



Dark Quest. I. Fast and Accurate Emulation of Halo Clustering Statistics and Its Application to Galaxy Clustering

Takahiro Nishimichi^{1,2} , Masahiro Takada¹ , Ryuichi Takahashi³ , Ken Osato⁴ , Masato Shirasaki⁵, Taira Oogi¹, Hironao Miyatake^{1,6,7}, Masamune Oguri^{1,4,8} , Ryoma Murata^{1,4} , Yosuke Kobayashi^{1,4}, and Naoki Yoshida^{1,4,8}

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

² Center for Gravitational Physics, Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan

³ Faculty of Science and Technology, Hirosaki University, 3 Bunkyo-cho, Hirosaki, Aomori 036-8561, Japan

⁴ Department of Physics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

⁵ Division of Theoretical Astronomy, National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

⁶ Institute for Advanced Research, Nagoya University, Nagoya 464-8601, Japan

⁷ Division of Particle and Astrophysical Science, Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan

⁸ Research Center for the Early Universe, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Received 2018 December 6; revised 2019 June 20; accepted 2019 July 24; published 2019 October 8

Abstract

We perform an ensemble of N -body simulations with 2048^3 particles for 101 flat w CDM cosmological models sampled based on a maximin distance sliced Latin hypercube design. By using the halo catalogs extracted at multiple redshifts in the range of $z = [0, 1.48]$, we develop DARK EMULATOR, which enables fast and accurate computations of the halo mass function, halo–matter cross-correlation, and halo autocorrelation as a function of halo masses, redshift, separations, and cosmological models based on principal component analysis and Gaussian process regression for the large-dimensional input and output data vector. We assess the performance of the emulator using a validation set of N -body simulations that are not used in training the emulator. We show that, for typical halos hosting CMASS galaxies in the Sloan Digital Sky Survey, the emulator predicts the halo–matter cross-correlation, relevant for galaxy–galaxy weak lensing, with an accuracy better than 2% and the halo autocorrelation, relevant for galaxy clustering correlation, with an accuracy better than 4%. We give several demonstrations of the emulator. It can be used to study properties of halo mass density profiles such as the concentration–mass relation and splashback radius for different cosmologies. The emulator outputs can be combined with an analytical prescription of halo–galaxy connection, such as the halo occupation distribution at the equation level, instead of using the mock catalogs to make accurate predictions of galaxy clustering statistics, such as galaxy–galaxy weak lensing and the projected correlation function for any model within the w CDM cosmologies, in a few CPU seconds.

Key words: large-scale structure of universe – methods: numerical – methods: statistical

1. Introduction

Cosmic large-scale structures are promising avenues to fundamental questions in cosmology. Various wide-area imaging or spectroscopic surveys of galaxies are ongoing and being planned, aimed at addressing the nature of dark matter and dark energy in the universe. These include the Subaru Hyper Suprime-Cam (HSC) Survey⁹ (Aihara et al. 2018), Dark Energy Survey,¹⁰ Kilo-Degree Survey,¹¹ Subaru Prime Focus Spectrograph (PFS; Takada et al. 2014), Dark Energy Spectroscopic Instrument (DESI),¹² Large Synoptic Survey Telescope (LSST),¹³ ESA satellite mission *Euclid*,¹⁴ and NASA satellite mission *WFIRST*.¹⁵ However, one of the most serious systematic effects in galaxy survey-based cosmology lies in the galaxy bias that generally states an inevitable uncertainty in the relation between distributions of dark matter and large-scale structure tracers (Kaiser 1984; see also

Desjacques et al. 2018, for a recent review). Since physical processes involved in galaxy formation and evolution are still impossible to solve from first principles, it is of critical importance to explore a practical route to extracting cosmological information from observables of galaxy surveys, yet being least affected by the galaxy bias uncertainty, in order to attain the full potential of ongoing and future galaxy surveys.

The growth of cosmic structures is driven mainly by the spatial inhomogeneities of dark matter, which are easier to describe analytically on large scales (Bernardeau et al. 2002) or via N -body numerical simulations down to small scales (Miyoishi & Kihara 1975; Davis et al. 1985) than the variety of astrophysical processes where baryons play a major role in order to form galaxies (e.g., Vogelsberger et al. 2014). In practice, however, we can observe only the projected or three-dimensional distribution of galaxies from galaxy surveys from which we have to infer the dark matter distribution. This is not an easy task and a major challenge that all wide-area galaxy surveys must confront. Nevertheless, there is a theory-motivated working hypothesis that we can employ to make a connection between galaxies and the dark matter distribution. Galaxies or galaxy clusters are believed to form inside dark matter halos, which are self-gravitative systems and correspond to the peaks of the primordial mass density field (Kaiser 1984). The distribution of halos with respect to the dark matter

⁹ <https://hsc.mtk.nao.ac.jp/ssp/>

¹⁰ <https://www.darkenergysurvey.org>

¹¹ <http://kids.strw.leidenuniv.nl>

¹² <https://www.desi.lbl.gov>

¹³ <https://www.lsst.org>

¹⁴ <https://www.euclid-ec.org>

¹⁵ <https://wfIRST.gsfc.nasa.gov>

distribution, referred to as halo bias, and its dependence on the halo mass and cosmological models can be predicted in the cold dark matter (CDM)-dominated structure formation scenario using analytical models (Bardeen et al. 1986; Mo & White 1996; Sheth & Tormen 1999; Sheth et al. 2001) and/or N -body simulations (Tinker et al. 2010). Here it is known that the large-scale bias of halos, and therefore galaxies, should approach a constant value, known as “linear bias,” for the adiabatic initial Gaussian conditions of structure formation due to the equivalence principle of gravity (e.g., Desjacques et al. 2018; see Dalal et al. 2008, for a counterexample, such as the primordial non-Gaussian initial condition). On small scales, the halo bias becomes scale-dependent and varies with cosmological models in a complex way due to nonlinearities of structure formation (McDonald 2006; McDonald & Roy 2009; Taruya et al. 2010; Sato & Matsubara 2011; Baldauf et al. 2012; Nishizawa et al. 2013). These distinct behaviors of halo bias over different scales have to be kept in mind in order not to have any bias in cosmological parameter inference.

Observationally, there are promising probes of galaxy surveys that help to link galaxies to the dark matter distribution or halos, at least in a statistical manner. Galaxy-galaxy or cluster-galaxy lensing, which can be measured by stacking shapes of background galaxies around the foreground tracers, allows us to probe the “average” projected matter (mostly dark matter) distribution around the tracers (e.g., Brainerd et al. 1996; dell’Antonio & Tyson 1996; Fischer et al. 2000; Sheldon et al. 2009). The large-scale galaxy-galaxy lensing signal gives a direct estimate of the linear bias of the galaxies (e.g., Hoekstra et al. 2001; Sheldon et al. 2004). However, the weak-lensing signal is generally noisy. Although the small-scale lensing signal has a higher signal-to-noise ratio (S/N), it probes the dark matter distribution inside the same halo, which is generally difficult to predict accurately. Nevertheless, the integrated lensing signal within the projected aperture of the virial radius can be used to infer the average halo mass of galaxies in a sample (e.g., Mandelbaum et al. 2006), which can in turn be used to infer the linear bias at large scales with the help of a theoretical model. The autocorrelation function of galaxies’ positions in the large-scale structure is another powerful probe of cosmology (e.g., Peebles 1980). It can be measured from a wide-area spectroscopic sample and is relatively easy to measure, i.e., with high S/Ns. If only the large-scale correlation signals are used and the linear bias is a priori assumed, the cosmological information can be extracted from the shape information (e.g., Tegmark et al. 2004). However, the small-scale correlations, which carry even higher S/Ns, cannot be interpreted easily, and the correlations of galaxies in the same halo, the so-called one-halo term, add a significant contribution to the measured signal, which complicates the cosmological analysis.

Although each observable alone has its own pros and cons, combining different clustering observables enables us to perform a robust cosmological analysis, e.g., obtain tighter constraints on cosmological parameters, yet simultaneously calibrate systematic errors such as the bias uncertainty that are otherwise difficult to calibrate with each observable alone (e.g., Oguri & Takada 2011; Yoo & Seljak 2012; Schaan et al. 2017, for similar discussion). Implementation of joint-probes cosmology to actual data can be found in various works (Seljak et al. 2005; Hikage et al. 2013; Mandelbaum et al. 2013; Reid et al. 2014; More et al. 2015b; Abbott et al. 2018;

Joudaki et al. 2018). Such analyses can be done by combining wide-area imaging and spectroscopic surveys over the same region of the sky; for instance, this is the case for the Subaru HSC and PFS surveys.

Hence, the purpose of this paper is to develop software to make accurate model predictions for clustering observables in preparation for high-precision cosmology achievable from ongoing and future wide-area galaxy surveys. Motivated by the fact that dark matter halos are building blocks of the large-scale structure and the places hosting galaxies, we build an “emulator,” dubbed DARK EMULATOR, that allows fast, accurate computations of “halo” clustering quantities—halo mass function (HMF), halo-matter cross-correlation function, and halo autocorrelation function—as a function of halo mass, separation, redshift, and cosmological models. To develop the emulator, we use a large number of N -body simulation realizations and their halo catalogs at multiple output redshifts for different cosmological models that cover a sufficiently broad range of models within flat-geometry, time-varying dark energy and CDM cosmologies (hereafter wCDM). These halo clustering quantities include all of the relevant physics, such as the linear halo bias, nonlinear bias, and halo exclusion effect. Since we use a limited number of N -body simulation realizations for sparsely sampled cosmological models in six-dimensional cosmological parameter space, we carefully propagate statistical uncertainties in halo clustering quantities to the model predictions (emulator outputs) by using principal component analysis (PCA) and Gaussian process regression (GPR) in a high-dimensional space of input and output data vectors.

The concept of our study is somewhat similar to emulators developed in previous studies that interpolate various quantities measured from simulations over the cosmological parameter space (Heitmann et al. 2006, 2010, 2016; Habib et al. 2007; Schneider et al. 2008; Lawrence et al. 2010, 2017; Agarwal et al. 2012, 2014; Petri et al. 2015; DeRose et al. 2019; Euclid Collaboration et al. 2019; Garrison et al. 2018; Liu & Madhavacheril 2019; McClintock et al. 2019; Wibking et al. 2019; Zhai et al. 2019). However, our study is quite different from these works in the sense that we do not make a one-to-one mapping between the input cosmological parameters to the final statistical quantities with the emulation process. We focus more on developing a machinery consisting of several building blocks, each of which works as a separate emulator, and combining them in an analytical manner to work together. Specifically, we focus on halo clustering statistics and do not employ any specific prescription to connect halos to galaxies, such as the halo occupation distribution (HOD; Zheng et al. 2005). Hence, to obtain predictions of galaxy clustering quantities that can be compared with the measurements, a user needs to adopt a prescription to model the galaxy-halo connection, especially the one-halo term contributions arising from galaxies in the same halo, and then combine the outputs of DARK EMULATOR to compute the desired statistical quantities. As a working example, we show how to analytically combine the outputs of DARK EMULATOR and the other small-scale physics prescriptions, such as the HOD model and the distribution of satellite galaxies inside a halo, at the equation level (e.g., Fourier transform and numerical integration) to compute clustering quantities of galaxies, such as galaxy-galaxy weak lensing and projected galaxy correlation function, for galaxies in a hypothetical sample. In this sense, our

approach might be regarded as a numerical simulation version of the halo model approach (Ma & Fry 2000; Peacock & Smith 2000; Seljak 2000; Scoccimarro et al. 2001; Valageas & Nishimichi 2011; also see Cooray & Sheth 2002, for a review). Thus, our emulator gives the flexibility that users can decide how to use the emulator outputs for their desired purpose. This study is the initial work of the DARK QUEST campaign project, and the final goal is to use the Dark Quest products to achieve accurate and robust cosmological analysis with wide-area galaxy surveys. Therefore, the requirements we impose for DARK EMULATOR give sufficiently accurate predictions for desired observables and are sufficiently fast to allow cosmological parameter inferences, such as a Markov chain Monte Carlo analysis, in a high-dimensional parameter space, e.g., six-dimensional cosmological parameters plus various nuisance parameter, including HOD parameters. We demonstrate how well we achieve these requirements.

The structure of the paper is as follows. We start with a brief review of the halo approach to the galaxy clustering and the relevant observables in Section 2. In Section 3, we summarize the details of the simulation setups, including the parameter sampling scheme, initial conditions, time evolution, and post-processing. We then discuss the details of each module that constitutes our emulator in Section 4 including the cross-validation tests. We focus on typical halos that host CMASS galaxies observed by the Sloan Digital Sky Survey (SDSS) at $z \sim 0.5$ in this section. We demonstrate how these modules can be combined to make predictions of various halo and galaxy statistics in Section 5. We summarize in Section 6 with comments on the actual situations where our codes can be applied. Convergence studies, our treatment of the massive neutrinos, the mass and redshift dependence of our modules, and an example HOD prescription implemented in the current version of the emulator are shown in appendices. Readers who are interested only in the final accuracies of DARK EMULATOR may go directly to Appendix F for the results of our validation study.

2. Halo Cosmology

Before going into the details of our method to construct DARK EMULATOR, we first describe the concept of our approach. In particular, we describe why we focus on statistical quantities of halos and how we can connect the halo statistics to observables for galaxies and galaxy clusters that can be used to extract cosmological information.

2.1. Galaxy Observables

Our final goal is to make predictions for clustering observables that are available from wide-area galaxy surveys. For example, the galaxy-galaxy weak-lensing signal is measured by cross-correlating the positions of foreground galaxies with the shapes of background galaxies and probes the average excess mass density profile around the lensing galaxies, $\Delta\Sigma_g(R)$. This signal reflects the three-dimensional galaxy-mass cross-correlation function, $\xi_{gm}(x)$, projected along the line-of-sight direction,

$$\Delta\Sigma_g(R) = \bar{\Sigma}_g(<R) - \Sigma_g(R), \quad (1)$$

where

$$\Sigma_g(R) = \bar{\rho}_{m0} \int_{-\infty}^{\infty} \xi_{gm}(\sqrt{R^2 + \pi^2}) d\pi, \quad (2)$$

$$\bar{\Sigma}_g(<R) = \frac{2}{R^2} \int_0^R \Sigma_g(y) y dy. \quad (3)$$

Here we denote by π and R separations in the line-of-sight and transverse directions, respectively, and $\bar{\rho}_{m0}$ is the present-day mean matter density. The use of $\bar{\rho}_{m0}$ is due to the fact that we define the surface mass density in the comoving coordinates rather than the physical coordinates. Similarly, the projected galaxy autocorrelation function is related to the three-dimensional galaxy autocorrelation function, $\xi_{gg}(r)$, via

$$w_{gg}(R) = 2 \int_0^{\pi_{\max}} \xi_{gg}(\sqrt{R^2 + \pi^2}) d\pi \quad (4)$$

for the projection width $[-\pi_{\max}, \pi_{\max}]$.

The simplest linear deterministic bias model, which connects the matter density field δ_m and the galaxy number density field δ_g as $\delta_g = b_g \delta_m$, leads to

$$\xi_{gm} = b_g \xi_{mm}, \quad \xi_{gg} = b_g^2 \xi_{mm}, \quad (5)$$

with a free parameter b_g , which is completely degenerate with the normalization of the linear matter power spectrum, σ_8 . Having both the lensing and clustering signals, one can break this degeneracy and infer the underlying matter correlation function $\xi_{mm}(x)$ by combining the two correlation functions:

$$\xi_{mm}(x) = \frac{[\xi_{gm}(x)]^2}{\xi_{gg}(x)}. \quad (6)$$

In reality, however, both nonlinear corrections and stochasticity can alter this relation. To quantify this, we introduce the cross-correlation coefficient (Tegmark & Peebles 1998) defined by

$$r_{gm}(x) \equiv \frac{\xi_{gm}(x)}{\sqrt{\xi_{gg}(x)\xi_{mm}(x)}}. \quad (7)$$

The departure of r_{gm} from unity characterizes the degree to which the linear deterministic relation is violated.

2.2. Halo Model Approach to Galaxy Bias

Dark matter halos are basic building blocks of large-scale structure and the sites harboring the formation of galaxies and galaxy clusters. Since the physical processes involved in galaxy formation are still difficult to resolve or simulate from first principles, dark matter halos could give us a practical route to connect the theory and observations of galaxy surveys. Hence, in this paper, we develop an emulator that primarily predicts statistical quantities of halos as a function of halo mass, redshift, clustering separation scale, and cosmological parameters (model). This halo model allows us to compute galaxy clustering statistics, e.g., by using an HOD prescription.

The HMF dn/dM is defined as the comoving number density of halos in the mass range $[M, M + dM]$ and at redshift z for a given cosmological model denoted by its parameters \mathbf{p} :

$$\frac{dn}{dM}(M, z, \mathbf{p}) dM. \quad (8)$$

This is the first important quantity that we are going to calibrate with simulations. We will build a module enabling the fast computation.

Two-point clustering properties of halos are characterized by the autocorrelation functions of two halo samples and the

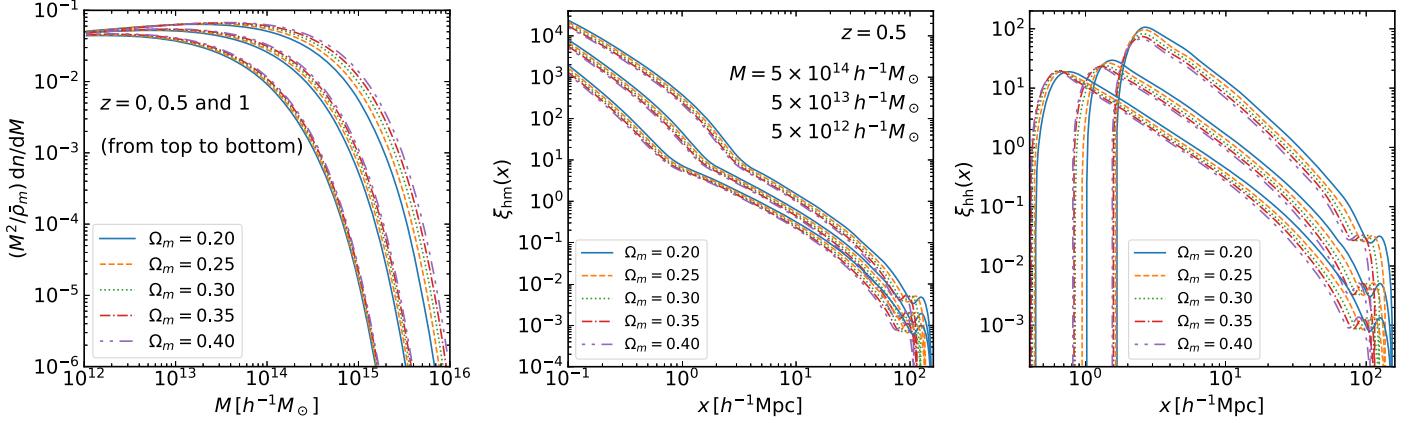


Figure 1. Demonstration of our halo modules in DARK EMULATOR to predict halo-related statistical quantities as a function of cosmological parameters. We show how the HMF (left panel), halo–matter cross-correlation function (middle), and halo autocorrelation function (right) vary with Ω_m for a flat-geometry cosmology but with σ_8 and other cosmological parameters being kept fixed to their fiducial values. The three sets of lines in the left panel show the mass functions at redshifts $z = 0, 0.5$, and 1 , and the three sets in the middle and right panels show the correlation functions for halos at three masses as indicated in the figure legend at redshift $z = 0.5$. Throughout this paper, we use the spherical overdensity (SO) mass, M_{200} , for the halo mass definition, where the mean mass overdensity within the halo boundary is 200 times $\bar{\rho}_{\text{mo}}$.

cross-correlation functions of halos and matter, which we denote as

$$\xi_{\text{hm}}(x; M_1, M_2, z, \mathbf{p}) \quad (9)$$

and

$$\xi_{\text{hh}}(x; M, z, \mathbf{p}), \quad (10)$$

respectively. Here we explicitly denote that the two halo samples in the halo–halo correlation function can have different masses, M_1 and M_2 . We will omit the arguments z and \mathbf{p} below for notational simplicity.

Figure 1 shows examples of the quantities of our main interest, halo statistical quantities predicted by DARK EMULATOR, that we develop in this paper. The figure shows the HMFs, halo–matter cross-correlation functions, and halo–halo autocorrelation functions with a varying density parameter Ω_m but fixing other parameters to their fiducial values. In doing so, we keep the spatial flatness, overall normalization σ_8 , and baryon fraction, and the change in Ω_m is compensated for by the density of dark energy Ω_{de} , as well as the Hubble parameter H_0 . We show in the left panel the prediction at three different redshifts, while in the middle and right panels, we consider three halo masses at $z = 0.5$. Each of these quantities can be computed very quickly by the emulator in ~ 100 ms on a typical modern laptop computer.

Once these statistical quantities of halos are given, we can compute galaxy observables, such as those shown in Equations (1) and (4), based on an empirical HOD model for the mean number of galaxies within a halo with mass M , $\langle N(M) \rangle$. We here employ the functional form originally proposed by Zheng et al. (2005) and then slightly generalized by More et al. (2015b). The explicit formulae, as well as the derivation of the resultant galaxy statistics, can be found in Appendix G.

On large scales, the galaxy statistics are computed as the weighted average of the corresponding halo statistics. Specifically, using the HOD that gives the average number of galaxies in halos of mass M , $\langle N(M) \rangle$, we can compute

$$\xi_{\text{gm}}(x) = \frac{1}{\bar{n}_g} \int dM \frac{dn}{dM}(M) \langle N(M) \rangle \xi_{\text{hm}}(x; M) \quad (11)$$

for the cross and

$$\xi_{\text{gg}}(x) = \frac{1}{\bar{n}_g^2} \int dM_1 \frac{dn}{dM}(M_1) \langle N(M_1) \rangle \times \int dM_2 \frac{dn}{dM}(M_2) \langle N(M_2) \rangle \xi_{\text{hh}}(x; M_1, M_2) \quad (12)$$

for the autocorrelation functions, where the mean galaxy number density, \bar{n}_g , is given by

$$\bar{n}_g = \int dM \frac{dn}{dM} \langle N(M) \rangle. \quad (13)$$

We now introduce the mass-dependent halo bias functions

$$\xi_{\text{hm}}(x; M) = b_{\text{hm}}(x; M) \xi_{\text{mm}}(x), \quad (14)$$

$$\xi_{\text{hh}}(x; M_1, M_2) = b_{\text{hh}}^{(2)}(x; M_1, M_2) \xi_{\text{mm}}(x) \quad (15)$$

for the cross- and autocorrelation functions, respectively. The galaxy correlation functions, Equations (11) and (12), can be rewritten as

$$\xi_{\text{gm}}(x) = b_g^{\text{cross}}(x) \xi_{\text{mm}}(x), \quad (16)$$

$$\xi_{\text{gg}}(x) = [b_g^{\text{auto}}(x)]^2 \xi_{\text{mm}}(x), \quad (17)$$

where the corresponding galaxy bias functions are computed as

$$b_g^{\text{cross}}(x) = \frac{1}{\bar{n}_g^2} \int dM \frac{dn}{dM}(M) \langle N(M) \rangle b_{\text{hm}}(x; M),$$

$$[b_g^{\text{auto}}(x)]^2 = \frac{1}{\bar{n}_g^2} \int dM_1 \frac{dn}{dM}(M_1) \langle N(M_1) \rangle \times \int dM_2 \frac{dn}{dM}(M_2) \langle N(M_2) \rangle b_{\text{hh}}^{(2)}(x; M_1, M_2). \quad (18)$$

$$(19)$$

Now the cross-correlation coefficient for galaxy and matter fields, Equation (7), reads

$$r_{\text{gm}}(x) = \frac{b_g^{\text{cross}}(x)}{b_g^{\text{auto}}(x)}. \quad (20)$$

The condition $r_{\text{gm}} = 1$ is trivially satisfied when the halo autobias function is written as a product of the two cross-bias

Table 1
Summary of Our Simulation Suites

Class	N_{part}	L_{box}	Cosmology	IC	N_{real}	Purpose
HR	2048 ³	1000	Fiducial	Random	28	Assessment of variance (HMF, HMCCF)
			20 models in slice 1	Fixed ^a	1	Test of ICs
			100 models in slices 1–5	Random	1	Emulator (HMF, HMCCF; slices 1–4 for training, slice 5 for validation)
LR	2048 ³	2000	Fiducial	Random	14	Assessment of variance (HACF, PROP)
			100 models in slices 1–5	Random	1	Emulator (HACF, PROP; slices 1–4 for training, slice 5 for validation) (slice 5 also for validation of HMF)
test	256 ³	250	Fiducial	Fixed ^{b,c}	1	Convergence study (same resolution as LR)
	512 ³		Fiducial	Fixed ^{b,c}	1	Convergence study (same resolution as HR)
	1024 ³		Fiducial	Fixed ^b	1	Convergence study
	2048 ³		Fiducial	Fixed ^b	1	Convergence study

Notes. We show the number of particles (N_{part}), comoving box size (L_{box} in h^{-1} Mpc), cosmological model, random number seeds used in initial conditions (IC), number of realizations per model or parameter set (N_{real}), and purpose of the simulations; calibration of either the HMF, halo–matter cross-correlation function (HMCCF), halo autocorrelation function (HACF), halo propagator (PROP), or other testing purposes, such as the IC of simulations.

^a Initial phases taken to be the same as one of the fiducial HR realization.

^b Exactly the same initial phases are employed for the six test simulations.

^c Initial conditions based on the Zel'dovich approximation are generated in addition to 2LPT. Also, five initial redshifts $1 + z_{\text{in}} = 15, 30, 60, 120, and 240 are tested.$

functions,

$$b_{\text{hh}}^{(2)}(r; M_1, M_2) = b_{\text{hm}}(r; M_1)b_{\text{hm}}(r; M_2), \quad (21)$$

for any halo mass M_1 and M_2 . This relation would hold at sufficiently large separations.¹⁶

On mildly nonlinear scales, one should take into account a violation of Equation (21). In particular, understanding the scale-dependent function $r_{\text{gm}}(x)$, and thus $b_{\text{hm}}(x; M)$ and $b_{\text{hh}}^{(2)}(x; M_1, M_2)$, is crucial to recovering the underlying matter correlation function out of galaxy observables. On strongly nonlinear scales, the one-halo term, namely, clustering contribution due to pairs of galaxy–galaxy or galaxy–matter within the same halo, dominates the signal over the two-halo term discussed so far. Especially, the galaxy–galaxy lensing signal defined in Equation (1) in this regime gives information on the average mass of halos in which lensing galaxies reside. For a simplistic scenario where HOD is a delta function at mass M , this halo mass information tells us how the bias function should behave on large scales, breaking the degeneracy between bias and the underlying clustering signal. In more realistic settings, the small-scale information inferred from the galaxy–galaxy lensing helps to determine the HOD parameters, and we can perform a quasi-bias-free analysis using the large-scale clustering signal.

None of these analyses can be realized unless we have good control of the model predictions for the halo clustering signals, including their dependence on mass and cosmological parameters. Therefore, the three quantities, dn/dM , ξ_{hh} , and ξ_{hm} , are our central interest in this paper (see Figure 1 for example plots of these quantities varying Ω_m). Our emulator models these quantities at the core and predicts the galaxy statistics by combining them with an HOD prescription in an analytical manner. In doing so, we pay attention to the evaluation speed of the statistics such that it is feasible to perform a Markov-chain Monte Carlo analysis of parameter inference in a high-

¹⁶ In Fourier space, however, a residual contribution is known to persist even on the large-scale limit, and this behaves as a non-Poissonian shot-noise term, k^0 (Seljak et al. 2009; Hamaus et al. 2010; Baldauf et al. 2013). Hence, it would not contribute significantly in configuration space.

dimensional parameter space, e.g., a space including cosmological parameters, as well as HOD parameters.

Two-dimensional projected clustering statistics can also be computed analytically based on the three-dimensional clustering signals predicted by our emulator. Alternatively, one might be tempted to project the matter particles, halos, or mock galaxies in a simulation box along a single chosen axis or direction and then measure the correlation signals in two dimensions to model the projected signals directly. One could further increase statistics by combining the results from multiple projection directions. In contrast to this conventional approach, we would like to emphasize that our procedure, which first measures the correlation functions in three dimensions and then performs projection by numerical integration along the line of sight, is more advantageous in the sense that we automatically access the information in all of the possible two-dimensional maps obtained by projection along all of the possible different directions. We also note that in our approach, the projection width can be chosen as desired once the full three-dimensional information is available. Since redshift-space distortion can impact the projected statistics when the width is small, we implement a simple model to account for this effect in the module that computes w_{gg} .

3. Simulation Ensemble

We summarize here the basic features of the DARK QUEST simulation suite. All of the simulations presented in this paper are listed in Table 1. More detailed explanations on each of the simulation suites will be given in the subsequent subsections. We also describe details of post-processing analyses.

3.1. Simulation Design

One of the key elements for an efficient emulator is the sampling scheme of the models in a high-dimensional input parameter space. It should be designed such that the hypervolume of interest is sampled as homogeneously as possible. Indeed, Latin hypercube designs (LHDs) have been employed in previous studies to show a good performance to construct the training data for emulators (e.g., Heitmann et al. 2009). An LHD is a design achieved by first selecting a hyper-rectangle, then dividing it into a

regular lattice and selecting only one sample in every lattice interval when projected into any one dimension.

The LHDs are one of the useful techniques employed in the literature of experimental design (see Garud & Karimi 2017 for a recent review). Imposing certain conditions, an LHD can have desirable space-filling and projection properties. Because of these, they are often employed in black-box experiments, where the dependence of the outcome on input variables is completely unknown. While our situation is slightly different (i.e., the relation between inputs and outputs can be approximately modeled using fitting formulae in cosmology), LHDs have been a standard tool for the development of emulators in cosmological settings. In many cases of cosmology, one wishes to emulate a considerably large number of outputs. An experimental design highly optimized to one output can sometimes give a significantly inferior performance on other outputs. An LHD is expected to give, albeit nonoptimal, a reasonable set of samples for all the outputs similarly to black-box experiments.

We here employ a variant of LHD, called maximin distance “sliced” LHD (SLHD), developed in Ba et al. (2015). This is a technique to realize a hierarchy of maximin distance (i.e., the minimum distance between different sampling points is maximized) LHDs: the whole samples are located to construct an LHD, and they are classified into subgroups called “slices” with the same number of samples, each of which independently satisfies the conditions for an LHD. In practice, a good space-filling property (i.e., a near-maximin design) is ensured by minimizing the following quantity:

$$\Phi(\mathbf{X}_N) = \frac{1}{2} \left(\phi_{\text{all}} + \frac{1}{m} \sum_{t=1}^m \phi_t \right), \quad (22)$$

where we denote by $\mathbf{X}_N = \{\mathbf{x}^{(i)} | i = 1, \dots, N, \mathbf{x}^{(i)} \in \mathbb{R}^n\}$ the locations of the whole N samples in the n -dimensional input parameter space, and ϕ_{all} and ϕ_t , respectively, stand for the cost function for the total and the t th slice,

$$\phi_{\text{all}}(\mathbf{X}_N) = \left(\frac{2}{N(N-1)} \sum_{i,j \in \mathbf{X}_N} \frac{1}{d^r(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})} \right)^{1/r}, \quad (23)$$

$$\phi_t(\mathbf{X}_t) = \left(\frac{2}{M(M-1)} \sum_{i,j \in \mathbf{X}_t} \frac{1}{d^r(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})} \right)^{1/r}, \quad (24)$$

where $\mathbf{X}_t = \{\mathbf{x}^{(i)} | i = (t-1)M + 1, \dots, tM, \mathbf{x}^{(i)} \in \mathbb{R}^n\}$ is the sample in the t th slice with $M = N/m$ members; the quantity $d(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})$ is the distance between two samples, $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(j)}$; and we use the standard Euclidean distance, $\sqrt{|\mathbf{x}^{(i)} - \mathbf{x}^{(j)}|^2}$, for simplicity. This is equivalent to putting a uniform prior when the input parameter space is sampled. Here the minimization of ϕ_{all} or ϕ_t at the limit of $r \rightarrow \infty$ is equivalent to the maximization of the minimum distance among the design points, as the name “maximin” suggests. An optimal SLHD is achieved by minimizing the mixture of ϕ_{tot} and ϕ_t with the former upweights according to the ratio of the number of sampling points in the whole and subsamples (i.e., Equation (22)). We use the parameter $r = 15$, which is the default value in the SLHD code.

This method allows a rather flexible design of samples, unlike standard single-slice LHDs, for which splitting the

samples into a training and a validation set can ruin the desirable space-filling or projection properties before splitting. For instance, the training and validation sets are chosen from different slices in our case, both covering the parameter space homogeneously, and the sample points in the two sets are guaranteed to be reasonably far (i.e., no sample in the training set is very close to any sample in the validation set). This is crucial for a stringent validation test because the accuracy of emulation can be more objectively tested by such validation samples. We sample $N = 100$ cosmological models in total with $m = 5$ slices, each of which is composed of $M = 20$ samples in an $n =$ six-dimensional parameter space. We consider the w CDM cosmology in the parameter range of

$$\begin{aligned} 0.0211375 < \omega_b < 0.0233625, \\ 0.10782 < \omega_c < 0.13178, \\ 0.54752 < \Omega_{de} < 0.82128, \\ 2.4752 < \ln(10^{10} A_s) < 3.7128, \\ 0.916275 < n_s < 1.012725, \\ -1.2 < w < -0.8, \end{aligned} \quad (25)$$

where $\omega_b \equiv \Omega_b h^2$ and $\omega_c \equiv \Omega_c h^2$ are the physical density parameters of baryons and CDM, where $h = H_0 / (100 \text{ km s}^{-1} \text{ Mpc}^{-1})$ is the Hubble parameter, $\Omega_{de} \equiv 1 - (\omega_b + \omega_c + \omega_\nu)/h^2$ is the dark energy density parameter assuming a flat geometry of the universe, A_s and n_s are the amplitude and tilt of the primordial curvature power spectrum normalized at 0.05 Mpc^{-1} , and w is the equation-of-state parameter of dark energy. As for the neutrino density $\omega_\nu \equiv \Omega_\nu h^2$, we fix to 0.00064, corresponding to 0.06 eV for the total mass of the three mass eigenstates (see Appendix D for our approximate treatment of massive neutrinos). When computing the distance in Equations (23) and (24), we linearly rescale the range of the six cosmological parameters in Equation (25) to [0,1]. Our 100 samples are shown in Figure 2, where the samples from the same slice are depicted by the same color.

The parameter range above is centered at the best-fit cosmological parameters to the *Planck* cosmic microwave background (CMB) data (Planck Collaboration et al. 2016): $(\omega_b, \omega_c, \Omega_{de}, \ln(10^{10} A_s), n_s, w) = (0.02225, 0.1198, 0.6844, 3.094, 0.9645, -1)$. The fiducial *Planck* cosmology gives, as derived parameters, $\Omega_m = 0.3156$ (the present-day matter density parameter) and $\sigma_8 = 0.831$ (the rms linear mass density fluctuations within a top-hat sphere of radius $8 h^{-1} \text{ Mpc}$). We should note that the range of each cosmological parameter covered by our SLHD is sufficiently broad such that the simulations can cover a range of cosmological models that ongoing large-scale structure surveys can probe. The parameter range shown in Equation (25) corresponds to a change of $\pm 5\%$ for ω_b and n_s , $\pm 10\%$ for ω_c , and $\pm 20\%$ for Ω_{de} , $\ln(10^{10} A_s)$, and w from their central values, which are much larger than the constraints by Planck Collaboration et al. (2016). As most of the large-scale structure probes are sensitive to a combination of σ_8 and Ω_m , we show in the top right panel of Figure 2 the range of SLHD models in this projected parameter space. Note that the current-generation galaxy surveys have put constraints on the combination of σ_8 and Ω_m at a precision of its 95% CL region comparable with or smaller than the supported range of our emulator. However, if the best-fit model inferred from the galaxy survey is away from

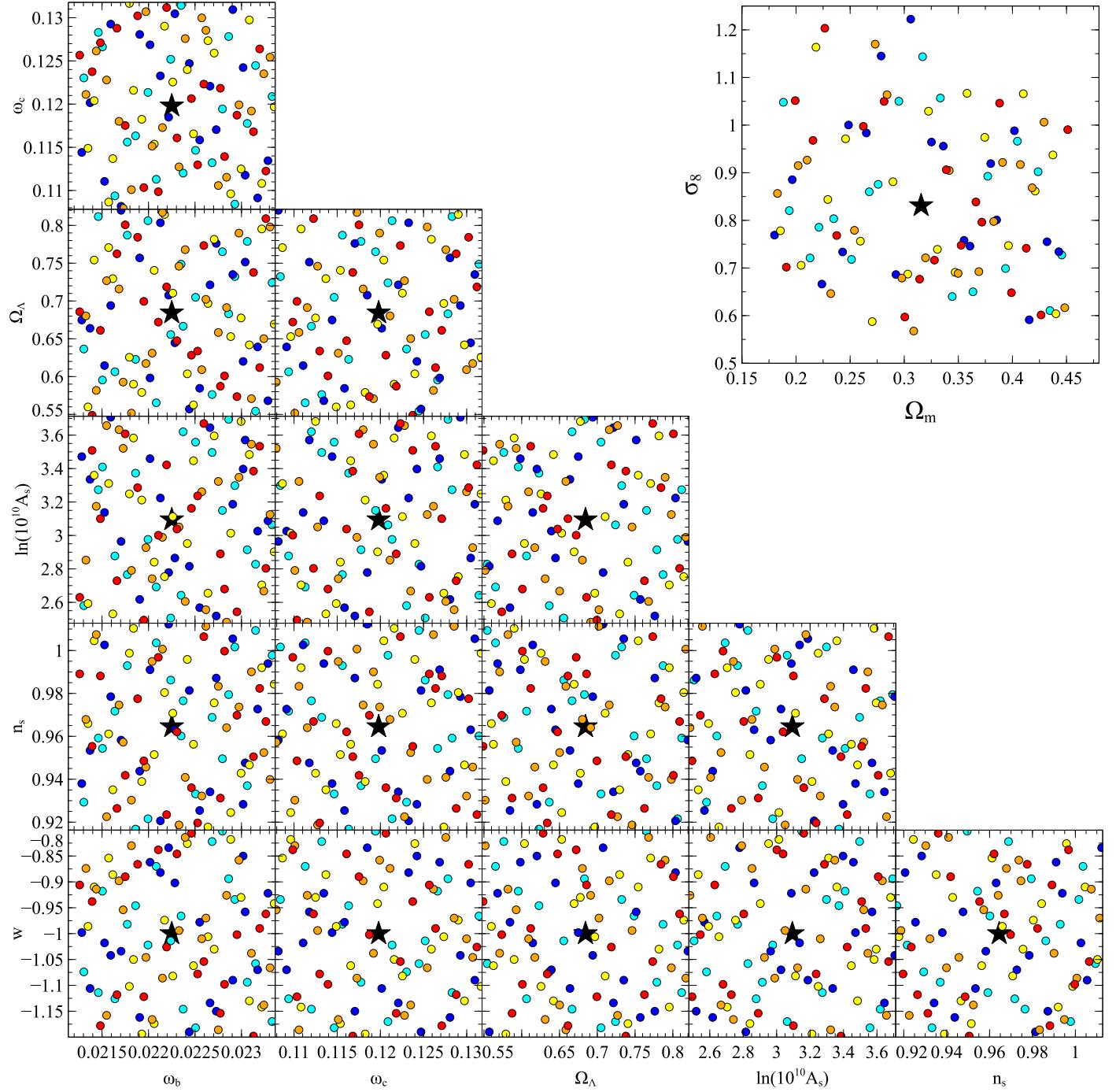


Figure 2. SLHD sampling scheme with five slices in the six-dimensional cosmological parameter space within the flat w CDM framework around the fiducial *Planck* cosmology (stars). The samples from the same slice are shown by the circles in the same color. In addition to the six varied parameters in the w CDM model, we show the projection of samples to the two-dimensional planes of the derived parameters Ω_m and σ_8 in the top right panel.

the fiducial *Planck* cosmology, the posterior region might be outside the supported region (e.g., see Hikage et al. 2019, for an example). In such cases, one needs to supply an alternative model or perform additional simulations so that the support range of our emulator can cover the range inferred from the actual data. Or, one could use an empirical approach to extrapolate the prediction outside the support range by using the halo model or other analytical method. This is beyond the scope of this paper and will be explored, if needed, in a separate paper.

3.2. Box Size and Resolution

The simulations presented here are performed with 2048^3 particles in comoving cubes with a side length of either $1 h^{-1}$ Gpc (hereafter high-resolution runs (HR)) or $2 h^{-1}$ Gpc (low-resolution runs (LR)). The mass of the simulation particle in HR (LR) simulations is 1.020×10^{10} (8.158×10^{10}) $h^{-1} M_\odot$ for the fiducial *Planck* cosmological model and varies with the value of Ω_m for different cosmological models. We perform one LR and HR simulation at every 100 SLHD sampling points. In addition, we have performed 28 (14) random realizations for

the fiducial *Planck* cosmology under the HR (LR) setting. The total volume of 28 or 112 ($h^{-1}\text{Gpc}$)³ for the HR or LR runs at the fiducial *Planck* cosmology is sufficiently large compared to the SDSS volume, which is ~ 4 ($h^{-1}\text{Gpc}$)³, corresponding to the comoving volume up to $z \sim 0.6$ over the solid angle of about 10,000 deg². In the following, we refer to each SLHD slice simply as “slice,” e.g., “slice 1.” We will use the 20 simulations in slice 5 for a cross-validation of the emulator and only the 80 simulations in slices 1–4 for the training. In addition, we run simulations with a smaller box size, 250 $h^{-1}\text{Mpc}$ for the fiducial *Planck* cosmology, with several different numbers of particles, 256³, 512³, 1024³, and 2048³, to assess a numerical convergence of our results. Note that the spatial resolution of the simulations with 512³ or 256³ particles in these small boxes is equivalent to that of the main HR or LR simulation, respectively.

As we will show later, the mass resolution of our HR simulations is sufficient to accurately estimate the HMF and halo–matter cross-correlation function in each halo mass bin down to the minimum mass of $\sim 10^{12} h^{-1} M_{\odot}$, smaller than a typical host halo mass of CMASS or LOW-Z galaxies (e.g., see Figure 4 in More et al. 2015b), where the LOW-Z galaxies roughly correspond to the SDSS luminous red galaxies (LRGs) in their figure. These simulations have already been used in Murata et al. (2018) to calibrate the mass–richness relation of the redMaPPer clusters by comparing the model predictions of stacked lensing and abundance with their measurements. In addition, the splashback features of halo edges traced by subhalos in the density and velocity space were investigated by Okumura et al. (2018b) using these simulations.

On the other hand, the LR simulations are mainly used to calibrate the halo–halo autocorrelation function, which is noisier than the halo–matter cross-correlation due to the larger shot noise, and thus the precise calibration requires bigger-box simulations. These simulations allow us to investigate large-scale phenomena: the alignment between the orientation of massive clusters and the large-scale structure surrounding them were studied in Okumura et al. (2018a) and Osato et al. (2018) using these simulations. We will show below in more detail how different statistical quantities are evaluated from these HR and LR simulations.

3.3. Initial Conditions

We generate initial conditions of individual N -body simulations using the second-order Lagrangian perturbation theory (2LPT; Scoccimarro 1998; Crocce et al. 2006) implemented by Nishimichi et al. (2009) and then parallelized in Valageas & Nishimichi (2011). We use the linear matter power spectrum computed by CAMB (Lewis et al. 2000) and generate Gaussian random fields from this spectrum. We compute displacements and velocities by 2LPT for each particle located on the regular lattice. The initial redshift is determined such that the rms displacement (at the linear order) is 25% of the mean interparticle distance in one dimension, and this depends on the box size and cosmological parameters. For the fiducial cosmological model, this condition roughly corresponds to $z = 59$ and 29 for the HR and LR runs, respectively. In Appendix B, we study how the results vary with the initial redshift, as well as how the results are altered if the Zel’dovich approximation (Zel’dovich 1970) is used instead of 2LPT to set up the initial conditions.

When we generate initial conditions for different cosmological models, we can adopt two ways regarding the randomness of the realization. The first possibility is to use the same random seed for different models as that for the fiducial *Planck* cosmology. This might be advantageous in the sense that the simulated large-scale structure shares the same randomness; thus, one can estimate how each Fourier mode grows in a different way depending on cosmological models by reducing the sample variance, i.e., the dependence of structure growth on cosmological models. Motivated by this, we perform a set of 20 simulations with a fixed random number seed for the HR simulations in slice 1. This random number seed is the same as one of the 28 realizations of the fiducial *Planck* model. However, a fixed random number seed across different cosmological models is not guaranteed to give a converged result in the final emulator in the sense that every simulation is affected by the same sample variance error that never goes away by sampling many cosmological models. By selecting a different random seed for each simulation, we hope that the sample variance will be reduced in the final results, to which the error in all of the simulations propagates in a Bayesian manner. We thus adopt varied random number seeds for the rest of our simulations. We will see in Appendix C how the emulation results can change against these difference choices of the initial seeds. The results shown in what follows are all based on the varied seed simulations, except for Appendix C.

3.4. Time Integration

Once we generate a random realization following the method described in the previous subsection, we simulate the distribution of particles using the parallel tree particle-mesh (PM) code GADGET2 (Springel 2005). We set the softening length to 5% of the mean interparticle distance in one dimension. We employ the number of fast Fourier transform (FFT) meshes that is two times larger than the number of particles in one dimension. Other configuration parameters were previously calibrated (e.g., Nishimichi et al. 2009; Valageas & Nishimichi 2011; Takahashi et al. 2012). The relevant parameters are ERRTOLINTACCURACY = 0.05 for the time-integral accuracy, MAXSIZETIMESTEP = 0.03 for the time-stepping criterion, MAXRMSDISPLACEMENTFAC = 0.25 for an additional limiter for the PM time step based on the rms particle displacement, and ERRTOLTHETA = 0.5 and ERRTOLFORCEACC = 0.001 for the tree opening criterion that controls the force accuracy. In the references above, the convergence of the matter power spectrum was intensively tested to confirm that the accuracy is better than the 1% level. Using N -body simulations with these carefully tuned parameters, Takahashi et al. (2012) provided revised parameters for the halofit formula (Smith et al. 2003). As will be shown below, the convergence of clustering signals of halos will be better once we adopt the number density-matching scheme for simulations with different spatial resolution. We thus believe that the parameters chosen to give a good accuracy on the matter power spectrum are already adequate for a calibration of halo clustering quantities without further modification.

We store outputs of each N -body realization in 21 redshift bins in the range $0 \leq z \leq 1.48$, equally stepped by the linear growth rate for the fiducial *Planck* model. They are 1.48, 1.35, 1.23, 1.12, 1.03, 0.932, 0.846, 0.765, 0.689, 0.617, 0.549, 0.484, 0.422, 0.363, 0.306, 0.251, 0.198, 0.147, 0.097, 0.048, and zero. We use the same redshifts to dump snapshots for

other cosmological models. Since the time evolution of the statistics relevant for our purpose is slow and monotonic, these 21 snapshots are sufficient to be interpolated to make a prediction at an arbitrary redshift in between.

3.5. Halo Catalogs

Since the aim of this paper is to accurately characterize halo clustering quantities, the identification of halos in each N -body simulation output is of crucial importance. There have already been comparison studies of different halo finders (e.g., Knebe et al. 2011). While halo properties appear to be relatively robust, the ability of finding substructures can differ significantly depending on which algorithm is to be used, especially near the center of halos (Pujol et al. 2014). We thus simply select probable host halos in which galaxies of interest reside and discard subhalos from our primary halo catalog when building an emulator of halo clustering quantities.

As our default choice to identify dark matter halos in each simulation output, we employ ROCKSTAR (Behroozi et al. 2013), which identifies dark matter halos and subhalos without distinction based on the clustering of N -body particles in phase space. We supplementarily use SUBFIND (Springel et al. 2001) to study the dependence of the halo statistics on the finder (see Appendix E). Throughout this paper, we adopt $M \equiv M_{200m} = 4\pi/3(R_{200m})^3(200\bar{\rho}_{m0})$ for the halo mass definition, where R_{200m} is the spherical halo boundary radius within which the mean mass density is 200 times $\bar{\rho}_{m0}$. Again note that $\bar{\rho}_{m0}$ in the above equation is due to our use of the comoving coordinate, and therefore R_{200m} is in the comoving length unit. We follow the default setting of the ROCKSTAR finder and define the center of each halo from the center-of-mass location of a subset of member particles in the inner part of the halo, which is selected to minimize the uncertainty caused by the Poisson noise and the positional dispersion of individual particles, which is larger at the outskirts. Our definition of halo mass includes all of the N -body particles within the boundary R_{200m} around the halo center (i.e., including particles even if they are not gravitationally bound by the halo). The halo mass defined in this way is more relevant for weak-lensing observables, which measure the total enclosed mass within a given aperture.

After we identify halo candidates by either ROCKSTAR or SUBFIND, we determine whether they are central or satellite halos. When the separation of two different halos (between their centers) is closer than R_{200m} of the more massive one, we mark the less massive one as a satellite halo. We keep halos with mass above $10^{12}h^{-1}M_\odot$ in the final halo catalog. The dependences of the halo clustering on the halo finder, mass definition, and central/satellite split criterion are presented in Appendix E.

3.6. Hybrid Fourier-direct Method to Measure the Correlation Signal

After simulations are done and halos are identified, we measure the clustering quantities. While correlation functions can be accurately estimated by direct pair counting, such a method can be computationally expensive due to its $\mathcal{O}(N^2)$ scaling to the number of particles N . Here we develop a hybrid method that combines the direct pair-counting method with a grid-based method that makes use of FFT. The former is used to measure the clustering signal on small scales, and the latter is

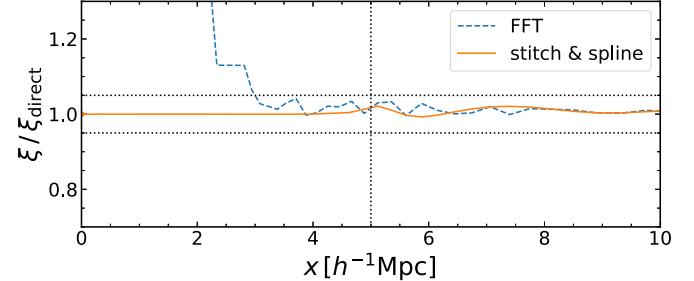


Figure 3. Accuracy of our hybrid Fourier-direct method. We plot with the dashed line the ratio of the halo–matter cross-correlation function measured with the FFT-based method to that from direct pair counting, which should give the most accurate result. Compared to the reference result based on direct pair counting, the FFT-based method shows overestimation at small pair separation. Also, it shows a noisy pattern at intermediate separations. Our final method, which combines the FFT with direct pair counting at $x = 5 h^{-1}$ Mpc (vertical dotted line) and smoothed by a cubic spline (see later discussion), is shown by the solid line. While a small residual can be seen near the switching scale of $5 h^{-1}$ Mpc, the overall behavior is within our target accuracy.

on large scales. The FFT method suffers from inaccuracy near the grid spacing but is robust for scales much larger than the grid size.

We measure the auto- and cross-correlation functions for halo–halo, halo–matter, and matter–matter pairs. We employ 1024^3 FFT grids for the large-scale signal and use the direct pair-counting method at scales below 5 or $10 h^{-1}$ Mpc for the HR or LR runs that have 1 or $2 h^{-1}$ Gpc for the box size, respectively. These switching scales roughly correspond to five times the FFT grid spacing, which is chosen so that the FFT method provides good accuracy at scales greater than the switching scale.

Figure 3 shows an example of our measurements of the halo–matter correlation function for a halo sample with mass larger than $10^{13} h^{-1} M_\odot$ at $z = 0$. For this exercise, we take one HR simulation for the fiducial *Planck* cosmology and measure the correlation function with direct pair counting up to $10 h^{-1}$ Mpc as a reference. Compared to this measurement, we show the result of the FFT-based method (dashed line). Two features can be found from the ratio. First, the FFT-based method starts to deviate from the reference rather quickly as the separation is decreased below $\sim 2 h^{-1}$ Mpc. This scale corresponds to about twice the grid size and thus simply reflects the resolution limit of FFT. Second, a noisy feature can be observed on intermediate scales up to $x \sim 8 h^{-1}$ Mpc. This is due to the discrete sampling of the pair separations (we can take only an integer vector in units of the grid spacing), together with the subtlety in the choice of the bin center, which we take as the geometric mean of the bin edges for simplicity. Note that this pattern appears to be almost the same for different random realizations and halo samples, supporting our interpretation above. To avoid the large error due to the first effect, we conservatively choose the switching scale to be $5 h^{-1}$ Mpc, indicated by the vertical dotted line. Furthermore, the stitched result is smoothed with a cubic spline function (see the next section for details) to reduce the second effect on intermediate scales while keeping the time-consuming pair-counting part to only a limited range of pair separations. Our default result that we will use in the emulator building is shown by the solid curve.

In our initial implementation, we accelerate the pair-counting method by first sorting both particles and halos into a coarse grid with 200^3 cells. We count pairs only in the same or

adjacent cells from which pairs can be closer than the matching scale. The code is then updated to employ a more sophisticated sort-tile-recursive (STR) R-tree scheme for more efficient spatial indexing (Mitsuhashi et al. 2016). Note that these different versions give identical results, and the difference lies only in the speed to perform the exact pair counting. Even with the help of these methods, the measurement of the matter-matter autocorrelation function is computationally expensive, so we cannot measure it from all the snapshots for all the models with 2048^3 particles. For this, we randomly select only 1/64 of the simulation particles in the measurement. This random selection increases the Poisson noise to the measured signal, but the typical error caused by this is not important over the scales of interest, roughly larger than $0.1 h^{-1}$ Mpc. While the main product of our emulator is the halo clustering quantities, we also provide the matter autocorrelation function for comparison, which can be used to estimate the effective bias function of halos or galaxies under consideration.

In the FFT-based measurement, we use the cloud-in-cells scheme (Hockney & Eastwood 1981) to assign matter particles or halos to each grid density estimate and then perform the FFT. We compute the product, $\delta_{1,k}\delta_{2,k}^*$, of two fields (where 1 and 2 denote halos and/or matter) at each wavenumber vector k and then Fourier transform it back to the configuration space. We take an average of the field product in each spherical shell to estimate the correlation function at the radial bin.

4. Emulation

DARK EMULATOR is constructed based on the DARK QUEST simulation suite and the analysis pipeline that was explained in the previous section. In this section, we discuss how we construct different modules, each of which predicts a statistical quantity of halos, and combine them to form DARK EMULATOR.

4.1. Overall Design

The final goal of our work is to build an N -body simulation-calibrated emulator that provides an accurate prediction of galaxy clustering quantities as a function of cosmological parameters and parameters needed to connect halos and galaxies for a given cosmological model. There are various ways to do this. One way is to adopt an *a priori* parametric prescription to connect halos and galaxies such as HOD, make mock catalogs of galaxies in each N -body simulation realization based on the assumed prescription, measure galaxy clustering quantities from the mocks, and build an emulator of galaxy quantities from the tabulated database with the model parameters in the prescription treated as the input variables in addition to the cosmological parameters. This approach was, for example, employed in Kwan et al. (2015; see also Zhai et al. 2019). However, an emulator built based on such a method may produce inaccurate results with uncertainties associated with galaxy-halo connection. For instance, there is no guarantee that a restricted HOD functional form assumed in the emulator can accurately describe the clustering properties for a sample of galaxies in a given survey. In addition, the radial profile of the satellite galaxies in a given host halo has not yet been well constrained. Furthermore, some of the central galaxies might be off-center from the true center. Therefore, variations in galaxy clustering properties cannot be

incorporated in an emulator that employs the restricted model of the halo-galaxy connection. Put another way, in this approach, it is very difficult to modify or change an emulator after its construction to include these variations and add flexibilities in the model predictions.

For this reason, we employ an alternative approach in this paper. The core function of our emulator is to predict several basic halo clustering quantities that are given as a function of cosmological parameters, halo mass, separation scale, and redshift. We will combine the “modules” analytically at the equation level, instead of using the mock catalogs, by employing a halo-galaxy connection prescription (e.g., HOD) to compute predictions of galaxy clustering quantities. The design of our emulator is illustrated in Figure 4. It is composed of three groups of modules, surrounded in the figure by rounded rectangular boxes, each of which has a number of functionalities as denoted by the text. The first group are LINEAR MODULES, which predict the statistical quantities of the linear matter perturbations (see Appendix A for details). The second group is the core part and predicts various statistical properties of dark matter halos (HALO MODULES). Finally, at the bottom of the figure, we have UTILITY MODULES, which combine the upper-level modules to compute observable quantities. The key ingredient in this group is the prescription to connect halos and galaxies (see the items in the inset). Another functionality implemented here is to compute the projected clustering quantities, such as the galaxy-galaxy weak-lensing correlation function, by directly projecting the three-dimensional correlation function along the line-of-sight direction by numerical integration. We also provide options to include possible baryonic corrections to the mass profile near the halo center, as well as redshift-space distortions (these effects will be presented in a separate paper). Although we assume a specific HOD prescription as a working example of halo-galaxy connection, a user can change it and adopt another prescription to have the galaxy clustering quantities from HALO MODULES. Thus, our method allows a flexible modification of the halo-galaxy connection without the need for additional training based on numerical realizations of mock galaxy catalogs.

4.2. Resolution Study and Matching Scheme

Because of limited numerical resources, such as memory and executive CPU time, we can run only a finite number of N -body simulation realizations, where the size of each simulation is mainly determined by the number of N -body particles. Even for a fixed number of particles, there is a trade-off between the resolution and the box size. While the former is responsible for the minimum length scale and the minimum mass of halos down to which the simulation results are accurate, the latter defines the number of Fourier modes available in each simulation and thus controls the statistical precision. The usual way to cover a wider dynamic range of predictions is to combine simulations performed in different box sizes and then stitch their results over separations or wavenumbers between neighboring box-size simulations. Indeed, such a method was used in previous works, such as Lawrence et al. (2010), Valageas & Nishimichi (2011), and Takahashi et al. (2012), where the main goal was to calibrate the matter power spectrum. An analytical model based on perturbative calculation was further combined at the large-scale limit in Lawrence et al. (2010) to suppress uncertainties due to the large

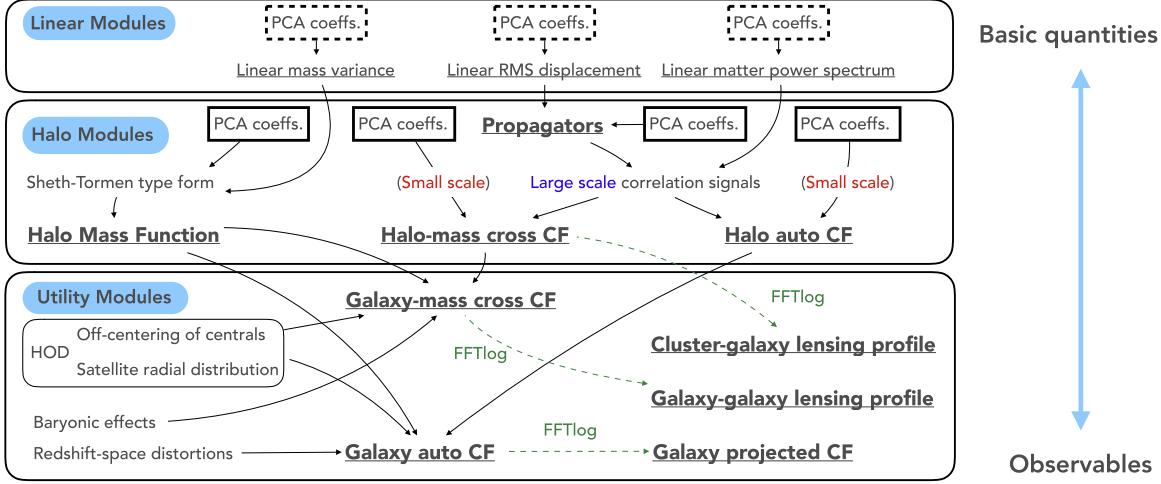


Figure 4. Layout of different modules of DARK EMULATOR. The cosmology dependences of the quantities in a square (i.e., “PCA coeffs.”) are modeled by the GP, and those underlined are physical quantities evaluated in each of the modules. The whole DARK EMULATOR code is made up of three groups of modules (enclosed by rounded rectangles). The first group of modules, shown at the top of the figure, is for linear theory quantities (LINEAR MODULES). The second group shows the modules for the abundance and clustering properties of halos (HALO MODULES). These are calibrated with a suite of N -body simulations and the core pieces of DARK EMULATOR. The modules at the bottom work on the outputs of the HALO MODULES and transform them into observable quantities (UTILITY MODULES). These mainly connect halos to galaxies using an analytical prescription and project the three-dimensional quantities onto the two-dimensional sky.

sample variance near the wavenumber corresponding to the box size.

In this subsection, we examine the numerical convergence of halo quantities using a set of simulations with different resolutions. We then discuss a strategy to combine the results of different simulations to predict the clustering signals over wider ranges of halo masses and length scales.

4.2.1. HMF

In Figure 5, we first examine the HMF for the fiducial *Planck* cosmology at three redshifts, $z = 1.02, 0.549$, and 0 , using four N -body simulations with different numerical resolutions. Note that the simulation with the lowest resolution among the four, which has 256^3 particles, has the same resolution as our main LR suite, whereas the second from the worst, with 512^3 particles, corresponds to the resolution of the HR suite. For reference, the HMF in Figure 5 is normalized by the fitting formula of Tinker et al. (2008), with a mass definition of 200 times the cosmic mean density. To have a fair comparison, we integrated the Tinker et al. (2008) HMF (hereafter Tinker HMF) over the halo masses in each mass bin, which are used when we measure the HMF from simulations.

We can see that the measured HMF better matches the Tinker HMF down to lower masses as the simulation resolution increases. The four simulations agree with each other at the high-mass end, although the curves are noisy due to the Poisson noise. Note that these simulations are done in a comoving box with a side length of $250 h^{-1}$ Mpc, which is smaller than our main simulations of 1 or $2 h^{-1}$ Gpc used for the emulator development. This suggests that our main simulations have much lower Poisson noise at such high-mass bins. The vertical arrows, from right to left for higher resolution, denote the halo mass corresponding to 100 N -body particles. The figure indicates that the simulation HMF at this mass scale is underestimated by about 10%, fairly independently of redshift. Thus, one needs at least several hundred particles to estimate the HMF to a percent accuracy.

We here propose a way to empirically correct for a systematic error in the estimated HMF due to numerical

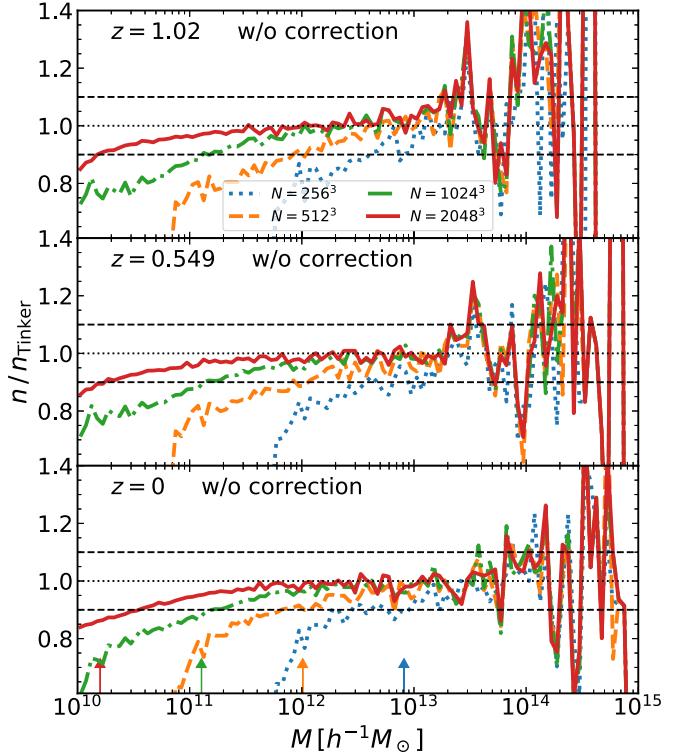


Figure 5. Resolution study for the HMF. Here we fix the simulation box size to $250 h^{-1}$ Mpc and compare the mass functions measured from simulations with different mass resolutions. We use simulations of a $1 h^{-1}$ Gpc box and 2048^3 particles (HR simulations) to build an emulator of HMF, which is equivalent, in terms of the resolution, to the 512^3 simulation in this plot. Plotted here is the ratio of the simulation result to the fitting formula in Tinker et al. (2008) at $z = 0$. The arrows on the horizontal axis denote halo mass, which corresponds to 100 particles for the halo mass definition of 200 times the cosmic mean density. The two horizontal dashed lines denote a $\pm 10\%$ fractional difference from the Tinker et al. (2008) mass function. The three panels show the ratio at $z = 1.02, 0.549$, and 0 (top to bottom).

resolution. Our method is motivated by the method in Warren et al. (2006), which was developed for halos that are identified by the friends-of-friends (FoF) method. They proposed that the

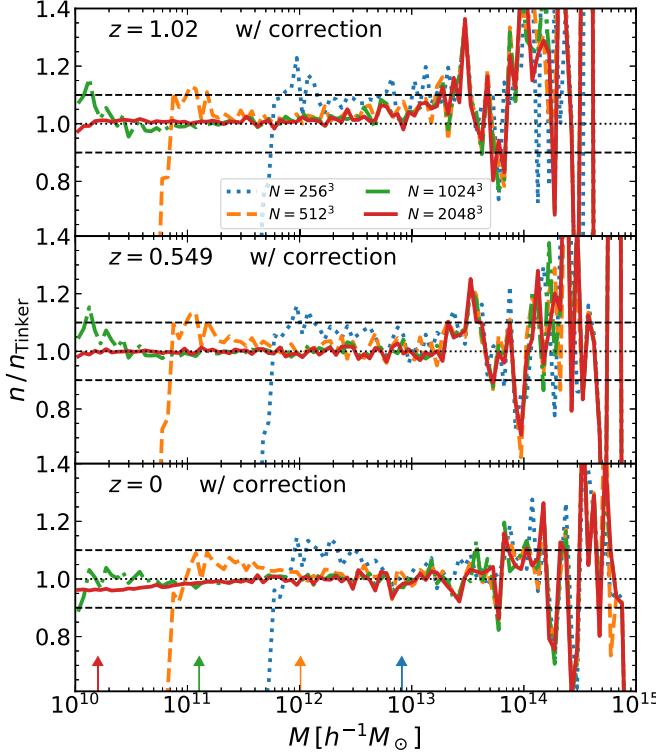


Figure 6. Similar to Figure 5, but the plot shows the mass functions when inaccuracies in individual halo masses due to limited mass resolution are corrected for according to Equation (27).

FoF mass of each halo is calibrated as

$$\tilde{M} = (1 - N_p^{-0.6})M, \quad (26)$$

where $N_p = M/m_p$ is the number of member particles, m_p is the N -body particle mass, and \tilde{M} is the corrected mass. Since the FoF algorithm tends to link physically unbound particles near the halo boundary when the mass resolution is poor, an FoF halo mass tends to be overestimated compared to the true mass. Hence, the FoF-based HMF tends to be overestimated for low halo masses that are affected by numerical resolution. The correction factor is applied to each FoF halo in such a way that the FoF mass is reduced to correct for the overestimation in HMF. This procedure was further confirmed in Crocce et al. (2010), where the method was applied to FoF halos in MICE simulations. On the other hand, our result in Figure 5 displays a rather opposite trend: our HMF is underestimated in low-resolution simulations, implying that the mass of each low-mass halo tends to be underestimated. Since we use the SO mass, the SO-based HMF is affected by the matter density field around the halo region, which tends to be underestimated in low-resolution simulations. This is different from the FoF finder, and thus the opposite trend is understandable.

We thus use the following equation to correct for the SO mass:

$$\tilde{M} = (1 + N_p^{-0.55})M, \quad (27)$$

where N_p is the number of particles within R_{200m} in the SO mass definition. We employ a slightly different power of N_p from Equation (26) to get a better calibration. After this correction, the HMFs from different resolution simulations

better agree with each other down to the resolution limit denoted by the vertical arrows, as shown in Figure 6. Below the mass limit, the SO halo mass becomes overcorrected, yielding an overestimation in HMF. These trends after the correction appear to be very similar at different redshifts. While a further refinement of the empirical function given by Equation (27) would be possible in principle, we do not use halos with less than 200 particles when we calibrate HMF. Note that we employed a slightly conservative threshold of 200 particles, as compared to the case of 100 particles discussed in Figure 6. Furthermore, we use only the HR simulations (resolution equivalent to the one with 512^3 particles here) for development of the HMF emulator. We can determine HMF accurately down to $\sim 10^{12} h^{-1} M_\odot$, and this number varies depending on the cosmological model, as we will discuss below.

4.2.2. Correlation Functions

Next, we check the clustering correlation functions of halos. In doing so, we need to consider subsamples of halos divided by halo discriminators, such as halo mass, and then consider the clustering correlations as a function of different subsamples. In the left plot of each panel in Figures 7 and 8, we show the halo–matter cross- and halo–halo autocorrelation functions for a mass threshold sample of halos with $M \geq 10^{12} h^{-1}$ or $5 \times 10^{12} h^{-1} M_\odot$ measured from the four simulations as in Figure 5. Here the threshold mass $10^{12} h^{-1} M_\odot$ corresponds to halos with more than 100 member particles for simulations with more than 512^3 particles, whereas it corresponds to halos with only ~ 10 particles for the 256^3 simulation. Note that we did not apply the correction from Equation (27) for halo masses in these figures. While all of the correlation functions agree with each other at large separations, the smaller scales clearly show the effect of numerical resolution; the measurements from a lower-resolution simulation start to deviate from those from a higher-resolution simulation on scales smaller than $\sim 1 h^{-1}$ Mpc. The comparison of Figures 7 and 8 reveals that the deviation is larger for a halo sample of smaller mass threshold. The inaccuracy is ascribed to several facts. In a lower-resolution simulation, halo masses around the mass threshold are not determined accurately on an individual halo basis due to the lack of numerical resolution, as discussed in Figure 5. Thus, the halo sample of a given mass threshold becomes different from that of a higher-resolution simulation. Moreover, the mass distribution around each halo in a lower-resolution simulation is simulated less accurately.

In this paper, we employ a slightly different sample of halos to develop the emulator. Rather than using the mass as the primary proxy of the different clustering strengths of halos, we consider mass threshold samples and label each sample in terms of the number density of halos above the threshold. We expect two advantages from this conversion. First, the cosmology dependence of the noise level of various statistics is weaker compared to the samples labeled by the mass. Indeed, we know that the mass of the heaviest halos available in each simulation can be quite different among different cosmological models and at different redshifts. Second, as is quite obvious from Figure 5, the masses inferred from simulations are quite sensitive to the resolution, especially at the low-mass end. To see this more qualitatively, we show in the right plot of each panel in Figures 7 and 8 the clustering signals for the mass threshold samples with a fixed number density in each

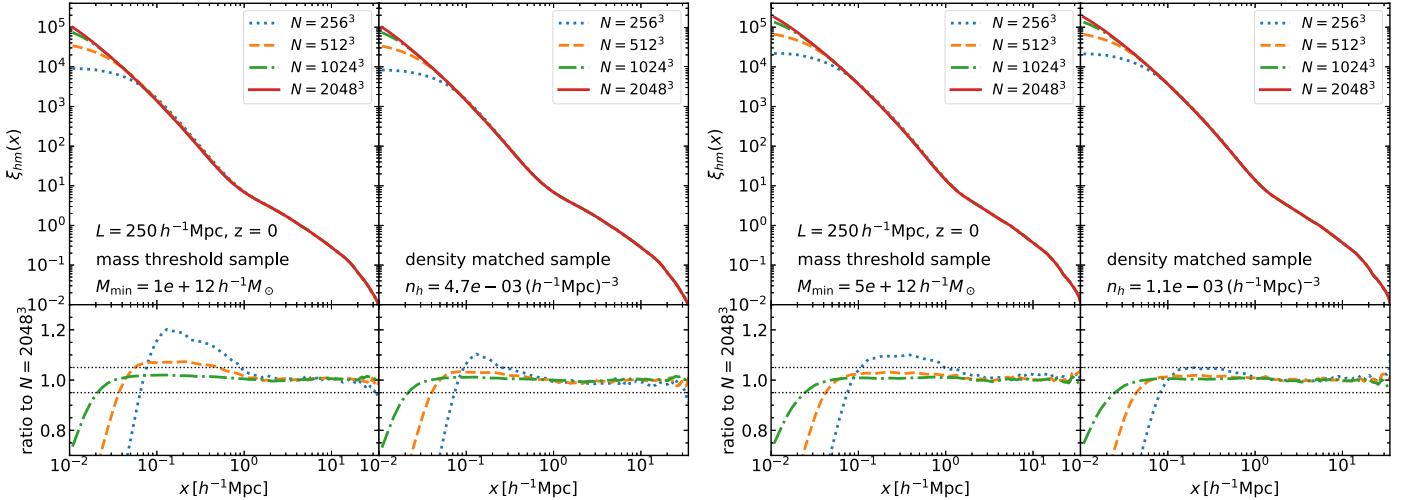


Figure 7. Resolution study for the halo–matter cross-correlation function using simulations with different mass/spatial resolutions in a small box ($250 h^{-1}$ Mpc per side), as in Figure 5. For the module of the halo–matter cross-correlation, we use the HR simulations ($1 h^{-1}$ Gpc box size and 2048^3 particles), which are equivalent to the 512^3 simulations in this plot. We show the results for the mass threshold samples in the two left panels, while in the right panels, the halo samples are chosen so as to have an equal number density above the mass threshold. We show in the bottom panels the ratio to the simulation with 2048^3 particles as a reference.

simulation. Here the number density is the same as that of the mass threshold sample for the highest-resolution simulation (2048^3) in the left plot, and the mass thresholds for other simulations are determined to match the target number density. Now an agreement between different resolution simulations is better than in the left panel, reflecting the fact that the number of halos in the sample is less affected by numerical resolution compared to the mass threshold sample. Nevertheless, the lowest-resolution simulation still exhibits a relatively large deviation for the halo–matter cross-correlation at small scales, especially for the sample corresponding to $10^{12} h^{-1} M_\odot$, because the matter distribution in high-density regions is less accurately simulated in such a low-resolution simulation. To avoid this inaccuracy, we use only the HR simulations to estimate the halo–matter cross-correlation function for different cosmological models.

On the other hand, Figure 8 shows a slightly better agreement among the four simulations with different resolutions, implying that the halo–halo autocorrelation is relatively robust against the numerical resolution. Note that larger scatters in the ratio on small scales ($\lesssim 1 h^{-1} \text{Mpc}$) are due to the halo exclusion effect, which states that the correlation signal is sharply suppressed on small scales due to the fact that no halo pair can exist below R_{200} radius of the larger one by construction in our halo sample. Thus, a slight miscalculation of the halo radii due to numerical resolution can lead to a large error in the correlation signal at a fixed scale around the typical R_{200} of the sample. Since our final product is the galaxy correlation function, and the one-halo term gives a dominant contribution around these scales, the scatter seen here does not largely affect the predictions of the galaxy autocorrelation function, as we will show later. Based on these results, we use the LR simulations to estimate the autocorrelation functions for different cosmological models.

In summary, we use the correlation functions of halos measured for halo samples with different number densities. When this is combined with the HMF module that gives the halo number density as a function of mass, one can compute the halo correlation function for a given mass threshold instead of the number density. Furthermore, we can compute the

correlation function of halos in an infinitesimally narrow mass bin by taking the numerical derivative of the correlation functions for a mass threshold halo sample with respect to the threshold mass.

4.2.3. Large-scale Limit

The clustering correlation functions of halos measured from each of the simulations become considerably noisy on very large scales around the baryon acoustic oscillation (BAO) scale due to the significant sample variance due to the finite simulation volume, even for our LR simulations of $2 h^{-1}$ Gpc size. To overcome this obstacle, we employ a semi-analytical approach based on the propagator (e.g., Crocce & Scoccimarro 2006a), which captures most of the expected linear and nonlinear effects around the BAO scale. We then stitch the prediction with the direct simulation results to obtain model predictions over a wide range of scales, as described below.

Figures 9–11 show the matter auto-, halo–matter cross-, and halo autocorrelation functions on large scales, respectively. The solid curves in the upper left panel of each figure depict the correlation functions measured from each of the 14 LR realizations for the fiducial *Planck* cosmology. Clearly, the realization-to-realization scatter is large. For comparison, the upper right panels show the linear theory predictions we computed using the same Gaussian random realizations as in the initial conditions of each simulation in order to properly take into account the sample variance effect (see Section 4.3.4 for our method to determine the linear bias parameter from the halo correlation functions). The scatter among the realizations seen in the linear predictions is comparable to that of the corresponding nonlinear counterparts, except for the halo–halo autocorrelation function with a low number density of $10^{-5} (h^{-1} \text{Mpc})^{-3}$ (i.e., the right panels of Figure 11). This suggests that the primary source of the scatter is indeed the sample variance in the initial conditions, and the shot noise adds only a moderate scatter for low-density samples of halos.

Since our varied cosmology simulation suite is, in principle, performed only once at each model, the large scatters in the measured correlation function make it difficult to construct an

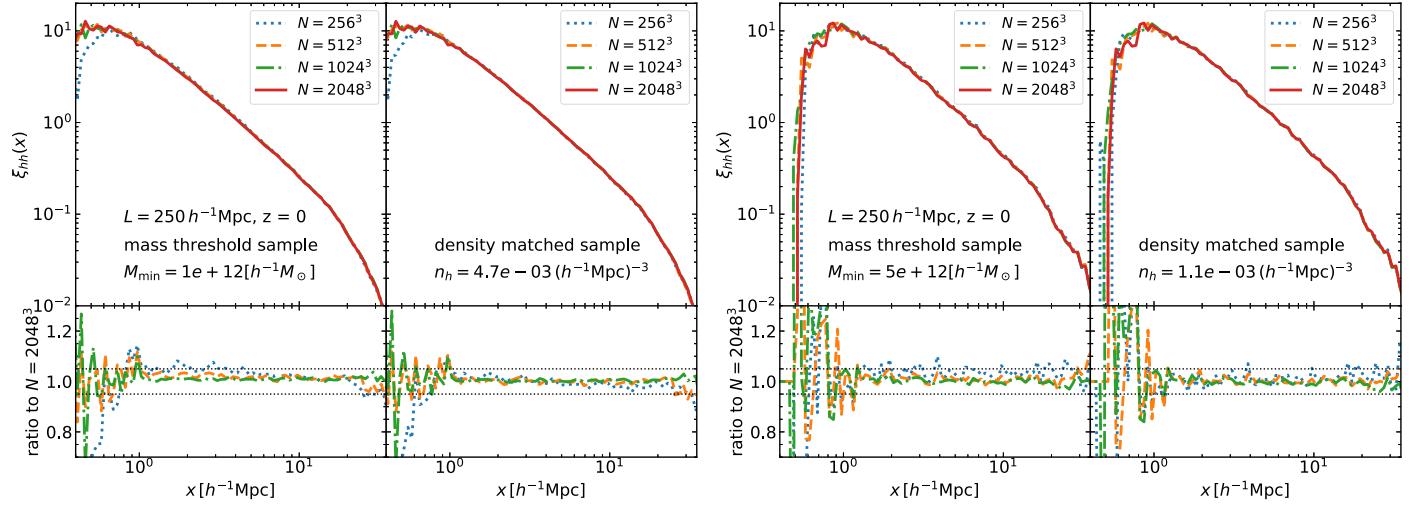


Figure 8. Resolution study for the halo–halo autocorrelation function from simulations with different mass/spatial resolutions, as in the previous figure. For the module that computes the halo–halo correlation, we use simulations of $2 h^{-1}$ Gpc box size and 2048^3 particles (hereafter LR simulations) that are equivalent to the 256^3 simulations in this plot.

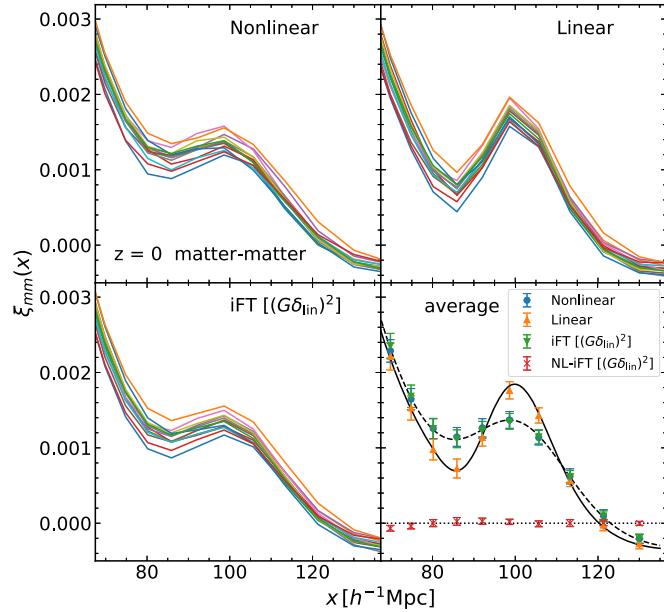


Figure 9. Matter autocorrelation function around the BAO scale. We show in the upper left panel the correlation function measured from the 14 LR simulations for the fiducial *Planck* cosmology at $z = 0$. The upper right panel shows the linear correlation function for the Gaussian random realizations that correspond to the initial conditions used in the simulations of the upper left panel. The lower left panel shows the results computed by taking the inverse Fourier transform (IFT) of the product of the propagator and the linear power spectrum, referred to as $iFT[G_m^2(k)P_{lin}(k)]$ here, for the same random realizations (see text for details). Finally, the lower right panel shows the average of the curves in the other panels over the 14 realizations. The crosses with error bars show the difference between the full nonlinear curves and the propagator-based model for the random realizations considered here. The solid and dashed curves denote analytical calculations for the linear theory and propagator model, respectively.

accurate emulator. Unlike the matter power spectrum, we cannot switch to a parameter-free perturbative calculation on these large scales because we have to know the halo bias that is not accurately described by a simple analytical prescription that often ignores the dependence on scale and cosmology. We need an appropriate method where the sample variance is sufficiently reduced and, at the same time, the large-scale bias of the halos under consideration is properly taken into account.

Another important effect on the BAO scale, in addition to the large-scale bias, is the damping of the BAO feature. This is clearly visible from comparison of the upper left and right panels in Figures 9–11. It is known that this effect is to a large extent due to the large-scale bulk motion of the cosmic fluid, which can be accurately modeled by the propagator (Crocce & Scoccimarro 2006a). In their paper, the propagator for the matter field is defined by

$$\left\langle \frac{\partial \Phi_{a,k}}{\partial \delta_{lin,k'}} \right\rangle \equiv \delta_D^3(\mathbf{k} - \mathbf{k}') G_a(k), \quad (28)$$

where Φ_a can be either the density or the velocity divergence field of matter. Note here and in what follows that the linear density field δ_{lin} and its power spectrum P_{lin} are always scaled by the linear growth factor to the same redshift as other quantities, such as Φ_a or G_a . The function $G_a(k)$ is called the (two-point) propagator, which shows a damping form very close to a Gaussian shape toward high k . This function can be interpreted to describe how much memory of the initial density field (δ_{lin}) persists in the final (nonlinear) fields (Φ_a). One can analytically show that this function is exactly a Gaussian with its variance equal to the inverse square of the rms displacement field in the case of the Zel'dovich dynamics for a Gaussian initial condition.

In most of the resummed perturbation theories, the leading-order contribution to the mixed power spectrum of two fields δ_a and δ_b is expressed as $G_a(k)G_b(k)P_{lin}(k)$, where the subscripts a and b can be the density or velocity divergence of matter or any tracers (e.g., Crocce & Scoccimarro 2006b, 2008; Bernardeau et al. 2008) and $P_{lin}(k)$ is the linear matter power spectrum. The IFT of this combination gives a reasonable prescription on the two-point correlation function around the BAO scale,

$$\xi_{a,b,\text{tree}}(r) = iFT[G_a(k)G_b(k)P_{lin}(k)], \quad (29)$$

where we use the subscript “tree” to indicate that this quantity is the tree-level result (i.e., the leading-order diagrams) of the resummed perturbation theories. Indeed, Equation (29) with a simple Gaussian approximation of the propagator can already

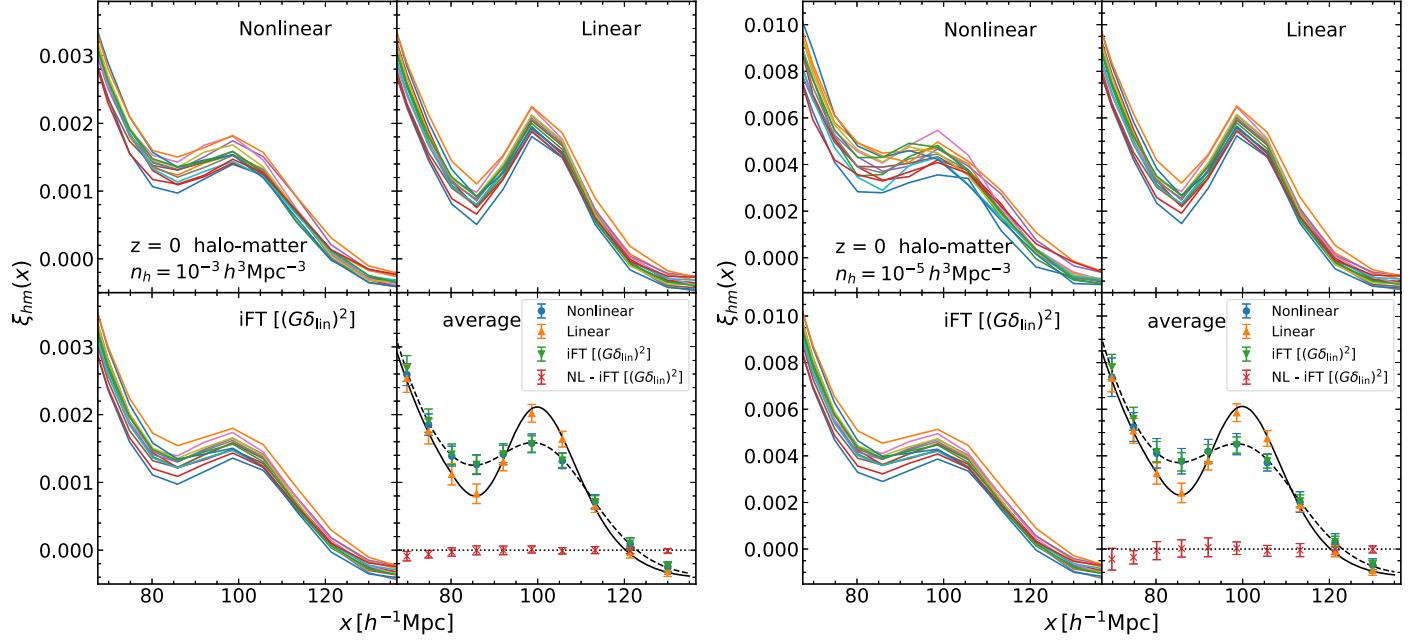


Figure 10. Similar to Figure 9, but for the halo-matter cross-correlation functions. Here we show the results for halo samples with number densities of 10^{-3} and $10^{-5} (h^{-1}\text{Mpc})^{-3}$ in the left and right panel, respectively.

explain the damping of the BAO peak in the matter correlation function very accurately (e.g., Matsubara 2008). By going into higher orders, a subpercent-level shift in the BAO peak location to a smaller separation scale can be realized (Crocce & Scoccimarro 2008). This will be important in interpreting the BAO-related distance measurements from actual observations.

We now consider the propagator for halos. One can define the propagator by simply replacing Φ_a with the density field of halos in a given sample. In what follows, we denote this by $G_a(k)$, where the subscript a is either the matter or the halo density field. In case of halos, the low- k limit of the function corresponds to the linear bias factor. A damping behavior at high k should be similar to that of the matter field, and this damping is responsible for the smearing of the BAO peak measured through the clustering of halos. We show the functions for matter and halos with different number densities and at different redshifts in Figure 12, and they are measured from the 14 LR realizations of the fiducial *Planck* cosmology. We estimate the function by taking

$$G_a(k) = \frac{P_{a,\text{lin}}(k)}{P_{\text{lin}}(k)}, \quad (30)$$

where $P_{a,\text{lin}}(k)$ is the cross-power spectrum of tracers “ a ” and the linear density field. In this estimator, we use the linear power spectrum $P_{\text{lin}}(k)$ measured from the linear density field used for the initial condition instead of the theoretical smooth function, and the sample variance is largely suppressed by taking this ratio. Indeed, the scatter among the 14 realizations seen in the figure is rather small, especially for low number density halo samples compared to the scatter in the corresponding autocorrelation function in Figure 11. The overall trend of this function looks very simple, as already discussed; it appears to be a Gaussian-like damping function with a linear bias factor at the low- k limit that depends on the halo number density. In addition, we can see that the damping starts at

smaller k at lower redshifts, reflecting the fact that the information in the initial density field remains more on larger scales and at higher redshifts.

To summarize, our strategy to describe the large-scale limit of matter or halo correlation functions is to emulate the function $G_a(k)$ for both matter and halo fields and substitute it into Equation (29). Likewise, we take the same combination for the random fields δ_{lin} used in the initial conditions, which we schematically denote as $\text{iFT}[(G\delta_{\text{lin}})^2]$. We show this model in the lower left panel of Figures 9–11 for the random realizations corresponding to the 14 simulations shown in the upper panels. The curves obtained in this way appear to be very similar to the direct simulation results in the upper left panels.

Finally, the averages of these curves are shown by the downward-pointing triangles with error bars in the lower right panel. They are almost indistinguishable from the circles for the nonlinear correlation function directly measured from the nonlinear fields. Indeed, their differences, shown by the crosses, are consistent with zero. A closer look at the scale dependence of this residual indicates a small pattern that would cause a small shift on the BAO peak toward a smaller scale. It tends to be positive around the inflection point of the correlation function (at around $90 h^{-1}$ Mpc) and negative at scales smaller than $80 h^{-1}$ Mpc. Since in most cases, these features are within the error bars, which correspond to the scatter among realizations, we simply ignore this small residual in the following discussion.

We also show the continuous limit of the model, Equation (29), by the dashed line. This is the expectation value of the downward-pointing triangles in the limit of an infinite number of realizations. Our final model for the large-scale correlation function is this line. With this procedure, we can reduce the sample variance significantly, since the prediction is based on the noiseless linear power spectrum P_{lin} . Our approach works well even in the case of the halo autocorrelation function for a halo sample with a small number density (see the upper left panel in the right part of Figure 11);

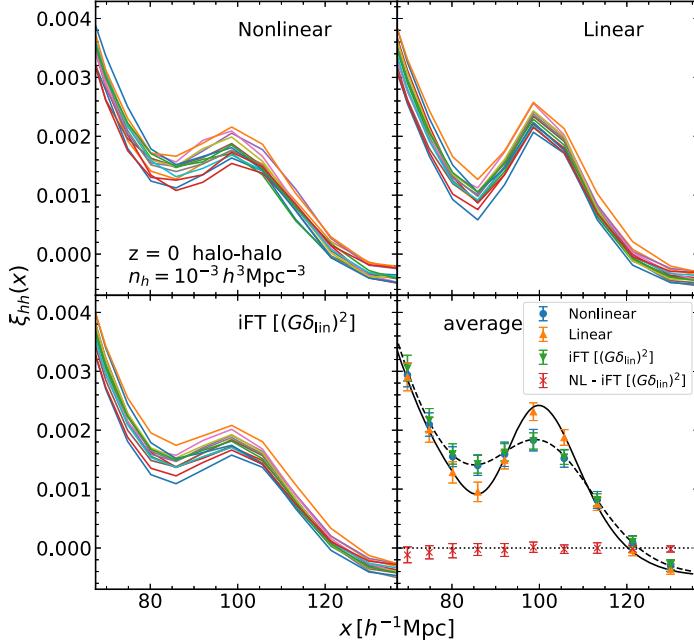


Figure 11. Similar to Figure 10, but for the halo–halo autocorrelation functions for the two number density–selected samples, as before.

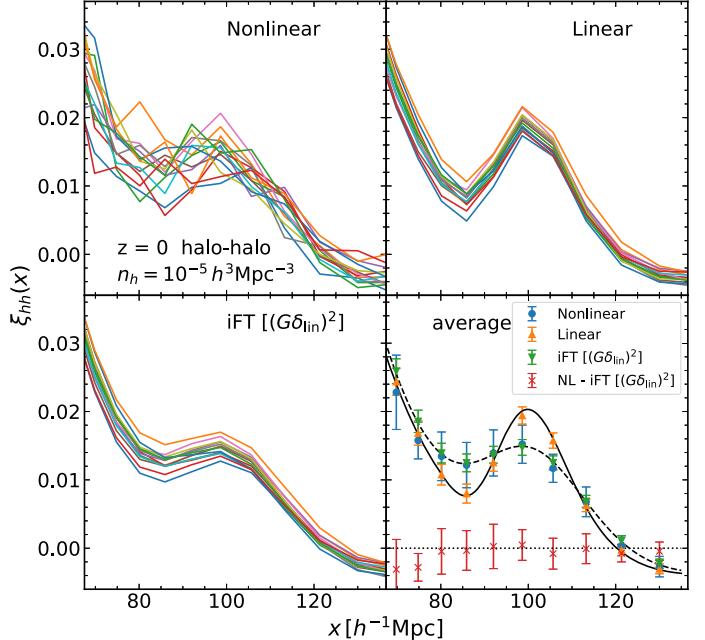
the unaccounted shot-noise effect only adds a random scatter, and no systematic trend can be seen in the residual. We will explain how we switch from the direct measurement of the correlation functions to the prescription based on the propagator explained here in Section 5.1.1.

4.3. Implementation Detail and Performance

We have so far described the building blocks of our DARK EMULATORS. Below, we describe how to model their cosmological dependences.

Our basic strategy for emulator development is as follows. First, we build a data vector for each of the four main halo functions (HMF, halo–matter cross-correlation, halo–halo autocorrelation, and propagator), including their dependence on redshift, separation, and number density, which can be translated into the halo mass threshold, from simulation realizations of each cosmological model. Second, we apply PCA to the data vector, which allows for a huge dimensionality reduction of the data vector by keeping only a handful of the most significant principal component (PC) coefficients (see also Lawrence et al. 2010, for a similar method for the matter power spectrum). In doing so, we use a public PCA package, EMPCA (Bailey 2012), which allows us to introduce a weighting to the input data vector. An advantage of this weighting method is that we can put a zero weight to missing data. Third, we apply GPR to the significant PC coefficients for different cosmological models in order to have a quick GP interpolation of the model prediction of each of the halo functions in an arbitrary cosmological model. As for the GPR, we use a public code, GEORGE (Ambikasaran et al. 2015). We adopt a stationary kernel function with either EXP-SQUARED, EXP, MATERN32, or MATERN52 and pick one for each PC coefficient based on the likelihood of explaining the data after optimization.

In building the emulator, we use multiple realizations for the fiducial *Planck* cosmology to estimate errors in the PC coefficients. Assuming that the errors are independent of



cosmology, we add the errors in square into the diagonal components of the GP kernel function. Unless otherwise stated, we use 80 simulations in slices 1–4 from either the HR or the LR suite. The remaining 20 models in slice 5, as well as the fiducial *Planck* model, are used for a cross-validation of the emulator outputs.

We describe the details of the actual implementation of the four main halo modules in the following subsections. The connection to the galaxy statistics will be explained in the subsequent section.

4.3.1. HMF

In this section, we describe how to build a module of the HMF. As shown in Figures 5 and 6, the fitting formula by Tinker et al. (2008) works very well, at least for the fiducial cosmology at $z = 0$. The fitting function we use in the following is a modified version of the earlier model in Press & Schechter (1974; see also Sheth & Tormen 2002), given by

$$\frac{dn}{dM} = f(\sigma_M) \frac{\bar{\rho}_m}{M} \frac{d \ln \sigma_M^{-1}}{dM}, \quad (31)$$

with

$$f(\sigma_M) = A \left[\left(\frac{\sigma_M}{b} \right)^a + 1 \right] \exp \left(-\frac{c}{\sigma_M^2} \right). \quad (32)$$

Here the mass variance σ_M^2 is given by

$$\sigma_M^2 = \int \frac{k^2 dk}{2\pi^2} P_{lin}(k; z) |\tilde{W}_R(k)|^2, \quad (33)$$

where $P_{lin}(k; z)$ is the linear matter power spectrum at redshift z , and $\tilde{W}_R(k)$ is the Fourier transform of a top-hat filter of radius R that is specified by an input halo mass M via $R = (3M/4\pi\bar{\rho}_{m,0})^{1/3}$. Tinker et al. (2008) showed that the HMF measured in the simulations is well fitted by the above

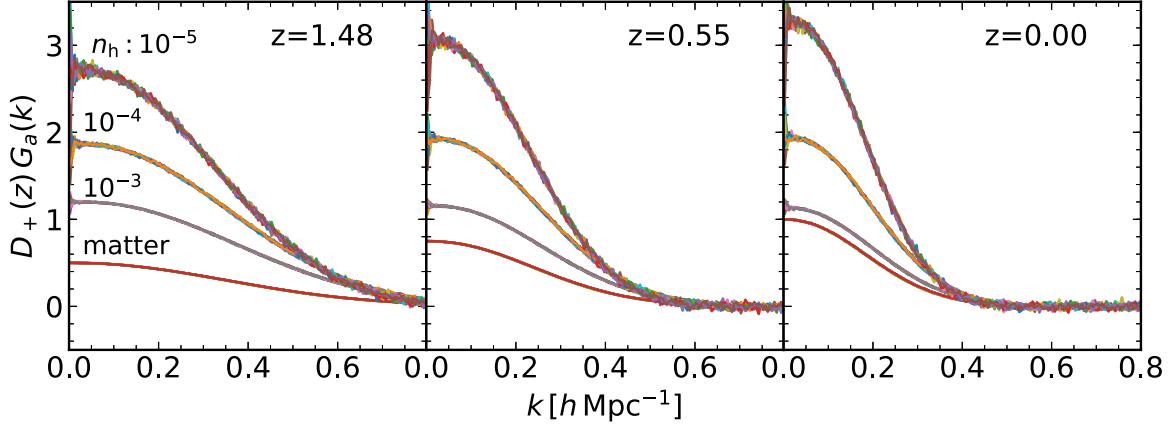


Figure 12. Propagator for the matter and halo density fields (see Equation (28) for the definition of the propagator). We here consider the halo samples with number densities $n_h = 10^{-5}, 10^{-4}$, and $10^{-3}(h^{-1}\text{Mpc})^{-3}$ and show the results at redshifts $z = 1.48, 0.55$, and 0 in the left, middle, and right panels, respectively. We multiply the linear growth factor $D_+(z)$ to reduce the dynamic range.

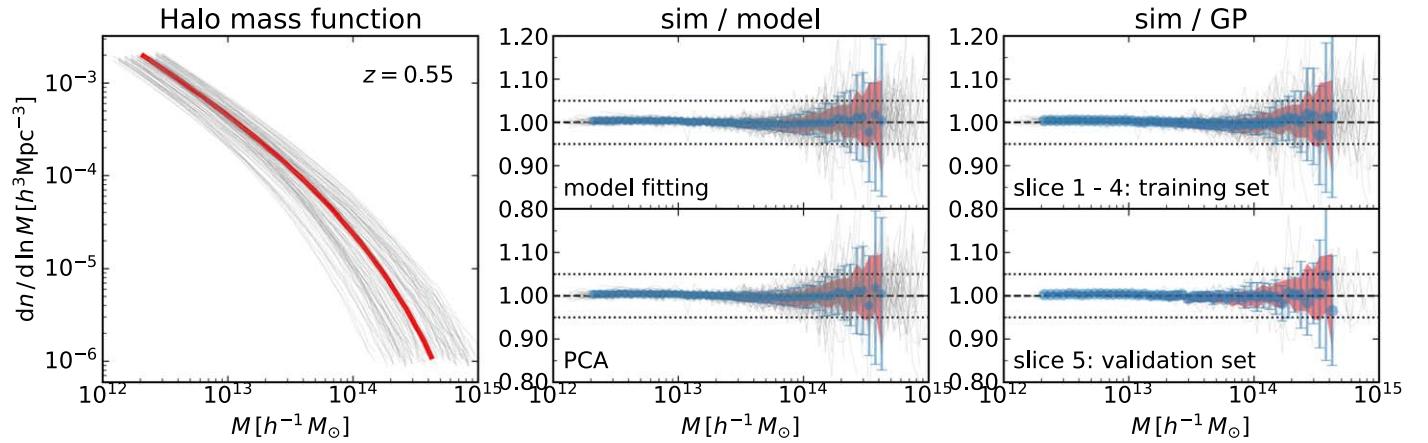


Figure 13. Modeling of the HMF. Left panel: variations in HMF at $z = 0.55$, which are measured from each simulation of 100 cosmological models in the HR simulation suite (each simulation has $1 h^{-1} \text{Gpc}$ on side). The red curve shows the HMF for the fiducial *Planck* cosmology. In the upper middle panel, we model the HMF in each simulation by a functional form of Tinker et al. (2008; Equation (32)), where we estimated the best-fitting parameters of A and a to the simulated HMF but used the same b and c in Equations (36) and (37). Each gray curve is the ratio of the simulated HMF to the best-fit Tinker HMF for each of 100 cosmological models. The point and error bar at each mass bin in this and the following panels denote the mean and scatter of the ratios at the mass bin. The shaded region in this and other plots denotes the statistical uncertainties in the HMF that are estimated from scatters of HMFs in the 28 realizations of *Planck* cosmology. The horizontal dotted lines denote $\pm 5\%$ in the fractional difference. In the lower middle panel, to model the redshift and mass dependence of the HMF in each cosmological model, we performed PCA to the best-fitting Tinker parameters, A and a , at each of 21 output redshifts over the range $0 < z < 1.48$, hence 42 data points in each cosmological model (see text for details). The plot shows that keeping the six most significant PC coefficients gives almost identical accuracy as compared to the results after the model fitting (upper panel). The loss of accuracy induced in this procedure is less than 1% in all cases. In the upper right panel, we perform the GPR to the PC coefficients at 80 sampling points in slices 1–4 in six-dimensional cosmological parameter space. The lower right panel is a validation test of the GP interpolation, i.e., our HMF emulator module, showing how the GP interpolation can reproduce the simulated HMF in each of 20 cosmological models in slice 5, which are not used in the GPR.

functional form with time-dependent coefficients:

$$A(z) = 0.186(1 + z)^{-0.14}, \quad (34)$$

$$a(z) = 1.47(1 + z)^{-0.06}, \quad (35)$$

$$b(z) = 2.57(1 + z)^{-\alpha}, \quad (36)$$

$$c(z) = 1.19, \quad (37)$$

$$\alpha = -\left(\frac{0.75}{\log_{10}(\Delta/75)}\right)^{1.2}. \quad (38)$$

The overdensity Δ is 200 in our halo mass definition.

The variation in the HMF over our 100 cosmological models can be found in the left panel of Figure 13 (gray curves). We employ the HR suite here and choose to show the HMF at $z = 0.55$ as a typical redshift of the CMASS galaxies. We also

show (red curve) the HMF for the fiducial *Planck* cosmology (the mean of 28 realizations). The 100 models are taken from the five SLHD slices (hereafter slice 1, 2, ..., 5), each of which consists of 20 different cosmological models, as described in Section 3.1. The variation in the HMF is quite large, and it is not so obvious whether or not the universal form of Equation (32) can explain it.

Before constructing an HMF module, we first test the accuracy of the Tinker HMF against our simulation suite (HR runs). Figure 14 compares the simulated HMF with the original Tinker HMF prediction (using the coefficients in Equations (34)–(37)) for each of 20 cosmological models in slice 5 at three redshift bins, $z = 0, 0.55$, and 1.48 . The comparison indicates a larger deviation of the Tinker formula from the simulation results as redshift increases. At $z = 0$, the ratio of our HR simulation suite to the

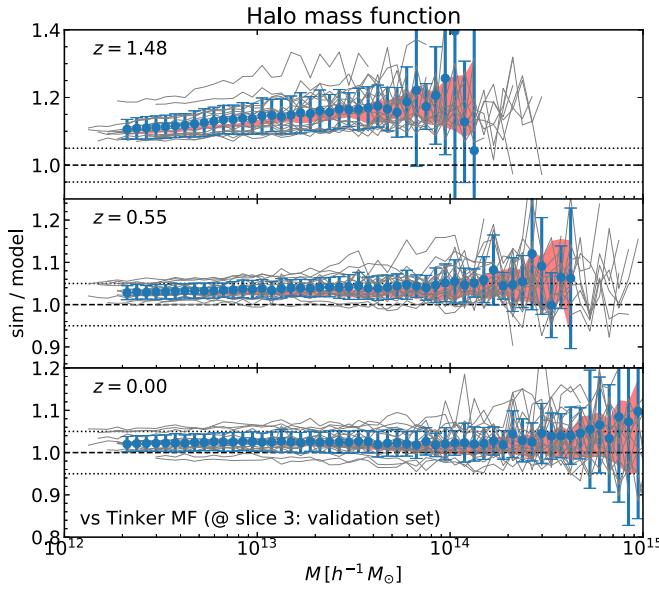


Figure 14. Comparison of our simulation HMF with the original Tinker fitting formula (Tinker et al. 2008).

Tinker MF typically shows 5% scatter from the Tinker MF, with the mean slightly biased from unity (by $\sim 2\%-3\%$). The overall slope of the HMF is already very well captured by this formula, and the error is mostly on the amplitude. At higher redshifts, both the bias and the scatter grow. At $z = 1.48$, the slope of the ratio shows a clear mass dependence with a larger bias toward the massive end. We leave further discussion on the inaccuracy of the original Tinker HMF formula to Appendix B, where we discuss that this discrepancy is mainly due to the fact that the Tinker HMF was calibrated against simulations using initial conditions based on the Zel'dovich approximation at an initial redshift that is not high enough.

From this exercise, we decide to update some of the parameters in the Tinker HMF. We drop the assumption of the HMF universality and allow the parameters A and a to vary as a function of cosmological models. For the parameters b and c , we keep the values given by Equations (36) and (37), where b determines the slope of the HMF at the low-mass end and c determines the cutoff at the high-mass end. Our HR suite is most accurate over the intermediate range of halo masses, where the overall amplitude and the slope are controlled by the parameters A and a , respectively. Therefore, we recalibrate these parameters for each of our simulations.

The upper middle panel of Figure 13 addresses how the functional form given by Equation (32) can fit the simulated HMF for each of 100 cosmological models with the updated parameters. Here we estimate the best-fitting parameters of a and A that reproduce the simulated HMF. In the fitting, we consider two sources of errors. For high-mass bins, we consider statistical uncertainties due to the Poisson noise of the number counts. We also include a phenomenological penalty term at the low-mass bins, where the correction (Equation (27)) plays a significant role. That is, in the model fitting, we include the following statistical errors in the number counts of halos at each mass bin:

$$\frac{\Delta N_h}{N_h} = \frac{1}{\sqrt{N_h}} + \frac{1}{N_p}, \quad (39)$$

where N_h is the number of halos at each mass bin, and N_p is the number of N -body member particles, M/m_p , at the logarithmic bin center. Moreover, we do not include any mass bin of halos that are defined by $N_p < 200$. The figure shows that our model HMF generally gives a very good fit to the simulated HMF for each of the 100 cosmological models, where the ratio is very close to unity well within the $\pm 5\%$ accuracy denoted by the horizontal dotted lines. At the high-mass end, the simulation data points are dominated by Poisson noise due to a too-small number of halos per bin. The circle and error bar at each mass bin are the average and scatter in the ratios of the best-fit HMF to the simulated HMF for the 100 cosmological models. For comparison, the red shaded region denotes scatters among the 28 realizations of *Planck* cosmology, giving an estimate of the sample variance for a volume of $1 (h^{-1} \text{ Gpc})^3$. The typical accuracy of the model as indicated by the error bars is $\sim 1\%$ (3%) at 10^{13} (10^{14}) $h^{-1} M_\odot$.

We then compress the data vector, $\mathbf{d} = (A_0, a_0, \dots, A_{20}, a_{20})$, which consists of the fitting parameters A and a at each of 21 redshifts (therefore, 42 data in total) for each cosmological model, using PCA. Combining all data vectors for 100 cosmological models, as well as 28 realizations of the fiducial *Planck* cosmology (therefore, $128 \times 42 = 5376$ data points in total), we decompose the data vector for the i th simulation, \mathbf{d}_i , into the PCs as

$$\mathbf{d}_i = \sum_{j=1}^n \alpha_{i,j}^{\text{HMF}} \mathbf{e}_j^{\text{HMF}}, \quad (40)$$

where $\mathbf{e}_j^{\text{HMF}}$ is the j th eigenvector with 42 components, which is independent of the cosmology or simulation realization, and $\alpha_{i,j}^{\text{HMF}}$ is the j th PC coefficient for the i th simulation. After various checks, we find that keeping the six most significant PC coefficients for each cosmological model, corresponding to $n = 6$ in Equation (40), is sufficient to keep the error induced in this step to a subpercent level. The accuracy level after applying the PCA method is not degraded to an extent easily visible by eye when we compare the lower and upper middle panels of Figure 13. In doing the PCA analysis, we downweight the components $A(z)$ by a factor of 10 compared with $a(z)$ to compensate for their different dynamic ranges.

Our next task is to collect the six PC coefficients ($\alpha_{i,j}^{\text{HMF}}$ in Equation (40)) for different cosmological models and perform GPR to interpolate each of the six PC coefficients between the sampled cosmological models, each of which is located at a particular position in six-dimensional cosmological parameter space. To do this, we apply the GPR to the PC coefficients for the 80 cosmological models included in slices 1–4 as the training set, excluding the 20 cosmological models in slice 5 (see Section 3.1 for details). Note that we do not include the fiducial *Planck* cosmology in this GPR either. The upper right panel of Figure 13 compares the GPR HMF with the simulated HMF for each of the 80 cosmological models in the training set. The GP does not perfectly reproduce the results of PCA at each of sampled cosmological models because we take into account the statistical uncertainties of the training data in the regression. Nevertheless, the plot shows that after applying the GPR, the rms in the ratios among different cosmological models are kept below $\sim 1\%$ (3%) on $M \lesssim 10^{13}$ (10^{14}) $h^{-1} M_\odot$.

The lower right panel of Figure 13 is the most important plot that gives an assessment of the performance of the HMF

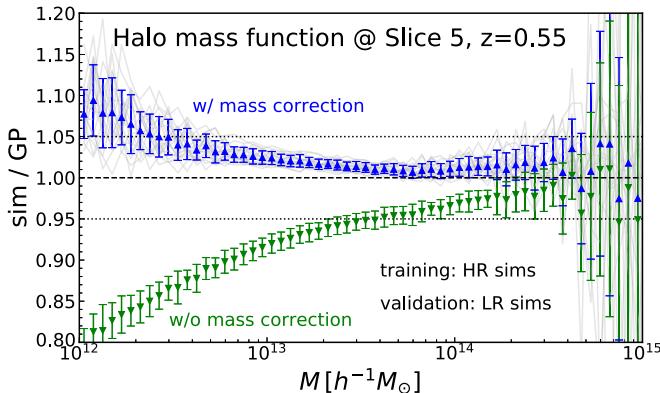


Figure 15. Comparison of the emulator prediction against the simulations in the LR suite. We consider the 20 cosmologies in slice 5 at $z = 0.55$ and show the ratio of the measurements from the LR suite to the emulator predictions. We show the results both with (upper) and without (lower) the mass correction (Equation (27)). Similar to Figure 13, we show the mean and scatter among the models by the error bars, and the individual cosmologies are shown by the gray solid curves.

emulator for an arbitrary cosmological model. The plot compares the GP-interpolated HMF with the simulated HMF for each of the 20 cosmological models in slice 5 that are not used in the GPR and serve as a cross-validation sample. The HMF emulator achieves great accuracy to predict the HMF, better than a few percent, in the amplitude up to halo masses of a few times $10^{14} h^{-1} M_\odot$. The performance is degraded for more massive halos, but the inaccuracy (averaged value denoted by the circle at each bin) is comparable with the statistical scatter. The good performance suggests that our GP method does not suffer from an overfitting to the training set.

While the performance of the emulator can be assessed fairly precisely for halo masses up to $\sim 10^{14} h^{-1} M_\odot$, the scatter among the models appears to be large for cluster-sized halos. While the large scatter could be simply due to the inaccuracy of the emulator, it can be partly due to the large Poisson noise, which can significantly affect the measurements used as the reference due to the small number of available cluster-scale halos in the simulations. As a final check of the accuracy of the emulator at the high-mass end, we compare the emulator prediction to the measurement from the LR simulations, which have bigger volume and are thus less affected by the Poisson noise.

In Figure 15, we show the ratio of the mass function measured from these LR simulations to the emulator prediction, which is trained based on the HR simulation suite. We show the mean and scatter among the 20 cosmological models in slice 5, with individual cosmology result (gray solid curve). We apply the correction of the mass based on the number of member particles (Equation (27)) for the upper set of curves, while we do not apply this to the lower set of curves. Since the size of the correction is pretty large for a mass less than several times $10^{13} h^{-1} M_\odot$ for these sets of simulations, we cannot derive a clear conclusion for these masses given the empirical nature of the correction. For more massive halos, the simulation results after the correction are very close to the emulator prediction. The scatter among the models is much smaller than what we can see in Figure 13, suggesting that the large scatter in the previous figure is indeed mainly due to the large Poisson noise in the reference simulations. However, a closer look at Figure 15 reveals that the mean of the ratio among the cosmological models for cluster-size halos is systematically

above unity by $\sim 1\%-2\%$. This would be a slight inaccuracy of the emulator due to the use of the HR simulation suite with a smaller volume. Since the size of this systematic error is comparable to those discussed for less massive halos based on the HR suite, we do not further consider the possibility of combining the measurements from the LR and HR simulation suites and instead stick to the emulator build based only on the HR suite for the HMF.

4.3.2. Halo–Matter Cross-correlation Function

We now discuss the halo–matter cross-correlation function. We first measure the cross-correlations for 13 mass threshold halo samples with different number densities at each of 21 redshifts from each simulation run, where we define the halo samples in 13 logarithmically spaced bins in the range of $n_h = [10^{-8.5}, 10^{-2.5}] (h^{-1} \text{Mpc})^{-3}$ (i.e., two bins per decade). For each measurement, we have 140 separation bins (40 logarithmic bins from 0.01 to $5 h^{-1}$ Mpc for the direct pair-counting method and 100 linear bins from 5 to $500 h^{-1}$ Mpc for the FFT method; see Section 3.6 for details). We thus have $13 \times 21 \times 140 = 38,220$ data points per simulation.

The left panel of Figure 16 shows variations in the halo–matter cross-correlations for 101 different cosmological models in the HR simulation suite for the mass threshold halo sample with a number density of $10^{-4} (h^{-1} \text{Mpc})^{-3}$ and at $z = 0.55$. This halo sample is chosen to roughly mimic the number density and large-scale bias of the CMASS galaxies. The plot displays rich cosmological dependences over scales ranging through the one-halo, two-halo terms to BAO scales. We then reduce the dimensionality of the data vector by first resampling the separation bins and then applying a PCA. The former is done using a cubic spline interpolation of the original data points up to $100 h^{-1}$ Mpc, with more data points around the one- and two-halo transition scale (i.e., around a Mpc scale), as depicted by the vertical dotted lines in the middle panel of Figure 16. We take 66 data points after the resampling. This procedure degrades the accuracy on small scales by no more than 3% ($\lesssim 40 h^{-1}$ Mpc). One might notice a small wiggly feature around $6 h^{-1}$ Mpc. This is due to the grid effect of the FFT method around the switching point to the direct pair-counting method, as described in Figure 3. Our spline function tries to remove this spurious feature to some extent by forcing the curve to be smooth. Since the raw simulation measurements employed here as the numerator still suffer from this artifact, the feature is still present in the ratio.

Our data vector still has $13 \times 21 \times 66 = 18,018$ components per simulation, which is quite large. We apply PCA to this data vector. As in the case for HFM, we combine all of the data vectors from 128 simulations (100 for the varied cosmological models plus the 28 *Planck* cosmology simulations), corresponding to $128 \times 18,018 = 2,306,304$ data points, and parameterize the halo–matter cross-correlation into its PCs as

$$\xi_{\text{hm}}(x, n_h, z)|_i = \sum_{a=1}^n \alpha_{i,a}^{\text{CCF}} e_a^{\text{CCF}}(x, n_h, z), \quad (41)$$

where $\xi_{\text{hm}}(x, n_h, z)|_i$ is the halo–matter cross-correlation at separation x for the halo sample with number density n_h and at redshift z in the i th simulation; e_a^{CCF} is the a th PC eigenvectors given as a function of separation x , n_h , and z (cosmology-independent); and $\alpha_{i,a}^{\text{CCF}}$ is the a th PC coefficient for the i th

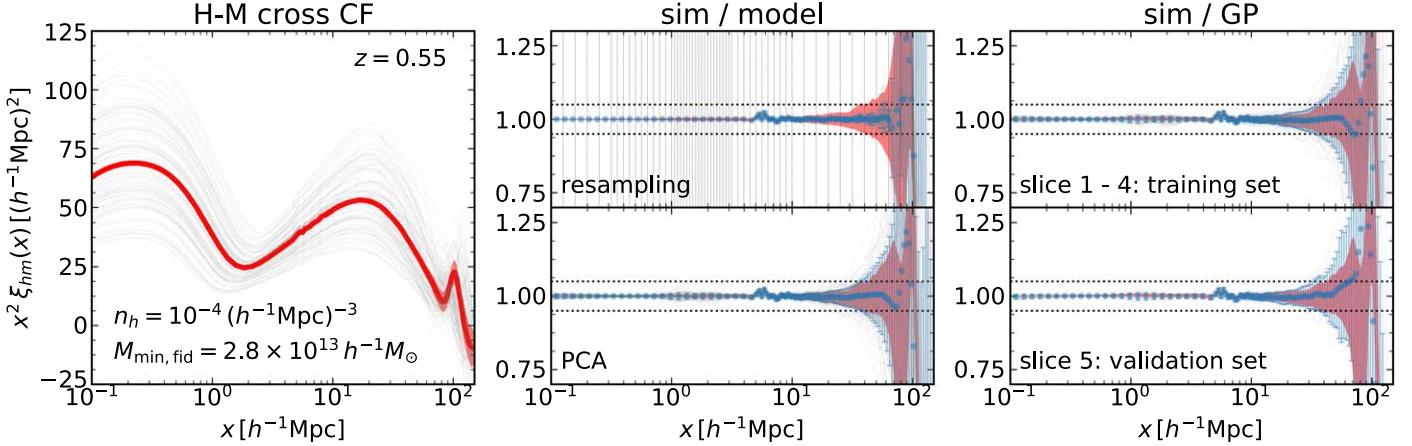


Figure 16. Modeling of halo–matter cross-correlation functions, similar to Figure 13. Here we consider the halo sample at $z = 0.55$ and with a number density $n_h = 10^{-4} (h^{-1}\text{Mpc})^{-3}$, corresponding to the mass threshold of $M \geq 2.8 \times 10^{13} h^{-1} M_\odot$ for the *Planck* cosmology, as an example. In the upper middle panel, we employ a resampling of separation bins, where the resampling points are denoted by the vertical dashed lines, and then model the cross-correlations by a cubic spline interpolation. The plot shows the interpolation results compared to the correlations directly measured from 100 simulations (see text for details). The lower middle panel shows the results when we model the cross-correlations using the PCA analysis for the data vectors, including cross-correlations in separation bins, halo sample bins, and 21 redshift bins (18,018 data points). Here we show the results obtained by using the five most significant PC coefficients (therefore, a huge dimension reduction from 18,018 to 5). The right panels show the emulator predictions for the training (upper) and validation (lower) cosmologies that are obtained after applying the GPR to the PC coefficients for the training simulations at 80 cosmological models. We show (red shaded region) the statistical uncertainties that are estimated from scatters in the 28 realizations of the fiducial *Planck* cosmology.

simulation. Thus, the eigenvectors $\{e_a^{\text{CCF}}\}$ describe dependences of the halo–matter cross-correlation on separation, halo sample (halo number density), and redshift. The different eigenvectors, e_a^{CCF} and e_b^{CCF} with $a \neq b$, are orthogonal to each other. The PC coefficients $\{\alpha_{i,a}^{\text{CCF}}\}$ describe the dependences on different simulations, i.e., cosmological models. In applying this PCA, we adopt the weight for each data point by $n_h x [1 - \exp(-x^2)]$ (x is in units of $h^{-1} \text{Mpc}$), such that the data points containing the halo sample of higher number density are upweighted, and the data points at large separation (x) are relatively upweighted; we empirically find the functional form of weight that satisfies these conditions. As shown in the upper right panel of Figure 16, we find that keeping the five significant PC coefficients reproduces the simulation results within a 5% accuracy on $x \lesssim 30 h^{-1} \text{Mpc}$ for each of 100 cosmological models, despite the huge dimensionality reduction (from 18,018 to 5). While we can see a relatively large deviation from unity at around $1 h^{-1} \text{Mpc}$ for a few cosmological models, the rms among the 100 models, as shown by the error bars, is typically below the 2% level on small scales up to $\sim 20 h^{-1} \text{Mpc}$ and reaches to 5% at $\sim 40 h^{-1} \text{Mpc}$. Since the halo–matter cross-correlation is a smooth function anyway, the PCA decomposition of the halo–matter cross-correlations works very well.

However, note that the PCA description cannot well describe the correlation at very large separations. The scatter of the curves around unity is consistent with the statistical error of the simulation data due to a finite number of simulation realizations or, equivalently, a finite simulation volume, as shown by the red shaded region, which is estimated from the 28 simulations at the fiducial *Planck* cosmology. We will instead use a different prescription, the propagator method, for the very large scale to overcome both sample variance in the simulation data and inaccurate modeling, as we will describe later.

We then perform GPR to model the cosmology dependence of the PC coefficients for 80 cosmological models in

slices 1–4, which is our training set, similar to Figure 13. Again, note that we do not include the 28 *Planck* cosmology realizations for this GPR. The upper right panel of Figure 16 shows that the GPR does not degrade the accuracy compared to the results after the PCA by more than 1% for the 80 cosmological models. The lower right panel gives a validation of our emulator; the GP interpolation reproduces equally well the cross-correlation function for each of 20 validation cosmological models in slice 5, which are not used in the GPR. The scatter is similar to the statistical error denoted by the shaded region (the sample variance of 28 *Planck* cosmology simulations). Thus, again, our GPR interpolation performs well without significant overfitting.

In Appendix F, we show the performance of the GP interpolation for different redshifts, as well as a different sample of halos that are characterized by the different number density in each cosmological model.

4.3.3. Halo Autocorrelation Function

We next discuss the halo autocorrelation functions. In this case, we generally need to consider two mass threshold halo samples of different number densities, and thus the dimension of the input data vector is even larger than that for the halo–matter cross-correlation functions. However, we cannot obtain a meaningful signal for halo samples with number densities lower than $\sim 10^{-6} (h^{-1}\text{Mpc})^{-3}$, unlike the halo–matter cross-correlation function due to the large Poisson noise. Hence, we consider here only eight bins with a high number density out of the 13 bins that we considered for the cross-correlation function. We thus consider 36 ($= 8(8+1)/2$) combinations for the two halo samples to form a matrix of autocorrelation functions at each of 21 redshifts and at each separation bin per simulation.

The left panel of Figure 17 shows variations in the halo autocorrelation functions for 101 cosmological models for the mass threshold halo sample with number density $n_h = 10^{-4} (h^{-1}\text{Mpc})^{-3}$ and at $z = 0.55$ that are now computed from the LR simulation suite. The plot shows rich cosmological

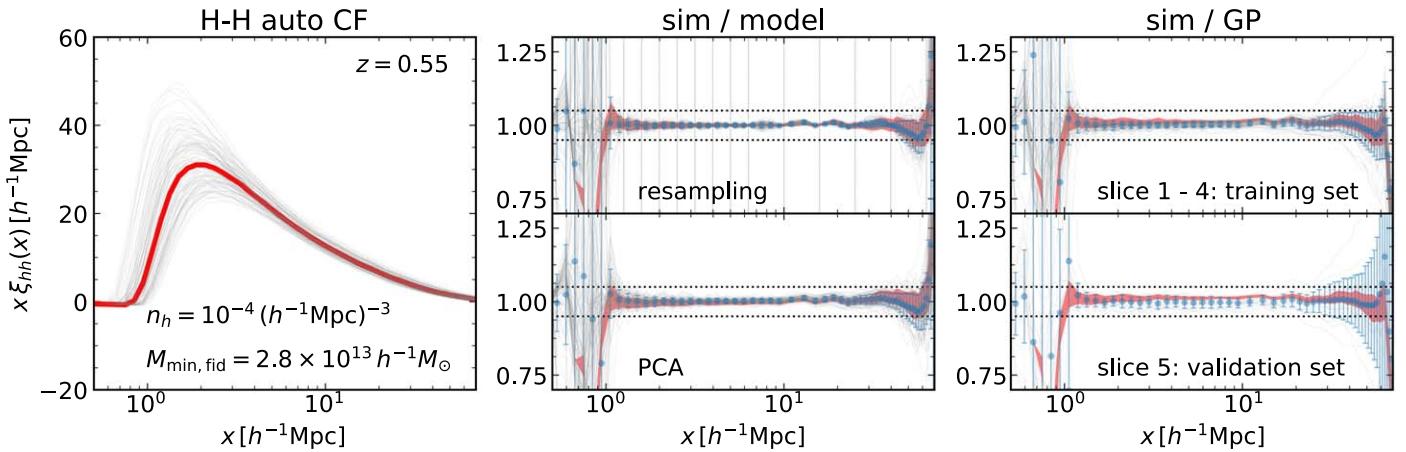


Figure 17. Modeling of the halo–halo autocorrelation functions, similar to Figure 16 for the halo–matter cross-correlation functions. As before, we consider halo samples with $n_h = 10^{-4} (h^{-1}\text{Mpc})^{-3}$ at $z = 0.55$. Variations in the function over the 101 cosmological models (left), the accuracy of our modeling procedures (resampling: upper middle; PCA: lower middle), and the performance after applying the GPR (training set: upper right; validation set: lower right) are shown, with the red shaded region indicating the scatters among the 14 simulations of the fiducial *Planck* cosmology.

dependences in the halo autocorrelations over the range of separation scales.

We then reduce the dimensionality of the data vector by first resampling the separation bins. We originally have 185 bins for each of the autocorrelation functions, 40 from the direct pair-counting method ($0.1 h^{-1}\text{Mpc} < x < 10 h^{-1}\text{Mpc}$) and 145 for the FFT method (up to $300 h^{-1}\text{Mpc}$). Since the autocorrelation function has a smooth shape at scales below BAO scale, we can significantly reduce the number of sampling points. As in the halo–matter cross-correlation case, we use a cubic spline interpolation to obtain the new data vector, and the vertical dotted lines in the upper middle panel denote the locations of new sampling points (21 points in total). The error level after this procedure is around 3% for the worst cases and 1%–2% in terms of the rms among the models on scales $1 \lesssim x/[h^{-1}\text{Mpc}] \lesssim 40$. The larger error on smaller scales is due to the significant halo exclusion effect.

Even after the above resampling of separation bins, we still have $21 \times 21 \times 36 = 15,876$ data points per simulation. The next task is to reduce the dimensionality of the data vector using PCA. As in the case of the halo–matter cross-correlations, we combine all of the data vectors from 114 simulations (100 cosmological models plus 14 *Planck* cosmology simulations), corresponding to $114 \times 21 \times 21 \times 36 = 1,809,864$ data points, and parameterize the halo autocorrelation into its PCs as

$$\xi_{hh}(x; n_1, n_2, z)|_i = \sum_{a=1}^n \alpha_{i,a}^{\text{ACF}} e_a^{\text{ACF}}(x, n_1, n_2, z), \quad (42)$$

where $\xi_{hh}(x, n_1, n_2, z)|_i$ is the halo correlation function at separation x for the two mass threshold halo samples with number densities $n_h = n_1$ and n_2 , respectively, and at redshift z in the i th simulation; e_a^{ACF} is the a th PC eigenvector given as a function of separation x , n_1 , n_2 , and z ; and $\alpha_{i,a}^{\text{ACF}}$ is the a th PC coefficient for the i th simulation. In applying the PCA, we adopt the weight, given as $n_1 n_2 x^2$, that is given as a function of the two number densities n_1 and n_2 and the separation x . However, since we use the LR suite here, there is a case that we cannot define a sample of halos with the highest number density, e.g., $n_h = 10^{-2.5} (h^{-1}\text{Mpc})^{-3}$, depending on redshifts and cosmological models. In such cases, we set the weight for

PCA to zero. After some experiments, we find that keeping the eight significant PC coefficients well reproduces the simulation results. To be more quantitative, we can maintain the error level of a few percent with a slight degradation toward the large scales ($\gtrsim 10 h^{-1}\text{Mpc}$), as shown in the lower middle panel of Figure 17. Thus, we made a huge dimensionality reduction of the data points (from 1.8×10^6 to 8). The error induced by this GPR is negligibly small except for very large scales, where we stitch to the propagator-based prescription, as we will describe below.

Next, we perform GPR to model the eight PC coefficients for 80 cosmological models: 80 different cosmological models in slices 1–4. The upper right plot of Figure 17 shows that the GP interpolation reproduces the simulation results for each of 80 cosmological models, which is the training set for the GPR, within 1%–5% rms accuracy, depending on the scale. The lower right panel gives a validation of our emulator; the GP interpolation gives 2%–3% rms accuracy on $1 \lesssim x/[h^{-1}\text{Mpc}] \lesssim 20$ for the 20 cosmological models in slice 5, which is the validation set.

It is worth noting that there are differences in the GPR results between the halo–matter cross- and halo autocorrelations. First, the small-separation data points have large scatters, which is not seen in the cross-correlation function. This is due to the fact that the autocorrelation function is suppressed on these small scales due to the halo exclusion effect. We eventually expect a zero-crossing of the autocorrelation function at the scale where the halo exclusion effect is dominant, and thus the ratio at the zero-crossing scale becomes large and noisy. Second, the accuracy of the emulator prediction is worse than that for the cross-correlation because of the larger noise of autocorrelation measurements due to the Poisson noise originating from the discreteness of halos.

In Appendix F, we show the performance and validation for the halo autocorrelation functions for halo samples of different number densities and redshifts, as well as the cross-correlations between two halo samples with different number densities.

4.3.4. Propagator

The final piece of our halo modules is the propagator that describes the large-scale correlation functions around BAO

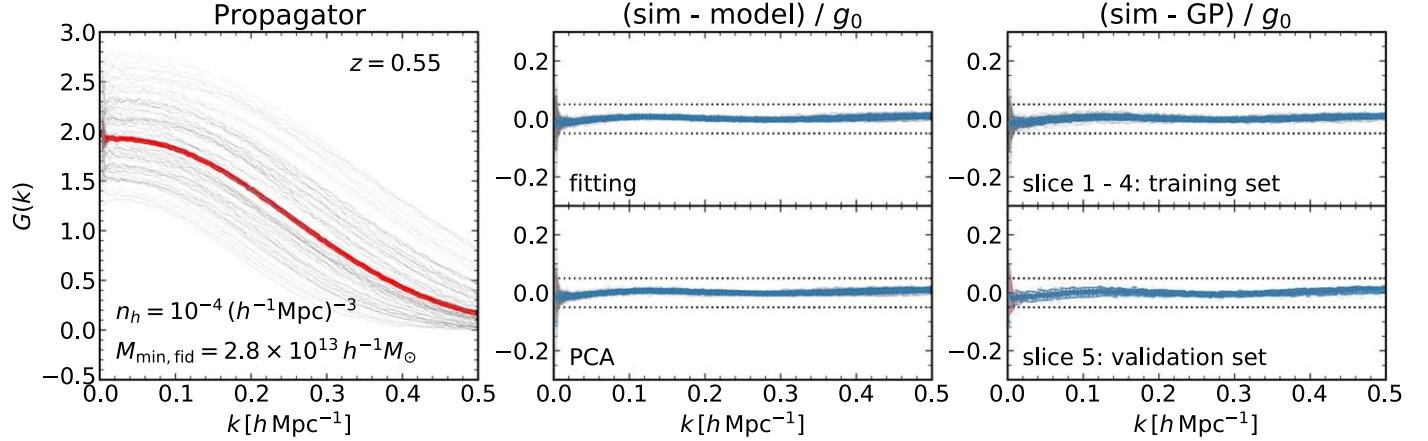


Figure 18. Modeling of the halo propagator, similar to Figures 16 and 17 for correlation functions. As before, we consider halo samples with $n_h = 10^{-4}(h^{-1}\text{Mpc})^{-3}$ at $z = 0.55$. Variations in the function over the 101 cosmological models (left), the accuracy of our modeling procedures (model fitting: upper middle; PCA: lower middle), and the performance after applying the GPR (training set: upper right; validation set: lower right) are shown, with the red shaded region indicating the scatter among the 14 simulations of the fiducial *Planck* cosmology.

scales. Since the halo–matter cross-correlation function involves the propagators of halo and matter, we here study both. We use the LR suite for this purpose, as the damping behavior of the propagator is known to be mainly due to the large-scale bulk motion. Since the estimation of the propagator is done by using the cross-correlation between the halo (or matter) density field and the linear density field used in setting up the initial conditions of the simulations (i.e., Equation (30)), the measured data do not show discreteness noise, unlike what we see in the halo autocorrelation functions. Therefore, we can measure it to a reasonably accurate precision for all 13 number density bins for the mass threshold samples from $10^{-2.5}$ to $10^{-8.5}(h^{-1}\text{Mpc})^{-3}$ used for the halo–matter cross-correlation functions.

Since the shape of the propagator is roughly a Gaussian with its width being the rms displacement, as shown in Figure 12, we parameterize it as

$$G_a(k) = (g_0 + g_2 k^2 + g_4 k^4) \exp\left(-\frac{\sigma_{\text{d,lin}}^2 k^2}{2}\right), \quad (43)$$

where g_0 can be interpreted as the linear bias at the large-scale limit and $\sigma_{\text{d,lin}}$ is the linear rms displacement for the matter field at the redshift under consideration, which is computed by the linear module. The other coefficients, g_2 and g_4 , are fitting parameters, introduced to capture a departure from the simple Gaussian form. In the above, the subscript a on the left-hand side stands for either m (matter) or h (halo), and the factor g_0 is set to unity for the matter propagator.

We show in Figure 18 our modeling detail of this function for halo samples with number density $10^{-4}(h^{-1}\text{Mpc})^{-3}$ at $z = 0.55$. As before, we show variations in the function for the 101 cosmological models in the left panel, with the red shaded region showing the mean and scatter of this function for the fiducial *Planck* cosmology (the shaded region is also shown in the other panels, though it is heavily overlapped with other symbols and thus difficult to see). We can see that the propagator is always a simple decaying function of wavenumber, with a strong dependence on cosmology in the amplitude and the typical wavenumber at which the curve is decaying.

We then fit the data using Equation (43) and show the residual in the upper middle panel. Here and in the other three

panels, we normalize the residual by g_0 , which is the low- k limit of this function, to see the importance of the residual relative to the overall amplitude of the function. While we see a small wiggly pattern in the residual, the typical amplitude of this pattern is below a few percent level, which is sufficiently small for our purpose.

At this point, we have three fitting parameters per halo sample, and thus $819(=3 \times 13 \times 21)$ data points per simulation for the halo propagator. As before, we reduce the dimensionality by applying the PCA. The data points are approximated by

$$\mathbf{d}_i^{\text{PRO}}(n_h, z) = \sum_{a=1}^n \alpha_{i,a}^{\text{PRO}} \mathbf{e}_a^{\text{PRO}}(n_h, z), \quad (44)$$

where $\mathbf{d}_i^{\text{PRO}}(n_h, z)$ is the vector formed with the three fitting parameters for the halo sample with number density n_h and at redshift z in the i th simulation, $\mathbf{e}_a^{\text{PRO}}$ is the a th PC eigenvector given as a function of n_h and z , and $\alpha_{i,a}^{\text{PRO}}$ is the a th PC coefficient for the i th simulation. In applying the PCA analysis, we adopt the weight, simply given as n_h . As before, since we use the LR runs here, there is a case that we cannot define a sample of halos with the highest number density, e.g., $n_h = 10^{-2.5}(h^{-1}\text{Mpc})^{-3}$, depending on redshifts and cosmological models. In such cases, we set the weight for PCA to zero. After some experiments, we find that keeping the four most significant PC coefficients well reproduces the simulation results, as shown in the lower middle panel of Figure 18. The extra error induced by the PCA is below the 1% level.

The remaining task is the same as before: train a GPR using the 80 cosmological models in slices 1–4 and validate the results using the remaining 20 models in slice 5 (as well as the fiducial *Planck* cosmology). Although the variance of the residuals among the models shown in the right panels seems to be somewhat larger than that in the middle panels, the prediction of the GP stays within the $\pm 5\%$ band shown by the horizontal dotted lines. More importantly, the accuracy of the GP for the validation set is not degraded compared to that for the training set, implying that there is no problem of overfitting to the training data.

The validation tests for other halo samples, as well as the matter propagator at various redshifts, can be found in

Appendix F. In the current implementation, we model the matter propagator by exactly following the same procedure for halos. The only difference is that we have one less free parameter (g_0 should always be unity for matter), and we need only two PCA components to ensure the accuracy.

5. Usage of the Emulator

We have explained how each of the basic modules are modeled and tested in the previous section. These correspond to the HALO MODULES in Figure 4, which predict the statistical properties of dark matter halos. Now, in this section, we show several demonstrations of how to use the emulator to predict properties of dark matter halos. Also, we show how to combine the predictions to compute the clustering statistics of galaxies (i.e., usage of the UTILITY MODULES at the bottom of Figure 4).

5.1. Halo Properties

5.1.1. Implementation Detail

One application of our emulator is to make predictions for halos in the mass range of galaxies to clusters. Since our modules that compute halo clustering properties directly predict the signals as a function of the cumulative halo number density, which is discretely sampled every 0.5 dex, it might not be so practically useful as it is. To obtain the predictions of halo correlation functions at a given halo mass, we have to interpolate over the sampled number densities and convert it to the halo mass in a given cosmological model. Hence, if one wants to predict the clustering signals as a function of halo mass, an inaccuracy in the conversion from the mass to the number density can be a new source of error.

In Figure 19, we study how the conversion from the cumulative halo number density to a target halo mass using the HMF module causes a possible error in the predictions of halo correlation functions. To do this, we consider the emulator outputs of halo correlation functions for two mass threshold samples with $M_{\min} = 10^{13}$ and $10^{14} h^{-1} M_{\odot}$ for the fiducial *Planck* cosmology at $z = 0.55$. Then we use the following method to propagate a possible error in the halo number density into an error in predicting the halo correlation functions as a function of the halo mass. (i) We first use the HMF module to compute the cumulative number density for the halo mass thresholds, $M_{\min} = 10^{13}$ and $10^{14} h^{-1} M_{\odot}$. (ii) We multiply the number density by a factor of 0.96, 0.98, 1.02, or 1.04, which is intended to mimic a possible error in the number density calibration by -4% , -2% , 2% , or 4% , respectively. (iii) We then obtain the emulator predictions of halo correlation functions by inserting the shifted values of halo number density in the emulator. Here a $\pm 4\%$ error in the cumulative halo number density is considered as a rather pessimistic case because Figure 13 shows that a typical error in the mass function is smaller in terms of the rms among the models ($\sim 1\%$ (3%) at $10^{13}(10^{14}) h^{-1} M_{\odot}$). In addition, the error in the HMF seen in Figure 13 would be partly canceled when we consider the error on the cumulative halo number density. The upper panel shows the ratios of the shifted halo–matter cross-correlation functions, $\xi_{\text{hm}}(x)$, to the fiducial prediction. The figure shows a constant shift in the two-halo regime, a slightly larger shift in the one-halo term, and a bump-like feature at transition scales between the two regimes. These are caused by changes in the linear bias and mass profile, respectively. The size of the fractional shift in the cross-correlation function is

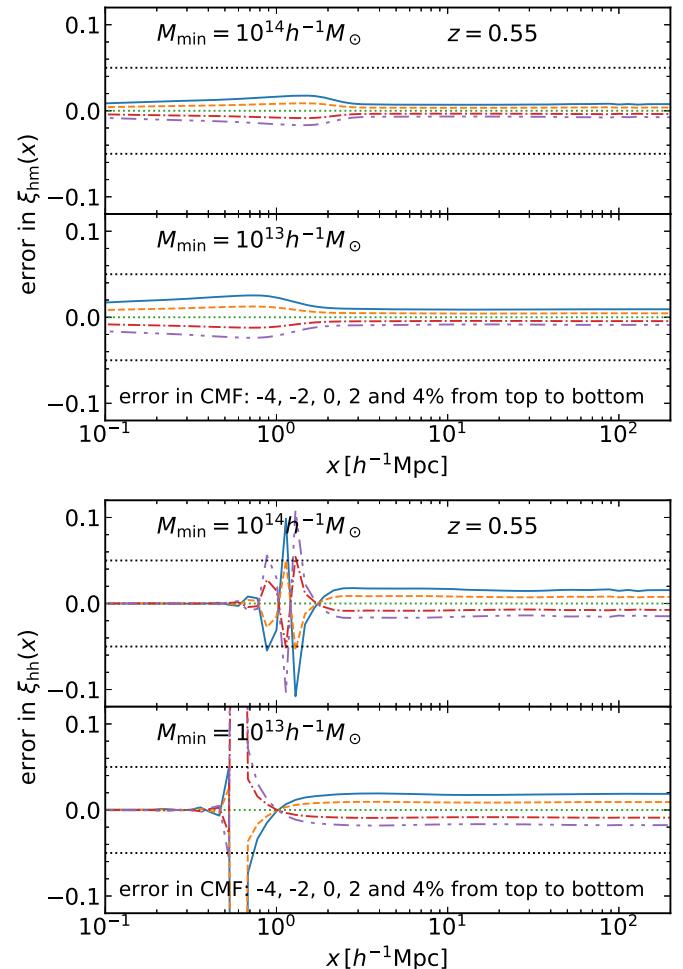


Figure 19. Impact of a possible error in the conversion between the cumulative halo number density and the mass threshold on the emulator predictions of halo correlation functions. As a working example, we here consider the cross-correlation functions, $\xi_{\text{hm}}(x)$, in the upper panel and the halo autocorrelation functions, $\xi_{\text{hh}}(x)$, in the lower panel for the two mass thresholds (10^{14} and $10^{13} h^{-1} M_{\odot}$) for the *Planck* cosmology at $z = 0.55$. Here we first use the HMF module to compute the cumulative halo number density for the mass threshold; shift the number density by -4% , -2% , 2% , or 4% ; and then input the shifted number density into the emulator to obtain the shifted predictions of $\xi_{\text{hm}}(x)$ and $\xi_{\text{hh}}(r)$, respectively (see text for details). The figure shows the ratio of the shifted correlation function to the fiducial prediction. The significant features around $x \approx 1 h^{-1}\text{Mpc}$ in $\xi_{\text{hh}}(x)$ are due to the halo exclusion effect that would not be present for the galaxy correlation function.

smaller than that on the cumulative mass function, with a slight decreasing trend toward higher masses. In the lower panel, the autocorrelation function, $\xi_{\text{hh}}(x)$, shows a larger shift in the two-halo regime, reflecting the fact that it scales as bias squared. A sharp feature can be found where the halo exclusion effect kicks in. Since the latter part is dominated by the one-halo term in the case of the galaxy correlation function, the final shift would be much smaller. Even with the pessimistic case of a 4% error in the cumulative mass function, the induced shift in the correlation functions are well within the $\pm 5\%$ band and mostly within the $\pm 3\%$ level. When we consider a realistic error on the cumulative mass function (i.e., a few percent or below), the error on the correlation function arising from this is smaller than the typical error in the emulator in both cases.

As another example of the applications, we show in Figure 20 the output of the emulator for the large-scale bias as a function of the halo number density (symbols). Here the

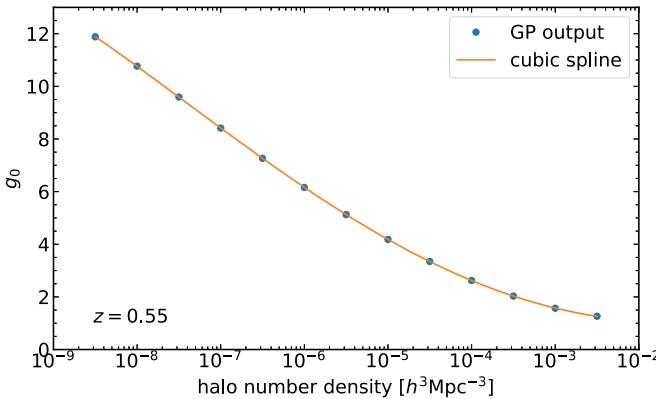


Figure 20. Halo bias as a function of the halo number density for the fiducial cosmological model at $z = 0.55$. The symbols are the direct output of our emulator, and the solid curve is its interpolation using the cubic spline.

large-scale bias is defined as the fitted parameter g_0 in Equation (43), which is the $k \rightarrow 0$ limit of the propagator (Equation (28)). The plot shows the result for the fiducial *Planck* cosmology and at $z = 0.55$. We interpolate these data points using the cubic spline function to make a prediction at any halo number density in the range shown here. Note that we use the logarithm of the halo number density, instead of the raw values of the number density, for which our sampling is uniform.

Once the spline interpolator is ready, we can compute the halo bias as a function of the halo mass or peak height, if one prefers, by first using the HMF module to convert the number density to the minimum halo mass and then taking a finite difference derivative to have the bias at a specific desired halo mass scale. For this derivative, we employ $\pm 1\%$ changes in the mass as the default step size. The dependence of the results on the step size is much weaker than the typical accuracy of the emulator, as shown in Figure 21. The result is shown in Figure 22 as a function of the peak height $\nu \equiv \delta_c/\sigma_M$, with $\delta_c = 1.686$. Now we show the results for different cosmologies at three different redshifts. We vary Ω_m , keeping the flatness and fixing the present-day amplitude of the linear matter power spectrum, σ_8 , to its fiducial value. We compare the results with the fitting formula by Tinker et al. (2010), as denoted by the thin solid line, which is independent of redshift or cosmology. The fractional differences from Tinker et al. (2010) at the three redshifts are plotted together in the bottom panel. Our emulator prediction is overall consistent with the fitting formula, with the accuracy no worse than 10% over all the ranges examined here. Such an inaccuracy of Tinker et al. (2010) is also pointed out in the previous work (e.g., Li et al. 2016). We confirm that the bias function is rather universal, with little dependence on cosmology or redshift.

Another interesting feature of the bias is its scale dependence around the BAO scale. We implement the same spline interpolation for the fitting parameters g_2 and g_4 in the propagator. This allows us to estimate the propagator at any halo mass. The tree-level calculation, Equation (29), gives us a prediction of the correlation functions, and we already show that the BAO scale is well described by this simple model, as illustrated in Figures 9–11. We show in Figure 23 the square root of the ratio of the halo and matter correlation functions at $z = 0$. We normalize it by the linear bias factor g_0 such that the ratio becomes unity when the bias is independent of scale. We consider four halo samples with different number densities as

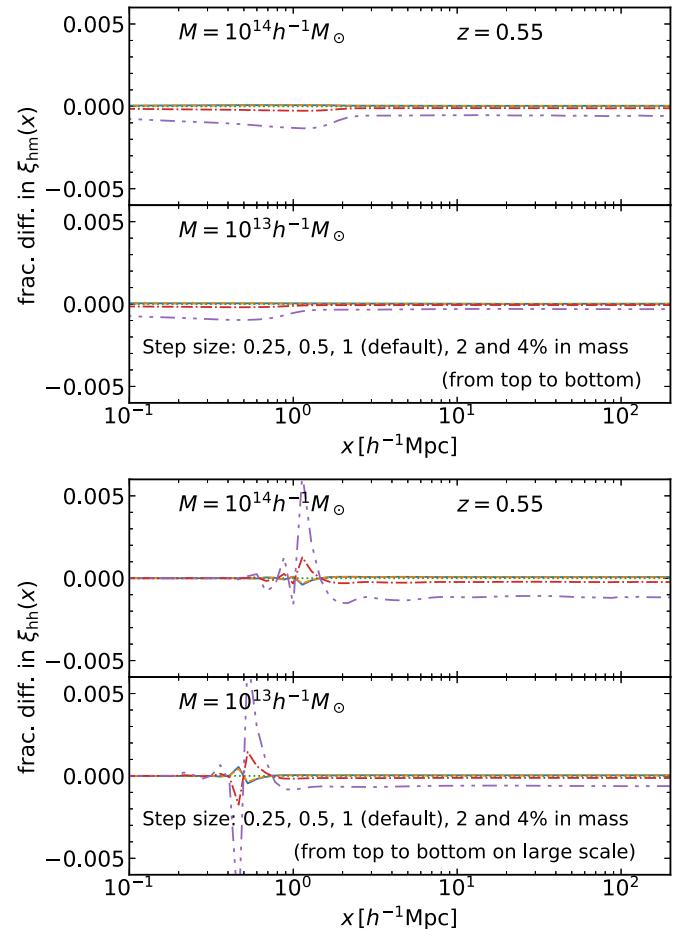


Figure 21. Stability of the finite difference evaluation of the correlation functions at a given halo mass. We show the fractional change in the halo–matter cross- (upper) and halo auto- (lower) correlation function. We employ the default step size of $\pm 1\%$ in mass as the reference for this figure. Notice the rather narrow range of the vertical axis.

written in the figure legend. The result indicates that the BAO peak structure can be boosted for low number density samples (i.e., when only massive halos are included in the sample). This is fully consistent with the expectation by the peak model (compare our results with Figures 7 and 8 of Desjacques et al. 2010). Also, this feature was previously found in numerical simulations (e.g., Angulo et al. 2014; Crocce et al. 2015). This kind of prediction is possible because our model has the freedom to control the damping of the BAO feature in terms of the two free parameters, g_2 and g_4 , in the propagator, Equation (43), in addition to the damping due to the typical random displacements of matter, $\sigma_{d,\text{lin}}$.

We implement the same cubic spline interpolation for other quantities, such as $\xi_{hm}(r; n_h)$ or $\xi_{hh}(r; n_1, n_2)$, with the latter using the bivariate cubic spline for two number densities, n_1 and n_2 . We do not find any sizable error originating from this interpolation, as all of the quantities vary rather smoothly with (the logarithm of) the halo number density. Analogously, the redshift dependence is interpolated with the cubic spline function. As is clear from Figure 22, the dependence of bias on redshift is weak, and thus the same interpolation scheme works fine. The situation is similar for the other interpolated quantities. We show in Figures 24 and 25 our interpolation of the halo–matter cross- or halo autocorrelation function over the number density and redshift. We first construct a data

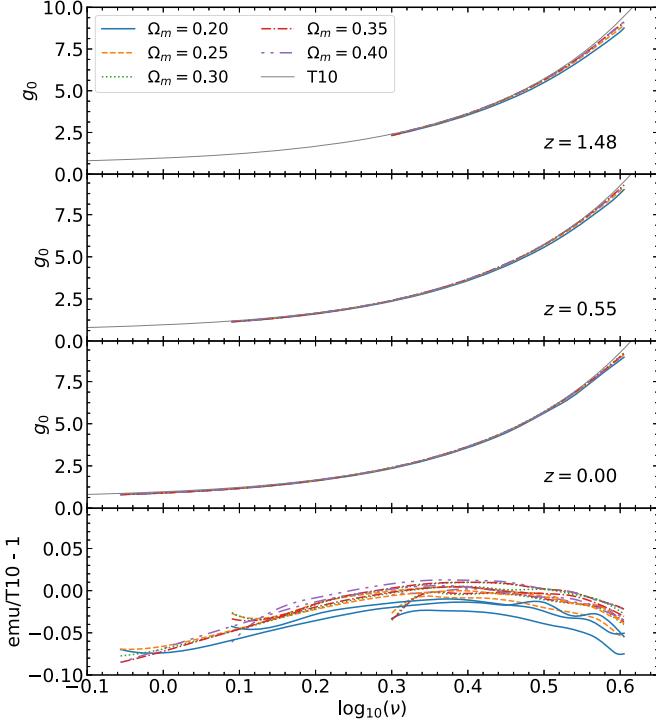


Figure 22. Halo bias as a function of the peak height for different cosmologies at various redshifts as shown in the legend. We vary Ω_m but keep σ_8 and other cosmological parameters fixed to their fiducial values. We also show the fitting formula of Tinker et al. (2010; thin solid line).

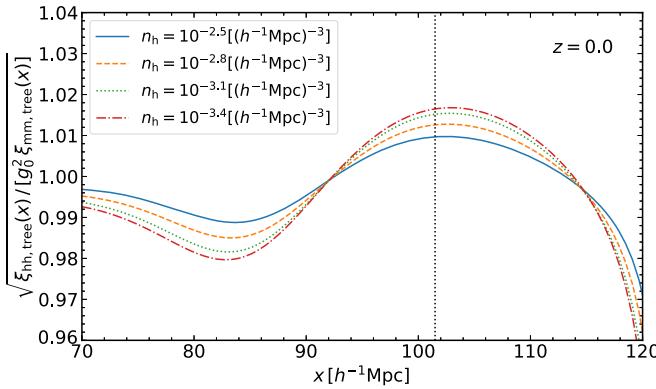


Figure 23. Scale dependence of halo bias around the BAO scale (vertical dotted line). We show the square root of the ratio of the halo and matter correlation functions, with the latter scaled by the square of the linear bias factor g_0 . We show with different lines four mass threshold halo samples with the number density listed in the figure legend, corresponding, respectively, to a threshold mass of 1.58×10^{12} , 3.34×10^{12} , 6.86×10^{12} , and $1.36 \times 10^{13} h^{-1} M_\odot$ at $z = 0$.

matrix at the locations, as depicted by the dots, based on the GP and PCA methods, and then we perform a cubic spline interpolation for each dimension.

While the halo–matter cross-correlation function and the propagator are very well determined down to a quite low halo number density, $10^{-8.5}(h^{-1}\text{Mpc})^{-3}$ —corresponding to very massive halos, thanks to the fact that they are given by the cross-correlation with the matter field—the halo autocorrelation function instead suffers from severe Poisson noise, especially at such a high-mass end. We thus switch to a simple scaling, $\xi_{hh}(r; n_1, n_2) = [g_0(n_1)/g_0(n_{\min})] \xi_{hh}(r; n_{\min}, n_2)$, when the number density n_1 is below the minimum number density

$n_{\min} = 10^{-5.75}(h^{-1}\text{Mpc})^{-3}$. In the above, the bias factor, $g_0(n_i)$, is computed again in the module that computes the propagator (i.e., the function plotted in Figure 20). We do the same when n_2 is below the threshold; we simply multiply the ratio of the large-scale bias one more time. While this might be a reasonable approximation on large scales, it cannot properly reproduce the correlation functions around scales where the halo exclusion effect is not negligible. Nevertheless, our current implementation does not lead to a severe error for a sample of galaxies such as the CMASS sample, because the small-scale correlation function is mainly described by the one-halo term.

Now we have predictions of ξ_{hm} , ξ_{hh} , and the propagator given as a function of halo number density and redshift within the ranges relevant for the resolution and output redshifts of our simulations. We combine all of these predictions to obtain a well-behaved prediction over a wide range of separations. We do this by smoothly stitching the two predictions as

$$\xi_{ab,\text{full}}(x) = D(x)\xi_{ab,\text{direct}}(x) + [1 - D(x)]\xi_{ab,\text{tree}}(x), \quad (45)$$

where “ab” is either “hm” or “hh,” and we use the damping function defined as

$$D(x) = \exp\left[-\left(\frac{x}{x_{\text{switch}}}\right)^4\right]. \quad (46)$$

We find that $x_{\text{switch}} = 60 h^{-1}\text{Mpc}$ provides a reasonably good model for both the auto- and cross-correlation functions over the range of scales demonstrated in Figure 26.

In summary, the current implementation of our emulator works in a parameter sampler as follows. When a new cosmological model is proposed, the code first calls a GP interpolator to evaluate the coefficients for the PCs for all of the statistics we consider here. Then, combining these coefficients with the eigenvectors, it computes the statistics at all of the redshifts and mass and separation bins to form a data table. One function call of our high-level interface to set the cosmological parameters does all the tasks up to here internally. Now, a user can further call other high-level functions prepared for each statistics. These functions accept a redshift, a halo mass (either a threshold mass or a target mass scale), and a set of separations at which the correlation function should be evaluated. In this final step, the code finds the values by calling a spline interpolator over the table created in the previous step. For users who wish to compute galaxy statistics, a separate module can be used with additional parameters describing the HOD model. This module internally calls the functions for the halo statistics and integrates them over the halo mass with the product of the mean HOD and the HMF as a weight. We also prepare functions computing projected statistics, which works similarly.

5.1.2. Demonstrations

Now we can compute the three main halo statistics—the HMF, halo–matter cross-correlation function, and halo auto-correlation function—for an arbitrary cosmological model that is covered by our sampled cosmological models within the flat w CDM cosmologies. Using the results, we can obtain how these halo statistical quantities vary with cosmological parameters, as demonstrated in Figure 1, which gives their dependences on Ω_m .

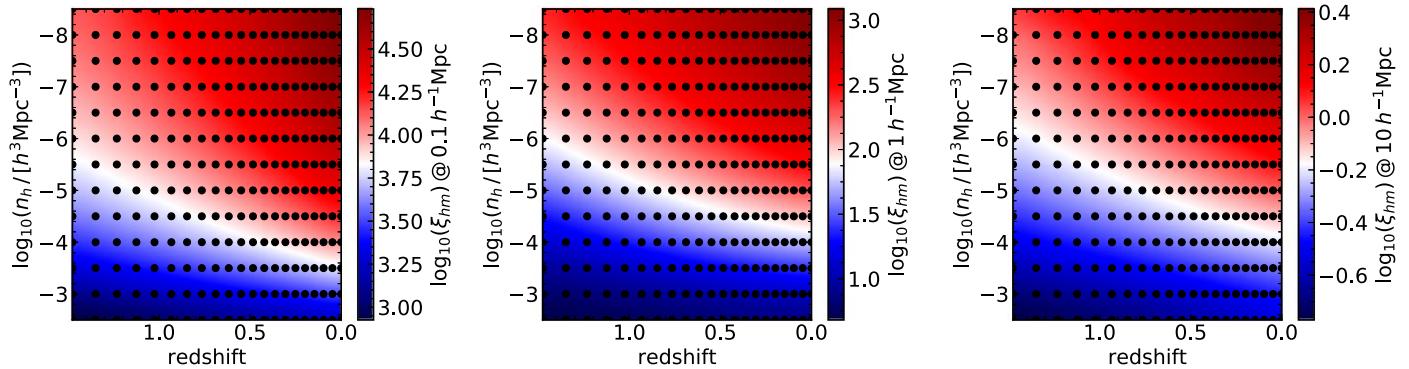


Figure 24. Interpolation of the tabulated halo–matter cross-correlation function over the halo number density and redshift. Each panel shows the interpolated result (color scale) at a fixed separation, as shown in the color bar label, and the dots show the location where the data table is available.

We can further predict in detail, for instance, the density profile of dark matter halos from our emulator. Properties and cosmological dependences of the mass density profiles around halos have been extensively studied (e.g., Navarro et al. 1996). The emulator output of the halo–matter cross-correlation on small separations can be used to study the mass density profiles for halos whose masses are in the range supported by the emulator.¹⁷

In Figure 27, we compare the profiles from the emulator (solid) with the best-fit Navarro–Frenk–White (NFW) profiles (dashed) for halos of different masses. For the fitting, we included the data over the range of radii from twice the softening scale to 80% of R_{200m} . The middle panel shows that the NFW profile gives a good fit to a fractional accuracy better than about 5% up to the virial radius (R_{200m}), beyond which the NFW profile no longer reproduces the simulation results.

Using the NFW fitting results, we can also study how the concentration parameter varies with halo mass, as well as cosmological models over different redshifts, where we define the concentration parameter, c_{200c} , by the ratio of the radius within which the density is 200 times the critical density to the scale radius determined by the NFW fit. Figure 28 shows c_{200c} as a function of the peak height, $\nu = \delta_c/\sigma_M$, for cosmologies with different Ω_m at three redshifts. Similar to the previous plots, we keep the spatial flatness and vary the normalization A_s such that σ_8 is kept unchanged for models with different Ω_m . While the relation seems to be universal at high redshift with little dependence on Ω_m , we can see a clear dependence at lower redshifts. The increasing trend of c_{200c} as decreasing redshift, as well as its positive Ω_m dependence, can be found in Diemer & Kravtsov (2015). Further study of the dependence of the concentration–mass relation on the cosmological parameters can be found in Kwan et al. (2013).¹⁸ In this way, our emulator approach automatically incorporates a possible nonuniversality of the concentration–mass relation. This is quite different in the standard analytical halo model approach, where one usually employs a simulation-calibrated scaling relation for the concentration. We would also like to note that such a calibration of the concentration is often done for a specific cosmological model.

¹⁷ Notice that the average spherical density profile of halos is equivalent to the positional cross-correlation function between halos and matter by definition.

¹⁸ We cannot make a direct comparison with their emulator because the Hubble parameter is automatically determined to match to the CMB constraint given the other parameters in their code. On the other hand, we here vary Ω_{de} , keeping the spatial flatness, and h is simultaneously changed to keep ω_b and ω_c fixed.

Now we focus on the halo mass density profiles at radii larger than R_{200m} in Figure 27, where NFW no longer gives a good fit. The figure shows a clear feature of the transition from the one-halo to the two-halo regime. This feature recently drew attention as a possible “physical” outer boundary of a halo associated with the first orbital apocenter of accreted matter after its infalling, dubbed the “splashback” feature (Diemer & Kravtsov 2014; also see More et al. 2016, for the first detection from observational data). This feature has already been studied with our DARK QUEST simulation suite in Okumura et al. (2018a, 2018b), with particular attention to the feature in the velocity statistics around halos.

The feature can be found from the bottom panel of Figure 27, where we show the logarithmic slope of the mass density profile. We obtain this by first fitting the emulator results by the functional form proposed in Diemer & Kravtsov (2014; hereafter the DK fit), and then we take the derivative. We do this for the separation range again from twice the softening scale but to four times the radius R_{200m} to cover both the one- and two-halo regimes. The best-fit model is shown in the top panel by the dotted lines (but they are difficult to distinguish from the solid lines; they are almost on top of each other), and the ratio of the emulator results is shown in the middle panel. The accuracy of the fit is similar or better than the NFW form, and it remains to be a good fit to much larger scales. Now we can see in the bottom panel that the derivative based on the DK fit shows a sharp dip with a slope steeper than the outer NFW slope (i.e., -3), marking the location of the splashback radius.

Figure 29 shows the splashback radius, R_{sp} , for various cosmological models and redshifts, where we define R_{sp} by the location of the minimum logarithmic slope of the DK fit. For clarity, we plot the ratio, R_{sp}/R_{200m} , as a function of the peak height. Overall, R_{sp} is similar to R_{200m} , with a slight decreasing trend as a function of the peak height. In addition, the ratio is higher for cosmological models with larger Ω_m . These trends are in qualitative agreement with the fitting formulae in More et al. (2015a), in which the dependence is encoded in the redshift-dependent density parameter $\Omega_m(z)$ in addition to the peak height ν (see also Adhikari et al. 2014)¹⁹

¹⁹ Note, however, that it was argued that the majority of the dependence of R_{sp}/R_{200m} comes from the accretion rate, and its distribution at different redshifts and for cosmologies should depend on how to define distinct halos and their mass accretion histories from N -body simulation outputs in great detail. This is beyond the scope of this paper.

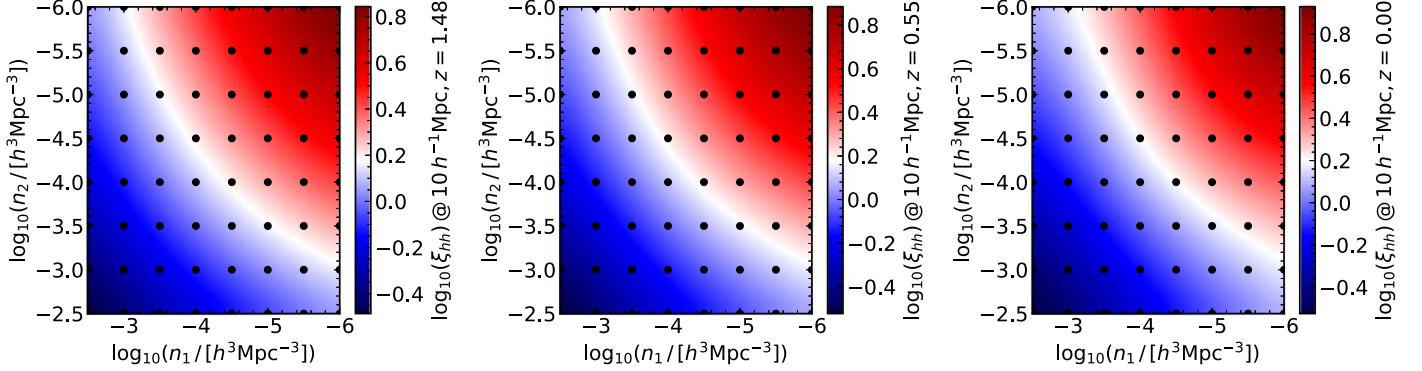


Figure 25. Similar to Figure 24, but for the halo autocorrelation function. We now fix both separation and redshift in each panel and show the interpolation over the two number densities, n_1 and n_2 .

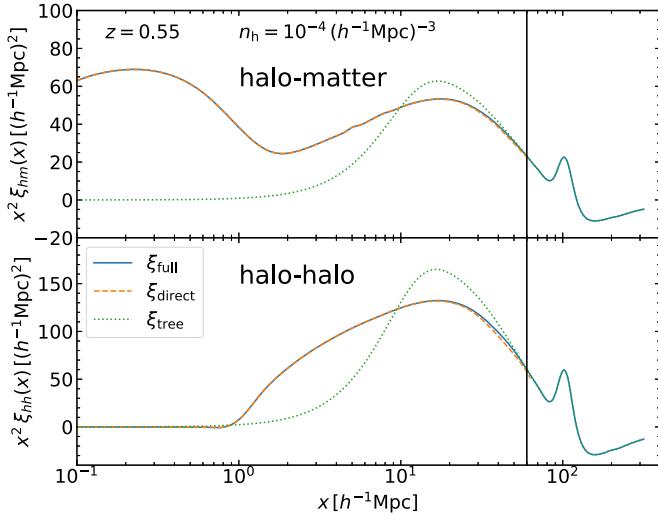


Figure 26. Stitching of the large- and small-scale predictions. The direct output of the PCA-GP modeling is shown by the dashed lines (ξ_{direct}), the large-scale model based on the emulated propagator is shown by the dotted lines (ξ_{tree}), and the final prediction is shown by the solid lines (ξ_{full}). We show in the upper panel the halo–matter cross-correlation function and in the lower panel the halo autocorrelation function, both for the halo number density $10^{-4} h^3 \text{Mpc}^{-3}$ at $z = 0.55$. The vertical solid line marks the stitching scale $x_{\text{switch}} = 60 h^{-1} \text{Mpc}$.

5.2. Projected Galaxy Clustering Statistics

We have introduced the emulation of halo clustering statistics in the previous sections. Since our emulator’s accuracy depends on the mass of halos, it would be useful to examine the accuracy for a galaxy sample whose clustering statistics is approximately given as a weighted sum of those of halos. Here we consider the following HOD parameters to make a representative galaxy mock catalog similar to the BOSS CMASS sample (based on a conservative volume-limited selection): $M_{\min} = 10^{13.94} h^{-1} M_{\odot}$, $\sigma_{\log M} = 0.63$, $M_{\text{l}} = 10^{14.49} h^{-1} M_{\odot}$, $\alpha = 1.19$, and $\kappa = 0.60$ for the HOD parameters (see Appendix G for definitions).

Our UTILITY MODULES combine the outputs of HALO EMULATORS to first make the galaxy clustering signals in three dimensions and then project them along the line of sight to obtain the relevant signals based on the FFTLog algorithm (Hamilton 2000). To test the accuracy of DARK EMULATOR in predicting galaxy clustering, we also generated the mock catalogs of galaxies; we populate central and satellite galaxies into halos taken from the halo catalog in each of 24 HR realizations of the fiducial *Planck* cosmology and then measure

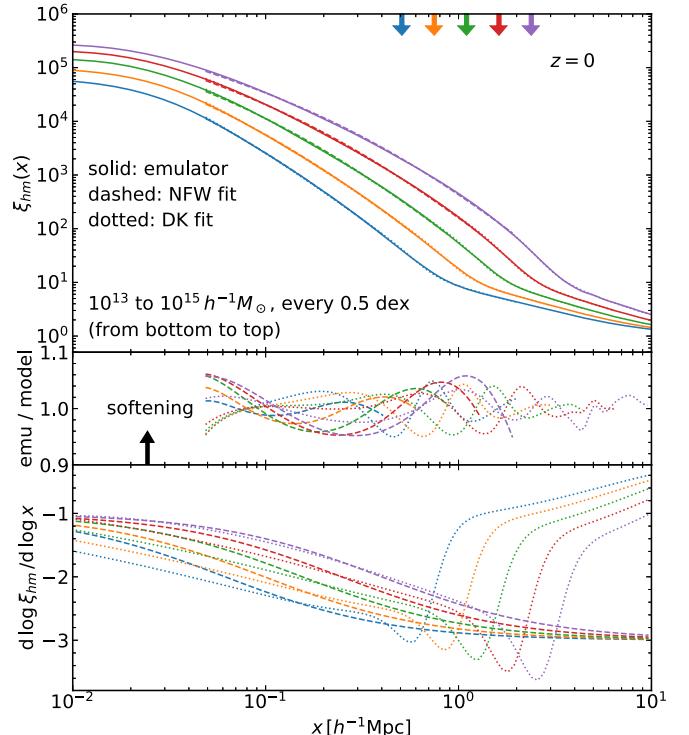


Figure 27. Top: model fit (dashed: NFW; dotted: DK) to the density profile around halos predicted by the emulator (solid). We stick to the fiducial *Planck* cosmology at $z = 0$ and consider various halo masses as shown in the legend. Middle: ratio of the emulator to the model fit in the top panel. Bottom: logarithmic derivative of the profile for the analytical fit. The radius R_{200m} for each sample is indicated by the downward-pointing arrows in the top panel. The softening length is shown by the upward-pointing arrow in the middle panel.

the galaxy–galaxy weak lensing and projected correlation function of galaxies from the mock catalogs. To be more precise, assuming the plane-parallel approximation, we project the matter and galaxy distributions along one of the three axes in each realization and then measure the galaxy–matter cross- and galaxy autocorrelation functions using the two-dimensional FFT, respectively. We use, as the prediction of the mock catalogs, the average of the 72 measurements (24 realizations \times 3 projection directions). Note that the fiducial *Planck* cosmology is not used in the GPR; thus, it should serve as a cross-validation test after the additional ingredients in UTILITY MODULES.

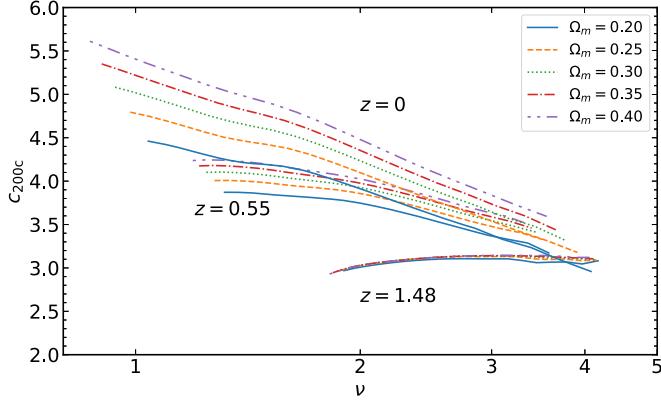


Figure 28. Concentration–peak height relation at various cosmological models and different redshifts.

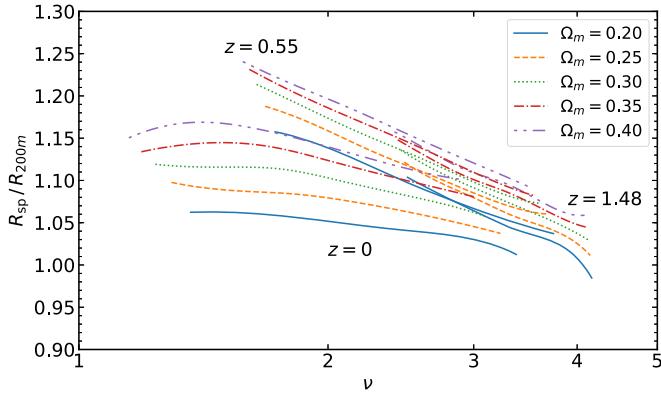


Figure 29. Splashback radius divided by R_{200m} .

The measurements from the mock catalogs are compared with the emulator predictions in Figure 30 for the galaxy-galaxy lensing (left panel) and the projected galaxy correlation function (right). In the upper panels, we show the emulator predictions with the solid lines for models with different Ω_m , as indicated in the figure legend. For the fiducial *Planck* cosmology, the dashed, dotted, and dotted-dashed lines show the different contributions in the model calculations; the contributions from central and satellite galaxies are shown for the galaxy-galaxy lensing, while the one- and two-halo term contributions are for the projected correlation function of galaxies. For the galaxy-galaxy lensing profile, we plot $\Delta\Sigma/\bar{\rho}_m$ for each of the different Ω_m models because it becomes the same dimension as that of w_p in the right panel. With this definition, both $\Delta\Sigma$ and w_p display a similar dependence on Ω_m ; increasing Ω_m leads to a smaller amplitude (we vary Ω_{de} and A_s , while the other four input cosmological parameters are fixed, to keep the spatial flatness, as well as the value of σ_8).

The lower panels explicitly compare the emulator-based predictions with the mock measurements, showing the ratio for each galaxy observable for the fiducial *Planck* cosmology. The gray shaded region around unity shows the statistical errors expected for the measurements. In the left panel, we assume, for the galaxy-galaxy weak-lensing measurement, the Subaru HSC first-year shape catalog (Mandelbaum et al. 2018) and the SDSS DR11 CMASS galaxies at redshifts around $z \simeq 0.484$ (Alam et al. 2015) for background galaxy shapes and foreground lensing galaxies, respectively, where the overlapping region of the two data sets is about 140 deg^2 . In the

right panel, we assume the projected correlation function of the CMASS galaxies for about 8500 deg^2 (More et al. 2015b). Each panel shows that the ratio is very close to unity, meaning a remarkable agreement between the emulator-based prediction and the mock measurement for each observable. Most importantly, the emulator-based predictions take just a few seconds of CPU time. The wiggly features in the ratio, especially for the projected correlation function, are due to imperfect accuracy in the numerical calculation, such as the numerical integration of the emulator outputs over halo masses. The gray shaded region gives statistical errors at each radial bin, estimated from the mock catalogs, where we assumed 30 bins over the range of $0.057 h^{-1}\text{Mpc} \leq R \leq 71 h^{-1}\text{Mpc}$ corresponding to $\Delta \log_{10} R \simeq 0.1$. The dark shaded region gives an overall requirement on the uncertainty in the model prediction of each observable. The requirement is estimated from the inverse of the total S/N integrated over all of the radial bins. Since we find $S/N \simeq 25$ and 35 for the weak-lensing and projected correlation function, respectively, the requirement on the overall factor in the model prediction, i.e., m for $\Delta\Sigma = (1 + m)\Delta\Sigma_{\text{emulator}}$ or $w_p = (1 + m)w_{p,\text{emulator}}$, is $m \lesssim 0.04$ or 0.029, respectively, such that an uncertainty in the model prediction does not exceed the overall statistical error by more than 1σ . The figure shows that the accuracy of DARK EMULATOR safely meets the requirements for the Subaru HSC and SDSS measurements. We note that, since variations in cosmological parameters cause a scale-dependent change in these observables, the requirements for such changes are less stringent.

5.3. Cross-correlation Coefficient

There remains another interesting and important check. It would be of great practical use if we could infer the underlying matter clustering properties from biased fields alone. One can compute the cross-correlation coefficient between matter and halo or between two different halo samples from our emulators.²⁰

First, we show in Figure 31 the cross-correlation coefficient for halo samples selected by mass at the fiducial *Planck* cosmology at $z = 0.484$. The curves are computed by DARK EMULATOR. We consider five different halo masses from 10^{13} to $1.6 \times 10^{14} h^{-1}M_\odot$. We show in the top panel three quantities, $\xi_{hh}(x)$, $\xi_{hm}(x)$, and $\xi_{mm}(x)$. We then take the ratio $\xi_{hm}(x)/\xi_{mm}(x)$ to examine the scale dependence of bias in the middle panel. Finally, the bottom panel depicts the cross-correlation coefficient. A similar plot can be found in Figure 32 for galaxies based on the HOD model described in the previous section. Here we also show the measurements from the mock galaxies distributed in the simulated halos based on the same HOD prescription (symbols with error bars). The emulator predictions are in agreement with the measurements from the mock galaxies.

The middle panels indicate that the scale dependence of bias is rather weak on scales larger than several $h^{-1}\text{Mpc}$ for all of the cases investigated here. On these scales, the cross-correlation

²⁰ To this end, we need an emulator to compute the matter autocorrelation function, which is not supported in the current version of DARK EMULATOR specifically designed for biased tracers. While the optimization for implementation detail or the final accuracies are not tested as stringently as the other modules, we have a development version of a module to do this. The figures in this section are based on this version, but it would be sufficient for demonstration purposes.

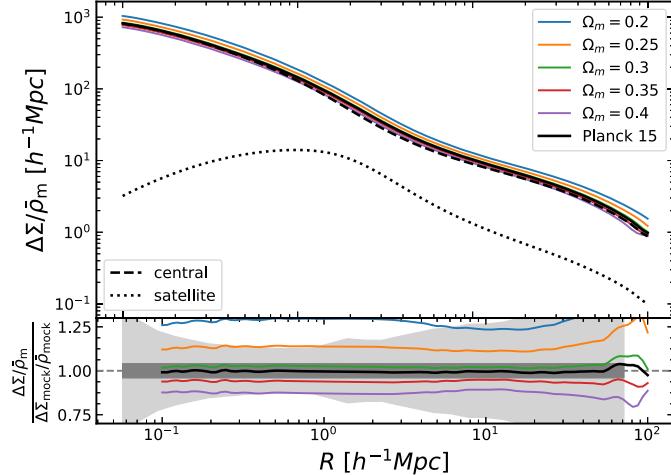


Figure 30. Example usage of DARK EMULATOR in combination with an HOD model for making model predictions of galaxy clustering observables. Left panel: excess surface mass density profile from galaxy–galaxy weak lensing ($\Delta\Sigma$). Right panel: projected correlation function of galaxies (w_p). In the upper panels, the solid lines show how the prediction varies with Ω_m but keeping other parameters fixed to the fiducial *Planck* values. The other lines show the different contributions to the total power as indicated by the figure legend (see text for details). The lower panels compare the emulator-based predictions with the “mock” signals measured from 72 mock realizations of projected maps of CMASS-type galaxies that are generated from the halo catalogs for the fiducial *Planck* cosmology (see text for details). The two results, obtained from totally different methods, are in remarkably nice agreement with each other. In the left panel, the gray shaded region shows the measurement errors expected when combining the Subaru HSC galaxies and the SDSS CMASS galaxies for background and foreground galaxies, respectively, where the overlapping region is about 140 deg^2 . In the right panel, we assume the measurement expected for the SDSS DR11 CMASS galaxies around $z = 0.484$ covering about 8500 deg^2 . The dark gray regions around unity give a requirement on the overall uncertainty in the model prediction, which is estimated from the inverse of the total S/N over $0.057 h^{-1}\text{Mpc} \leq R \leq 71 h^{-1}\text{Mpc}$: the requirements are about 0.04 and 0.029, corresponding to S/N = 25 and 35 for $\Delta\Sigma$ and w_p , respectively. The black line shows that the emulator predictions safely meet the requirements over the range of separation bins.

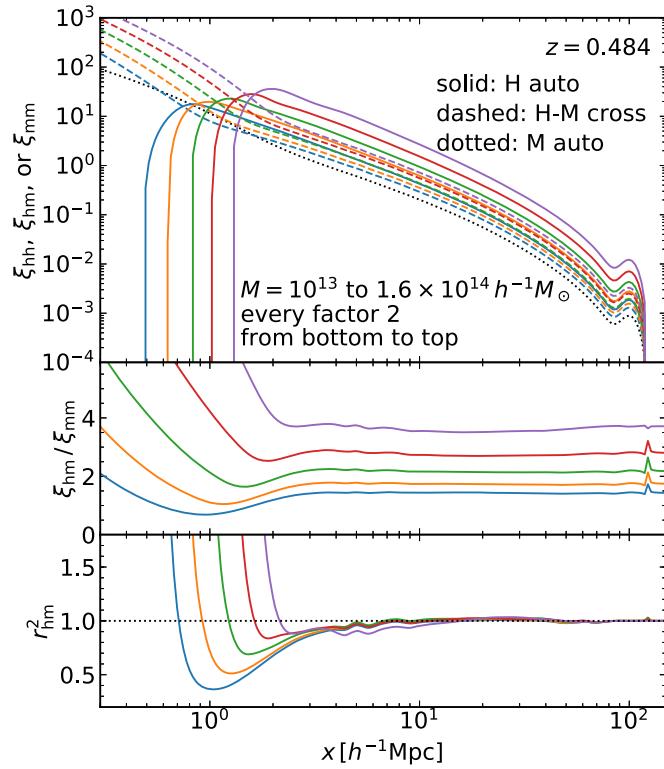
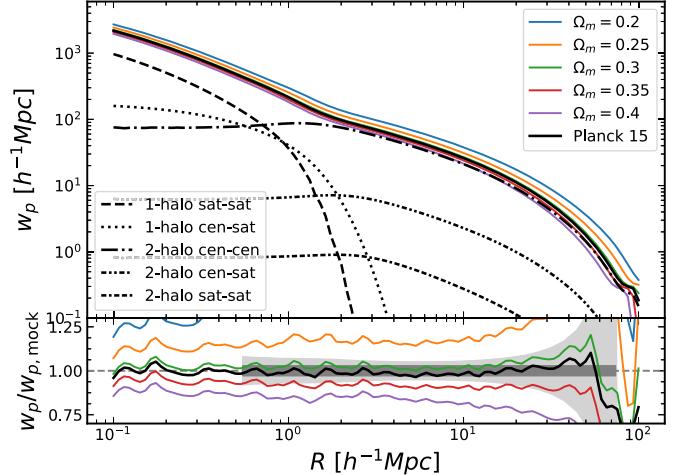


Figure 31. Various halo clustering statistics at $z = 0.484$. In the top panel, the solid (dashed) lines show the halo auto- (halo–matter cross-) correlation function for the masses indicated in the figure legend. We also show the matter correlation function (dotted line). The middle panel depicts the halo bias defined by the ratio $\xi_{\text{hm}} / \xi_{\text{mm}}$. The bottom panel shows the square of the cross-correlation coefficient for the halo samples.

coefficient is close to unity within $\sim 10\%$. Based on the results, we may extract the underlying matter clustering statistics by combining the auto- and cross-correlation functions of biased

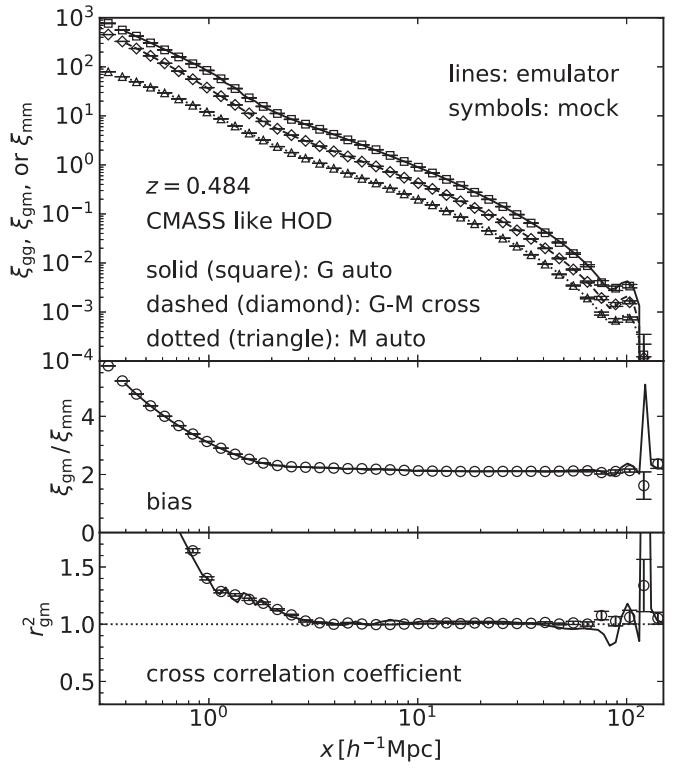


Figure 32. Similar to Figure 31, but for galaxy clustering with the HOD prescription that is the same as in Figure 30. We also show here the measurements of the same quantities from the mock galaxies distributed following the same HOD model (circles with error bars).

tracers. This statement would be true as long as we consider a simple model for the galaxy–halo connection as the HOD model considered here. Further studies are warranted to fully explore the potential to reconstruct the underlying matter clustering from real data for a wider class of galaxy populations.

Table 2
Summary of Model Parameters

Class	Parameter	Prior Range	Definition
Cosmology	$\Omega_b h^2$	[0.0211375, 0.0233625]	Physical baryon density parameter
	$\Omega_{\text{cdm}} h^2$	[0.10782, 0.13178]	Physical CDM density parameter
	Ω_{de}	[0.54752, 0.82128]	DE density parameter
	$\ln(10^{10} A_s)$	[2.4752, 3.7128]	Amplitude of primordial power spectrum
	n_s	[0.916275, 1.012725]	Spectral tilt of primordial power spectrum
	w	[-1.2, -0.8]	Equation-of-state parameter of DE
Common	z	[0, 1.47619]	Redshift
Halo	M or M_{th}^{a}	[10^{12} , 10^{16}]	Halo mass ^b in [$h^{-1} M_\odot$]
	M' or M_{th}'	[10^{12} , 10^{16}]	Halo mass of the second halo for the halo–halo power spectrum
	n	[$10^{-8.5}$, $10^{-2.5}$ ^c]	Halo number density ^c in [$(h^{-1}\text{Mpc})^{-3}$]
	n'	[$10^{-8.5}$, $10^{-2.5}$]	Number density of the second halo for the halo–halo power spectrum
HOD ^d	Seven-parameter model (default) $\{M_{\text{min}}, \sigma_{\log M}, \alpha_{\text{inc}}, M_{\text{inc}}, \kappa, M_1, \alpha\}$		
Profile ^e	NFW model (default) $\{c(M, z), f_{\text{off}}, R_{\text{off}}\}$		

Notes.

^a The halo–matter and halo–halo power spectra can be output for a sample of halos with a given number density n , at a given mass M , or with masses greater than a given mass threshold M_{th} (see text).

^b The emulator employs M_{200m} for halo mass definition.

^c A warning message can be output if the input number density is too high for an input set of cosmology parameters and redshift, i.e., if the input number density is outside the support of the emulator.

^d We employ the HOD given by seven parameters as a default prescription for halo and galaxy connection. A user can replace this module with another prescription if needed.

^e We assume that the distribution of satellite galaxies in their host halo follows a normalized NFW model, where the halo mass and concentration follow the fitting formula in Diemer & Kravtsov (2015), for our default model. Another option is to distribute satellite galaxies following the matter distribution around a halo as predicted by the halo–matter cross-correlation function. We also include a possibility that a fraction f_{off} of central galaxies is offset from the true halo center and assume that the normalized distribution, with respect to the true center, is a Gaussian with width radius R_{off} . A user can replace this module if needed.

5.4. Summary of the Current Code

Finally, we summarize the functionalities of DARK EMULATOR in this section.

First, the input parameters for the code are listed in Table 2. The items in the “Cosmology” class are the six cosmological parameters of the wCDM cosmologies considered in this paper. In the “Common” class, we have redshift z as a common parameter for all of the modules. Quantities calculated from linear theory are evaluated at $z = 0$ and then properly scaled by the linear growth factor. The third class is “Halo,” relevant for HALO MODULES. The parameters in this class specify a halo sample in terms of either a mass range or a specific mass scale at which the desired halo clustering quantities are evaluated. The number density and mass threshold can be converted to each other using the module that computes the HMF. Next, we have the “HOD” class. The parameters here determine how many galaxies (centrals and satellites) are populated in the halos. Finally, we have a set of parameters that model variations in the locations of the galaxies inside a halo. For instance, we allow central galaxies to be off from the true halo center using the two off-centering parameters. The satellite galaxies are assumed to follow either the NFW profile or the average halo mass profile, where the latter is equivalent to the halo mass cross-correlation for halos of each mass range that is an output of our emulator. In the case of the NFW profile, we assume the concentration–mass relation calibrated by Diemer & Kravtsov (2015).

The HOD model, as well as the profile of galaxies, can be modified easily when needed. This way, we allow the code to

have the flexibility to support various galaxy populations, possibly beyond the model currently implemented. It has been suggested that the clustering statistics could depend on a secondary parameter beyond the halo masses. The so-called halo assembly bias is not implemented in the current code. We study the impact of such effects on cosmological analyses in a separate paper. Our code can also account for the effect of the residual redshift-space distortions on the projected statistics with a finite projection width assuming linear theory, as well as a modification of the mass profile around halos due to baryonic effects in a parametric manner. Again, these effects are studied in full detail in a separate paper.

Finally, the outputs of the emulator are summarized in Table 3. The first set of outputs is the three linear quantities based on LINEAR MODULES. The primary outputs of the emulator are the abundance and clustering of halos and matter. These include the abundance of halos in a given mass range, halo–matter cross-correlation function, halo autocorrelation function, and propagators of matter and halo. These quantities are then combined and projected based on analytical calculations to eventually have the items in the “Derived” class. The connection between halos and galaxies as specified in the UTILITY MODULES (i.e., HOD and profile) is reflected to the final galaxy statistics.

6. Summary

In this paper, we have performed an N -body simulation ensemble, dubbed DARK QUEST, and developed an emulator enabling a fast computation of halo clustering quantities from

Table 3
Summary of the Emulator Output

Class	Output	Definition
Linear	$\sigma^2(M, z)$	Linear mass variance
	$\sigma_d(z)$	rms linear displacement in one dimension
	$P_{\text{lin}}(k, z)$	Linear matter power spectrum
Primary	$n(M_{\min}, M_{\max}; z)$	Number density of halos in the mass range $[M_{\min}, M_{\max}]$
	$\xi_{\text{hm}}(x; M, z)$	3D halo-matter cross-correlation for halos
	$\xi_{\text{hh}}(x; M, M', z)$	3D halo autocorrelation for halos of masses M and M'
	$G_m(k; z)$	Propagator for the matter density field
	$G_h(k; M, z)$	Propagator for the halo density field specified by mass M
Derived	$\Sigma_{\text{hm}}(R; M, z)$	Surface mass density profile of halos of M
	$\Delta\Sigma_{\text{hm}}(R; M, z)$	Excess surface mass density profile around halos of M
	$\Delta\Sigma_{\text{gg}}(R; z)$	Excess surface mass density profile around galaxies
	$w_{\text{gg}}(R; z)$	Projected correlation function of galaxies

the simulation outputs, named DARK EMULATOR. The main features of our products are as follows.

1. We employed 2048^3 particles in either 1 or 2 Gpc h^{-1} comoving boxes, covering 100 six-parameter $w\text{CDM}$ cosmological models sampled via the SLHD around a fiducial ΛCDM cosmology. The mass density fields and catalogs of halos with $M_{200} \gtrsim 10^{12} h^{-1} M_\odot$ (slightly depending on cosmological models) were extracted at 21 redshifts in the range of $z = [0, 1.48]$. The parameter space covers a sufficiently broad range of parameters that are consistent with the existing cosmology data sets.
2. We used the DARK QUEST data sets to build DARK EMULATOR. It models the HMF, halo-matter cross-correlation, and halo autocorrelation based on the GPR after significant dimension reduction via the PCA. The predicted halo clustering properties are easily combined assuming a model for the halo-galaxy connection, such as an HOD prescription, to compute the galaxy statistics.
3. We carefully validated the accuracy of the DARK EMULATOR predictions (outputs) using validation samples of N -body simulations for cosmological models that are not used in the emulator development. The validation samples are also located following the LHD and are maximin design by themselves, in combination with the training samples. Thus, they allow us to test the accuracy at distant points from the nearest training data and, at the same time, uniformly cover the whole domain of the parameter space.
4. We achieved 1%–2% accuracy for the HMF (the rms error over the 20 models), except for the massive end ($M \gtrsim 10^{14} h^{-1} M_\odot$), where the Poisson error is significant in both the training and the validation sets. The accuracy for the halo-matter cross-correlation function for a halo sample with number density $10^{-4} (h^{-1} \text{Mpc})^{-3}$, which resembles the typical host halos of LRGs or CMASS-like galaxies, was shown to be $\sim 2\%$ over the comoving separation $0.1 h^{-1} \text{Mpc} < x < 30 h^{-1} \text{Mpc}$ (again in terms of the rms error). The halo-halo autocorrelation function

for the same halo sample has a slightly larger error, $\sim 3\%-4\%$, reflecting the shot-noise error. The accuracy gets worse at $x \lesssim 1 h^{-1} \text{Mpc}$, where the halo exclusion effect is significant and thus does not contribute much to galaxy clustering signals. In all cases, the biggest discrepancy between the prediction and the validation set is not worse than 5% over the ranges of halo masses and separations.

5. The accuracy of the emulator depends on the halo mass and slightly on the redshift. This can be checked in Appendix F. We find overall that the validation accuracy scales consistently with the sample variance error estimated from the multiple random realizations prepared for the fiducial cosmology. Thus, further significant improvement of the accuracy would be possible only by using more simulations (with a larger box size in addition), while further refinement of the implementation detail would not at this moment.
6. We introduced a special treatment based on the propagator to large-scale clustering signals where the large sample variance prevents us from an accurate modeling or validation test. The propagator encodes the large-scale bias, as well as the damping of the BAO feature, and it can be measured accurately, as the sample variance mostly cancels in its estimator. A module that emulates the propagator is also trained and validated to ensure the accuracy of the predictions of the correlation functions on large separations.
7. We demonstrated that the DARK EMULATOR outputs can be used to study detailed properties of the mass density profiles around halos, such as the concentration-mass relation and the splashback feature. The emulator can predict their dependence on redshift, halo mass, and cosmological models.
8. We also demonstrated that the emulator outputs can be used to predict, as an example, the projected galaxy correlation function and the galaxy-galaxy weak-lensing profile when combined with a prescription of the halo-galaxy connection, such as an HOD model. In doing this, we can easily incorporate variants of small-scale effects such as the off-centering of galaxies with respect to the halo center, the incompleteness selection of galaxies, and the distribution of satellite galaxies in their host halo based on a Fourier space implementation. The DARK EMULATOR modules extensively use the FFTLog algorithm that enables a fast computation of converting the three-dimensional correlation functions to the projected correlation functions.
9. The evaluation time for the halo and galaxy statistics is typically of order ~ 100 ms and a few seconds, respectively, on a standard laptop computer available today. The latter is slower because it usually involves integrals over the halo masses.
10. The cross-correlation coefficients between halos and matter are shown to be quite close to unity on large scales. This remains the same for galaxies populated into halos based on the HOD description.

The current implementation and accuracy are likely sufficient for ongoing wide-area galaxy surveys such as the Subaru HSC survey (see Figure 30 for its validation). It is still not clear how the cosmological information can be extracted from the cosmological dependences of halo clustering quantities that are measured from such a galaxy survey, even after marginalizing

over nuisance parameters that model the small-scale clustering in the one-halo term. To address this, one has to use realistic mock catalogs that resemble the actual galaxy survey; measure clustering observables of interest from the mock catalogs, including all realistic small-scale effects; and make a hypothetical parameter inference from the comparison of the emulator predictions with the mock measurements, including marginalization of the nuisance parameters (see Hand et al. 2017, for a similar discussion). This kind of study can assess the power and usefulness of DARK EMULATOR for precision cosmology and gives a validation of the parameter inference method–cosmology challenges. This is our ongoing project and will be presented in the future (Miyatake et al. in preparation). Our implementation to incorporate the residual redshift-space distortion, as well as the baryonic effects, into the mass profile around halos will also be presented and tested in that paper.

However, the current emulator would not meet the accuracy required for future surveys such as *LSST*, *Euclid*, and *WFIRST*. As already mentioned above, a naive and straightforward way is to accumulate more simulation data and reduce the statistical uncertainties on the training data. Since the current implementation of DARK EMULATOR includes several approximate treatments, the systematic error from them can be a problem with the improved statistical error. These include the following.

1. The sample variance error on the measured statistical signals is assumed to be diagonal (i.e., no off-diagonal covariance) and independent of cosmological models.
2. The PCA coefficients are modeled by GPR one by one, ignoring the correlation between them.
3. The metric in the cosmological parameter space is assumed to be stationary (i.e., independent of the location in the space).
4. The functional forms assumed in the HMF and the propagator might be insufficient for ultimate precision.
5. Although the damping of the BAO peak is already included, a possible “shift” of the BAO scale due to nonlinearity is ignored.
6. Extra dependence of the halo clustering properties other than the mass dependence, i.e., halo assembly bias, is not considered at all.
7. The current emulator supports halos with mass $\gtrsim 10^{12} h^{-1} M_{\odot}$. This should be improved to model, e.g., emission line galaxies that form in less massive halos.
8. The suite of N -body simulations and halo catalogs can also be used to study the intrinsic alignment (IA) of halo shapes and their dependences on cosmological models, halo mass, and redshifts. The IA is not only one of the major systematic errors in high-precision weak-lensing measurements but can also be a new cosmological probe as it arises from large-scale structures. This is our future project and will be presented elsewhere.

Nevertheless, we are optimistic for such a challenge. As we stressed, we designed DARK EMULATOR to cover a sufficiently broad range of cosmological models within w CDM cosmologies, which are much broader than the models favored by the *Planck* CMB measurements, because we want to keep a broader range of applications of DARK EMULATOR to problems that users might want to study. If the range of cosmological parameters is narrowed down, and if specific requirements for given clustering observables for a future survey under consideration are given, some of the approximations, such as

the cosmology-independent modeling of the statistical error or the stationary metric, would be even more appropriate. We can also design a new set of N -body simulations to run in the new narrower parameter space and then construct an emulator that can meet the requirements. We believe that the methods and techniques developed in this paper would be useful to explore such an N -body simulation suite and then develop a sufficiently accurate emulator enabling one to predict the clustering observables that one wants to use for precision cosmology. The current version of DARK EMULATOR will be made public in the near future.

We appreciate useful comments by Salman Habib and Katrin Heitmann on the numerical simulations and Gaussian process. We also thank Surhud More for useful discussion during the early stage of this work and providing us with the HOD model for the SDSS galaxies. We appreciate useful discussion with Shiro Ikeda and Naonori Ueda on Bayesian techniques and experimental design. This research was supported by the World Premier International Research Center Initiative (WPI), MEXT, Japan. This work was supported in part by the MEXT Grant-in-Aid for Scientific Research on Innovative Areas (Nos. JP15H05887, JP15H05892, JP15H05893, and JP15K21733), Japan Science and Technology Agency CREST JPMHCR1414, MEXT Priority Issue 9 on Post-K Computer (Elucidation of the Fundamental Laws and Evolution of the Universe), and JICFuS. This work was also supported by JSPS KAKENHI grant Nos. JP17K14273 (T.N.), JP15H03654 (M.T.), JP17H01131 (R.T.), JP16J01512 (K.O.), JP18H04358 (M.S.), JP18H04350 (H.M.), JP18K03693 (M.O.), and JP17J00658 (R.M.). Numerical computations were carried out on the Cray XC30 and XC50 at the Center for Computational Astrophysics, National Astronomical Observatory of Japan.

Appendix A Linear Modules

While the linear theory predictions of cosmological structure formation can be obtained accurately and quickly using public Boltzmann codes such as CMBFAST (Seljak & Zaldarriaga 1996), CAMB (Lewis et al. 2000), and CLASS (Blas et al. 2011; Lesgourgues 2011), the computation time is still nonnegligible, e.g., for parameter inference using Markov chain Monte Carlo in a high-dimensional parameter space. Within HALO MODULES, we need to evaluate the linear power spectrum $P_{\text{lin}}(k)$, the mass variance σ_M^2 , and the rms displacement σ_d for the input set of cosmological parameters, halo mass, and redshift. In particular, the latter two involve an integral over wavenumber. To speed up the computations, we developed an emulator module that allows for a quick computation of these quantities based on PCA and GPR methods, as we did in HALO MODULES.²¹

We first sample 400 sets of cosmological parameters, each of which is taken within the range of parameters given by Equation (25), using the SLHD scheme. In this case, we generate 10 slices with 40 samples each. We use the CLASS code to compute the relevant linear theory quantities for each model, although we have used CAMB for the initial conditions of our N -body simulations. While it is known that the results of CAMB and CLASS can differ slightly depending on their

²¹ See Fendt & Wandelt (2007a, 2007b) for a similar attempt to speed up the calculation of linear power spectra.

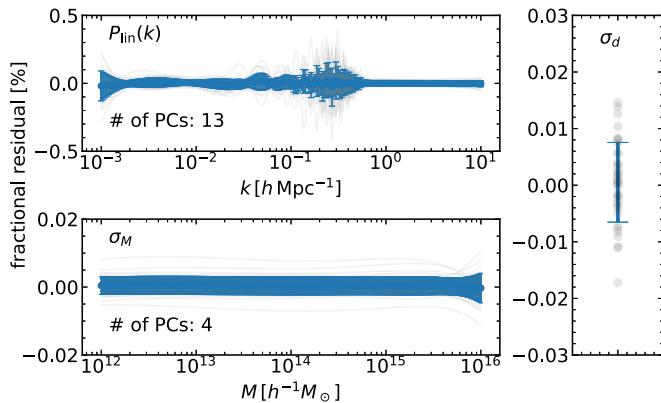


Figure 33. Cross-validation study of the LINEAR MODULES. We show the fractional residual in percent, comparing the emulator prediction to the direct evaluation by CLASS for each of the 40 validation cosmological models in an SLHD slice by gray lines or circles in each panel (upper left: $P_{\text{lin}}(k)$; lower left: σ_M ; right: σ_d). In each panel, we also show (error bars) the scatters among the 40 models (1σ level).

accuracy parameters, the difference is typically at a subpercent level, which is much below our target accuracy here. We use the 360 cosmological models in nine slices as the training data and test the accuracy of the emulator using the remaining 40 models in the last slice as a validation sample. We reduce the dimensionality of the data vector for $P_{\text{lin}}(k)$ and σ_M by keeping only the most significant PCs. We originally sample $P_{\text{lin}}(k)$ (σ_M) by 200 points over wavenumber (401 points over mass) and keep only 13 (4) PCs. Since σ_d is a single number, we use it as it is. Finally, the coefficients of the PC eigenvectors (or σ_d) are modeled by GPR.

The accuracy of our model is assessed by cross-validation and shown in Figure 33. The accuracy is always better than 1%, with $P_{\text{lin}}(k)$ generating the biggest error of $\sim 0.3\%-0.4\%$. The accuracy for the other two quantities is even better, typically at the 0.01% level. The bigger error on $P_{\text{lin}}(k)$ is attributed to the characteristic features of the BAOs. This is contrasted with the rather smooth and monotonic dependence of σ_M on M . The current implementation of the LINEAR MODULES enables us to evaluate all of these quantities in a few milliseconds for an input cosmological model, which is a negligible time in the whole calculation of DARK EMULATOR.

Appendix B Initial Conditions of N -body Simulations

B.1. Optimal Initial Redshift

In this appendix, we discuss how the choice of redshift used to set the initial conditions of N -body simulation affects the results of late-time clustering. While a higher initial redshift is preferable to reduce the transient effect arising from the fact that the initial conditions do not follow the growing solution precisely at higher orders (Crocce & Scoccimarro 2006b), the regular lattice pre-initial configuration can excite spurious modes (e.g., Marcos et al. 2006; Joyce & Marcos 2007; Garrison et al. 2016) if the starting redshift is very high. The latter effect can be understood via “particle linear theory” (Marcos et al. 2006): the growing solution of particles close to the lattice configuration is different from the fluid growing solution in a direction-dependent manner. The net effect after averaging over the direction is to slow down the growing

modes compared to what the fluid linear theory predicts, and this becomes more important toward larger wavenumbers. While a numerical method to correct for this effect was recently proposed by Garrison et al. (2016), we here adopt a simpler approach. Since both the effects suppress the structure growth, we choose an initial redshift so that the N -body simulation produces the highest power spectrum at late times.

We first focus on the evolution of the matter density contrast at early epochs. We generate particle distributions from an identical random realization of the linear density field at different initial redshifts ($1 + z_{\text{in}} = 15, 30, 60, 120$, and 240) using either the Zel'dovich approximation or the 2LPT. We implement this numerical experiment using N -body simulations in a cubic volume with a side length of $L = 250 h^{-1} \text{Mpc}$ employing two different resolutions, one with 512^3 and the other with 256^3 particles, corresponding to the resolution of HR and LR runs, respectively. We run N -body simulations assuming the different initial conditions and then measure the matter power spectrum at later epochs. In doing this, we store the snapshots of N -body simulations at different epochs starting from $z = 49.75$ corresponding to the linear growth factor $D_+ = 0.025$ down to $z = 0$ at every interval of $\Delta D_+ = 0.025$ (40 snapshots in total).

Figure 34 shows the ratio of the measured power spectra to the prediction by a Eulerian perturbation theory that is computed on a grid basis assuming the same random realization of the simulation up to the two-loop order using FFT (GRIDSPT in Taruya et al. 2018). Note that in the computation, we have included odd-order contributions to the power spectrum, which should vanish in the ensemble average sense but are present in a finite volume (or a given realization) where we have a limited number of Fourier modes. We show the evolution as a function of the linear growth factor $D_+(z)$ normalized to unity at present. Each line starts at the initial growth rate corresponding to the initial redshift denoted in the legend. The figure shows two overall trends. First, for some results, the growth in the power spectrum is sharply suppressed compared to the perturbation theory prediction soon after the initial redshift, even if the perturbation theory should be accurate at very early epochs, especially at small wavenumbers. The higher the initial redshift with which we start the simulation, the greater suppression the growth has. The figure also shows that the effect is more important at higher wavenumbers (note the different plotting ranges in different panels): it is only at an $\sim 0.2\%$ level at $k = 0.03 h \text{ Mpc}^{-1}$, and it reaches to $\sim 1\%$ at $k = 0.53 h \text{ Mpc}^{-1}$ for the simulation with 512^3 particles (left panel). Comparing the left and right panels shows that the sudden drop in the power is about twice as large for the case with 256^3 particles than that with 512^3 . The same trend can be seen for both the 2LPT and the Zel'dovich Approximation (ZA) initial conditions; thus, this effect is not associated with the accuracy of the Lagrangian perturbation theory that is used to set up the initial displacement field. Furthermore, although we do not show it here, the suppression in the power persists even if we choose more stringent parameters to control the accuracy of the N -body simulations (in both the force computation and time stepping). All of these features indicate that this effect is ascribed to the particle discreteness effect.

Second, Figure 34 shows that the ratio gradually decreases with time after the first sharp decrease, meaning that the structure grows slowly compared to the perturbation theory

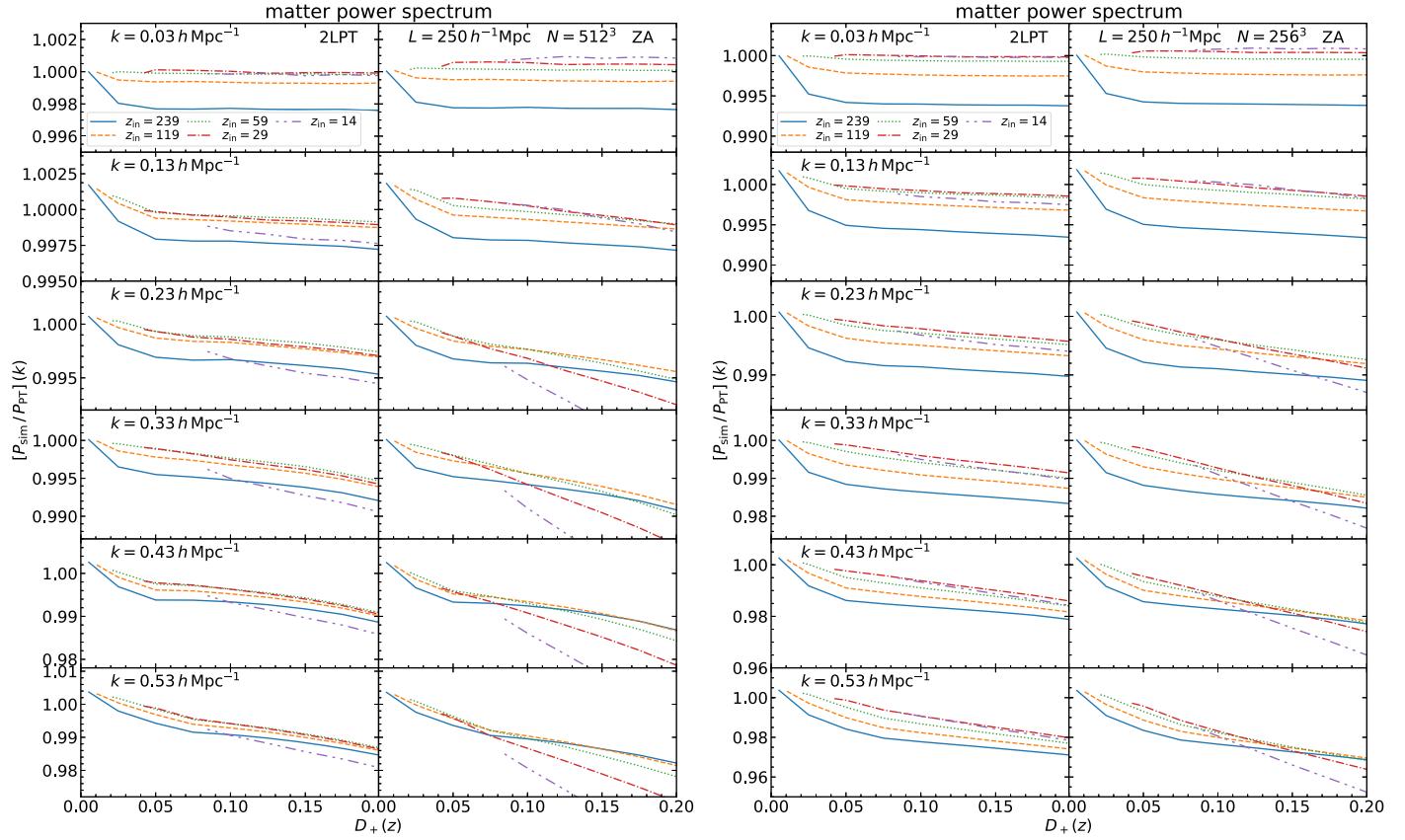


Figure 34. Time evolution of the matter power spectra in N -body simulations relative to the perturbation theory predictions (see text for details), where we use the simulations with 512^3 or 256^3 in a cubic volume with a side length of $250 h^{-1}\text{Mpc}$ in the left or right panel, respectively. Note that the simulations have the same resolution as those of the HR (LR) runs. We plot the ratio as a function of the linear growth rate that is normalized to unity at present. Different types of lines correspond to different initial redshifts of the simulations, as denoted in the legend. In each panel, we show the results of 2LPT and ZA initial conditions in the left and right columns, respectively, which are used to set up the initial displacements of N -body particles.

(i.e., the slope of the ratio is negative). While the slope is almost zero for $k = 0.03 h \text{ Mpc}^{-1}$, it becomes increasingly negative toward larger wavenumbers. The apparent slow growth of the power spectrum in the simulations is at least partly due to a breakdown of the perturbation theory at later epochs, at $k \gtrsim 0.1 h \text{ Mpc}^{-1}$; that is, the perturbation theory overpredicts the power spectrum amplitudes at such high k values. More important is the difference in the slope of the curves for simulations with different z_{in} . It can be seen that the dependence is subtle for 2LPT and quite significant for ZA. In general, the slope is steeper for a smaller z_{in} . This is the transient effect due to the fact that we truncate the perturbative calculation for the displacement field at a finite order.

Now the question is how these systematic effects due to the initial conditions affect the outputs of N -body simulations at the late epochs, $z \lesssim 1.5$, in which we are most interested. To explicitly study this, Figure 35 shows the power spectrum at $z = 0$ measured from the simulations at various wavenumbers as a function of $1 + z_{\text{in}}$ (left panel). For simplicity, we show only the results using the 2LPT initial conditions here. Importantly, the figure shows that the systematic effects that we find at the early-time evolution persist even at late times. For the simulations with 512^3 particles (upward-pointing triangles), the power spectrum has the greatest amplitudes at all wavenumbers when the initial redshift $z_{\text{in}} = 59$ is employed. The peak redshift is shifted toward lower redshift for the simulations with 256^3 particles (downward-pointing

triangles). This peak structure is a result of the competition of the two systematic effects that we have discussed above. The suppressed power for the higher z_{in} than the peak redshift is due to the particle discreteness effect, while the inaccuracy for the lower z_{in} is due to the insufficient nonlinear evolution in simulations. Since the former effect should scale as the typical initial displacement of particles from the regular lattice in units of the lattice interval, we plot in the right panel the same power spectrum as a function of $\sigma_d(z_{\text{in}})/\Delta_{\text{part}}$, where σ_d is the rms of the initial particle displacements and Δ_{part} is the mean interparticle distance. The peak locations in the power spectrum amplitude for the two resolutions are almost identical when plotted as a function of this combination. To be more quantitative, the peak location appears when the rms displacement is about 20%–30% of the interparticle separation. We thus simply adopt a z_{in} that gives $\sigma_d(z_{\text{in}})/\Delta_{\text{part}} = 0.25$ for the main simulations, HR and LR, presented in this paper; these correspond to $z_{\text{in}} = 59$ and 29, respectively. Note that these conclusions hold for neighboring cosmological models around the fiducial *Planck* model, but z_{in} could vary significantly depending on cosmological modes.

B.2. Impact on Halo Statistics

Figure 36 shows how the HMF measured from simulations at late times varies with initial redshifts. The HMF does not largely vary with different initial redshifts as long as the 2LPT

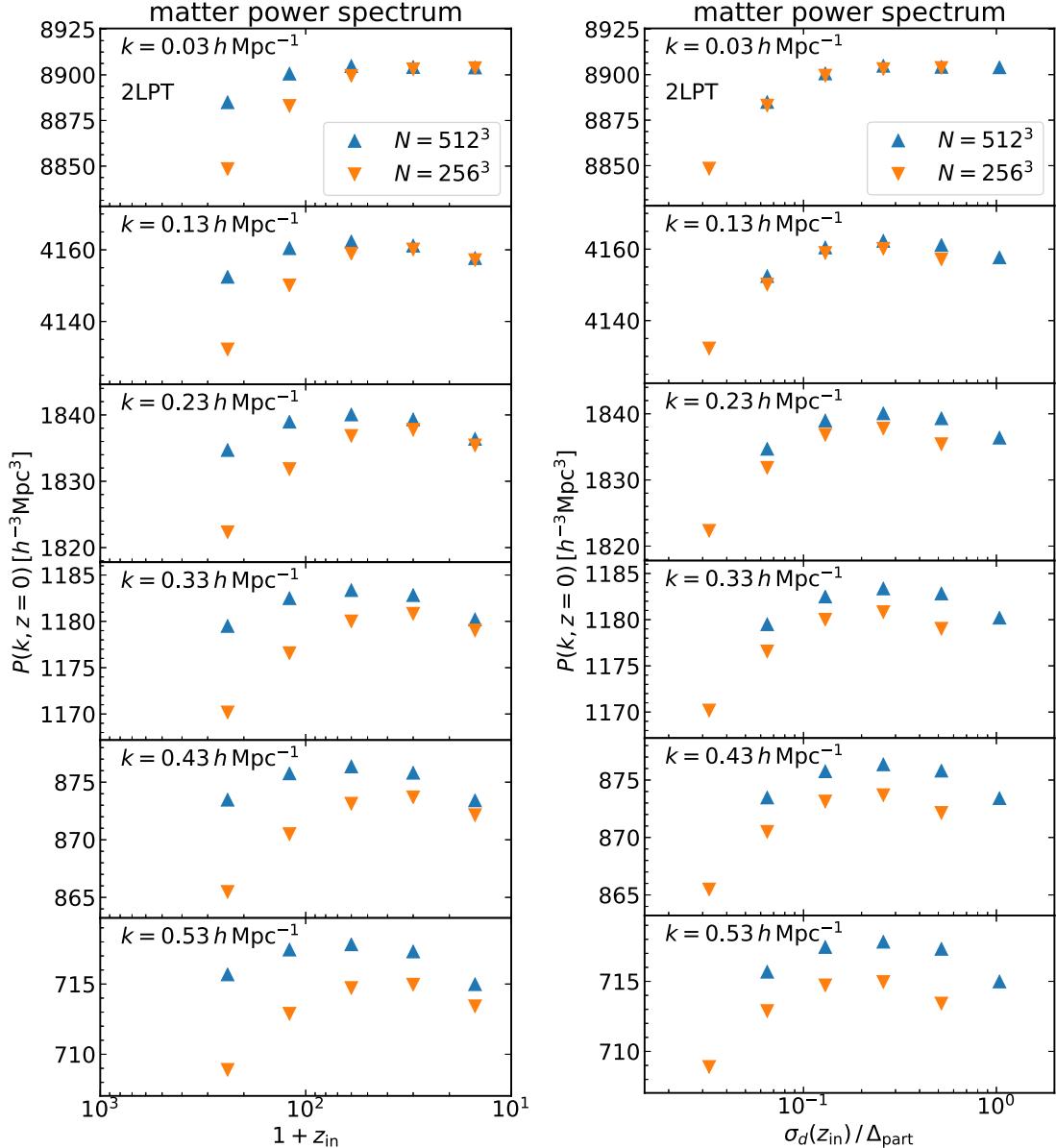


Figure 35. Dependences of the matter power spectrum amplitudes at $z = 0$ on the initial redshift of the simulation for different wavenumbers. The left plot shows the results as a function of the initial redshift of the simulation, whereas the right plot shows the results as a function of $\sigma_d/\Delta_{\text{part}}$, where σ_d is the rms of the initial particle displacements at the initial redshift and Δ_{part} is the mean interparticle distance. Even for the same initial random seeds, the power spectrum amplitudes vary for different initial redshifts, and the amplitude peaks at a particular initial redshift (see text for details).

instead of ZA is used to set up the initial conditions. Similarly, Figures 37 and 38 show how the halo–matter cross- and halo–halo autocorrelation functions vary when using simulations with different initial redshifts. Here we consider halo samples with a number density of $10^{-3}(h^{-1}\text{Mpc})^{-3}$ instead of the

fiducial value of $10^{-4}(h^{-1}\text{Mpc})^{-3}$ (i.e., including less massive halos compared to the fiducial analysis) to investigate the case where the systematic effects would be more important. The figures show that the results are well converged for a range of redshifts if the 2LPT initial conditions are used.

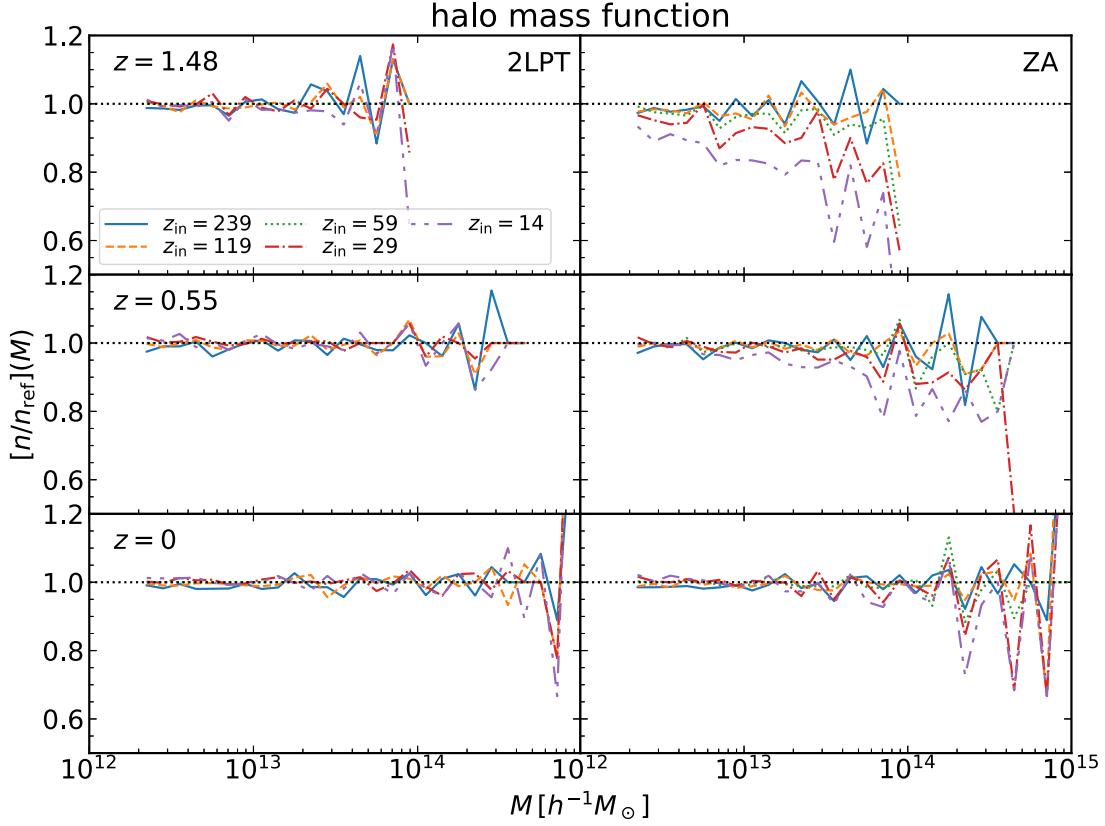


Figure 36. Dependence of the HMF on the initial redshift. For our fiducial setting, we adopt $z_{\text{in}} = 59$ for the HR simulation (equivalent to 512^3 particles for a box of $250 h^{-1}\text{Mpc}$) and use 2LPT to set up the initial displacements. The panels in the left column show the results for 2LPT and different initial redshifts relative to the fiducial result, while those in the right column are the results for the ZA initial conditions and different initial redshifts (but with the same resolution). The line styles indicate the initial redshift, as shown in the legend.

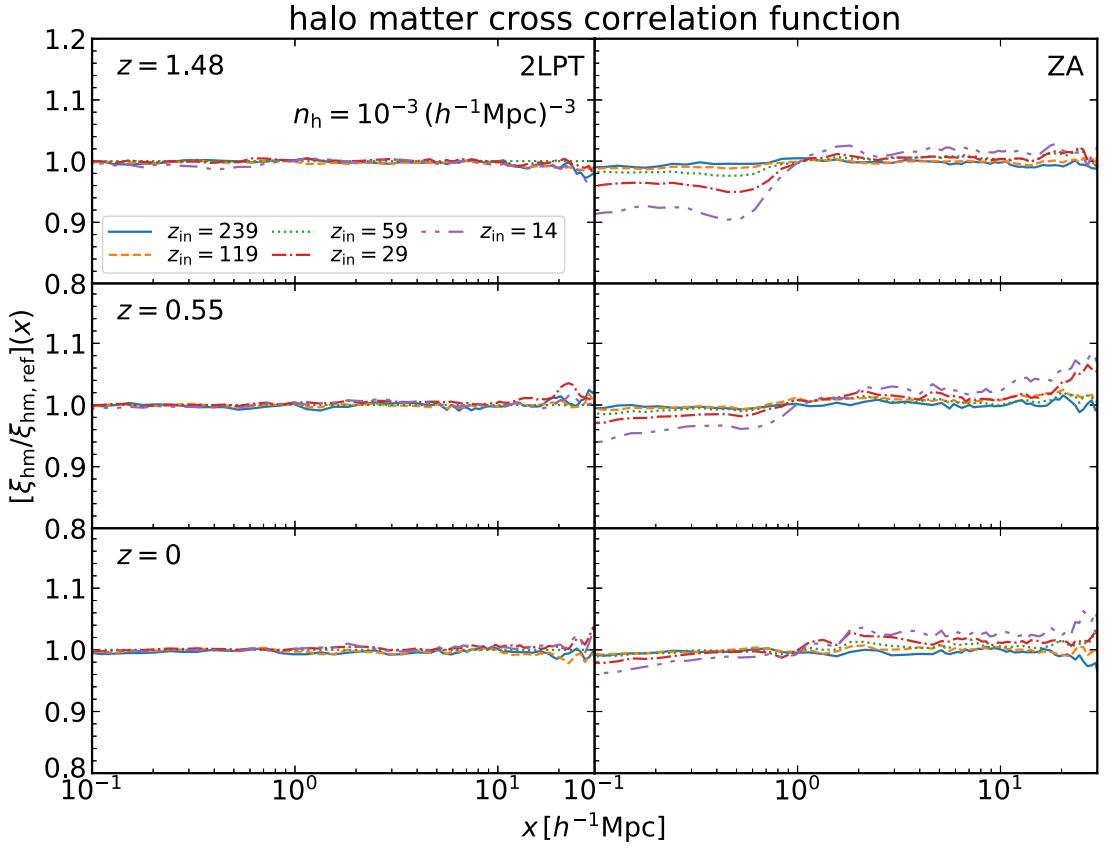


Figure 37. Similar to Figure 36, but for the halo–matter cross-correlation functions. Here we consider a sample of halos with a number density of $10^{-3} (h^{-1}\text{Mpc})^{-3}$.

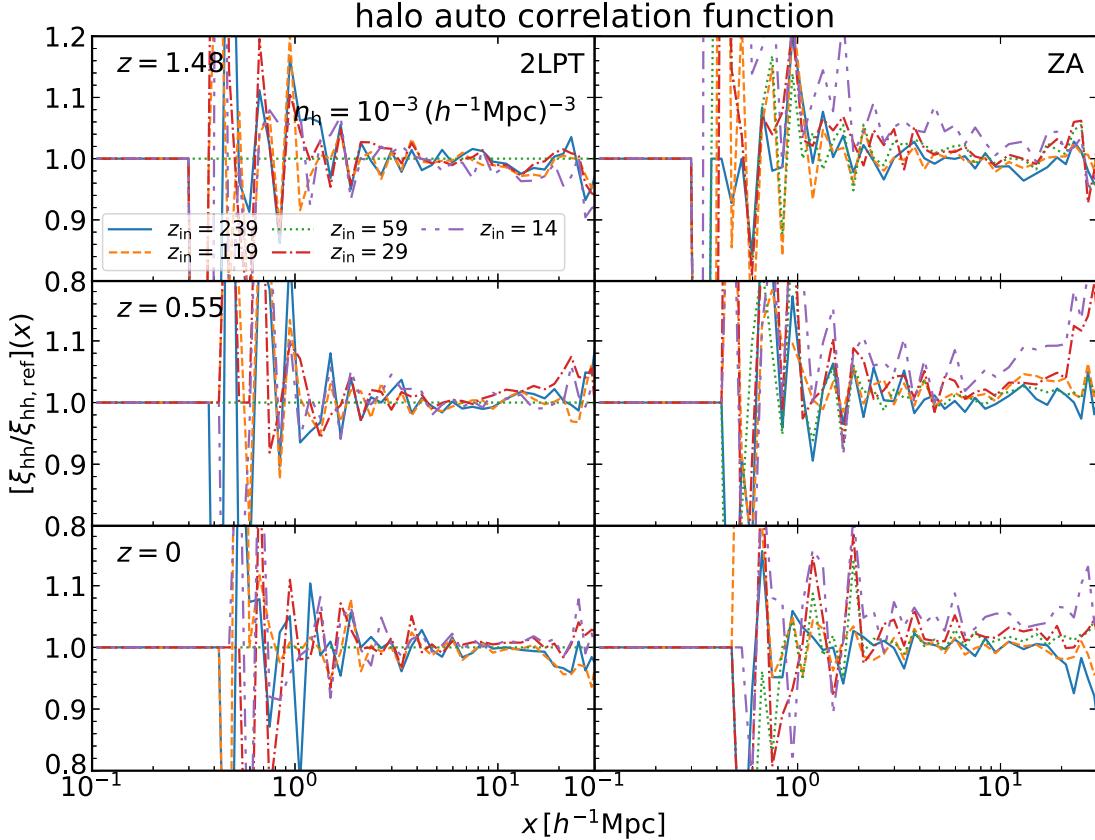


Figure 38. Similar to Figure 36, but for the halo autocorrelation functions. Here we consider a sample of halos with $10^{-3} (h^{-1}\text{Mpc})^{-3}$ in computation of the autocorrelation functions.

Appendix C Random versus Fixed Phase Simulations

For the Gaussian random initial conditions, the amplitude $|\delta_{\text{lin},k}|$ and the phase θ_k of the initial condition, when expressed as $\delta_{\text{lin},k} = |\delta_{\text{lin},k}|e^{i\theta_k}$, follow the Rayleigh or uniform probability distributions, respectively, for each k mode, where the width of the Rayleigh distribution is set by the initial power spectrum. A random number seed is often used to generate $|\delta_{\text{lin},k}|$ and θ_k for each k mode, i.e., the initial density field in an N -body simulation for a given cosmological model.

When studying dependences of nonlinear structure formation on cosmological models with N -body simulations, there is a choice of whether or not one keeps the same random seeds to set up the initial conditions for different cosmological models. A possible advantage of using the common random seeds is to reduce the sample variance contamination when comparing the clustering quantities between different cosmological models. We expect the advantage for neighboring cosmological models around a specific target model. However, for two models that are sufficiently far from each other, nonlinear evolution via complex mode coupling could produce significantly different results so that the naively expected variance “cancellation” is ruined. Thus, there is no guarantee that using the same random seeds leads to converged estimations of statistical quantities from a limited number of simulation realizations.

To test whether a development of the emulator benefits from the fixed-seed simulations, we compare the GPR results obtained from two sets of 20 simulations performed on slice 1; one set contains 20 simulations using the same fixed seed, while the other set is from 20 simulations with varied random

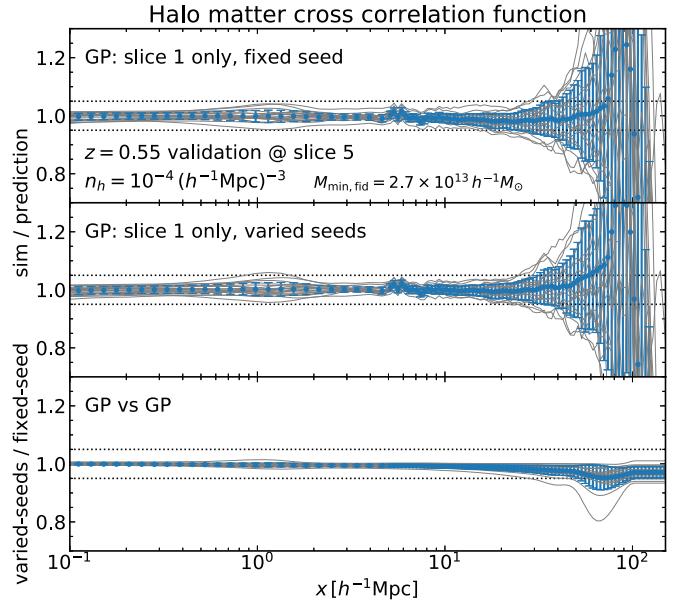


Figure 39. Comparison of the accuracy of the emulators built by using the 20 different cosmology simulations with fixed or varied random seeds in slice 1. Here we show the fractional difference of the emulator prediction for the halo-matter cross-correlation relative to the direct measurement from each of 20 cosmological models in slice 5, which is a validation sample that is not used in building the emulator considered here.

seeds. To do this, we use the HR runs and compare the results for the halo–matter cross-correlation function. For simplicity, we do not repeat the hyperparameter optimization for this

purpose and reuse the one optimized for our whole sample. In Figure 39, we show a cross-validation test for the GP models with the two sets (the fixed and varied seeds; upper and middle panels, respectively) at the other 20 models in slice 5. The ratio of the GP models to the simulations in slice 5 is generally very close to unity. The two panels look quite similar. We also show in the lower panel the ratio of the two GP models. The difference is mostly below the 1% level, except at very large scales ($\sim 70 h^{-1}\text{Mpc}$). Note that we use a different prescription to calibrate the clustering statistics on large scales (see Section 4.2.3).

From this exercise, we conclude that the difference in the choice of the random number seeds does not largely affect the accuracy of our emulator. For simplicity and to be more conservative, we employ the method using different random number seeds for each of the cosmological models for our main results (development of the emulator).

Appendix D Effect of Massive Neutrinos

Massive neutrinos can impact the growth of cosmological fluctuations; thus, the large-scale structure observables may provide us with a unique opportunity to constrain the sum of the three mass eigenstates (Bond et al. 1980). While a proper treatment of massive neutrinos, including their impact on nonlinear structure formation, would be important for such cosmological tests (e.g., Saito et al. 2008), we here restrict ourselves to a cosmological model with neutrinos of small mass scales as implied from oscillation experiments, $\sum m_\nu = 0.06 \text{ eV}$, and treat them only at the level of the linear transfer function. More precisely, we compute the transfer function of the total matter fluctuations, including massive neutrinos at $z = 0$ using CAMB (Lewis et al. 2000), and multiply it with the linear growth factor to scale back to the initial redshift. When we compute the linear growth factor, we ignore the scale-dependent growth due to the massive neutrinos and assume the $w\text{CDM}$ model with the neutrino density included in the matter content throughout this paper. After generating the initial particle distribution based on this scaled transfer function, we consistently ignore massive neutrinos (and radiation/massless neutrinos) and eventually solve the time evolution of the particle distribution down to $z = 0$ in an N -body simulation.

We give a validation of our treatment using the linear theory. Figure 40 shows the ratio of the matter power spectrum calculated by two methods. The numerator is the one computed by the linear Boltzmann solver CAMB at the redshift indicated by the figure legend. On the other hand, the denominator is the one computed similarly by CAMB but at $z = 0$ and then scaled to the redshift of interest by multiplying the square of the linear growth factor computed without massive neutrinos. This latter one is effectively the underlying linear power spectrum for our simulations. The ratio is by definition unity at $z = 0$ and grows with increasing the redshift, reaching an $\sim 3\%$ deviation at $z \sim 30$. Our target redshifts are rather low, $z \lesssim 1.5$, and the deviation stays well below the 1% level. While nonlinearity can, in principle, bring the sizable difference at earlier epochs to later epochs through mode coupling, it would be a higher-order effect, as the difference is at most at a few percent level from the beginning.

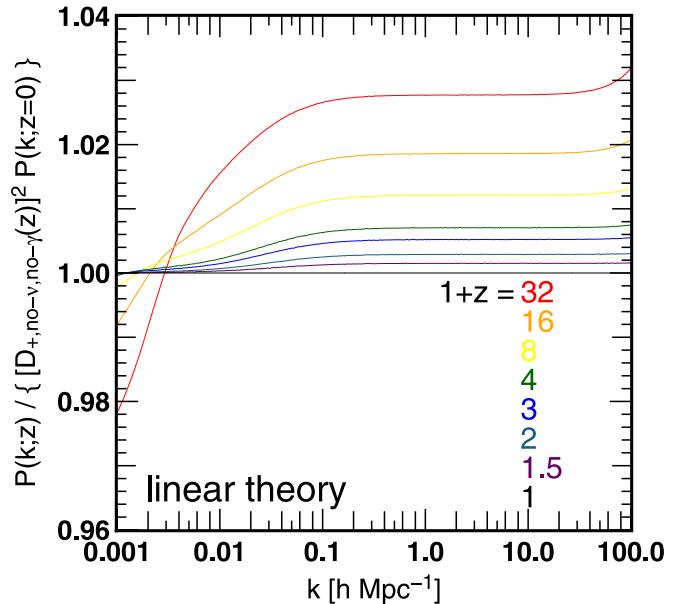


Figure 40. Fractional ratio of the linear matter power spectra, where the effect of massive neutrinos with 0.06 eV is included by an approximated method (see text for details), relative to the spectra directly computed with CAMB at different redshifts.

Appendix E Dependence on the Halo Finder

The main target of this paper is to present the statistics of central halos after removing substructures. However, the definition of central halos is rather ambiguous. We here examine two things: first, the dependence on the criterion to separate substructures, and second, the algorithm to identify a list of possible central-halo candidates. As discussed in the main text, we remove a halo if it is within the radius R_{200m} of a bigger halo in our default setting. In addition, we discard a halo as a fake central halo if the exact spherical mass within R_{200m} is larger than that determined by ROCKSTAR by 30%. This fraction is a parameter that can alter the properties of the central halos remaining after these screening procedures. We also examine the SUBFIND algorithm (Springel et al. 2001) in addition to ROCKSTAR employed in the main text.

We first examine in Figure 41 the HMF after removal of substructures. In the upper panel, we show the HMF obtained by ROCKSTAR, while the lower panel shows that by SUBFIND at $z = 0$. In both cases, we use the test simulation with 512^3 particles in $(250 h^{-1}\text{Mpc})^3$ and divide the results with the reference result based on ROCKSTAR with the maximum allowed mass increase of 30%. The upper panel shows that the parameter can change the mass function more severely near the low-mass end. The change can reach the $\sim 5\%$ level in the worst case with a fraction of 10%. When this fraction is larger, the change is only moderate, $\sim 3\%$ maximum at the low-mass end. The same exercise is presented for SUBFIND in the lower panel. Here the mass increase is based on the change of the mass from the bound mass determined by SUBFIND. Note that the reference mass function in the denominator is still the one with the ROCKSTAR finder. Interestingly, we can match the SUBFIND mass function with ROCKSTAR by adjusting the parameter. A maximum mass increase of $\sim 70\%-80\%$ with

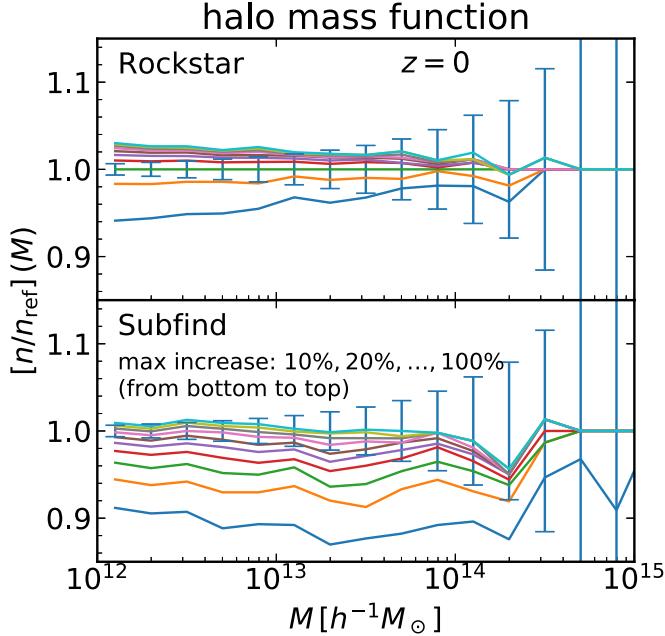


Figure 41. Dependence of the HMF on the maximum allowed mass increase by including particles not dynamically associated with the halo of interest and on the halo finder (upper: ROCKSTAR; lower: SUBFIND). We normalize the mass function by that for the reference catalog based on the ROCKSTAR finder with the fraction parameter 0.3. We also show (error bars) the Poisson noise level for the reference halo catalog.

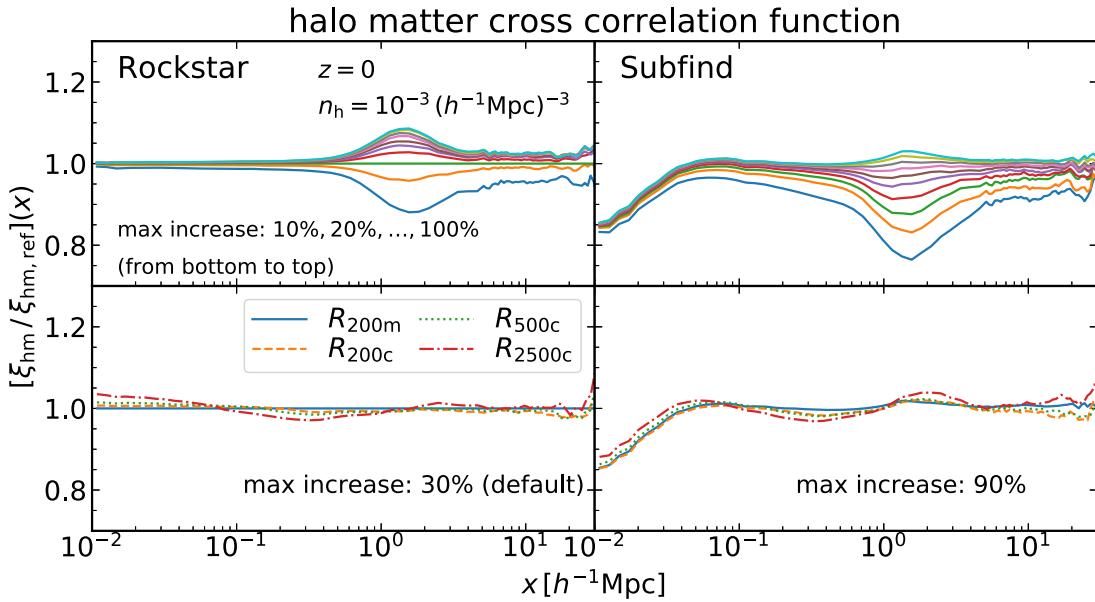


Figure 42. Dependence of the halo–matter cross-correlation function on the maximum allowed mass increase by including particles not dynamically associated with the halo of interest and on the halo finder (left: ROCKSTAR; right: SUBFIND). We normalize the mass function by that for the reference catalog based on the ROCKSTAR finder with the fraction parameter 0.3.

SUBFIND gives an almost identical result to that of ROCKSTAR with the default parameter.

We perform similar tests for the correlation functions in Figures 42 and 43 for the halo–matter cross- and halo–halo autocorrelation functions in the upper panels for the two finders. The most significant effect on the cross-correlation function appears at $\sim 1 h^{-1}\text{Mpc}$, near the halo boundary. This is natural because our parameter controls the exclusion of substructures near the outskirts of a halo. Another notable thing is the innermost part (i.e., $x \lesssim 0.05 h^{-1}\text{Mpc}$) for the

SUBFIND finder. This is because of a different algorithm employed to define the halo center (the center of mass of particles in the core region versus the most bound particle). However, we do not pay much attention here because the scale is close to the softening length ($0.024 h^{-1}\text{Mpc}$) and the main target of the emulator is on a somewhat larger scale ($\gtrsim 0.1 h^{-1}\text{Mpc}$). In the case of the autocorrelation function, the parameter can alter the overall bias factor. This acts, in a sense, as an assembly bias effect, as the halo population, especially the recent merger history, is altered by changing the

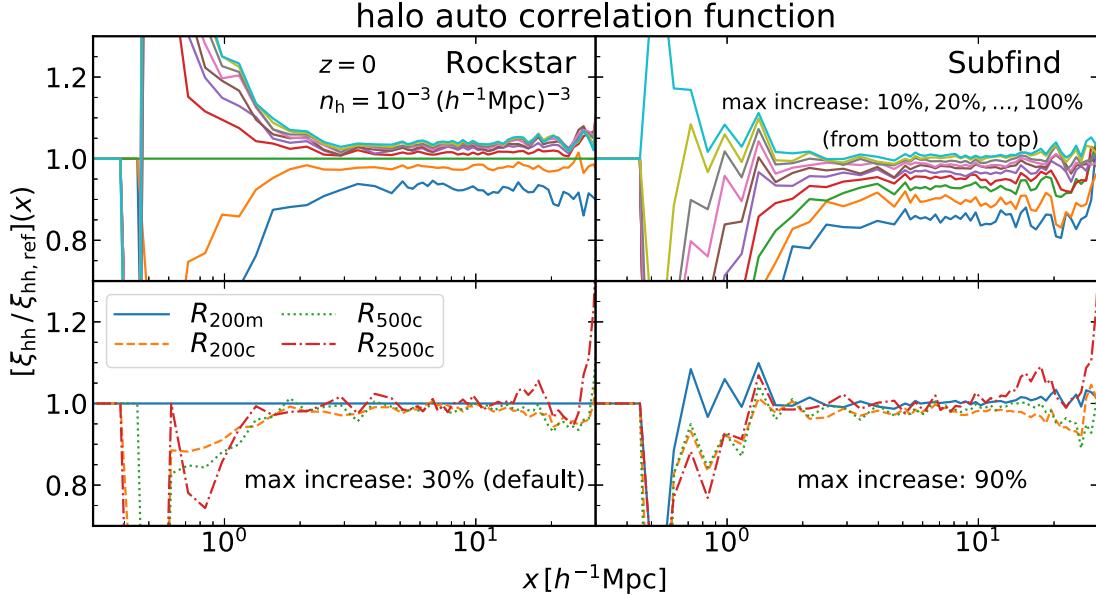


Figure 43. Same as Figure 42 but for the halo autocorrelation function.

criterion to regard a structure at the outskirts of a halo as a substructure or not.

An important message here is again that the correlation functions from the SUBHALO groups can be matched to those from ROCKSTAR by adjusting the parameter that separates a subgroup from the central halo. In other words, the dependence of the result to the halo-finding algorithm is subdominant and can be absorbed by a parameter that defines the central halos. While one has to bear in mind the possible dependence of the correlation functions on the precise definition of the central halos, we speculate that such a dependence would also be absorbed by HOD parameters when galaxy clustering is considered. A structure discarded as a substructure in one algorithm but treated as a central halo in another might be accounted for by populating a satellite galaxy there. We postpone further explicit tests of this point to a future investigation.

Finally, we examine the dependence of the cross- and autocorrelation functions on the outer boundary of the central halos that defines substructures. We consider R_{200c} , R_{500c} , and R_{2500c} , in addition to the default choice of R_{200m} . Here the numbers in the subscript before “c” indicate that the interior density is that number times the critical density. We show in the lower panels of Figures 42 and 43 the results normalized by the default setting. One can see trends similar to the one when we vary the parameter that controls the maximum allowed mass increase, but with a smaller variation. The typical change is

within the target accuracy of this study (i.e., a few percent), except the case with a rather extreme choice of R_{2500c} .

From the analyses presented in this appendix, we conclude that the clustering properties of halos predicted by HALO MODULES are robust against halo-finding algorithms, but the parameter that determines the central/satellite separation can affect the results.

Appendix F Dependence of the Performance on Redshift and Halo Number Density

We have focused on how the emulator performs against simulations at $z = 0.55$ and for halos with a number density $10^{-4}(h^{-1}\text{Mpc})^{-3}$ in the main text. We summarize our findings at different redshifts and for different halo samples in this appendix.

We first show a cross-validation study of the HMF in Figure 44. Each panel corresponds to the lower right panel of Figure 13, where the accuracy of the emulator is tested for the 20 cosmologies in slice 5, which are not used in the GPR. We can confirm that the scatter among the thin solid lines (i.e., different cosmologies) scales similarly to the width of the red shaded regions (the scatter of the HMF for the different realizations of the fiducial *Planck* cosmology). Thus, we conclude that the modeling is reasonably accurate given the uncertainties in the simulation data.

Next, we show in Figure 45 a similar cross-validation study for the halo–matter cross-correlation function at various number

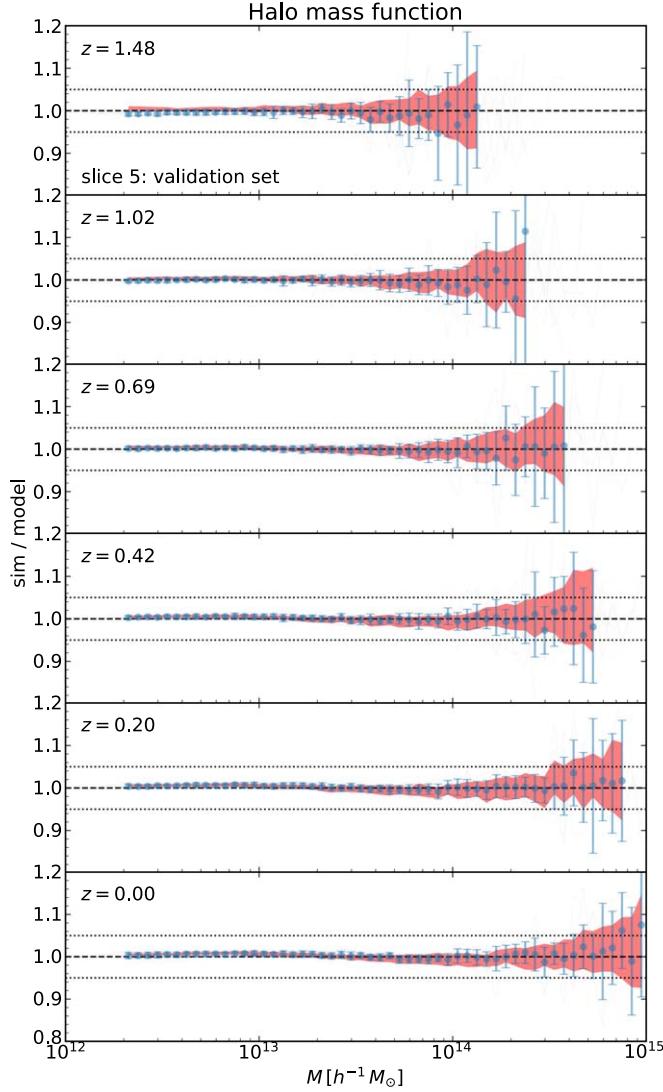


Figure 44. Redshift dependence of the cross-validation test for the HMF.

densities (rows) and different redshifts (columns). The overall trend is that the accuracy is degraded as decreasing the number density reflecting the bigger uncertainties in the simulation data due to a larger noise. On the other hand, no clear dependence on redshift is found.

The halo-halo correlation function is tested in Figure 46. This time, the upper three rows are for the autocorrelation function of

the same halo samples, and the remaining three rows are for two halo samples with different number densities, as indicated in the figure legend. A similar trend, a bigger scatter for low-density samples, can be found. Finally, the propagators are shown in Figure 47. These plots are a practical guide to the accuracy of the current code depending on the halo number densities (or halo masses) in actual use.

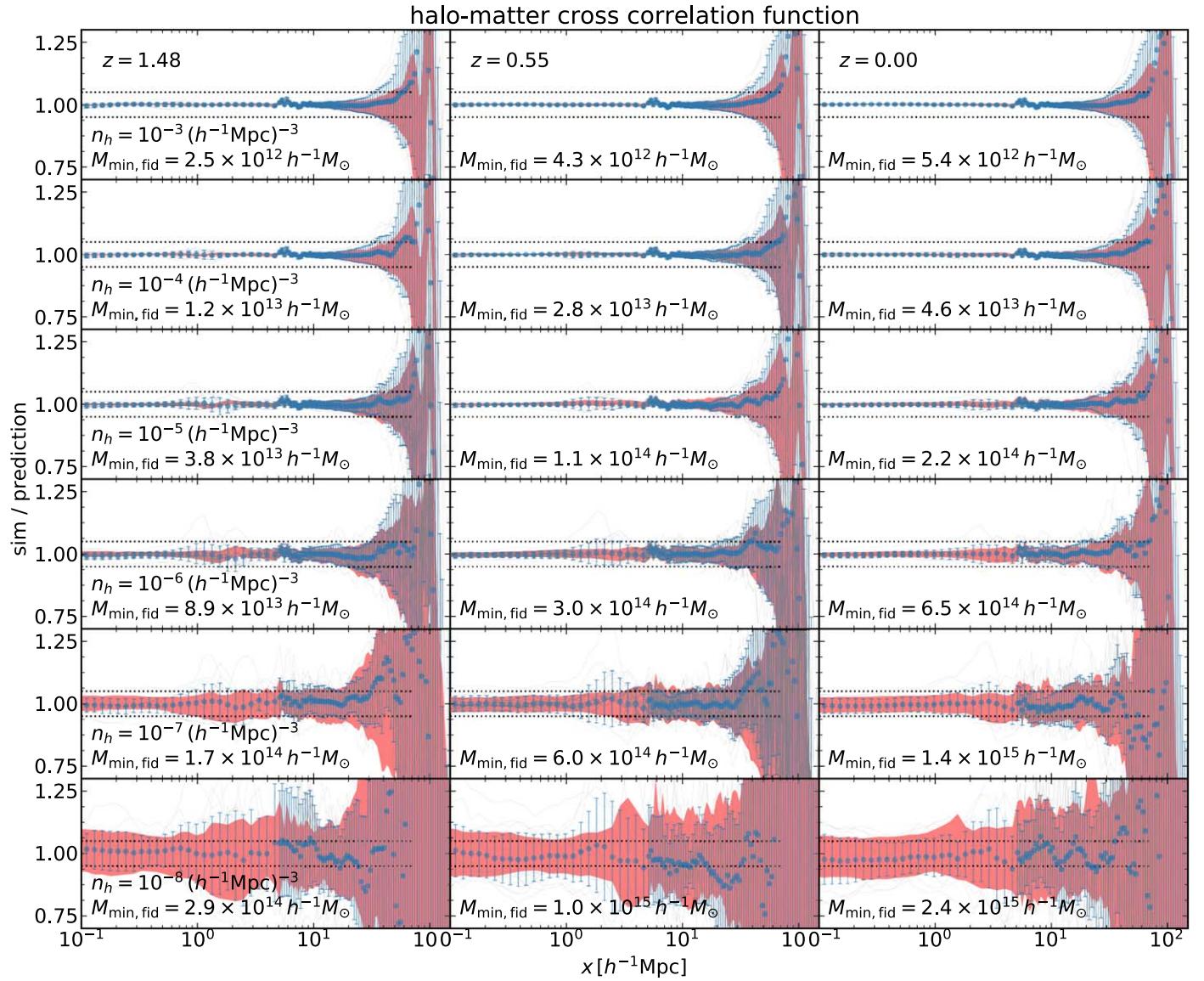


Figure 45. Redshift and number density dependence of the cross-validation test for the halo–matter cross-correlation function.

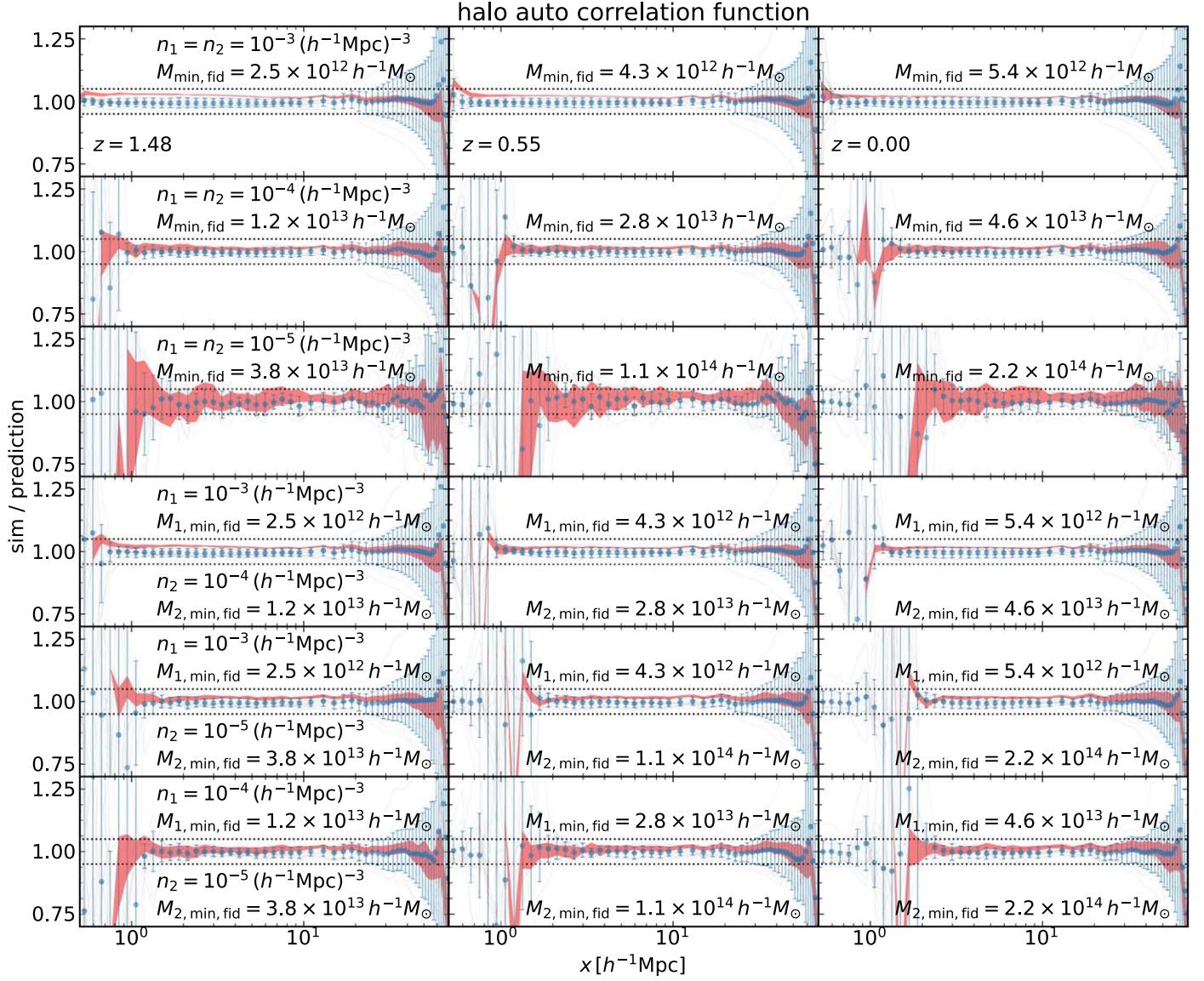


Figure 46. Redshift and number density dependence of the cross-validation test for the halo autocorrelation function.

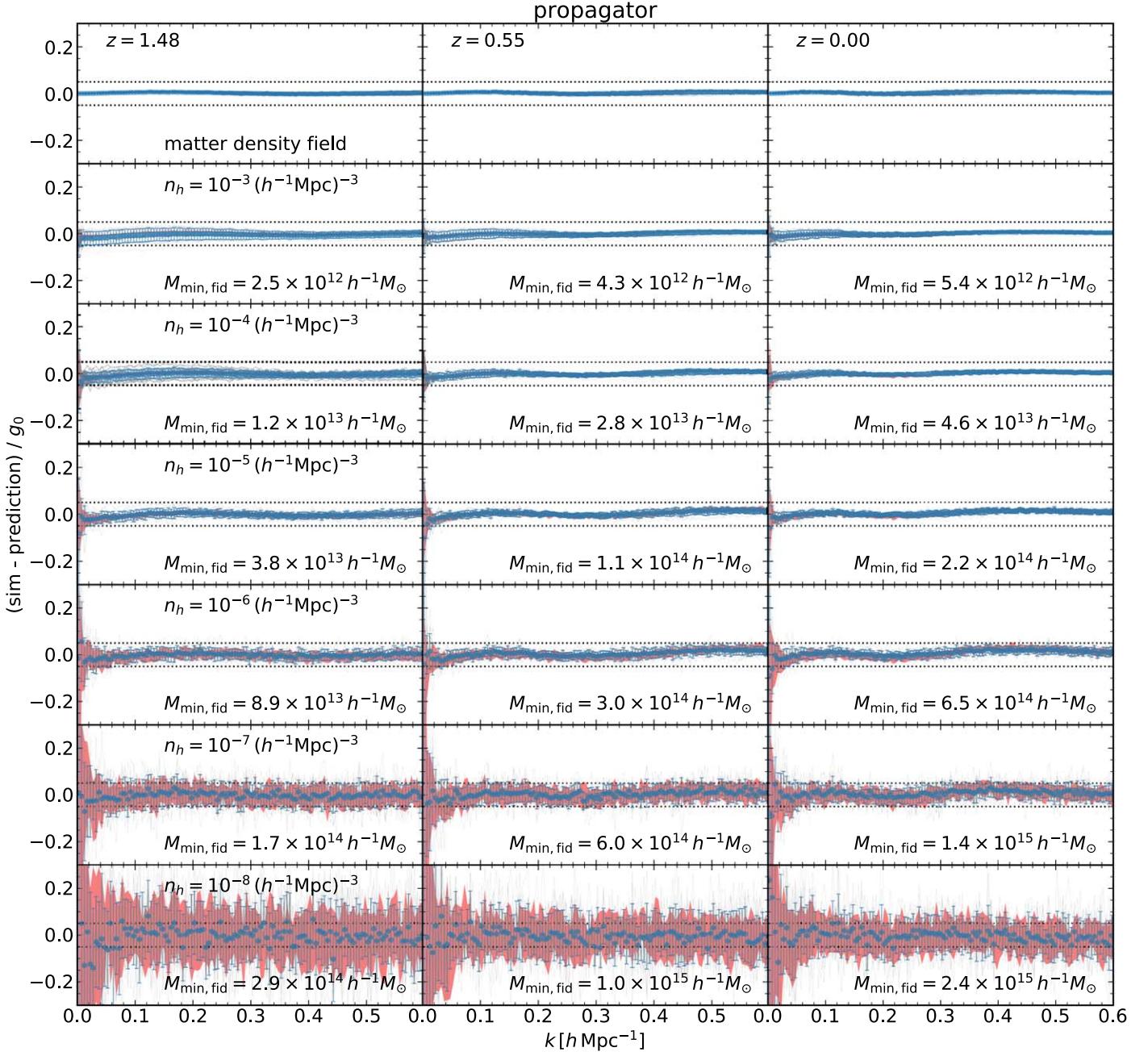


Figure 47. Halo mass and redshift dependence of the performance of the propagator module. We also show the matter propagator in the top row.

Appendix G HOD Model

As we stressed in the main text, one can insert one's own module into DARK EMULATOR to model how halos are related to galaxies under consideration. Here, as a working example, we show an HOD model (Jing et al. 1998; Peacock & Smith 2000; Seljak 2000; Scoccimarro et al. 2001; Zheng et al. 2005) to predict the abundance, clustering, and lensing signal of galaxies. In particular, we employ the model in More et al. (2015b). A module based on this HOD prescription is provided in DARK EMULATOR as a default package.

We adopt the HOD model with an explicit split of the halo occupation into central and satellite galaxies,

$$\langle N \rangle_M = \langle N_c \rangle_M + \langle N_s \rangle_M, \quad (47)$$

where the average $\langle \dots \rangle_M$ is taken for halos with mass M . The mean HOD for the central galaxies is given by

$$\langle N_c \rangle_M = f_{\text{inc}}(M) \frac{1}{2} \left[1 + \text{erf} \left(\frac{\log M - \log M_{\min}}{\sigma_{\log M}} \right) \right], \quad (48)$$

where $\text{erf}(x)$ is the error function and M_{\min} and $\sigma_{\log M}$ are model parameters. The function $f_{\text{inc}}(M)$ accounts for a potential incompleteness of the central galaxies that models a possibility that a central galaxy in some halos can be missed due to an imperfect selection effect or galaxies under consideration do not necessarily occupy halos at the center even for sufficiently massive halos (e.g., Masaki et al. 2013). We assume a log-linear functional form given by

$$f_{\text{inc}}(M) = \max [0, \min [1, 1 + \alpha_{\text{inc}}(\log M - \log M_{\text{inc}})]], \quad (49)$$

where α_{inc} and M_{inc} are model parameters. The mean HOD for the satellite galaxies is given by

$$\langle N_s \rangle_M \equiv \langle N_c \rangle_M \lambda_s(M) = \langle N_c \rangle_M \left[\frac{M - \kappa M_{\min}}{M_1} \right]^\alpha, \quad (50)$$

where κ , M_1 , and α are model parameters, and we have defined the notation $\lambda_s(M) \equiv [(M - \kappa M_{\min})/M_1]^\alpha$ for convenience in the following discussion. We assume that the distribution of central galaxies, N_c , follows the Bernoulli distribution (i.e., can take only zero or 1) with mean $\langle N_c \rangle_M$. On the other hand, we populate satellite galaxies to a halo only when a central galaxy exists. The conditional distribution of N_s in a halo with mass M that has a central galaxy is given by the Poisson distribution with mean $\lambda_s(M)$. Our HOD model is fully specified by seven parameters: $\{M_{\min}, \sigma_{\log M}, \alpha_{\text{inc}}, M_{\text{inc}}, \kappa, M_1, \alpha\}$.

Once the HOD model is specified, the mean number density of galaxies in a sample is computed as

$$\bar{n}_g = \int dM \frac{dn}{dM} [\langle N_c \rangle_M + \langle N_s \rangle_M]. \quad (51)$$

Here the HMF dn/dM is given by DARK EMULATOR for a given cosmological model.

Now we consider the galaxy-galaxy weak lensing that measures the excess surface mass density profile around lensing galaxies. The galaxy-galaxy weak-lensing profile for lens galaxies at redshift z_l can be expressed in terms of the

galaxy-matter power spectrum (e.g., Murata et al. 2018) as

$$\Delta\Sigma(R; z_l) = \bar{\rho}_m \int_0^\infty \frac{k dk}{2\pi} P_{gm}(k; z_l) J_2(kR), \quad (52)$$

where $\bar{\rho}_m$ is the present-day mean matter density and $J_2(x)$ is the second-order Bessel function. Under the HOD model described above, we further make some assumptions on the location of central and satellite galaxies within the host halo. First, we allow some fraction (f_{off}) of the central galaxies to be located off of the true halo center following a Gaussian distribution with width R_{off} . In this case, this off-centering effect can be expressed in terms of a kernel in Fourier space (Oguri & Takada 2011; Hikage et al. 2013):

$$\mathcal{H}_{\text{off}}(k; M, f_{\text{off}}, R_{\text{off}}) \equiv 1 - f_{\text{off}} + f_{\text{off}} \exp \left[-\frac{1}{2} (kR_{\text{off}})^2 \right]. \quad (53)$$

We then introduce a function $\tilde{u}_s(k; M)$ for the normalized radial profile of satellite galaxies, again in Fourier space. For this, we assume an NFW profile, which is specified by a given model of the halo matter-concentration relation, denoted as $c(M, z)$, for halos of a given mass. We adopt the fitting formula in Diemer & Kravtsov (2015) to compute $c(M, z)$ for a given cosmological model, and $c(M, z)$ is not a free parameter in our default setting. Alternatively, we provide an option to distribute satellite galaxies following the mass distribution given by $\propto 1 + \xi_{\text{hm}}(x)$, which can be computed by one of our HALO MODULES. With these assumptions, the galaxy-matter cross-power spectrum is given as

$$P_{gm}(k) = \frac{1}{\bar{n}_g} \int dM \frac{dn}{dM} [\langle N_c \rangle_M \mathcal{H}_{\text{off}}(k; M, f_{\text{off}}, R_{\text{off}}) + \langle N_s \rangle_M \tilde{u}_s(k; M)] P_{hm}(k; M). \quad (54)$$

In Equations (52) and (54), dn/dM and $P_{hm}(k; M)$ are given by DARK EMULATOR, and the radial profiles of off-centering central galaxies and satellite galaxies are modeled by the parameters $\{c(M, z), f_{\text{off}}, R_{\text{off}}\}$. Thus, one can compute the galaxy-galaxy weak-lensing profile from DARK EMULATOR for a given cosmological model, e.g., once nine parameters for connecting halos to galaxies are specified: seven parameters for HOD plus two profile parameters.

Next we consider the projected correlation function of galaxies, $w_{gg}(R)$, which is defined in terms of the three-dimensional correlation function as

$$w_{gg}(R; z) = 2 \int_0^{\pi_{\max}} d\pi \xi_{gg}(\sqrt{R^2 + \pi^2}; z). \quad (55)$$

Here

$$\xi_{gg}(x) \equiv \int_0^\infty \frac{k^2 dk}{2\pi^2} P_{gg}(k) j_0(kx), \quad (56)$$

where $j_0(x)$ is the zeroth-order spherical Bessel function and $P_{gg}(k)$ is the galaxy autopower spectrum. Using the above model ingredients, we can express the galaxy autopower

spectrum as

$$\begin{aligned}
 P_{gg}(k) = & \frac{1}{\bar{n}_g^2} \int dM \frac{dn}{dM} [2 \langle N_s \rangle_M \mathcal{H}_{\text{off}}(k; M) \tilde{u}_s(k; M) \\
 & + \langle N_c \rangle_M \lambda_s(M)^2 \tilde{u}_s(k; M)^2] \\
 & + \frac{1}{\bar{n}_g^2} \left[\int dM \frac{dn}{dM} \{ \langle N_c \rangle_M \mathcal{H}_{\text{off}}(k; M) \right. \\
 & \left. + \langle N_s \rangle_M \tilde{u}_s(k; M) \} \right] \\
 & \times \left[\int dM' \frac{dn}{dM'} \{ \langle N_c \rangle_{M'} \mathcal{H}_{\text{off}}(k; M') \right. \\
 & \left. + \langle N_s \rangle_{M'} \tilde{u}_s(k; M') \} \right] P_{hh}(k; M, M'). \quad (57)
 \end{aligned}$$

Note that in deriving the equation above, we have used the relations

$$\begin{aligned}
 \langle N_c N_s \rangle_M &= \langle N_s \rangle_M, \\
 \langle N_s (N_s - 1) \rangle_M &= \langle N_c \rangle_M \lambda_s(M)^2, \quad (58)
 \end{aligned}$$

which follow from our assumptions on the distribution of N_c and N_s described above. Once again, the projected correlation function of galaxies can be computed for the same set of model parameters as those of the galaxy–galaxy weak-lensing profile (i.e., nine parameters for the halo–galaxy connection).

ORCID iDs

Takahiro Nishimichi <https://orcid.org/0000-0002-9664-0760>

Masahiro Takada <https://orcid.org/0000-0002-5578-6472>

Ryuichi Takahashi <https://orcid.org/0000-0001-6021-0147>

Ken Osato <https://orcid.org/0000-0002-7934-2569>

Masamune Oguri <https://orcid.org/0000-0003-3484-399X>

Ryoma Murata <https://orcid.org/0000-0002-8232-189X>

Naoki Yoshida <https://orcid.org/0000-0001-7925-238X>

References

- Abbott, T. M. C., Abdalla, F. B., Alarcon, A., et al. 2018, *PhRvD*, **98**, 043526
- Adhikari, S., Dalal, N., & Chamberlain, R. T. 2014, *JCAP*, **11**, 019
- Agarwal, S., Abdalla, F. B., Feldman, H. A., Lahav, O., & Thomas, S. A. 2012, *MNRAS*, **424**, 1409
- Agarwal, S., Abdalla, F. B., Feldman, H. A., Lahav, O., & Thomas, S. A. 2014, *MNRAS*, **439**, 2102
- Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, *PASJ*, **70**, S4
- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, **219**, 12
- Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., & O’Neil, M. 2015, *ITPAM*, **38**, 252
- Angulo, R. E., White, S. D. M., Springel, V., & Henriques, B. 2014, *MNRAS*, **442**, 2131
- Ba, S., Brenneman, W. A., & Myers, W. R. 2015, *Technometrics*, **57**, 479
- Bailey, S. 2012, *PASP*, **124**, 1015
- Baldauf, T., Seljak, U., Desjacques, V., & McDonald, P. 2012, *PhRvD*, **86**, 083540
- Baldauf, T., Seljak, U., Smith, R. E., Hamaus, N., & Desjacques, V. 2013, *PhRvD*, **88**, 083507
- Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, *ApJ*, **304**, 15
- Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013, *ApJ*, **762**, 109
- Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, *PhR*, **367**, 1
- Bernardeau, F., Crocce, M., & Scoccimarro, R. 2008, *PhRvD*, **78**, 103521
- Blas, D., Lesgourgues, J., & Tram, T. 2011, *JCAP*, **7**, 034
- Bond, J. R., Efstathiou, G., & Silk, J. 1980, *PhRvL*, **45**, 1980
- Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, *ApJ*, **466**, 623
- Cooray, A., & Sheth, R. 2002, *PhR*, **372**, 1
- Crocce, M., Castander, F. J., Gaztañaga, E., Fosalba, P., & Carretero, J. 2015, *MNRAS*, **453**, 1513
- Crocce, M., Fosalba, P., Castander, F. J., & Gaztañaga, E. 2010, *MNRAS*, **403**, 1353
- Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, *MNRAS*, **373**, 369
- Crocce, M., & Scoccimarro, R. 2006a, *PhRvD*, **73**, 063520
- Crocce, M., & Scoccimarro, R. 2006b, *PhRvD*, **73**, 063519
- Crocce, M., & Scoccimarro, R. 2008, *PhRvD*, **77**, 023533
- Dalal, N., Doré, O., Huterer, D., & Shirokov, A. 2008, *PhRvD*, **77**, 123514
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *ApJ*, **292**, 371
- dell’Antonio, I. P., & Tyson, J. A. 1996, *ApJL*, **473**, L17
- DeRose, J., Wechsler, R. H., Tinker, J. L., et al. 2019, *ApJ*, **875**, 69
- Desjacques, V., Crocce, M., Scoccimarro, R., & Sheth, R. K. 2010, *PhRvD*, **82**, 103529
- Desjacques, V., Jeong, D., & Schmidt, F. 2018, *PhR*, **733**, 1
- Diemer, B., & Kravtsov, A. V. 2014, *ApJ*, **789**, 1
- Diemer, B., & Kravtsov, A. V. 2015, *ApJ*, **799**, 108
- Euclid Collaboration, Knabenhans, M., Stadel, J., et al. 2019, *MNRAS*, **484**, 5509
- Fendt, W. A., & Wandelt, B. D. 2007a, arXiv:0712.0194
- Fendt, W. A., & Wandelt, B. D. 2007b, *ApJ*, **654**, 2
- Fischer, P., McKay, T. A., Sheldon, E., et al. 2000, *AJ*, **120**, 1198
- Garrison, L. H., Eisenstein, D. J., Ferrer, D., et al. 2018, *ApJS*, **236**, 43
- Garrison, L. H., Eisenstein, D. J., Ferrer, D., Metchnik, M. V., & Pinto, P. A. 2016, *MNRAS*, **461**, 4125
- Garud, Sushant S., Karimi, I. A., & Kraft, M. 2017, *Computers & Chemical Engineering*, **106**, 71
- Habib, S., Heitmann, K., Higdon, D., Nakhleh, C., & Williams, B. 2007, *PhRvD*, **76**, 083503
- Hamaus, N., Seljak, U., Desjacques, V., Smith, R. E., & Baldauf, T. 2010, *PhRvD*, **82**, 043515
- Hamilton, A. J. S. 2000, *MNRAS*, **312**, 257
- Hand, N., Seljak, U., Beutler, F., & Vlah, Z. 2017, *JCAP*, **10**, 009
- Heitmann, K., Bingham, D., Lawrence, E., et al. 2016, *ApJ*, **820**, 108
- Heitmann, K., Higdon, D., Nakhleh, C., & Habib, S. 2006, *ApJL*, **646**, L1
- Heitmann, K., Higdon, D., White, M., et al. 2009, *ApJ*, **705**, 156
- Heitmann, K., White, M., Wagner, C., Habib, S., & Higdon, D. 2010, *ApJ*, **715**, 104
- Hikage, C., Mandelbaum, R., Takada, M., & Spergel, D. N. 2013, *MNRAS*, **435**, 2345
- Hikage, C., Oguri, M., Hamana, T., et al. 2019, *PASJ*, **71**, 43
- Hockney, R. W., & Eastwood, J. W. 1981, *Computer Simulation Using Particles* (New York: McGraw-Hill)
- Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2001, *ApJL*, **558**, L11
- Jing, Y. P., Mo, H. J., & Börner, G. 1998, *ApJ*, **494**, 1
- Joudaki, S., Blake, C., Johnson, A., et al. 2018, *MNRAS*, **474**, 4894
- Joyce, M., & Marcos, B. 2007, *PhRvD*, **76**, 103505
- Kaiser, N. 1984, *ApJL*, **284**, L9
- Knebe, A., Knollmann, S. R., Muldrew, S. I., et al. 2011, *MNRAS*, **415**, 2293
- Kwan, J., Bhattacharya, S., Heitmann, K., & Habib, S. 2013, *ApJ*, **768**, 123
- Kwan, J., Heitmann, K., Habib, S., et al. 2015, *ApJ*, **810**, 35
- Lawrence, E., Heitmann, K., Kwan, J., et al. 2017, *ApJ*, **847**, 50
- Lawrence, E., Heitmann, K., White, M., et al. 2010, *ApJ*, **713**, 1322
- Lesgourgues, J. 2011, arXiv:1104.2932
- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, **538**, 473
- Li, Y., Hu, W., & Takada, M. 2016, *PhRvD*, **93**, 063507
- Liu, J., & Madhavacheril, M. S. 2019, *PhRvD*, **99**, 083508
- Ma, C.-P., & Fry, J. N. 2000, *ApJ*, **543**, 503
- Mandelbaum, R., Miyatake, H., Hamana, T., et al. 2018, *PASJ*, **70**, S25
- Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006, *MNRAS*, **368**, 715
- Mandelbaum, R., Slosar, A., Baldauf, T., et al. 2013, *MNRAS*, **432**, 1544
- Marcos, B., Baertschiger, T., Joyce, M., Gabrielli, A., & Sylos Labini, F. 2006, *PhRvD*, **73**, 103507
- Masaki, S., Hikage, C., Takada, M., Spergel, D. N., & Sugiyama, N. 2013, *MNRAS*, **433**, 3506
- Matsubara, T. 2008, *PhRvD*, **77**, 063530
- McClintock, T., Rozo, E., Becker, M. R., et al. 2019, *ApJ*, **872**, 53
- McDonald, P. 2006, *PhRvD*, **74**, 103512
- McDonald, P., & Roy, A. 2009, *JCAP*, **8**, 020
- Mitsuhashi, R., Kawashima, H., Nishimichi, T., & Tatebe, O. 2016, in 2016 IEEE International Conf. on Big Data, BigData 2016 (Piscataway, NJ: IEEE), 2938, <https://ieeexplore.ieee.org/document/7840944>
- Miyoshi, K., & Kihara, T. 1975, *PASJ*, **27**, 333
- Mo, H. J., & White, S. D. M. 1996, *MNRAS*, **282**, 347
- More, S., Diemer, B., & Kravtsov, A. V. 2015a, *ApJ*, **810**, 36
- More, S., Miyatake, H., Mandelbaum, R., et al. 2015b, *ApJ*, **806**, 2
- More, S., Miyatake, H., Takada, M., et al. 2016, *ApJ*, **825**, 39

- Murata, R., Nishimichi, T., Takada, M., et al. 2018, *ApJ*, **854**, 120
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, *ApJ*, **462**, 563
- Nishimichi, T., Shirata, A., Taruya, A., et al. 2009, *PASJ*, **61**, 321
- Nishizawa, A. J., Takada, M., & Nishimichi, T. 2013, *MNRAS*, **433**, 209
- Oguri, M., & Takada, M. 2011, *PhRvD*, **83**, 023008
- Okumura, T., Nishimichi, T., Umetsu, K., & Osato, K. 2018a, *PhRvD*, **98**, 023523
- Okumura, T., Nishimichi, T., Umetsu, K., & Osato, K. 2018b, *PhRvD*, **98**, 023523
- Osato, K., Nishimichi, T., Oguri, M., Takada, M., & Okumura, T. 2018, *MNRAS*, **477**, 2141
- Peacock, J. A., & Smith, R. E. 2000, *MNRAS*, **318**, 1144
- Peebles, P. J. E. 1980, The Large-scale Structure of the Universe (Princeton, NJ: Princeton Univ. Press)
- Petri, A., Liu, J., Haiman, Z., et al. 2015, *PhRvD*, **91**, 103511
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, **594**, A13
- Press, W. H., & Schechter, P. 1974, *ApJ*, **187**, 425
- Pujol, A., Gaztañaga, E., Giocoli, C., et al. 2014, *MNRAS*, **438**, 3205
- Reid, B. A., Seo, H.-J., Leauthaud, A., Tinker, J. L., & White, M. 2014, *MNRAS*, **444**, 476
- Saito, S., Takada, M., & Taruya, A. 2008, *PhRvL*, **100**, 191301
- Sato, M., & Matsubara, T. 2011, *PhRvD*, **84**, 043501
- Schaan, E., Krause, E., Eifler, T., et al. 2017, *PhRvD*, **95**, 123512
- Schneider, M. D., Knox, L., Habib, S., et al. 2008, *PhRvD*, **78**, 063529
- Scoccimarro, R. 1998, *MNRAS*, **299**, 1097
- Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, *ApJ*, **546**, 20
- Seljak, U. 2000, *MNRAS*, **318**, 203
- Seljak, U., Hamaus, N., & Desjacques, V. 2009, *PhRvL*, **103**, 091303
- Seljak, U., Makarov, A., Mandelbaum, R., et al. 2005, *PhRvD*, **71**, 043511
- Seljak, U., & Zaldarriaga, M. 1996, *ApJ*, **469**, 437
- Sheldon, E. S., Johnston, D. E., Frieman, J. A., et al. 2004, *AJ*, **127**, 2544
- Sheldon, E. S., Johnston, D. E., Scranton, R., et al. 2009, *ApJ*, **703**, 2217
- Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *MNRAS*, **323**, 1
- Sheth, R. K., & Tormen, G. 1999, *MNRAS*, **308**, 119
- Sheth, R. K., & Tormen, G. 2002, *MNRAS*, **329**, 61
- Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, *MNRAS*, **341**, 1311
- Springel, V. 2005, *MNRAS*, **364**, 1105
- Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, *MNRAS*, **328**, 726
- Takada, M., Ellis, R. S., Chiba, M., et al. 2014, *PASJ*, **66**, R1
- Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, *ApJ*, **761**, 152
- Taruya, A., Nishimichi, T., & Jeong, D. 2018, *PhRvD*, **98**, 103532
- Taruya, A., Nishimichi, T., & Saito, S. 2010, *PhRvD*, **82**, 063522
- Tegmark, M., & Peebles, P. J. E. 1998, *ApJL*, **500**, L79
- Tegmark, M., Strauss, M. A., Blanton, M. R., et al. 2004, *PhRvD*, **69**, 103501
- Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, *ApJ*, **688**, 709
- Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, *ApJ*, **724**, 878
- Valageas, P., & Nishimichi, T. 2011, *A&A*, **527**, A87
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *Natur*, **509**, 177
- Warren, M. S., Abazajian, K., Holz, D. E., & Teodoro, L. 2006, *ApJ*, **646**, 881
- Wibking, B. D., Salcedo, A. N., Weinberg, D. H., et al. 2019, *MNRAS*, **484**, 989
- Yoo, J., & Seljak, U. 2012, *PhRvD*, **86**, 083504
- Zel'dovich, Y. B. 1970, *A&A*, **5**, 84
- Zhai, Z., Tinker, J. L., Becker, M. R., et al. 2019, *ApJ*, **874**, 95
- Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, *ApJ*, **633**, 791