

# Constraining $M_\nu$ with the Bispectrum II: the Information Content of the Galaxy Bispectrum

CHANGHOON HAHN,<sup>1,2,\*</sup> FRANCISCO VILLAESCUSA-NAVARRO,<sup>3,4</sup> AND ...

<sup>1</sup>*Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley CA 94720, USA*

<sup>2</sup>*Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA*

<sup>3</sup>*Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*

<sup>4</sup>*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton NJ 08544, USA*

(Dated: DRAFT --- 129f0a1 --- 2020-06-23 --- NOT READY FOR DISTRIBUTION)

## ABSTRACT

*Keywords:* cosmology: cosmological parameters — cosmology: large-scale structure of Universe. — cosmology: theory

## 1. INTRODUCTION

**Neutrino mass background** (condensed version of paper 1) The discovery of the lower bound on the sum of neutrino masses ( $M_\nu \gtrsim 0.06$  eV), provides conclusive evidence of physics beyond the Standard Model of particle physics (??). A more precise measurement of  $M_\nu$  has the potential to distinguish between the ‘normal’ and ‘inverted’ neutrino mass hierarchy scenarios and further reveal the physics of neutrinos. However, neutrino oscillation experiments, which are insensitive to  $M_\nu$ , and other upcoming laboratory experiments (*e.g.* double beta decay and tritium beta decay) are not sufficient to distinguish between the mass hierarchies (??). Through the cosmic neutrino background, neutrinos affect the expansion history and the growth of cosmic structure. Measuring these effects provides complementary and potentially more precise measurements of  $M_\nu$ .

Neutrinos, in the early Universe, are relativistic and contribute to the energy density of radiation. Later as they become non-relativistic, they contribute to the energy density of matter. This transition affects the expansion history of the Universe and leaves imprints observable in the cosmic microwave background (CMB) anisotropy spectrum (??). Massive neutrinos also impact the growth of structure. On large scales, neutrino perturbations are indistinguishable from perturbations of cold dark matter (CDM). However, on scales smaller than their free-streaming scale, neutrinos do not contribute to the clustering and thereby reduce the amplitude of the total matter power spectrum. In addition, they also reduce the growth rate of CDM perturbations at late times. This combined suppression of the small-scale matter power spectrum leaves measurable imprints on the CMB as well as large-scale structure. For more on details the effect of neutrinos in cosmological observables, we refer readers to ?? and ?.

\* hahn.changhoon@gmail.com

TODO

TODO

transition to why LSS neutrino mass measurement is important (condensed version of paper 1)

The tightest cosmological constraints on  $M_\nu$  currently come from combining CMB data with other cosmological probes. Temperature and large angle polarization data from the *Planck* satellite places an upper bound of  $M_\nu < 0.54$  eV with 95% confidence level (?). Adding the Baryon Acoustic Oscillation (BAO) to the *Planck* likelihood breaks geometrical degeneracies (among  $M_\nu$ ,  $h$ ,  $\Omega_m$ ) and significantly tightens the upper bound to  $M_\nu < 0.16$  eV. CMB lensing further tightens the bound to  $M_\nu < 0.13$  eV, though not as significantly. Future improvements will likely continue to come from combining CMB data on large scales with clustering/lensing data on small scales and low redshifts, where the suppression of power by neutrinos is strongest (?). CMB experiments, however, measure the combined quantity  $A_s e^{-2\tau}$ , where  $\tau$  is the optical depth of reionization. Hence, improvements in neutrino mass constraints obtained from comparing the power spectrum on small and large scales will heavily rely on a better determination of  $\tau$  (???). The best constraints on  $\tau$  currently come from *Planck* —  $\tau = 0.054 \pm 0.007$ . However, most upcoming ground-based CMB experiments (*e.g.* CMB-S4) will not observe scales larger than  $\ell < 30$ , and therefore will not directly constrain  $\tau$  (?). While the upcoming CLASS experiment aims to improve  $\tau$  constraints (?), proposed future space-based experiments such as LiteBIRD<sup>1</sup> and LiteCORe<sup>2</sup>, which have the greatest potential to precisely measure  $\tau$ , have yet to be confirmed. CMB data, however, is not the only way to improve  $M_\nu$  constraints. The imprint of neutrinos on 3D clustering of galaxies can be measured to constrain  $M_\nu$  and with the sheer cosmic volumes mapped, upcoming surveys such as DESI<sup>3</sup>, PFS<sup>4</sup>, EUCLID<sup>5</sup>, and WFIRST<sup>6</sup> will be able tightly constrain  $M_\nu$  (?????).

TODO

why the bispectrum: condensed version of paper1 A major limitation of using 3D clustering is obtaining accurate theoretical predictions beyond linear scales, for bias tracers, and in redshift space. Simulations have made huge strides in accurately and efficiently modeling nonlinear structure formation with massive neutrinos (*e.g.* Brandbyge et al. 2008; ?; ?; ?; ?; ?). In conjunction, new simulation based ‘emulation’ models that exploit the accuracy of  $N$ -body simulations while minimizing the computing budget have been applied to analyze small-scale galaxy clustering with remarkable success (*e.g.* ???McClintock et al. 2018; Zhai et al. 2018; ?). Developments on these fronts have the potential to unlock the information content in nonlinear clustering to constrain  $M_\nu$ .

Various works have examined the impact of neutrino masses on nonlinear clustering of matter in real-space (*e.g.* Brandbyge et al. 2008; ?; ?; ?; Viel et al. 2010; ?; ?; ?; ?) and in redshift-space (???). Most recently, using a suite of more than 1000 simulations, ? examined the impact of  $M_\nu$  on the redshift-space matter and halo power spectrum to find that the imprint of  $M_\nu$  and  $\sigma_8$  on the power spectrum are degenerate and differ by  $< 1\%$  (see also Figure ??). The strong  $M_\nu - \sigma_8$  degeneracy poses a serious limitation on constraining  $M_\nu$  with the power spectrum. However, information in the nonlinear regime cascades from the power spectrum to higher-order statistics — *e.g.* the bispectrum.

<sup>1</sup> <http://litebird.jp/eng/>

<sup>2</sup> <http://www.core-mission.org/>

<sup>3</sup> <https://www.desi.lbl.gov/>

<sup>4</sup> <https://pfs.ipmu.jp/>

<sup>5</sup> <http://sci.esa.int/euclid/>

<sup>6</sup> <https://wfirst.gsfc.nasa.gov/>

**Table 1.** The QUIJOTE suite includes 15,000 standard  $N$ -body simulations at the fiducial cosmology to accurately estimate the covariance matrices. It also includes sets of 500 simulations at 13 other cosmologies, where only one parameter is varied from the fiducial value (underlined), to estimate derivatives of observables along the cosmological parameters.

Name	$M_\nu$	$\Omega_m$	$\Omega_b$	$h$	$n_s$	$\sigma_8$	ICs	realizations
Fiducial	0.0	0.3175	0.049	0.6711	0.9624	0.834	2LPT	15,000
Fiducial ZA	0.0	0.3175	0.049	0.6711	0.9624	0.834	Zel’dovich	500
$M_\nu^+$	<u>0.1</u> eV	0.3175	0.049	0.6711	0.9624	0.834	Zel’dovich	500
$M_\nu^{++}$	<u>0.2</u> eV	0.3175	0.049	0.6711	0.9624	0.834	Zel’dovich	500
$M_\nu^{+++}$	<u>0.4</u> eV	0.3175	0.049	0.6711	0.9624	0.834	Zel’dovich	500
$\Omega_m^+$	0.0	<u>0.3275</u>	0.049	0.6711	0.9624	0.834	2LPT	500
$\Omega_m^-$	0.0	<u>0.3075</u>	0.049	0.6711	0.9624	0.834	2LPT	500
$\Omega_b^+$	0.0	0.3175	<u>0.051</u>	0.6711	0.9624	0.834	2LPT	500
$\Omega_b^-$	0.0	0.3175	<u>0.047</u>	0.6711	0.9624	0.834	2LPT	500
$h^+$	0.0	0.3175	0.049	<u>0.6911</u>	0.9624	0.834	2LPT	500
$h^-$	0.0	0.3175	0.049	<u>0.6511</u>	0.9624	0.834	2LPT	500
$n_s^+$	0.0	0.3175	0.049	0.6711	<u>0.9824</u>	0.834	2LPT	500
$n_s^-$	0.0	0.3175	0.049	0.6711	<u>0.9424</u>	0.834	2LPT	500
$\sigma_8^+$	0.0	0.3175	0.049	0.6711	0.9624	<u>0.849</u>	2LPT	500
$\sigma_8^-$	0.0	0.3175	0.049	0.6711	0.9624	<u>0.819</u>	2LPT	500

In fact, the bispectrum has a comparable signal-to-noise ratio to the power spectrum on nonlinear scales (??). [summary of paper 1 results](#)

paragraph on others including galaxy bias for  $M_\nu$  constraints, but they’re all in the perturbation theory framework so none extend to nonlinear scales. In this paper we include galaxy bias in a simulation-based approach with emulation and LFI in mind. We use HODs, which are (a sentence on hods)

Furthermore, although  $M_\nu$  is not included in their analyses, [Sefusatti et al. \(2006\)](#) and [Yankelevich & Porciani \(2019\)](#) have shown that including the bispectrum significantly improves constraints on cosmological parameters. Including  $M_\nu$ , [Chudaykin & Ivanov \(2019\)](#) find that the bispectrum significantly improves constraints for  $M_\nu$ . Their forecasts, however, do not include the constraining power on nonlinear scales (Section ??). No work to date has quantified the total information content and constraining power of the full redshift-space bispectrum down to nonlinear scales — especially for  $M_\nu$ .

## 2. THE QUIJOTE SIMULATION SUITE

We use a subset of simulations from the QUIJOTE suite, a set of over 43,000  $N$ -body simulations that spans over 7000 cosmological models and contains, at a single redshift, over 8.5 trillion parti-

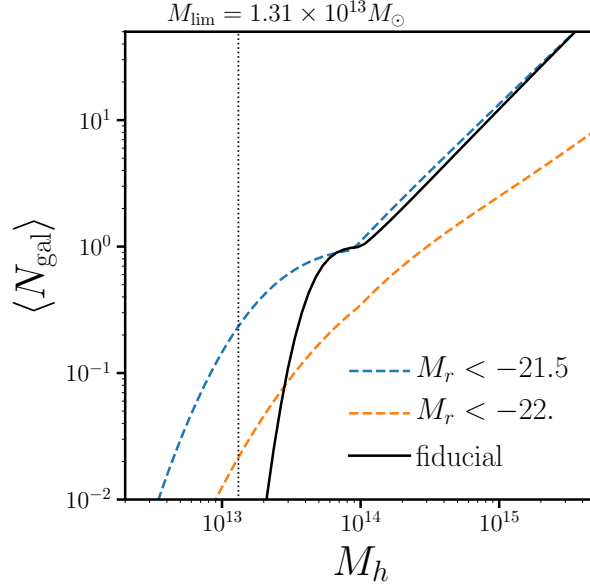
TODO

cles (Villaescusa-Navarro et al. 2019). The QUIJOTE suite was designed to quantify the information content of cosmological observables and also to train machine learning algorithms. Hence, the suite includes enough realizations to accurately estimate the covariance matrices of high-dimensional observables such as the bispectrum as well as the derivatives of these observables with respect to cosmological parameters. For the derivatives, the suite includes sets of simulations run at different cosmologies where only one parameter is varied from the fiducial cosmology:  $\Omega_m=0.3175$ ,  $\Omega_b=0.049$ ,  $h=0.6711$ ,  $n_s=0.9624$ ,  $\sigma_8=0.834$ , and  $M_\nu=0.0$  eV. Along  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$ , the fiducial cosmology is adjusted by either a small step above or below the fiducial value:  $\{\Omega_m^+, \Omega_m^-, \Omega_b^+, \Omega_b^-, h^+, h^-, n_s^+, n_s^-, \sigma_8^+, \sigma_8^-\}$ . Along  $M_\nu$ , because  $M_\nu \geq 0.0$  eV and the derivative of certain observable with respect to  $M_\nu$  is noisy, QUIJOTE includes sets of simulations for  $\{M_\nu^+, M_\nu^{++}, M_\nu^{+++}\} = \{0.1, 0.2, 0.4\}$  eV. See Table 1 for a summary of the QUIJOTE simulations used in this work.

The initial conditions for all the simulations were generated at  $z = 127$  using second-order perturbation theory for simulations with massless neutrinos ( $M_\nu = 0.0$  eV) and the Zel’dovich approximation for massive neutrinos ( $M_\nu > 0.0$  eV). The initial conditions with massive neutrinos take their scale-dependent growth factors/rates into account using the Zennaro et al. (2017) method, while for the massless neutrino case we use the traditional scale-independent rescaling. From the initial conditions, the simulations follow the gravitational evolution of  $512^3$  dark matter particles, and  $512^3$  neutrino particles for massive neutrino models, to  $z = 0$  using GADGET-III TreePM+SPH code (Springel 2005). Simulations with massive neutrinos are run using the “particle method”, where neutrinos are described as a collisionless and pressureless fluid and therefore modeled as particles, same as CDM (Brandbyge et al. 2008; Viel et al. 2010). Halos are identified using the Friends-of-Friends algorithm (FoF; Davis et al. 1985) with linking length  $b = 0.2$  on the CDM + baryon distribution. We limit the halo catalogs to halos with masses above  $M_{\text{lim}} = 3.2 \times 10^{13} h^{-1} M_\odot$ . For the fiducial cosmology, the halo catalogs have  $\sim 156,000$  halos ( $\bar{n}_h \sim 1.56 \times 10^{-4} h^3 \text{Gpc}^{-3}$ ) with  $\bar{n}P_0(k = 0.1) \sim 3.23$ . We refer readers to Villaescusa-Navarro et al. (2019) and Hahn et al. (2019) for further details on the QUIJOTE simulations.

### 3. HALO OCCUPATION DISTRIBUTION

We are interested in quantifying the information content of the galaxy bispectrum. For a perturbation theory approach, this involves incorporating a bias model for galaxies (*e.g.* Sefusatti et al. 2006; Yankelevich & Porciani 2019; Chudaykin & Ivanov 2019). Perturbation theory approaches, however, break down on small scales and limit the constraining power from nonlinear regime. Instead, in our simulation based approach we use the halo occupation distribution (HOD) framework (*e.g.* Zheng et al. 2005; Leauthaud et al. 2012; Tinker et al. 2013; Zentner et al. 2016; Vakili & Hahn 2019). HOD models statistically populate galaxies in dark matter halos by specifying the probability of a given halo hosting a certain number of galaxies. This statistical prescription for connecting galaxies to halos has been remarkably successful in reproducing the observational statistics of galaxies (*e.g.* galaxy clustering) and, as a result, is the standard approach for constructing simulated galaxy mock catalogs in galaxy clustering analyses to estimate covariance matrices and test systematic effects (*e.g.* Rodríguez-Torres et al. 2016, 2017; Beutler et al. 2017). More importantly, HOD models in simulations for build galaxy clustering emulators (see the Aemulus project McClintock et al. 2018; Zhai



**Figure 1.** Our fiducial halo occupation (black) parameterized using the standard Zheng et al. (2007) HOD model. We derive the parameter values of our fiducial HOD model (Eq. 4) by modifying the best-fit HOD parameters of the SDSS  $M_r < -21.5$  and  $< -22$ . samples from Zheng et al. (2007) to accommodate the  $M_{\text{lim}} = 3.2 \times 10^{13} h^{-1} M_{\odot}$  halo mass limit of the QUIJOTE simulations (black dashed). We include the best-fit halo occupations of the SDSS  $M_r < -21.5$  (blue dashed) and  $< -22$ . samples (orange dashed) from Zheng et al. (2007) for reference. Our fiducial HOD sample has a galaxy number density of  $\bar{n}_g \sim 1.63 \times 10^{-4} h^3/\text{Mpc}^3$ .

et al. 2018). Emulation, as we mention above, is one of the most promising approaches for modeling small scale galaxy clustering and is what we’re trying to forecast in this work.

In the simplest HOD models, the probability of a given halo hosting  $N$  galaxies of a certain class is dictated by its halo mass —  $P(N|M_h)$ . We use the standard  $P(N|M_h)$  model from Zheng et al. (2007), which has been ubiquitously used in galaxy clustering analyses (e.g. Sinha et al. 2018, many more). The model specifies the mean number of galaxies in a halo as

$$\langle N_{\text{gal}} \rangle = \langle N_{\text{cen}} \rangle + \langle N_{\text{sat}} \rangle \quad (1)$$

with mean central galaxy occupation

$$\langle N_{\text{cen}} \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M_h - \log M_{\text{min}}}{\sigma_{\log M}} \right) \right] \quad (2)$$

and mean satellite galaxy occupation

$$\langle N_{\text{sat}} \rangle = \langle N_{\text{cen}} \rangle \left( \frac{M_h - M_0}{M_1} \right)^\alpha. \quad (3)$$

The mean number of centrals in a halo transitions smoothly from 0 to 1 for halos with mass  $M_h > M_{\text{min}}$ . The width of the transition is dictated by  $\sigma_{\log M}$ , which reflects the scatter between stellar

mass/luminosity and halo mass (?). For  $M_h > M_{\min}$ ,  $\langle N_{\text{sat}} \rangle$  follows a power law with slope  $\alpha$ .  $M_0$  is the halo mass cut-off for satellite occupation and  $M_h = M_0 + M_1$  is the typical mass scale for halos to host one satellite galaxy. The numbers of centrals and satellites for each halo are drawn from Bernoulli and Poisson distribution, respectively. Central galaxies are placed at the center of the halo while position and velocity of the satellite galaxies are sampled from a Navarro et al. (1997) (NFW) profile.

The halo occupation in the Zheng et al. (2007) model depends solely on  $M_h$ . Simulations, however, find evidence that secondary halo properties such as concentration or formation history correlate with spatial distribution of halos — a phenomenon referred to as “halo assembly bias” (Sheth & Tormen 2004; Gao et al. 2005; Harker et al. 2006; Wechsler et al. 2006). A model that only depends on  $M_h$ , does not account for this halo assembly bias and may not sufficiently describe the connection between galaxies and halos. Moreover, if unaccounted for in the HOD model, and thus not marginalized over, halo assembly bias may impact the cosmological parameter constraints. Zentner et al. (2016) and Vakili & Hahn (2019) recently examined evidence for assembly bias in the observed clustering measurements of the Sloan Digital Sky Survey (SDSS) DR 7 main galaxy sample. short description of the papers and how they compare with a model with assembly bias. However, they find little evidence for assembly bias in the galaxy clustering of the SDSS  $M_r < -21.5$  and  $-22$  samples. Therefore, in this work we use the standard Zheng et al. (2007) HOD model and assume it is sufficient for modeling the galaxy–halo connection. TODO

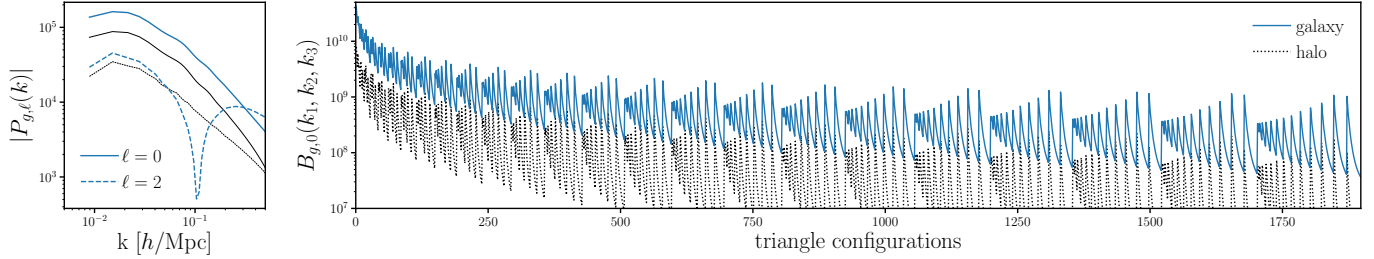
For the fiducial parameter values of the HOD model we used values motivated by best-fit HOD parameters from the literature, namely the Zheng et al. (2007) fits to the SDSS  $M_r < -21.5$  and  $-22$  samples:

$$\{M_{\min}, \sigma_{\log M}, \log M_0, \alpha, \log M_1\} = \{13.65, 0.2, 14., 1.1, 14.\}. \quad (4)$$

In Figure 1 we present the halo occupation of our fiducial HOD parameters (black). We include the best-fit halo occupations of the SDSS  $M_r < -21.5$  (blue) and  $-22$  (orange) samples from Zheng et al. (2007) for comparison. We also mark the halo mass limit,  $M_{\text{lim}}$ , of the QUIJOTE simulations (black dotted). At  $M_h \sim 10^{13} M_{\odot}$ , the best-fit halo occupations of the SDSS samples extend below  $M_{\text{lim}}$  — *i.e.* they have halos below  $M_{\text{lim}}$  that host galaxies. This prevents us from directly using the values from the literature and instead, we reduce  $\sigma_{\log M}$  to 0.2 dex. We confirm using QUIJOTE simulations with higher mass resolution ( $1024^3$  CDM particles) that  $M_{\text{lim}}$  does not impact the observables or their derivatives in our analysis for our fiducial HOD parameters.

As we mention above,  $\sigma_{\log M}$  reflects the scatter between stellar mass/luminosity and halo mass. The high  $\sigma_{\log M}$  in the  $M_r < -21.5$  and  $-22$  SDSS samples is caused by the turnover in this relation at high stellar mass/luminosity. Our fiducial halo occupation, with its lower  $\sigma_{\log M}$ , results in a galaxy sample with a tighter scatter than the samples selected based on  $M_r$  or  $M_*$  cuts, *e.g.* used in SDSS and BOSS. Hence, such a sample would require selecting based on observable galaxy properties that correlate more strongly with  $M_h$  than luminosity or  $M_*$ . Alpaslan et al. in prep. find that  $L_{\text{sat}}$  is more correlated to  $M_h$  than luminosity or  $M_*$ ; however, it has yet to be used for galaxy sample selection. Regardless, in this work we focus on quantifying the information content of the galaxy bispectrum and not on analyzing an observed galaxy sample. We therefore opt for a more conservative set of TODO





**Figure 2.** The redshift-space galaxy power spectrum multipoles ( $P_\ell^g$ ; left) and bispectrum monopole ( $B_0^g$ ; right) of our HOD galaxy sample at the fiducial HOD parameters and cosmology (blue). We include for comparison the halo  $P_{h,\ell}$  and  $B_h$  at the fiducial cosmology (black). The  $B_0^g$  above is averaged over 15,000  $N$ -body realizations and we include all 1898 triangle configurations with  $k_1, k_2, k_3 \geq k_{\max} = 0.5 h/\text{Mpc}$ . The configurations are ordered by looping through  $k_3$  in the inner most loop and  $k_1$  in the outer most loop satisfying  $k_1 \leq k_2 \leq k_3$ . The HOD sample has galaxy bias  $b_g \sim 2.55$ , significantly higher than the halo bias  $b_h \sim 1.85$  as reflected in the higher  $P_\ell^g$  and  $B_g$  amplitudes.

HOD parameters with respect to  $M_{\text{lim}}$ . For our fiducial halo occupation at the fiducial cosmology, the galaxy catalog has  $\bar{n}_g \sim 1.63 \times 10^{-4} h^3 \text{Gpc}^{-3}$  and linear bias of  $b_g \sim 2.55$ .

We construct galaxy catalogs using 22,000  $N$ -body simulations of the QUIJOTE suite: 15,000 at the fiducial cosmology and 500 at the 14 other cosmologies listed in Table 1 of ?. We use the fiducial HOD parameters for these catalogs. In addition, we construct galaxy catalogs from 500  $N$ -body simulations at the fiducial cosmology using 10 additional sets of HOD parameters. We use these catalogs to estimate the derivatives of our observables with respect to the 5 HOD parameters — a pair per parameter. For each pair we vary one HOD parameter above and below the fiducial value by step sizes:

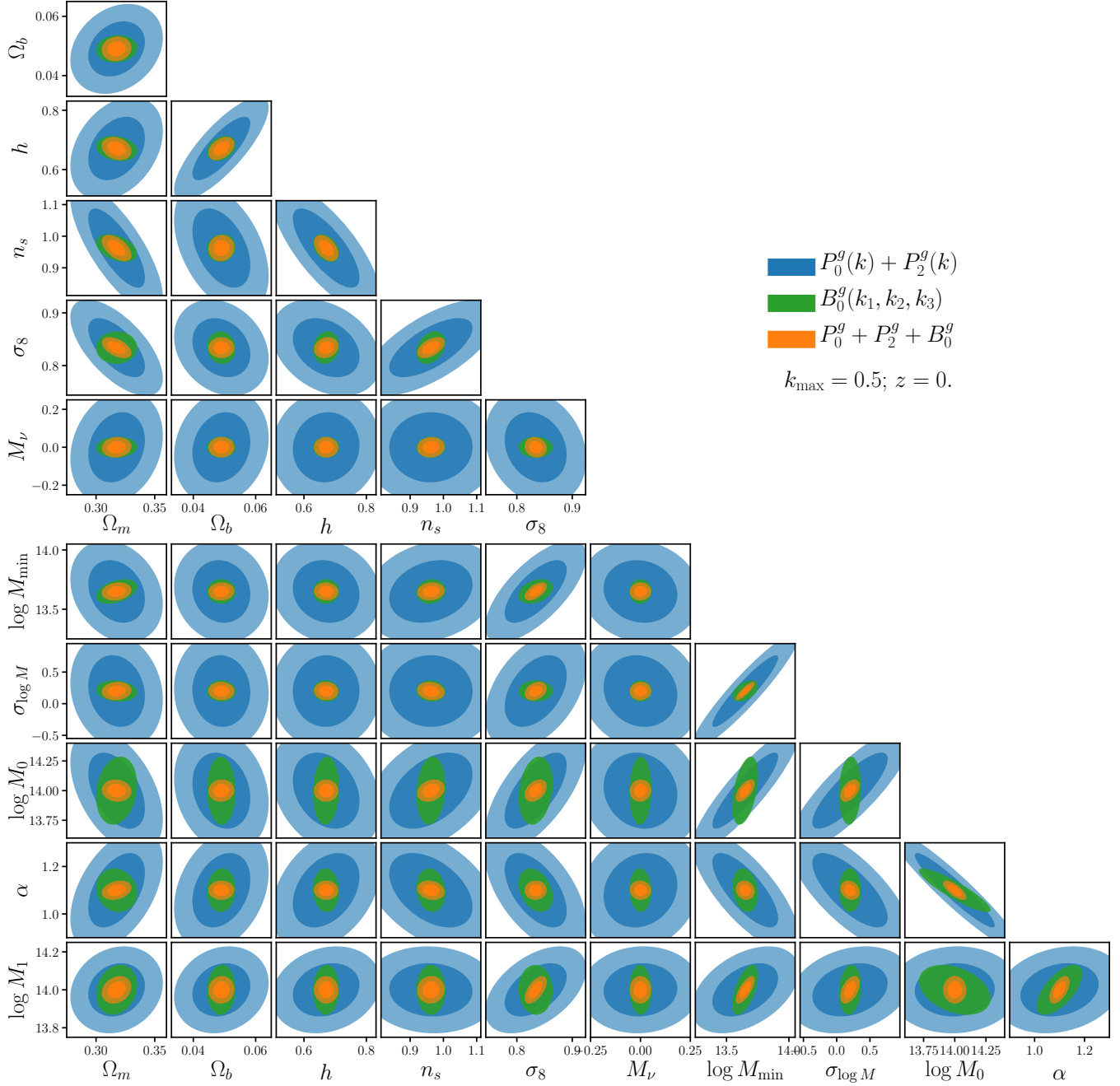
$$\{\Delta M_{\min}, \Delta \sigma_{\log M}, \Delta \log M_0, \Delta \alpha, \Delta \log M_1\} = \{0.05, 0.2, 0.2, 0.2, 0.2\}. \quad (5)$$

We generate one set of HOD realizations for the galaxy catalogs used to estimate the covariance matrices (fiducial cosmology and HOD) and apply redshift-space distortions along the  $z$ -axis. We generate 5 sets of HOD realizations for the galaxy catalogs used to estimate the derivatives and use apply redshift-space distortions along all 3 directions. *In total, we use 195,000 galaxy catalogs in our analysis.* These galaxy catalogs are publicly available at [where to access the galaxy catalogs](#).

TODO

#### 4. RESULTS

We present the Fisher matrix constraints for  $M_\nu$  and other cosmological parameters from the redshift-space galaxy  $P_\ell^g$  (blue),  $B_0^g$  (green), and combined  $P_\ell^g + B_0^g$  (orange) in Figure 3. These constraints marginalize over the Zheng et al. (2007) HOD parameters (bottom panels) and extends to  $k_{\max} = 0.5 h/\text{Mpc}$ . The contours mark the 68% and 95% confidence intervals. With the redshift-space  $P_\ell^g$  alone, we derive the following  $1\sigma$  constraints for  $\{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\}$ : **CH: 0.03443, 0.01219, 0.14069, 0.16552, 0.07701, 0.58119** With the redshift-space  $B_0^g$  alone, we get: **CH: 0.01422, 0.00371, 0.03409, 0.03543, 0.02494, 0.10599**. *The galaxy bispectrum substantially improves the constraints on all cosmological parameters over the power spectrum.*



**Figure 3.** Fisher matrix constraints for  $M_\nu$  and other cosmological parameters for the redshift-space galaxy  $P_\ell^g$  (blue),  $B_0^g$  (green), and combined  $P_\ell^g$  and  $B_0^g$  (orange) for  $k_{\max} = 0.5 h/\text{Mpc}$ . Our forecasts marginalizes over the Zheng et al. (2007) HOD parameters:  $\{M_{\min}, \sigma_{\log M}, \log M_0, \alpha \log M_1\}$  (bottom panels). The contours mark the 68% and 95% confidence intervals. The bispectrum substantially improves constraints on all of the cosmological parameters over the power spectrum.  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$  constraints improve by factors of **CH: 1.9, 2.6, 3.1, 3.6, and 2.6**, respectively. For  $M_\nu$ , the bispectrum improves  $\sigma_{M_\nu}$  from **CH: 0.2968 to 0.0572 eV** — over a factor of  $\sim 5$  improvement over the power spectrum.



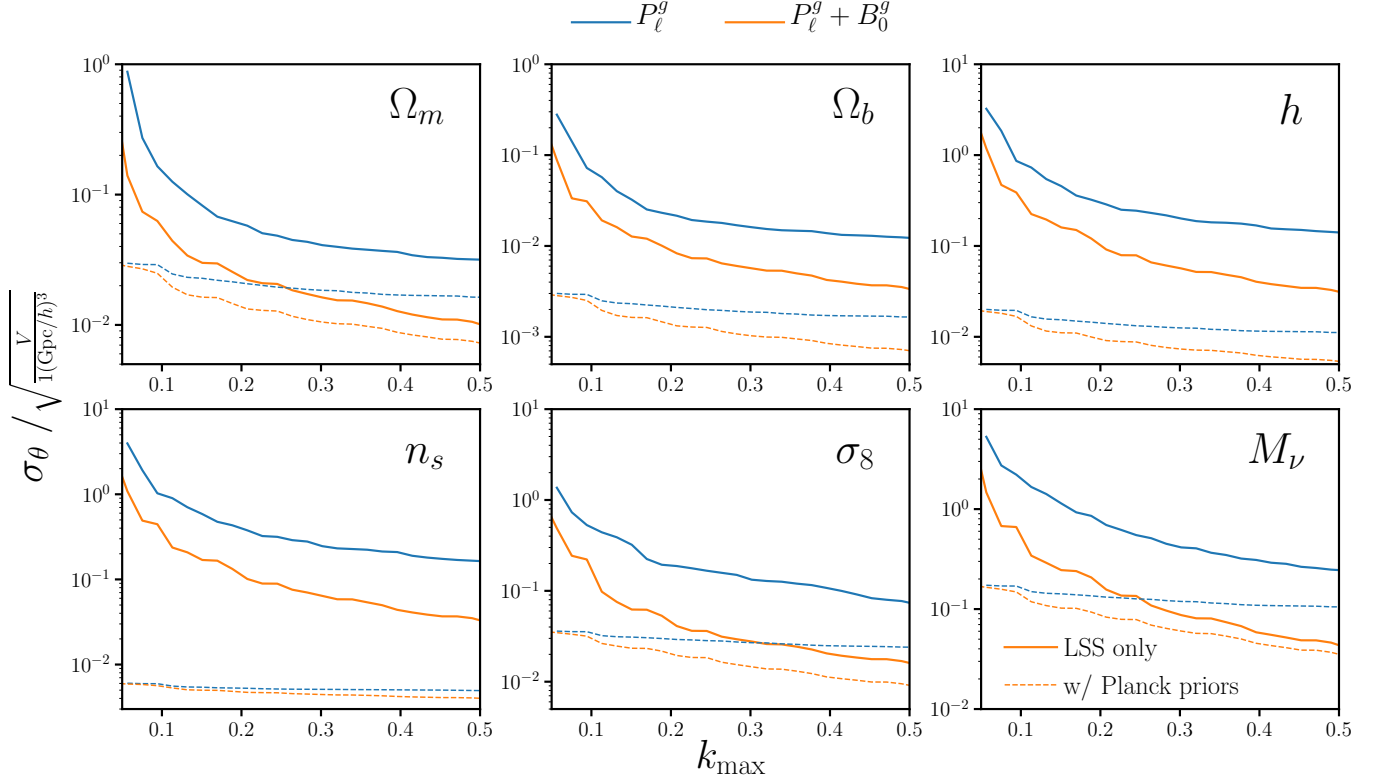
**Table 2.** Marginalized Fisher parameter constraints from the redshift-space  $P_\ell$ ,  $B_0$ , and  $P_\ell + B_0$ . We list constraints for cosmological parameters  $M_\nu$ ,  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$  as well as HOD and nuisance parameters.

	$k_{\max} = 0.2$			$k_{\max} = 0.5$		
	$P_\ell$	$B_0$	$P_\ell + B_0$	$P_\ell$	$B_0$	$P_\ell + B_0$
$M_\nu$						
$\Omega_m$						
$\Omega_b$						
$h$						
$n_s$						
$\sigma_8$						
$M_{\min}$						
$\sigma_{\log M}$						
$\log M_0$						
$\alpha$						
$\log M_1$						

With  $P_\ell^g$  and  $B_0^g$ , we derive even better constraints by breaking a number of parameter degeneracies. Among the cosmological parameters, the  $\Omega_m - \sigma_8$  degeneracy is broken and leads to significant improvements in both  $\Omega_m$  and  $\sigma_8$  constraints. Meanwhile, for the HOD parameters, degeneracies with  $\log M_0$ ,  $\alpha$ , and  $\log M_1$  are substantially reduced. Combining  $P_\ell^g$  and  $B_0^g$ , we get the following  $1\sigma$  constraints for  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$ , and  $M_\nu$ : **CH: 0.01033, 0.00336, 0.03150, 0.03361, 0.01641, and 0.10299** With  $P_\ell^g$  and  $B_0^g$  combined, we improve  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$  constraints by factors of **CH: 3.3, 3.6, 4.5, 4.9, and 4.7** and  $M_\nu$  constraint by a factor of 5.6.

In Figure 4, we present the marginalized  $1\sigma$  constraints,  $\sigma_\theta(k_{\max})$ , of the cosmological parameters  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$ , and  $M_\nu$  as a function of  $k_{\max}$  for  $P_\ell^g$  (blue) and the combined  $P_\ell^g + B_0^g$  (orange). Again, these constraints are marginalized over the Zheng et al. (2007) HOD parameters (Eq. 4). For both  $P_\ell^g$  and  $P_\ell^g + B_0^g$ , parameter constraints improve at higher  $k_{\max}$ . More importantly, *the galaxy bispectrum significantly improves constraints on all cosmological parameters throughout the  $k_{\max}$  range and not only at high  $k_{\max}$* . Even for  $k_{\max} \sim 0.2 h/\text{Mpc}$ , including  $B_0^g$  improves  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$  and  $M_\nu$  constraints by factors of **CH: 2.6, 2.4, 2.7, 3.2, 4.1, and 3.6**.

We also present  $\sigma_\theta(k_{\max})$  for  $P_\ell^g$  (blue dashed) and  $P_\ell^g + B_0^g$  (orange dashed) *with priors from Planck*. Once we include *Planck* priors,  $P_\ell^g$  constraints do not improve for  $k_{\max} \gtrsim 0.12 h/\text{Mpc}$ . However, the constraining power of  $P_\ell^g + B_0^g$  continues to increase for  $k_{\max} > 0.15 h/\text{Mpc}$ . At  $k_{\max} = 0.2 h/\text{Mpc}$ ,  $B_0^g$  improves the  $P_\ell^g + \text{Planck}$  priors constraints on  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$  and  $M_\nu$  constraint by factors of **CH: 1.5, 1.4, 1.4, 1.1, 1.3, and 1.4**. At  $k_{\max} = 0.5 h/\text{Mpc}$ ,  $B_0^g$  improves the  $P_\ell^g + \text{Planck}$  priors constraints on  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$  and  $M_\nu$  constraint by factors of **CH: 2.1, 2.1, 2.0, 1.2, 2.2, and 2.2**. Even with *Planck* priors, the galaxy bispectrum significantly improves cosmological constraints. In fact, we emphasize that the constraints in Figure 4 are for a  $1 (\text{Gpc}/h)^3$  box. Hence,



**Figure 4.** Marginalized  $1\sigma$  constraints,  $\sigma_\theta$ , of the cosmological parameters  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$ , and  $M_\nu$  as a function of  $k_{\max}$  for the redshift-space  $P_\ell^g$  (blue) and combined  $P_\ell^g + B_0^g$  (orange). Even after marginalizing over HOD parameters (Eq. 4), the galaxy bispectrum *significantly* improves cosmological parameter constraints above  $k_{\max} > 0.1 h/\text{Mpc}$ . Constraints from  $P_\ell^g$  and  $P_\ell^g + B_0^g$  improve with higher  $k_{\max}$ . Throughout  $0.2 < k_{\max} < 0.5$ , including the bispectrum improves  $\{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\}$  by **CH: X, Y %**. When we include *Planck* priors (dotted), the improvement from  $B_0^g$  is even more evident. The constraining power of  $P_\ell^g$  complete saturates for  $k_{\max} \gtrsim 0.12 h/\text{Mpc}$ . Adding  $B_0^g$  not only improves constraints, but the constraints continue to improve for higher  $k_{\max}$ . At  $k_{\max} = 0.2$  and  $0.5 h/\text{Mpc}$ , the  $P_\ell^g + B_0^g$  improves the  $M_\nu$  constraint by **CH: X, Y %** over  $P_\ell^g$ . We emphasize that the constraints above are for  $1 (h/\text{Mpc})^3$  box and thus underestimate the constraining power of upcoming galaxy clustering surveys.

they *underestimate* the constraining power contribution from galaxy clustering that we expect from upcoming galaxy redshift surveys, which will probe a much larger volume (*e.g.* DESI, Euclid). With more constraining power coming from galaxy clustering, improvements from including  $B_0^g$  to  $P_\ell^g$  and *Planck* will be larger.

In the previous paper of the series, [Hahn et al. \(2020\)](#) presents the full information content of the redshift-space halo bispectrum,  $B_0^h$ . For  $B_0^h$  to  $k_{\max} = 0.5 h/\text{Mpc}$ , [Hahn et al. \(2020\)](#) find  $1\sigma$  constraints of 0.012, 0.004, 0.04, 0.036, 0.014, and 0.057 for  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$  and  $M_\nu$ . In comparison, we find that  $B_0^g$  produces comparable constraints for  $\Omega_b$ ,  $h$ , and  $n_s$ . On the other hand,  $B_0^g$  has less constraining power than  $B_0^h$  for  $\Omega_m$ ,  $\sigma_8$ , and  $M_\nu$ . This is the same for  $k_{\max} = 0.2 h/\text{Mpc}$ . When we compare the signal-to-noise ratios (SNR) of  $B_0^g$  and  $B_0^h$ , estimated from the covariance matrix (*e.g.*

?Sefusatti et al. 2006; ?), we find lower SNR for  $B_0^g$ , consistent with the constraints. We also find that the increase in SNR with  $k_{\max}$  is lessened for  $B_0^g$ . This demonstrates that marginalizing over HOD parameters reduces some of the constraining power of  $B_0^g$ . Fingers-of-god (FoG), also contributes to this reduction. **CH: elaborate on how FoG.** Nevertheless,  $B_0^g$  significantly improves parameters constraints over  $P_\ell^g$ . In fact, marginalizing over HOD parameters and FoG reduces the constraining power of the power spectrum more so than the bispectrum. Therefore, we find larger improvements in the parameter constraints from  $B_0^g$  over  $P_\ell^g$  than from  $B_0^h$  over  $P_\ell^h$ .

A number of previous works have quantified the information content of the bispectrum: (*e.g.* Scoccimarro et al. 2004; Sefusatti et al. 2006; Sefusatti & Komatsu 2007; Song et al. 2015; Tellarini et al. 2016; Yamauchi et al. 2017; Karagiannis et al. 2018; Yankelevich & Porciani 2019; Chudaykin & Ivanov 2019; Coulton et al. 2019; Reischke et al. 2019). We, however, focus our comparison to Sefusatti et al. (2006), Yankelevich & Porciani (2019), and Chudaykin & Ivanov (2019), which provide bispectrum forecasts for full sets of cosmological parameters. Sefusatti et al. (2006) present  $\Lambda$ CDM forecasts for a joint likelihood analysis of  $B_0^g$  with  $P^g$  and WMAP. For  $k_{\max} = 0.2 h/\text{Mpc}$ , they find that including  $B_0^g$  improves constraints on  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$  by 1.6, 1.2, 1.5, 1.4, and 1.5 times from the  $P^g$  and WMAP constraints. In comparison, for  $k_{\max} = 0.2 h/\text{Mpc}$  and with *Planck* priors, we find  $B_0^g$  improves constraints by 1.5, 1.4, 1.4, 1.1, and  $1.3\times$ , which is in good agreement with Sefusatti et al. (2006). We note, however, that there are some significant differences in our analyses. First, Sefusatti et al. (2006) uses the WMAP likelihood while we use priors from *Planck*. Furthermore, in our simulation-based approach, we marginalizes over the HOD parameters. On the other hand, Sefusatti et al. (2006) use a perturbation theory approach and marginalize over the linear and quadratic bias terms ( $b_1, b_2$ ). Nevertheless, the improvement Sefusatti et al. (2006) find in parameter constraints including  $B_0^g$  is in good agreement with our results.

Next, Yankelevich & Porciani (2019) present  $\Lambda$ CDM,  $w$ CDM and  $w_0w_a$ CDM Fisher forecasts for a Euclid-like survey? over  $0.65 < z < 2.05$ . Focusing only on their  $\Lambda$ CDM forecasts, they find that for  $k_{\max} = 0.15 h/\text{Mpc}$ ,  $P^g + B_0^g$  produces constraints on  $\Omega_{\text{cdm}}$ ,  $\Omega_b$ ,  $A_s$ ,  $h$ ,  $n_s$  that are  $\sim 1.3\times$  tighter than  $P^g$  alone. With *Planck* priors, they find parameter constraints improve by a factor of  $\sim 1.1$ . In contrast, we find even at  $k_{\max} = 0.15 h/\text{Mpc}$  significantly larger improvement in the parameter constraints from including  $B_0^g$ . However, we similarly find that the improvement decreases once we include *Planck* priors (Figure 4).

Although Yankelevich & Porciani (2019) find significantly less improvement in parameter constraints from  $B_0^g$ , we emphasize that Yankelevich & Porciani (2019) present forecasts for a significantly different galaxy sample — *i.e.* the Euclid survey over  $0.65 < z < 2.05$ . For instance, their  $z = 0.7$  redshift bin has  $\bar{n}_g = 2.76 \times 10^{-3} h^3\text{Gpc}^{-3}$  and linear bias of  $b_g = 1.18$ . Meanwhile our galaxy sample is at  $z = 0$  with  $\bar{n}_g \sim 1.63 \times 10^{-4} h^3\text{Gpc}^{-3}$  and linear bias of  $b_g \sim 2.55$  (Section 3). Furthermore, while we use the HOD framework, they use a bias expansion with linear, non-linear, and tidal bias ( $b_1$ ,  $b_2$ , and  $b_{s^2}$ ). They also marginalize over 56 nuisance parameters since they jointly analyze 14  $z$  bins, each with 4 nuisance parameters. Lastly, unlike our simulation-based approach, Yankelevich & Porciani (2019) use perturbation theory models and, therefore, limit their forecast to  $k_{\max} = 0.15 h/\text{Mpc}$  due to theoretical uncertainties. Nevertheless, when they estimate the constrain-

ing power beyond  $k_{\text{max}} > 0.15 h/\text{Mpc}$  using Figure of Merit they find that the constraining power of  $B_0^g$  relative to  $P^g$  increases for higher  $k_{\text{max}}$  consistent with our results.

Finally, Chudaykin & Ivanov (2019) present  $M_\nu + \Lambda\text{CDM}$  forecasts for the power spectrum and bispectrum of a Euclid-like survey over  $0.5 < z < 2.1$ . For  $\omega_{\text{cdm}}$ ,  $\omega_b$ ,  $h$ ,  $n_s$ ,  $A_s$ , and  $M_\nu$  they find  $\sim 1.2, 1.5, 1.4, 1.3$ , and  $1.1\times$  tighter constraints from  $P_\ell^g$  and  $B_0^g$  than from  $P_\ell^g$  alone. For  $M_\nu$ , they find a factor of 1.4 improvement, from 0.038 eV to 0.028 eV. With *Planck*, they get  $\sim 2, 1.1, 2.3, 1.5, 1.1$ , and  $1.3\times$  tighter constraints for  $\omega_{\text{cdm}}$ ,  $\omega_b$ ,  $h$ ,  $n_s$ ,  $A_s$ , and  $M_\nu$  from including  $B_0^g$ . Overall, Chudaykin & Ivanov (2019) find significant improvements from including  $B_0^g$  — consistent with our results. However, they find more modest improvements than we find in Figures 3 and 4.

Again, there are significant differences between our analyses. First, like Yankelevich & Porciani (2019), Chudaykin & Ivanov (2019) present forecasts for a Euclid-like survey, which is significantly different than our galaxy sample. Their  $z = 0.6$  redshift bin, for instance, has  $\bar{n}_g = 3.83 \times 10^{-3} h^3 \text{Gpc}^{-3}$  and linear bias of  $b_g = 1.14$ . Next, they include the Alcock-Paczynski (AP) effect for  $P_\ell^g$  but not for  $B_0^g$ . They find that including the AP effect significantly improves  $P_\ell^g$  constraints (e.g. tightens  $M_\nu$  constraints by  $\sim 30\%$ ); this reduces the improvement they report from including  $B_0^g$ .

Another difference between our analyses is that although Chudaykin & Ivanov (2019) use a more accurate Markov-Chain Monte-Carlo (MCMC) approach to derive parameter constraints, they neglect the non-Gaussian contributions to both  $P_\ell^g$  and  $B_0^g$  covariance matrices and also do not include the covariance between  $P_\ell^g$  and  $B_0^g$  for the joint constraints. We find that neglecting the covariance between  $P_\ell^g$  and  $B_0^g$  overestimates constraints by CH: XXXX for our  $k_{\text{max}} = 0.2 h/\text{Mpc}$  constraints. Lastly, Chudaykin & Ivanov (2019) use a one-loop and tree-level perturbation theory to model  $P_\ell^g$  and  $B_0^g$ , respectively. Rather than imposing a  $k_{\text{max}}$  cutoff to restrict their forecasts to scales where their perturbation theory models can be trusted, they use a theoretical error covariance model approach from ?. With a tree-level  $B_0^g$  model, theoretical errors quickly dominate at  $k_{\text{max}} \gtrsim 0.1 h/\text{Mpc}$ , where one- and two-loop contribute significantly (e.g. ?). So effectively, their forecasts do not include the constraining power on those scales. In fact, if we restrict our forecast to  $k_{\text{max}} = 0.25 h/\text{Mpc}$  for  $P_\ell^g$  and  $k_{\text{max}} = 0.1 h/\text{Mpc}$  for  $B_0^g$ , our  $\Omega_{\text{m}}$ ,  $\Omega_{\text{b}}$ ,  $h$ ,  $n_s$ ,  $\sigma_8$ , and  $M_\nu$  constraints improve by 1.3, 1.2, 1.2, 1.4, 2.0, and  $1.7\times$  from including  $B_0^g$ , roughly consistent with Chudaykin & Ivanov (2019).

Among the various differences between our forecast and previous works, we emphasize that we use a simulation-based approach. This, combined with the immense number of simulations, is what allows us to go beyond previous perturbation theory approaches and accurately quantify the constraining power in the nonlinear regime. Furthermore, this allows us to derive, for the first time, the total information content of the redshift-space galaxy power spectrum and bispectrum down to nonlinear scales and demonstrate the constraining power of the galaxy bispectrum for  $M_\nu$ .

A simulation-based approach, however, has a few caveats, which we discuss below. First, our forecasts rely on the stability and convergence of the covariance matrix and numerical derivatives. For our constraints we use 195000 galaxy catalogs (Section 3): 15,000 for the covariance matrices and 180,000 for the derivatives with respect to 11 parameters. To ensure that our results are robust, we conduct the same set of convergence tests as Hahn et al. (2020). First, we test whether our results have sufficiently converged by deriving our constraints using different numbers of galaxy catalogs

to estimate the covariance matrix and derivatives:  $N_{\text{cov}}$  and  $N_{\text{deriv}}$ . For  $N_{\text{cov}}$ , we find  $< XXXX\%$  variation  $\sigma_\theta$  for  $N_{\text{cov}} > 12000$ . For  $N_{\text{deriv}}$ , we find  $< XXXX\%$  variation  $\sigma_\theta$  for  $N_{\text{cov}} > 12000$ . vary by  $< 10\%$ , we conclude that the convergence of the covariance matrix or derivatives do not significantly impact our forecast. **CH: fill this in once we have the convergence test.**

We evaluate the  $P_\ell^g$  and  $B_0^g$  derivatives with respect to  $M_\nu$  using simulations at the  $\{\theta_{\text{ZA}}, M_\nu^+, M_\nu^{++}, M_\nu^{+++}\}$  cosmologies. To test the stability of the  $M_\nu$  derivatives, we compare parameter constraints derived using two different sets of cosmologies:  $\{\theta_{\text{ZA}}, M_\nu^+\}$  and  $\{\theta_{\text{ZA}}, M_\nu^+\}$ . **CH: check this once the multiple HOD seeds finish**

Furthermore, as we note in Section 3, our fiducial HOD parameter values are chosen based on the Zheng et al. (2007) fits to the SDSS  $M_r < -21.5$  and  $-22$  samples. However, due to the halo mass limit of QUIJOTE, we use a tighter  $\sigma_{\log M}$  scatter. This means that our HOD galaxy catalogs have a different selection function than samples based on commonly used  $M_r$  or  $M_*$  cuts (e.g. SDSS or BOSS). To test the impact of this difference, **CH: in Appendix ??, we compare  $\partial P_\ell^g / \partial \sigma_{\log M}$  and  $\partial B_0^g / \partial \sigma_{\log M}$  at  $\sigma_{\log M} = 0.2$  (our fiducial value) to the derivatives evaluated at  $\sigma_{\log M} =$ , estimated using higher resolution QUIJOTE simulations. Something more after we do the comparison. CH: do we find a lot of degeneracies that are broken with  $\sigma_{\log M}$ ?**

Besides their convergence and stability, our forecasts are derived from Fisher matrices. We, therefore, assume that the posterior is approximately Gaussian. When posteriors are highly non-elliptical or asymmetric, Fisher forecasts significantly underestimate the constraints (Wolz et al. 2012). We note that in this paper we do not derive actual parameter constraints. Instead, we focus on quantifying the information content and constraining power of  $B_0^g$  relative to  $P_\ell^g$ . Hence, we do not explore beyond the Fisher forecast. In a later paper of the series, when we analyze the SDSS-III BOSS data using a simulation-based approach, we will use a robust method to sample the posterior.

Besides the caveats above, a number of extra steps and complications remain between this work and a full galaxy bispectrum analysis using a simulation based approach. For instance, we use the basic Zheng et al. (2007) HOD model, which does not include assembly bias. Zentner et al. (2016) and Vakili & Hahn (2019) find little evidence for assembly bias in the galaxy clustering of the SDSS  $M_r < -21.5$  and  $-21$  samples. ? also found that the basic HOD is sufficient to reproduce several galaxy clustering statistics (e.g. projected 2PCF, 2PCF, group multiplicity function) of high luminosity galaxies in the Illustris and EAGLE hydrodynamic simulations. While the basic HOD is likely sufficient for the forecast we present, many works have demonstrated that assembly bias impacts galaxy clustering for lower luminosity/mass samples both using observations (????Vakili & Hahn 2019) and hydrodynamic simulations (??). **CH: include ??.**

In addition to assembly bias, central and satellite velocity biases can also impact galaxy clustering (?Guo et al. 2015). Central galaxies, in both observations and simulations, are not at rest in the centers of the host halos (e.g. ???). Similarly, satellite galaxies in simulations do not have the same velocities as the underlying dark matter (e.g. ??????). This velocity bias for centrals reduces the Kaiser effect; while for satellites, it reduces the FoG effect. For the high luminosity ( $M_r < -21$  and  $< -21.5$ ) SDSS samples, Guo et al. (2015) find little satellite velocity bias. While they find some central velocity bias, their constraints are based on galaxy clustering on very small scales



( $\sim 0.1 - 25h^{-1}\text{Mpc}$ ). More recently, ? found that removing central and satellite velocity biases for the Illustris and EAGLE simulations had little impact on various clustering measurements of their high luminosity sample. Although assembly bias and velocity bias do not likely impact the forecasts we present, for lower luminosity/mass galaxy samples and for higher precision measurements of observations they must be included. Therefore, when we analyze observations with a simulation-based approach later in the series, we will use a decorated HOD framework (such as *e.g.* Vakili & Hahn 2019; ?, CH: others) that includes both assembly bias and velocity biases.

Our analysis also does not include baryonic effects. Although they have been typically neglected in galaxy clustering analyses, baryonic effects, such as feedback from active galactic nuclei (AGN), can impact the matter distribution at cosmological distances (*e.g.* ?????). For AGN feedback in particular, various works have an impact on the matter power spectrum (*e.g.* ???????). Although there is no consensus on the magnitude of the effect, ultimately, a more effective AGN feedback increases the impact on the matter clustering (?). In state-of-the-art hydrodynamical simulations (TNG, EAGLE, and BAHAMAS), ? find  $\lesssim 1\%$  impact on the matter power spectrum at  $k \lesssim 0.5 h/\text{Mpc}$ . For the matter bispectrum, ? find that the effect of baryons is peaked at  $k = 3 h/\text{Mpc}$  and, similarly, a  $\lesssim 1\%$  effect at  $k \lesssim 0.5 h/\text{Mpc}$ . Despite the growing evidence of baryon impacting the matter clustering, the effect is mainly seen on scales smaller than what is probed by galaxy clustering analyses with spectroscopic redshift surveys. We, therefore, do not include baryonic effects in our forecasts and do not consider it further in the series.

alcock pacinzsky effect

beyond the monopole Yankelevich & Porciani (2019) 50% more information beyond monopole

we take broad triangle bins, Yankelevich & Porciani (2019) finds higher constraining power for narrower bins.  $\Delta k = 3k_f$  has  $\sim 10\%$  less constrainig power according to their FoM estimate. the need for data compression

no geometry obviously. repeat of how realistic geometry lowers the signal to noise.

**CH: restatement and discussson of key findings** The improvements of  $B_0^g$  come from breaking parameter degeneracies. Most notably,  $\Omega_m$  and  $\sigma_8$  degeneracies are broken by  $P_\ell^g + B_0^g$ , which improves  $\Omega_m - \sigma_8$ ,  $\Omega_m - M_\nu$ ,  $\sigma_8 - M_\nu$  that we care about.

A key part of the improvement from  $B_0^g$  comes from the degeneracies that are broken among HOD parameters. Furthermore, including  $B_0^g$  allows us to break a number of degeneracies among HOD parameters. This is interesting because there are many questions regarding HOD parameters that are still up in the air. For instance, the impact and importance of assembly bias, which can impact cosmological analyses. Moreover, tighter constraints on the halo occupation in general allows us to better understand the galaxy-halo connection, which will improve our understanding of galaxy formation and evolution. While we consider a relatively simplistic halo occupation model, our results clearly demonstrate the advantages of higher-order statistic bispectrum over the standard two-point statistics.

what's next in the paper series

- data compression
- outline the overall forward modeling approach that will culminate in an analysis of BOSS.



- paragraph outline the obvious extensions to DESI and PFS with rough numbers

## 5. SUMMARY

## ACKNOWLEDGEMENTS

It's a pleasure to thank Mehmet Alpaslan, Arka Banerjee, Joseph DeRose, Elena Massara, Jeremy L. Tinker, Roman Scoccimarro, Digvijay Wadekar, Risa Wechsler ... for valuable discussions and comments.

## APPENDIX

## REFERENCES

- Beltz-Mohrmann, G. D., Berlind, A. A., & Szewciw, A. O. 2020, [Monthly Notices of the Royal Astronomical Society](#), 491, 5771
- Beutler, F., Seo, H.-J., Saito, S., et al. 2017, [Monthly Notices of the Royal Astronomical Society](#), 466, 2242
- Brandbyge, J., Hannestad, S., Haugbølle, T., & Thomsen, B. 2008, [Journal of Cosmology and Astro-Particle Physics](#), 08, 020
- Chaves-Montero, J., Angulo, R. E., Schaye, J., et al. 2016, [Monthly Notices of the Royal Astronomical Society](#), 460, 3100
- Chudaykin, A., & Ivanov, M. M. 2019, [arXiv:1907.06666 \[astro-ph, physics:hep-ph\]](#), [arXiv:1907.06666 \[astro-ph, physics:hep-ph\]](#)
- Coulton, W. R., Liu, J., Madhavacheril, M. S., Böhm, V., & Spergel, D. N. 2019, [Journal of Cosmology and Astro-Particle Physics](#), 2019, 043
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, [The Astrophysical Journal](#), 292, 371
- Gao, L., Springel, V., & White, S. D. M. 2005, [Monthly Notices of the Royal Astronomical Society](#), 363, L66
- Guo, H., Zheng, Z., Zehavi, I., et al. 2015a, [Monthly Notices of the Royal Astronomical Society](#), 453, 4368
- . 2015b, [Monthly Notices of the Royal Astronomical Society](#), 446, 578
- Hahn, C., Villaescusa-Navarro, F., Castorina, E., & Scoccimarro, R. 2020, [Journal of Cosmology and Astroparticle Physics](#), 03, 040
- Harker, G., Cole, S., Helly, J., Frenk, C., & Jenkins, A. 2006, [Monthly Notices of the Royal Astronomical Society](#), 367, 1039
- Karagiannis, D., Lazanu, A., Liguori, M., et al. 2018, [Monthly Notices of the Royal Astronomical Society](#), 478, 1341
- Lazanu, A., & Liguori, M. 2018, [Journal of Cosmology and Astro-Particle Physics](#), 2018, 055
- Leauthaud, A., Tinker, J., Bundy, K., et al. 2012, [The Astrophysical Journal](#), 744, 159
- McClintock, T., Rozo, E., Becker, M. R., et al. 2018, [arXiv:1804.05866 \[astro-ph\]](#), [arXiv:1804.05866 \[astro-ph\]](#)
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, [The Astrophysical Journal](#), 490, 493
- Reischke, R., Desjacques, V., & Zaroubi, S. 2019, [arXiv:1909.03761 \[astro-ph\]](#), [arXiv:1909.03761 \[astro-ph\]](#)
- Rodríguez-Torres, S. A., Chuang, C.-H., Prada, F., et al. 2016, [Monthly Notices of the Royal Astronomical Society](#), 460, 1173
- Rodríguez-Torres, S. A., Comparat, J., Prada, F., et al. 2017, [Monthly Notices of the Royal Astronomical Society](#), 468, 728
- Scoccimarro, R., Sefusatti, E., & Zaldarriaga, M. 2004, [Physical Review D](#), 69, 103513
- Sefusatti, E., Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, [Physical Review D](#), 74, [arXiv:astro-ph/0604505](#)
- Sefusatti, E., & Komatsu, E. 2007, [Physical Review D](#), 76, 083004
- Sheth, R. K., & Tormen, G. 2004, [Monthly Notices of the Royal Astronomical Society](#), 350, 1385

- Sinha, M., Berlind, A. A., McBride, C. K., et al. 2018, [Monthly Notices of the Royal Astronomical Society](#), 478, 1042
- Song, Y.-S., Taruya, A., & Oka, A. 2015, [Journal of Cosmology and Astro-Particle Physics](#), 2015, 007
- Springel, V. 2005, [Monthly Notices of the Royal Astronomical Society](#), 364, 1105
- Tellarini, M., Ross, A. J., Tasinato, G., & Wands, D. 2016, [Journal of Cosmology and Astro-Particle Physics](#), 2016, 014
- Tinker, J. L., Leauthaud, A., Bundy, K., et al. 2013, [The Astrophysical Journal](#), 778, 93
- Vakili, M., & Hahn, C. 2019, [The Astrophysical Journal](#), 872, 115
- Viel, M., Haehnelt, M. G., & Springel, V. 2010, [Journal of Cosmology and Astro-Particle Physics](#), 06, 015
- Villaescusa-Navarro, F., Hahn, C., Massara, E., et al. 2019, [arXiv:1909.05273 \[astro-ph\]](#), [arXiv:1909.05273 \[astro-ph\]](#)
- Wechsler, R. H., Zentner, A. R., Bullock, J. S., Kravtsov, A. V., & Allgood, B. 2006, [The Astrophysical Journal](#), 652, 71
- Wolz, L., Kilbinger, M., Weller, J., & Giannantonio, T. 2012, [Journal of Cosmology and Astroparticle Physics](#), 2012, 009
- Yamauchi, D., Yokoyama, S., & Takahashi, K. 2017, [Physical Review D](#), 95, 063530
- Yankelevich, V., & Porciani, C. 2019, [Monthly Notices of the Royal Astronomical Society](#), 483, 2078
- Zennaro, M., Bel, J., Villaescusa-Navarro, F., et al. 2017, [Monthly Notices of the Royal Astronomical Society](#), 466, 3244
- Zentner, A. R., Hearin, A., van den Bosch, F. C., Lange, J. U., & Villarreal, A. 2016, [arXiv:1606.07817 \[astro-ph\]](#), [arXiv:1606.07817 \[astro-ph\]](#)
- Zhai, Z., Tinker, J. L., Becker, M. R., et al. 2019, [The Astrophysical Journal](#), 874, 95
- Zheng, Z., Coil, A. L., & Zehavi, I. 2007, [The Astrophysical Journal](#), 667, 760
- Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, [The Astrophysical Journal](#), 633, 791