance was improved, the SSC has an appreciable effect and weakens the constraints on the parameters.

We have carried out Fisher forecasts on cosmological parameters and the hydrostatic mass bias parameter through the tSZ power spectrum measurement combined with the results of *Planck* CMB anisotropy measurements. In particular, the hydrostatic mass bias parameter, which cannot be constrained with only primary CMB anisotropies, can be tightly constrained through the tSZ power spectrum. By masking clusters, the constraints on the parameters related to the primordial power spectrum, i.e., A_s and n_s , and the hydrostatic mass bias parameter improve. On the other hand, the SSC becomes important because it persists after masking. For example, the constraint on the hydrostatic mass bias with covariance including SSC is 10% larger than that with covariance excluding SSC. Since the effect of SSC remains at all scales, it is critical to incorporate the SSC for accurate estimates on parameters for both ongoing and upcoming observational surveys of the tSZ effect. Moreover, the cluster masking loses the information from massive galaxy clusters. In order to compensate for the loss, a joint analysis with cluster counts is thought to be a practical solution [9,67,68]. The method we developed in this paper increases the potential of tSZ cosmology. Massive clusters with significant tSZ signals on an individual cluster basis can be used for cosmology, e.g., via the number counts of the clusters. On the other hand, when those clusters are masked from the power spectrum measurement, the tSZ power spectrum can be used to estimate the cosmological parameters. Thus, we showed that massive clusters and other tSZ clusters can play complementary roles in cosmology. This direction is worth exploring further in more detail.

ACKNOWLEDGMENTS

K. O. is supported by JSPS Overseas Research Fellowships. This work is also supported in part by the World Premier International Research Center Initiative

(WPI Initiative), MEXT, Japan, Reischauer Institute of Japanese Studies at Harvard University, and JSPS KAKENHI Grants No. JP15H05887, No. JP15H05893, No. JP15H05896, No. JP15K21733, and No. JP19H00677.

-
- [1] R. A. Sunyaev and Y. B. Zel'dovich, *Astrophys. Space Sci.* **7**, 3 (1970).
- [2] R. A. Sunyaev and Y. B. Zel'dovich, *Comments Astrophys. Space Phys.* **4**, 173 (1972), http://articles.adsabs.harvard.edu/cgi-bin/get_file?pdfs/ComAp/0004/1972ComAp...4.173S.pdf.
- [3] R. A. Sunyaev and Y. B. Zel'dovich, *Mon. Not. R. Astron. Soc.* **190**, 413 (1980).
- [4] J. E. Carlstrom, G. P. Holder, and E. D. Reese, *Annu. Rev. Astron. Astrophys.* **40**, 643 (2002).
- [5] T. Kitayama, *Prog. Theor. Exp. Phys.* **2014**, 06B111 (2014).
- [6] Z. Haiman, J. J. Mohr, and G. P. Holder, *Astrophys. J.* **553**, 545 (2001).
- [7] E. Komatsu and U. Seljak, *Mon. Not. R. Astron. Soc.* **336**, 1256 (2002).
- [8] M. Takada and S. Bridle, *New J. Phys.* **9**, 446 (2007).
- [9] M. Oguri and M. Takada, *Phys. Rev. D* **83**, 023008 (2011).
- [10] L. Salvati, M. Douspis, and N. Aghanim, *Astron. Astrophys.* **614**, A13 (2018).
- [11] Planck Collaboration, *Astron. Astrophys.* **594**, A22 (2016).
- [12] Planck Collaboration, *Astron. Astrophys.* **594**, A24 (2016).
- [13] D. S. Swetz *et al.*, *Astrophys. J. Suppl. Ser.* **194**, 41 (2011).
- [14] J. Dunkley *et al.*, *Astrophys. J.* **739**, 52 (2011).
- [15] B. A. Benson *et al.*, *SPIE Int. Soc. Opt. Eng.* **9153**, 91531P (2014).
- [16] E. M. George *et al.*, *Astrophys. J.* **799**, 177 (2015).
- [17] Simons Observatory Collaboration, *J. Cosmol. Astropart. Phys.* **02** (2019) 056.
- [18] K. N. Abazajian *et al.*, [arXiv:1610.02743](https://arxiv.org/abs/1610.02743).
- [19] J. C. Hill and E. Pajer, *Phys. Rev. D* **88**, 063526 (2013).
- [20] E. Komatsu and T. Kitayama, *Astrophys. J. Lett.* **526**, L1 (1999).
- [21] A. Cooray and R. Sheth, *Phys. Rep.* **372**, 1 (2002).
- [22] L. D. Shaw, D. Nagai, S. Bhattacharya, and E. T. Lau, *Astrophys. J.* **725**, 1452 (2010).
- [23] H. Trac, P. Bode, and J. P. Ostriker, *Astrophys. J.* **727**, 94 (2011).
- [24] I. G. McCarthy, A. M. C. Le Brun, J. Schaye, and G. P. Holder, *Mon. Not. R. Astron. Soc.* **440**, 3645 (2014).
- [25] K. Dolag, E. Komatsu, and R. Sunyaev, *Mon. Not. R. Astron. Soc.* **463**, 1797 (2016).
- [26] B. Bolliet, B. Comis, E. Komatsu, and J. F. Macías-Pérez, *Mon. Not. R. Astron. Soc.* **477**, 4957 (2018).
- [27] W. Fang, K. Kadota, and M. Takada, *Phys. Rev. D* **85**, 023007 (2012).
- [28] L. Van Waerbeke, G. Hinshaw, and N. Murray, *Phys. Rev. D* **89**, 023508 (2014).
- [29] N. Battaglia, J. C. Hill, and N. Murray, *Astrophys. J.* **812**, 154 (2015).
- [30] A. Hojjati, T. Tröster, J. Harnois-Déraps, I. G. McCarthy, L. van Waerbeke, A. Choi, T. Erben, C. Heymans, H. Hildebrandt, G. Hinshaw, Y.-Z. Ma, L. Miller, M. Viola, and H. Tanimura, *Mon. Not. R. Astron. Soc.* **471**, 1565 (2017).
- [31] K. Osato, S. Flender, D. Nagai, M. Shirasaki, and N. Yoshida, *Mon. Not. R. Astron. Soc.* **475**, 532 (2018).
- [32] R. Makiya, S. Ando, and E. Komatsu, *Mon. Not. R. Astron. Soc.* **480**, 3928 (2018).
- [33] K. Osato, M. Shirasaki, H. Miyatake, D. Nagai, N. Yoshida, M. Oguri, and R. Takahashi, *Mon. Not. R. Astron. Soc.* **492**, 4780 (2020).
- [34] R. Makiya, C. Hikage, and E. Komatsu, *Publ. Astron. Soc. Jpn.* **72**, 26 (2020).
- [35] M. Takada and W. Hu, *Phys. Rev. D* **87**, 123504 (2013).
- [36] B. Horowitz and U. Seljak, *Mon. Not. R. Astron. Soc.* **469**, 394 (2017).
- [37] L. D. Shaw, O. Zahn, G. P. Holder, and O. Doré, *Astrophys. J.* **702**, 368 (2009).
- [38] Planck Collaboration, *Astron. Astrophys.* **641**, A1 (2020).
- [39] Planck Collaboration, *Astron. Astrophys.* **641**, A6 (2020).
- [40] The comoving distance also serve as the indicator of the cosmic time.
- [41] D. Blas, J. Lesgourgues, and T. Tram, *J. Cosmol. Astropart. Phys.* **07** (2011) 034.
- [42] G. L. Bryan and M. L. Norman, *Astrophys. J.* **495**, 80 (1998).
- [43] J. F. Navarro, C. S. Frenk, and S. D. M. White, *Astrophys. J.* **462**, 563 (1996).
- [44] J. F. Navarro, C. S. Frenk, and S. D. M. White, *Astrophys. J.* **490**, 493 (1997).
- [45] A. Klypin, G. Yepes, S. Gottlöber, F. Prada, and S. Heß, *Mon. Not. R. Astron. Soc.* **457**, 4340 (2016).
- [46] S. Bocquet, A. Saro, K. Dolag, and J. J. Mohr, *Mon. Not. R. Astron. Soc.* **456**, 2361 (2016).
- [47] J. L. Tinker, B. E. Robertson, A. V. Kravtsov, A. Klypin, M. S. Warren, G. Yepes, and S. Gottlöber, *Astrophys. J.* **724**, 878 (2010).
- [48] D. Nagai, A. V. Kravtsov, and A. Vikhlinin, *Astrophys. J.* **668**, 1 (2007).
- [49] Planck Collaboration, *Astron. Astrophys.* **550**, A131 (2013).
- [50] D. Suto, H. Kawahara, T. Kitayama, S. Sasaki, Y. Suto, and R. Cen, *Astrophys. J.* **767**, 79 (2013).
- [51] K. Nelson, E. T. Lau, and D. Nagai, *Astrophys. J.* **792**, 25 (2014).
- [52] X. Shi, E. Komatsu, K. Nelson, and D. Nagai, *Mon. Not. R. Astron. Soc.* **448**, 1020 (2015).

- [53] F. Vazza, M. Angelinelli, T. W. Jones, D. Eckert, M. Brügger, G. Brunetti, and C. Gheller, *Mon. Not. R. Astron. Soc.* **481**, L120 (2018).
- [54] E. Medezinski, N. Battaglia, K. Umetsu, M. Oguri, H. Miyatake, A. J. Nishizawa, C. Sifón, D. N. Spergel, I. N. Chiu, Y.-T. Lin, N. Bahcall, and Y. Komiyama, *Publ. Astron. Soc. Jpn.* **70**, S28 (2018).
- [55] H. Miyatake *et al.*, *Astrophys. J.* **875**, 63 (2019).
- [56] E. T. Lau, D. Nagai, and K. Nelson, *Astrophys. J.* **777**, 151 (2013).
- [57] In some literature, another definition $Y'_{500} = Y_{500}/D_A^2$ is employed. In this definition, the unit is normally arcmin² and the quantity is computed by integrating observed SZ flux in angular space. For detailed discussions, see Sec. 3.1 in Ref. [49].
- [58] R. E. Angulo, V. Springel, S. D. M. White, A. Jenkins, C. M. Baugh, and C. S. Frenk, *Mon. Not. R. Astron. Soc.* **426**, 2046 (2012).
- [59] A. Cooray, *Phys. Rev. D* **64**, 063514 (2001).
- [60] M. Takada and B. Jain, *Mon. Not. R. Astron. Soc.* **395**, 2065 (2009).
- [61] M. Sato, T. Hamana, R. Takahashi, M. Takada, N. Yoshida, T. Matsubara, and N. Sugiyama, *Astrophys. J.* **701**, 945 (2009).
- [62] A. J. S. Hamilton, C. D. Rimes, and R. Scoccimarro, *Mon. Not. R. Astron. Soc.* **371**, 1188 (2006).
- [63] M. S. Madhavacheril *et al.*, *Phys. Rev. D* **102**, 023534 (2020).
- [64] R. Takahashi, S. Soma, M. Takada, and I. Kayo, *Mon. Not. R. Astron. Soc.* **444**, 3473 (2014).
- [65] M. Tegmark, A. N. Taylor, and A. F. Heavens, *Astrophys. J.* **480**, 22 (1997).
- [66] M. Tegmark, *Phys. Rev. Lett.* **79**, 3806 (1997).
- [67] M. Takada and D. N. Spergel, *Mon. Not. R. Astron. Soc.* **441**, 2456 (2014).
- [68] E. Schaan, M. Takada, and D. N. Spergel, *Phys. Rev. D* **90**, 123523 (2014).