PHYSICAL REVIEW D **93,** 063507 (2016)

Separate universe consistency relation and calibration of halo bias

Yin Li,[1,2] Wayne Hu,[3] and Masahiro Takada[2]
[1]*Berkeley Center for Cosmological Physics, Department of Physics and Lawrence Berkeley National Laboratory,*
University of California, Berkeley, California 94720, USA
[2]*Kavli Institute for the Physics and Mathematics of the Universe (WPI), UTIAS,*
The University of Tokyo, Chiba 277-8583, Japan
[3]*Kavli Institute for Cosmological Physics, Department of Astronomy & Astrophysics,*
Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA
(Received 6 November 2015; published 10 March 2016)

The linear halo bias is the response of the dark matter halo number density to a long-wavelength fluctuation in the dark matter density. Using abundance matching between separate universe simulations which absorb the latter into a change in the background, we test the consistency relation between the change in a one-point function, the halo mass function, and a two-point function, the halo-matter cross-correlation in the long-wavelength limit. We find excellent agreement between the two at the 1%–2% level for average halo biases between $1 \lesssim \bar{b}_1 \lesssim 4$ and no statistically significant deviations at the 4%–5% level out to $\bar{b}_1 \approx 8$. The halo bias inferred assuming instead a universal mass function is significantly different and inaccurate at the 10% level or more. The separate universe technique provides a way of calibrating the linear halo bias efficiently for even highly biased rare halos in the Λ cold dark matter model. Observational violation of the consistency relation would indicate new physics, e.g. in the dark matter, dark energy, or primordial non-Gaussianity sectors.

DOI: 10.1103/PhysRevD.93.063507

I. INTRODUCTION

Dark matter halos, which host observable galaxies and galaxy clusters, are biased tracers of the underlying dark matter density field of the large-scale structure of the Universe [1]. Therefore, understanding the mass, redshift, and scale dependence of halo bias is important for extracting cosmological information, on e.g. dark energy, massive neutrinos, and the statistics of the primordial perturbations [2–5], from ongoing and future wide-area galaxy surveys such as the Dark Energy Survey [6]; Dark Energy Spectrograph Instrument [7]; the Subaru Hyper Suprime-Cam/Prime Focus Spectrograph Survey [8,9]; and ultimately Large Synoptic Survey Telescope [10], Euclid [11], and Wide-Field Infrared Survey Telescope [12].

Whereas near the nonlinear scale a single definition of the halo bias does not suffice due to a host of effects that influence the clustering of halos ([13,14], see Ref. [15] for a recent review), the linear response of dark matter halos to the dark matter density field is much better understood. In particular, under the peak-background split approach [16], the halo bias can be modeled through the halo mass function. Under the assumption that it is a universal function of the variance of the dark matter density field, this provides a simple expression for the halo bias [17–23].

More directly, the halo bias can be measured from the cross-correlation of halos with the dark matter distribution in the large-scale limit—the clustering bias [2,22,24–26]. Previous works [22,27–29] have shown that the universal mass function bias approximates the clustering bias, at least at the 10% level, but were inconclusive beyond this level partly because the two biases were not always self-consistently estimated from the mass functions and the clustering correlations in same simulations. References [27,28] claimed evidence for inconsistency near this level. Consistency between the bias and the mass function is important for dark energy tests that utilize both the abundance and clustering of halos (e.g. Refs. [30,31]).

In this paper, we consider a related but alternative way of understanding and calibrating the linear halo bias. As in the peak-background split approach, the linear halo bias is modeled as the response of the number density of halos, or halo mass function, to a change in the background dark matter density field. Unlike the universal mass function implementation, this linearized change in the background is modeled throughout the whole past temporal history of the density fluctuation using the separate universe simulation approach developed in Refs. [32,33] (see also Refs. [34–37]). The induced change in the mass function yields the response of halo number densities to the background dark matter density, or "response bias". Defined in this way, the response bias is quite general in the sense that it does not assume the universality of the halo mass function and it includes all the effects of mergers and mass accretion that are correlated with the background density mode. It can also be easily extended to baryonic and galaxy formation effects using simulations that include them.

We furthermore use a consistent set of simulations to address whether the response bias matches the clustering bias and also compare the results with the fitting formula of clustering bias in Ref. [22]. Observational violation of this consistency relation would indicate new physics where the dark matter, dark energy, primordial non-Gaussianity, or other effects provide alternate means of producing a mass function response to the dark matter density fluctuation.

The outline of this paper is as follows. In Sec. II, we define response bias and clustering bias in a ΛCDM cosmology, give a brief review of the separate universe simulation, and then propose the abundance matching method for calibrating the response bias. We present results and tests of the consistency of response and clustering biases in Sec. III. We discuss the results in Sec. IV. In the Appendixes, we present robustness checks on the bias results and compare them with inferences from the universal mass function assumption.

II. HALO BIAS

A. Halo response vs clustering bias

Dark matter halos of a given mass M are biased tracers of the underlying dark matter density field. On large scales where the dark matter density fluctuations $\delta = \delta\rho_m/\rho_m$ are still in the linear regime $|\delta| \ll 1$, we can think of biasing as the linearized response of the halo number density to changes in the dark matter density, implicitly of some linear wave number k,

$$b_1(M) \equiv \frac{\mathrm{d}\delta_\mathrm{h}}{\mathrm{d}\delta} = \frac{\mathrm{d}\ln n_{\ln M}}{\mathrm{d}\delta}, \qquad (1)$$

where the mass function $n_{\ln M}(M)$ is the differential number density of halos per logarithmic mass interval. We will call this quantity the response bias.

This definition of the linear density bias is quite general as it includes any effect that is correlated with the change in δ, as designated by the total derivative in Eq. (1). For example, the halo density in a given mass range can change due to mass accretion, minor mergers, and major mergers. A change in δ could also be correlated with changes in the dark energy or massive neutrino density that could likewise influence halo numbers through their impact on the history of structure formation, e.g. the halo accretion and merger history [4,38–41]. Intrinsic non-Gaussian correlation between long-wavelength initial curvature fluctuations and small-scale power in the density field can also change the response in a scale-dependent way [2].

On the other hand, we can define the linear density bias directly via cross-correlation of halos with the cold dark matter distribution,

$$b_1(M) = \lim_{k\to 0} \frac{P_{\mathrm{h}\delta}(k;M)}{P_{\delta\delta}(k)}, \qquad (2)$$

where

$$\langle \delta_\mathrm{h}^*(\mathbf{k})\delta(\mathbf{k}')\rangle = (2\pi)^3\delta(\mathbf{k}-\mathbf{k}')P_{\mathrm{h}\delta}(k),$$
$$\langle \delta^*(\mathbf{k})\delta(\mathbf{k}')\rangle = (2\pi)^3\delta(\mathbf{k}-\mathbf{k}')P_{\delta\delta}(k). \qquad (3)$$

We will call this form for b_1 the "clustering bias". Equations (1) and (2) characterize the same physical quantity since the mass function response can come from any effect that is correlated with δ. Uncorrelated changes in the halo density, e.g. from stochasticity in the bias, can affect the autocorrelation of halos but by definition do not change the cross-correlation.

In this paper, we focus on the most fundamental response, that of the direct influence of the long-wavelength dark matter density fluctuation on the halo number density in the Lambda cold dark matter (ΛCDM) cosmology with Gaussian initial conditions. The critical assumption that we seek to test is the extent to which this local number density depends only on the local mean dark matter density. In this case, the equivalence of Eqs. (1) and (2) forms a consistency relation between the change in a one-point function, the halo mass function, and a two-point function, the halo-matter cross-correlation in the long-wavelength limit. Validation of this consistency relation would allow two alternate means of calibrating the bias in simulations. Observational tests of this consistency can in principle uncover new physics beyond ΛCDM where the dark matter, dark energy, or primordial non-Gaussianity provides alternate means of producing a mass function response to δ.

Specifically, as detailed in the next section, we will use separate universe (SU) simulations to test this consistency relation. In this approach, the fluctuation in the dark matter density is characterized by changes to cosmological parameters or spatially constant background densities to match the mean fluctuation $\delta_\mathrm{b} = \delta$. This should be compared with the well-known peak-background or universal mass function approach to quantifying b_1 through the mass function $n_{\ln M}$. Here, it is assumed that the mass function can be described as a universal function of the peak height $\nu = \delta_\mathrm{c}/\sigma(M)$, the ratio of the collapse threshold of halos δ_c relative to the rms linear density fluctuations in a radius that encloses the mass M at the background density $\sigma(M)$. Changing the collapse threshold via shifting the background $\delta_\mathrm{c} \to \delta_\mathrm{c} - \delta_\mathrm{b}$ then changes the number density of halos, providing an approximation for b_1 through Eq. (1).

Both the separate universe and the universal mass function approach seek to characterize the response bias through replacing δ with a change in the background δ_b. However, the former does not rely on the existence of a universal mass function or the idea of a strict collapse threshold of dark matter halos. All types of responses of the mass function to the background, including the highly nonlinear processes of the merger history of halos, etc., are automatically included in the simulations. Although we only test N-body effects and dark matter halos here, this in principle applies to baryonic effects and galaxy tracers through simulations that incorporate them. We present the separate universe approach in the main text and its comparison to the universal mass function approach in Sec. II B.

B. Separate universe technique

To calibrate numerically the response of the halo mass function to a background mode, we use the SU simulation technique [32–35]. We follow Ref. [32] and refer the reader there for details.

In summary, the long-wavelength density fluctuation $\delta_{\rm b}$ is absorbed into the background density $\bar{\rho}_{mW}$ of a separate universe,

$$\bar{\rho}_{mW} = \bar{\rho}_m(1+\delta_{\rm b}), \tag{4}$$

where the quantities with subscript "W" denote the quantities in a separate universe.

The separate universe consequently has a different expansion history, and accordingly we need to change cosmological parameters for the flat ΛCDM cosmology, to the first order of $\delta_{\rm b}$, as

$$\frac{\delta h}{h} \equiv \frac{H_{0W}-H_0}{H_0} = -\frac{5\Omega_m}{6}\frac{\delta_{\rm b}}{D}, \tag{5}$$

where the linear growth rate is normalized as $\lim_{a\to 0} D = a$. Since $\delta_{\rm b}/D$ is independent of time, the SU is characterized by a simple constant shift in parameters. Similarly, the other parameters need to be changed to

$$\frac{\delta\Omega_{\rm m}}{\Omega_{\rm m}} = \frac{\delta\Omega_\Lambda}{\Omega_\Lambda} = -\delta\Omega_K = -2\frac{\delta h}{h}. \tag{6}$$

Thus, in the presence of a $\delta_{\rm b} > 0$, the properties of smaller-scale structures including the abundance of halos experience the accelerated growth of a closed universe.

Finally, the separate universes have to be compared at the same time, which corresponds to a different value of the scale factor,

$$a_W \simeq a\left(1-\frac{\delta_{\rm b}}{3}\right). \tag{7}$$

Because of this difference, the SU simulations are most naturally set up as a Lagrangian approach where the simulation volumes match in their comoving rather than physical volume (cf. Ref. [32] for an alternative method that matches physical volumes at a specific time). This splits the response of the mass function into two pieces. The first corresponds to the change due to the growth of structures, including processes such as shell crossing, mass accretion, and the merger of halos,

$$b_1^{\rm L}(M) \equiv \frac{\partial \ln n^{\rm L}_{\ln M}}{\partial \delta_{\rm b}} = \left.\frac{\partial \ln n_{\ln M}}{\partial \delta_{\rm b}}\right|_{V_{\rm c}}, \tag{8}$$

where $|_{V_{\rm c}}$ denotes the separate universe response at fixed comoving volume. "L" superscripts refer to that fact that this generalizes the concept of Lagrangian bias to the whole volume rather than individual N-body particles or halos. The second is due to the change in the physical volume and hence physical densities due to Eq. (7) or

$$\frac{\partial \ln a_W^3}{\partial \delta_{\rm b}} = -1. \tag{9}$$

The sum of these two effects is then the Eulerian response bias

$$b_1(M) \equiv b_1^{\rm L}(M) + 1. \tag{10}$$

It is important to note that this is a definition and hence is exact, rather than an approximation that relies on halo number conservation. This is the growth-dilation derivative technique developed in Ref. [33] as applied to the mass function response. Calibrating the response bias with separate universe simulations therefore amounts to determining the derivative of the Lagrangian mass function $n^{\rm L}_{\ln M}$ with respect to the background density fluctuation $\delta_{\rm b}$ in Eq. (8).

C. Abundance matching

Much of the response of the Lagrangian mass function $n^{\rm L}_{\ln M}$ to $\delta_{\rm b}$ comes from small changes in the mass of individual halos rather than a change in the net number of halos in the volume. Therefore, measuring the response by binning halos into finite mass ranges is very inefficient (see Appendix A 2), since the mass change associated with a small $\delta_{\rm b}$ only shifts the mass of halos near bin edges.

Given the pairs of SU simulations with the same Gaussian random fields, in principle the same halos could be identified in each and the response calculated from the average change in the mass. However, in practice, the identity of halos can be easily affected by mergers. Even for those halos for which a one-to-one correspondence exists, their change in mass is not uniquely determined by M due to differences in the environment around halos of the same M which introduces scatter into the mapping. This suggests that we need to find a statistic that does not rely on a one-to-one correspondence between SU halos in mass of which the ensemble average recovers the desired response in numbers.

Abundance matching of the cumulative number density or mass function of halos above a given mass threshold $M_{\rm th}$ provides such a statistic [42,43]. Defining

$$n(M_{\rm th};\delta_{\rm b}) \equiv \int_{M_{\rm th}}^{\infty} \frac{{\rm d}M}{M} n^{\rm L}_{\ln M}(M;\delta_{\rm b}), \tag{11}$$

we change the threshold $M_{\rm th}(\delta_{\rm b})$ to keep the cumulative number density in the comoving volume fixed when varying $\delta_{\rm b}$,

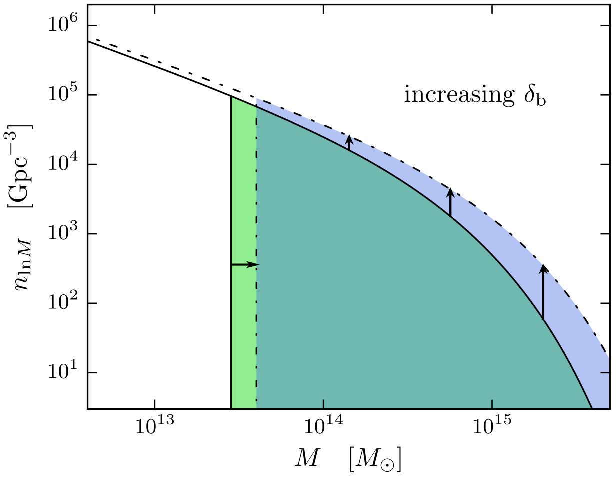

FIG. 1. Abundance matching relates the number density weighted bias above threshold mass M_th to the shift of that threshold. The halo abundance above M_th grows in proportion to the bias function when increasing δ_b, which we can compensate by moving M_th accordingly. This figure graphically illustrates Eq. (14).

$$\frac{\mathrm{d}n(M_\mathrm{th};\delta_\mathrm{b})}{\mathrm{d}\delta_\mathrm{b}} = 0. \tag{12}$$

We use $(\ldots;p)$ to denote a quantity for which we omit the parameter p where no confusion should arise.

Abundance matching balances two effects to keep the number density the same, as illustrated in Fig. 1. The first is the boundary effect of halos moving across a threshold shifted by s due to the change in $\mathrm{d}\delta_\mathrm{b}$,

$$\mathrm{d}\ln M_\mathrm{th} \equiv s(M_\mathrm{th})\mathrm{d}\delta_\mathrm{b}. \tag{13}$$

The second is the integrated change in the mass function itself, which is the effect we want to extract for estimating the response bias. Abundance matching sets these to be equal,

$$n^\mathrm{L}_{\ln M}(M_\mathrm{th})s(M_\mathrm{th}) = \int_{M_\mathrm{th}}^\infty \frac{\mathrm{d}M}{M}\frac{\partial n^\mathrm{L}_{\ln M}}{\partial\delta_\mathrm{b}}, \tag{14}$$

which also follows algebraically from Eqs. (11) and (12).

Measuring the mass shift s associated with matching the abundance therefore provides a way of estimating the average response bias above threshold,

$$\begin{aligned}\bar{b}^\mathrm{L}_1(M_\mathrm{th};\infty) &\equiv \frac{1}{n(M_\mathrm{th})}\int_{M_\mathrm{th}}^\infty \frac{\mathrm{d}M}{M} b^\mathrm{L}_1 n^\mathrm{L}_{\ln M} \\ &= \frac{1}{n(M_\mathrm{th})}\int_{M_\mathrm{th}}^\infty \frac{\mathrm{d}M}{M}\frac{\partial \ln n^\mathrm{L}_{\ln M}}{\partial\delta_\mathrm{b}} n^\mathrm{L}_{\ln M} \\ &= \frac{n^\mathrm{L}_{\ln M}(M_\mathrm{th})s(M_\mathrm{th})}{n(M_\mathrm{th})}.\end{aligned} \tag{15}$$

We emphasize that such an estimation of the response bias does not rely on any assumption on the universality of the halo mass function.

Note that measuring this quantity also defines the average bias in a finite mass bin,

$$\begin{aligned}\bar{b}^\mathrm{L}_1(M_1,M_2) &\equiv \frac{\int_{M_1}^{M_2}\mathrm{d}\ln M b^\mathrm{L}_1 n^\mathrm{L}_{\ln M}}{\int_{M_1}^{M_2}\mathrm{d}\ln M n^\mathrm{L}_{\ln M}} \\ &= \frac{n^\mathrm{L}_{\ln M}(M_1)s(M_1) - n^\mathrm{L}_{\ln M}(M_2)s(M_2)}{n(M_1)-n(M_2)}.\end{aligned} \tag{16}$$

In the limit that $M_2 \to M_1$ from above, this quantity is simply the Lagrangian bias or mass function response itself $b^\mathrm{L}_1(M_1)$ and is equivalent to replacing the formal definition in terms of derivatives,

$$b^\mathrm{L}_1(M) = -\frac{\partial s}{\partial\ln M} - s\frac{\partial\ln n^\mathrm{L}_{\ln M}}{\partial\ln M}. \tag{17}$$

with a finite difference approximation. Since the clustering bias also must be explicitly estimated from finite mass binning, it is in fact Eq. (16) that should be directly compared with it. As a shorthand convention, we plot the average bias in a bin as

$$b^\mathrm{L}_1(M) \approx \bar{b}^\mathrm{L}_1(M_1;M_2) \tag{18}$$

using the average mass of halos in the bin

$$M \equiv \frac{\int_{M_1}^{M_2}\mathrm{d}\ln M M n^\mathrm{L}_{\ln M}}{\int_{M_1}^{M_2}\mathrm{d}\ln M n^\mathrm{L}_{\ln M}}. \tag{19}$$

Following our notational convention, we also take

$$\bar{b}^\mathrm{L}_1(M) = \bar{b}^\mathrm{L}_1(M;\infty) \tag{20}$$

when no confusion will arise.

To measure these response bias quantities directly, we need the estimators of the cumulative mass function $n(M)$, the threshold mass shift $s(M)$, and the differential mass function $n^\mathrm{L}_{\ln M}(M)$ in the Lagrangian volume. We consider their explicit construction in Sec. III C.

III. METHODOLOGY AND RESULTS

In this section, we describe the methodology to calibrate the model ingredients needed to estimate response and clustering halo biases using suites of simulations in the fiducial cosmology and its separate universe pairs. We then show the main results that establish their consistency.

A. Simulations

We simulate the fiducial ΛCDM cosmology specified in Table I. Each pair of separate universe simulations has the

TABLE I. Parameters of baseline flat ΛCDM model [5].

$\Omega_{\rm m}$	$\Omega_{\rm b}$	h	$n_{\rm s}$	σ_8
0.310	0.04508	0.703	0.964	0.785

same realizations of the initial Gaussian random density field, in order to reduce the sample variance in the change of the mass function.

We set up the initial conditions using `CAMB` [44,45], and `2LPTIC` [46], with 1024^3 particles at $a_{\rm i} = 0.02$. We then employ `L-Gadget2` [47] with a 2048^3 TreePM grid to produce the simulations. For calibrating the response bias, we employ $N_{\rm sim} = 32$ simulations with $V_{\rm c} = (500\ {\rm Mpc}/0.703)^3$ for each of 3 $\delta_{\rm b} = 0, \pm 0.01$ at $z = 0$. The separate universe variations all have the same comoving volume $V_{\rm c}$ in ${\rm Mpc}^3$ (see Sec. II B).

The $\delta_{\rm b} = \pm 0.01$ pairs are used in abundance matching, and the $\delta_{\rm b} = 0$ simulations are used to calibrate the mass function (see Sec. II B). Since measuring the clustering bias for rare high mass halos requires more numbers than the response bias, we supplement these with $N_{\rm sim} = 25$ simulations with $V_{\rm c} = (1\ {\rm Gpc}/0.703)^3$ fiducial simulations at $\delta_{\rm b} = 0$. The particle masses for the two box sizes are $1.4 \times 10^{10} M_\odot$ and $1.1 \times 10^{11} M_\odot$ respectively, which limits the minimum halo mass that we can robustly identify as we shall now discuss.

B. Halo finding and catalog

While the choices made in halo finding can affect the mass function and bias results, for tests of the correspondence between the SU response bias and clustering bias, what is important is that we apply the *same* halo finding technique to each. In practice, we use an algorithm similar to that in Ref. [48] to identify halos as spherical overdense regions centered around local density peaks as we now describe.

We first locate local density maxima by assigning particles to a 1024^3 grid, using the nearest-grid-point scheme. We find local density maximum grid points that are denser than their six immediate neighbors. Starting at the center of mass associated with each local maxima, we grow a halo until the enclosed mass reaches an effective overdensity of

$$\Delta_W = \frac{\Delta}{1+\delta_{\rm b}} = \frac{200}{1+\delta_{\rm b}} \tag{21}$$

defining a trial radius $r_{\rm tr}$. The $1 + \delta_{\rm b}$ factor makes sure that the spherical overdensity is 200 times the global mean matter density. We refine the center of the halo by locating the center of mass iteratively in shrinking radii from $r_{\rm tr}/3$ to $r_{\rm tr}/15$ or until only 20 particles remain. We then regrow the halo around this center until the overdensity criteria Eq. (21) is exactly satisfied, with sub-particle resolution. To achieve this, we assume the mass of the last particle is uniformly distributed in a spherical mass shell lying between the last two particles and interpolate to the exact radius r. The mass of all particles within r gives the halo mass M.

Each simulation provides a catalog of the positions and masses of these halos. We ignore halos with < 100 particles when creating the catalog. We retain halos with 100–400 particles to eliminate edge effects in the mass function determinations below but only report results for halos with ≥ 400 particles [48] (see also Appendix A 3). To remove subhalos in the catalog, starting from the most massive halos, we compare pairs of halos in descending order in mass and discard the smaller halo of the pair if the center of one resides in the other.

C. Halo mass functions and mass shift

As discussed in Sec. II C, we measure the response bias through an abundance matching technique to reduce the shot noise in its determination. This technique requires us to estimate the cumulative and differential mass function in the fiducial model as well as the mass shift from matching the $\pm\delta_{\rm b}$ pairs of SU simulations. We show here that these can be robustly estimated without binning the halo catalogs in mass. Coarse binning would miss the small changes in mass due to $\delta_{\rm b}$, whereas fine binning would be subject to severe shot noise.

We start by combining the halo catalogs of all $N_{\rm sim}$ simulations of the same $\delta_{\rm b}$ and $V_{\rm c}$ into a single halo catalog ordered from highest to lowest mass $i > j$ for $M_i < M_j$ with total number $N_{\rm tot}$. We construct a table for the cumulative abundance above a given mass object in the catalog as

$$\begin{aligned} \ln \mathbf{M} &= [\ln M_1, \ldots \ln M_{N_{\rm tot}}]^{\rm T}, \\ \mathbf{n} &= \frac{[1/2, \ldots, N_{\rm tot} - 1/2]^{\rm T}}{N_{\rm sim} V_{\rm c}}, \end{aligned} \tag{22}$$

which we will denote as the data vector $\mathbf{n}(\ln \mathbf{M}; \delta_{\rm b}, V_{\rm c})$. Here, we count the halo with mass M_i as one-half above and one-half below M_i due to discreteness and recall $V_{\rm c}$ is the comoving volume in ${\rm Mpc}^3$ and is fixed in the SU simulations when varying $\delta_{\rm b}$.

Next, we construct a data vector of mass shifts by abundance matching. Since we have rank ordered the vector from highest to lowest mass, at a given i, the abundances match by definition

$$n_i(\ln M_i^+; +\delta_{\rm b}, V_{\rm c}) = n_i(\ln M_i^-; -\delta_{\rm b}, V_{\rm c}) \tag{23}$$

but relate to different masses. Note that the total length of the vectors can differ, and so the matching stops at $i = \min(N_{\rm tot}^+, N_{\rm tot}^-)$. We then form the elements of the mass shift data vector as

$$s_i = \frac{\ln M_i^+ - \ln M_i^-}{2\delta_\mathrm{b}},$$
$$\ln M_i = \frac{\ln M_i^+ + \ln M_i^-}{2}, \tag{24}$$

which we denote as $\mathbf{s}(\ln \mathbf{M}; V_\mathrm{c})$.

We then estimate the underlying smooth functions $\hat{n}(\ln M; \delta_\mathrm{b} = 0, V_\mathrm{c})$ and $\hat{s}(\ln M; V_\mathrm{c})$ from these data vectors using the penalized spline technique described in detail in Appendix A 1, with two spline knots per dex in mass

$$\ln \hat{n}(\ln M) = \mathcal{S}\{\ln \mathbf{n}(\ln \mathbf{M})\}, \tag{25}$$

$$\hat{s}(\ln M) = \mathcal{S}\{\mathbf{s}(\ln \mathbf{M})\}, \tag{26}$$

where $S\{\}$ denotes the smoothing operator. Finally, we estimate the differential mass function as the derivative of $\hat{n}(\ln M)$,

$$\hat{n}_{\ln M}(\ln M) = -\frac{\mathrm{d}\hat{n}(\ln M)}{\mathrm{d}\ln M}. \tag{27}$$

Using mock catalogs drawn from a known mass function, we demonstrate in Appendix A 1 that the bias of estimators in Eqs. (25) and (27), if any, is better than the subpercent level and much smaller than the statistical error. To quantify the statistical error, we sample with replacement from the N_sim simulations to make a bootstrap resampled construction of $\hat{n}$, $\ln \hat{n}_{\ln M}$, and $\hat{s}$. By repeating this procedure 100 times, we measure the bootstrap error as the standard deviation of the resamples.

We present the mass function measurement in Fig. 2 as well as the fitting function from Ref. [48], with the latter labeled as "T08" in this paper. Their difference is consistent with the stated precision of the fitting formula but is typically much larger than the bootstrap error. Figure 3 shows the mass shift estimate from all pairs of separate universe simulations. The bootstrap error is of the order of a few percent or better over the mass range $6 \times 10^{12} \sim 2 \times 10^{15} M_\odot$. Note the turn located between $10^{14} M_\odot$ and $10^{15} M_\odot$ corresponds to the transition between polynomial and exponential regions in the halo mass function in Fig. 2.

D. Response vs clustering bias

From the estimates of the mass functions and the shift of threshold mass, we construct the response bias cumulative from a threshold $\bar{b}_1(M) = \bar{b}_1(M; \infty)$ using Eq. (15) as shown in Fig. 4. We compare this result to the fitting formula for $b_1(M)$ from Ref. [22] integrated over the self-consistent mass function from Ref. [48]. Our results are systematically low by ~2% at the low mass end and differ by up to 6% at the high mass end.

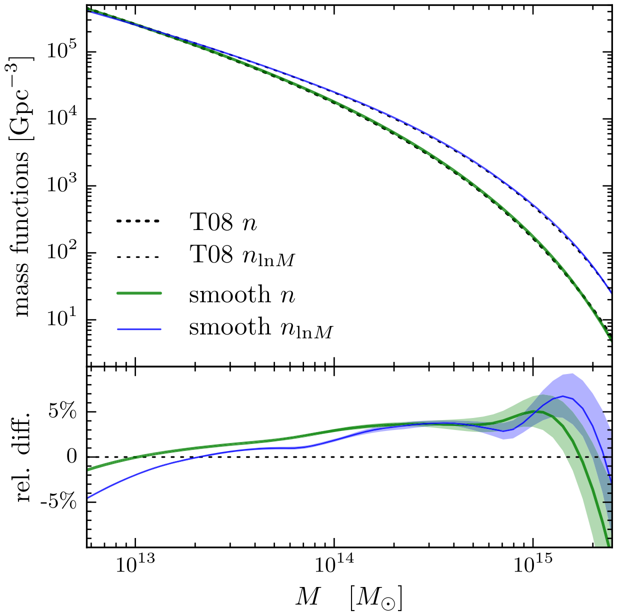

FIG. 2. Cumulative (thick solid green) and differential (thin solid blue) mass functions at $z = 0$ calibrated by penalized-spline smoothing the cumulative number density of all $(500\ \mathrm{Mpc}/h)^3$ fiducial simulations. Shaded regions show the standard deviation of bootstrap resamples. The T08 fitting mass functions [48] (dashed black) are also shown for reference with the lower panel showing the difference for each case.

In Fig. 5, we show the average bias in five logarithmically spaced mass bins per dex plotted as $b_1(M) = \bar{b}_1(M_1, M_2)$ using Eqs. (16) and (18). We compare this to the unbinned $b_1(M)$ from Ref. [22] for reference.

To calibrate clustering bias, we follow Eq. (2) and measure the auto matter power spectrum $P_{\delta\delta}$ and the cross halo-matter power spectrum $P_{\mathrm{h}\delta}$. We bin halos in either the

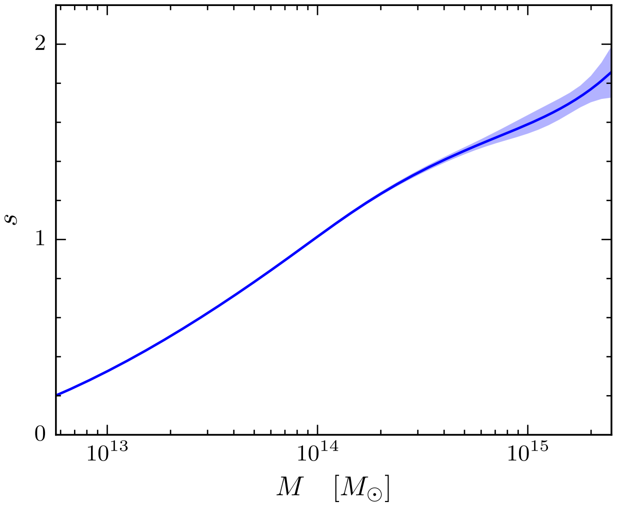

FIG. 3. Threshold mass shift as a response of varying δ_b at fixed cumulative abundance at $z = 0$. The solid blue line and shaded region show the smoothed estimate and the bootstrap error.

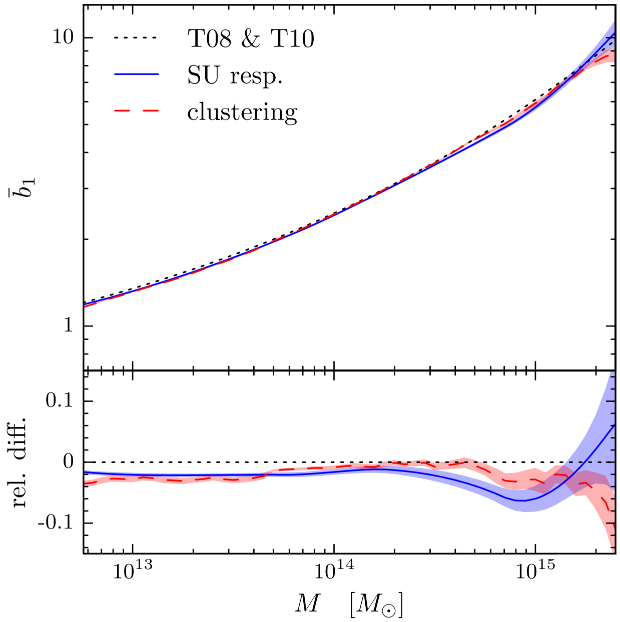

FIG. 4. Average bias for halos with mass $> M$. The solid blue line and shaded area show the SU response bias with bootstrap errors, whereas the dashed red line and shaded area show the same for the clustering bias. The dotted line shows the bias of T10 [22] integrated over the mass function of T08 [48] for comparison.

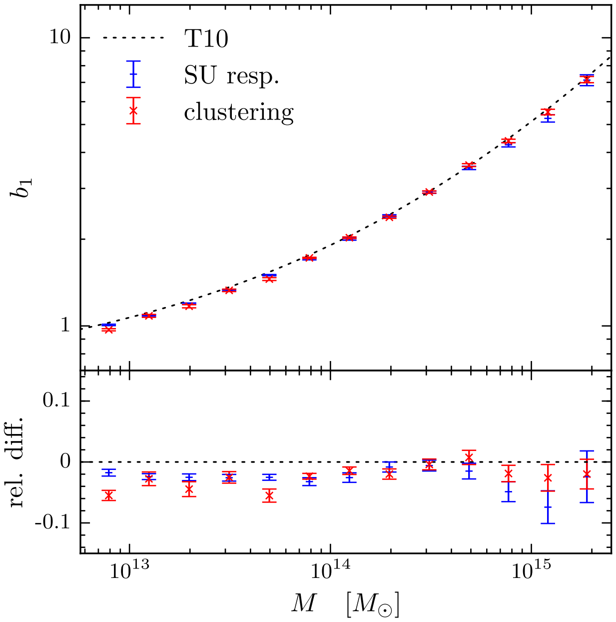

FIG. 5. Average bias for halos in mass bins. Blue + points show the SU response bias with bootstrap errors centered on average masses (19), and red × points show the same for the clustering bias. The dotted line shows the fitting formula of the clustering bias from T10 [22].

same five logarithmic mass bins per dex or cumulative above threshold and assign the particles or halos in each bin to a 256^3 grid with the cloud-in-cell (CIC) scheme and apply the Fast Fourier Transform (FFT) before deconvolving the CIC window.

For halos in a mass bin $[M_1, M_2]$, we can estimate the clustering bias following Eq. (2),

$$\hat{\bar{b}}_1(M_1, M_2) = \frac{\sum_{|\mathbf{k}|<k_{\max}} \langle \delta_{\rm h}^*(\mathbf{k})\delta(\mathbf{k})\rangle}{\sum_{|\mathbf{k}|<k_{\max}} \langle \delta^*(\mathbf{k})\delta(\mathbf{k})\rangle}, \tag{28}$$

where the average is over the $N_{\rm sim}$ simulations of the same volume. This quantity matches its response bias analog in Eq. (16) since linearity in $\delta_{\rm h}$ implicitly weights the statistic by number density. We only use large-scale modes up to $k_{\max} = 0.03h/\text{Mpc}$ and show the scale dependence on $k_{\max}$ in Appendix A 3. We conclude that $k_{\max}$ is at most a source of systematic error that is comparable to our statistical error.

Given the lack of high mass halos in the $(500\ \text{Mpc}/h)^3$ simulation volumes, we combine these estimates with the $(1\ \text{Gpc}/h)^3$ simulations according to the expected inverse shot variance weight, i.e. eight times higher weight for the larger volume simulations down to their eight times higher minimum mass. In Appendix A 3, we show results from the two sets separately to test for resolution and volume effects. To estimate the errors, we bootstrap resample with the $N_{\rm sim}$ of each set.

We compare the clustering and response bias in Figs. 4 and 5. The agreement in the $1 \lesssim \bar{b}_1 \lesssim 4$ region is an excellent 1%–2%. For the higher bias of rarer halos, the statistical errors for both quantities increase, but the agreement is better than the 4%–5% level for $\bar{b}_1 \lesssim 8$. The bias in mass bins is slightly noisier but still consistent within the bootstrap errors for $1 \lesssim b_1 \lesssim 8$.

In addition to abundance matching, we also measure the response bias directly from the change of number counts in the same set of mass bins. We present the comparison between the two methods in Appendix A 2, both to demonstrate the robustness of abundance matching and to show its statistical efficiency.

IV. DISCUSSION

The linear halo bias is the response of the halo number density to a change in the long-wavelength dark matter density as manifest in the cross-correlation between the clustering of halos and the dark matter. In this paper, we have used the separate universe simulation technique to calibrate the response bias of halos by treating the long-wavelength density mode as a change in the background density in a separate universe. By using pairs of SU simulations with the same realizations of the initial Gaussian random seeds, we can reduce sample variance effects when comparing the mass functions in two separate universes.

Rather than comparing the mass functions at each mass bin in the SU simulations, we introduced an alternative method, the abundance matching method for the comparison, where we adjust the mass threshold so as to have the same cumulative abundance of halos above the mass threshold in the separate universes. We show how to calibrate the response bias from the mass threshold shift and the mass functions. The method can robustly extract the effect of subtle changes in the mass of individual halos, caused by the different merger and accretion histories in the paired SU simulations, and thus outperform the direct method by a factor of 3–5 in statistical power.

We found agreement between the response and clustering biases at the 1%–2% level for average biases $1 \lesssim \bar{b}_1 \lesssim 4$ and no significant deviations at the 4%–5% level out to $\bar{b}_1 \sim 8$. This excellent agreement provides a precise test of the consistency relation between the changes in a one-point function, the halo mass function, and a two-point function, the halo-matter cross-correlation in the large-scale limit that can in principle test for new physics in the dark matter, dark energy, or primordial non-Gaussianity sectors. Our results are systematically lower than the bias given by the T10 fitting formula [22] by 2% and differ by up to 6% at high mass end.

Our method can be easily extended to including other effects in halo bias beyond the flat ΛCDM cosmology. It would be straightforward to apply SU techniques in cosmological hydrodynamical simulations for studying effects of baryonic physics on the large-scale halo bias. Further, massive neutrinos and/or dark energy change the growth of long-wavelength dark matter perturbation and will in turn cause changes in the response of the halo mass function. Primordial non-Gaussianity causes additional mode coupling between the long- and short-wavelength modes, inducing a characteristic scale-dependent effect on the halo bias at large scales [2]. Different halos of the same mass can have different large-scale bias if the halos experience different assembly histories—the so-called assembly bias [39,49]. A generalization of the SU simulation technique can give a better handle on calibrating these modifications in the halo bias by reducing the sample variance effects for both the long-wavelength and short-wavelength modes.

ACKNOWLEDGMENTS

We thank Emanuele Castorina, Marilena LoVerde, Surhud More, Fabian Schmidt, and Ravi Sheth for useful discussion. W. H. is supported by the U.S. Department of Energy Contract No. DE-FG02-13ER41958, NASA ATP NNX15AK22G, and by the Kavli Institute for Cosmological Physics at the University of Chicago through Grant No. NSF PHY-1125897 and an endowment from the Kavli Foundation and its founder Fred Kavli. M. T. is supported in part by a Grants-in-Aid for Scientific Research from the JSPS Promotion of Science (Grants No. 23340061 and No. 26610058), MEXT Grant-in-Aid for Scientific Research on Innovative Areas (Grants No. 15H05893 and No. 15H05892), and by JSPS Program for Advancing Strategic International Networks to Accelerate the Circulation of Talented Researchers. This work was completed in part with the computation and storage resources provided by the University of Chicago Research Computing Center.

APPENDIX A: ROBUSTNESS OF TECHNIQUES

In this Appendix, we describe our smoothing procedure and demonstrate its robustness when applied as a mass function estimator in Appendix A 1. Appendix A 2 shows the robustness and statistical power of the abundance matching technique compared with the direct measurement of abundance changes in fixed mass bins. We test the dependence of the clustering bias on k_{max}, the resolution, and the volume in Appendix A 3.

1. Spline smoothing robustness

The halo abundance and mass shift measured from a simulation are defined at a discrete set of masses of its constituent halos. Instead of the commonly adopted method that bins the noisy data in mass, we smooth the cumulative mass function and mass shift and demonstrate its advantage and robustness below.

Among all the twice differentiable functions that model our discrete observations (x_i, y_i), $i = 1, \ldots, n$, we look for the $f(x) = \hat{f}(x)$ that minimizes

$$\sum_{i=1}^{n}[y_i - f(x_i)]^2 + \lambda \int_{x_1}^{x_n} f''(x)^2 \mathrm{d}x. \tag{A1}$$

The first term is the residual sum of squares, which encourages $\hat{f}(x)$ to fit the data well, while the second one is a penalty term that suppresses variability. The nonnegative smoothing parameter λ controls the tradeoff between fidelity and smoothness or bias and variance. When $\lambda = 0$, the resulting $\hat{f}(x)$ becomes the interpolating spline, while when $\lambda \to \infty$, it converges to the linear least squares.

It can be shown that the solution that minimizes Eq. (A1) is a natural cubic spline with knots at x_i (see e.g. Ref. [50]), known as a smoothing spline. This procedure is nonparametric but is computationally intense for a large number of data points. In practice, we can greatly improve the performance and avoid overfitting by using a smaller number of knots. This latter approach is sometimes referred to as the penalized spline.

Consider the function estimates of the form

$$f(x) = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{b}(x) \equiv \sum_{j=1}^{m} \beta_j b_j(x), \tag{A2}$$

where $\mathbf{b}^{\mathrm{T}}(x) \equiv [b_1(x), \ldots, b_m(x)]$ are the basis functions for natural cubic splines with m knots. So we can write Eq. (A1) in terms of the bases

$$|\mathbf{y} - \mathbf{B}\boldsymbol{\beta}|^2 + \lambda \boldsymbol{\beta}^{\mathrm{T}}\Omega\boldsymbol{\beta}, \tag{A3}$$

where $B_{ij} \equiv b_j(x_i)$ and $\Omega_{jk} \equiv \int b_j''(x) b_k''(x) \mathrm{d}x$, with $i = 1, \ldots, n$ and $j, k = 1, \ldots, m$. The coefficients $\boldsymbol{\beta}^{\mathrm{T}} \equiv [\beta_1, \ldots, \beta_m]$ that minimize Eq. (A3) are

$$\hat{\boldsymbol{\beta}} = (\mathbf{B}^{\mathrm{T}}\mathbf{B} + \lambda\Omega)^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{y}, \tag{A4}$$

and thus our function estimate

$$\begin{aligned} \hat{f}(x) &= \mathbf{b}^{\mathrm{T}}(x)(\mathbf{B}^{\mathrm{T}}\mathbf{B} + \lambda\Omega)^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{y} \\ &\equiv \mathcal{S}\{\mathbf{y}(\mathbf{x})\}, \end{aligned} \tag{A5}$$

where $\mathcal{S}\{\}$ denotes the smoothing operator that maps discrete data to the estimate of a continuous function. The fitted values at $\mathbf{x}^{\mathrm{T}} \equiv [x_i, \ldots, x_n]$ are

$$\hat{\mathbf{y}} \equiv \hat{f}(\mathbf{x}) = \mathbf{S}\mathbf{y}, \tag{A6}$$

where the matrix $\mathbf{S} \equiv \mathbf{B}(\mathbf{B}^{\mathrm{T}}\mathbf{B} + \lambda\Omega)^{-1}\mathbf{B}^{\mathrm{T}}$ acts linearly on the data $\mathbf{y}^{\mathrm{T}} \equiv [y_i, \ldots, y_n]$.

To avoid either overfitting or oversmoothing, we choose the smoothing parameter λ by cross-validation. Specifically, the criterion of the leave-one-out cross-validation (LOOCV) is widely used [50]. In LOOCV, we successively take each data point i as a validation point for the smoothing operation trained on the remaining $n - 1$ data points. We choose the value of λ that minimizes the sum over the squared residuals for these points,

$$\sum_{i=1}^{n} [y_i - \hat{f}_\lambda^{(-i)}(x_i)]^2 = \sum_{i=1}^{n} \left[\frac{y_i - \hat{f}_\lambda(x_i)}{1 - [\mathbf{S}_\lambda]_{ii}}\right]^2, \tag{A7}$$

where the superscript $^{(-i)}$ indicates the fit leaving the ith observation (x_i, y_i) out and the subscript $_\lambda$ makes the λ dependence explicit. The equality in Eq. (A7) [50] allows this procedure to be performed without explicitly obtaining $\hat{f}_\lambda^{(-i)}$ for each point.

In this paper, we utilize this penalized spline method to smooth discrete data sets, including halo catalogs in fiducial simulations and the shift of the threshold mass when matching the abundance between paired separate universe simulations. This procedure avoids problems with binning halos in mass as well as taking derivatives of noisy data.

To verify the robustness, we test our smoothing estimator on mock data, drawn from a known distribution. For this purpose, we use the fitting formula for the halo mass function in Ref. [48] to generate 1000 mock catalogs. The minimum mass in the catalogs is $1.4 \times 10^{12} M_\odot$, corresponding to the smallest halos that our halo finder keeps (100 particles). We also introduce a maximum mass $10^{16} M_\odot$ since there is a negligible probability of obtaining even one such halo in the ΛCDM cosmology. We populate catalogs with total number $\hat{N}_{\mathrm{halo}}$ drawn from a Poisson distribution, with the mean as the mean number of halos in a volume of $4\ \mathrm{Gpc}^3/h^3$, the same as that of all fiducial simulations combined. For each halo in the catalog, we use the inverse cumulative distribution function algorithm to draw its mass and form a realization of the cumulative number density $n_i(\ln M_i)$.

We employ the smoothing algorithm described above to provide an estimate of the underlying smooth function $\hat{n}(\ln M)$ from the discrete data. The smoothing function needs to handle both the polynomial and exponential regions of the mass function. To achieve this, we take the natural logarithm of both the cumulative number density n_i and the mass M_i, $i = 1, \ldots, \hat{N}_{\mathrm{halo}}$, before applying the smoothing operation in Eq. (A5) with 2 knots per dex in mass

$$\ln \hat{n}(\ln M) = \mathcal{S}\{\ln \mathbf{n}(\ln \mathbf{M})\}, \tag{A8}$$

where $\hat{n}(\ln M)$ is the function estimate. Thus, we can estimate the mass function by taking the derivative of the smooth cumulative mass function estimator

$$\hat{n}_{\ln M}(\ln M) = -\frac{\mathrm{d}\hat{n}(\ln M)}{\mathrm{d}\ln M}. \tag{A9}$$

Note that we include halos with 100–400 particles for smoothing, to avoid the enhanced error near the edge, but only trust and present results for halos with ≥ 400 particles.

We set up the robustness test to exactly parallel to our estimation of halo mass functions. Figure 6 shows that the bias of the smoothing estimator, if any, is at the subpercent level, much smaller than the statistical error per catalog.

2. Abundance matching robustness and performance

In Sec. II C, we demonstrate the abundance matching technique for the response bias calibration from separate universe simulations and show the results in Sec. III C. Abundance matching efficiently makes use of the mass information of almost all the halos, whereas in a direct measurement of abundance changes within a set of fixed mass bins, only halos near bin edges are shifted into neighboring bins and counted. Here, we compare the bias measured with both methods, from the same set of simulations, both as a test of robustness of abundance matching and as a demonstration of its statistical power.

Let us denote the number counts of halos with mass in $[M_1, M_2]$ in all separate universe realizations of the same δ_{b} and V_{c} by $\Delta N(+\delta_{\mathrm{b}}, V_{\mathrm{c}})$ and $\Delta N(-\delta_{\mathrm{b}}, V_{\mathrm{c}})$. Following Eq. (1), the average bias in this mass bin is

$$\hat{\bar{b}}_1(M_1, M_2) = \frac{\ln(\Delta N(+\delta_{\mathrm{b}}, V_{\mathrm{c}})/\Delta N(-\delta_{\mathrm{b}}, V_{\mathrm{c}}))}{2\delta_{\mathrm{b}}}. \tag{A10}$$

We show the response bias by both methods in Fig. 7, where the statistical consistency verifies the robustness of the abundance matching technique. Its advantage over the

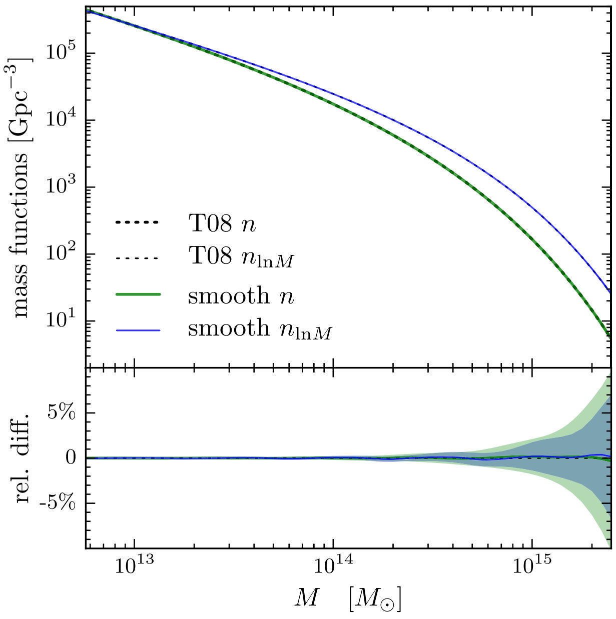

FIG. 6. Robustness of the smoothing procedure verified by comparing smoothed abundance estimates from 1000 mocks drawn from the fitting mass function T08 [48] to the function itself (solid). We generate each mock catalog for halos between $1.4 \times 10^{12} M_\odot$ and $10^{16} M_\odot$, in a volume of $4\ \mathrm{Gpc}^3/h^3$, the same as that of all fiducial $(500\ \mathrm{Mpc}/h)^3$ simulations combined. Lines and shaded regions show the mean and scatter of the estimated cumulative (thick green) and differential (thin blue) mass functions.

direct calibration is obvious with the greatly reduced errors by a factor of 3–5.

3. Clustering bias robustness

The calibration of the clustering bias depends on the $k_{\max}$ cut on the large-scale modes as well as the resolution and volume of the simulations. Repeating the bias estimation in Eq. (28) with different $k_{\max}$, we present the scale dependence in Fig. 8 for $V_c = (500\ \mathrm{Mpc}/h)^3$ and $V_c = (1\ \mathrm{Gpc}/h)^3$ separately. As $k_{\max}$ approaches the nonlinear scale, the bias increases with $k_{\max}$ for the most massive halos and slightly decreases for $\lesssim 10^{13} M_\odot$ halos, similar to the trend demonstrated in Fig. 2 of Ref. [51]. These trends are also stable between the two volumes which have different mass resolutions.

In the main text, we compromise between losing modes, increasing the statistical errors, and using more modes but increasing the systematic bias by choosing $k_{\max,\mathrm{fid}} = 0.021\ \mathrm{Mpc}^{-1}$. Taking the measurement with this choice as the fiducial values, we can quantify the possible systematic bias of using a different $k_{\max}$ by the deviation averaged over mass bins,

$$\frac{1}{N_{\mathrm{bin}}} \sum_{i}^{N_{\mathrm{bin}}} \frac{[b_1(M_i; k_{\max}) - b_1(M_i; k_{\max,\mathrm{fid}})]^2}{\sigma_{b_1}(M_i; k_{\max,\mathrm{fid}})^2}. \tag{A11}$$

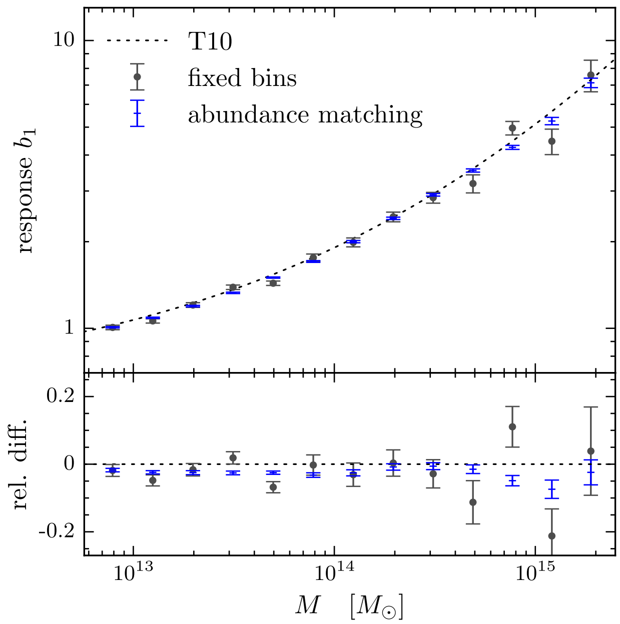

FIG. 7. Abundance matching robustness and efficiency. Blue + points show the response bias with bootstrap errors centered on average masses [Eq. (19)] by abundance matching, and grey • points show the same by the direct measurement of abundance changes within fixed mass bins. The two methods give consistent results, while the former has much reduced errors by a factor of 3–5. The dotted line shows the fitting formula of the clustering bias from T10 [22].

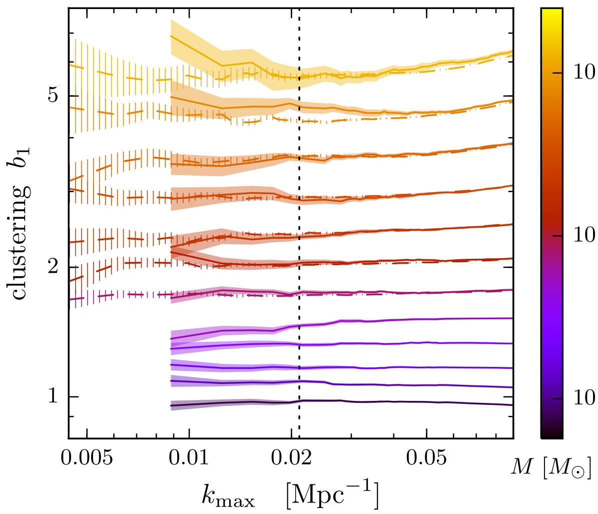

FIG. 8. Dependence of clustering bias calibration on $k_{\max}$, in the $V_c = (500\ \mathrm{Mpc}/h)^3$ simulations (solid, shaded) and $V_c = (1\ \mathrm{Gpc}/h)^3$ (dashed, hatched) at $z = 0$. Shown are the mean and bootstrap errors for the 0.2 dex mass bins centered from $7.9 \times 10^{12} M_\odot$ to $1.2 \times 10^{15} M_\odot$. Larger $k_{\max}$ gives more modes and thus smaller variance but also introduces bias due to scale dependence approaching the nonlinear scale. We choose to use modes below the dashed line (see the text for discussion of robustness).

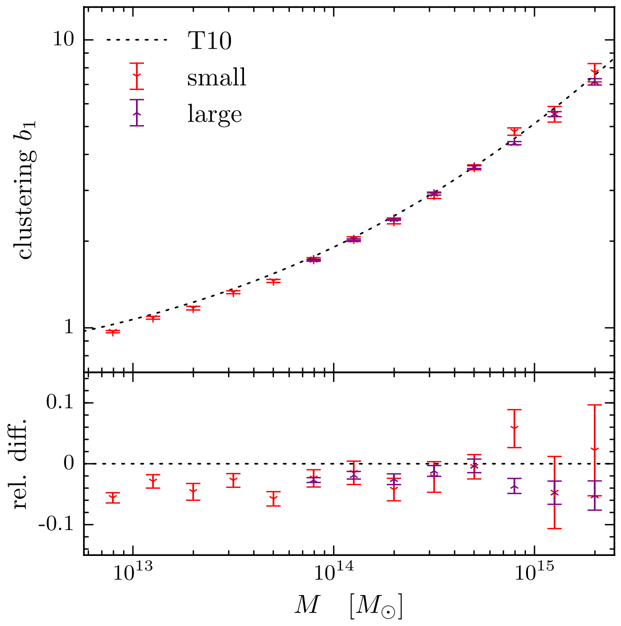

FIG. 9. Clustering bias robustness to simulation volume $V_{\rm c} = (500\ {\rm Mpc}/h)^3$ (small) and $(1\ {\rm Gpc}/h)^3$ (large). Overlapping points show the level of robustness to the 400 particle criteria for the minimum halo mass in the large volume and fluctuations due to the lack of high mass halos in the small volume.

For $V_{\rm c} = (500\ {\rm Mpc}/h)^3$, the k range where this average variance is below 1 is from 0.013 $\rm Mpc^{-1}$ to 0.03 $\rm Mpc^{-1}$; for $V_{\rm c} = (1\ {\rm Gpc}/h)^3$, there is a very similar range from 0.015 $\rm Mpc^{-1}$ to 0.035 $\rm Mpc^{-1}$. Given the substantial range in the linear regime over which results are stable, we conclude that the systematic error due to $k_{\rm max}$ is at most comparable to our statistical error.

With the fiducial $k_{\rm max,fid} = 0.021\ {\rm Mpc}^{-1}$, we show in Fig. 9 the results for $b_1(M)$ of the two volume types separately. In the main text, we combined the volumes (cf. Fig. 5). For most of the mass bins, the clustering bias measured from the large $(1\ {\rm Gpc}/h)^3$ volume simulations agrees well with that from the small $(500\ {\rm Mpc}/h)^3$ ones, confirming that 400 particles are enough to resolve halos for estimating the clustering bias. The small volume estimates fluctuate substantially at the high mass end due to having very few high mass halos in such volumes. In fact, the high point at $\sim 8 \times 10^{14} M_\odot$ can be traced back to Fig. 8 as a statistical fluctuation of the $k_{\rm max,fid} = 0.021\ {\rm Mpc}^{-1}$ modes that is not present at higher $k_{\rm max}$.

APPENDIX B: Universal mass function

As explained in Sec. II A, the response bias is often approximated by assuming a universal mass function (UMF) rather than the more exact separate universe approach introduced in the main text. In addition to the universality assumption, the mass function is typically fit to a specific functional form motivated by spherical collapse and the excursion set approach (e.g. Refs. [22,27]). To separate the roles of these assumptions, we calibrate the universal form nonparametrically and compare the results to the clustering bias, both measured from the same halo catalog.

1. UMF response bias

The universality assumption restricts the halo mass function in the following form,

$$n_{\ln M}(M) = \frac{\bar\rho_m}{M} \nu f(\nu) \frac{\partial \ln \nu}{\partial \ln M}, \tag{B1}$$

where the multiplicity function $\nu f(\nu)$ captures the mass fraction (per $\ln \nu$) contained in halos of peak height $\nu \equiv \delta_{\rm c}/\sigma(M)$. Here, $\delta_{\rm c}$ is the linear threshold of spherical collapse and is usually taken as the Einstein-de Sitter value $\delta_{\rm c} = 1.686$ due to its weak cosmology dependence. The rms of the linear density fluctuation is computed as usual,

$$\sigma^2(M) = \int \frac{d^3k}{(2\pi)^3} P_{\rm lin}(k) |W(kR)|^2, \tag{B2}$$

where $M = 4\pi\bar\rho_m R^3/3$ is the enclosed mass, $P_{\rm lin}(k)$ is the linear power spectrum, and the top-hat window function is

$$W(x) = \frac{3}{x^3}(\sin x - x \cos x). \tag{B3}$$

In the UMF response bias approach, the shift in the background density is viewed as an effective change in the collapse threshold $\delta_{\rm c} \to \delta_{\rm c} - \delta_{\rm b}$ or in the peak height

$$\nu = \frac{\delta_{\rm c} - \delta_{\rm b}}{\sigma}. \tag{B4}$$

Thus, the linear bias becomes

$$\begin{aligned} n_{\ln M} b_1^{\rm L} &= \frac{\bar\rho_m}{M} \frac{\partial}{\partial \delta_{\rm b}} \left(\nu f(\nu) \frac{\partial \ln \nu}{\partial \ln M} \right) \\ &= -\frac{\bar\rho_m}{M} \frac{1}{\delta_{\rm c}} \frac{d \nu f(\nu)}{d \ln \nu} \frac{\partial \ln \nu}{\partial \ln M} \\ &= -\frac{\partial \mu n_{\ln M}}{\partial \ln M} - \mu n_{\ln M}, \end{aligned} \tag{B5}$$

where we have introduced a shorthand,

$$\mu = \frac{1}{\delta_{\rm c}} \left(\frac{\partial \ln \nu}{\partial \ln M} \right)^{-1}, \tag{B6}$$

and the average UMF bias above $M_{\rm th}$ becomes

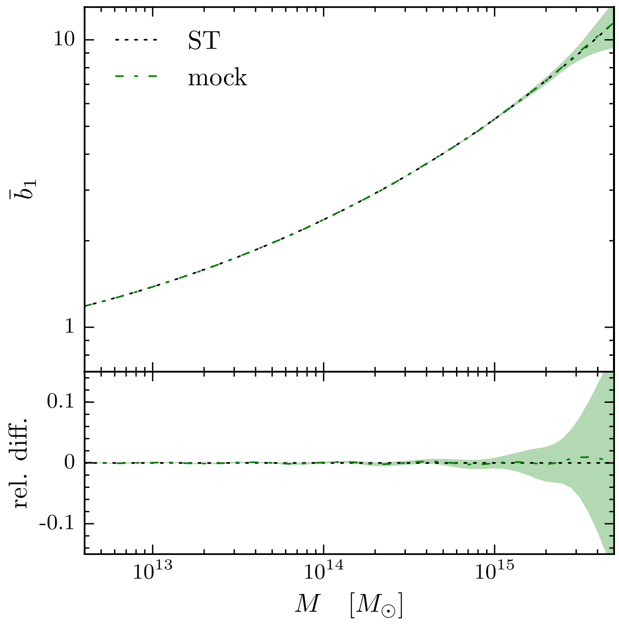

FIG. 10. Robustness of the UMF response bias estimator verified using 1000 mock catalogs of the Sheth-Tormen (ST) mass function compared with the analytic ST bias. The dotted-dashed line shows the mean of the estimated $\bar{b}_1$ matches the analytic result (dashed) to $\lesssim 1\%$, well within the scatter of the estimated $\bar{b}_1$ (shaded region).

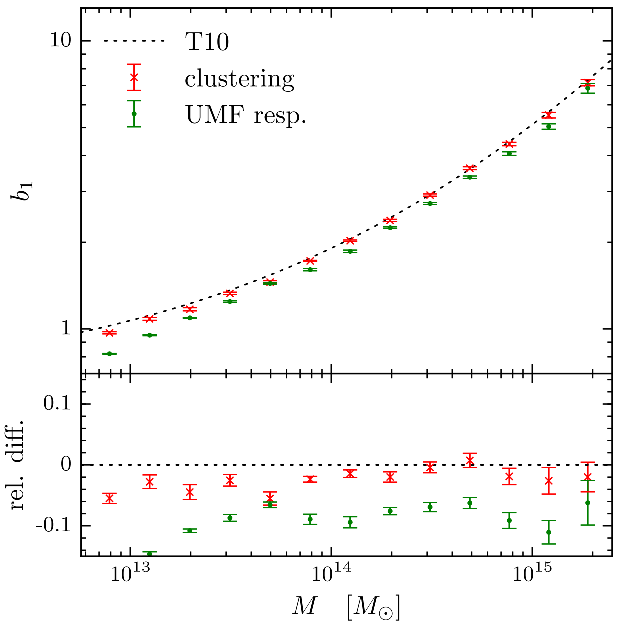

FIG. 12. Average bias for halos in mass bins. Green dots show the UMF response bias with bootstrap errors centered on average masses (19), and red × points show the same for the clustering bias. The dotted line shows the fitting formula of clustering bias from T10 [22].

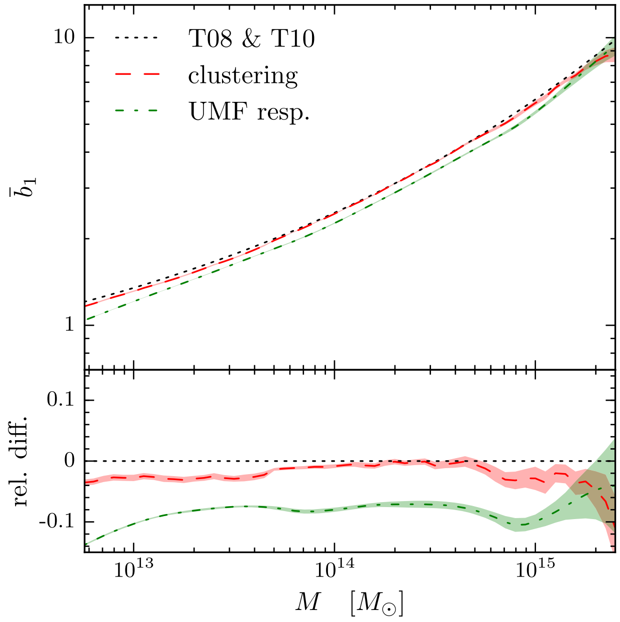

FIG. 11. Average bias for halos with mass $> M$. the dotted-dashed green line and shaded area show the UMF response bias with bootstrap errors, whereas the dashed red line and shaded area show the same for the clustering bias. The dotted line shows the clustering bias of T10 [22] integrated over the mass function of T08 [48] for comparison.

$$\begin{aligned}\bar{b}_1^{\mathrm{L}}(M_{\mathrm{th}}) &\equiv \frac{1}{n(M_{\mathrm{th}})}\int_{M_{\mathrm{th}}}^{\infty}\frac{\mathrm{d}M}{M}\,b_1^{\mathrm{L}} n_{\ln M}\\ &= \frac{\mu(M_{\mathrm{th}})n_{\ln M}(M_{\mathrm{th}})}{n(M_{\mathrm{th}})} - \bar{\mu}(M_{\mathrm{th}}),\\ \bar{\mu}(M_{\mathrm{th}}) &\equiv \frac{1}{n(M_{\mathrm{th}})}\int_{M_{\mathrm{th}}}^{\infty}\frac{\mathrm{d}M}{M}\,\mu n_{\ln M}.\end{aligned} \tag{B7}$$

Given $\bar{b}_1^{\mathrm{L}}(M_{\mathrm{th}})$, we can difference to get the average bias in a finite mass bin,

$$\begin{aligned}\bar{b}_1^{\mathrm{L}}(M_1, M_2) &\equiv \frac{\int_{M_1}^{M_2}\mathrm{d}\ln M\, b_1^{\mathrm{L}} n_{\ln M}}{\int_{M_1}^{M_2}\mathrm{d}\ln M\, n_{\ln M}}\\ &= \frac{\bar{b}_1^{\mathrm{L}}(M_1)n(M_1) - \bar{b}_1^{\mathrm{L}}(M_2)n(M_2)}{n(M_1) - n(M_2)}.\end{aligned} \tag{B8}$$

We should emphasize that Eqs. (B7) and (B8) describe a nonparametric procedure to calibrate the UMF response bias quantities. In deriving them, we do not assume any functional form for the multiplicity function $\nu f(\nu)$ (cf. Refs. [22,27]), as such assumptions can introduce a systematic bias into the measurement. On the other hand, by doing so, we can no longer make the connection to excursion set methods based on either a fixed or moving barrier [19].

Similar to the SU response bias calibration, here we also need the cumulative and differential mass functions.

In addition, we need to estimate the number density weighted $\bar{\mu}$ above threshold mass $M_{\rm th}$ to quantify the UMF response bias.

2. Bias comparisons

In Sec. III C, we have explained how to make continuous estimates of $n_{\ln M}$ and n from the discrete halo catalog measured from simulations. Following the same reasoning, we can construct the estimator for $\bar{\mu}(M_{\rm th})$. Similar to Eq. (22), we start from a halo catalog and arrange the cumulative sum in descending order in mass,

$$\ln \mathbf{M} = [\ln M_1, \dots \ln M_{N_{\rm tot}}]^{\rm T},$$
$$\bar{\boldsymbol{\mu}} = \left[\frac{\mu(M_1)/2}{1/2}, \dots, \frac{\sum_{i=1}^{N_{\rm tot}-1} \mu(M_i) + \mu(M_{N_{\rm tot}})/2}{N_{\rm tot} - 1/2}\right]^{\rm T}. \tag{B9}$$

Recall that the factor of $1/2$ arises from partitioning discrete points. From these data vectors, we can obtain a smooth estimate of $\ln \bar{\mu}$ using a penalized spline (see Appendix A 1) with 2 spline knots per dex in mass,

$$\ln \hat{\bar{\mu}}(\ln M) = \mathcal{S}\{\ln \bar{\boldsymbol{\mu}}(\ln \mathbf{M})\}, \tag{B10}$$

where $S\{\}$ is the smoothing operator.

From these estimates of the mass functions and $\bar{\mu}$, we construct the UMF response biases from Eqs. (B7) and (B8). To verify our estimator for the UMF response bias, we test it on 1000 mocks from the Sheth-Tormen mass function [19], drawn in the same way as explained in Appendix A 1, and compare the result to that analytically derived assuming universality. We show this comparison in Fig. 10 and find that our estimator is accurate to the subpercent level, well below the statistical scatter of each catalog.

Using simulations from the same set of $V_{\rm c} = (500\ {\rm Mpc}/h)^3$, we compare the UMF response biases to the clustering bias (Sec. III D) in Figs. 11 and 12. The UMF response bias is systematically lower than the clustering $\bar{b}_1$ by 5%–10% for $1 \lesssim \bar{b}_1 \lesssim 7$ or lower by $\gtrsim 6\%$ than the clustering b_1 for most of the measured mass range.

The fitting functions for the clustering bias from T08 and T10 are also added as references. For both $\bar{b}_1$ and b_1, the UMF response biases are systematically lower than the fitting functions by ~8%.

We conclude that the UMF response bias is statistically inconsistent with the clustering bias, at least for halos identified at $\Delta = 200$. Given the excellent agreement between the clustering bias and the SU response bias, the UMF response bias is also inconsistent as an approximation of the latter.

[1] N. Kaiser, Astrophys. J. **284**, L9 (1984).

[2] N. Dalal, O. Doré, D. Huterer, and A. Shirokov, Phys. Rev. D **77**, 123514 (2008).

[3] M. Biagetti, V. Desjacques, A. Kehagias, and A. Riotto, Phys. Rev. D **90**, 045022 (2014).

[4] M. LoVerde, Phys. Rev. D **90**, 083530 (2014).

[5] S. More, H. Miyatake, R. Mandelbaum, M. Takada, D. N. Spergel, J. R. Brownstein, and D. P. Schneider, Astrophys. J. **806**, 2 (2015).

[6] http://www.darkenergysurvey.org.

[7] http://desi.lbl.gov.

[8] http://www.naoj.org/Projects/HSC/index.html.

[9] M. Takada *et al.*, Publ. Astron. Soc. Jpn. **66**, R1 (2014).

[10] http://www.lsst.org.

[11] http://sci.esa.int/euclid/.

[12] http://wfirst.gsfc.nasa.gov.

[13] P. McDonald, Phys. Rev. D **74**, 103512 (2006); **74**, 129901(E) (2006).

[14] K. C. Chan and R. Scoccimarro, Phys. Rev. D **86**, 103519 (2012).

[15] F. Schmidt, D. Jeong, and V. Desjacques, Phys. Rev. D **88**, 023515 (2013).

[16] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser, Astrophys. J. **379**, 440 (1991).

[17] H. J. Mo and S. D. M. White, Mon. Not. R. Astron. Soc. **282**, 347 (1996).

[18] H. J. Mo, Y. P. Jing, and S. D. M. White, Mon. Not. R. Astron. Soc. **284**, 189 (1997).

[19] R. Sheth and G. Tormen, Mon. Not. R. Astron. Soc. **308**, 119 (1999).

[20] R. K. Sheth, H. J. Mo, and G. Tormen, Mon. Not. R. Astron. Soc. **323**, 1 (2001).

[21] M. C. Martino and R. K. Sheth, Mon. Not. R. Astron. Soc. **394**, 2109 (2009).

[22] J. L. Tinker, B. E. Robertson, A. V. Kravtsov, A. Klypin, M. S. Warren, G. Yepes, and S. Gottlöber, Astrophys. J. **724**, 878 (2010).

[23] S. Bhattacharya, K. Heitmann, M. White, Z. Lukić, C. Wagner, and S. Habib, Astrophys. J. **732**, 122 (2011).

[24] W. Hu and A. V. Kravtsov, Astrophys. J. **584**, 702 (2003).

[25] U. Seljak and M. S. Warren, Mon. Not. R. Astron. Soc. **355**, 129 (2004).

[26] S. Saito, T. Baldauf, Z. Vlah, U. Seljak, T. Okumura, and P. McDonald, Phys. Rev. D **90**, 123522 (2014).

[27] M. Manera, R. K. Sheth, and R. Scoccimarro, Mon. Not. R. Astron. Soc. **402**, 589 (2010).

[28] M. Manera and E. Gaztanaga, Mon. Not. R. Astron. Soc. **415**, 383 (2011).

[29] T. Baldauf, U. Seljak, V. Desjacques, and P. McDonald, Phys. Rev. D **86**, 083540 (2012).

[30] M. Lima and W. Hu, Phys. Rev. D **70**, 043504 (2004).

[31] M. Oguri and M. Takada, Phys. Rev. D **83**, 023008 (2011).

[32] Y. Li, W. Hu, and M. Takada, Phys. Rev. D **89**, 083519 (2014).

[33] Y. Li, W. Hu, and M. Takada, Phys. Rev. D **90**, 103530 (2014).

[34] E. Sirko, Astrophys. J. **634**, 728 (2005).

[35] T. Baldauf, U. Seljak, L. Senatore, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 10 (2011) 031.

[36] C. Wagner, F. Schmidt, C.-T. Chiang, and E. Komatsu, Mon. Not. R. Astron. Soc. **448**, L11 (2015).

[37] M. Takada and W. Hu, Phys. Rev. D **87**, 123504 (2013).

[38] R. K. Sheth and G. Tormen, Mon. Not. R. Astron. Soc. **350**, 1385 (2004).

[39] R. H. Wechsler, A. R. Zentner, J. S. Bullock, A. V. Kravtsov, and B. Allgood, Astrophys. J. **652**, 71 (2006).

[40] N. Dalal, M. White, J. R. Bond, and A. Shirokov, Astrophys. J. **687**, 12 (2008).

[41] K. Ichiki and M. Takada, Phys. Rev. D **85**, 063521 (2012).

[42] A. V. Kravtsov, A. A. Berlind, R. H. Wechsler, A. A. Klypin, S. Gottlöber, B. Allgood, and J. R. Primack, Astrophys. J. **609**, 35 (2004).

[43] R. M. Reddick, R. H. Wechsler, J. L. Tinker, and P. S. Behroozi, Astrophys. J. **771**, 30 (2013).

[44] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000).

[45] C. Howlett, A. Lewis, A. Hall, and A. Challinor, J. Cosmol. Astropart. Phys. 04 (2012) 027.

[46] http://cosmo.nyu.edu/roman/2LPT/.

[47] V. Springel *et al.*, Nature (London) **435**, 629 (2005).

[48] J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz, Astrophys. J. **688**, 709 (2008).

[49] H. Miyatake, S. More, M. Takada, D. N. Spergel, R. Mandelbaum, E. S. Rykoff, and E. Rozo, Phys. Rev. Lett. **116**, 041301 (2016).

[50] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning* (Springer, New York, 2013).

[51] A. J. Nishizawa, M. Takada, and T. Nishimichi, Mon. Not. R. Astron. Soc. **433**, 209 (2013).