



CHORUS

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

Super-sample covariance of the power spectrum, bispectrum, halos, voids, and their cross covariances

Adrian E. Bayer, Jia Liu, Ryo Terasawa, Alexandre Barreira, Yici Zhong, and Yu Feng

Phys. Rev. D **108**, 043521 — Published 16 August 2023

DOI: [10.1103/PhysRevD.108.043521](https://doi.org/10.1103/PhysRevD.108.043521)

Super-sample covariance of the power spectrum, bispectrum, halos, voids, and their cross-covariances

Adrian E. Bayer,^{1,2,3,4,*} Jia Liu,^{4,†} Ryo Terasawa,^{4,5} Alexandre Barreira,^{6,7} Yici Zhong,^{5,8} and Yu Feng¹

¹*Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA*

²*Department of Physics, University of California, Berkeley, CA 94720, USA*

³*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

⁴*Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS,*

The University of Tokyo, Kashiwa, Chiba 277-8583, Japan

⁵*Department of Physics, School of Science, The University of Tokyo,*

7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

⁶*Excellence Cluster ORIGINS, Boltzmannstraße 2, 85748 Garching, Germany*

⁷*Ludwig-Maximilians-Universität, Schellingstraße 4, 80799 München, Germany*

⁸*Department of Astrophysical Sciences, Peyton Hall,*

Princeton University, Princeton, NJ 08544, USA

(Dated: July 19, 2023)

We study the effect of super-sample covariance (SSC) on the power spectrum and higher-order statistics: bispectrum, halo mass function, and void size function. We also investigate the effect of SSC on the cross-covariance between the statistics. We consider both the matter and halo fields. Higher-order statistics of the large-scale structure contain additional cosmological information beyond the power spectrum and are a powerful tool to constrain cosmology. They are a promising probe for ongoing and upcoming high precision cosmological surveys such as DESI, PFS, Rubin Observatory LSST, Euclid, SPHEREx, SKA, and Roman Space Telescope. Cosmological simulations used in modeling and validating these statistics often have sizes that are much smaller than the observed Universe. Density fluctuations on scales larger than the simulation box, known as super-sample modes, are not captured by the simulations and in turn can lead to inaccuracies in the covariance matrix. We compare the covariance measured using simulation boxes containing super-sample modes to those without. We also compare with the Separate Universe approach. We find that while the power spectrum, bispectrum and halo mass function show significant scale- or mass-dependent SSC, the void size function shows relatively small SSC. We also find significant SSC contributions to the cross-covariances between the different statistics, implying that future joint-analyses will need to carefully take into consideration the effect of SSC. To enable further study of SSC, our simulations have been made publicly available at [🔗](#).

I. INTRODUCTION

Ongoing and upcoming cosmological missions such as DESI¹ [1], PFS² [2], Rubin Observatory LSST³ [3], Euclid⁴ [4], SPHEREx⁵ [5], SKA⁶ [6], and Roman Space Telescope⁷ [7] will probe ever larger volumes of cosmic structure in the small-scale, nonlinear regime. These data contain rich information that can be used to constrain fundamental physics, such as dark energy, dark matter, and neutrino mass. To fully realize the potential of these surveys, many higher-order (or non-Gaussian) statistics have been proposed to extract additional information beyond the power spectrum (2-point function). These include, for example, the bispectrum (3-point function), halo mass function, void size function, probabil-

ity distribution function, marked power spectrum, and wavelet scattering transform [8–41]. They have been studied intensively in recent years and are becoming standard tools for cosmological inferences. Moreover, there is increased interest in joint analysis, in which second and higher-order statistics are combined to maximize the information gain (see e.g. [36, 42–45]).

Models for higher-order statistics usually rely on simulations for validation of analytic theories, calibration of semi-analytic models, or as the base of simulation-based inferences. To compute the covariance matrix of higher-order statistics, one typically requires a large set of simulations with different random initial conditions [46]. Such simulations assume periodic boundary conditions and are normally much smaller than the typical observed volumes of the Universe. Importantly, the mean density of these simulations is the cosmic one, and so by construction, they do not take into account the effects of perturbations with wavelengths longer than the size of the simulation. These so-called “super-sample modes” can however contribute sizeably to the covariance matrix; this effect is called the super-sample covariance (SSC) effect and must be carefully included to achieve accurate results.

To make contact with past literature, SSC has been

* abayer@berkeley.edu

† jia.liu@ipmu.jp

¹ <https://www.desi.lbl.gov>

² <https://pfs.ipmu.jp/index.html>

³ <https://www.lsst.org>

⁴ <https://www.euclid-ec.org>

⁵ <https://www.jpl.nasa.gov/missions/spherex>

⁶ <https://www.skatelescope.org>

⁷ <https://roman.gsfc.nasa.gov>

studied as the “beat-coupling” (BC) effect in the mildly nonlinear regime using perturbation theory, as “halo sample variance” (HSV) in the highly nonlinear regime using the halo model, and was sometimes called the DC mode effect as an analogy between the constant background fluctuation and constant electric Direct Current [47, 48]. It was first studied in the context of the power spectrum [49, 50], and its effects have since been quantified using direct simulations [51], perturbation theory [52–55], and separate universe simulations [56–58]. It has also been studied in relation to cluster counts [12, 59, 60], the matter bispectrum [9, 61–63], the matter one-point probability density function [64], the redshift space galaxy power spectrum [65–67], and the lensing power spectrum [68]. Furthermore, the effects of baryons on the SSC has been studied in [69–71]. The SSC effect associated with gravitational potential perturbations in cosmological with local primordial non-Gaussianity (i.e. $f_{\text{NL}} \neq 0$) has also recently been studied [72]. Fast, approximate, methods exist to account for the SSC in forecasts for upcoming lensing and photometric surveys [73–75].

In this paper, we study the effects of SSC for the power spectrum and several higher-order statistics: the bispectrum, halo mass function, and void size function. To do so, we compare the statistics measured using small periodic boxes, which ignore SSC, to those using equally-sized boxes that are embedded in a much larger simulation, which include SSC. We study the effect of SSC in both the total matter field and the halo field. We also validate our results against the separate universe (SU) approach, in which the SSC contribution is calculated semi-analytically using the response of the statistics to certain changes in the cosmological parameters.

Our work is the first to investigate the effects of SSC for voids. Cosmic voids have been studied intensively in recent years [20, 22, 36, 37, 76–83] and have achieved cosmological constraints with observational data [42, 84–86]. The bias parameters of voids have also been recently studied using SU simulations [87, 88]. Our work is also the first to study the effects of SSC on the cross-covariance between the combinations of all of these statistics.

The paper is organized as follows. Section II outlines the methods employed to compute the SSC of the power spectrum, bispectrum, halo mass function, and void size function. Section III presents the results for SSC of these statistics and their cross-covariances. We conclude in Section IV.

II. METHOD

In this section, we describe the methods used to run the N-body simulations, to compute the statistics, and to compute the SSC. We also briefly describe the SU approach.

A. Covariance

The covariance matrix between an observable \mathcal{O}_α and another observable \mathcal{O}_β is given by

$$C_{\alpha\beta} = \langle (\mathcal{O}_\alpha - \langle \mathcal{O}_\alpha \rangle) (\mathcal{O}_\beta - \langle \mathcal{O}_\beta \rangle) \rangle, \quad (1)$$

where $\langle \rangle$ denotes the mean over realizations. The α and β subscripts can refer to different bins of particular statistic, or two completely different statistics. The covariance can be estimated by evaluating Eqn. 1 using an ensemble of simulations with different random realizations of the initial conditions.

We quantify the SSC effect by comparing the following two sets of simulations:

- **sub-boxes** that are embedded in a much larger simulation, where the effect of SSC is properly captured;
- **small** boxes that are of the same size and resolution as the sub-boxes, but are independently simulated with periodic boundary conditions and have no super-sample modes.

Because the super-sample modes are only present in the former and not in the latter, the SSC is given by

$$C_{\text{SSC}} = C_{\text{sub}} - C_{\text{small}}, \quad (2)$$

where C_{sub} is the covariance computed using sub-boxes and C_{small} is that using small boxes.

B. N-body simulations

We use **FastPM** [89, 90], a particle-mesh (PM) N-body simulation, to simulate a big box of side-length 5 Gpc/h with 2048^3 matter particles. We then split this big box into $8^3 = 512$ sub-boxes. We compare these to 512 independent, periodic, small boxes of size $5000/8=625$ Mpc/h, each with 256^3 particles. The resolution of the small boxes is chosen to match the big box. We consider a maximum scale cut of $k_{\text{max}} = 0.8 h/\text{Mpc}$ as scales with lower k than this are well modeled by our simulations. In all cases, we begin the simulations at $z = 9$ and take 60 steps to $z = 0$. The resolution of the force mesh is 2 times the number of particles. For simplicity, we consider and discuss only the results at $z = 0$, which is when SSC is expected to be the strongest.

Our cosmological parameters are $h = 0.6774$, $\Omega_m = 0.3089$, $\Omega_b = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$, $M_\nu = 0$.

To identify halos, we use the Friends-of-Friends (FoF) algorithm with a linking length of 0.2. We generate the particle and halo overdensity fields using the Cloud-in-Cell (CIC) method with $N_{\text{mesh}} = 256$ using **nbodykit** [91]. Further, we consider the halo field in real and not in redshift space. We compensate the field for window effects before calculating the statistics [92]. For matter the overdensity field is computed as $\delta = (\rho - \bar{\rho})/\bar{\rho}$,

where ρ is the matter density, while for halos it is computed as $(n - \bar{n})/\bar{n}$, where n is the number density. In order to compute the overdensity field in the sub-boxes, there are two choices for the mean density $\bar{\rho}$ (or \bar{n}): using the “**global**” mean of the big box, or the “**local**” mean of the sub-box. Realistically, for weak lensing surveys it is appropriate to use the global mean as the mean density can be directly calculated from the cosmological model, moreover, the measured weak-lensing shear field is sensitive to the global-mean density. However, for galaxy surveys, since one does not know how to predict from first principles the total number of galaxies, the local mean is what is most appropriate as we measure the galaxy statistics w.r.t. the observed galaxy number density in the survey. We consider both cases in our analysis.

C. Statistics

We now define the various statistics considered in this work, and the method used to compute them. We plot the statistics averaged over all the sub-boxes in Fig. 1.

- The **power spectrum** $P(k)$ is defined as the Fourier transform of the 2-point correlation function $\xi(x_1 - x_2) \equiv \langle \delta(x_1)\delta(x_2) \rangle$, where δ is the overdensity field. Defining the fundamental frequency of the box as $k_F \equiv 2\pi/625 h/\text{Mpc}$, we use 15 linear bins between bin edges $[0, 80k_F]$, with $\Delta k = 5k_F$.
- The **bispectrum** $B(k_1, k_2, k_3)$ is defined as the Fourier transform of the 3-point correlation function $\langle \delta(x_1)\delta(x_2)\delta(x_3) \rangle$. We consider three particular configurations of the bispectrum: equilateral ($k_1 = k_2 = k_3$), isosceles ($k_1 = k_2 \neq k_3$), and squeezed ($k_1 = k_2, k_3 \sim 0$). More concretely, we consider the squeezed mode as $k_3 = 3.6 \times 10^{-2} h/\text{Mpc}$. We compute the bispectrum using the `bskit` package [93], which employs the FFT-based bispectrum estimators of [94, 95]. The k binning is the same as for the power spectrum.
- The **halo mass function** (HMF), denoted $dn^H/d\ln M$, is defined as the comoving number density of halos n^H per unit of log halo mass $\ln M$. We consider 7 logarithmic bins bounded by $M_{\min} = 10^{14} M_\odot/h$ and $M_{\max} = 10^{16} M_\odot/h$.
- The **void size function** (VSF), denoted dn^V/dR , corresponds to the comoving number density of voids n^V per unit of void radius R . We consider spherical voids in smoothed density fields. The $N_{\text{mesh}} = 256^3$ field is smoothed with top-hat filters of size R_{filter} in 7 linear bins between $R_{\min} = 12.2 \text{ Mpc}/h$ and $R_{\max} = 100.1 \text{ Mpc}/h$, with $\Delta R = 6d_{\text{grid}}$ where $d_{\text{grid}} \approx 2.44 \text{ Mpc}/h$ is the grid size. We search hierarchically, first finding the largest voids and then the, more abundant, smaller voids. Voids of size $R = R_{\text{filter}}$ are defined as local minima in the R_{filter} -filtered field, with values

lower than a predefined threshold δ_{th} , unless they overlap with existing larger voids. In this work we investigate thresholds of $\delta_{\text{th}} = -0.3$ and -0.5 . The void finding algorithm was developed by [96] and we use the implementation in `Pylians3` [97].

D. Separate universe simulations

We now briefly summarize an alternative method to compute the SSC using Separate Universe (SU) simulations; we refer the reader to [56, 57] for more details. In this approach, the effect of a super-survey density mode that is constant inside the box and has amplitude δ_b is mimicked by adjusting the cosmological parameters such that $\bar{\rho}_m \rightarrow \bar{\rho}_m (1 + \delta_b)$, where $\bar{\rho}_m$ is the mean physical matter density; if the fiducial cosmology is a spatially flat universe, this implies that the separate universe has non-zero curvature ($\Omega_k \neq 0$). The response of any summary statistic \mathcal{O} to δ_b is computed by considering the difference between simulations run with different δ_b . The SSC is then approximated by

$$C_{\text{SSC-SU}}^{ij} \simeq \sigma_b^2 \frac{d\mathcal{O}_i}{d\delta_b} \frac{d\mathcal{O}_j}{d\delta_b}, \quad (3)$$

where σ_b^2 is the variance of the linear matter density fluctuations on the size of the survey described by a window function W ,

$$\sigma_b^2 \equiv \frac{1}{V_W^2} \int \frac{d^3\mathbf{k}}{(2\pi)^3} |W(\mathbf{k})|^2 P_{\text{lin}}(k), \quad (4)$$

where $V_W = \int d^3\mathbf{x} W(\mathbf{x})$ is the survey volume and $P_{\text{lin}}(k)$ is the linear matter power spectrum. The window function used in this work corresponds to a 3d cube of side-length $625 \text{ Mpc}/h$, giving $\sigma_b^2 = 6.8 \times 10^{-5}$. Concretely, we evaluate the responses using finite difference methods on simulations with $\delta_b = \pm 0.03$, and averaging over 20 realizations of the initial conditions.

The δ_b mode modifies the background expansion history, which implies some care when choosing the box size of the simulations with $\delta_b \neq 0$. In our simulations here, we choose the comoving box size to match at all times in Mpc units. This corresponds to the “growth-dilation” methods in the notation of [56], or equivalently, with our SU simulations we measure the so-called “growth-only” responses in the language of [57]. Importantly, when identifying halos in the simulations, the FoF linking length in the separate universe needs to be rescaled by the ratio of the scale factors in the two simulations to guarantee matching halo definitions (see e.g. [98] for a discussion).

Since the simulations with different δ_b values can be initialized with the same random phases of the initial conditions, the SU approach has the significant advantage of converging with much fewer simulations than the sub-box approach (discussed in Section II B). We note, however, that our SU simulations account only for the

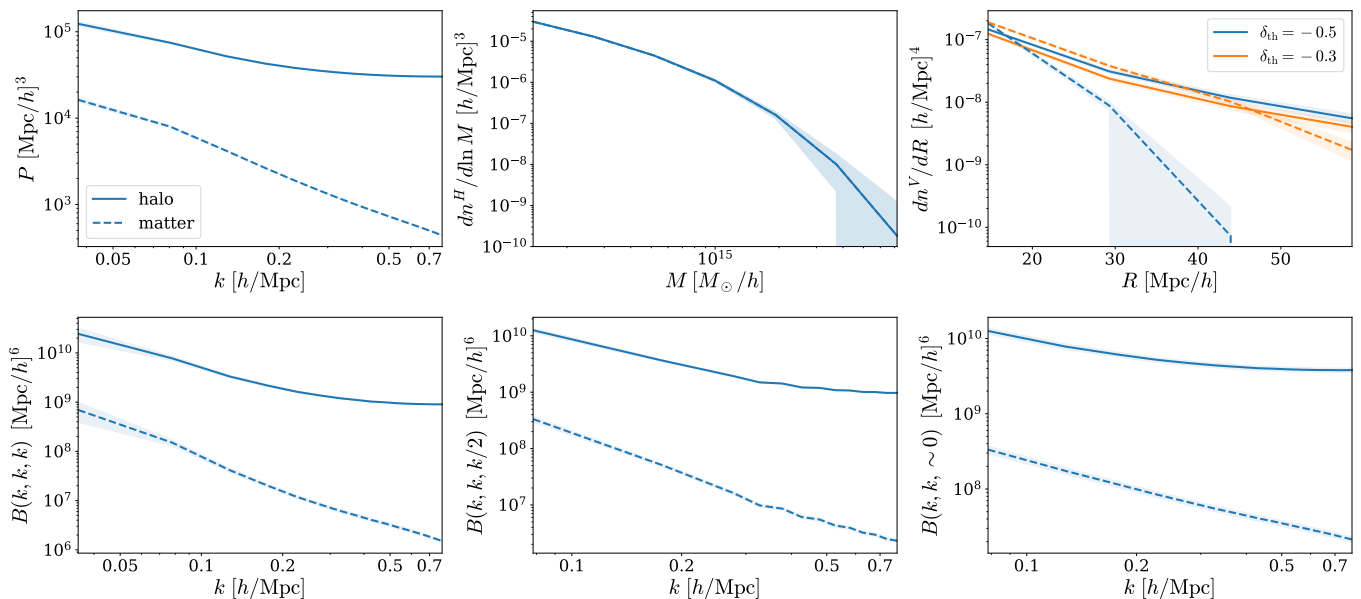


FIG. 1: Power spectrum (top left), halo mass function (top middle), void size function (top right), and three bispectrum configurations (bottom) averaged over all small boxes. Error bands correspond to the standard deviation across the small boxes, $\sqrt{C_{\text{small}}}$. We show results for the halo (solid) and matter (dashed) field. For the void size function, we show results for two density thresholds used for void searching, $\delta_{\text{th}} = -0.5$ (blue) and $\delta_{\text{th}} = -0.3$ (orange).

impact of isotropic density perturbations as super-survey modes, i.e., they do not account in particular for the effect of super-survey tidal fields [99–101]. Here, we consider angular averaged spectra in real space, for which the impact of super-survey tidal fields averages out, but we note that for analyses in redshift space [65–67] and weak-lensing applications [68] this is not the case and the super-survey tidal fields can have a non-negligible effect. Further, the super-survey tidal fields contribute also to the SSC effect of halo and void counts, although in a weak manner since this happens only at second order – this is because the tidal field is a tensor, and thus must be contracted at least once (forming a quadratic quantity) if it is to affect any scalar statistic like counts [102]. On the other hand, the SSC calculated using the sub-box approach automatically includes both the effects of density and tidal fields.

III. RESULTS

Here we show the effect of SSC for individual statistics as well as their cross-covariances. In all plots, error bars are computed using bootstrapping and correspond to the 95% confidence interval.

A. Matter field statistics

Fig. 2 shows the results of the SSC contribution for the power spectrum, void size function, halo mass func-

tion, and three bispectrum configurations. Each statistic contains two panels, the upper of which shows the diagonal term of the covariance computed with and without SSC, using sub- and small boxes, respectively. For sub-boxes, we show results using both the local mean and the global mean density. The lower panel shows the ratio between the sub and small box, which is equal to $C_{\text{sub}}/C_{\text{small}} = 1 + C_{\text{SSC}}/C_{\text{small}}$ (using Eqn. 2). Shown also is the result from the SU approach (marked by the red crosses and pluses for the local and global mean cases respectively), which agrees reasonably well overall with the SSC estimated from the sub-box approach.

For the case of our power spectrum and bispectrum results, we note also that while the SSC does not depend to first order on the size of the wavenumber bins, other contributions to the covariance typically do, which can have an impact on the exact values of $C_{\text{sub}}/C_{\text{small}}$ (note this does not mean there is a dependence of the signal-to-noise on the bin size). This does not have an impact on the main takeaway points of our results, but it is useful to keep in mind especially when comparing quantitatively to results obtained previously in the literature.

1. Power spectrum

The SSC of the matter power spectrum can be seen to be a just under a 100% effect at $k = 0.7 h/\text{Mpc}$ in the local mean case. However, the SSC is much larger when using the global mean, with the ratio $C_{\text{sub}}/C_{\text{small}}$ increasing with k to a factor of ~ 4 at $k = 0.7 h/\text{Mpc}$.

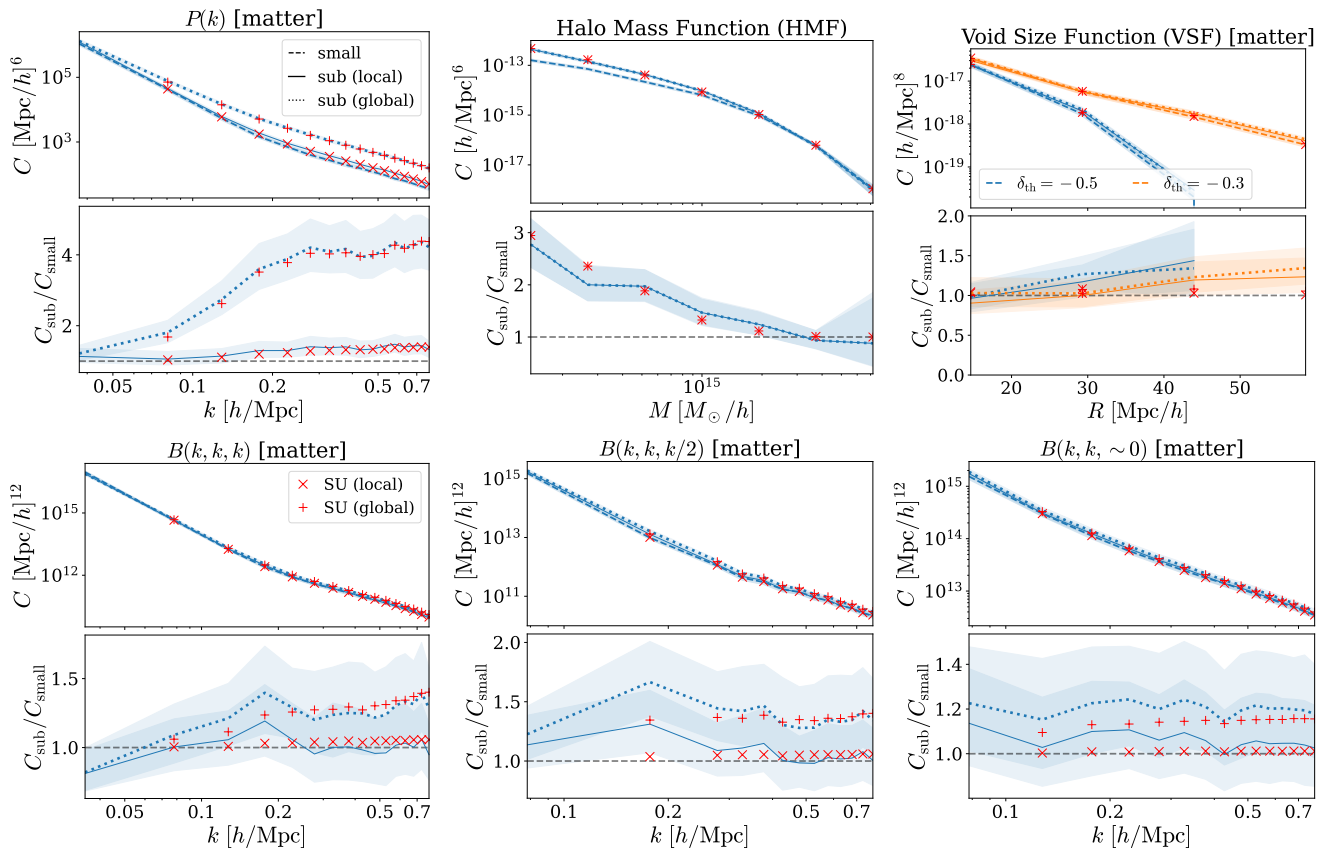


FIG. 2: Covariances for the power spectrum (top left), halo mass function (top middle), void size function (top right), and three bispectrum configurations (bottom) in the matter field. Each statistic contains two panels, the top of which shows the diagonal term of the covariance computed in the small box (dashed), and in the sub-box using local (solid) and global (dotted) mean. The lower panel shows the ratio between the sub- and small boxes, where the dashed horizontal grey line indicates no SSC. For the void size function, we show results for two density thresholds used for void searching, $\delta_{\text{th}} = -0.5$ (blue) and $\delta_{\text{th}} = -0.3$ (orange). We also show separate universe results for SSC computed using local and global mean densities, marked in “x” and “+”, respectively. Shaded regions correspond to 95% confidence intervals.

This can be explained as follows: the local mean density is modified with respect to the global mean by the background density δ_b , as $\bar{\rho}_{\text{local}} = \bar{\rho}_{\text{global}}(1 + \delta_b)$. Thus the power spectrum with respect to the local mean is given by $P_{\text{local}}(k) = (1 + \delta_b)^{-2} P_{\text{global}}(k)$, where $P_{\text{global}}(k)$ is the power spectrum with respect to the global mean. The local and global responses are then related as,

$$\frac{d \ln P_{\text{local}}(k)}{d \delta_b} \approx \frac{d \ln P_{\text{global}}(k)}{d \delta_b} - 2, \quad (5)$$

where we use the fact that $\delta_b = 0.03 \ll 1$. It can be shown with perturbation theory that the global response is close to 2 for the scales considered in this paper [50, 51, 56], hence the local response is much suppressed in comparison to the global. Recall, the SSC of the power spectrum referenced to the global density is what is relevant to weak lensing analysis, and this strong response is ultimately the reason why SSC is the most important piece of the off-diagonal covariance in cosmic shear 2-

point function studies [68, 103].

2. Halo Mass Function

We find the SSC has very little contribution to the counts of massive halos $\gtrsim 10^{15} M_{\odot}/h$, while it increases towards less massive halos, with the ratio $C_{\text{sub}}/C_{\text{small}}$ becoming roughly a factor of 3 for masses of $\sim 10^{14} M_{\odot}/h$. The fluctuation in the number density of halos $\delta n = \delta n(M)$ in a mass bin M is a biased tracer of the underlying matter field δ_m

$$\frac{\delta n}{\bar{n}} = b \delta_m, \quad (6)$$

where $b = b(M)$ is the halo bias in the mass bin. The diagonal term of the sub-box HMF covariance divided by the shot noise $C_{\text{small}} = \bar{n}/V$ is thus [59],

$$C_{\text{sub}}/C_{\text{small}} = 1 + \sigma_b^2 b^2 \bar{N}, \quad (7)$$

where $\bar{N} = \bar{n}V$ is the number (or abundance) of halos. The second term on the right is the SSC contribution. While massive halos tend to be more biased (by a factor of few compared to low mass halos), their abundance is exponentially suppressed. Thus, the most massive halos are in the shot-noise dominated regime with little contribution from SSC.

3. Void size function

Recall, we consider spherical voids with density thresholds $\delta_{\text{th}} = -0.5$ and -0.3 , whose results are shown on the top right panel of Fig. 2. The SSC contribution for voids is generally small for all void radii shown. Following from the discussion above for the HMF, this is as expected since voids are approximately 100 times less abundant than the halos, and their bias values remain of order $0 - 10$ [87, 88]. Thus, the covariance is strongly dominated by shot noise and the SSC effect is negligible. Note further that for certain void radii, and contrary to the case of halos, the void bias can be zero, in which case even an infinite abundance would have no SSC.

It can, however, be seen that the error bars from the small versus sub-box analysis allows for an $\mathcal{O}(10 - 100\%)$ SSC effect, with a slight increase with R , while the SU approach suggests a negligible SSC for all R . This is likely a numerical artifact due to the void finder assuming periodic boundary conditions: the covariance of the sub-boxes, which are not periodic, could be spuriously large, especially for large voids that extend beyond the boundary of the box. On the other hand, the SU results are based on the difference between two periodic boxes, and are thus free of such boundary effects. We leave correcting this boundary effect for future work.

4. Bispectrum

All three configurations of the bispectrum that we consider have a much smaller SSC than the power spectrum, typically a factor 1–2 effect for $k \lesssim 0.7h/\text{Mpc}$. Similarly to the power spectrum, the global case has a larger effect than local. This agrees within error bars with [61], which considered the equilateral and isosceles cases using an analytical response approach and simulations. Also [63] found the SSC effect to be small for the matter bispectrum. Our results are also in good agreement with [62] which analytically derived that the SSC effect is small for the squeezed matter bispectrum, as a few multiplicative factors get suppressed by the soft (low k) modes. More generally speaking, one can intuit that the matter bispectrum has a smaller SSC than the power spectrum as follows. While the power spectrum only has a Gaussian piece and a connected 4-pt function non-Gaussian piece (which includes the SSC), the bispectrum consists of additional non-Gaussian disconnected pieces (proportional to the square of the bispectrum and the power spec-

trum times the trispectrum) which increase the covariance overall. Thus, the connected 6-pt function (which is where the bispectrum’s SSC enters) is small in comparison to the lower-order contributions to the bispectrum covariance.

B. Halo field statistics

Fig. 3 shows the results for statistics computed from the halo field, which is a biased tracer of the matter field. Here, we apply a minimum mass cut of $M_{\text{min}} = 10^{14}M_{\odot}/h$. The halo field SSC shows qualitative similarities to that of the matter field, namely larger SSC on smaller scales for the power spectrum and the bispectrum, and a small SSC for voids.

One noteworthy difference concerns the relative size of the local and global cases for the power spectrum, which is now comparable. This is because the local and global responses for halos are related as

$$\frac{d \ln P_{\text{local}}^h(k)}{d\delta_b} \simeq \frac{d \ln P_{\text{global}}^h(k)}{d\delta_b} - 2b. \quad (8)$$

Unlike in the case of the local matter response (Eqn. 5), where the global response nearly cancels with the -2 term, the global halo response is corrected by $-2b_1$. In our work the bias is $b_1 \approx 2.5$, which leads to a negative local halo response (see Fig. 12 of [104] and derivation therein), so much so that the local effect is now comparable to the global effect after taking the square of the response in Eqn. 3. The exact value of the response is mass and redshift dependent, as halo bias increases when considering more massive halos and/or halos at higher redshift. The SSC of the halo bispectrum will be similarly affected by the bias and mass cuts, the exact manner in which requires future theoretical study. For the setup considered in this work, we find that the global halo bispectrum SSC remains small for all configurations, while the local SSC becomes large for the equilateral and isosceles configurations.

C. Cross-covariance

In this section, we study the SSC contribution to the cross-covariance of the statistics. Higher-order statistics typically have large off-diagonal terms in the covariance compared to the power spectrum and are usually studied jointly with other statistics. Therefore, it is important to study not only the variances of individual statistics, but also their cross-covariance. To focus on the off-diagonal terms, we normalize the covariance using the diagonal term to obtain the correlation matrix

$$r_{ij} \equiv C_{ij} / \sqrt{C_{ii}C_{jj}}, \quad (9)$$

where C is the covariance matrix with indices run through the bins of all the statistics studied here. Fig. 4

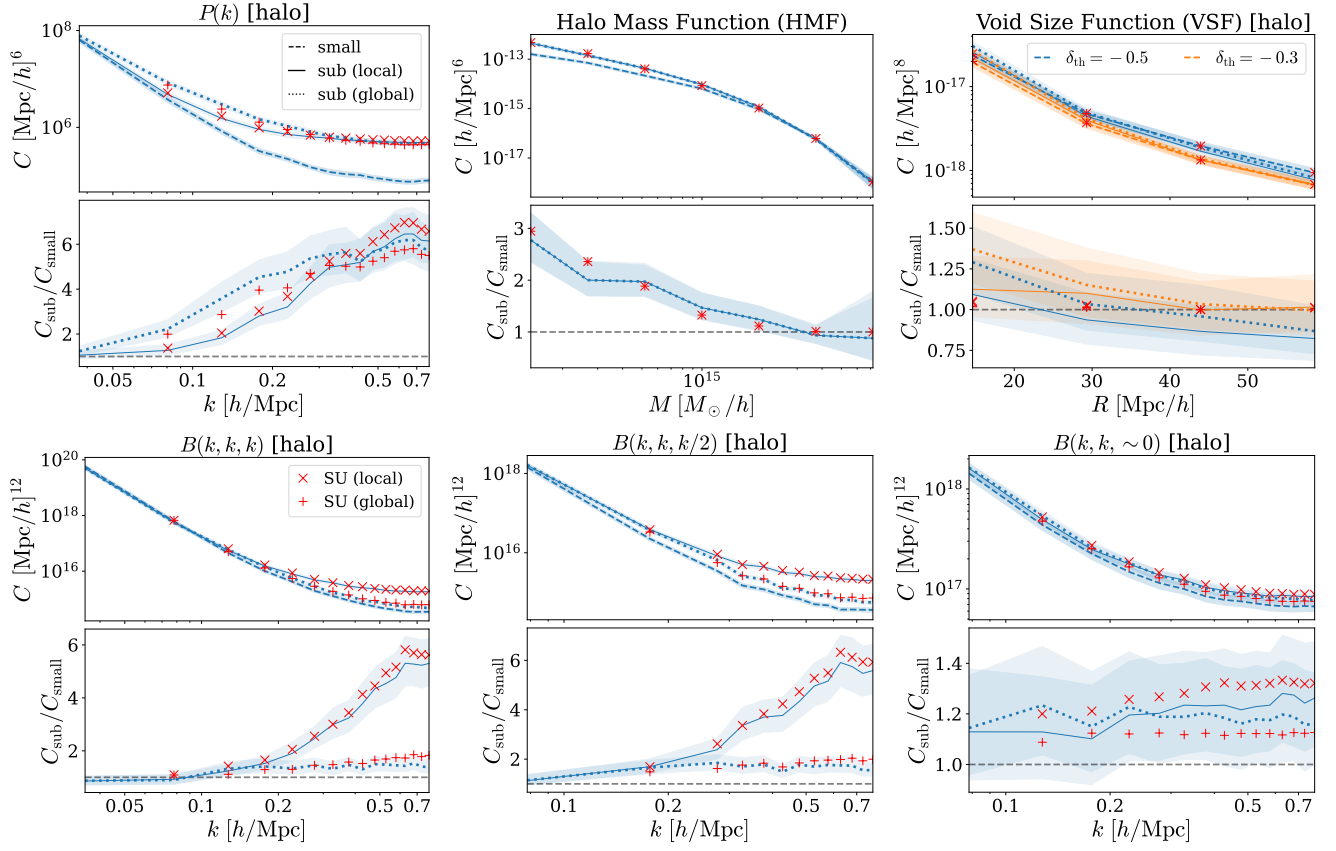


FIG. 3: Same as Fig. 2 but for the halo field, with a minimum mass cut $M_{\min} = 10^{14} M_{\odot}/h$. The halo mass function plot (top middle) is duplicated here for completeness.

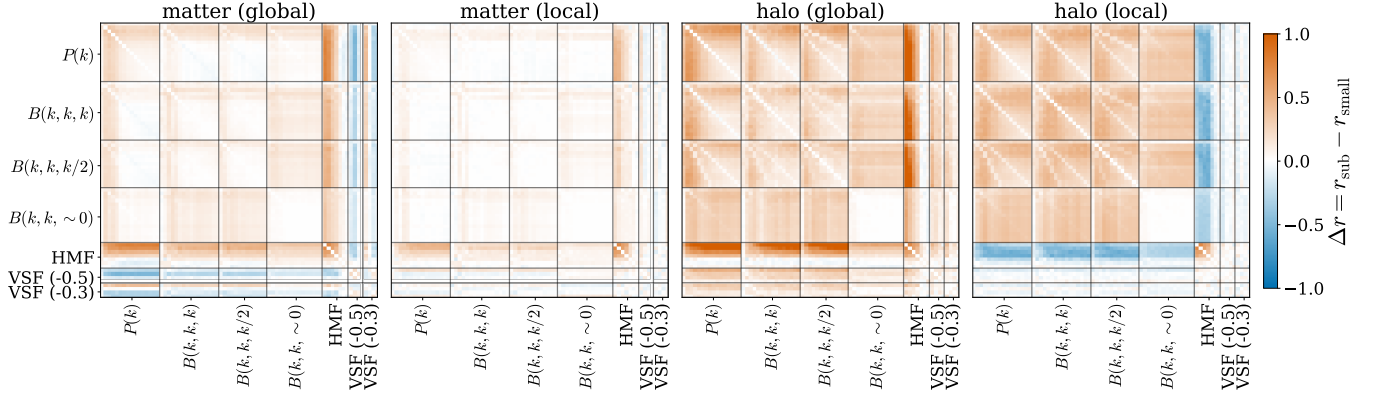


FIG. 4: Difference in correlation coefficient between the sub and small box $\Delta r \equiv r_{\text{sub}} - r_{\text{small}}$. From left to right: matter global, matter local, halo global, halo local.

shows the difference in correlation coefficient between the sub- and small box, $\Delta r \equiv r_{\text{sub}} - r_{\text{small}}$, for the both the matter and halo fields, using local and global mean densities.

For the power spectrum and the bispectrum, the SSC contribution to the cross-covariance is positive in all cases. The amplitude is the smallest in the local matter case, due to the local mean response cancellation dis-

cussed in Sec. III A 1. The local and global halo cases see comparable contributions, also similarly to what was observed for the diagonal term in Sec. III B.

For the HMF, we see a large contribution of SSC to the cross-covariance with other statistics. The effect is positive in all cases, except for the halo local-mean case.

For the VSF, we observe a relatively small contribution of SSC to the cross-covariance with other statistics,

except for the global-mean matter field. This is consistent with the observation of almost negligible SSC in VSF variances in Sec. III A 3, where we discussed that the SSC is low due to the low abundance of voids.

IV. CONCLUSIONS

We study the effect of SSC on the power spectrum, bispectrum, halo mass function, and void size function, as well as on the cross-covariance between them. We consider both the total matter and halo fields. We compare the covariance that includes the SSC (computed using 625 Mpc/h *sub*-boxes of a 5 Gpc/h simulation), to the covariance without (computed using 625 Mpc/h *small* periodic boxes). We now summarize our main results together with additional discussion:

- We see an increasing impact of SSC on smaller scales for the matter and halo power spectrum, reaching a factor of a few compared to the covariance that ignores SSC from $k \approx 0.2h/\text{Mpc}$ and beyond. The exception is the case of the power spectrum referenced to the local-mean density, for which the effect is less than 100%. This is in agreement with previous studies [49–52, 56].
- For the halo mass function, the SSC has little effect on massive halos above $10^{15}M_{\odot}/h$, as they are dominated by shot noise. However, the effect of SSC increases to a factor of 2–3 for lower mass halos, consistently with previous studies [12, 59, 60].
- For the void size function, we found a relatively small SSC effect, due to the low number density of voids compared to that of halos for a given survey volume, and also the low bias. This is an attractive feature of voids, making their covariance simpler to approximate without considering the SSC. In this work we considered spherical voids, however one might consider different void finding algorithms, such as VIDE [105]. However, as long as the void abundance and bias are comparable to those we considered, our general conclusions should hold for other void definitions as well.
- While the *matter* bispectrum receives less SSC contribution, an approximately 50% effect, in good agreement with previous studies [61, 62], the *halo* bispectra shows dependence on the bispectrum configuration and the choice of local or global mean density. Concretely, the level of SSC remains relatively low ($\approx 10\%$ – 100% level) for all three halo bispectrum configurations (equilateral, isosceles, and squeezed) when using the global-mean density and for squeezed bispectrum using the local mean. However, the SSC contribution is a many hundred percent effect for the equilateral and isosceles halo bispectra referenced to the local mean.
- For the cross-covariances, we see non-negligible contribution of SSC, in particular for the halo field statistics. We also observe a negative effect of SSC (or reducing the off-diagonal terms) for HMF \times other statistics in the halo-local mean case, and VSF \times other statistics in some radius bins. This indicates the importance of including the effect of SSC in future joint-statistic analysis.

In summary, our work shows that future cosmological analyses with the power spectrum and higher-order statistics, as well as their joint analysis, should need to carefully consider the effect of SSC. Analyses where many higher-order statistics are combined can have large data vector sizes, which puts pressure on simulation-based methods for the covariance because of the need to have a sufficiently converged covariance matrix that is stable under inversion (which is what is needed in parameter inference analyses). This therefore strongly motivates more simulation-based works like ours here towards a robust understanding of the super-sample covariance and cross-covariance of higher-order statistics.

The level of impact of SSC depends on the box/survey size, halo sample, and redshift – it would be fruitful future work to investigate these dependencies in detail. In particular, we considered a box size of 625 Mpc/h, a value in the ballpark of simulation volumes currently used for covariance estimation (see e.g. Table 1 of [106]). To translate our results to different survey volumes, the scaling of the super-sample covariance with window volume can be determined from Eqn. 3. The responses to δ_b are volume independent, thus the σ_b^2 term is solely responsible for volume dependence. Translating between different window volumes thus simply requires reevaluating σ_b . To gain some intuition, it can be seen that if the power spectrum were taken out of the integral in Eqn. 4, the scaling would be exactly $C_{\text{SSC}} \sim \sigma_b^2 \sim 1/V_W$, just like for other covariance contributions. However, performing the integral of P_{lin} convolved with the window function causes some corrections to this $1/V_W$ scaling.

It would also be interesting to assess the impact of the SSC at the level of final parameter posteriors in simulated likelihood inference analyses for ongoing and future surveys. Furthermore, it would be fruitful to use our simulations to quantify the SSC of other higher-order statistics. To enable such future works, our simulations have been made publicly available at [GitHub](https://github.com/HalfDomeSims/ssc)⁸.

ACKNOWLEDGMENTS

We thank Masahiro Takada, Eiichiro Komatsu, Yue Nan, Uroš Seljak, James Sullivan, and Francisco Villaescusa-Navarro for insightful discussion. This work was supported by JSPS KAKENHI Grants 23K13095 and

⁸ <https://github.com/HalfDomeSims/ssc>

23H00107 (to JL). AEB thanks Kavli IPMU for hosting him during the duration of this project. AB acknowledges support from the Excellence Cluster ORIGINS which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2094-390783311. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, op-

erated under Contract No. DE-AC02-05CH11231 using NERSC award HEP-ERCAP0023125. This work was partly performed at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611. We use `FastPM` [89, 90] to simulate large-scale structure. We use `nbodykit` [91] to compute overdensity fields and power spectra. The bispectrum is computed using the `bskit` package [93], which employs the FFT-based bispectrum estimators of [94, 95]. We use `Pylians3`⁹ [97] to find voids and compute the void size function.

-
- [1] DESI Collaboration *et al.*, The desi experiment part i: Science,targeting, and survey design, 2016, arXiv:1611.00036.
- [2] M. Takada *et al.*, Publications of the Astronomical Society of Japan **66**, R1 (2014).
- [3] LSST Science Collaboration *et al.*, arXiv e-prints , arXiv:0912.0201 (2009), [arXiv:0912.0201].
- [4] R. Laureijs *et al.*, ArXiv e-prints (2011), [arXiv:1110.3193].
- [5] O. Doré *et al.*, arXiv e-prints , arXiv:1412.4872 (2014), [arXiv:1412.4872].
- [6] P. E. Dewdney, P. J. Hall, R. T. Schilizzi and T. J. L. W. Lazio, IEEE Proceedings **97**, 1482 (2009).
- [7] D. Spergel *et al.*, Wide-field infrared survey telescope-astrophysics focused telescope assets wfirst-afta final report, 2013, arXiv:1305.5422.
- [8] M. Takada and B. Jain, Monthly Notices of the Royal Astronomical Society **348**, 897 (2004), [arXiv:astro-ph/0310125].
- [9] E. Sefusatti, M. Crocce, S. Pueblas and R. Scoccimarro, Phys. Rev. D **74**, 023522 (2006), [arXiv:astro-ph/0604505].
- [10] J. Bergé, A. Amara and A. Réfrégier, The Astrophysical Journal **712**, 992 (2010), [arXiv:0909.0529].
- [11] I. Kayo, M. Takada and B. Jain, Monthly Notices of the Royal Astronomical Society **429**, 344 (2013), [arXiv:1207.6322].
- [12] E. Schaan, M. Takada and D. N. Spergel, Phys. Rev. D **90**, 123523 (2014), [arXiv:1406.3330].
- [13] J. Liu *et al.*, Phys. Rev. D **91**, 063507 (2015), [arXiv:1412.0757].
- [14] X. Liu *et al.*, Monthly Notices of the Royal Astronomical Society **450**, 2888 (2015), [arXiv:1412.3683].
- [15] T. Kacprzak *et al.*, ArXiv e-prints (2016), [arXiv:1603.05040].
- [16] H. Shan *et al.*, Monthly Notices of the Royal Astronomical Society **474**, 1116 (2018), [arXiv:1709.07651].
- [17] N. Martinet *et al.*, Monthly Notices of the Royal Astronomical Society **474**, 712 (2018), [arXiv:1709.07678].
- [18] J. Liu and M. S. Madhavacheril, Phys. Rev. D **99**, 083508 (2019).
- [19] Z. Li, J. Liu, J. M. Z. Matilla and W. R. Coulton, Phys. Rev. D **99**, 063527 (2019), [arXiv:1810.01781].
- [20] C. D. Kreisch *et al.*, Monthly Notices of the Royal Astronomical Society **488**, 4413 (2019), [arXiv:1808.07464].
- [21] W. R. Coulton, J. Liu, M. S. Madhavacheril, V. Böhm and D. N. Spergel, Journal of Cosmology and Astroparticle Physics **5**, 043 (2019), [arXiv:1810.02374].
- [22] M. Sahlén, Physical Review D **99** (2019).
- [23] G. A. Marques *et al.*, Journal of Cosmology and Astroparticle Physics **2019**, 019 (2019), [arXiv:1812.08206].
- [24] V. Ajani *et al.*, Constraining neutrino masses with weak-lensing starlet peak counts, 2020, arXiv:2001.10993.
- [25] C. Hahn, F. Villaescusa-Navarro, E. Castorina and R. Scoccimarro, Journal of Cosmology and Astroparticle Physics **2020**, 040–040 (2020).
- [26] C. Hahn and F. Villaescusa-Navarro, Constraining m_ν with the bispectrum ii: The total information content of the galaxy bispectrum, 2020, arXiv:2012.02200.
- [27] J.-P. Dai, L. Verde and J.-Q. Xia, Journal of Cosmology and Astroparticle Physics **2020**, 007 (2020), [arXiv:2002.09904].
- [28] C. Uhlemann, O. Friedrich, F. Villaescusa-Navarro, A. Banerjee and S. Codis, Monthly Notices of the Royal Astronomical Society **495**, 4006 (2020), [arXiv:1911.11158].
- [29] E. Allys *et al.*, Phys. Rev. D **102**, 103506 (2020), [arXiv:2006.06298].
- [30] D. Gualdi, S. Novell, H. Gil-Marín and L. Verde, arXiv e-prints , arXiv:2009.02290 (2020), [arXiv:2009.02290].
- [31] J. Harnois-Déraps *et al.*, arXiv e-prints , arXiv:2012.02777 (2020), [arXiv:2012.02777].
- [32] A. Banerjee and T. Abel, Monthly Notices of the Royal Astronomical Society **500**, 5479 (2021), [arXiv:2007.13342].
- [33] E. Massara, F. Villaescusa-Navarro, S. Ho, N. Dalal and D. N. Spergel, Phys. Rev. Lett. **126**, 011301 (2021), [arXiv:2001.11024].
- [34] S. Cheng, Y.-S. Ting, B. Ménard and J. Bruna, Monthly Notices of the Royal Astronomical Society **499**, 5902 (2020), [arXiv:2006.08561].
- [35] S. Cheng and B. Ménard, Monthly Notices of the Royal Astronomical Society **507**, 1012 (2021), [arXiv:2103.09247].
- [36] A. E. Bayer *et al.*, ApJ **919**, 24 (2021), [arXiv:2102.05049].
- [37] C. D. Kreisch *et al.*, ApJ **935**, 100 (2022), [arXiv:2107.02304].

⁹ <https://pylians3.readthedocs.io/>

- [38] A. E. Bayer, A. Banerjee and U. Seljak, *Phys. Rev. D* **105**, 123510 (2022), [arXiv:2108.04215].
- [39] G. Valogiannis and C. Dvorkin, *arXiv e-prints*, arXiv:2108.07821 (2021), [arXiv:2108.07821].
- [40] G. Valogiannis and C. Dvorkin, *arXiv e-prints*, arXiv:2204.13717 (2022), [arXiv:2204.13717].
- [41] M. Eickenberg *et al.*, *arXiv e-prints*, arXiv:2204.07646 (2022), [arXiv:2204.07646].
- [42] N. Hamaus *et al.*, *J. Cosmology Astropart. Phys.* **2020**, 023 (2020), [arXiv:2007.07895].
- [43] A. Dvornik *et al.*, *arXiv e-prints*, arXiv:2210.03110 (2022), [arXiv:2210.03110].
- [44] S. Pandey *et al.*, *Phys. Rev. D* **106**, 043520 (2022), [arXiv:2105.13545].
- [45] DES Collaboration *et al.*, *arXiv e-prints*, arXiv:2207.05766 (2022), [arXiv:2207.05766].
- [46] F. Villaescusa-Navarro *et al.*, *The Astrophysical Journal* **250**, 2 (2020), [arXiv:1909.05273].
- [47] E. Sirko, *ApJ* **634**, 728 (2005), [arXiv:astro-ph/0503106].
- [48] N. Y. Gnedin, A. V. Kravtsov and D. H. Rudd, *ApJS* **194**, 46 (2011), [arXiv:1104.1428].
- [49] A. J. S. Hamilton, C. D. Rimes and R. Scoccimarro, *Monthly Notices of the Royal Astronomical Society* **371**, 1188 (2006), [arXiv:astro-ph/0511416].
- [50] M. Takada and W. Hu, *Phys. Rev. D* **87**, 123504 (2013), [arXiv:1302.6994].
- [51] R. de Putter, C. Wagner, O. Mena, L. Verde and W. J. Percival, *Journal of Cosmology and Astroparticle Physics* **2012**, 019 (2012), [arXiv:1111.6596].
- [52] R. Takahashi *et al.*, *ApJ* **700**, 479 (2009), [arXiv:0902.0371].
- [53] T. Baldauf, U. Seljak, L. Senatore and M. Zaldarriaga, *Journal of Cosmology and Astroparticle Physics* **2011**, 031 (2011), [arXiv:1106.5507].
- [54] A. Barreira and F. Schmidt, *Journal of Cosmology and Astroparticle Physics* **2017**, 053 (2017), [arXiv:1703.09212].
- [55] A. Barreira and F. Schmidt, *Journal of Cosmology and Astroparticle Physics* **2017**, 051 (2017), [arXiv:1705.01092].
- [56] Y. Li, W. Hu and M. Takada, *Phys. Rev. D* **89**, 083519 (2014), [arXiv:1401.0385].
- [57] C. Wagner, F. Schmidt, C. T. Chiang and E. Komatsu, *MNRAS* **448**, L11 (2015), [arXiv:1409.6294].
- [58] R. Terasawa, R. Takahashi, T. Nishimichi and M. Takada, *Phys. Rev. D* **106**, 083504 (2022), [arXiv:2205.10339].
- [59] W. Hu and A. V. Kravtsov, *ApJ* **584**, 702 (2003), [arXiv:astro-ph/0203169].
- [60] O. H. E. Philcox, D. N. Spergel and F. Villaescusa-Navarro, *Phys. Rev. D* **101**, 123520 (2020), [arXiv:2004.09515].
- [61] K. C. Chan, A. Moradinezhad Dizgah and J. Noreña, *Phys. Rev. D* **97**, 043532 (2018), [arXiv:1709.02473].
- [62] A. Barreira, *Journal of Cosmology and Astroparticle Physics* **2019**, 008 (2019), [arXiv:1901.01243].
- [63] W. R. Coulton *et al.*, *arXiv e-prints*, arXiv:2206.15450 (2022), [arXiv:2206.15450].
- [64] C. Uhlemann *et al.*, *arXiv e-prints*, arXiv:2210.07819 (2022), [arXiv:2210.07819].
- [65] K. Akitsu, M. Takada and Y. Li, *Phys. Rev. D* **95**, 083522 (2017), [arXiv:1611.04723].
- [66] K. Akitsu and M. Takada, *Phys. Rev. D* **97**, 063527 (2018), [arXiv:1711.00012].
- [67] Y. Li, M. Schmittfull and U. Seljak, *Journal of Cosmology and Astroparticle Physics* **2018**, 022 (2018), [arXiv:1711.00018].
- [68] A. Barreira, E. Krause and F. Schmidt, *Journal of Cosmology and Astroparticle Physics* **2018**, 015 (2018), [arXiv:1711.07467].
- [69] I. Mohammed and U. Seljak, *MNRAS* **445**, 3382 (2014), [arXiv:1407.0060].
- [70] A. Barreira *et al.*, *MNRAS* **488**, 2079 (2019), [arXiv:1904.02070].
- [71] A. Halder and A. Barreira, *arXiv e-prints*, arXiv:2201.05607 (2022), [arXiv:2201.05607].
- [72] E. Castorina and A. Moradinezhad Dizgah, *Journal of Cosmology and Astroparticle Physics* **2020**, 007 (2020), [arXiv:2005.14677].
- [73] F. Lacasa and J. Grain, *A&A* **624**, A61 (2019), [arXiv:1809.05437].
- [74] S. Gouyou Beauchamps *et al.*, *A&A* **659**, A128 (2022), [arXiv:2109.02308].
- [75] F. Lacasa *et al.*, *arXiv e-prints*, arXiv:2209.14421 (2022), [arXiv:2209.14421].
- [76] E. Platen, R. van de Weygaert and B. J. T. Jones, *Monthly Notices of the Royal Astronomical Society* **387**, 128–136 (2008).
- [77] E. G. P. Bos, R. van de Weygaert, K. Dolag and V. Pettorino, *Monthly Notices of the Royal Astronomical Society* **426**, 440–461 (2012).
- [78] P. M. Sutter, G. Lavaux, B. D. Wandelt and D. H. Weinberg, *The Astrophysical Journal* **761**, 187 (2012).
- [79] E. Jennings, Y. Li and W. Hu, *Monthly Notices of the Royal Astronomical Society* **434**, 2167–2181 (2013).
- [80] A. Pisani *et al.*, *Phys. Rev. D* **92**, 083531 (2015).
- [81] E. Paillas *et al.*, *Monthly Notices of the Royal Astronomical Society* **470**, 4434 (2017), [arXiv:1609.00101].
- [82] S. Contarini *et al.*, *Monthly Notices of the Royal Astronomical Society* **488**, 3526–3540 (2019).
- [83] G. Verza, A. Pisani, C. Carbone, N. Hamaus and L. Guzzo, *Journal of Cosmology and Astroparticle Physics* **2019**, 040–040 (2019).
- [84] N. Hamaus *et al.*, *A&A* **658**, A20 (2022), [arXiv:2108.10347].
- [85] S. Contarini *et al.*, *arXiv e-prints*, arXiv:2205.11525 (2022), [arXiv:2205.11525].
- [86] M. Bonici *et al.*, *arXiv e-prints*, arXiv:2206.14211 (2022), [arXiv:2206.14211].
- [87] D. Jamieson and M. Loverde, *Phys. Rev. D* **100**, 123528 (2019), [arXiv:1909.05313].
- [88] K. C. Chan, Y. Li, M. Biagetti and N. Hamaus, *ApJ* **889**, 89 (2020), [arXiv:1909.03736].
- [89] Y. Feng, M.-Y. Chu, U. Seljak and P. McDonald, *Monthly Notices of the Royal Astronomical Society* **463**, 2273 (2016), [arXiv:1603.00476].
- [90] A. E. Bayer, A. Banerjee and Y. Feng, *Journal of Cosmology and Astroparticle Physics* **2021**, 016 (2021).
- [91] N. Hand *et al.*, *The Astronomical Journal* **156**, 160 (2018).
- [92] Y. P. Jing, *ApJ* **620**, 559 (2005), [arXiv:astro-ph/0409240].
- [93] S. Foreman, W. Coulton, F. Villaescusa-Navarro and A. Barreira, *Monthly Notices of the Royal Astronomical Society* **498**, 2887 (2020), [arXiv:1910.03597].
- [94] R. Scoccimarro, *ApJ* **544**, 597 (2000), [arXiv:astro-ph/0004086].

- [95] E. Sefusatti, M. Crocce, R. Scoccimarro and H. M. P. Couchman, *MNRAS* **460**, 3624 (2016), [arXiv:1512.07295].
- [96] A. Banerjee and N. Dalal, *Journal of Cosmology and Astroparticle Physics* **11**, 015 (2016), [arXiv:1606.06167].
- [97] F. Villaescusa-Navarro, *Pylians: Python libraries for the analysis of numerical simulations*, *Astrophysics Source Code Library*, record ascl:1811.008, 2018, arXiv:1811.008.
- [98] A. Barreira, G. Cabass, F. Schmidt, A. Pillepich and D. Nelson, *JCAP* **12**, 013 (2020), [arXiv:2006.09368].
- [99] A. S. Schmidt, S. D. M. White, F. Schmidt and J. Stücker, *MNRAS* **479**, 162 (2018), [arXiv:1803.03274].
- [100] S. Masaki, T. Nishimichi and M. Takada, *MNRAS* **496**, 483 (2020), [arXiv:2003.10052].
- [101] K. Akitsu, Y. Li and T. Okumura, *Journal of Cosmology and Astroparticle Physics* **2021**, 041 (2021), [arXiv:2011.06584].
- [102] S. Saito *et al.*, *Phys. Rev. D* **90**, 123522 (2014), [arXiv:1405.1447].
- [103] A. Barreira, E. Krause and F. Schmidt, *JCAP* **10**, 053 (2018), [arXiv:1807.04266].
- [104] T. Baldauf, U. Seljak, L. Senatore and M. Zaldarriaga, *Journal of Cosmology and Astroparticle Physics* **2016**, 007 (2016), [arXiv:1511.01465].
- [105] P. M. Sutter *et al.*, *Vide: The void identification and examination toolkit*, 2014, arXiv:1406.1191.
- [106] T. Kacprzak, J. Fluri, A. Schneider, A. Refregier and J. Stadel, *JCAP* **02**, 050 (2023), [arXiv:2209.04662].