Prepared for submission to JCAP

Generating Mock Catalogs for the Baryon Oscillation Spectroscopic Survey: An Approximate N-Body approach

Tomomi Sunayama[a], Nikhil Padmanabhan[a], Katrin Heitmann[b], Salman Habib[b], Steve Rangel[b,c]

[a]Yale University, New Haven, CT
[b] Argonne National Laboratory, Lemont, IL
[c] Northwestern University, Evanston, IL

E-mail: tomomi.sunayama@yale.edu, nikhil.padmanabhan@yale.edu,heitmann@anl.gov, habib@anl.gov,steverangel@u.northwestern.edu

Abstract. Precision measurements of the large scale structure of the Universe require large numbers of high fidelity mock catalogs to accurately assess, and account for, the presence of systematic effects. We introduce and test a scheme for generating mock catalogs rapidly using suitably derated N-body simulations. Our aim is to reproduce the large scale structure and the gross properties of dark matter halos with high accuracy, while sacrificing the details of the halo's internal structure. By adjusting global and local time-steps in an N-body code, we demonstrate that we recover halo masses to better than 2% and the power spectrum (both in real and redshift space, for $k = 1h\mathrm{Mpc}^{-1}$) to better than 1%, while requiring a factor of 4 less CPU time. We also calibrate the redshift spacing of outputs required to generate simulated light cones. We find that outputs separated by $\Delta z = 0.05$ allow us to interpolate particle positions and velocities to reproduce the real and redshift space power spectra to better than 1% (out to $k = 1h\mathrm{Mpc}^{-1}$). We apply these ideas to generate a suite of simulations spanning a range of cosmologies, motivated by the Baryon Oscillation Spectroscopic Survey (BOSS) but broadly applicable to future large scale structure surveys including eBOSS and DESI. As an initial demonstration of the utility of such simulations, we calibrate the shift in the BAO position as a function of galaxy bias with higher precision than has been possible before. This paper also serves to document these simulations, which we make publically available.

Keywords: cosmology; large-scale structure of Universe, cosmological parameters, galaxies; halos, statistics

Contents

1 Introduction

Large-volume spectroscopic surveys of the Universe [1–3] are revolutionizing our understanding of cosmology and structure formation. Based on these successes, a new generation of surveys [1, 4, 5] is being planned that will improve our constraints by an order of magnitude (or more). This unprecedented improvement in statistical precision places stringent demands on the theoretical modeling and analysis techniques; simulations will play an essential in meeting these requirements.

One of the challenges for simulations are the varied roles they play, and the different requirements these impose on the simulations. At one extreme, simulations are necessary for estimating the errors on the measurements. This typically requires very large volumes to simulate entire surveys thousands of times, but have lower accuracy requirements. Motivated by these considerations, a number of recent studies have investigated methods designed to produce mock catalogs with reduced accuracy, but much higher throughput compared to the full N-body simulations [6–17].

An open question still is the effect of changing the input cosmology used to generate the covariance matrix on cosmological inferences, and how best to implement such variations (but see [18] for recent work on this). More recently, the impact of super-survey modes (modes outside the survey volume) on inferred errors has been shown to be potentially larger than previously appreciated [19] and is an area of active study.

At the other extreme, simulations are crucial for calibrating the theoretical models used to fit the data. Examples here are quantifying shifts in the baryon acoustic oscillation distance scale due to nonlinear evolution and galaxy bias [20–25], or templates used to fit the full shape of the galaxy correlation function. For such applications, one ideally requires high fidelity simulations. The volume requirements are significantly reduced from that for covariance matrices, but still need to much larger than survey volumes to keep systematic errors below statistical errors.

An intermediate application are the generation of mock catalogs that capture the observational characteristics of surveys (eg. geometry, selection effects). The importance of these cannot be underestimated, since the effects of many observational systematics can only be quantitatively estimated by simulating them. These issues will get progressively more important for the next generations of surveys which will move away from highly complete and pure samples that have mostly been used for cosmological studies to date.

The simplest way to generate approximate density fields is to use analytic approximations such as Lagrangian perturbation theory followed by prescriptions to put in halos in a way that better matches results from N-body simulations [26, 27], or simply to run lower resolution N-body codes with a small number of time-steps [11], or a combination of the two approaches [12]. These methods are successful in capturing the large-scale density field but lose information at small scales. Because of their speed, they can be used to produce large numbers of simulations required to build sample covariance matrices, at error levels ranging from $5-10\%$ (depending on the quantities being predicted). It is difficult to estimate, however, what the loss of accuracy implies for tests of systematic errors, which may need to be modeled at the $< 1\%$ level.

The approach we take here is to reduce the small-scale accuracy of a high-resolution N-body code by coarsening its temporal resolution. For the code we consider here, the time-stepping consists of two components, (i) a long time step for solving for evolution under the long-range particle-mesh (PM) force, and (ii) a set of underlying sub-cycled time steps for a short-range particle-particle interaction, computed either via a tree-based algorithm, or by direct particle-particle force evaluations. The idea is to reduce the number of both types of time steps while preserving enough accuracy to correctly describe the large scale distribution of galaxies, as modeled by a halo occupation distribution (HOD) approach. Our first goal in this paper is, therefore, to quantitatively understand the impact of the temporal resolution on the halo density field and how best to accurately reproduce the details of the halo density field on large scales, sacrificing small scale structure information. This allows to generate a suite of large volume simulations, spanning a range of cosmologies. This paper presents the details of these simulations and outlines future applications.

This paper is organized as follows. Sec. 2 briefly describes the Hardware/Hybrid Accelerated Cosmology Code (HACC) N-body framework we use to generate our simulations, focusing on the flexibility in the time-stepping that we exploit here. Sec. 3 presents a sequence of convergence tests where we evaluate the effects of time-stepping on the halo density field. Sec. 4 discusses interpolating between saved time steps, necessary for constructing light-cone outputs. Sec. 5 presents two example applications of these simulations : generating mock catalogs that match the BOSS galaxy sample, and calibrating shifts in the baryon acoustic oscillation (BAO) scale. We conclude in Sec. 6 by outlining possible future directions.

Unless specified, all simulations and calculations in this paper assume a ΛCDM cosmology with $\Omega_m = 0.2648$, $\Omega_\Lambda = 0.7352$, $\Omega_b h^2 = 0.02258$, $n_s = 0.963$, $\sigma_8 = 0.8$ and $h = 0.71$.

2 HACC

All simulations in this paper were carried out using the HACC (Hardware/Hybrid Accelerated Cosmology Code) framework. HACC provides an advanced, architecture-agile, extreme-scale N-body capability targeted to cosmological simulations. It is descended from an approach originally developed for the heterogeneous architecture of Roadrunner [28, 29], the first computer to break the petaflop performance barrier.

HACC's flexible code architecture combines MPI with a variety of more local programming models, (e.g., OpenCL, OpenMP) and is easily adaptable to different platforms. HACC has demonstrated scaling on the entire IBM BG/Q Sequoia system up to 1,572,864 cores with an equal number of MPI ranks, attaining 13.94 PFlops at 69.2% of peak and 90% parallel efficiency (for details, see [30]). Examples of science results obtained using HACC include 64-billion particle runs for baryon acoustic oscillations predictions for the BOSS Lyman-α forest [31], high-statistics predictions for the halo profiles of massive clusters [32], and 0.5 and 1.1 trillion particle runs at high mass resolution. A recent overview of the HACC framework can be found in [33].

HACC uses a hybrid parallel algorithmic structure, splitting the force calculation into a specially designed grid-based long/medium range spectral PM component that is common to all computer architectures, and an architecture-specific short-range solver. Modular code design combined with particle caching allows the short-range solvers to be 'hot-swappable' on-node; they are blind to the parallel implementation of the long-range solver. The short-range solvers can use direct particle-particle interactions, i.e., a $\mathrm{P^3M}$ algorithm [34], as on (Cell or GPU) accelerated systems, or use tree methods on conventional or many-core architectures. (This was the case for the simulations reported here.) In all cases, the time-stepping scheme is based on a symplectic method with (adaptive) sub-cycling of the short-range force. The availability of multiple algorithms within the HACC framework allows us to carry out careful error analyses, for example, the $\mathrm{P^3M}$ and the TreePM versions agree to within 0.1% for the nonlinear power spectrum test in the code comparison suite of [35]. As already discussed, an important feature of the work presented here is the ability to carry out error-controlled approximate simulations at high throughput. In order to understand how we implement this, we provide some details on the HACC time-stepping algorithm. Evolution is viewed as a symplectic map on phase space: $\zeta(t) = \exp(-t\mathbf{H})\zeta(0)$ where, ζ is a phase-space vector $(\mathbf{x}, \mathbf{v})$, H is the (self-consistent) Hamiltonian, and the operator, $\mathbf{H} = [H, \;]_P$, denotes the action of taking the Poisson bracket with the Hamiltonian. Suppose that the Hamiltonian can be written as the sum of two parts; then by using the Campbell-Baker-Hausdorff (CBH) series we can build an integrator for the time evolution; repeated application of the CBH formula yields

$$\exp(-t(\mathbf{H}_1 + \mathbf{H}_2)) = \exp(-(t/2)\mathbf{H}_1)\exp(-t\mathbf{H}_2)\exp(-(t/2)\mathbf{H}_1) + O(t^3),$$

a second order symplectic integrator. In the basic PM application, the Hamiltonian H_1 is the free particle (kinetic) piece while H_2 is the one-particle effective potential; corresponding respectively to the 'stream' and 'kick' maps $M_1 = \exp(-t\mathbf{H}_1)$ and $M_2 = \exp(-t\mathbf{H}_2)$. In the stream map, the particle position is drifted using its known velocity, which remains unchanged; in the kick map, the velocity is updated using the force evaluation, while the position remains unchanged. This symmetric 'split-operator' step is termed SKS (stream-kick-stream). A KSK scheme constitutes an alternative second-order symplectic integrator.

In the presence of both short and long-range forces, we split the Hamiltonian into two parts, $H_1 = H_{sr} + H_{lr}$ where H_{sr} contains the kinetic and particle-particle force interaction (with an associated map M_{sr}), whereas, $H_2 = H_{lr}$ is just the long range force, corresponding to the map M_{lr}. Since the long range force varies relatively slowly, we construct a single time-step map by sub-cycling M_{sr}: $M_{full}(t) = M_{lr}(t/2)(M_{sr}(t/n_c))^{n_c}M_{lr}(t/2)$, the total map being a usual second-order symplectic integrator. This corresponds to a KSK step, where the S is not an exact stream step, but has enough M_{sr} steps composed together to obtain the required accuracy. (We take care that the time-dependence in the self-consistent potential is treated correctly; HACC uses the scale factor, a, as the time variable.) The code therefore has two degrees to tune its time-steps : the length of the full time step (t above), and the number of sub-cycles for the short range force (n_c above). As discussed later below, we will use the flexibility in the sub-cycling as a way of reducing the number of time steps such that the loss of accuracy only affects the resolution at very small scales, which, as discussed previously, are not of interest in the current set of simulations.

3 Time Step Tuning

In this section, we systematically examine how reducing the number of time steps affects individual gross halo properties (i.e., halo masses, positions, and velocities), as well as aggregate statistics like the mass function and spatial clustering. We run a set of convergence tests with boxes of size $(256h^{-1}\mathrm{Mpc})^3$ with 256^3 particles. These runs have the same particle mass as the full $(4000h^{-1}\mathrm{Mpc})^3$ volume simulations we present later. We run these with the following time step options : 450/5, 300/3, 300/2, 150/3 and 150/2 where the first number is the number of long time-steps, while the second is the number of subcycles. The 450/5 case has been independently verified to give fully converged results and is the baseline against which we compare all other results. Each simulation is started from the same initial conditions and evolved down to $z = 0.15$. We demonstrate that the 300/2 case, corresponding to $\Delta a \approx 0.003$ reproduces the full resolution simulation for all the large scale properties we consider, and is our choice for the mocks presented in Sec. 5.

3.1 Matching Halos : An Algorithm

In order to compare detailed halo properties, we need to match individual halos between our reference (450/5) run and test runs with degraded time steps. We first discuss the algorithm used for identifying the corresponding halos in the two cases and then compare halo mass, position, and velocity for the matched halos. From this quantitative comparison, we find that the simulations with 300 global time steps have significantly less scatter in the measured quantities, compared to the baseline determined by the 450/5 simulation, than do the samples with 150 global steps. In addition, we find that the differences between the different sub-cycling choices are almost negligible.

3.1.1 Algorithm

All simulations share the same particle initial conditions, allowing us to match halos in different runs by matching their individual particle content. Given a halo in simulation A, we consider all halos in simulation B that between them hold all the particles belonging to the halo in simulation A. Given this list of possible matches, we choose the run B halo with the largest number of common particles with the reference halo in run A. To avoid spurious matches, we also require that the fraction of common particles (relative to simulation A) exceeds a given threshold. To illustrate how this matching algorithm works, we use the samples from the 300/2 simulation and the 450/5 simulation, and adopt a threshold of 50% as our default choice.

The matching algorithm described above is unidirectional - multiple halos in run A may have particles resident in a single halo in run B. In our simulations, this happens at the 1-2% level, adopting a particle matching threshold of 50%. We refer to these cases as 'multiply-booked' halos. Figure 1 compares halo masses matching the 450/5 simulation to the 300/2 simulation for the case of multiply-booked halos, as well as the rest. The top left panel shows the mass scatter for all the matched halos between the two simulations, while the top right panel shows the mass scatter only for the non multiply-booked halos. The bottom left panel shows the mass scatter for individual multiply-booked halos, while the bottom right panel plots the summed halo mass for the corresponding halos. The overall behavior represented in Figure 1 is straightforward to interpret.

As the top left panel shows, there are low-mass halos in the 450/5 simulation matched to high-mass halos in the 300/2 simulation. The same trend is observed for the case of multiply-booked halos (bottom left panel), but not for the non-multiply-booked halos (top right). Furthermore, the disagreement for halo masses between the two simulations are resolved by adding the corresponding halo masses. This implies that there are multiple halos in the 450/5 simulation which are merged into one halo in the 300/2 simulation. The smaller number of time steps in the 300/2 simulations reduces substructure as well as the compactness of the halos compared to the 450/5 simulation. Thus, for a small fraction of halos in the 450/5 simulation, individual halos can be merged into a single halo in the 300/2 simulation.

Figure 2 shows the number densities of the unmatched halos in the 450/5 simulation when compared to the 300/2 simulation at $z = 0.15$. There are three reasons that halos can turn up as unmatched. In the first case, particles forming a halo in simulation A may not form a component of a halo in simulation B (no common particles). Second, if the fraction of common particles over the total number of particles in each halo is less than the threshold of 50%, these halos will be eliminated from the matching set. Finally, for the case of multiply-booked halos, we remove all but the one with the largest number of common particles. In Figure 2, we show each type of unmatched number density as a function of halo mass. The first case occurs only for low halo masses, where low effective resolution in a simulation can lead to halo drop out (halos are too "fuzzy" to meet the Friends-of-Friends (FOF) overdensity criterion), and falls off steeply with rising halo mass. Most of the unmatched halos arise due to their not passing the threshold criterion. The fraction of unmatched halos due to being "multiply-booked" is similar to the threshold case, albeit at a lower level.

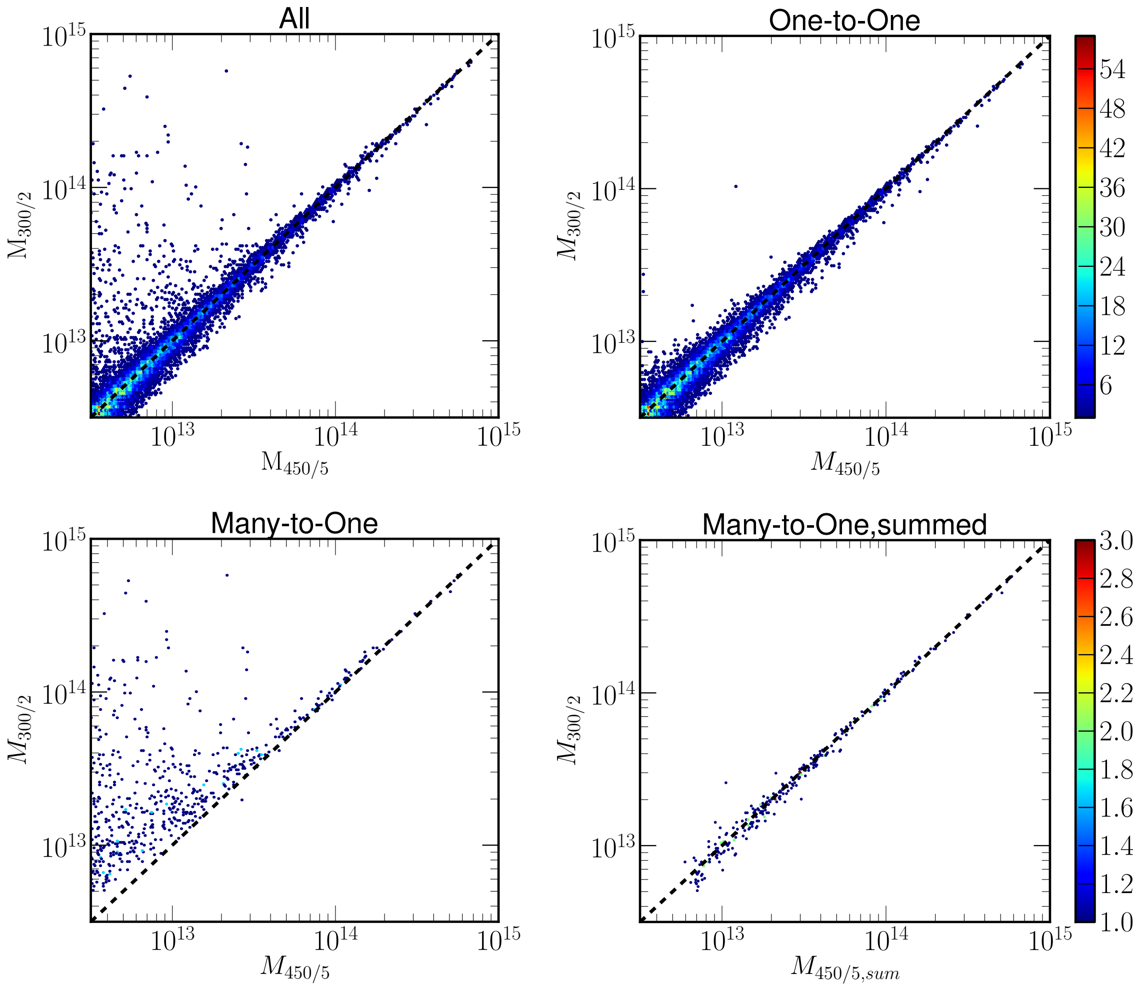

Figure 1. Comparing the halo mass of matched halos in the 450/5 simulation (x-axis) to the 300/2 simulation (y-axis) at $z = 0.15$. Panels correspond to halos with different matching criteria imposed: all the matched halos (top left), the vast majority of matched halos having one-to-one correspondence (top right), matched halos not having one-to-one correspondence called "multiply-booked" halos (bottom left), and the multiply-booked halos whose corresponding halo masses are added (bottom right). The results shown in these panels imply that the low-mass scatter between the 450/5 simulation and the 300/2 simulation shown in the top left panel arises when physically associated halos in the 450/5 simulation are merged into one halo in the 300/2 simulation due to an effectively worse resolution in this case.

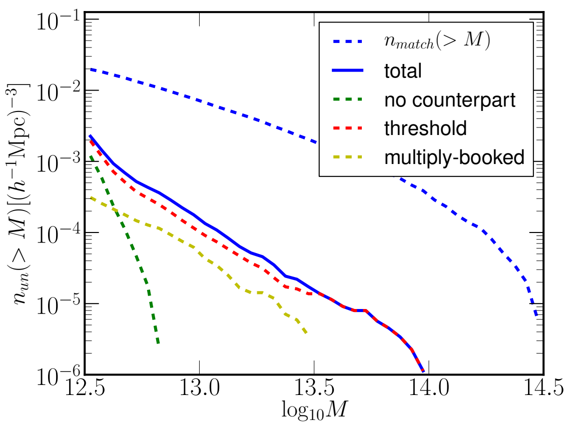

Figure 2. Itemization of unmatched halos (from the 450/5 and 300/2 simulations at $z = 0.15$) shown as cumulative number densities of the unmatched halos arising from each procedure in the matching algorithm. The solid blue line is the total number density of the unmatched halos. The dashed green line shows halos with no counterpart – none of the particles were identified as belonging to a halo in the comparison simulation; this is significant only at low halo mass. The dashed red line shows halos eliminated because of not meeting the matching threshold (i.e., the halos do not have enough of a fraction of the same particles). The dashed cyan line is for the halos eliminated because multiple halos correspond to one halo (see text).

The trends described above persist at different redshifts. As we reduce the number of total long timesteps taken, the unmatched fraction increases, due to the lower resolution of the simulations, while the number of subcycles does not noticeably change these results. In all the halo comparisons that follow, we restrict ourselves to halos with a one-to-one correspondence (i.e. not multiply-booked halos).

3.1.2 Halo Properties

We now systematically compare halo properties (i.e., halo mass, position, and velocity) for halos in the lower resolution runs that were successfully matched to those in the 450/5 simulation. We are interested in correctly describing the large-scale distribution of galaxies using an HOD approach; this requires that only the dark matter halo locations and masses be estimated sufficiently accurately.

The comparison of halo mass for different time-stepping schemes to the 450/5 simulation at $z = 0.15$ is shown in Figure 3. We take all the matched halos whose masses are between $10^{12.5}\mathrm{M}_\odot$ to $10^{13.0}\mathrm{M}_\odot$, $10^{13.0}\mathrm{M}_\odot$ to $10^{13.5}\mathrm{M}_\odot$, and $10^{13.5}\mathrm{M}_\odot$ to $10^{14.0}\mathrm{M}_\odot$, and compute their means and standard deviations for $\mathrm{M}/\mathrm{M}_{450/5}$, where $M_{450/5}$ is a halo mass for the 450/5 simulation and M corresponds to a halo in the samples generated with different time-stepping schemes. Figure 3 shows that halos generated from the simulations with small number of time steps have systematically lower FOF masses than those in the 450/5 simulation. (The same linking length ($b = 0.168$) is used in the FOF algorithm to define halos for all the simulations.) For the case of the 300/2 simulation, the deviation from the FOF masses in the 450/5 simulation is about 1.7%, while for the case of the 150/3 and 150/2 simulations, the deviations are about 5.7% and 8.5% respectively. Under these circumstances, the FOF mass is highest in the 450/5 simulation, decreasing systematically with increase in loss of temporal resolution.

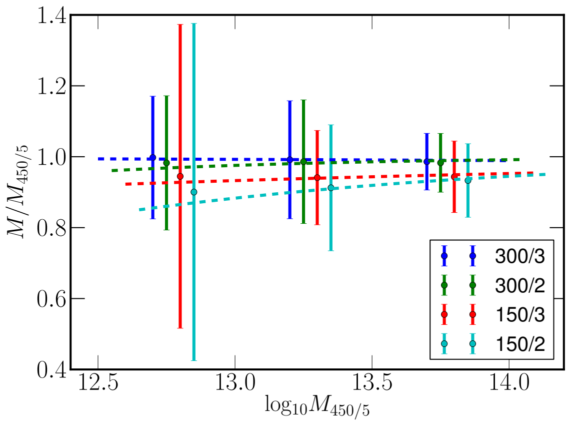

Figure 3. Comparison of halo mass (FOF, $b = 0.168$) for matched halos between the 450/5 simulation and coarsened time-stepping schemes at $z = 0.15$. We take all the matched halos whose masses are between $10^{12.5}\mathrm{M}_\odot$ to $10^{13.0}\mathrm{M}_\odot$, $10^{13.0}\mathrm{M}_\odot$ to $10^{13.5}\mathrm{M}_\odot$, and $10^{13.5}\mathrm{M}_\odot$ to $10^{14.0}\mathrm{M}_\odot$, and compute the mean and the standard deviation for $\mathrm{M}/\mathrm{M}_{450/5}$ where $M_{450/5}$ is a halo mass for the 450/5 simulation and M is for the simulations with different number of time steps corresponding to different colors in the plot. The x-positions have been displaced to avoid overlapping the error bars. Halo masses decrease systematically as the time resolution is coarsened.

Figure 4 shows the differences in positions (left panel) and velocities (right panel) for the matched halos at $z = 0.15$. Simulations with a smaller number of global time steps (150) display significantly more scatter; they also show a small bias in the velocity. With 300 global time steps, the results are much improved; the velocity bias is almost entirely removed and the scatter is significantly reduced. The standard deviation in the differences in halo distances is matched is better than $200h^{-1}$kpc in these cases, and better than 50km/s in velocities. As a reference, the standard deviation of velocities for the full resolution simulation (450/5) is about 300km/s. The distributions are very close to Gaussian shown as dashed lines in the figure. As is clear from Figure 4, the difference between 3 and 2 sub-cycles is insignificant for our purposes. We observe the same trend in halo properties discussed here at different redshifts.

As shown in Figure 2, the fraction of unmatched halos in the 300/2 simulation to the 450/5 simulation is less than 5% on most of halo mass ranges, which implies that the 300/2 simulation has almost the same number of halos as in the 450/5 simulation. Furthermore, Figure 1 shows that the halo masses in the 450/5 and 300/2 simulations have linear relation with the slope being one. So, most of halos in the 300/2 simulation have the same mass as the ones in the 450/5 simulation. Since the number of sub-cycles do not affect to halo positions and velocities as shown in Figure 4, the 300/2 time step is our choice to save the simulation time while keeping the halo properties almost identical to the 450/5 simulation.

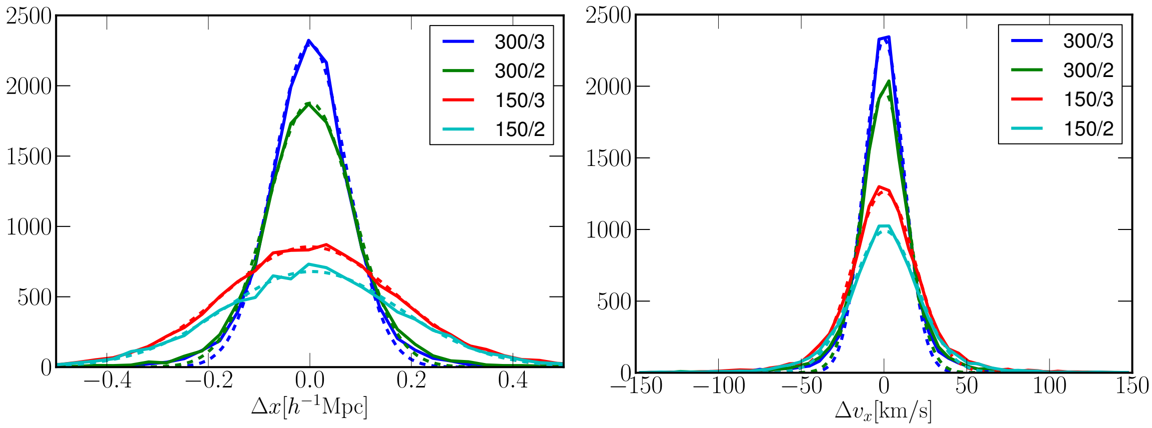

Figure 4. Comparison of the positions (left) and velocities (right) of halos matched across simulations with different time steps. The reference simulation is 450/5 while the colors correspond to 300/3 (blue), 300/2 (green), 150/3 (red), and 150/2 (cyan). The dashed lines are Gaussian fit.

The results shown in Figures 2, 1, and 4, show that the the 300/2 option has a low ratio of unmatched halos (less than 5%), excellent halo mass correlation to the 450/5 simulation (the small mass bias can be easily corrected as described below), and sufficiently small scatter in halo position. This time-stepping option is therefore a good candidate for generating mock catalogs efficiently, while maintaining high accuracies. In terms of the time savings alone, this will result in an increased capacity to generate high quality catalogs by a factor of four, which is quite significant. We will consider memory and storage savings further below in Section 4.

3.1.3 Halo Mass Adjustment and Resulting Observables

Halos generated by the de-tuned simulations have systematically lower masses than the halos in the 450/5 simulation as shown in Figure 3. In the following, we describe how to implement a systematic mass correction by matching to the 450/5 results; we also display the resulting observables including mass functions and power spectra.

To undertake the mass calibration, we first take all the matched halos between the 450/5 simulation and the de-tuned simulations and compute means for each mass bin. We consider only the matched halos because the aim of the mass adjustment is to correct systematic mass differences for the halos that are theoretically identical in the different runs. After computing the means for each mass bin, we fit them to a functional form that brings the reassigned halo mass, M_{re}, close to the average halo mass for the 450/5 simulation. For our simulations, we find that the following simple form suffices for this task:

$$M_{re} = M(1.0 + \alpha(M/10^{12.0}[\mathrm{M}_{\odot}])^{\beta}, \tag{3.1}$$

	α	β
300/3	0.005	0.175
300/2	0.07	-0.47
150/3	0.101	-0.162
150/2	0.315	-0.411

Table 1. Mass reassignment parameters α and β of Eq. 3.1 for simulations run with different numbers of time steps (the results are shown at $z = 0.15$).

where M_{re} is the reassigned halo mass, M is the original halo mass, and α and β are free parameters. The α and β values for the simulations with different numbers of time steps are listed in Table 1 (at $z = 0.15$). The best-fit parameters α and β are functions of redshift. For the case of the 300/2 simulation, the best fit parameters are $\alpha(z) = 0.123z + 0.052$ and $\beta(z) = -0.154z - 0.447$.

Given the mass corrections, we now compute mass functions using the results from the different time-stepping schemes, as shown in Figure 5, where we use the 450/5 simulation at $z = 0.15$ as the reference. In Figure 5, we show the ratio $n(> M)/n_{450/5}(> M)$, where $n_{450/5}(> M)$ is a cumulative mass function for the 450/5 simulation and $n(> M)$ is a cumulative mass function for the other cases. We compare the results before and after mass adjustment. While the mass functions from the 250/3 and 150/2 simulations are suppressed by more than 10% on all mass ranges before correction, they are significantly improved afterwards, especially for halo masses greater than $10^{13.0}\mathrm{M}_\odot$. For simulations with 300 global time steps, the mass adjustment is especially effective at small halo masses.

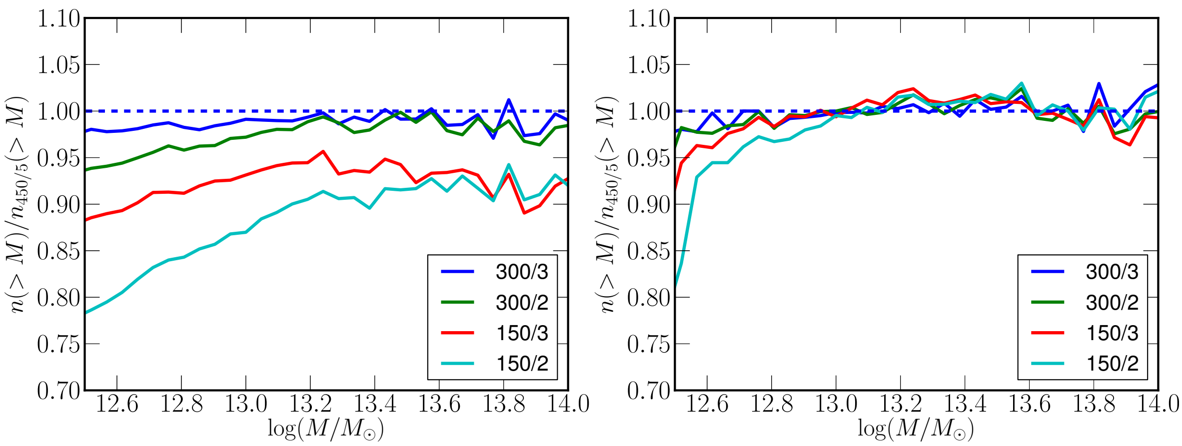

Figure 5. Comparison of cumulative mass functions in different simulations taking the 450/5 simulation as a reference. Lines, from top to bottom, correspond to the time stepping choices, 300/3 (blue), 300/2 (green), 150/3 (red), and 150/2 (cyan) respectively. The left panel shows the cumulative mass functions for unadjusted masses (as described in the text), while the right panel shows the post-correction results. A simple mass recalibration allows one to successfully recover the mass functions, even in the extreme case of the 150/2 simulation, for which the original result differed by more than 10% (on all mass scales).

We next compute the halo-matter cross power spectra between halo and matter density fields in both real and redshift space, as shown in Figure 6. This figure shows the ratio $P_{hm}/P_{hm,450/5}$ at $z = 0.15$, where $P_{hm,450/5}$ is the cross power spectrum for the 450/5 simulation and P_{hm} is the cross power spectrum for other time steps. For the dark matter density field, we use the output of the 450/5 simulation for all the halo samples. Note that the dark matter density fields are in real-space for both cases. In this way, the ratio $P_{hm}/P_{hm,450/5}$ in real-space is equivalent to the ratio of halo bias between the 450/5 simulation and the simulations with other time-steps. To select halos, we apply the soft-mass cut method using the probability given by

$$\langle N_{halo}(M)\rangle = \frac{1}{2}\mathrm{erfc}\left(\frac{\log(\mathrm{M_{cut}}/M)}{\sqrt{2}\sigma}\right), \tag{3.2}$$

where we set $\mathrm{M_{cut}} = 10^{13.0}[\mathrm{M}_\odot]$ and $\sigma = 0.5$. This probability has a similar form to the HOD technique so that the probability gradually becomes one as halo mass increases. We use this method to avoid noise from halos scattering across sharp halo mass boundaries. The errors calculated here are not due to sample variance as we generate 10 samples from one full sample with the soft-mass cut method. The results show that as the time stepping is coarsened, the ratio of the cross power spectra increases, especially in redshift-space, where we observe large deviations from unity on small scales for the 150/2 and 150/3 simulations. This is due to the overall smaller halo velocities for those simulations, as shown in Figure 4. For the simulations with the 300 global time steps, overall agreement with the 450/5 simulation is almost at the 1% level on any scale in both real-space and redshift-space. Based on these convergence tests, we conclude that the 300/2 option meets the error requirements.

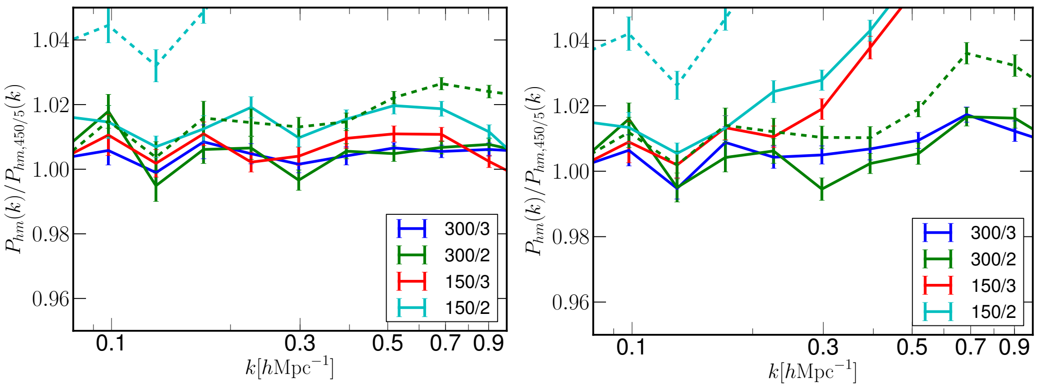

Figure 6. Ratio of halo-matter cross power spectra as a function of time steps with respect to the 450/5 simulation at $z = 0.15$. We use the real-space halo density field for the left panel and the redshift-space halo density field for the right panel; the dark matter density fields used here are in real-space for both cases. The left panel shows that agreements with the 450/5 simulation are all within 2%. The dashed line shows the case without mass corrections for the 150/2 simulation. In the right panel, the large discrepancy of the cross power spectra for the simulations with 150 global steps on small scales is mainly due to the systematically small velocities shown in Figure 4. Note that the halos are selected based on the soft mass-cut method with $M_{cut} = 13.0$ and $\sigma = 0.5$.

4 Constructing Light Cones

One can choose to construct mock catalogs either from a single static snapshot in redshift or to take the lightcone evolution of the halo and galaxy distribution into account. There are two broad approaches to the latter problem. The first is to incrementally build up the lightcone while the simulation is running, while the second stitches together static snapshots at different redshifts to construct an approximate lightcone. Both methods have advantages and disadvantages. Building the lightcone "on the fly" gets the redshift evolution correct by construction, but requires simulation post-processing codes (like halo finders) to be run concurrently with the simulation, making it nontrivial to change input parameters to these analyses. Using static time snapshots can be more flexible, but requires a relatively dense sampling of timesteps. We take this approach here; this section discusses how we interpolate between different snapshots and the requirements on the time step sampling. We note that these results are more broadly applicable that just for this simulation suite.

Our algorithm for constructing a lightcone is :

1. For every snapshot, construct a spherical shell (or part thereof) with a radius centered on the comoving distance to the snapshot redshift and a width equal to its redshift extent.

2. The redshifts of each halo are determined by their radial distance from the origin.

3. Interpolate the position and velocity (see below) of the halo from the snapshot redshift to the halo redshift. Note that changing the position of the halo strictly changes its redshift, but this is a small effect (mostly $< 0.1\%$).

4. For the case of a halo crossing its boundary of the shell, we choose the halo whose distance from the boundary is closer before shifting.

We interpolate the positions of halos using a simple linear rule :

$$\vec{x}|_{z=z_{pos}} = \vec{x}|_{z=z_{snap}} + \vec{v}_{pec}|_{z=z_{snap}} \Delta t, \tag{4.1}$$

where z_{snap} is the redshift of the snapshot, z_{pos} is the redshift corresponding to its radial position, $\vec{v}_{pec}$ is its peculiar velocity, and Δt is the time elapsed between z_{snap} and z_{pos}. We test this using the calibration simulations used in the previous simulations, restricting ourselves to halos identified across two or more timesteps. The left panel of Fig. 7 plots the x-component of the distance between halos identified in the $z = 0.25$ and $z = 0.15$ simulations, both before and after shifting halos from $z = 0.25$ to $z = 0.15$ using the above equation. The improvement is clearly visible; the scatter after shifting is half of the original scatter with similar results for the other components and different redshift slices. The right panel of the same figure shows the ratio of the power spectrum of the shifted halos to its expected value. We consider shifts in redshift ranging from $\Delta z = 0.05$ to $\Delta z = 0.25$. Except for the $\Delta z = 0.25$ case at high k, we recover the expected power spectrum to better than 1% out to $k = 1h\text{Mpc}^{-1}$; the accuracy is a few tenths of a percent for most cases over most of that range. We get similar results for shifting to different redshifts.

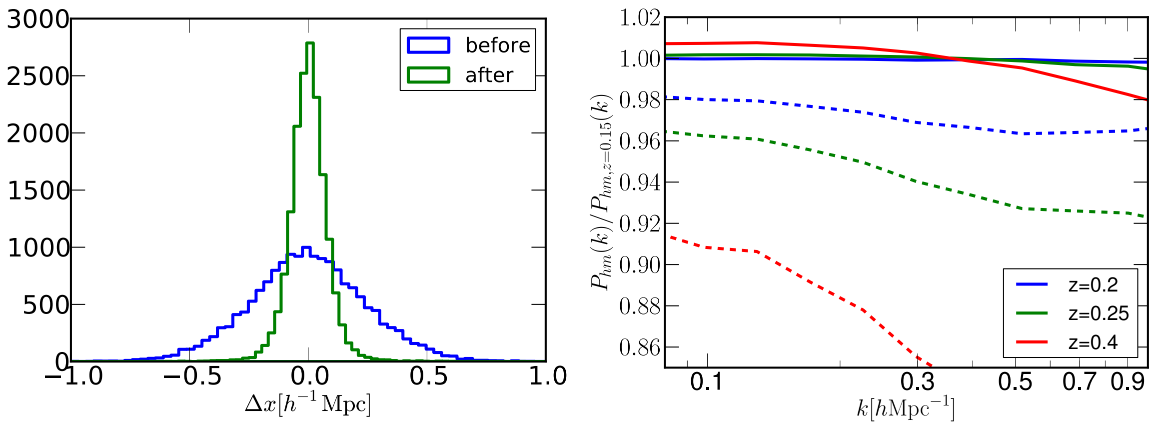

Figure 7. Left: Comparison of the x-component of the distance before and after shifting the positions of halos from $z = 0.25$ to $z = 0.15$. We assume that the peculiar velocity is constant between the snapshots (described in Eq. 4.1) to compute the distances to shift. The plot shows that the scatter after shifting shrinks to half of the original scatter. Right: The ratio of the power spectrum of the shifted halos to its expected value. We shift halos from the redshift shown in the plot ($z = 0.4$, 0.25, and 0.2) to $z = 0.15$ and compute the halo-matter cross power spectra with the matter density field at $z = 0.15$. The denominator is the cross power spectra at $z = 0.15$, which is an expected value after shifting. The solid line shows the results after shifting, while the dashed lines are before shifting. We recover the expected power spectrum to better than 1% out to $k = 1h\text{Mpc}^{-1}$ for the case of $\Delta z < 0.25$.

Working in redshift space requires interpolating both the positions and the velocities of halos. Since we do not store accelerations of halos, we simply linearly interpolate the halo velocities between two redshifts, z_1 and z_2 :

$$\vec{v}|_{z=z_{pos}} = \vec{v}|_{z=z_1} \frac{z_{pos} - z_2}{z_1 - z_2} + \vec{v}|_{z=z_2} \frac{z_{pos} - z_1}{z_2 - z_1}, \tag{4.2}$$

where $z_1 < z_{pos} < z_2$ and z_{pos} is the corresponding redshift for positions. Fig.8 uses the redshift slices at $z = 0.15$ and 0.25 to predict the values at $z = 0.2$; we recover the true velocities with a scatter of 38.8km/s, which is improved from the original scatter of 57.4km/s as shown in the left panel. The right panel of the same figure plots the ratio of the predicted angle-averaged power spectrum to the expected value. While the accuracy is degraded compared to the real space case, we still recover the power spectrum to 1% out to $k = 0.8h\text{Mpc}^{-1}$. The fall off at large k is expected since the errors in the velocities introduce a random scatter in the positions of the galaxies, washing out the signal on small scales.

These results suggest using a redshift spacing of $\Delta z = 0.1$ or better between different simulation outputs. For the simulation suite discussed in the next section, we choose to be conservative and store outputs every $\Delta z = 0.05$ over the region on interest, corresponding to a maximum shift in redshift of 0.025 (smaller than all the cases considered here).

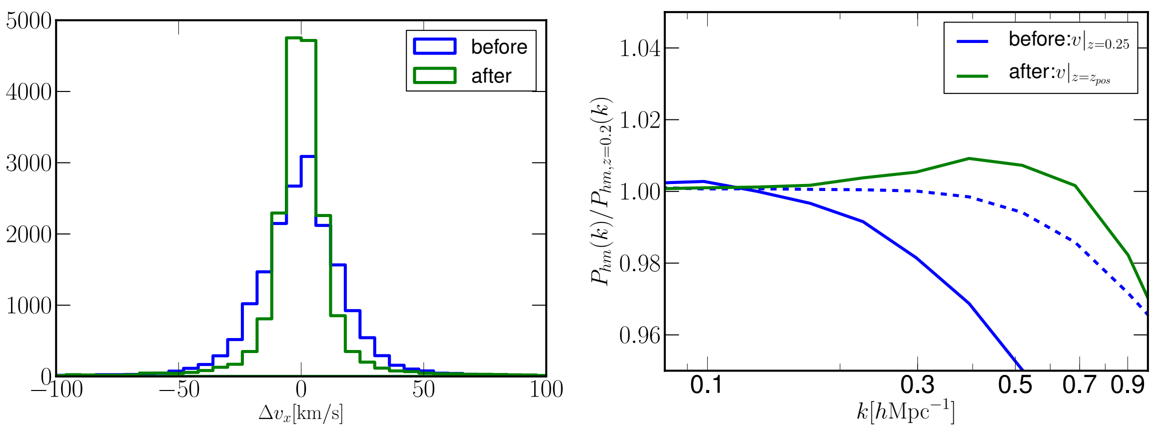

Figure 8. Left: Comparisons of the x-component of velocities to the ones at $z = 0.2$ before and after linear interpolation. We estimate velocities at $z = 0.2$ by linearly interpolating velocities at $z = 0.25$ and $z = 0.15$. The blue line shows the original scatter of the velocities between $z = 0.25$ and $z = 0.2$, while the green line shows the scatter after linear interpolation. Right: The ratio of the power spectrum of the shifted halos to its expected value in redshift space. We shift halos from $z = 0.25$ to $z = 0.2$ and use original velocities at $z = 0.25$ (blue line) and the linearly interpolated velocities (green line) to define redshift space. The denominator is the redshift space cross power spectra at $z = 0.2$, which is an expected value after shifting. As a reference, the result of shifting from $z = 0.25$ to $z = 0.2$ in real space is shown as a dashed line. We recover the expected power spectrum to better than 1% out to $k = 0.8h\text{Mpc}^{-1}$ by linearly interpolating velocities.

5 BOSS Mock Catalogs

As a concrete implementation of the approach discussed above, we construct catalogs designed to mimic the Baryon Oscillation Spectroscopic Survey (hereafter, BOSS) galaxy samples. BOSS ([36]), part of the SDSS-III project ([37]), is a spectroscopic survey that has achieved percent level distance measurements using the baryon acoustic oscillation technique ([38]). The low redshift ($z < 0.7$) distance measurements use two galaxy samples : the LOWZ ($z < 0.45$) and CMASS ($z < 0.7$) samples ([38, 39]). We describe the construction of the CMASS sample below. The construction of the LOWZ sample is analogous, and the same simulations (as different time slices) can be used in its construction.

We choose a simulation volume large enough to build a full-sky mock catalog. Since the CMASS sample extends to $z \sim 0.7$, we choose a simulation side of $4000h^{-1}$Mpc, corresponding to a comoving

	WMAP-7	WMAP-7	WMAP-7	Planck	Planck	Planck
Ω_m	0.265	0.265	0.265	0.315	0.315	0.315
$\Omega_b h^2$	0.02258	0.02258	0.02258	0.02202	0.02202	0.02202
h	0.71	0.71	0.71	0.673	0.673	0.673
n_s	0.963	0.963	0.963	0.96	0.96	0.96
σ_8	0.8	0.83	0.77	0.8	0.83	0.77

Table 2. Cosmological parameters used to run the simulation. For each model, we run one full simulation (meaning with very accurate time stepping) out to $z = 0$, and 5 realizations that have reduced time steps out to $z = 0.15$.

distance to $z \sim 0.8$ from the center of the box. We start the simulations at $z = 200$ using Zel'dovich initial conditions and are run down to $z = 0.15$ using the time-stepping procedure described earlier in the paper. We store outputs every $\Delta z \sim 0.05$ starting at $z \sim 0.75$; in addition, we store outputs uniformly spaced by $\Delta z = 0.2$ between $z = 1$ and $z = 2$ as well as at $z \sim 2.5$, 3 and 4. The close spacing at low redshift enables us to make lightcones using the method described in the previous section. We run a friends-of-friends halo finder on each of the outputs with a linking length of $b = 0.168$, keeping halos down to 40 particles, or $10^{12.6} M_\odot$. By comparison, the characteristic halo mass of the BOSS galaxies is $10^{13} M_\odot$, which we resolve with 100 particles. Each output stores central positions, mean velocities and halo masses , and 1% of halo particles with minimum of 5 particles per halo and 1% of all particles randomly sampled. We run the simulations with a range of cosmologies centered at WMAP-7 ([40]) and Planck ([41]) with varied σ_8, shown in Table 2. For each model, we save one full simulation (meaning with very accurate time stepping) out to $z = 0$, and 5 realizations that have reduced time steps out to $z = 0.15$. In this section, we use four realizations from the WMAP-7 cosmology with $\sigma_8 = 0.8$. For this case, one realization was failed to save.

To generate the galaxy mock catalogs, we first populate halos with galaxies using a halo occupational distribution (HOD) approach. The HOD functional form (based on a number of free parameters, 5 in our case) provides probabilities for the number of central and satellite galaxies based on the masses of halos that host those galaxies. A halo hosts a central galaxy with probability $< N_{cen}(M) >$ and a number of satellite galaxies given by a Poisson distribution with mean $< N_{sat}(M) >$:

$$< N_{cen}(M) >= \frac{1}{2}\mathrm{erfc}\left[\frac{\ln(M_{cut}/M)}{\sqrt{2}\sigma}\right], \tag{5.1}$$

and

$$< N_{sat}(M) >= N_{cen}(M)\left(\frac{M - \kappa M_{cut}}{M_1}\right)^{\alpha}, \tag{5.2}$$

where M_{cut}, M_1, σ, κ, and α are free parameters and M is the halo mass. We assume that $N_{sat}(M)$ is zero when $M < \kappa M_{cut}$ and halos do not host satellite galaxies without a central galaxy [42]. The total number of galaxies hosted by each halo is a sum of the number of central and satellite galaxies. Equations 5.1 and 5.2 are not the only possible functional form for the HOD, and it is trivial to change this. However, these forms are known to successfully reproduce the clustering of the BOSS galaxies [43] and are therefore a convenient choice.

After assigning a number of galaxies to each halo, we distribute those galaxies within the halo. The central galaxy is always at the center of the halo, while the distribution of satellite galaxies follow a spherically symmetric NFW profile specified by:

$$\rho(r) = \frac{4\rho_s}{\frac{cr}{R_{\rm vir}}(1 + \frac{cr}{R_{\rm vir}})^2}, \tag{5.3}$$

where ρ_s is the density at the characteristic scale $r_s = R_{\rm vir}/c$, $R_{\rm vir}$ is the virial radius for the halo and c is the concentration parameter. $R_{\rm vir}$ is the virial radius of the halo. We use a cosmic emulator [44] to generate a table of concentration-mass relation for halos at each redshift with the given cosmology.

The velocity of the central galaxy is equal to the host halo velocity. We assume that satellite galaxies are randomly moving inside the host halos. Therefore, the velocities of the satellite galaxies

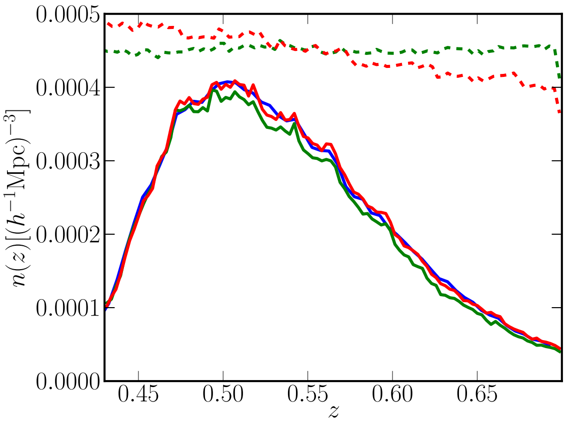

Figure 9. Normalized redshift distribution of galaxies from DR11 (North) in Ref. [45] (blue solid line), and a comparison of galaxy number densities before fitting to DR11 redshift selection function (dashed lines) and after (solid lines). The green and red lines are from the mocks at $z = 0.55$ and the lightcone output respectively. The HOD parameters used to generate the mock catalogs can be found in text.

are the sum of their host halo velocity and a random virial component. For this random component, we draw a Gaussian distribution with zero mean and variance given by:

$$< v_x^2 >=< v_y^2 >=< v_z^2 >= \frac{1}{3}\frac{GM}{R_{vir}}. \tag{5.4}$$

Following the above procedures, we generate two galaxy mocks from the simulation at $z = 0.55$ and from the lightcone sample described in Section 4. Using the simulation at $z = 0.55$ is the way to generate galaxy mocks described in Ref. [26, 27] and used in [38], The reason we also generate a galaxy mock from the lightcone sample is to explore whether those two galaxy mocks exhibit any differences or not. We use the following HOD parameters, $M_{cut} = 12.9$, $M_1 = 14.0$, $\alpha = 1.013$, $\kappa = 1.0$, $\sigma = 0.85$, to populate the halos with galaxies. Note that the goal here is not to completely fit to the observed DR11 correlation functions, but to show that our simulations have capability to compute correlation functions for BOSS.

Once generating the galaxy mocks, the next step is matching the number density $n(z)$ to the redshift selection function of Ref. [45]. For each redshift bin, we randomly subsample galaxies. Figure 9 shows redshift distributions of galaxies before and after subsampling for BOSS CMASS North Galactic Cap, which correspond to dashed and solid lines respectively. The blue solid line is the redshift distribution of galaxies for BOSS, while the green and red line represents $n(z)$ from the mock at $z = 0.55$ and the lightcone sample.

After matching the redshift distribution of galaxies, we finally compute correlation functions for those galaxy mocks. In Figure 10, we compare those correlation functions to the one from DR11. The left panel shows the monopole terms of the correlation functions, and the right panel is the quadrupole terms. We do not see siginificant differences between the galaxy mocks from the simulation at $z = 0.55$ and the lightcone sample for both monopole and quadrupole terms. For the monopole terms, the mocks agree relatively well with the one from DR11 on $r \in [40h^{-1}\mathrm{Mpc}, 80h^{-1}\mathrm{Mpc}]$. Note that we do not try to fit to the acoustic peak hear, because the cosmologies for our simulation and DR11 are different. On the other hand, we fail to fit the quedrupole terms to the observed quadrupole term from DR11 by overestimating the power, which also happens in Ref. [12].

One of the other applications of those mocks is to use the mock correlation functions as a template function, which we discuss in detail in Section 5.1, to fit to an observed correlation function. As a demonstration, we fit our mock correlation functions to the one from DR11. We use the mock correlation function as a template shown in Eq. 5.6 and measure the value of α, which is the scale dilation parameter measuring the relative distance scales between observations and the fiducial model.

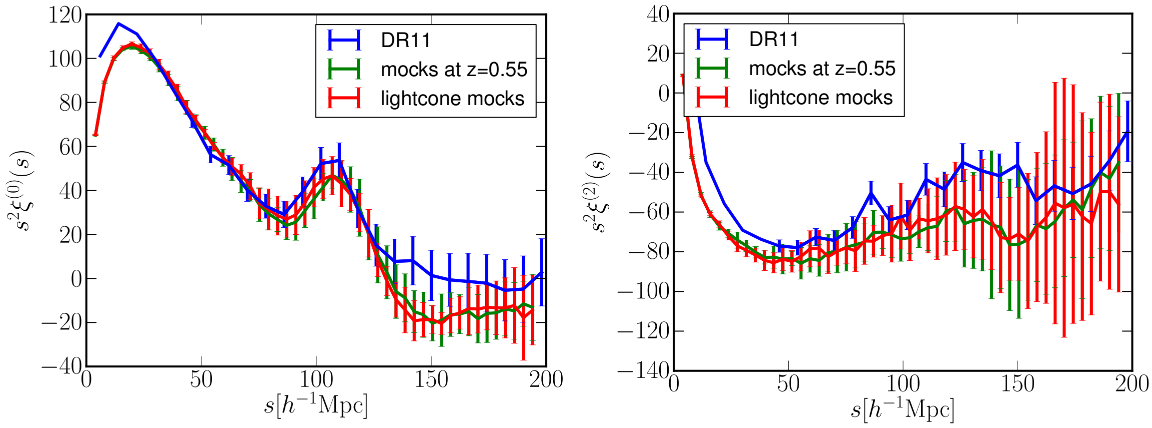

Figure 10. Correlation function monopoles $\xi^{(0)}(s)$ (left) and quadrupoles $\xi^{(2)}(s)$ (right) of the mocks (green and red) and DR11 in Ref. [26] (blue) at $z = 0.55$. The HOD parameters used to generate the mock catalogs can be found in text.

This parameter α is related to the ratio:

$$\alpha \equiv \left(\frac{D_\mathrm{V}(z)}{r_\mathrm{d}}\right)\left(\frac{r_\mathrm{d,fid}}{D_\mathrm{V}^\mathrm{fid}(z)}\right), \tag{5.5}$$

where r_d is the projection of the sound horizon at the drag epoch and $D_\mathrm{V} \equiv \left[cz(1+z)^2 D_\mathrm{A}(z)^2 H^{-1}(z)\right]^{1/3}$ where D_A is the angular diameter distance and $H(z)$ is the Hubble parameter. The ratio $\left(\frac{r_\mathrm{d,fid}}{D_\mathrm{V}^\mathrm{fid}(z)}\right)$ for our simulations is 13.44 at $z = 0.57$, while the ratio $\left(\frac{D_\mathrm{V}(z)}{r_\mathrm{d}}\right)$ from DR11 is 13.91 ± 0.13. So, the estimated value for α is 1.025 ± 0.010.

Figure 11 shows the χ^2 values as a function of the parameter α in the left panel and compares the monopole correlation function from DR11 and the fitted function in the right panel. We obtain the value of 1.033 ± 0.012, which is within 1σ from the estimated value for α. This demonstrates that our simulations can be used as a template function to measure the shift in the acoustic peak.

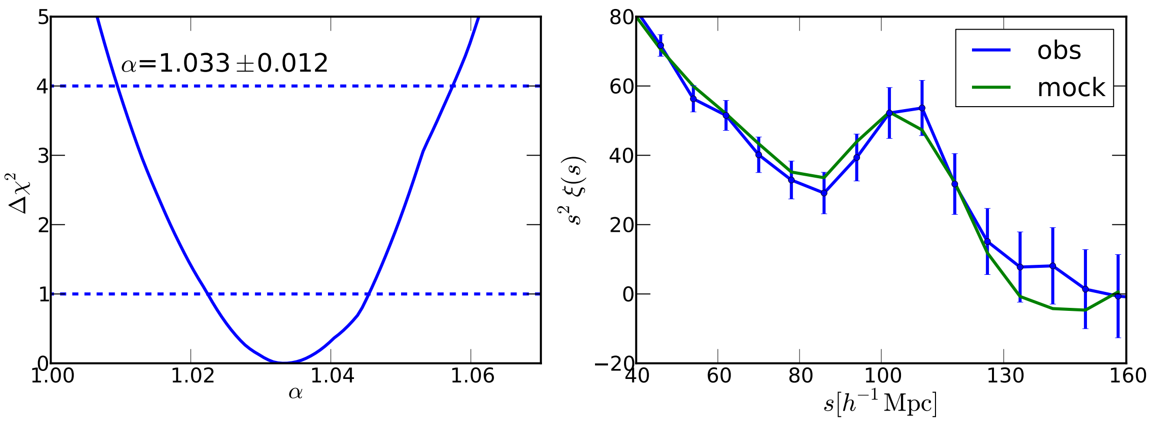

Figure 11. Left: Plot of $\Delta\chi^2$ vs. α for the data from DR11. The dashed lines (from top to bottom) correspond to 2σ and 1σ for the value of α. The best fit value of α is 1.033 ± 0.012. Right: The monopole correlation functions of the data from DR11 (blue line) and of the lightcone galaxy mock fitted with $\alpha = 1.033$.

5.1 Testing the shift of the BAO peaks

As a second application of this suite of simulations, we explore how the baryon acoustic oscillation (hereafter BAO) changes as a function of changing galaxy bias. These shifts, predicted by basic perturbative arguments (e.g.. [21, 23, 25]), must be calibrated at the sub percent level for future BAO experiments as a function of the underlying cosmology and galaxy type. Measuring these shifts requires large simulation volumes to robustly measure them from under the statistical error. The approximations we describe in this paper allow such large volumes to be run, without sacrificing the gross details (positions, masses, velocities) of the halos themselves (which, in turn, allow one to construct mock galaxy populations). Note that our goal here is to demonstrate the statistical power of these simulations, deferring an exhaustive study to later work.

In order to construct samples of different galaxy biases, we construct realizations of HODs with the same parameters as in Sec 5.1, except for M_{cut} which we vary from 12.5 to 14.1 in steps of 0.2. The resulting nine HODs span a range of biases from 1.5 to 3.5. In order to estimate the covariance matrix for galaxy correlation functions, we subdivide each of our $4 \times (4h^{-1}\mathrm{Gpc})^3$ volumes into $4 \times 64 = 256$ subvolumes. For simplicity, we analyze each of these subvolumes individually, except for the Mcut=13.9 and 14.1 samples where we analyze groups of 4 sub volumes to reduce the noise in the measurements. [1] We focus here on real-space measurements.

We measure the BAO scale using the methodology in Ref. [45]. Specifically, we describe the observed correlation function by ξ_{fit}:

$$\xi_{\mathrm{fit}}(r) = B\xi_t(\alpha r) + A_0 + A_1/r + A_2/r^2, \tag{5.6}$$

where ξ_t is a template correlation function, B is the galaxy bias squared and $A_{0,1,2}$ represent nuisance parameters to account for shot noise, nonlinear evolution of the matter density field and the mapping from matter to galaxies (NEED TO FIND A BETTER PHRASE FOR THE LAST; the galaxy distribution as a biased tracer of matter). The template correlation function $\xi_t(r)$ is given by Fourier transforming $P_t(k)$:

$$P_t(r) = (P_{\mathrm{lin}}(k) - P_{nw}(k))e^{-\frac{k^2\Sigma^2}{2}} + P_{nw}(k), \tag{5.7}$$

where $P_{\mathrm{lin}}(k)$ is the linear power spectrum and $P_{nw}(k)$ is the no-wiggle power spectrum described in Ref. [46], and Σ is a nonlinear parameter which accounts for broadening the BAO peak due to non-linear evolution. We set $\Sigma = 5h^{-1}\mathrm{Mpc}$. We determine the parameters by minimizing $\chi^2 = (\xi_{\mathrm{HOD}} - \xi_{\mathrm{fit}})^{\mathrm{T}} C^{-1} (\xi_{\mathrm{HOD}} - \xi_{\mathrm{fit}})$ where C^{-1} is the inverse covariance matrix and ξ_{HOD} is a correlation function which ξ_{fit} is fitted to. We consider the measured correlation function from $60h^{-1}\mathrm{Mpc}$ to $160h^{-1}\mathrm{Mpc}$ in bins of $4h^{-1}\mathrm{Mpc}$ for a total of 25 data points. Since all parameters except α are linear, we perform the minimization on a grid of α values, computing the minimum value of the other parameters directly. The parameter α measures the shift of the BAO peak from its original position predicted by the linear perturbation theory. Note that $\alpha = 1$ implies that there is no shift of the BAO peak.

The left panel of Figure 12 compares the correlation function computed from one of the full boxes with $M_{cut} = 12.9$ (which corresponds to the galaxy bias of 1.81) and the model correlation function described in Eq. 5.6 with the best-fit parameters. The best fit value of α for this mock is 1.003. Note that the error bars shown in the left panel of Figure 12 are computed from the covariance matrix. The right panel of Figure 12 shows the distribution of $\alpha - 1$ for the case of $M_{cut} = 12.9$ for the 256 samples. The mean value of $\alpha - 1$ is 0.0068 ± 0.0012.

[1] In these cases, we scale the covariance matrix by a factor of 1/4.

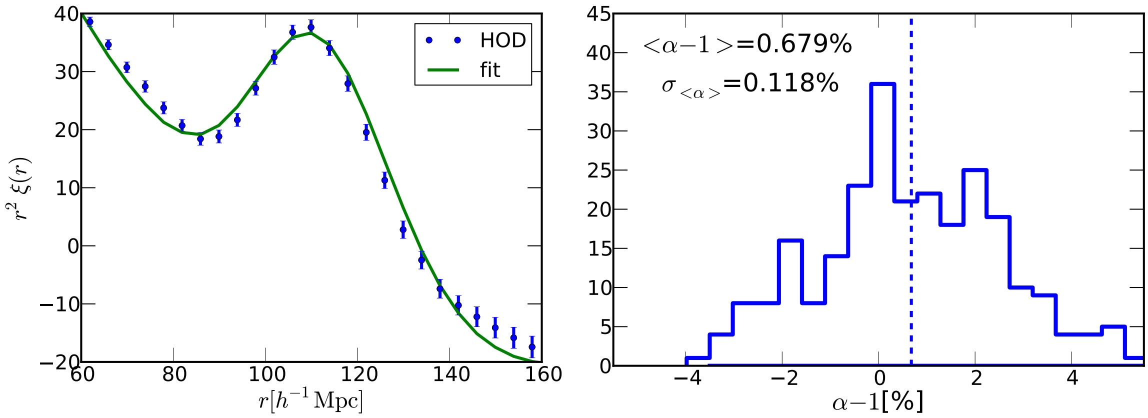

Figure 12. Left: The correlation function $\xi_{\rm HOD}(r)$ computed from the full HOD galaxy mock (labeled as "data") with $M_{cut} = 12.9$ (which corresponds to the bias value of 1.81) and the template correlation function with the best-fit parameters (labeled as "model"). The best fit value of α for this mock is 1.003. The error bars are computed from the covariance matrix. Right: Distribution of the values $\alpha - 1$ for the case of $M_{cut} = 12.9$. The dashed line corresponds to the mean value of $\alpha - 1$, which is 0.006.

Figure 13 shows the measured shifts in the BAO scale as a function of galaxy bias, for the nine HODs described above. We detect the shift in the BAO scale at high significance ($\sim 6\sigma$ for most samples) but do not detect a strong variation with the galaxy bias. Indeed, the shift is consistent with being constant for biases ¿ 3.0 and only increases for extremely large biases. These results agree with the trends in Ref. [47], although note that the absolute values of the shifts differ since Ref. [47] consider samples at $z = 1$, while the results here are at $z = 0.15$. We also observe that our results are roughly consistent with what one expects from perturbation theory ([21, 23, 25]), although we find a weaker mass dependence than what was predicted by Ref. [23].

This work can be extended by using reconstruction on the sample to verify that it does indeed reduce these biases both in real-space and in redshift-space, comparing to the perturbation theory results, testing redshift evolution and cosmology dependence in the shift of the peak.

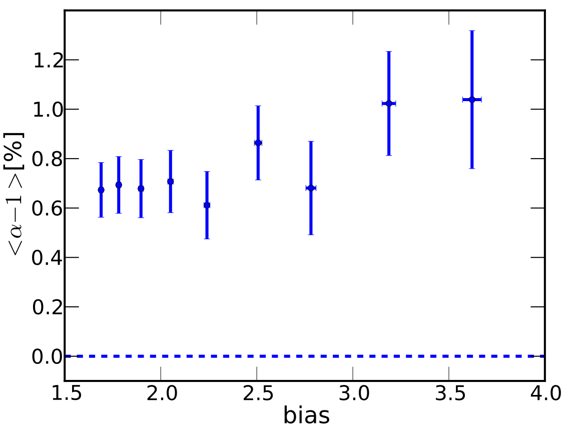

Figure 13. The shift in the BAO scale as a function of galaxy bias. Note that $\alpha - 1$ implies no shift in the BAO scale. We detect a significant shift ($\sim 6\sigma$ for most cases) for all the samples we consider. The shift is consistent with being constant for biases ¿ 3.0 and only increases for extremely large biases. These results agree with the trends in Ref. [47]. The third point from left uses the HOD parameter of $M_{cut} = 12.9$, which are the same samples used in Figure 12.

6 Discussion

Precision required for current and future galaxy spectroscopic surveys to test the expansion and structure formation histories of the Universe requires an accurate understanding of systematic effects. In this paper we have presented a quantitative study of the impact of time step sizes on the halo and matter density fields. Our code has two adjustable time stepping parameters - a global time step and a number of sub-cycles (responsible for a particle-particle interactions) to track that particle trajectories on small scales. We consider cases where we increase the length of each time step by factors of 1.5 and 3 respectively, as well as reducing the number of sub-cycles. Our fiducial choice is to use using 300 global time steps corresponding to $\Delta a(z) = 0.003$ and 2 sub-cycles (increasing the length of the global time step by a factor of 1.5 and that of the sub-cycles by a factor of 2.5), resulting in a reduction of the simulation run time by 4 times less. We keep the mass resolution constant; the results here are based on a particle mass of $6.86 \times 10^{10} h\mathrm{M}_{\odot}$.We summarize the key results below:

(a) The halo masses tend to be underestimated in these cases, as one might expect because reducing the number of time steps produces halos with less substructure and a more diffuse distribution of mass. However, this trend may be calibrated with smaller simulations and corrected, recovering the halo masses to 98%. The halo mass function is correctly recovered fully for masses above $10^{12.7} h^{-1}\mathrm{M}_{\odot}$ corresponding to 100 particles. Note that we run the halo finder with identical parameters as in the full resolution runs. It may however be possible to get similar results by changing the parameters of the halo finder, as was done in [26].

(b) The halo positions and velocities are recovered with a scatter of $0.08[h^{-1}\mathrm{Mpc}]$ and 12.8[km/s] respectively for the simulation of our fiducial choice.

(c) The clustering of these halos is correctly recovered to better than 1% on scales below $k < 1[h\mathrm{Mpc}^{-1}]$ in real-space and $k < 0.5[h\mathrm{Mpc}^{-1}]$.

(d) We find that the number of sub-cycles makes almost no difference to any of our final results.

We also consider the redshift sampling required to construct light cone outputs. We first compare the distances for the halos at different redshifts before and after shifting their positions from one redshift to the another. Moving halos over a $\Delta z = 0.1$ interval correctly reduces the standard deviation of those distances from $0.25[h^{-1}\mathrm{Mpc}]$ to $0.09[h^{-1}\mathrm{Mpc}]$. Moreover, the power spectra are correctly recovered to better than 1% for $k < 1h\mathrm{Mpc}^{-1}$ for $\Delta z < 0.25$. This worsens in redshift space to 2% up to $k < 0.2h\mathrm{Mpc}^{-1}$ by assuming that the change in velocity across the simulations is negligible and using the original velocity to compute redshift space. The agreement is, however, improved to 1% for $k < 0.8h\mathrm{Mpc}^{-1}$ by using the velocity linearly interpolated between the snapshots and the true redshift. Our fiducial choice to construct light cone outputs is $\Delta z = 0.05$.

This work is a natural extension of the approaches described in [12, 16, 17, 26, 27]. The primary goal for those papers was to generate the large numbers of simulations required for estimating covariance matrices. We quantify, in detail, the impact of size of the time step on large scale observables; our suite of simulations are better designed for testing for systematic errors in theory and analysis techniques. As a proof of principle, we present a set of eight full BOSS (both North and South Galactic caps simultaneously) simulations. The time savings presented in this paper allowed to extend this to 50 simulations, across a range of cosmologies centered at WMAP-7 ([40]) and Planck ([41]) with varied σ_8. These results from these will be presented in future publications.

Acknowledgement

TS and NP are supported by a DOE Early Career Grant. This work was supported in part by the facilities and staff of the Yale University Faculty of Arts and Sciences High Performance Computing Center, and by resources at the National Energy Research Scientific Computing Center. TS would like to thank Andrew Szymkowiak for useful discussions. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Argonne National Laboratory's work was supported under U.S. Department of Energy contract DE-AC02-06CH11357.

References

[1] D. Spergel, N. Gehrels, J. Breckinridge, M. Donahue, A. Dressler, B. S. Gaudi, T. Greene, O. Guyon, C. Hirata, J. Kalirai, N. J. Kasdin, W. Moos, S. Perlmutter, M. Postman, B. Rauscher, J. Rhodes, Y. Wang, D. Weinberg, J. Centrella, W. Traub, C. Baltay, J. Colbert, D. Bennett, A. Kiessling, B. Macintosh, J. Merten, M. Mortonson, M. Penny, E. Rozo, D. Savransky, K. Stapelfeldt, Y. Zu, C. Baker, E. Cheng, D. Content, J. Dooley, M. Foote, R. Goullioud, K. Grady, C. Jackson, J. Kruk, M. Levine, M. Melton, C. Peddie, J. Ruffa, and S. Shaklan, *Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA Final Report*, *ArXiv e-prints* (May, 2013) [arXiv:1305.5422].

[2] D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, L. Carey, M. A. Carr, F. J. Castander, B. Chen, P. L. Colestock, A. J. Connolly, J. H. Crocker, I. Csabai, P. C. Czarapata, J. E. Davis, M. Doi, T. Dombeck, D. Eisenstein, N. Ellman, B. R. Elms, M. L. Evans, X. Fan, G. R. Federwitz, L. Fiscelli, S. Friedman, J. A. Frieman, M. Fukugita, B. Gillespie, J. E. Gunn, V. K. Gurbani, E. de Haas, M. Haldeman, F. H. Harris, J. Hayes, T. M. Heckman, G. S. Hennessy, R. B. Hindsley, S. Holm, D. J. Holmgren, C.-h. Huang, C. Hull, D. Husby, S.-I. Ichikawa, T. Ichikawa, Ž. Ivezić, S. Kent, R. S. J. Kim, E. Kinney, M. Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, J. Korienek, R. G. Kron, P. Z. Kunszt, D. Q. Lamb, B. Lee, R. F. Leger, S. Limmongkol, C. Lindenmeyer, D. C. Long, C. Loomis, J. Loveday, R. Lucinio, R. H. Lupton, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, A. Meiksin, A. Merelli, D. G. Monet, J. A. Munn, V. K. Narayanan, T. Nash, E. Neilsen, R. Neswold, H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nonino, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. L. Peterson, D. Petravick, J. R. Pier, A. Pope, R. Pordes, A. Prosapio, R. Rechenmacher, T. R. Quinn, G. T. Richards, M. W. Richmond, C. H. Rivetta, C. M. Rockosi, K. Ruthmansdorfer, D. Sandford, D. J. Schlegel, D. P. Schneider, M. Sekiguchi, G. Sergey, K. Shimasaku, W. A. Siegmund, S. Smee, J. A. Smith, S. Snedden, R. Stone, C. Stoughton, M. A. Strauss, C. Stubbs, M. SubbaRao, A. S. Szalay, I. Szapudi, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, S.-i. Wang, M. Watanabe, D. H. Weinberg, B. Yanny, N. Yasuda, and SDSS Collaboration, *The Sloan Digital Sky Survey: Technical Summary*, AJ **120** (Sept., 2000) 1579–1587, [astro-ph/0006396].

[3] C. Blake, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom, T. Davis, M. J. Drinkwater, K. Forster, D. Gilbank, M. Gladders, K. Glazebrook, B. Jelliffe, R. J. Jurek, I.-H. Li, B. Madore, D. C. Martin, K. Pimbblet, G. B. Poole, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, T. K. Wyder, and H. K. C. Yee, *The WiggleZ Dark Energy Survey: the growth rate of cosmic structure since redshift z=0.9*, MNRAS **415** (Aug., 2011) 2876–2891, [arXiv:1104.2948].

[4] M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn, D. Eisenstein, B. Flaugher, K. Honscheid, R. Kron, O. Lahav, P. McDonald, N. Roe, D. Schlegel, and representing the DESI collaboration, *The DESI Experiment, a whitepaper for Snowmass 2013*, *ArXiv e-prints* (Aug., 2013) [arXiv:1308.0847].

[5] R. Laureijs, J. Amiaux, S. Arduini, J. . Augueres, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, and et al., *Euclid Definition Study Report*, *ArXiv e-prints* (Oct., 2011) [arXiv:1110.3193].

[6] P. Monaco, T. Theuns, G. Taffoni, F. Governato, T. Quinn, and J. Stadel, *Predicting the Number, Spatial Distribution, and Merging History of Dark Matter Halos*, ApJ **564** (Jan., 2002) 8–14, [astro-ph/0109322].

[7] P. Monaco, T. Theuns, and G. Taffoni, *The pinocchio algorithm: pinpointing orbit-crossing collapsed hierarchical objects in a linear density field*, MNRAS **331** (Apr., 2002) 587–608, [astro-ph/0109323].

[8] P. Fosalba, E. Gaztañaga, F. J. Castander, and M. Manera, *The onion universe: all sky lightcone simulations in spherical shells*, MNRAS **391** (Nov., 2008) 435–446, [arXiv:0711.1540].

[9] K. Riebe, A. M. Partl, H. Enke, J. Forero-Romero, S. Gottlöber, A. Klypin, G. Lemson, F. Prada, J. R. Primack, M. Steinmetz, and V. Turchaninov, *The MultiDark Database: Release of the Bolshoi and MultiDark cosmological simulations*, *Astronomische Nachrichten* **334** (Aug., 2013) 691–708.

[10] M. Crocce, F. J. Castander, E. Gaztanaga, P. Fosalba, and J. Carretero, *The MICE Grand Challenge Lightcone Simulation II: Halo and Galaxy catalogues*, *ArXiv e-prints* (Dec., 2013) [arXiv:1312.2013].

[11] S. Tassev, M. Zaldarriaga, and D. J. Eisenstein, *Solving large scale structure in ten easy steps with COLA*, JCAP **6** (June, 2013) 36, [arXiv:1301.0322].

[12] M. White, J. L. Tinker, and C. K. McBride, *Mock galaxy catalogues using the quick particle mesh method*, MNRAS **437** (Jan., 2014) 2594–2606, [arXiv:1309.5532].

[13] P. Monaco, E. Sefusatti, S. Borgani, M. Crocce, P. Fosalba, R. K. Sheth, and T. Theuns, *An accurate tool for the fast generation of dark matter halo catalogues*, MNRAS **433** (Aug., 2013) 2389–2402, [arXiv:1305.1505].

[14] T. Hamana, S. Colombi, and Y. Suto, *Two-point correlation functions on the light cone: Testing theoretical predictions against N-body simulations*, A&A **367** (Feb., 2001) 18–26, [astro-ph/0010287].

[15] M. Sato, T. Hamana, R. Takahashi, M. Takada, N. Yoshida, T. Matsubara, and N. Sugiyama, *Simulations of Wide-Field Weak Lensing Surveys. I. Basic Statistics and Non-Gaussian Effects*, ApJ **701** (Aug., 2009) 945–954, [arXiv:0906.2237].

[16] F.-S. Kitaura, G. Yepes, and F. Prada, *Modelling baryon acoustic oscillations with perturbation theory and stochastic halo biasing*, MNRAS **439** (Mar., 2014) L21–L25, [arXiv:1307.3285].

[17] C.-H. Chuang, F.-S. Kitaura, F. Prada, C. Zhao, and G. Yepes, *EZmocks: extending the Zel'dovich approximation to generate mock galaxy catalogues with accurate clustering statistics*, *ArXiv e-prints* (Sept., 2014) [arXiv:1409.1124].

[18] B. Kalus, W. J. Percival, and L. Samushia, *Do we need model-dependent covariances when we test cosmological models with galaxy power spectra?*, arXiv:1504.0397.

[19] M. Takada and W. Hu, *Power spectrum super-sample covariance*, Phys. Rev. D **87** (June, 2013) 123504, [arXiv:1302.6994].

[20] H.-J. Seo, E. R. Siegel, D. J. Eisenstein, and M. White, *Nonlinear Structure Formation and the Acoustic Scale*, ApJ **686** (Oct., 2008) 13–24, [arXiv:0805.0117].

[21] M. Crocce and R. Scoccimarro, *Nonlinear evolution of baryon acoustic oscillations*, Phys. Rev. D **77** (Jan., 2008) 023533, [arXiv:0704.2783].

[22] R. E. Smith, R. Scoccimarro, and R. K. Sheth, *Motion of the acoustic peak in the correlation function*, Phys. Rev. D **77** (Feb., 2008) 043525, [astro-ph/0703620].

[23] N. Padmanabhan and M. White, *Calibrating the baryon oscillation ruler for matter and halos*, Phys. Rev. D **80** (Sept., 2009) 063508, [arXiv:0906.1198].

[24] H.-J. Seo, J. Eckel, D. J. Eisenstein, K. Mehta, M. Metchnik, N. Padmanabhan, P. Pinto, R. Takahashi, M. White, and X. Xu, *High-precision Predictions for the Acoustic Scale in the Nonlinear Regime*, ApJ **720** (Sept., 2010) 1650–1667, [arXiv:0910.5005].

[25] B. D. Sherwin and M. Zaldarriaga, *Shift of the baryon acoustic oscillation scale: A simple physical picture*, Phys. Rev. D **85** (May, 2012) 103523, [arXiv:1202.3998].

[26] M. Manera, R. Scoccimarro, W. J. Percival, L. Samushia, C. K. McBride, A. J. Ross, R. K. Sheth, M. White, B. A. Reid, A. G. Sánchez, R. de Putter, X. Xu, A. A. Berlind, J. Brinkmann, C. Maraston, B. Nichol, F. Montesano, N. Padmanabhan, R. A. Skibba, R. Tojeiro, and B. A. Weaver, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: a large sample of mock galaxy catalogues*, MNRAS **428** (Jan., 2013) 1036–1054, [arXiv:1203.6609].

[27] M. Manera, L. Samushia, R. Tojeiro, C. Howlett, A. J. Ross, W. J. Percival, H. Gil-Marín, J. R. Brownstein, A. Burden, and F. Montesano, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: mock galaxy catalogues for the low-redshift sample*, *ArXiv e-prints* (Jan., 2014) [arXiv:1401.4171].

[28] S. Habib, A. Pope, L. Luki'c, D. Daniel, P. Fasel, N. Desai, K. Heitmann, C.-H. Chuang, L. Ankeny, G. Mark, S. Bhattacharya, and J. Ahrens, *Hybrid petacomputing meets cosmology: The roadrunner universe project*, *Journal of Physics: Conference Series* **180** (2009), no. 1 012019.

[29] A. Pope, S. Habib, Z. Lukic, D. Daniel, P. Fasel, N. Desai, and K. Heitmann, *The accelerated universe*, *Computing in Science and Engg.* **12** (July, 2010) 17–25.

[30] S. Habib, V. Morozov, H. Finkel, A. Pope, K. Heitmann, K. Kumaran, T. Peterka, J. Insley, D. Daniel, P. Fasel, N. Frontiere, and Z. Lukic, *The Universe at Extreme Scale: Multi-Petaflop Sky Simulation on the BG/Q*, *ArXiv e-prints* (Nov., 2012) [arXiv:1211.4864].

[31] M. White, A. Pope, J. Carlson, K. Heitmann, S. Habib, P. Fasel, D. Daniel, and Z. Lukic, *Particle Mesh Simulations of the Lyα Forest and the Signature of Baryon Acoustic Oscillations in the Intergalactic Medium*, ApJ **713** (Apr., 2010) 383–393, [arXiv:0911.5341].

[32] S. Bhattacharya, S. Habib, K. Heitmann, and A. Vikhlinin, *Dark Matter Halo Profiles of Massive Clusters: Theory versus Observations*, ApJ **766** (Mar., 2013) 32, [arXiv:1112.5479].

[33] S. Habib, A. Pope, H. Finkel, N. Frontiere, K. Heitmann, D. Daniel, P. Fasel, V. Morozov, G. Zagaris, T. Peterka, V. Vishwanath, Z. Lukic, S. Sehrish, and W.-k. Liao, *HACC: Simulating Sky Surveys on State-of-the-Art Supercomputing Architectures*, *ArXiv e-prints* (Oct., 2014) [arXiv:1410.2805].

[34] R. W. Hockney and J. W. Eastwood, *Computer simulation using particles*. 1988.

[35] K. Heitmann, P. M. Ricker, M. S. Warren, and S. Habib, *Robustness of Cosmological Simulations. I. Large-Scale Structure*, ApJS **160** (Sept., 2005) 28–58, [astro-ph/0411795].

[36] K. S. Dawson, D. J. Schlegel, C. P. Ahn, S. F. Anderson, É. Aubourg, S. Bailey, R. H. Barkhouser, J. E. Bautista, A. Beifiori, A. A. Berlind, V. Bhardwaj, D. Bizyaev, C. H. Blake, M. R. Blanton, M. Blomqvist, A. S. Bolton, A. Borde, J. Bovy, W. N. Brandt, H. Brewington, J. Brinkmann, P. J. Brown, J. R. Brownstein, K. Bundy, N. G. Busca, W. Carithers, A. R. Carnero, M. A. Carr, Y. Chen, J. Comparat, N. Connolly, F. Cope, R. A. C. Croft, A. J. Cuesta, L. N. da Costa, J. R. A. Davenport, T. Delubac, R. de Putter, S. Dhital, A. Ealet, G. L. Ebelke, D. J. Eisenstein, S. Escoffier, X. Fan, N. Filiz Ak, H. Finley, A. Font-Ribera, R. Génova-Santos, J. E. Gunn, H. Guo, D. Haggard, P. B. Hall, J.-C. Hamilton, B. Harris, D. W. Harris, S. Ho, D. W. Hogg, D. Holder, K. Honscheid, J. Huehnerhoff, B. Jordan, W. P. Jordan, G. Kauffmann, E. A. Kazin, D. Kirkby, M. A. Klaene, J.-P. Kneib, J.-M. Le Goff, K.-G. Lee, D. C. Long, C. P. Loomis, B. Lundgren, R. H. Lupton, M. A. G. Maia, M. Makler, E. Malanushenko, V. Malanushenko, R. Mandelbaum, M. Manera, C. Maraston, D. Margala, K. L. Masters, C. K. McBride, P. McDonald, I. D. McGreer, R. G. McMahon, O. Mena, J. Miralda-Escudé, A. D. Montero-Dorta, F. Montesano, D. Muna, A. D. Myers, T. Naugle, R. C. Nichol, P. Noterdaeme, S. E. Nuza, M. D. Olmstead, A. Oravetz, D. J. Oravetz, R. Owen, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, J. K. Parejko, I. Pâris, W. J. Percival, I. Pérez-Fournon, I. Pérez-Ràfols, P. Petitjean, R. Pfaffenberger, J. Pforr, M. M. Pieri, F. Prada, A. M. Price-Whelan, M. J. Raddick, R. Rebolo, J. Rich, G. T. Richards, C. M. Rockosi, N. A. Roe, A. J. Ross, N. P. Ross, G. Rossi, J. A. Rubiño-Martin, L. Samushia, A. G. Sánchez, C. Sayres, S. J. Schmidt, D. P. Schneider, C. G. Scóccola, H.-J. Seo, A. Shelden, E. Sheldon, Y. Shen, Y. Shu, A. Slosar, S. A. Smee, S. A. Snedden, F. Stauffer, O. Steele, M. A. Strauss, A. Streblyanska, N. Suzuki, M. E. C. Swanson, T. Tal, M. Tanaka, D. Thomas, J. L. Tinker, R. Tojeiro, C. A. Tremonti, M. Vargas Magaña, L. Verde, M. Viel, D. A. Wake, M. Watson, B. A. Weaver, D. H. Weinberg, B. J. Weiner, A. A. West, M. White, W. M. Wood-Vasey, C. Yeche, I. Zehavi, G.-B. Zhao, and Z. Zheng, *The Baryon Oscillation Spectroscopic Survey of SDSS-III*, AJ **145** (Jan., 2013) 10, [arXiv:1208.0022].

[37] D. J. Eisenstein, D. H. Weinberg, E. Agol, H. Aihara, C. Allende Prieto, S. F. Anderson, J. A. Arns, É. Aubourg, S. Bailey, E. Balbinot, and et al., *SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems*, AJ **142** (Sept., 2011) 72, [arXiv:1101.1529].

[38] L. Anderson, É. Aubourg, S. Bailey, F. Beutler, V. Bhardwaj, M. Blanton, A. S. Bolton, J. Brinkmann, J. R. Brownstein, A. Burden, C.-H. Chuang, A. J. Cuesta, K. S. Dawson, D. J. Eisenstein, S. Escoffier, J. E. Gunn, H. Guo, S. Ho, K. Honscheid, C. Howlett, D. Kirkby, R. H. Lupton, M. Manera, C. Maraston, C. K. McBride, O. Mena, F. Montesano, R. C. Nichol, S. E. Nuza, M. D. Olmstead, N. Padmanabhan, N. Palanque-Delabrouille, J. Parejko, W. J. Percival, P. Petitjean, F. Prada, A. M. Price-Whelan, B. Reid, N. A. Roe, A. J. Ross, N. P. Ross, C. G. Sabiu, S. Saito, L. Samushia, A. G. Sánchez, D. J. Schlegel, D. P. Schneider, C. G. Scoccola, H.-J. Seo, R. A. Skibba, M. A. Strauss, M. E. C. Swanson, D. Thomas, J. L. Tinker, R. Tojeiro, M. V. Magaña, L. Verde, D. A. Wake, B. A. Weaver, D. H. Weinberg, M. White, X. Xu, C. Yèche, I. Zehavi, and G.-B. Zhao, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples*, MNRAS **441** (June, 2014) 24–62, [arXiv:1312.4877].

[39] R. Tojeiro, A. J. Ross, A. Burden, L. Samushia, M. Manera, W. J. Percival, F. Beutler, J. Brinkmann,

J. R. Brownstein, A. J. Cuesta, K. Dawson, D. J. Eisenstein, S. Ho, C. Howlett, C. K. McBride, F. Montesano, M. D. Olmstead, J. K. Parejko, B. Reid, A. G. Sánchez, D. J. Schlegel, D. P. Schneider, J. L. Tinker, M. V. Magaña, and M. White, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: galaxy clustering measurements in the low-redshift sample of Data Release 11*, MNRAS **440** (May, 2014) 2222–2237, [arXiv:1401.1768].

[40] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright, *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, ApJS **192** (Feb., 2011) 18, [arXiv:1001.4538].

[41] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al., *Planck 2015 results. XIII. Cosmological parameters*, *ArXiv e-prints* (Feb., 2015) [arXiv:1502.0158].

[42] Z. Zheng, A. A. Berlind, D. H. Weinberg, A. J. Benson, C. M. Baugh, S. Cole, R. Davé, C. S. Frenk, N. Katz, and C. G. Lacey, *Theoretical Models of the Halo Occupation Distribution: Separating Central and Satellite Galaxies*, ApJ **633** (Nov., 2005) 791–809, [astro-ph/0408564].

[43] M. White, M. Blanton, A. Bolton, D. Schlegel, J. Tinker, A. Berlind, L. da Costa, E. Kazin, Y.-T. Lin, M. Maia, C. K. McBride, N. Padmanabhan, J. Parejko, W. Percival, F. Prada, B. Ramos, E. Sheldon, F. de Simoni, R. Skibba, D. Thomas, D. Wake, I. Zehavi, Z. Zheng, R. Nichol, D. P. Schneider, M. A. Strauss, B. A. Weaver, and D. H. Weinberg, *The Clustering of Massive Galaxies at z ˜ 0.5 from the First Semester of BOSS Data*, ApJ **728** (Feb., 2011) 126, [arXiv:1010.4915].

[44] J. Kwan, S. Bhattacharya, K. Heitmann, and S. Habib, *Cosmic Emulation: The Concentration-Mass Relation for wCDM Universes*, ApJ **768** (May, 2013) 123, [arXiv:1210.1576].

[45] L. Anderson, E. Aubourg, S. Bailey, F. Beutler, A. S. Bolton, J. Brinkmann, J. R. Brownstein, C.-H. Chuang, A. J. Cuesta, K. S. Dawson, D. J. Eisenstein, K. Honscheid, E. A. Kazin, D. Kirkby, M. Manera, C. K. McBride, O. Mena, R. C. Nichol, M. D. Olmstead, N. Padmanabhan, N. Palanque-Delabrouille, W. J. Percival, F. Prada, A. J. Ross, N. P. Ross, A. G. Sanchez, L. Samushia, D. J. Schlegel, D. P. Schneider, H.-J. Seo, M. A. Strauss, D. Thomas, J. L. Tinker, R. Tojeiro, L. Verde, D. H. Weinberg, X. Xu, and C. Yeche, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Measuring D_A and H at z=0.57 from the Baryon Acoustic Peak in the Data Release 9 Spectroscopic Galaxy Sample*, *ArXiv e-prints* (Mar., 2013) [arXiv:1303.4666].

[46] D. J. Eisenstein and W. Hu, *Baryonic Features in the Matter Transfer Function*, ApJ **496** (Mar., 1998) 605–614, [astro-ph/9709112].

[47] K. T. Mehta, H.-J. Seo, J. Eckel, D. J. Eisenstein, M. Metchnik, P. Pinto, and X. Xu, *Galaxy Bias and Its Effects on the Baryon Acoustic Oscillation Measurements*, ApJ **734** (June, 2011) 94, [arXiv:1104.1178].