Draft version June 18, 2020
Typeset using LaTeX **preprint** style in AASTeX62

Constraining M_ν with the Bispectrum II: the Information Content of the Galaxy Bispectrum

ChangHoon Hahn,[1,2,*] Francisco Villaescusa-Navarro,[3,4] and ...

[1] *Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley CA 94720, USA*
[2] *Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA*
[3] *Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*
[4] *Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton NJ 08544, USA*

(Dated: `DRAFT --- 831204e --- 2020-06-17 --- NOT READY FOR DISTRIBUTION`)

ABSTRACT

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

1. INTRODUCTION

TODO

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_ν 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_ν, and other upcoming laboratory experiments (*e.g.* double beta decay and tritium beta decay) are not sufficient to distinguish between the mass hierarches (**??**). 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_ν.

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 **?**.

TODO

* hahn.changhoon@gmail.com

transition to why LSS neutrino mass measurement is important (condensed version of paper 1) The tightest cosmological constraints on M_ν 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_ν, h, $\Omega_{\rm m}$) 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 τ 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 τ (**???**). The best constraints on τ 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 τ (**?**). While the upcoming CLASS experiment aims to improve τ constraints (**?**), proposed future space-based experiments such as LiteBIRD[1] and LiteCOrE[2], which have the greatest potential to precisely measure τ, have yet to be confirmed. CMB data, however, is not the only way to improve M_ν constraints. The imprint of neutrinos on 3D clustering of galaxies can be measured to constrain M_ν and with the sheer cosmic volumes mapped, upcoming surveys such as DESI[3], PFS[4], EUCLID[5], and WFIRST[6] will be able tightly constrain M_ν (**?????**).

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_ν.

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_ν on the redshift-space matter and halo power spectrum to find that the imprint of M_ν and σ_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_ν with the power spectrum. However, information in the nonlinear regime cascades from the power spectrum to higher-order statistics — *e.g.* the bispectrum.

[1] http://litebird.jp/eng/
[2] http://www.core-mission.org/
[3] https://www.desi.lbl.gov/
[4] https://pfs.ipmu.jp/
[5] http://sci.esa.int/euclid/
[6] 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_ν	Ω_m	Ω_b	h	n_s	σ_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_ν^+	0.1 eV	0.3175	0.049	0.6711	0.9624	0.834	Zel'dovich	500
M_ν^{++}	0.2 eV	0.3175	0.049	0.6711	0.9624	0.834	Zel'dovich	500
M_ν^{+++}	0.4 eV	0.3175	0.049	0.6711	0.9624	0.834	Zel'dovich	500
Ω_m^+	0.0	0.3275	0.049	0.6711	0.9624	0.834	2LPT	500
Ω_m^-	0.0	0.3075	0.049	0.6711	0.9624	0.834	2LPT	500
Ω_b^+	0.0	0.3175	0.051	0.6711	0.9624	0.834	2LPT	500
Ω_b^-	0.0	0.3175	0.047	0.6711	0.9624	0.834	2LPT	500
h^+	0.0	0.3175	0.049	0.6911	0.9624	0.834	2LPT	500
h^-	0.0	0.3175	0.049	0.6511	0.9624	0.834	2LPT	500
n_s^+	0.0	0.3175	0.049	0.6711	0.9824	0.834	2LPT	500
n_s^-	0.0	0.3175	0.049	0.6711	0.9424	0.834	2LPT	500
σ_8^+	0.0	0.3175	0.049	0.6711	0.9624	0.849	2LPT	500
σ_8^-	0.0	0.3175	0.049	0.6711	0.9624	0.819	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

TODO

paragraph on others including galaxy bias for M_ν 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_ν 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_ν, Chudaykin & Ivanov (2019) find that the bispectrum significantly improves constraints for M_ν. 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_ν.

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-

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: $\Omega_{\rm m}$=0.3175, $\Omega_{\rm b}$=0.049, h=0.6711, n_s=0.9624, σ_8=0.834, and M_ν=0.0 eV. Along $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, and σ_8, the fiducial cosmology is adjusted by either a small step above or below the fiducial value: $\{\Omega_{\rm m}^+, \Omega_{\rm m}^-, \Omega_{\rm b}^+, \Omega_{\rm b}^-, h^+, h^-, n_s^+, n_s^-, \sigma_8^+, \sigma_8^-\}$. Along M_ν, because $M_\nu \geq 0.0$ eV and the derivative of certain observable with respect to M_ν 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 $M_{\rm lim} = 3.2 \times 10^{13} h^{-1} M_\odot$. For the fiducial cosmology, the halo catalogs have $\sim$156,000 halos ($\bar{n}_h \sim 1.56 \times 10^{-4}\ h^3{\rm 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

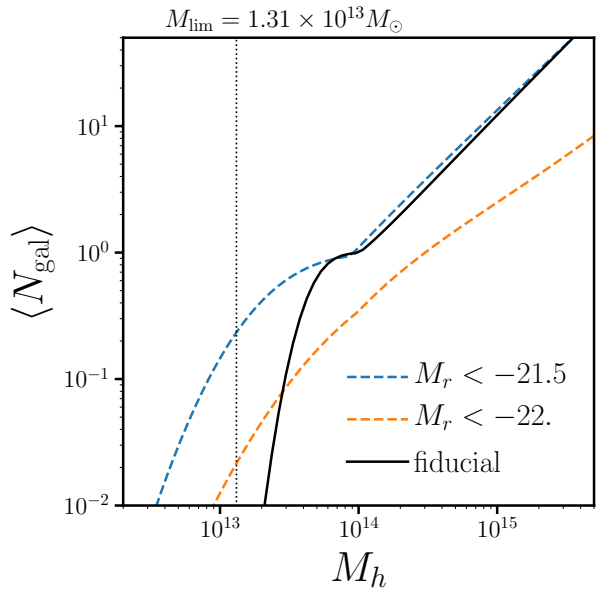

Figure 1. Our fiducial halo occupation (black) parameterized using the standard Zheng et al. (2007) HOD model. We derive the parameter values of our fiducial HOD model (Eq. 4) by modifying the best-fit HOD parameters of the SDSS $M_r < -21.5$ and $< -22.$ samples from Zheng et al. (2007) to accommodate the $M_{\rm lim} = 3.2 \times 10^{13} h^{-1} M_\odot$ 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 $\overline{n}_g \sim 1.63 \times 10^{-4}\ h^3/{\rm Mpc}^3$.

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 TODO

$$\langle N_{\rm gal} \rangle = \langle N_{\rm cen} \rangle + \langle N_{\rm sat} \rangle \tag{1}$$

with mean central galaxy occupation

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

and mean satellite galaxy occupation

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

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

mass/luminosity and halo mass (**?**). For $M_h > M_{\rm min}$, $\langle N_{\rm sat}\rangle$ follows a power law with slope α. 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 soley 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 sufficient 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 -21 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:

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

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, $M_{\rm lim}$, 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 $M_{\rm lim}$ — *i.e.* they have halos below $M_{\rm lim}$ 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 $M_{\rm lim}$ 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 strong with M_h than luminosity or M_*. Alpaslan et al. in prep. find that $L_{\rm sat}$ is more correlated to M_h than luminosity or M_*; however, it has yet to be used for galaxy sample selection. TODO 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

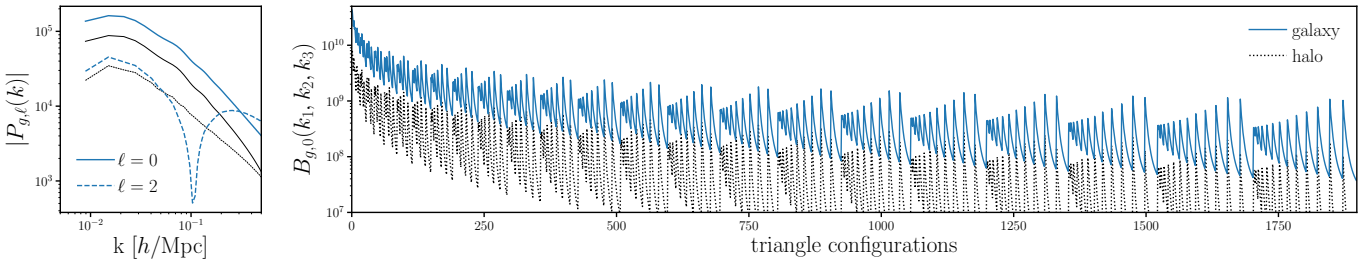

Figure 2. The redshift-space galaxy power spectrum multipoles (P_ℓ^g; left) and bispectrum monopole (B_0^g; right) of our HOD galaxy sample at the fidicual HOD parameters and cosmology (blue). We include for comparison the halo $P_{h,\ell}$ and B_h at the fiducial cosmology (black). The B_0^g above is averaged over 15,000 N-body realizations and we include all 1898 triangle configurations with $k_1, k_2, k_3 \geq k_{\rm max} = 0.5\,h/{\rm Mpc}$. The configurations are ordered by looping through k_3 in the inner most loop and k_1 in the outer most loop satisfying $k_1 \leq k_2 \leq k3$. 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_ℓ^g and Bg amplitudes.

HOD parameters with respect to $M_{\rm lim}$. For our fiducial halo occupation at the fiducial cosmology, the galaxy catalog has $\bar{n}_g \sim 1.63 \times 10^{-4}\ h^3{\rm Gpc}^{-3}$ 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 paramter above and below the fiducial value by step sizes:

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

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_ν and other cosmological parameters from the redshift-space galaxy P_ℓ^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 $k_{\rm max} = 0.5\,h/{\rm Mpc}$. The contours mark the 68% and 95% confidence intervals. With the redshift-space P_ℓ^g alone, we derive the following 1σ constraints for $\{\Omega_{\rm m}, \Omega_{\rm b}, h, n_s, \sigma_8, M_\nu\}$: **CH:** numbers. With the redshift-space B_0^g alone, we get: **CH:** numbers. *The galaxy bispectrum substantially improves the constraints on all cosmological parameters over the power spectrum.*

With P_ℓ^g and B_0^g, we derive even better constraints by breaking a number of parameter degeneracies. Among the cosmological parameters, the $\Omega_{\rm m} - \sigma_8$ degeneracy is broken and leads to significant

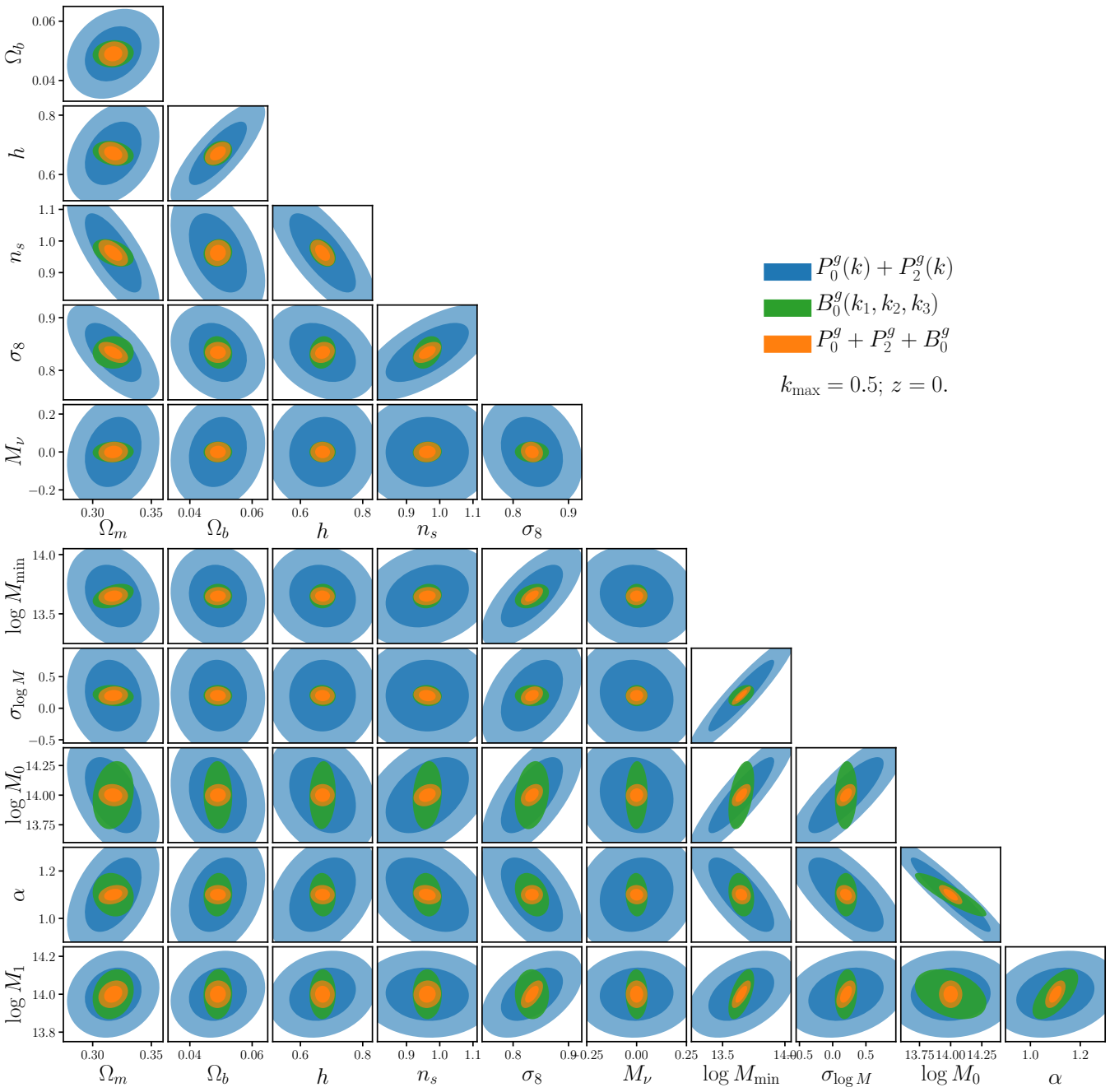

Figure 3. Fisher matrix constraints for M_ν and other cosmological parameters for the redshift-space galaxy P_ℓ^g (blue), B_0^g (green), and combined P_ℓ^g and B_0^g (orange) for $k_{\rm max} = 0.5\,h/{\rm Mpc}$. Our forecasts marginalizes over the Zheng et al. (2007) HOD parameters: $\{M_{\rm min}, \sigma_{\log M}, \log M_0, \alpha \log M_1\}$ (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. $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, and σ_8 constraints improve by factors of **CH:** 1.9, 2.6, 3.1, 3.6, and 2.6, respectively. For M_ν, the bispectrum improves σ_{M_ν} from **CH:** 0.2968 to 0.0572 eV — over a factor of ~ 5 improvement over the power spectrum.

Table 2. Marginalized Fisher parameter constraints from the redshift-space P_ℓ, B_0, and $P_\ell + B_0$. We list constraints for cosmological parameters M_ν, Ω_m, Ω_b, h, n_s, and σ_8 as well as HOD and nuisance parameters.

	P_ℓ	B_0	$P_\ell + B_0$
M_ν			
$\Omega_{\rm m}$			
$\Omega_{\rm b}$			
h			
n_s			
σ_8			
$M_{\rm min}$			
$\sigma_{\log M}$			
$\log M_0$			
α			
$\log M_1$			

improvements in both $\Omega_{\rm m}$ and σ_8 constraints. Meanwhile, for the HOD parameters, degeneracies with $\log M_0$, α, and $\log M_1$ are substantially reduced. Combining P_ℓ^g and B_0^g, we get the following 1σ constraints for $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, σ_8, and M_ν: **CH:** numbers numbers. *By combining P_ℓ^g and B_0^g, we improve $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, and σ_8 constraints by factors of* **CH:** *1.9, 2.6, 3.1, 3.6, and 2.6 and M_ν constraint by a factor of* **CH:** *number.*

In Figure 4, we present the marginalized 1σ constraints, $\sigma_\theta(k_{\rm max})$, of the cosmological parameters $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, σ_8, and M_ν as a function of $k_{\rm max}$ for P_ℓ^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_ℓ^g and $P_\ell^g + B_0^g$, parameter constraints improve at higher $k_{\rm max}$. More importantly, *the galaxy bispectrum significantly improves constraints on all cosmological parameters throughout our $k_{\rm max}$ range.* For $k_{\rm max} \geq 0.2\,h/{\rm Mpc}$, the B_0^g consistently improves constraints by **CH:** numbers.

We also present $\sigma_\theta(k_{\rm max})$ for P_ℓ^g (blue dashed) and $P_\ell^g + B_0^g$ (orange dashed) *with priors from Planck*. Once we include Planck priors, P_ℓ^g constraints do not improve for $k_{\rm max} \gtrsim 0.12\,h/{\rm Mpc}$. However, the constraining power of $P_\ell^g + B_0^g$ continues to increase for $k_{\rm max} > 0.15\,h/{\rm Mpc}$. At $k_{\rm max} = 0.2$ and $0.5\,h/{\rm Mpc}$, the $P_\ell^g + B_0^g$ improves $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, σ_8 and M_ν constraint by **CH:** X, Y %. 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 $({\rm Gpc}/h)^3$ box. Hence, 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_ℓ^g and planck would likely be larger.

signal-to-noise is lower for the galaxy power and bispectrums than the halo power and bispectra — fingers of god. **CH:** figure?

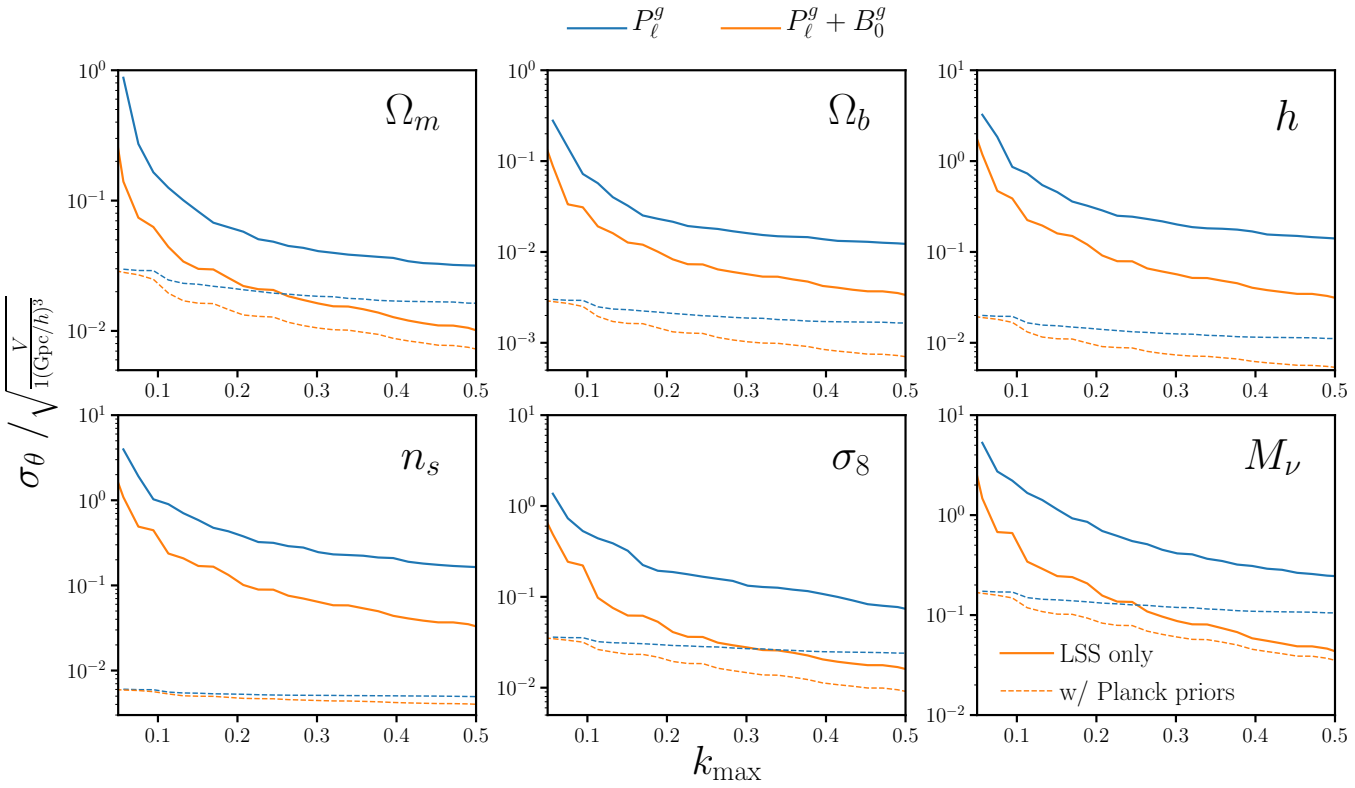

Figure 4. Marginalized 1σ constraints, σ_θ, of the cosmological parameters $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, σ_8, and M_ν as a function of $k_{\rm max}$ for the redshift-space P_ℓ^g (blue) and combined $P_\ell^g + B_0^g$ (orange). Even after marginalizing over HOD parameters (Eq. 4), the galaxy bispectrum *significantly* improves cosmological parameter constraints above $k_{\rm max} > 0.1\,h/{\rm Mpc}$. Constraints from P_ℓ^g and $P_\ell^g + B_0^g$ improve with higher $k_{\rm max}$. *Throughout* $0.2 < k_{\rm max} < 0.5$*, including the bispectrum improves* $\{\Omega_{\rm m}, \Omega_{\rm b}, h, n_s, \sigma_8, M_\nu\}$ *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_ℓ^g complete saturates for $k_{\rm max} \gtrsim 0.12\,h/{\rm Mpc}$. Adding B_0^g not only improves constraints, but the constraints continue to improve for higher $k_{\rm max}$. At $k_{\rm max} = 0.2$ and $0.5\,h/{\rm Mpc}$, the $P_\ell^g + B_0^g$ improves the M_ν constraint by **CH:** X, Y % over P_ℓ^g. We emphasize that the constraints above are for 1 $({\rm Gpc}/h)^3$ box and thus underestimate the constraining power of upcoming galaxy clustering surveys.

comparison to the literature.

- Compare to PT papers with bias models.
- a discussion on the impact of off diagonal terms of the covariance matrix.

We derive, for the first time, the total information cotent of the redshift-space galaxy power spectrum and bispectrum. Our simulation based approach, allows us to go beyond previous perturbation theory approaches and accurately quantify the constraining power in the nonlinear regime. They, however, rely on the stability and convergence of our 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 **?**.

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

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

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_ℓ^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.

CH: anything else that could be wrong that we've checked?

CH: interesting findings

- re-emphasize the impact of bispectrum in breaking HOD parameter degeneracies.
- Most noteably, $\Omega_{\rm m}$ and σ_8 degeneracies are broken by $P_\ell^g + B_0^g$, which improves $\Omega_{\rm m} - \sigma_8$, $\Omega_{\rm m} - M_\nu$, $\sigma_8 - M_\nu$ that we care about.
- HOD parameter degeneracies are broken. Some of them drastically. **CH:** what are the implications of this?
- **CH:** do we want H0 forecast and discussion on H0 tension.

CH: what's missing in the analysis: all the graints of salt/ future works

- how do the unrealistic hod parameters we use map onto real galaxy samples. We tried our best to get the SDSS-like sample. The main "unrealistic" parameter choice is $\sigma_{\log M}$ — use the high resolution simulations to compare the derivative w.r.t sigma logM and make some statements regarding this. **CH:** appendix?
- we use a simplistic HOD model which doesn't take into account assembly bias. Overview the impact of assembly bias in previous galaxy clustering results and discuss how it may analogously apply to our constraints.
- including some extra nuisance parameters doesn't do anything for cosmological constraints. We tried b1, Asn, Bsn. **CH:** try b2, gamma2
- alcock pacinzsky effect

- baryonic effects are not considered and some papers have looked at this recently.
- no geometry obviously. repeat of how realistic geometry lowers the signal to noise.

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

Beutler, F., Seo, H.-J., Saito, S., et al. 2017, Monthly Notices of the Royal Astronomical Society, 466, 2242

Brandbyge, J., Hannestad, S., Haugbø lle, T., & Thomsen, B. 2008, Journal of Cosmology and Astro-Particle Physics, 08, 020

Chudaykin, A., & Ivanov, M. M. 2019, arXiv:1907.06666 [astro-ph, physics:hep-ph], arXiv:1907.06666 [astro-ph, physics:hep-ph]

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

Harker, G., Cole, S., Helly, J., Frenk, C., & Jenkins, A. 2006, Monthly Notices of the Royal Astronomical Society, 367, 1039

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

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

Sefusatti, E., Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, Physical Review D, 74, arXiv:astro-ph/0604505

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

Springel, V. 2005, Monthly Notices of the Royal Astronomical Society, 364, 1105

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

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. 2018, arXiv:1804.05867 [astro-ph], arXiv:1804.05867 [astro-ph]

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