PHYSICAL REVIEW D **87,** 123504 (2013)

Power spectrum super-sample covariance

Masahiro Takada
Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), The University of Tokyo, Chiba 277-8583, Japan

Wayne Hu
Kavli Institute for Cosmological Physics, Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA
and Department of Astronomy and Astrophysics, University of Chicago, Chicago, Illinois 60637, USA
(Received 3 March 2013; published 5 June 2013)

We provide a simple, unified approach to describing the impact of super-sample covariance, or beat coupling, on power spectrum estimation in a finite-volume survey. For a wide range of survey volumes, the sample variance that arises from modes that are larger than the survey dominates the covariance of power spectrum estimators for modes much smaller than the survey. The deeply nonlinear version of this effect is known as halo sample variance. We show that all variants are unified by the matter trispectrum of squeezed configurations and that such configurations obey a consistency relation which relates them to the response of the power spectrum to a change in the background density. Our method also applies to statistics that are based on radial projections of the density field such as weak lensing shear. While we use the halo model for an analytic description to expose the nature of the effect, the consistency description enables an accurate calibration of the full effect directly from simulations. It also suggests that super-sample covariance may be viewed as an additional interesting signal rather than excess noise.

DOI: 10.1103/PhysRevD.87.123504 PACS numbers: 98.80.−k, 95.36.+x, 98.65.Dx

I. INTRODUCTION

Cosmic acceleration is perhaps the most tantalizing problem in modern cosmology. To unlock its mysteries, a number of ambitious wide-field optical and infrared galaxy surveys have been proposed. These range from ground-based imaging and spectroscopic surveys such as the Subaru Hyper Suprime-Cam Survey [1], the Dark Energy Survey [2], the Kilo-Degrees Survey [3], the Large Synoptic Survey Telescope [4], the Baryon Oscillation Spectrograph Survey [5], the Extended BOSS [6], the BigBOSS [7], and the Subaru Prime Focus Spectrograph Survey [8] (see also [9]) to space-based optical and near-infrared missions such as the Euclid project [10] and the WFIRST project [11]. Each of these surveys approaches the nature of cosmic acceleration using multiple large-scale structure probes: e.g. weak gravitational lensing, baryon acoustic oscillations, and clustering statistics of large-scale structure tracers such as galaxies and clusters, including redshift-space distortion effects (see Ref. [12] for a recent review).

To attain the full potential of such surveys, it is important to understand the statistical properties of large-scale structure probes and the matter density field that underlies them. Its two-point correlation function or the Fourier-transformed counterpart, the power spectrum, is the simplest and most commonly used statistical quantity to extract cosmological information from the large-scale structure probes. The statistical precision of power spectrum measurements is determined by their covariance matrix that itself contains two contributions, the measurement noise and sample variances caused by an incomplete sampling of the fluctuations due to a finite-volume survey.

Even though the initial density field is nearly Gaussian, the sample variance of large-scale structure probes gets substantial non-Gaussian contributions from the nonlinear evolution of large-scale structure [13–15]. Most of the useful cosmological information from such probes lies in the weakly or deeply nonlinear regime in the matter distribution. In the nonlinear regime, the different Fourier modes are no longer independent but rather are correlated with each other. Hence, off-diagonal entries in the power spectrum covariance matrix appear which are described by the connected 4-point correlation function of the matter distribution, the matter trispectrum. To realize the statistical power carried by power spectrum, it is important to accurately model the matter trispectrum in the nonlinear regime. The standard methods to study the power spectrum sample variance are based on simulations of the Λ-dominated cold dark matter (ΛCDM) model [13,14,16–24] which can be used to test and calibrate semianalytic descriptions such as the halo model approach that can encompass a larger range of model parameters [15,22,25–27].

One further complication arises. Nonlinear evolution couples short-wavelength modes relevant for the power spectrum measurement with the very long-wavelength modes outside a survey volume, the so-called super-survey modes. The super-survey modes are tricky to consider, because their effect vanishes for power spectrum measurements of simulations with periodic boundary conditions that have no contribution of modes outside the simulation box. Super-sample covariance on the power spectrum was originally pointed out in Ref. [16] and called beat coupling (BC) (see also [28], which studied the super-sample

variance effect on the number counts of halos), and then was followed by many studies [17,19,20,22,25,27,29]. Later the effect of super-survey modes for power spectra in the deeply nonlinear regime was studied using halo bias theory [30] in the halo model approach and called halo sample variance [19,22,25] (see also [28]). These studies have shown that super-sample variance dominates the non-Gaussian errors in the weakly or deeply nonlinear regime. Super-sample variance contributes to a saturation in the information content carried by the power spectrum amplitude [16,19,21,22,31] (see also [32], for tests on real data). The saturated information content has triggered a further discussion on how to recover the information content that was originally contained in the initial Gaussian field, by using the nonlinear transformation method (e.g. [33,34]) or including the higher-order moment information [22,35].

However, it is still unclear how all of these super-sample variance effects are described by the matter trispectrum and how they can best be quantified and treated in parameter estimation. In this paper, we provide a systematic study of common origin of all such effects, unify their description, and show how they can be directly quantified.

The structure of this paper is as follows. In Sec. II, we develop a simple, unified approach to describe the super-sample covariance of power spectrum estimation in a finite-volume survey and show its relation to the response of the power spectrum to a change in the background density. In Sec. III, we use the halo model to compute the power spectrum covariance for a ΛCDM model and verify that we can recover the beat coupling and halo sample variance results from our description. Section IV is devoted to discussion. We provide an appendix that extends the formalism to projected density field statistics such as weak lensing shear.

II. POWER SPECTRUM COVARIANCE IN A SURVEY WINDOW

In Sec. II A we review the construction of power spectrum estimators in a finite-volume survey and discuss their relation to the underlying true power spectrum. We relate the covariance of these estimators to the matter trispectrum in Sec. II B and identify where the additional sample variance from modes larger than the survey arises. In Sec. II C, we present a trispectrum consistency relation that straightforwardly relates this effect to the response of the power spectrum to a change in the background density and outline how it can be easily calibrated in simulations. As such, it may also be considered as signal rather than noise, carried by a single additional parameter, the variable background density.

A. Power spectrum estimator

Assume we measure the underlying density fluctuation field $\delta(\mathbf{x})$ through a survey window function $W(\mathbf{x})$ which is 1 in the measured region and 0 in the unmeasured region. Note that by construction we implicitly assume that the observable is the density fluctuation itself or equivalently the true mean density is known. This is appropriate for statistics related to weak lensing that probes the differential gravitational field, i.e. the tidal field. If the density fluctuation field is measured relative to the mean of the survey region there is a correction factor that we will describe in Sec. II C (see also Ref. [29]).

Regardless of the complexity of the survey geometry, we can always define the observed field as

$$\delta_W(\mathbf{x}) = \delta(\mathbf{x})W(\mathbf{x}), \tag{1}$$

whose Fourier transform is a convolution

$$\tilde{\delta}_W(\mathbf{k}) = \int \frac{d^3\mathbf{q}}{(2\pi)^3}\tilde{W}(\mathbf{q})\tilde{\delta}(\mathbf{k}-\mathbf{q}), \tag{2}$$

where $\tilde{W}(\mathbf{k})$ is the Fourier transform of the survey window function. We have here employed the continuous limit of discrete Fourier transforms under the approximation that the total volume for the Fourier transform is much greater than the survey region (see Refs. [22,25] for a pedagogical derivation of power spectrum estimator and the covariance based on the discrete Fourier decomposition).

Let us next define an estimator of the power spectrum as

$$\hat{P}(k_i) \equiv \frac{1}{V_W}\int_{|\mathbf{k}|\in k_i}\frac{d^3\mathbf{k}}{V_{k_i}}\tilde{\delta}_W(\mathbf{k})\tilde{\delta}_W(-\mathbf{k}), \tag{3}$$

where the integral is over a shell in k space of width Δk and volume $V_{k_i} \simeq 4\pi k_i^2\Delta k$ for $\Delta k/k_i \ll 1$. Here the effective survey volume is defined as

$$V_W \equiv \int d^3\mathbf{x}W(\mathbf{x}). \tag{4}$$

Given the definition of the power spectrum

$$\langle\tilde{\delta}(\mathbf{k})\tilde{\delta}(\mathbf{k}')\rangle = (2\pi)^3\delta_D^3(\mathbf{k}+\mathbf{k}')P(k), \tag{5}$$

the ensemble average of its estimator is

$$\langle\hat{P}(k_i)\rangle = \frac{1}{V_W}\int_{|\mathbf{k}|\in k_i}\frac{d^3\mathbf{k}}{V_{k_i}}\int\frac{d^3\mathbf{q}}{(2\pi)^3}|\tilde{W}(\mathbf{q})|^2P(\mathbf{k}-\mathbf{q}). \tag{6}$$

Thus, the observed power spectrum is given as a convolution of the underlying power spectrum with the window function combining the density modes separated by less than the Fourier width of the window. In the general case, one would deconvolve the window in constructing an unbiased estimator [36,37].

In this paper, we are interested in the effect of the global survey geometry on the power spectrum covariance, not in the effect of masked regions at small spatial scales. In this case, the window function has a width of $\sim 1/L$ in Fourier space, where $L \sim V_W^{1/3}$ or more generally its smallest dimension. When we focus on wave number modes satisfying $k \gg 1/L$, we find that the power spectrum estimator of Eq. (3) is unbiased:

$$\langle \hat{P}(k_i)\rangle \simeq \frac{1}{V_W}\int_{|\mathbf{k}|\in k_i}\frac{d^3\mathbf{k}}{V_{k_i}}P(k)\int\frac{d^3\mathbf{q}'}{(2\pi)^3}|\tilde{W}(\mathbf{q}')|^2$$
$$\simeq P(k_i)\frac{1}{V_W}\int\frac{d^3\mathbf{q}'}{(2\pi)^3}|\tilde{W}(\mathbf{q}')|^2 = P(k_i). \quad (7)$$

Here we have used that $P(|\mathbf{k}-\mathbf{q}|)\simeq P(k)$ over the integration range of $d^3\mathbf{q}$ which the window function supports and also assumed that $P(k)$ is not a rapidly varying function within the k-bin. In the third equality on the right-hand side, we have used the general identity for the window function:

$$V_W = \int d^3\mathbf{x}W^n(\mathbf{x})$$
$$= \int\left[\prod_{a=1}^{n}\frac{d^3\mathbf{q}_a}{(2\pi)^3}\tilde{W}(\mathbf{q}_a)\right](2\pi)^3\delta_D^3(\mathbf{q}_{1\ldots n}), \quad (8)$$

where $\mathbf{q}_{1\ldots n}=\mathbf{q}_1+\ldots\mathbf{q}_n$ here and below. For $n=2$, $V_W=\int|\tilde{W}(\mathbf{q})|^2d^3\mathbf{q}/(2\pi)^3$.

B. Power spectrum covariance

Now consider the power spectrum covariance, which can be defined in terms of the power spectrum estimator as

$$C_{ij}\equiv \mathrm{Cov}[P(k_i),P(k_j)] = \langle\hat{P}(k_i)\hat{P}(k_j)\rangle - \langle\hat{P}(k_i)\rangle\langle\hat{P}(k_j)\rangle. \quad (9)$$

Here we consider sample covariance only. A real measurement will have measurement noise covariance, but we do not consider its effect in this paper. In the same $k\gg 1/L$ limit, the covariance becomes

$$C_{ij}\simeq\frac{1}{V_W}\left[\frac{(2\pi)^3}{V_{k_i}}2P(k_i)^2\delta^K_{ij}+\bar{T}^W(k_i,k_j)\right], \quad (10)$$

where δ^K_{ij} is the Kronecker delta function; $\delta^K_{ij}=1$ if $k_i=k_j$ to within the bin width, otherwise $\delta^K_{ij}=0$. The second term, proportional to $\bar{T}^W(k_i,k_j)$, is the non-Gaussian contribution arising from the connected 4-point function or trispectrum (see also [16], for the similar derivation),

$$\langle\tilde{\delta}(\mathbf{k}_1)\tilde{\delta}(\mathbf{k}_2)\tilde{\delta}(\mathbf{k}_3)\tilde{\delta}(\mathbf{k}_4)\rangle_c$$
$$= (2\pi)^3\delta_D^3(\mathbf{k}_{1234})T(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4), \quad (11)$$

convolved with the survey window function:

$$\bar{T}^W(k_i,k_j) = \frac{1}{V_W}\int_{|\mathbf{k}|\in k_i}\frac{d^3\mathbf{k}}{V_{k_i}}\int_{|\mathbf{k}|'\in k_j}\frac{d^3\mathbf{k}'}{V_{k_j}}$$
$$\times\int\left[\prod_{a=1}^{4}\frac{d^3\mathbf{q}_a}{(2\pi)^3}\tilde{W}(\mathbf{q}_a)\right](2\pi)^3\delta_D^3(\mathbf{q}_{1234})$$
$$\times T(\mathbf{k}+\mathbf{q}_1,-\mathbf{k}+\mathbf{q}_2,\mathbf{k}'+\mathbf{q}_3,-\mathbf{k}'+\mathbf{q}_4). \quad (12)$$

The convolution with the window function means that different 4-point configurations separated by less than the Fourier width of the window function and involving contributions from super-survey modes contribute in principle. In deriving Eq. (10), we have again used the general window function identity Eq. (8) with $n=4$ to eliminate one factor of $1/V_W$ in the power spectrum term under the same approximation as in Eq. (7). If the same slowly varying approximation were true of the trispectrum, it can be taken out of the integral and we would obtain the standard result [13]

$$C_{ij}\approx\frac{1}{V_W}\frac{(2\pi)^3}{V_{k_i}}2P^2(k_i)\delta^K_{ij}+C^{T0}_{ij}, \quad (13)$$

where

$$C^{T0}_{ij}=\frac{1}{V_W}\int_{|\mathbf{k}|\in k_i}\frac{d^3\mathbf{k}}{V_{k_i}}\int_{|\mathbf{k}'|\in k_j}\frac{d^3\mathbf{k}'}{V_{k_j}}T(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}'), \quad (14)$$

so that the whole covariance scales as $1/V_W$. Furthermore the first term can be understood from Gaussian statistics: the number of independent k modes in the shell is given as

$$N_{\mathrm{mode}}(k_i)=\frac{V_{k_i}V_W}{2(2\pi)^3}\simeq\frac{2\pi k_i^2\Delta k V_W}{(2\pi)^3}, \quad (15)$$

such that

$$C^{\mathrm{G}}_{ij}=\frac{1}{N_{\mathrm{mode}}(k_i)}P^2(k_i)\delta^K_{ij}. \quad (16)$$

The factor 2 in the denominator of N_{mode} arises from the reality condition of the density field, i.e. $\tilde{\delta}^*_W(\mathbf{k})=\tilde{\delta}_W(-\mathbf{k})$. The difference in bin width Δk scaling between the Gaussian and non-Gaussian terms means that the latter becomes more prominent in the variance when averaging over wide bins or equivalently when considering the covariance of narrow bins.

For the non-Gaussian term, there are additional effects if the trispectrum at $k\gg 1/L$ has structure on the scale of the survey $\Delta k\lesssim 1/L$. We next consider the general origin and description of such a term.

C. Super-sample covariance and trispectrum consistency

The trispectrum term that governs the additional effects are so-called squeezed quadrilaterals where two pairs of sides are nearly equal and opposite. By analogy with similar effects for primordial non-Gaussianity [38], these configurations should be determined by the response of the power spectrum to a rescaling of the background. We call this separate-universe ansatz the *trispectrum consistency relation*.

To see this in Eq. (12), we can make the change of variables $\mathbf{k}+\mathbf{q}_1\leftrightarrow\mathbf{k}$ and $\mathbf{q}_1+\mathbf{q}_2\leftrightarrow\mathbf{q}_{12}$ under the delta function condition $\mathbf{q}_{1234}=\mathbf{0}$ and the approximation that $q_{12}\ll k,k'$. The term of interest therefore is

$$\lim_{q_{12}\to 0} T(\mathbf{k}, -\mathbf{k}+\mathbf{q}_{12}, \mathbf{k}', -\mathbf{k}'-\mathbf{q}_{12}). \quad (17)$$

In this limit, the 4-point configuration describes the connection between $P(k)$ and $P(k')$ through a shared infinite wavelength mode $\mathbf{q}_{12}$. This mode acts like a background density or constant mode to the short wavelengths $\mathbf{k}$ and $\mathbf{k}'$. It follows therefore that the squeezed trispectrum can be characterized by the response of $P(k)$ to a fluctuation in the background density δ_b through

$$\bar{T}(\mathbf{k}, -\mathbf{k}+\mathbf{q}_{12}, \mathbf{k}', -\mathbf{k}'-\mathbf{q}_{12}) \approx T(\mathbf{k}, -\mathbf{k}, \mathbf{k}', -\mathbf{k}') + \frac{\partial P(k)}{\partial \delta_b}\frac{\partial P(k')}{\partial \delta_b} P^{\rm L}(q_{12}). \quad (18)$$

The overbar here refers to an angle average over the direction of $\mathbf{q}_{12}$ since any directional dependence cannot be quantified by a purely constant mode. Here $P^{\rm L}$ is the linear power spectrum and is designated as such to remind the reader that for this relation to be applicable δ_b must be a mode in the linear regime. On the other hand, there is no such restriction on the k modes of $P(k)$.

In terms of the power spectrum covariance, this relation has the direct interpretation that the measured power in the k and k' bins are correlated by the underlying background fluctuation δ_b that they share. Trispectrum consistency then implies that the covariance is

$$C_{ij} = C_{ij}^{\rm G} + C_{ij}^{T0} + (\sigma_W^L)^2 \frac{\partial P(k_i)}{\partial \delta_b}\frac{\partial P(k_j)}{\partial \delta_b}, \quad (19)$$

where we have introduced the variance of the background density field δ_b in the survey window, defined as

$$(\sigma_W^L)^2 \equiv \frac{1}{V_W^2}\int \frac{d^3\mathbf{q}}{(2\pi)^3}|\tilde{W}(\mathbf{q})|^2 P^{\rm L}(q). \quad (20)$$

In reducing Eq. (12) with the consistency trispectrum Eq. (18) to Eq. (19), we used the following identity for the window function from the convolution theorem:

$$\int \frac{d^3\mathbf{q}_1}{(2\pi)^3}\tilde{W}(\mathbf{q}_1)\tilde{W}(\mathbf{q}-\mathbf{q}_1) = \int d^3\mathbf{x} W(\mathbf{x})^2 e^{i\mathbf{x}\cdot\mathbf{q}} = \int d^3\mathbf{x} W(\mathbf{x}) e^{i\mathbf{x}\cdot\mathbf{q}} = \tilde{W}(\mathbf{q}). \quad (21)$$

We call the δ_b type covariance term the super-sample covariance of the power spectrum estimator. Note that because the background correlates changes in $\hat{P}(k)$ for all k, it can appear as the dominant contribution to the (co) variance when the measurement involves a large number of k-modes. It scales with the volume of the survey only through $(\sigma_W^L)^2$ whereas the other terms scale like white noise V_W^{-1}. Until the survey becomes much larger than the turnover of matter power spectrum on the horizon scale of matter-radiation equality, its relative contribution remains important [22,28] (see also Sec. III B).

As in the case of super-survey mode effects for number counts [39], their impact on the power spectrum should perhaps not be considered as excess noise but rather a new signal: the fractional power spectrum response is a template that changes the measured power spectrum coherently across bins as

$$P(k) \to P(k)\left(1 + \frac{\partial \ln P(k)}{\partial \delta_b}\delta_b\right). \quad (22)$$

The uncertainty is carried by a single unknown parameter δ_b drawn from a Gaussian of variance $(\sigma_W^L)^2$. Use of the covariance formalism in data analysis would premarginalize over this parameter, but it may alternately be a fitting parameter for the power spectrum that has a prior given by σ_W^L. This point of view may also be more appropriate for surveys that are not volume limited, i.e. where the window is not strictly 0 or 1. Loss of information on other cosmological parameters of interest that also change the power spectrum only comes through degeneracies with this single mode [see also [25,27], for the similar discussion].

It is also easy in this language to account for the alternate definition of the density fluctuation field as relative to the mean of the survey region. In this case the observed power would be rescaled as $P_W(k) = P(k)/(1+\delta_b)^2$ and the trispectrum consistency would take the same form as Eqs. (18) and (19) but with the response to the background mode altered to be

$$\frac{\partial P(k)}{\partial \delta_b} \to \frac{\partial P_W(k)}{\partial \delta_b} \approx \frac{\partial P(k)}{\partial \delta_b} - 2P(k), \quad (23)$$

which generalizes the treatment in Ref. [29]. Given that the fractional response function $\partial \ln P(k)/\partial \delta_b$ is typically positive and order unity, the response in $P_W(k)$ can be significantly reduced compared with $P(k)$. Note again that if the observable of interest does not depend on the mean density, which is the case for weak lensing shear that probes the differential gravitational field (tidal field), $P(k)$ rather than $P_W(k)$ is the relevant quantity.

The fractional response function $\partial \ln P(k)/\partial \delta_b$ is a quantity that can be directly calculated in simulations since introducing δ_b is equivalent to simulating a Friedman-Robertson-Walker universe with different background densities [40–42]. For example, in principle for its evaluation at $z = 0$ only two simulations would be required to calibrate all such effects deep into the nonlinear regime. Nonetheless, we find it illustrative to highlight the main features of power spectrum super-sample covariance, as well as derive Eq. (19) explicitly, analytically through the halo model. This will also allow us to make contact with the existing literature on this effect, BC in second order perturbation theory and "halo sample variance" (HSV) in the deeply nonlinear regime. There are also terms related converting $P(k)$ to observables related to changes in the distance-redshift relation if the background mode encompasses the entire volume out to the observer. These in

principle can affect the interpretation of baryon acoustic oscillation features (see [43] for related sub-survey effects). We leave a more precise calibration and observational considerations to future work.

To be comprehensive, in the Appendix we also develop the analogous formulas to describe the power spectrum covariance for a two-dimensional, density field that is obtained by projecting the three-dimensional field, weighted with a selection function, along the line of sight. The formulation developed in this paper can be straightforwardly extended to covariance theory for higher-order correlation functions (see [22], for the attempt to model the BC effect on the lensing bispectrum covariance), and the super-sample variance effects should be similarly described by the response of the higher-order correlation to a change in the background density.

III. HALO MODEL APPROACH

We have seen in the previous section that the power spectrum estimator covariance in the survey window is generally described by the matter trispectrum and the specific effects of super-sample covariance by its squeezed configurations. The matter trispectrum itself can be approximated in the halo model by considering correlations induced between dark matter halos via perturbation theory and within halos via the universal density profile [15,44]. Although the halo model is an empirical model to describe the nonlinear clustering, it gives a fairly accurate prediction compared to the simulations, e.g. to within a 10%–20% accuracy in the power spectrum amplitude at scales of interest [22]. Evaluating the covariance using the halo model can therefore illustrate the general consistency construction that the additional super-sample covariance terms are equivalent to the response of the power spectrum to a background density mode. We develop the halo model formalism in Sec. III A and illustrate it in the ΛCDM context in Sec. III B.

A. Formalism

In the halo model [44–47], the power spectrum itself is described as

$$P(k) = I_2^0(k, k) + [I_1^1(k)]^2 P^{\rm L}(k), \tag{24}$$

where the first term involves two points correlated by being in the same halo and the second two points in separate halos that are themselves correlated by the linear power spectrum. We use the general notation [15]

$$I_\mu^\beta(k_1, k_2, \ldots, k_\mu) \equiv \int dM \frac{dn}{dM}\left(\frac{M}{\bar\rho_m}\right)^\mu b_\beta \tilde u_M(k_1)\tilde u_M(k_2)\cdots\tilde u_M(k_\mu), \tag{25}$$

where M is the halo mass, dn/dM is the halo mass function, $b_0 = 1$, $b_1 = b(M)$ is the halo bias, and $\tilde u_M(k)$ is the Fourier transform of the halo density profile normalized so that $\tilde u_M(0) = 1$. We have here assumed linear halo bias in that $b_\beta = 0$ for $\beta \geq 2$ (see [15], for a possible extension of the halo model to including the nonlinear halo bias). This approximation does not affect the main results of this paper. Note that for a halo bias that satisfies the peak-background consistency relation

$$\int dM \frac{dn}{dM}\left(\frac{M}{\bar\rho_m}\right) b(M) = 1, \tag{26}$$

$I_1^1(0) = 1$, the 2-halo term as $k \to 0$ is simply the linear power spectrum. Furthermore the same peak-background consistency of the bias [30] says that

$$\frac{\partial I_2^0}{\partial \delta_b} = I_2^1, \tag{27}$$

which will be useful in relating the power spectrum covariance to its response to a background mode δ_b (see Sec. II C).

Likewise, the halo model approach tells us that the matter trispectrum arises from contributions involving one to four halos:

$$T = T^{1h} + (T_{22}^{2h} + T_{13}^{2h}) + T^{3h} + T^{4h}, \tag{28}$$

where $T^{1h}, \ldots, T^{4h}$ denote the 1-, 2-, 3- and 4-halo terms that arise from correlations between four points that reside in the 1 halo (the same halo) and from 2 to 4 different halos, respectively. Using the notations defined in Ref. [15], the different halo terms are given as

$$\begin{aligned}
T^{1h}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) &= I_4^0(k_1, k_2, k_3, k_4),\\
T_{22}^{2h}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) &= P^{\rm L}(k_{12}) I_2^1(k_1, k_2) I_2^1(k_3, k_4)\\
&\quad + 2\,\text{perm},\\
T_{13}^{2h}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) &= P^{\rm L}(k_1) I_1^1(k_1) I_3^1(k_2, k_3, k_4) + 3\,\text{perm},\\
T^{3h}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) &= B^{\rm PT}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_{34}) I_1^1(k_1) I_1^1(k_2) I_2^1(k_3, k_4)\\
&\quad + 5\,\text{perm},\\
T^{4h}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) &= T^{\rm PT}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) I_1^1(k_1) I_1^1(k_2)\\
&\quad \times I_1^1(k_3) I_1^1(k_4).
\end{aligned} \tag{29}$$

The 2-halo term has two contributions, T_{13}^{2h} and T_{22}^{2h} where one or two point(s) among the four points are in the first halo, with the remaining points in the second halo. Here $B^{\rm PT}$ and $T^{\rm PT}$ are the matter bispectrum and trispectrum given based on perturbation theory [48]

$$\begin{aligned}
B^{\rm PT}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) &= 2F_2(\mathbf{k}_1, \mathbf{k}_2) P^{\rm L}(k_1) P^{\rm L}(k_2) + 2\,\text{perm},\\
T^{\rm PT}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) &= 4[F_2(\mathbf{k}_{13}, -\mathbf{k}_1) F_2(\mathbf{k}_{13}, \mathbf{k}_2) P^{\rm L}(k_{13})\\
&\quad \times P^{\rm L}(k_1) P^{\rm L}(k_2) + 11\,\text{perm}]\\
&\quad + 6[F_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) P^{\rm L}(k_1) P^{\rm L}(k_2) P^{\rm L}(k_3)\\
&\quad + 3\,\text{perm}],
\end{aligned} \tag{30}$$

where

$$F_2(\mathbf{k}_1,\mathbf{k}_2) \equiv \frac{5}{7} + \frac{1}{2}\left(\frac{1}{k_1^2}+\frac{1}{k_2^2}\right)(\mathbf{k}_1\cdot\mathbf{k}_2) + \frac{2}{7}\frac{(\mathbf{k}_1\cdot\mathbf{k}_2)^2}{k_1^2k_2^2},$$
$$F_3(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) \equiv \frac{7}{18}\frac{\mathbf{k}_{12}\cdot\mathbf{k}_1}{k_1^2}[F_2(\mathbf{k}_2,\mathbf{k}_3)+G_2(\mathbf{k}_1,\mathbf{k}_2)] + \frac{1}{18}\frac{k_{12}^2(\mathbf{k}_1\cdot\mathbf{k}_2)}{k_1^2k_2^2}[G_2(\mathbf{k}_2,\mathbf{k}_3)+G_2(\mathbf{k}_1,\mathbf{k}_2)],$$
$$G_2(\mathbf{k}_1,\mathbf{k}_2) \equiv \frac{3}{7} + \frac{1}{2}\left(\frac{1}{k_1^2}+\frac{1}{k_2^2}\right)(\mathbf{k}_1\cdot\mathbf{k}_2) + \frac{4}{7}\frac{(\mathbf{k}_1\cdot\mathbf{k}_2)^2}{k_1^2k_2^2}. \tag{31}$$

Strictly speaking, the mode-coupling kernels, F_n and G_n, are exact only for an Einstein-de Sitter model ($\Omega_{\rm m0}=1$) and have a very weak dependence on the density parameters for models with $\Omega_{\rm m0}\neq 1$ [48]. Here we employ the Einstein-de Sitter approximation for simplicity. This impacts how we make the comparison to the response of the power spectrum to a background mode.

If $k, k' \gg q_i$ each halo term of the trispectrum can be approximated as

$$\begin{aligned}
T^{1h}(\mathbf{k},-\mathbf{k}+\mathbf{q}_{12},\mathbf{k}',-\mathbf{k}'-\mathbf{q}_{12}) &\simeq T^{1h}(k,k,k',k'),\\
T_{22}^{2h}(\mathbf{k},-\mathbf{k}+\mathbf{q}_{12},\mathbf{k}',-\mathbf{k}'-\mathbf{q}_{12}) &\simeq T_{22}^{2h}(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}') + P^{\rm L}(\mathbf{q}_{12})I_2^1(k,k)I_2^1(k',k'),\\
T_{13}^{2h}(\mathbf{k},-\mathbf{k}+\mathbf{q}_{12},\mathbf{k}',-\mathbf{k}'-\mathbf{q}_{12}) &\simeq T_{13}^{2h}(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}'),\\
T^{3h}(\mathbf{k},-\mathbf{k}+\mathbf{q}_{12},\mathbf{k}',-\mathbf{k}'-\mathbf{q}_{12}) &\simeq T^{3h}(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}') + 4I_2^1(k,k)[I_1^1(k')]^2P^{\rm L}(q_{12})P^{\rm L}(k')F_2(\mathbf{q}_{12},\mathbf{k}')\\
&\quad + 4I_2^1(k',k')[I_1^1(k)]^2P^{\rm L}(q_{12})P^{\rm L}(k)F_2(-\mathbf{q}_{12},\mathbf{k}),\\
T^{4h}(\mathbf{k},-\mathbf{k}+\mathbf{q}_{12},\mathbf{k}',-\mathbf{k}'-\mathbf{q}_{12}) &\simeq T^{4h}(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}') + 8[I_1^1(k)I_1^1(k')]^2P^{\rm L}(\mathbf{q}_{12})P^{\rm L}(k)P^{\rm L}(k')[F_2(\mathbf{q}_{12},-\mathbf{k})F_2(\mathbf{q}_{12},\mathbf{k}')\\
&\quad + F_2(\mathbf{q}_{12},-\mathbf{k})F_2(\mathbf{q}_{12},-\mathbf{k}')],
\end{aligned} \tag{32}$$

where we have used approximations such as $|\mathbf{k}+\mathbf{q}_1|\simeq k$.

Inserting the halo model expressions for the matter trispectrum into Eq. (12), we find the non-Gaussian term can be broken up as

$$C_{ij}^{\rm NG} = C_{ij}^{\rm T0} + C^{\rm SSC}, \tag{33}$$

where C_{ij}^{T0} was given in Eq. (14) and is the standard non-Gaussian term involving only oppositely directed Fourier modes [13]. The last term $C_{ij}^{\rm SSC}$ is the sum of the pieces involving $P^{\rm L}(q_{12})$ in the T_{22}^{2h}, T^{3h}, and T^{4h} terms of Eq. (32). We shall see that the combined terms are exactly the super-sample covariance term of Eq. (19).

To see this relationship let us begin with the remaining $P^{\rm L}$ piece of the T_{22}^{2h} term. It can be simplified as

$$\begin{aligned}
C_{ij}^{\rm HSV} &= \frac{1}{V_W^2}\int_{|\mathbf{k}|\in k_i}\frac{d^2\mathbf{k}}{V_{k_i}}\int_{|\mathbf{k}|'\in k_j}\frac{d^2\mathbf{k}'}{V_{k_j}}I_2^1(k,k)I_2^1(k',k')\int\left[\prod_{a=1}^{4}\frac{d^3\mathbf{q}_a}{(2\pi)^3}\tilde{W}(\mathbf{q}_a)\right](2\pi)^3\delta_D^3(\mathbf{q}_{1234})P^{\rm L}(\mathbf{q}_{12})\\
&= \frac{1}{V_W^2}\int_{|\mathbf{k}|\in k_i}\frac{d^2\mathbf{k}}{V_{k_i}}\int_{|\mathbf{k}|'\in k_j}\frac{d^2\mathbf{k}'}{V_{k_j}}I_2^1(k,k)I_2^1(k',k')\int\frac{d^3\mathbf{q}_1}{(2\pi)^3}\frac{d^3\mathbf{q}_2}{(2\pi)^3}\frac{d^3\mathbf{q}}{(2\pi)^3}\tilde{W}(\mathbf{q}_1)\tilde{W}(\mathbf{q}-\mathbf{q}_1)\tilde{W}(\mathbf{q}_2)\tilde{W}(\mathbf{q}-\mathbf{q}_2)P^{\rm L}(q)\\
&\simeq I_2^1(k_i,k_i)I_2^1(k_j,k_j)\frac{1}{V_W^2}\int\frac{d^3\mathbf{q}}{(2\pi)^3}|\tilde{W}(\mathbf{q})|^2P^{\rm L}(q),
\end{aligned} \tag{34}$$

where we have used the variable change, $\mathbf{q}_{12}\to\mathbf{q}$, in the 2nd line on the right-hand side, and we have use the identity [Eq. (21)] for the window function. Note that the remaining integral over $P^{\rm L}$ is simply the variance of the linear density field convolved with the window function, Eq. (20). The HSV term can now be simply expressed as

$$C_{ij}^{\rm HSV} = (\sigma_W^L)^2 I_2^1(k_i,k_i)I_2^1(k_j,k_j). \tag{35}$$

We have labeled this term "HSV" for halo sample variance as it takes the form originally pointed out in Ref. [19] in their Eq. (18) (see also Sec. 3.4 in Ref. [22]), although those works focused on the two-dimensional case (i.e. lensing power spectrum).

The super-sample covariance consistency relation Eq. (19) is now manifest with Eq. (27) in the 1-halo dominated regime

$$\frac{\partial \ln P(k)}{\partial \delta_b} \approx \frac{I_2^1(k,k)}{I_2^0(k,k)}. \tag{36}$$

In the halo model this consistency arises from compatibility of the bias with the mass function. The number density

of rare halos responds to the background mode more strongly according to its bias and so the response function Eq. (36) involves the average bias of the halos contributing to the bin in k.

Next let us consider the remaining $P^{\rm L}$ piece to the T^{4h} term. This can be similarly simplified as

$$C_{ij}^{\rm BC} = \frac{1}{V_W^2}\int_{|\mathbf{k}|\in k_i}\frac{d^2\mathbf{k}}{V_{k_i}}\int_{|\mathbf{k}|'\in k_j}\frac{d^2\mathbf{k}'}{V_{k_j}}\int\frac{d^3\mathbf{q}_1}{(2\pi)^3}\frac{d^3\mathbf{q}_2}{(2\pi)^3} \times\frac{d^3\mathbf{q}}{(2\pi)^3}\tilde{W}(\mathbf{q}_1)\tilde{W}(\mathbf{q}-\mathbf{q}_1)\tilde{W}(\mathbf{q}_2)\tilde{W}(\mathbf{q}-\mathbf{q}_2)P^{\rm L}(q) \times\{8[I_1^1(k)I_1^1(k')]^2P^{\rm L}(k)P^{\rm L}(k') \times[F_2(\mathbf{q},-\mathbf{k})F_2(\mathbf{q},\mathbf{k}')+F_2(\mathbf{q},-\mathbf{k})F_2(\mathbf{q},-\mathbf{k}')]\} \simeq\left(\frac{68}{21}\right)^2[I_1^1(k_i)I_1^1(k_j)]^2P^{\rm L}(k_i)P^{\rm L}(k_j)(\sigma_W^L)^2, \quad (37)$$

where we have used the identity [Eq. (21)] for the window function and used the azimuthal-angle average of the perturbation theory kernel F_2 as

$$\int_{-1}^{1}\frac{d\mu_{12}}{2}F_2(\mathbf{q}_1,\mathbf{q}_2)=\frac{5}{7}+\frac{2}{21}=\frac{17}{21}. \quad (38)$$

This type of term was called beat coupling (BC) in Ref. [16]. Our expression differs from their Eq. (94) only through the inclusion of the I_1^1 terms, which mediate the transition to the 1-halo regime. Reference [16] also treated beat coupling in the deeply nonlinear regime by employing a model of the nonlinear matter trispectrum based on hyper extended perturbation theory [49] and derived the similar formula to Eq. (37) for the nonlinear regime, where the covariance terms have the nonlinear matter power spectra instead of the linear spectra and have a different prefactor from 17/21. However, the halo model differs from hyper extended perturbation theory in the nonlinear regime, and Refs. [19,27] showed that the HSV effect based on the halo model [Eq. (34)] gives a much better agreement to the power spectrum covariance in the nonlinear regime measured from the simulations of weak lensing field than predicted from the BC effect using the hyper extended perturbation theory. For this reason, we will hereafter use the label "BC" to refer the super-sample covariance in the weakly nonlinear regime only.

With these simplifications, the BC term is also exactly what we obtain from the consistency relation in the perturbative regime where the 2-halo term dominates the power spectrum. The background response,

$$\frac{\partial\ln P(k)}{\partial\delta_b}=\frac{68}{21}, \quad (39)$$

can be derived by considering the growth in a separate universe with rescaled mean density in the Einstein-de Sitter approximation that matches the assumptions for the F_2 kernel [see Eq. 122 in Ref. [42]].

The final term is the P^L piece of the T^{3h} term:

$$C_{ij}^{\rm HSV-BC} = \frac{1}{V_W^2}\int_{|\mathbf{k}|\in k_i}\frac{d^2\mathbf{k}}{V_{k_i}}\int_{|\mathbf{k}|'\in k_j}\frac{d^2\mathbf{k}'}{V_{k_j}}\int\frac{d^3\mathbf{q}_1}{(2\pi)^3}\frac{d^3\mathbf{q}_2}{(2\pi)^3} \times\frac{d^3\mathbf{q}}{(2\pi)^3}\tilde{W}(\mathbf{q}_1)\tilde{W}(\mathbf{q}-\mathbf{q}_1)\tilde{W}(\mathbf{q}_2)\tilde{W}(\mathbf{q}-\mathbf{q}_2) \times P^{\rm L}(q)\{4[I_1^1(k)]^2I_1^2(k',k')P^{\rm L}(k)F_2(\mathbf{q},\mathbf{k}') + 4[I_1^1(k')]^2I_1^2(k,k)P^{\rm L}(k')F_2(\mathbf{q},\mathbf{k})\} \simeq(\sigma_W^L)^2\left\{\frac{68}{21}[I_1^1(k_i)]^2I_2^1(k_j,k_j)P^{\rm L}(k_i)+(i\leftrightarrow j)\right\}. \quad (40)$$

This contribution takes the form of the cross term between HSV and BC.

We can now see that the sum of all three terms,

$$C_{ij}^{\rm SSC}=C_{ij}^{\rm HSV}+C_{ij}^{\rm HSV-BC}+C_{ij}^{\rm BC}, \quad (41)$$

reproduces the general expression for the super-sample covariance of Eq. (19) with the total halo model response

$$\frac{\partial\ln P(k)}{\partial\delta_b}\approx\frac{\frac{68}{21}[I_1^1(k)]^2P^{\rm L}(k)+I_2^1(k,k)}{[I_1^1(k)]^2P^{\rm L}(k)+I_2^0(k,k)}. \quad (42)$$

Thus the halo model obeys the consistency relation for the trispectrum of Eq. (18) and illustrates that super-sample covariance can be described by the response of the power spectrum to the background density. Beyond the halo model, it can be accurately calibrated directly from simulations.

B. ΛCDM examples

To illustrate the power spectrum covariance for a ΛCDM model, we employ cosmological parameters that are consistent with the WMAP 9-yr result [50]: $\Omega_{\rm m0}h^2=0.136$, $\Omega_{\rm b0}h^2=0.0226$ and $\Omega_{\rm m0}=0.278$ for the density parameters assuming a flat universe and $n_s=0.972$ and $A_s=2.41\times10^{-9}$ at $k=0.002$ Mpc^{-1} for the primordial power spectrum parameters. We use the fitting formula in Ref. [51] to compute the transfer function for the model. Note $\sigma_8=0.83$. As for the halo model ingredients, the halo mass function, the halo mass profile, and the halo bias, we used the same modes as in Ref. [52].

The fundamental building block for the super-sample covariance effect is the variance of the density field averaged over the survey window, $(\sigma_W^L)^2$. Our approach is not limited to simple window geometries since the power spectrum at high k should respond to the background density in the same way regardless of geometry. Here we consider spherical- or cylinder-shaped geometries as working examples for which the window functions are

$$|\tilde{W}(\mathbf{k})|=3\frac{j_1(kr)}{kr}V_W,\quad V_W=(4\pi/3)r^3,$$
$$|\tilde{W}(\mathbf{k})|=2\frac{J_1(k_\perp r)}{k_\perp r}\frac{\sin(k_\parallel l/2)}{k_\parallel l/2}V_W,\quad V_W=(\pi r^2)l, \quad (43)$$

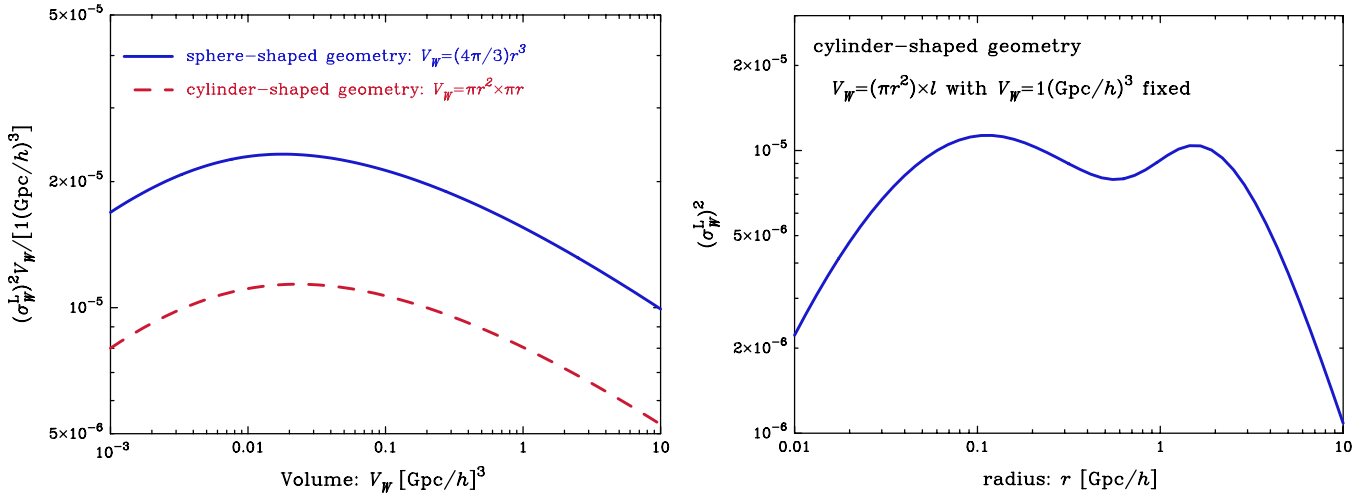

FIG. 1 (color online). The variance of the linear mass density field convolved with the survey window function, $(\sigma_W^L)^2$ for a ΛCDM model and $z = 0$. Left: Variance as a function of survey volume for spherical $V_W = (4\pi/3)r^3$ and cylindrical windows $V_W = \pi r^2 l$; $l = \pi r$. We show the variance multiplied by the volume to illustrate that for a wide range in survey volumes the relative impact of super-sample variance and the other terms which scale as $1/V_W$ remains the same within a factor of 2 due to the flatness of $P^{\mathrm{L}}(k)$ around matter-radiation equality. Right: Cylindrical windows of fixed volume $V_W = \pi r^2 l = 1(\mathrm{Gpc}/h)^3$. Optimizing for both minimum variance and a nonelongated geometry (see text) yields $l = \pi r$ as the best choice for volumes that exceed the matter-radiation turnover. As shown in the left panel, this choice has a lower variance than the spherical window of the same volume.

where $k^2 = k_\perp^2 + k_\parallel^2$. The left panel of Fig. 1 shows how the variance scales with survey window volume compared to $1/V_W$, which is the scaling of the standard covariance terms in T0 [Eqs. (13) and (14)]. The main result is that for both geometries the scaling of $(\sigma_W^L)^2$ differs from the standard scaling only by a factor of 2 across the whole 4 orders of magnitude in volume. This is because the matter power spectrum $P^{\mathrm{L}}(k)$ is relatively flat on the corresponding scales. It is only for volumes significantly greater than $1(\mathrm{Gpc}/h)^3$ that the super-sample variance effect declines relative to the other terms.

For a cylinder window, we employ a specific shape given by $l = \pi r$. This is motivated by the study in the right panel of Fig. 1 for the impact of the aspect ratio r/l or r at a fixed volume $1(\mathrm{Gpc}/h)^3$. For a general aspect ratio and volume, the variance scales as $k_\parallel k_\perp^2 P^{\mathrm{L}}(k)$. Since the maximum $k_\parallel k_\perp^2 \propto l^{-1} r^{-2} = \mathrm{const}$, minimizing the variance involves minimizing $P^{\mathrm{L}}(k)$ at the maximum k. Increasing from small r or a "tube" geometry, $k \propto 1/r$ and so the variance increases until the peak of the power spectrum is reached. It then declines until the minimum k is achieved when $l \approx \pi r$. The volume then becomes flattened into a "pill" geometry where $k \sim 1/l \propto r^{-2}$. The variance increases until the peak of $P^{\mathrm{L}}(k)$ is crossed in the opposite direction and then declines thereafter. Since either the extreme tube or pill cases have undesirable effects of washing out structure in $P^{\mathrm{L}}(k)$ via the convolution with the window and the smallest dimension cannot be in the nonlinear regime, this implies that for a cylindrical geometry the best shape for minimizing super-sample variance is $l \approx \pi r$. At this minimum point, the cylinder has a smaller variance by a factor of ∼2 compared with a sphere of the same volume.

The second building block of the super-sample covariance effect is the response of the nonlinear $P(k)$ to a fractional change in the background density δ_b. In Fig. 2, we study this response function $\partial \ln P/\partial \delta_b$ as predicted by the halo model [Eq. (42)]. In the one and two halo regimes, the response function is given by the HSV and BC effects respectively. Summing up the two terms gives a plateaulike feature in the response function up to a certain wave number, $k \sim 1\ h/\mathrm{Mpc}$ for $z = 0$ or somewhat smaller wave numbers for the higher redshift. This implies that the super-sample variance can be absorbed by a multiplicative, constant factor in the power spectrum amplitude, up to the certain wave number. In the halo model, higher wave number corresponds to smaller, less biased halos where the response decreases. Since halos of a given size are more rare at high redshift, there is an increase in response at higher redshift which creates a ridge at the transition from BC to HSV domination.

We can now put these two pieces together to form the contribution to the power spectrum covariance from the super-sample covariance terms. Figure 3 shows the non-Gaussian diagonal, or variance, term relative to the Gaussian expectation for a $V_W = 1(\mathrm{Gpc}/h)^3$ spherical survey with bins of $\Delta \log k = 0.1$. As expected, the non-Gaussian contributions are larger than Gaussian for $k \gtrsim 0.1\ h/\mathrm{Mpc}$ where the linear to nonlinear transition occurs or equivalently where the 2-halo to 1-halo transition occurs. Where the curve crosses unity specifically depends on the binning scheme since the Gaussian variance terms are suppressed as the number of independent k modes increases.

For the halo model, the BC term dominates over other non-Gaussian contributions in the weakly nonlinear

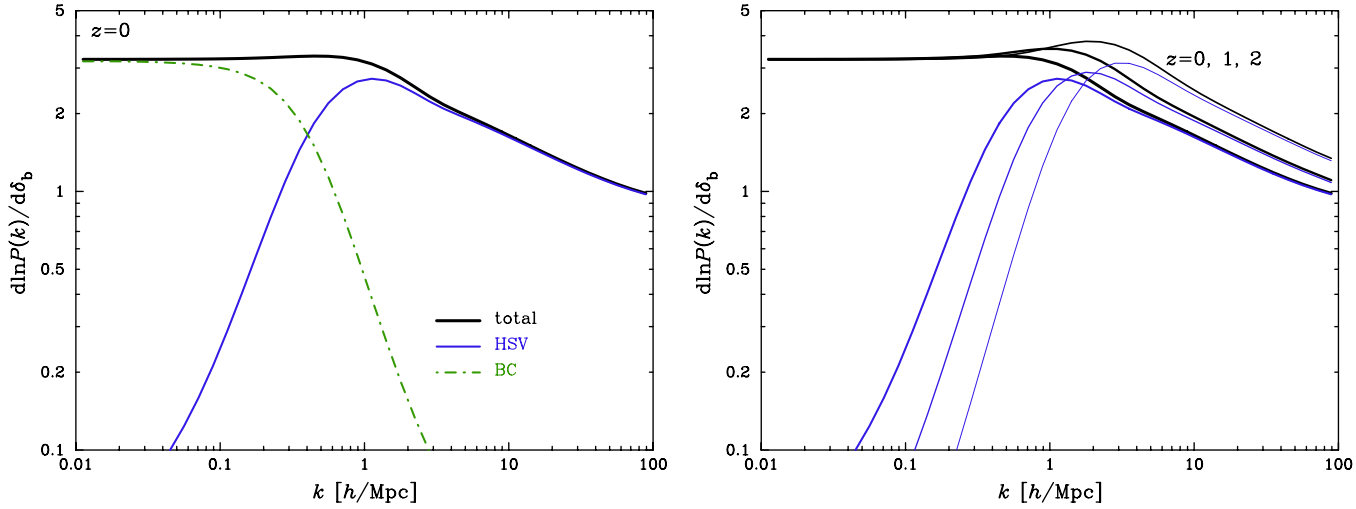

FIG. 2 (color online). The response function of the power spectrum to the super-survey mode, $\partial \ln P(k)/\partial \delta_b$ for the halo model [Eq. (42)]. Left: $z = 0$. Thin solid and dot-dashed curves are the halo sample variance and the beat-coupling term respectively. The total power shows a plateau up to $k \sim 1\ h/\mathrm{Mpc}$, having $\partial \ln P/\partial \delta_b \simeq 3$ for the amplitude. Right: $z = 0$, 1 and 2. At higher redshifts the plateau develops a small ridge due to the larger response of more rare halos.

regime and the HSV term in the deeply nonlinear regime with a smooth transition in between, which is mediated by the cross term. The total SSC contribution therefore dominates or is comparable to the standard T0 term everywhere.

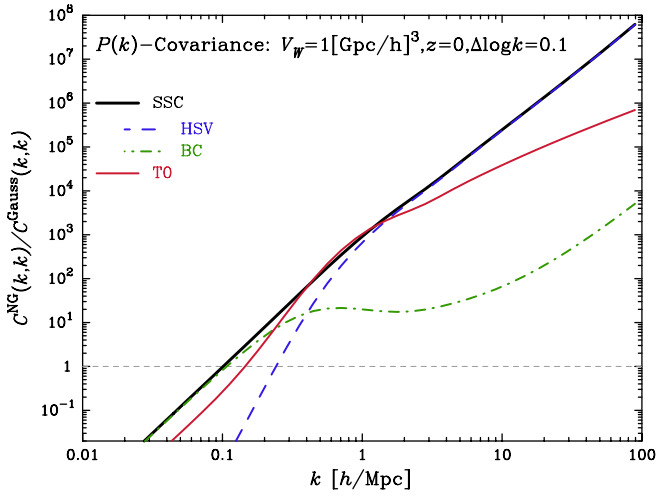

FIG. 3 (color online). The diagonal term of power spectrum covariance as a function of k in a ΛCDM model for a spherical survey volume of $V_W = 1(\mathrm{Gpc}/h)^3$, redshift $z = 0$, and k-binning of $\Delta \ln k = 0.1$, respectively. Plotted here is the non-Gaussian covariance contributions relative to the Gaussian term. At $k \gtrsim 0.1\ h/\mathrm{Mpc}$, the non-Gaussian errors start to dominate over the Gaussian term for the k-binning. The bold solid curve is the halo model prediction for the super-sample covariance (SSC), which is broken down into the dominant components on the weakly nonlinear and fully nonlinear scales respectively: the halo sample variance (HSV; dashed) and the beat-coupling effect (BC; dot-dashed) [see Eq. (41)]. The red (light) solid curve is the standard trispectrum term (T0) which is subdominant or comparable throughout.

These results are of course limited by the accuracy of the halo model itself which does not directly suffice for future surveys. Weak lensing cosmology, which is the primary science driver of planned imaging surveys, needs to use the power spectrum information up to $k \sim 1\ h/\mathrm{Mpc}$ to attain the full potential [53]. Galaxy clustering based cosmology aims at using the information up to at least a few $0.1\ \mathrm{Mpc}/h$ to capture the baryon acoustic oscillations as well as to measure the redshift-space distortion effect [9,54]. Thus an accurate simulation-based calibration of the background response $d \ln P/d\delta_b$ across this regime is important and will be presented in a separate paper.

IV. DISCUSSION

In this paper we have developed a simple, unified approach to describe the super-sample covariance of power spectrum estimation in a finite-volume survey. We show that the previously known effects of beat coupling and halo sample variance are both just limiting cases of the general response of the power spectrum to a change in the background density.

Formally, all power spectrum covariance effects are described by the matter trispectrum due to the two-point nature of power spectrum estimators. The super-sample covariance term arises due to the convolution of modes across the Fourier width of the survey window or "squeezed" trispectrum configurations which connect pairs of closely separated short-wavelength modes through a long-wavelength or super-survey mode. We show that there exists a consistency relation between these squeezed trispectra and the response of the power spectrum to a change in the background density. This consistency relation also exposes why this term is a covariance: the power

spectrum responds coherently across bins to a single unknown background density. Our formulation is general and applicable to any survey geometry since the response to a background mode does not depend on the detailed geometry of the window.

To make contact with the literature, we used the halo model trispectrum to illustrate these effects. We find that in the weakly nonlinear regime the response is exactly what is known as beat coupling. In the fully nonlinear regime it is exactly what is known as halo sample variance. The joint effect is the dominant non-Gaussian covariance for a wide range of survey volumes and its accurate calibration will be important for high-precision cosmology using large-scale structure probes. Our description also exposes the fact that accurate calibration is straightforward. Since the effect is the response of the power spectrum to a change in the background density, it only requires running two simulations with different background parameters to calibrate at any given redshift.

Our construction also exposes the possibility that power spectrum super-sample covariance need not be considered as an additional source of noise at all, but rather an additional signal of known shape but unknown amplitude given by the background density in the survey. By including it in the covariance, one premarginalizes the impact of the unknown amplitude. Alternately, one can fit for its amplitude given the template power spectrum response under a prior given by its expected variance in the survey window. This approach also directly exposes its impact on cosmological parameter estimation as different parameters will have different degeneracies with this background mode (see Ref. [27] for the similar discussion). It can also protect against unlikely realizations of the data by examining the dependence of results on the prior.

In principle, a joint fit could recover some information on super-survey modes from the observable sub-survey power spectrum albeit limited by degeneracies with global cosmological parameters. This opens up an interesting possibility to explore very large-wavelength fluctuations based on the background mode-coupling to short-wavelength observables.

ACKNOWLEDGMENTS

We thank Gary Bernstein, Nick Gnedin, Bhuvnesh Jain, Issha Kayo, Elisabeth Krause, Andrey Kravtsov, Marilena Loverde, Surhud More, Roland de Putter, Ryuiichi Takahashi, and Jochen Weller for useful discussions. M. T. was supported by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan, by the FIRST program "Subaru Measurements of Images and Redshifts (SuMIRe)," CSTP, Japan, and by Grant-in-Aid for Scientific Research from the JSPS Promotion of Science (23340061). W. H. was supported by U.S. Department of Energy Contract No. DE-FG02-90ER-40560, the David and Lucile Packard Foundation, and the Kavli Institute for Cosmological Physics at the University of Chicago through Grants No. NSF PHY-0114422 and No. NSF PHY-0551142 and an endowment from the Kavli Foundation and its founder Fred Kavli. W. H. acknowledges the hospitality of Kavli IPMU where this work was initiated.

APPENDIX: POWER SPECTRUM COVARIANCE FOR PROJECTED DENSITY FIELDS

The covariance formalism developed in the main paper can be directly applied to any statistic that is derived from the matter power spectrum. In particular, many applications such as weak lensing involve a weighted 2D projection of the density field. We use the flat-sky approach developed in Refs. [15,27] for the two-point statistics of a single projected field but the formalism can easily be generalized to cross spectra of multiple fields [52,55,56] or all-sky statistics [57–60].

We define the projected density field as

$$\Sigma(\boldsymbol{\theta}) = \int d\chi f(\chi)\delta(\chi\boldsymbol{\theta}, \chi), \tag{A1}$$

where we have assumed a spatially flat universe with radial coordinate $\chi(z) = \int_0^z dz'/H(z)$ and $f(\chi)$ is the radial selection function. For the weak lensing field, $f(\chi)$ is given by Eq. (4) in Ref. [27]. Taking into account the survey window function, the observed field is given as

$$\Sigma_{\mathcal{W}}(\boldsymbol{\theta}) = \mathcal{W}(\boldsymbol{\theta})\Sigma(\boldsymbol{\theta}), \tag{A2}$$

where $\mathcal{W}(\boldsymbol{\theta})$ is the survey window function on the sky; $\mathcal{W}(\boldsymbol{\theta}) = 1$ if the pixel in the direction $\boldsymbol{\theta}$ on the sky is in the survey region or contains data, otherwise $\mathcal{W}(\boldsymbol{\theta}) = 0$.

Under the flat-sky approximation, the Fourier-transformed field becomes

$$\tilde{\Sigma}_{\mathcal{W}}(\mathbf{l}) = \int \frac{d^2\mathbf{l}'}{(2\pi)^2}\tilde{\mathcal{W}}(\mathbf{l}-\mathbf{l}')\tilde{\Sigma}(\mathbf{l}'). \tag{A3}$$

Similarly to Eq. (3), the power spectrum estimator for the projected field is defined as

$$\hat{C}(l_i) \equiv \frac{1}{\Omega_{\mathcal{W}}}\int_{|\mathbf{l}|\in l_i}\frac{d^2\mathbf{l}}{\Omega_{l_i}}\tilde{\Sigma}_{\mathcal{W}}(\mathbf{l})\tilde{\Sigma}_{\mathcal{W}}(-\mathbf{l}), \tag{A4}$$

where $\Omega_{\mathcal{W}}$ is the effective survey area defined as $\Omega_{\mathcal{W}} \equiv \int d^2\boldsymbol{\theta}\mathcal{W}(\boldsymbol{\theta})$, and $\Omega_{l_i} = \int_{|\mathbf{l}|\in l_i} d^2\mathbf{l} \simeq 2\pi l_i \Delta l$ when $\Delta l/l_i \ll 1$.

In the limit that the angular mode of l is much greater than the width of the window function, the power spectrum estimator is unbiased in a sense that the ensemble average gives the underlying true power spectrum:

$$\langle\hat{C}(l_i)\rangle = C(l_i). \tag{A5}$$

As in the 3D case, we can derive the covariance of the 2D power spectrum including the effect of the window function:

$$\mathcal{C}_{ij} \equiv \langle \hat{C}(l_i)\hat{C}(l_j)\rangle - C(l_i)C(l_j) = \frac{1}{\Omega_{\mathcal{W}}}\left[\frac{(2\pi)^2}{\Omega_{l_i}} C(l_i)\delta^K_{ij} + \bar{\mathcal{T}}^{\mathcal{W}}(l_i, l_j)\right], \tag{A6}$$

where

$$\bar{\mathcal{T}}^{\mathcal{W}}(l_i,l_j) = \frac{1}{\Omega_{\mathcal{W}}}\int_{|\mathbf{l}|\in l_i}\frac{d^2\mathbf{l}}{\Omega_{l_i}}\int_{|\mathbf{l}'|\in l_j}\frac{d^2\mathbf{l}'}{\Omega_{l_j}} \times \int\left[\prod_{a=1}^{4}\frac{d^2\mathbf{q}_a}{(2\pi)^2}\tilde{W}(\mathbf{q}_a)\right](2\pi)^2\delta_D^2(\mathbf{q}_{1234}) \times \mathcal{T}(\mathbf{l}+\mathbf{q}_1, -\mathbf{l}+\mathbf{q}_2, \mathbf{l}'+\mathbf{q}_3, -\mathbf{l}'+\mathbf{q}_4). \tag{A7}$$

Using the Limber approximation, the angular power spectrum and trispectrum can be related to their 3D counterparts via

$$C(l) \approx \int d\chi f^2(\chi)\chi^{-2}P(\mathbf{k};\chi),$$
$$\mathcal{T}(\mathbf{l}_1,\mathbf{l}_2,\mathbf{l}_3,\mathbf{l}_4) \approx \int d\chi f^4(\chi)\chi^{-6}T(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4;\chi), \tag{A8}$$

where $\mathbf{k}_i = \mathbf{l}_i/\chi$.

The trispectrum consistency relation then implies

$$\mathcal{C}_{ij} = \mathcal{C}^{\mathrm{G}}_{ij} + \mathcal{C}^{\mathrm{T0}}_{ij} + \mathcal{C}^{\mathrm{SSC}}_{ij}, \tag{A9}$$

where

$$\mathcal{C}^{\mathrm{G}}_{ij} = \frac{1}{\Omega_{\mathcal{W}}}\frac{(2\pi)^2}{\Omega_{l_i}}C(l_i)\delta^K_{ij},$$
$$\mathcal{C}^{\mathrm{T0}}_{ij} = \frac{1}{\Omega_{\mathcal{W}}}\int_{|\mathbf{l}|\in l_i}\frac{d^2\mathbf{l}}{\Omega_{l_i}}\int_{|\mathbf{l}'|\in l_j}\frac{d^2\mathbf{l}'}{\Omega_{l_j}}\mathcal{T}(\mathbf{l},-\mathbf{l},\mathbf{l}',-\mathbf{l}'),$$
$$\mathcal{C}^{\mathrm{SSC}}_{ij} = \frac{1}{\Omega^2_{\mathcal{W}}}\int d\chi f(\chi)^4\chi^{-6}\frac{\partial P(k_i;\chi)}{\partial\delta_b}\frac{\partial P(k_j;\chi)}{\partial\delta_b} \times \int\frac{d^2\mathbf{l}}{(2\pi)^2}P^{\mathrm{L}}(k;\chi)|\tilde{W}(\mathbf{l})|^2, \tag{A10}$$

where we have again used the Limber relation $\mathbf{k}_i = \mathbf{l}_i/\chi$. With the halo model expressions for the response function from Eq. (42) these relations reproduce the results in the previous work, Eq. (18) in Ref. [19] and Eq. (17) in Ref. [27] for the HSV and BC contributions, respectively. In principle, background modes also change the $k \leftrightarrow l$ relation through perturbations to the distance redshift relation. However such modes would have to be constant out to the maximum distance in the integral and are typically much smaller than the δ_b considered here.

[1] http://www.naoj.org/Projects/HSC/index.html.
[2] http://www.darkenergysurvey.org.
[3] http://www.astro-wise.org/projects/KIDS/.
[4] http://www.lsst.org/lsst/.
[5] http://cosmology.lbl.gov/BOSS/.
[6] http://www.sdss3.org/future/eboss.php.
[7] http://bigboss.lbl.gov.
[8] http://sumire.ipmu.jp/pfs/intro.html.
[9] R. Ellis *et al.* (The PFS Team), arXiv:1206.0737.
[10] http://sci.esa.int/science-e/www/area/index.cfm?fareaid=102.
[11] http://wfirst.gsfc.nasa.gov.
[12] D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo, Phys. Rep. (in press).
[13] R. Scoccimarro, M. Zaldarriaga, and L. Hui, Astrophys. J. **527**, 1 (1999).
[14] W. Hu and M. White, Astrophys. J. **554**, 67 (2001).
[15] A. Cooray and W. Hu, Astrophys. J. **554**, 56 (2001).
[16] A. J. S. Hamilton, C. D. Rimes, and R. Scoccimarro, Mon. Not. R. Astron. Soc. **371**, 1188 (2006).
[17] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, Phys. Rev. D **74**, 023522 (2006).
[18] E. Semboloni, L. van Waerbeke, C. Heymans, T. Hamana, S. Colombi, M. White, and Y. Mellier, Mon. Not. R. Astron. Soc. **375**, L6 (2007).
[19] M. Sato, T. Hamana, R. Takahashi, M. Takada, N. Yoshida, T. Matsubara, and N. Sugiyama, Astrophys. J. **701**, 945 (2009).
[20] R. Takahashi, N. Yoshida, M. Takada, T. Matsubara, N. Sugiyama, I. Kayo, A. J. Nishizawa, T. Nishimichi, S. Saito, and A. Taruya, Astrophys. J. **700**, 479 (2009).
[21] H.-R. Yu, J. Harnois-Déraps, T.-J. Zhang, and U.-L. Pen, Mon. Not. R. Astron. Soc. **421**, 832 (2012).
[22] I. Kayo, M. Takada, and B. Jain, Mon. Not. R. Astron. Soc. **429**, 344 (2012).
[23] J. Harnois-Déraps and U.-L. Pen, Mon. Not. R. Astron. Soc. **423**, 2288 (2012).
[24] M. Manera *et al.*, Mon. Not. R. Astron. Soc. **428**, 1036 (2012).
[25] M. Takada and S. Bridle, New J. Phys. **9**, 446 (2007).
[26] M. C. Neyrinck and I. Szapudi, Mon. Not. R. Astron. Soc. **375**, L51 (2007).
[27] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. **395**, 2065 (2009).
[28] W. Hu and A. V. Kravtsov, Astrophys. J. **584**, 702 (2003).
[29] R. de Putter, C. Wagner, O. Mena, L. Verde, and W. Percival, J. Cosmol. Astropart. Phys. 04 (2012) 019.
[30] H. J. Mo, Y. P. Jing, and S. D. M. White, Mon. Not. R. Astron. Soc. **284**, 189 (1997).
[31] T.-J. Zhang, H.-R. Yu, J. Harnois-Déraps, I. MacDonald, and U.-L. Pen, Astrophys. J. **728**, 35 (2011).
[32] J. Lee and U.-L. Pen, Astrophys. J. Lett. **686**, L1 (2008).

[33] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, Astrophys. J. Lett. **698**, L90 (2009).
[34] H.-J. Seo, M. Sato, S. Dodelson, B. Jain, and M. Takada, Astrophys. J. Lett. **729**, L11 (2011).
[35] M. Takada and B. Jain, Mon. Not. R. Astron. Soc. **348**, 897 (2004).
[36] M. Tegmark *et al.*, Astrophys. J. **606**, 702 (2004).
[37] C. Hikage, M. Takada, T. Hamana, and D. Spergel, Mon. Not. R. Astron. Soc. **412**, 65 (2011).
[38] J. Maldacena, J. High Energy Phys. 05 (2003) 013.
[39] M. Lima and W. Hu, Phys. Rev. D **70**, 043504 (2004).
[40] E. Sirko, Astrophys. J. **634**, 728 (2005).
[41] N. Y. Gnedin, A. V. Kravtsov, and D. H. Rudd, Astrophys. J. Suppl. Ser. **194**, 46 (2011).
[42] T. Baldauf, U. Seljak, L. Senatore, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 10 (2011) 031.
[43] B. D. Sherwin and M. Zaldarriaga, Phys. Rev. D **85**, 103523 (2012).
[44] A. Cooray and R. Sheth, Phys. Rep. **372**, 1 (2002).
[45] J. A. Peacock and R. E. Smith, Mon. Not. R. Astron. Soc. **318**, 1144 (2000).
[46] U. Seljak, Mon. Not. R. Astron. Soc. **318**, 203 (2000).
[47] C. Ma and J. N. Fry, Astrophys. J. **543**, 503 (2000).
[48] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Phys. Rep. **367**, 1 (2002).
[49] R. Scoccimarro and J. A. Frieman, Astrophys. J. **520**, 35 (1999).
[50] G. Hinshaw *et al.*, arXiv:1212.5226.
[51] D. J. Eisenstein and W. Hu, Astrophys. J. **511**, 5 (1999).
[52] M. Oguri and M. Takada, Phys. Rev. D **83**, 023008 (2011).
[53] D. Huterer and M. Takada, Astropart. Phys. **23**, 369 (2005).
[54] M. Takada, E. Komatsu, and T. Futamase, Phys. Rev. D **73**, 083520 (2006).
[55] W. Hu and B. Jain, Phys. Rev. D **70**, 043009 (2004).
[56] T. Baldauf, R. E. Smith, U. Seljak, and R. Mandelbaum, Phys. Rev. D **81**, 063531 (2010).
[57] W. Hu, Phys. Rev. D **65**, 023003 (2001).
[58] N. Sehgal, P. Bode, S. Das, C. Hernandez-Monteagudo, K. Huffenberger, Y.-T. Lin, J. P. Ostriker, and H. Trac, Astrophys. J. **709**, 920 (2010).
[59] R. de Putter and M. Takada, Phys. Rev. D **82**, 103522 (2010).
[60] M. R. Becker, arXiv:1210.3069.