

Constraining M_ν with the Bispectrum I: Breaking Parameter Degeneracies

CHANGHOON HAHN,^{1,2,*} FRANCISCO VILLAESCUSA-NAVARRO,³ EMANUELE CASTORINA,^{2,1} AND
ROMAN SCOCCIMARRO⁴

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*Center for Cosmology and Particle Physics, Department of Physics, New York University, NY 10003, New York, USA*

(Dated: DRAFT --- 112c8e2 --- 2019-09-20 --- NOT READY FOR DISTRIBUTION)

ABSTRACT

Massive neutrinos suppress the growth of structure below their free-streaming scale and leave an imprint on large-scale structure. Measuring this imprint allows us to constrain the sum of neutrino masses, M_ν , a key parameter in particle physics beyond the Standard Model. However, degeneracies among cosmological parameters, especially between M_ν and σ_8 , limit the constraining power of standard two-point clustering statistics. In this work, we investigate whether we can break these degeneracies and constrain M_ν with the next higher-order correlation function — the bispectrum. We first examine the redshift-space halo bispectrum of 800 N -body simulations from the HADES suite and demonstrate that the bispectrum helps break the M_ν – σ_8 degeneracy. Then using 22,000 N -body simulations of the Quijote suite, we quantify for the first time the full information content of the redshift-space halo bispectrum down to nonlinear scales using a Fisher matrix forecast of $\{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\}$. For $k_{\text{max}}=0.5 \ h/\text{Mpc}$, the bispectrum provides Ω_m , Ω_b , h , n_s , and σ_8 constraints 1.9, 2.6, 3.1, 3.6, and 2.6 times tighter than the power spectrum. For M_ν , the bispectrum improves the 1σ constraint from 0.2968 to 0.0572 eV — over 5 times tighter than the power spectrum. Even with priors from *Planck*, the bispectrum improves M_ν constraints by a factor of 2.7. Although we reserve marginalizing over a more complete set of bias parameters to the next paper of the series, these constraints are derived for a $(1 \ h^{-1}\text{Gpc})^3$ box, a substantially smaller volume than upcoming surveys. Thus, our results demonstrate that the bispectrum offers significant improvements over the power spectrum, especially for constraining M_ν .

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

* hahn.changhoon@gmail.com

1. INTRODUCTION

The lower bound on the sum of neutrino masses ($M_\nu \gtrsim 0.06$ eV), discovered by neutrino oscillation experiments, provides conclusive evidence of physics beyond the Standard Model of particle physics (Forero et al. 2014; Gonzalez-Garcia et al. 2016). 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. Neutrino oscillation experiments, however, are insensitive to the absolute neutrino mass scales. Other laboratory experiments sensitive to M_ν (*e.g.* double beta decay and tritium beta decay experiments) have the potential to place upper bounds of $M_\nu < 0.2$ eV in upcoming experiments (Bonn et al. 2011; Drexlin et al. 2013). However, these upper bound alone are not sufficient to distinguish between the mass hierarchies. Neutrinos, through the cosmic neutrino background, affect the expansion history and the growth of cosmic structure. Measuring these effects with cosmological observables 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 (Lesgourgues & Pastor 2012, 2014). 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 Lesgourgues & Pastor (2012, 2014) and Gerbino (2018).

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 (Planck Collaboration et al. 2018). Adding the Baryon Acoustic Oscillation (BAO) to the *Planck* likelihood breaks geometrical degeneracies (among M_ν , h , Ω_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. Combining these probes, however, introduces a strong correlation between M_ν and A_s , the scale amplitude. Since CMB experiments measure the combined quantity $A_s e^{-2\tau}$, where τ is the optical depth of reionization, this leads to a strong degeneracy between M_ν and τ (Allison et al. 2015; Liu et al. 2016; Archidiacono et al. 2017).

The $\tau - M_\nu$ degeneracy will continue to be a major bottleneck for M_ν constraints from CMB data. 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 (Abazajian et al. 2016). Only the upcoming CLASS experiment aims to improve τ constraints (Watts et al. 2018). Meanwhile, proposed future space-

based experiments such as LiteBIRD¹ and LiteCORe², 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 DESI³, PFS⁴, EUCLID⁵, and WFIRST⁶ will be able tightly constrain M_ν (Audren et al. 2013; Font-Ribera et al. 2014; Petracca et al. 2016; Sartoris et al. 2016; Boyle & Komatsu 2018).

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; Villaescusa-Navarro et al. 2013; Castorina et al. 2015; Adamek et al. 2017; Emberson et al. 2017; Villaescusa-Navarro et al. 2018). 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.* Heitmann et al. 2009; Kwan et al. 2015; Euclid Collaboration et al. 2018; McClintock et al. 2018; Zhai et al. 2018; Wibking et al. 2019). 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; Saito et al. 2008; Wong 2008; Saito et al. 2009; Viel et al. 2010; Agarwal & Feldman 2011; Bird et al. 2012; Castorina et al. 2015; Banerjee & Dalal 2016) and in redshift-space (Marulli et al. 2011; Castorina et al. 2015; Upadhye et al. 2016). Most recently, using a suite of more than 1000 simulations, Villaescusa-Navarro et al. (2018) 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 1). 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. In fact, the bispectrum has a comparable signal-to-noise ratio to the power spectrum on nonlinear scales (Sefusatti & Scoccimarro 2005; Chan & Blot 2017). 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 4.2). 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 .

In this work, we examine the effect of massive neutrinos on the redshift-space halo bispectrum using $\sim 23,000$ N -body simulations with different neutrino masses. We first demonstrate that the bispectrum helps break the $M_\nu - \sigma_8$ degeneracy found in the power spectrum. Then we present the

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/>

full information content of the bispectrum for all triangle configurations down to $k_{\text{max}} = 0.5 \, h/\text{Mpc}$ using Fisher forecasts where we estimate the covariance matrix and derivatives with the large set of simulations from the Quijote suite (Villaescusa-Navarro et al. 2019). We then explore forecasts with different sets of nuisance and bias parameters as well as with and without priors from Planck and discuss some caveats of our results. This paper is the first of a series of papers that aim to demonstrate the potential of galaxy bispectrum analyses in constraining M_ν . In particular, with this series, we aim to establish the potential for constraining cosmological parameters (especially M_ν) with galaxy bispectrum analyses that extend to nonlinear scales using simulation based emulation models for upcoming galaxy surveys. In this paper, we focus on the halo bispectrum and consider some simple bias models. Parameter constraints, however, are derived from the galaxy distribution. In the subsequent paper, we will include a more realistic galaxy bias model that describes the distribution of central and satellite galaxies in halos in our forecasts to quantify the full information content of the *galaxy* bispectrum. In the series, we will also present methods to tackle challenges that come with analyzing the full galaxy bispectrum, such as data compression for reducing the dimensionality of the bispectrum.

In Section 2, we describe the two N -body simulation suites, HADES and Quijote, and the halo catalogs constructed from them. We then describe in Section 3, how we measure the bispectrum of these simulations. Afterwards, we use the redshift-space halo bispectra to demonstrate the distinct imprint of M_ν on the bispectrum, which allows it to break the degeneracy between M_ν and σ_8 , in Section 4.1. Finally, in Section 4.2 we present the full information content of the halo bispectrum with a Fisher forecast of cosmological parameters and demonstrate how the bispectrum significantly improves the constraints on the cosmological parameters: Ω_m , Ω_b , h , n_s , σ_8 , and *especially* M_ν .

2. HADES AND QUIJOTE SIMULATION SUITES

The HADES⁷ and Quijote⁸ suites are sets of, 43000 total, N -body simulations run on multiple cosmologies, including those with massive neutrinos ($M_\nu > 0 \, \text{eV}$). In this work, we use a subset of the HADES and Quijote simulations. Below, we briefly describe these simulations; a summary of the simulations can be found in Table 1. The HADES simulations start from Zel’dovich approximated initial conditions generated at $z = 99$ using the Zennaro et al. (2017) rescaling method and follow the gravitational evolution of $N_{\text{cdm}} = 512^3$ CDM, plus $N_\nu = 512^3$ neutrino particles for $M_\nu > 0 \, \text{eV}$ cosmologies, to $z = 0$. They are run using the GADGET-III TreePM+SPH code (Springel 2005) in a periodic $(1 \, h^{-1}\text{Gpc})^3$ box. All of the HADES simulations share the following cosmological parameter values, which are in good agreement with Planck constraints Ade et al. (2016): $\Omega_m=0.3175$, $\Omega_b=0.049$, $\Omega_\Lambda=0.6825$, $n_s=0.9624$, $h=0.6711$, and $k_{\text{pivot}} = 0.05 \, h\text{Mpc}^{-1}$.

The HADES suite includes N -body simulations with degenerate massive neutrinos of $M_\nu = 0.06$, 0.10 , and $0.15 \, \text{eV}$. These simulations 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). HADES also includes simulations with massless neu-

⁷ <https://franciscovillaescusa.github.io/hades.html>

⁸ <https://github.com/franciscovillaescusa/Quijote-simulations>

trino and different values of σ_8 to examine the $M_\nu - \sigma_8$ degeneracy. The σ_8 values in these simulations were chosen to match *either* σ_8^m or σ_8^c (σ_8 computed with respect to total matter, CDM + baryons + ν , or CDM + baryons) of the massive neutrino simulations: $\sigma_8 = 0.822, 0.818, 0.807$, and 0.798 . Each model has 100 independent realizations and we focus on the snapshots saved at $z = 0$. Halos closely trace the CDM+baryon field rather than the total matter field and neutrinos have negligible contribution to halo masses (*e.g.* Ichiki & Takada 2012; Castorina et al. 2014; LoVerde 2014; Villaescusa-Navarro et al. 2014). Hence, dark matter halos are identified in each realization 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 catalogs to halos with masses above $M_{\text{lim}} = 3.2 \times 10^{13} h^{-1} M_\odot$. We refer readers to Villaescusa-Navarro et al. (2018) for more details on the HADES simulations.

In addition to HADES, we use simulations from the Quijote suite, a set of 42,000 N -body simulations that in total contain more than 8 trillion (8×10^{12}) particles over a volume of $42000(h^{-1}\text{Gpc})^3$. These simulations were designed to quantify the information content of different cosmological observables using Fisher matrix forecasting technique (Section 4.2). They are therefore constructed to accurately calculate the covariance matrices of observables and the derivatives of observables with respect to cosmological parameters: Ω_m , Ω_b , h , n_s , σ_8 , and M_ν .

To calculate covariance matrices, Quijote includes $N_{\text{cov}}=15,000$ N -body simulations at a 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). It also includes sets of 500 N -body simulations run at different cosmologies where only one parameter is varied from the fiducial cosmology at a time for the derivatives. Along Ω_m , Ω_b , h , n_s , and σ_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_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 500 simulations for $M_\nu = 0.1$, 0.2 , and 0.4 eV. Table 1 lists the cosmologies included in the Quijote suite.

The initial conditions for all Quijote simulations were generated at $z = 127$ using 2LPT for simulations with massless neutrinos and the Zel'dovich approximation for massive neutrinos. The suite also includes a set of simulations at the fiducial $M_\nu = 0$ eV cosmology with Zel'dovich approximation initial conditions, which we use later in Section 4.2. Like HADES, the initial conditions of simulations with massive neutrinos take their scale-dependent growth factors/rates into account using the Zennaro et al. (2017) method. From the initial conditions, all of 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 (same as HADES). Halos are then identified using the same FoF scheme and mass limit as HADES. For the fiducial cosmology, the halo catalogs have $\sim 156,000$ halos ($\bar{n} \sim 1.56 \times 10^{-4} h^3 \text{Mpc}^{-3}$) with $\bar{n}P_0(k=0.1) \sim 3.23$. We refer readers to Villaescusa-Navarro et al. (2019) for further details on the Quijote simulations.

3. BISPECTRUM

We are interested in breaking parameter degeneracies that limit the constraining power on M_ν of two-point clustering analyses using three-point clustering statistics — *i.e.* the bispectrum. In this section, we describe the bispectrum estimator used throughout the paper. We focus on the bispectrum monopole ($\ell = 0$) and use an estimator that exploits the speed of Fast Fourier Transforms (FFTs).

Table 1. Specifications of the HADES and Quijote simulation suites.

Name	M_ν (eV)	Ω_m	Ω_b	h	n_s	σ_8^m ($10^{10}h^{-1}M_\odot$)	σ_8^c ($10^{10}h^{-1}M_\odot$)	ICs	realizations
HADES suite									
Fiducial	0.0	0.3175	0.049	0.6711	0.9624	0.833	0.833	Zel'dovich	100
	0.06	0.3175	0.049	0.6711	0.9624	0.819	0.822	Zel'dovich	100
	0.10	0.3175	0.049	0.6711	0.9624	0.809	0.815	Zel'dovich	100
	0.15	0.3175	0.049	0.6711	0.9624	0.798	0.806	Zel'dovich	100
	0.0	0.3175	0.049	0.6711	0.9624	0.822	0.822	Zel'dovich	100
	0.0	0.3175	0.049	0.6711	0.9624	0.818	0.818	Zel'dovich	100
	0.0	0.3175	0.049	0.6711	0.9624	0.807	0.807	Zel'dovich	100
	0.0	0.3175	0.049	0.6711	0.9624	0.798	0.798	Zel'dovich	100
Quijote suite									
Fiducial	0.0	0.3175	0.049	0.6711	0.9624	0.834	0.834	2LPT	15,000
Fiducial ZA	0.0	0.3175	0.049	0.6711	0.9624	0.834	0.834	Zel'dovich	500
M_ν^+	<u>0.1</u>	0.3175	0.049	0.6711	0.9624	0.834	0.834	Zel'dovich	500
M_ν^{++}	<u>0.2</u>	0.3175	0.049	0.6711	0.9624	0.834	0.834	Zel'dovich	500
M_ν^{+++}	<u>0.4</u>	0.3175	0.049	0.6711	0.9624	0.834	0.834	Zel'dovich	500
Ω_m^+	0.0	<u>0.3275</u>	0.049	0.6711	0.9624	0.834	0.834	2LPT	500
Ω_m^-	0.0	<u>0.3075</u>	0.049	0.6711	0.9624	0.834	0.834	2LPT	500
Ω_b^+	0.0	0.3175	<u>0.051</u>	0.6711	0.9624	0.834	0.834	2LPT	500
Ω_b^-	0.0	0.3175	<u>0.047</u>	0.6711	0.9624	0.834	0.834	2LPT	500
h^+	0.0	0.3175	0.049	<u>0.6911</u>	0.9624	0.834	0.834	2LPT	500
h^-	0.0	0.3175	0.049	<u>0.6511</u>	0.9624	0.834	0.834	2LPT	500
n_s^+	0.0	0.3175	0.049	0.6711	<u>0.9824</u>	0.834	0.834	2LPT	500
n_s^-	0.0	0.3175	0.049	0.6711	<u>0.9424</u>	0.834	0.834	2LPT	500
σ_8^+	0.0	0.3175	0.049	0.6711	0.9624	<u>0.849</u>	<u>0.849</u>	2LPT	500
σ_8^-	0.0	0.3175	0.049	0.6711	0.9624	<u>0.819</u>	<u>0.819</u>	2LPT	500

Top: The HADES suite includes sets of 100 N -body simulations with degenerate massive neutrinos of $M_\nu = 0.06, 0.10$, and 0.15 eV as well as sets of simulations with $M_\nu = 0.0$ eV and $\sigma_8 = 0.822, 0.818, 0.807$, and 0.798 to examine the $M_\nu - \sigma_8$ degeneracy. **Bottom:** The Quijote suite includes 15,000 N -body simulations at the fiducial cosmology to accurately estimate the covariance matrices. It also includes sets of 500 simulations at 13 different cosmologies, where only one parameter is varied from the fiducial value (underlined), to estimate derivatives of observables along the cosmological parameters.

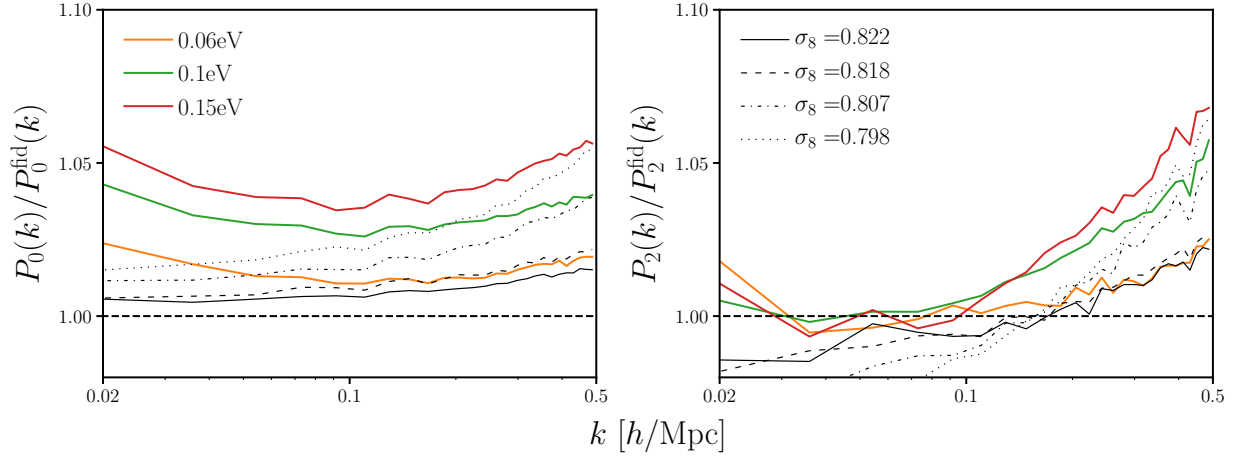


Figure 1. Impact of M_ν and σ_8 on the redshift-space halo power spectrum monopole, P_0 , and quadrupole, P_2 , of the HADES simulation suite, which includes simulations with $M_\nu = 0.0, 0.06, 0.10$, and 0.15 eV (orange, green, and red) as well as massless neutrino simulations with $\sigma_8 = 0.822, 0.818, 0.807$, and 0.798 (black solid, dashed, dot-dashed, and dotted). P_ℓ at each of these cosmology is averaged over 100 N -body realizations. While neutrinos suppress small-scale matter power spectrum, the halo P_ℓ increases with higher M_ν because cosmologies with higher M_ν have lower σ_8 , which translates into larger halo bias for a fixed halo mass limit. M_ν and σ_8 produce almost identical effects on halo clustering on small scales ($k > 0.1$ h/Mpc). This degeneracy can be partially broken through the quadrupole; however, M_ν and σ_8 produce almost the same effect on two-point clustering — within a few percent.

Our estimator is similar to the estimators described in Sefusatti (2005), Scoccimarro (2015), and Sefusatti et al. (2016); we also follow their formalism in our description below. Although Sefusatti et al. (2016) and Scoccimarro (2015) respectively describe estimators in real- and redshift-space, since we focus on the bispectrum monopole, we note that there is no difference.

To measure the bispectrum of our halo catalogs, we begin by interpolating the halo positions to a grid, $\delta(\mathbf{x})$, and Fourier transforming the grid to get $\delta(\mathbf{k})$. We use a fourth-order interpolation to get interlaced grids, which has advantageous anti-aliasing properties that allow unbiased measurements up to the Nyquist frequency (Hockney & Eastwood 1981; Sefusatti et al. 2016). Then using $\delta(\mathbf{k})$ we measure the bispectrum monopole as

$$\hat{B}_{\ell=0}(k_1, k_2, k_3) = \frac{1}{V_B} \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \delta_D(\mathbf{q}_{123}) \delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{q}_3) - B_{\ell=0}^{\text{SN}}. \quad (1)$$

δ_D is the Dirac delta function and hence $\delta_D(\mathbf{q}_{123}) = \delta_D(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3)$ ensures that the \mathbf{q}_i triplet actually forms a closed triangle. Each of the integrals above represent an integral over a spherical shell in k -space with radius δk centered at \mathbf{k}_i :

$$\int_{k_i} d^3 q \equiv \int_{k_i - \delta k/2}^{k_i + \delta k/2} dq q^2 \int d\Omega. \quad (2)$$

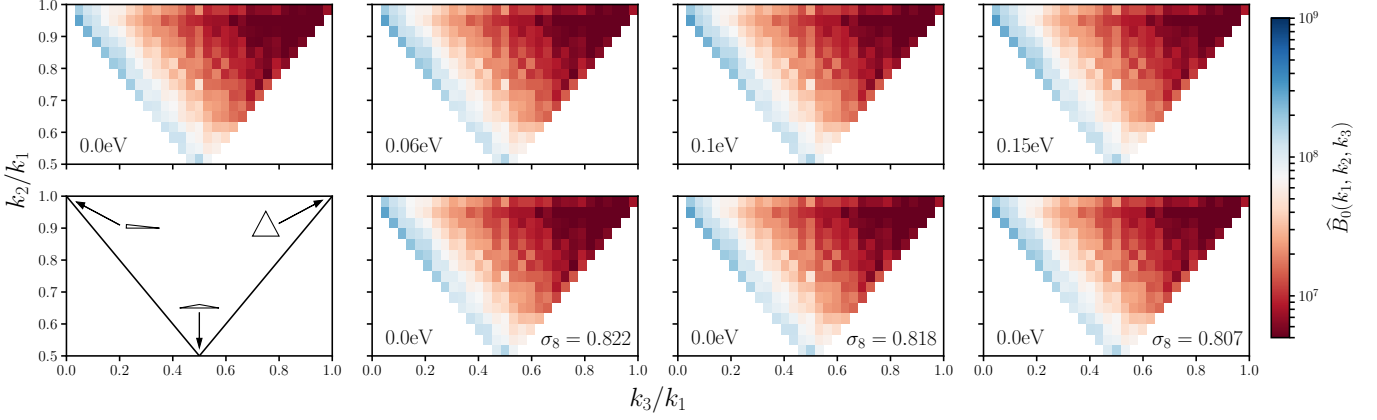


Figure 2. The redshift-space halo bispectrum, $\hat{B}_0(k_1, k_2, k_3)$, as a function of triangle configuration shape for $M_\nu = 0.0, 0.06, 0.10$, and 0.15 eV (upper panels) and $M_\nu = 0.0$ eV, $\sigma_8 = 0.822, 0.818$, and 0.807 (lower panels). The \hat{B}_0 for each cosmology (each panel) is averaged over 100 N -body realizations. The HADES simulations of the top and bottom panels in the three right-most columns, have matching σ_8 values (Section 2). We describe the triangle configuration shape by the ratio of the triangle sides: k_3/k_1 and k_2/k_1 . As we describe schematically in the lower leftmost panel, in each panel, the upper left bin contains squeezed triangles ($k_1 = k_2 \gg k_3$), the upper right bin contains equilateral triangles ($k_1 = k_2 = k_3$), and the bottom center bin contains folded triangles ($k_1 = 2k_2 = 2k_3$). We include all 1898 triangle configurations with $k_1, k_2, k_3 \leq k_{\text{max}} = 0.5 h/\text{Mpc}$ and use the \hat{B}_0 estimator in Section 3.

V_B is a normalization factor proportional to the number of triplets \mathbf{q}_1 , \mathbf{q}_2 , and \mathbf{q}_3 that can be found in the triangle bin defined by k_1 , k_2 , and k_3 with width δk :

$$V_B = \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \delta_D(\mathbf{q}_{123}). \quad (3)$$

Lastly, $B_{\ell=0}^{\text{SN}}$ is the correction for the Poisson shot noise, which contributes due to the self-correlation of individual objects:

$$B_{\ell=0}^{\text{SN}}(k_1, k_2, k_3) = \frac{1}{\bar{n}} (P_0(k_1) + P_0(k_2) + P_0(k_3)) + \frac{1}{\bar{n}^2} \quad (4)$$

where \bar{n} is the number density of objects (halos) and P_0 is the power spectrum monopole.

In order to evaluate the integrals in Eq. 1, we take advantage of the plane-wave representation of the Dirac delta function and rewrite the equation as

$$\hat{B}_{\ell=0}(k_1, k_2, k_3) = \frac{1}{V_B} \int \frac{d^3 x}{(2\pi)^3} \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{q}_3) e^{i\mathbf{q}_{123} \cdot \mathbf{x}} - B_{\ell=0}^{\text{SN}} \quad (5)$$

$$= \frac{1}{V_B} \int \frac{d^3 x}{(2\pi)^3} \prod_{i=1}^3 I_{k_i}(\mathbf{x}) - B_{\ell=0}^{\text{SN}} \quad (6)$$

where

$$I_{k_i}(\mathbf{x}) = \int_k d^3 q \delta(\mathbf{q}) e^{i\mathbf{q} \cdot \mathbf{x}}. \quad (7)$$

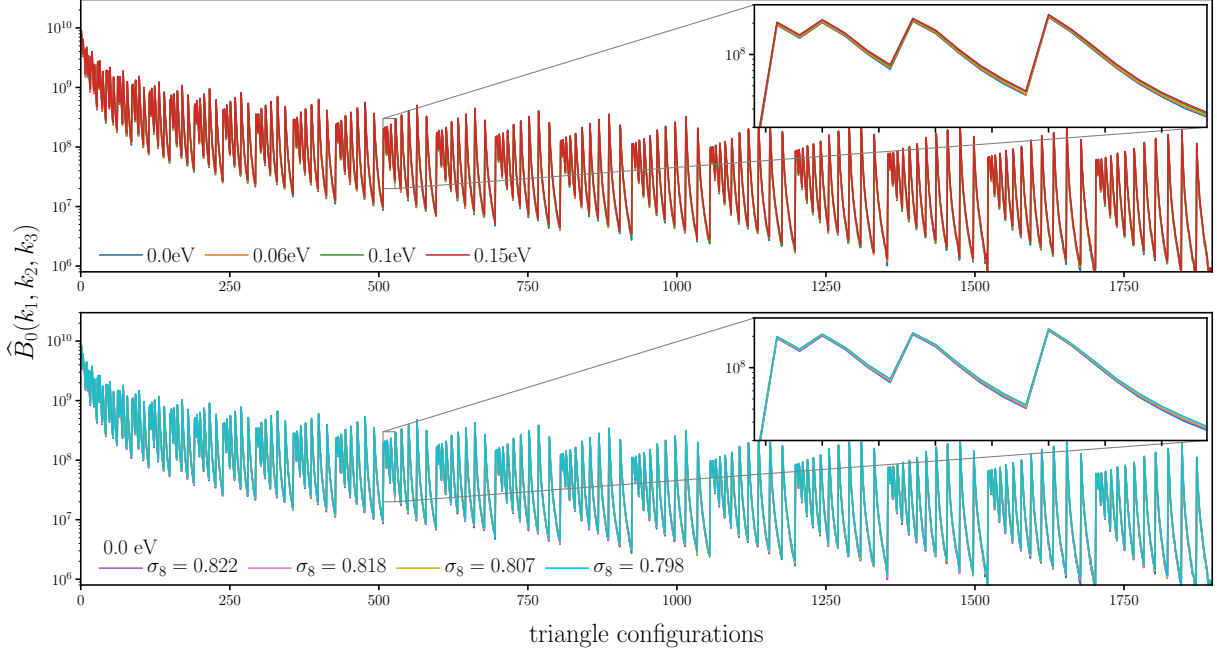


Figure 3. The redshift-space halo bispectrum, $\hat{B}_0(k_1, k_2, k_3)$, as a function of triangle configurations for $M_\nu = 0.0, 0.06, 0.10$, and 0.15 eV (top panel) and $M_\nu = 0.0$ eV, $\sigma_8 = 0.822, 0.818, 0.807$, and 0.798 (lower panel). Each \hat{B}_0 is averaged over 100 N -body realizations. We include all possible triangle configurations with $k_1, k_2, k_3 \leq k_{\max} = 0.5$ h/Mpc where we order the configurations by looping through k_3 in the inner most loop and k_1 in the outer most loop satisfying $k_1 \geq k_2 \geq k_3$ (see Appendix A for more details). In the insets of the panels we zoom into triangle configurations with $k_1 = 0.320$, $0.170 \leq k_2 \leq 0.302$, and $0.094 \leq k_3 \leq 0.302$ h/Mpc .

At this point, we measure $\hat{B}_{\ell=0}(k_1, k_2, k_3)$ by calculating the I_{k_i} s with inverse FFTs and summing over in real space. For $\hat{B}_{\ell=0}$ measurements throughout the paper, we use $\delta(\mathbf{x})$ grids with $N_{\text{grid}} = 360$ and triangle configurations defined by k_1, k_2, k_3 bins of width $\Delta k = 3k_f = 0.01885$ h/Mpc , three times the fundamental mode $k_f = 2\pi/(1000$ $h^{-1}\text{Mpc})$ given the box size⁹.

We present the redshift-space halo bispectrum of the HADES simulations measured using the estimator above in two ways: one that emphasizes the triangle shape dependence (Figure 2) and the other that emphasizes the amplitude (Figure 3). In Figure 2, we plot $\hat{B}_0(k_1, k_2, k_3)$ as a function of k_2/k_1 and k_3/k_1 , which describe the triangle configuration shapes. In each panel, the colormap of the $(k_2/k_1, k_3/k_1)$ bins represent the weighted average \hat{B}_0 amplitude of all triangle configurations in the bins. The upper left bins contain squeezed triangles ($k_1 = k_2 \gg k_3$), the upper right bins contain equilateral triangles ($k_1 = k_2 = k_3$), and the bottom center bins contain folded triangles ($k_1 = 2k_2 = 2k_3$) as schematically highlighted in the lower leftmost panel. We include all possible 1898 triangle configurations with $k_1, k_2, k_3 < k_{\max} = 0.5$ h/Mpc . \hat{B}_0 in the upper panels are HADES models with $(M_\nu, \sigma_8) = (0.0$ eV, $0.833)$ (fiducial), $(0.06$ eV, $0.822)$, $(0.10$ eV, $0.815)$, and $(0.15$ eV, $0.806)$.

⁹ The code that we use to evaluate $\hat{B}_{\ell=0}$ is publicly available at <https://github.com/changhoonhahn/pySpectrum>

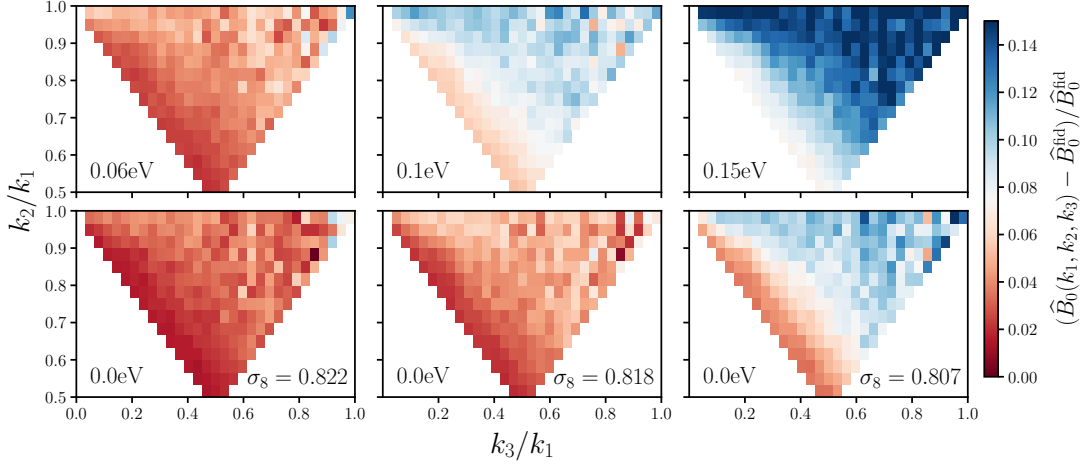


Figure 4. The shape dependence of the M_ν and σ_8 imprint on the redshift-space halo bispectrum, $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$. We align the $M_\nu = 0.06, 0.10$, and 0.15 eV HADES simulations in the upper panels with $M_\nu = 0.0$ eV $\sigma_8 = 0.822, 0.818$, and 0.807 simulations on the bottom such that the top and bottom panels in each column have matching σ_8^c , which produce mostly degenerate imprints on the redshift-space power spectrum. \hat{B}_0^{fid} is measured from the fiducial HADES simulation: $M_\nu = 0.0$ eV and $\sigma_8 = 0.833$. The difference between the top and bottom panels highlight that M_ν leaves a imprint distinct from σ_8 on elongated and isosceles triangles, bins along the bottom left and bottom right edges, respectively. *The imprint of M_ν has a distinct shape dependence on the bispectrum that cannot be replicated by varying σ_8 .*

\hat{B}_0 in the lower panels are HADES models with $M_\nu = 0.0$ eV and $\sigma_8 = 0.822, 0.818$, and 0.807 . The top and bottom panels of the three right-most columns have matching σ_8 values (Section 2).

Next, in Figure 3 we plot $\hat{B}_0(k_1, k_2, k_3)$ for all possible triangle configurations with $k_1, k_2, k_3 < k_{\text{max}} = 0.5$ h/Mpc where we order the configurations by looping through k_3 in the inner most loop and k_1 in the outer most loop with $k_1 \geq k_2 \geq k_3$. In the top panel, we present \hat{B}_0 of HADES models with $M_\nu = 0.0, 0.06, 0.10$, and 0.15 eV; in the lower panel, we present \hat{B}_0 of HADES models with $M_\nu = 0.0$ eV and $\sigma_8 = 0.822, 0.818$, and 0.807 . We zoom into triangle configurations with $k_1 = 0.320$, $0.170 \leq k_2 \leq 0.302$, and $0.094 \leq k_3 \leq 0.302$ h/Mpc in the insets of the panels. For further details on the redshift-space bispectrum, we refer to Appendix A.

4. RESULTS

4.1. Breaking the $M_\nu - \sigma_8$ degeneracy

One major bottleneck of constraining M_ν with the power spectrum alone is the strong $M_\nu - \sigma_8$ degeneracy. The imprints of M_ν and σ_8 on the power spectrum are degenerate and for models with the same σ_8^c , the power spectrum only differ by $< 1\%$ (see Figure 1 and Villaescusa-Navarro et al. 2018). The HADES suite, which has simulations with $M_\nu = 0.0, 0.06, 0.10$, and 0.15 eV as well as $M_\nu = 0.0$ eV simulations with matching σ_8^c , provides an ideal set of simulations to disentangle the impact of M_ν and examine the degeneracy between M_ν and σ_8 (Section 2 and Table 1). We measure the bispectrum of the HADES simulations (Figure 2 and 3) and present how the bispectrum can significantly improve M_ν constraints by breaking the $M_\nu - \sigma_8$ degeneracy.

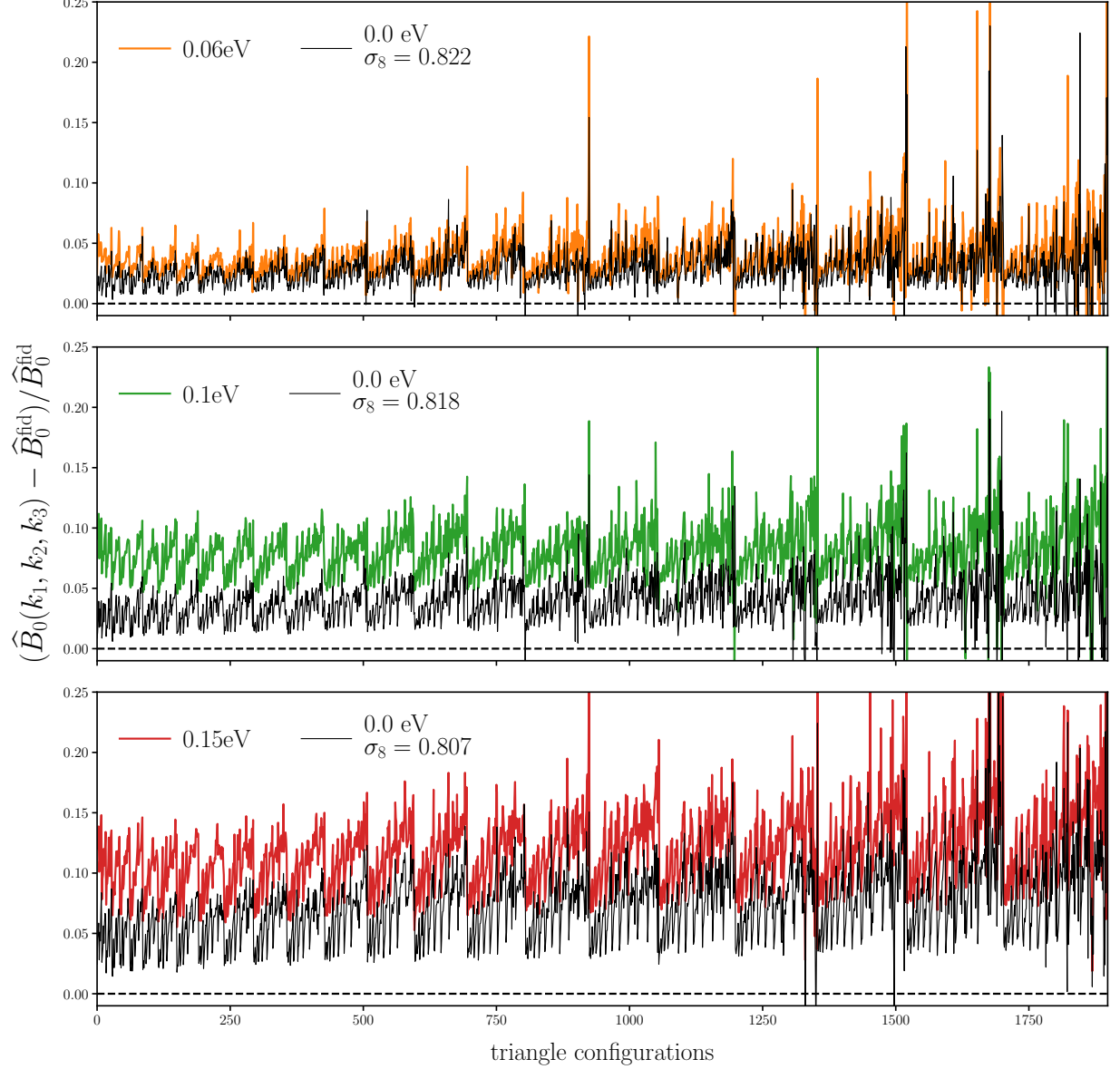


Figure 5. The impact of M_ν and σ_8 on the redshift-space halo bispectrum, $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$, for all 1898 triangle configurations with $k_1, k_2, k_3 \leq 0.5 h/\text{Mpc}$. We compare $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ of the $M_\nu = 0.06$ (top), 0.10 (middle), and 0.15 eV (bottom) HADES simulations to $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ of $M_\nu = 0.0$ eV, $\sigma_8 = 0.822, 0.818,$ and 0.807 simulations. The imprint of M_ν on the bispectrum has a significantly different amplitude than the imprint of σ_8 . For instance, $M_\nu = 0.15$ eV (red) has a $\sim 5\%$ stronger impact on the bispectrum than $M_\nu = 0.0$ eV $\sigma_8 = 0.798$ (black) even though their power spectra only differ by $< 1\%$ (Figure 1). *The distinct imprint of M_ν on the bispectrum illustrate that the bispectrum can break the degeneracy between M_ν and σ_8 that degrade constraints from two-point analyses.*

We begin by examining the triangle shape dependent imprint of M_ν on the redshift-space halo bispectrum versus σ_8 alone. In Figure 4, we present the fractional residual, $(\Delta\hat{B}_0 = \hat{B}_0 - \hat{B}_0^{\text{fid}})/\hat{B}_0^{\text{fid}}$, as a function of k_2/k_1 and k_3/k_1 for $M_\nu = 0.06, 0.10,$ and 0.15 eV in the upper panels and 0.0 eV

$\sigma_8 = 0.822, 0.818,$ and 0.807 in the bottom panels. The simulations in the top and bottom panels of each column have matching σ_8^c . Overall, as M_ν increases, the amplitude of the bispectrum increases for all triangle shapes (top panels). This increase is due to halo bias (Villaescusa-Navarro et al. 2018, see also Figure 1). Since we impose a fixed M_{lim} on our halos, lower values of σ_8 translate to a larger halo bias, which boosts the amplitude of the bispectrum. Within the overall increase in amplitude, however, there is a significant triangle dependence. Equilateral triangles (upper left) have the largest increase. For $M_\nu = 0.15$ eV, the bispectrum is $\sim 15\%$ higher than \hat{B}_0^{fid} for equilateral triangles while only $\sim 8\%$ higher for folded triangles (lower center). The noticeable difference in $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ between equilateral and squeezed triangles (upper left) is roughly consistent with the comparison in Figure 7 of Ruggeri et al. (2018). They, however, fix A_s in their simulations and measure the real-space halo bispectrum so we refrain from any detailed comparisons.

As σ_8 increases with $M_\nu = 0.0$ eV, the bispectrum also increases overall for all triangle shapes (bottom panels). However, the comparison of the top and bottom panels in each column reveals significant differences in $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ for M_ν versus σ_8 alone. Between $M_\nu = 0.15$ eV and $\{0.0 \text{ eV}, \sigma_8 = 0.807\}$ cosmologies, there is an overall $\gtrsim 5\%$ difference in the bispectrum. In addition, the shape dependence of the $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ increase is different for M_ν than σ_8 . This is particularly clear in the differences between 0.1 eV (top center panel) and $\{0.0 \text{ eV}, \sigma_8 = 0.807\}$ (bottom right panel): near equilateral triangles in the two panels have similar $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ while triangle shapes near the lower left edge from the squeezed to folded triangles have significantly different $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$. Hence, M_ν leaves an imprint on the bispectrum with a distinct triangle shape dependence than σ_8 alone. In other words, the triangle shape dependent imprint of M_ν on the bispectrum cannot be replicated by varying σ_8 — unlike the power spectrum.

We next examine the amplitude of the M_ν imprint on the redshift-space halo bispectrum versus σ_8 alone as a function of all triangle configurations. We present $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ for all 1898 possible triangle configurations with $k_1, k_2, k_3 < k_{\text{max}} = 0.5 \text{ h/Mpc}$ in Figure 5. We compare $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ of the $M_\nu = 0.06, 0.10,$ and 0.15 eV HADES models to the $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ of $M_\nu = 0.0$ eV $\sigma_8 = 0.822, 0.818,$ and 0.807 models in the top, middle, and bottom panels, respectively. The comparison confirms the difference in overall amplitude of varying M_ν and σ_8 (Figure 4). For instance, $M_\nu = 0.15$ eV (red) has a $\sim 5\%$ stronger impact on the bispectrum than $M_\nu = 0.0$ eV $\sigma_8 = 0.798$ (black) even though their power spectra differ by $< 1\%$ (Figure 1).

The comparison in the panels of Figure 5 confirms the difference in the configuration dependence in $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ between M_ν versus σ_8 . The triangle configurations are ordered by looping through k_3 in the inner most loop and k_1 in the outer most loop such that $k_1 \geq k_2 \geq k_3$. In this ordering, k_1 increases from left to right (see Figure 10 and Appendix A). $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ of M_ν expectedly increases with k_1 : for small k_1 (on large scales), neutrinos behave like CDM and therefore the impact is reduced. However, $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ of M_ν has a smaller k_1 dependence than $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ of σ_8 . The distinct imprint of M_ν on the redshift-space halo bispectrum illustrates that the bispectrum can break the degeneracy between M_ν and σ_8 . Therefore, by including the bispectrum, we can more precisely constrain M_ν than with the power spectrum.

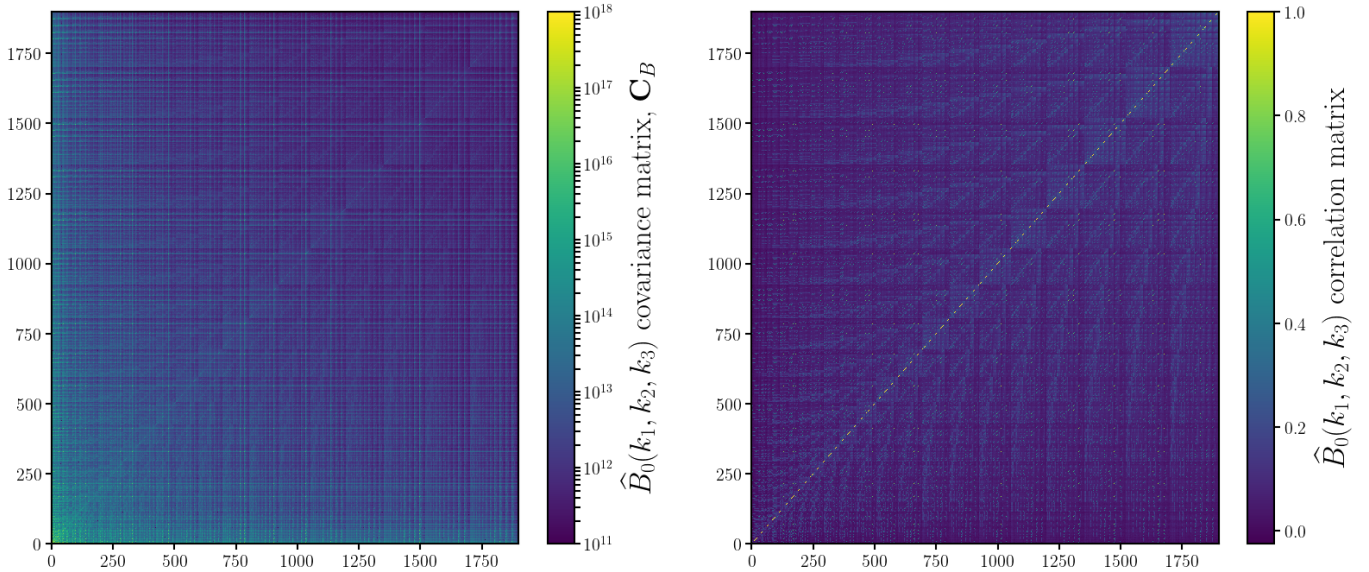


Figure 6. Covariance and correlation matrices of the redshift-space halo bispectrum estimated using $N_{\text{cov}} = 15,000$ realizations of the Quijote simulation suite at the fiducial cosmology: $\Omega_{\text{m}}=0.3175$, $\Omega_{\text{b}}=0.049$, $h=0.6711$, $n_s=0.9624$, $\sigma_8=0.834$, and $M_\nu=0.0$ eV. We include all possible triangle configurations with $k_1, k_2, k_3 \leq k_{\text{max}} = 0.5$ h/Mpc and order the configurations (bins) in the same way as Figures 3 and 5. We use the covariance matrix above for the Fisher matrix forecasts presented in Section 4.2.

4.2. M_ν and other Cosmological Parameter Forecasts

We demonstrate in the previous section with the HADES simulations, that the bispectrum helps break the M_ν – σ_8 degeneracy, a major challenge in precisely constraining M_ν with the power spectrum. While this establishes the bispectrum as a promising probe for M_ν , we are ultimately interested in determining the constraining power of the bispectrum for an analysis that include cosmological parameters beyond M_ν and σ_8 — *i.e.* Ω_{m} , Ω_{b} , h , and n_s . The Quijote simulation suite is *specifically* designed to answer this question through Fisher matrix forecast.

First, the Quijote suite includes $N_{\text{cov}} = 15,000$ N -body realizations run at a fiducial cosmology: $M_\nu=0.0$ eV, $\Omega_{\text{m}}=0.3175$, $\Omega_{\text{b}}=0.049$, $n_s=0.9624$, $h=0.6711$, and $\sigma_8=0.834$ (see Table 1). This allows us to robustly estimate the 1898×1898 covariance matrix \mathbf{C} of the *full* bispectrum (Figure 6). Second, the Quijote suite includes 500 N -body realizations evaluated at 13 different cosmologies, each a small step away from the fiducial cosmology parameter values along one parameter (Section 2 and Table 1). We apply redshift-space distortions along 3 different directions for these 500 realizations, which then effectively gives us $N_{\text{deriv.}}=1,500$ realizations. These simulations allow us to precisely estimate the derivatives of the bispectrum with respect to each of the cosmological parameters.

Since their introduction to cosmology over two decades ago, Fisher information matrices have been ubiquitously used to forecast the constraining power of future experiments (*e.g.* Jungman et al. 1996; Tegmark et al. 1997; Dodelson 2003; Heavens 2009; Verde 2010). Defined as

$$F_{ij} = - \left\langle \frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j} \right\rangle, \quad (8)$$

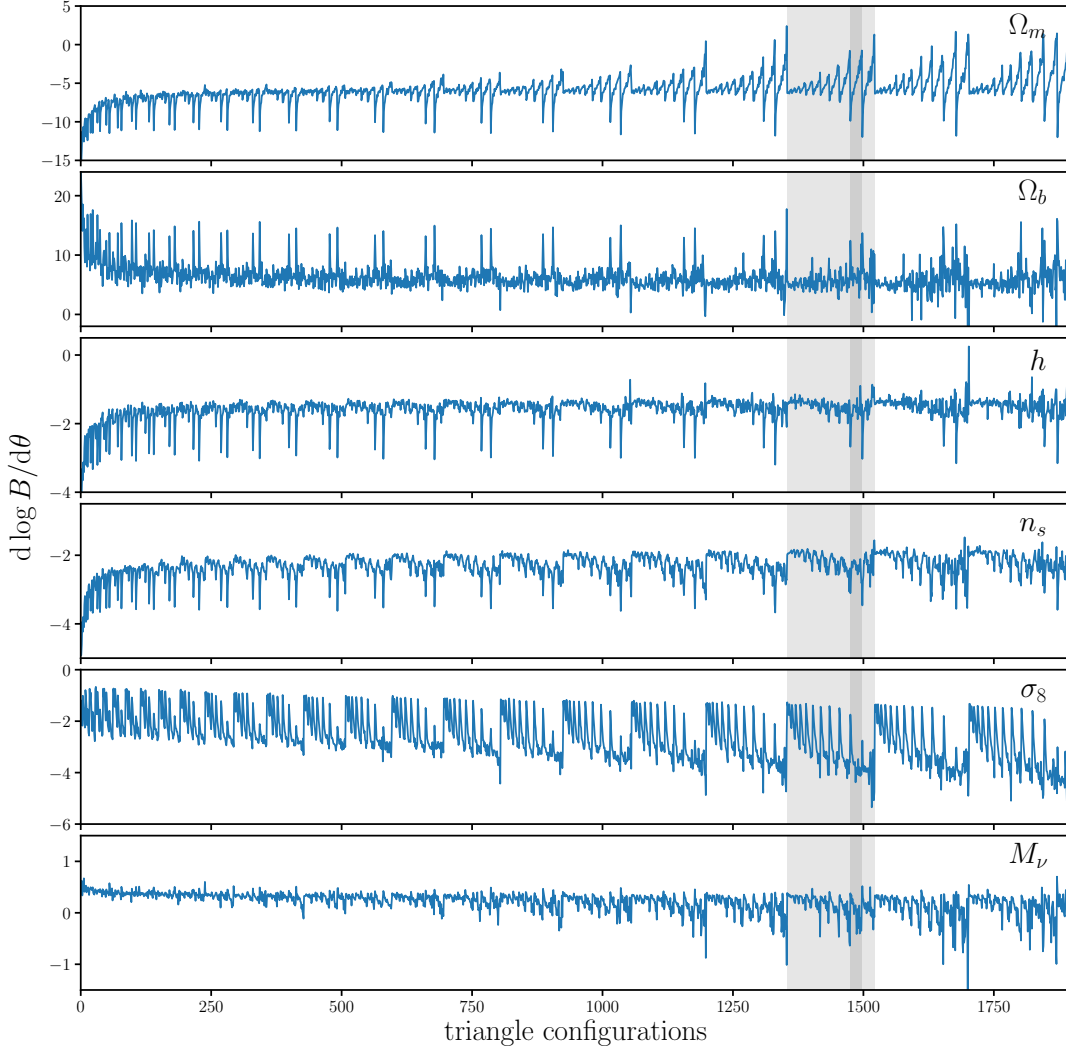


Figure 7. Derivatives of the redshift-space halo bispectrum, $d \log B_0 / d\theta$ with respect to Ω_m , Ω_b , h , n_s , σ_8 , and M_ν as a function of all 1898 triangle configurations with $k_1, k_2, k_3 \leq 0.5 \, h/\text{Mpc}$ (top to bottom panels). We estimate the derivatives at the fiducial parameters using $N_{\text{deriv}} = 1,500$ N -body realizations from the Quijote suite. The configurations above are ordered in the same way as Figures 3 and 5 (Appendix A). **The shaded region marks configurations with $k_1 = 72k_f$; the darker shaded region marks configurations with $k_1 = 72k_f$ and $k_2 = 69k_f$.** By using N -body simulations for the derivatives, we rely on fewer assumptions and approximations than analytic methods (*i.e.* perturbation theory). *Furthermore, with their accuracy on small scales, these derivatives enable us to quantify, for the first time, the information content of the bispectrum in the nonlinear regime.*

where \mathcal{L} is the likelihood, the Fisher matrix for the bispectrum can be written as

$$F_{ij} = \frac{1}{2} \text{Tr} \left[\mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_i} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_j} + \mathbf{C}^{-1} \left(\frac{\partial B_0}{\partial \theta_i} \frac{\partial B_0}{\partial \theta_j}^T + \frac{\partial B_0}{\partial \theta_i}^T \frac{\partial B_0}{\partial \theta_j} \right) \right]. \quad (9)$$

Since we assume that the B_0 likelihood is Gaussian, including the first term in Eq. 9 runs the risk of incorrectly including information from the covariance already included in the mean (Carron 2013).

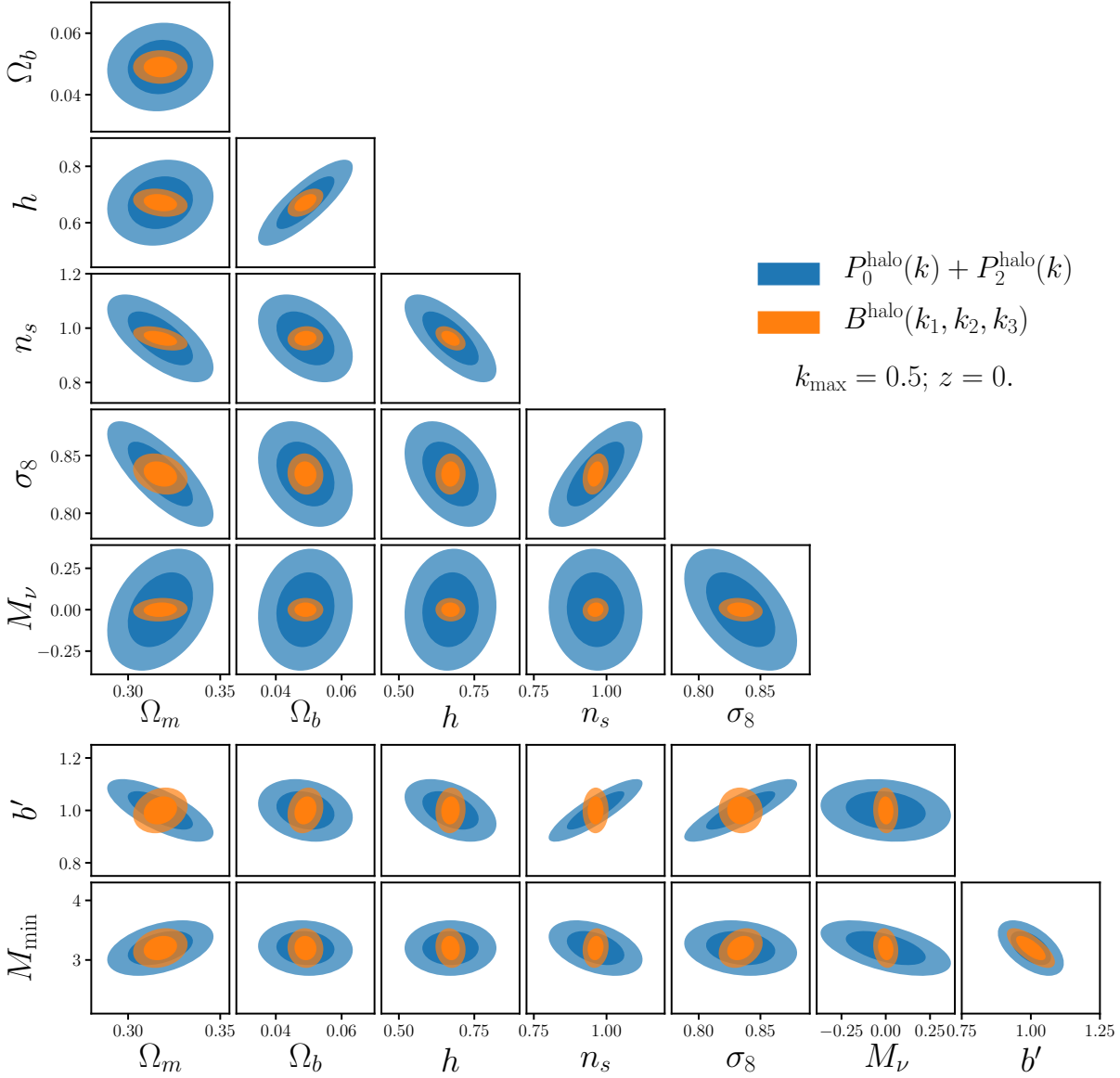


Figure 8. Fisher matrix constraints for M_ν and other cosmological parameters for the redshift-space halo bispectrum monopole (orange). We include Fisher parameter constraints for the redshift-space halo power spectrum monopole and quadrupole in blue for comparison. The contours mark the 68% and 95% confidence intervals. We set $k_{\text{max}} = 0.5 \, h/\text{Mpc}$ for both power spectrum and bispectrum. We include in our forecasts b' and M_{min} , a free amplitude scaling factor and halo mass limit, respectively. They serve as a simplistic bias model and we marginalize over them so that our constraints do not include extra constraining power from the difference in bias/number density in the different Quijote cosmologies. The bispectrum *substantially* improves constraints on all of the cosmological parameters over the power spectrum. Constraints on Ω_m , Ω_b , h , n_s , and σ_8 improve by factors of 1.9, 2.6, 3.1, 3.6, and 2.6, respectively. For M_ν , the bispectrum improves σ_{M_ν} from 0.2968 to 0.0572 eV — over a factor of ~ 5 improvement over the power spectrum.

We, therefore, conservatively neglect the first term and calculate the Fisher matrix as,

$$F_{ij} = \frac{1}{2} \text{Tr} \left[\mathbf{C}^{-1} \left(\frac{\partial B_0}{\partial \theta_i} \frac{\partial B_0}{\partial \theta_j}^T + \frac{\partial B_0}{\partial \theta_i}^T \frac{\partial B_0}{\partial \theta_j} \right) \right], \quad (10)$$

directly with \mathbf{C} and $\partial B_0/\partial \theta_i$ along each cosmological parameter from the Quijote simulations.

For Ω_m , Ω_b , h , n_s , and σ_8 , we estimate

$$\frac{\partial B_0}{\partial \theta_i} \approx \frac{B_0(\theta_i^+) - B_0(\theta_i^-)}{\theta_i^+ - \theta_i^-}, \quad (11)$$

where $B_0(\theta_i^+)$ and $B_0(\theta_i^-)$ are the average bispectrum of the 1,500 realizations at θ_i^+ and θ_i^- , respectively. Meanwhile, for M_ν , where the fiducial value is 0.0 eV and we cannot have negative M_ν , we use the Quijote simulations at M_ν^+ , M_ν^{++} , $M_\nu^{+++} = 0.1, 0.2, 0.4$ eV (Table 1) to estimate

$$\frac{\partial B_0}{\partial M_\nu} \approx \frac{-21B_0(\theta_{\text{fid}}^{\text{ZA}}) + 32B_0(M_\nu^+) - 12B_0(M_\nu^{++}) + B_0(M_\nu^{+++})}{1.2}, \quad (12)$$

which provides a $\mathcal{O}(\delta M_\nu^2)$ order approximation. Since the simulations at M_ν^+ , M_ν^{++} , and M_ν^{+++} are generated from Zel'dovich initial conditions, we use simulations at the fiducial cosmology also generated from Zel'dovich initial conditions ($\theta_{\text{fid}}^{\text{ZA}}$). By using these N -body simulations, instead of analytic methods (*e.g.* perturbation theory), we exploit the accuracy of the simulations in the non-linear regime and rely on fewer assumptions and approximations. In fact, these N -body simulation estimated derivatives are key in enabling us to quantify, for the first time, the total information content of the redshift-space bispectrum in the non-linear regime. We present the bispectrum derivatives for all triangle configurations in Figure 7 and discuss subtleties of the derivatives and tests of convergence and stability in Appendix B.

We present the constraints on M_ν and other cosmological parameters $\{\Omega_m, \Omega_b, h, n_s, \sigma_8\}$ derived from the redshift-space halo bispectrum Fisher matrix for $k_{\text{max}} = 0.5 h/\text{Mpc}$ in Figure 8. We include Fisher constraints for the redshift-space halo power spectrum monopole and quadrupole with the same k_{max} for comparison (blue). The shaded contours mark the 68% and 95% confidence intervals. We include in our Fisher constraints the following nuisance parameters: b' and M_{min} . b' is a scaling factor on the bispectrum amplitude, analogous to linear bias. And M_{min} is the halo mass limit, which we choose as a nuisance parameter to address the difference in the number densities among the Quijote cosmologies, which impacts the derivatives $\partial B_0/\partial \theta_i$. For instance, the σ_8^+ and σ_8^- cosmologies have halo $\bar{n} = 1.586 \times 10^{-4}$ and $1.528 \times 10^{-4} h^3 \text{Mpc}^{-3}$. These parameters serve as a simplistic bias model and by marginalizing over them we aim to ensure that our Fisher constraints do not include extra constraining power from the difference in bias or number density. b' is a multiplicative factor so $\partial B_0/\partial b' = B_0$. $\partial B_0/\partial M_{\text{min}}$, we estimate numerically using B_0 evaluated at $M_{\text{min}}^+ = 3.3 \times 10^{13} h^{-1} M_\odot$ and $M_{\text{min}}^- = 3.1 \times 10^{13} h^{-1} M_\odot$ with all other parameters set to the fiducial value. **In the power spectrum constraints that we include for comparison, we also include b' and M_{min} , where b' in this case is a scaling factor on the power spectrum amplitude.**

The bispectrum substantially improves constraints on all parameters over the power spectrum. For $k_{\text{max}} = 0.5 h/\text{Mpc}$, the bispectrum tightens the marginalized 1σ constraints, σ_θ , of Ω_m , Ω_b , h ,

Table 2. Marginalized Fisher parameter constraints from the redshift-space halo power spectrum (top) and bispectrum (bottom) for different k_{max} . We list constraints for cosmological parameters M_ν , Ω_m , Ω_b , h , n_s , and σ_8 as well as nuisance parameters b' and M_{min} .

	k_{max} (h/Mpc)	M_ν (eV)	Ω_m	Ω_b	h	n_s	σ_8	b'	M_{min} ($10^{13} h^{-1} M_\odot$)
		0.0	0.3175	0.049	0.6711	0.9624	0.834	1.	3.2
$P_{\ell=0,2}$	0.2	± 0.711	± 0.037	± 0.015	± 0.179	± 0.213	± 0.085	± 0.223	± 1.827
	0.3	± 0.509	± 0.029	± 0.013	± 0.151	± 0.167	± 0.047	± 0.117	± 0.725
	0.4	± 0.383	± 0.026	± 0.012	± 0.137	± 0.152	± 0.039	± 0.102	± 0.470
	0.5	$\pm \mathbf{0.297}$	± 0.023	± 0.012	± 0.123	± 0.130	± 0.037	± 0.096	± 0.376
+Planck priors	0.5	± 0.077	± 0.012	± 0.0012	± 0.0084	± 0.0044	± 0.017	± 0.033	± 0.212
B_0	0.2	± 0.256	± 0.029	± 0.014	± 0.142	± 0.138	± 0.049	± 0.264	± 1.299
	0.3	± 0.130	± 0.021	± 0.008	± 0.077	± 0.074	± 0.023	± 0.143	± 0.654
	0.4	± 0.078	± 0.015	± 0.006	± 0.052	± 0.047	± 0.016	± 0.088	± 0.367
	0.5	$\pm \mathbf{0.057}$	± 0.012	± 0.004	± 0.040	± 0.036	± 0.014	± 0.070	± 0.268
+Planck priors	0.5	± 0.043	± 0.009	± 0.0008	± 0.0064	± 0.0043	± 0.010	± 0.068	± 0.253

n_s , and σ_8 by factors of ~ 1.9 , 2.6 , 3.1 , 3.6 , 2.6 over the power spectrum. For M_ν , the bispectrum improves the constraint from $\sigma_{M_\nu} = 0.2968$ to 0.0572 eV — over a factor of 5 improvement over the power spectrum. This $\sigma_{M_\nu} = 0.0572$ eV constraint is from the bispectrum alone and only for a $1h^{-1}\text{Gpc}$ box. For a larger volume, V , σ_θ scales roughly as $\propto 1/\sqrt{V}$. The \bar{n} of our halo catalogs ($\sim 1.56 \times 10^{-4} h^3\text{Mpc}^{-3}$) is also significantly lower than, for instance, \bar{n} of the SDSS-III BOSS LOWZ + CMASS sample ($\sim 3 \times 10^{-4} h^3\text{Mpc}^{-3}$; Alam et al. 2015). A higher number density reduces the shot noise term and thus would further improve the constraining power of the bispectrum (see Eq. 4). We list the precise marginalized Fisher parameter constraints of both cosmological and nuisance parameters for P_ℓ and B_0 in Table 2.

Even below $k_{\text{max}} < 0.5 h/\text{Mpc}$, the bispectrum significantly improves cosmological parameter constraints. We compare σ_θ of Ω_m , Ω_b , h , n_s , σ_8 , and M_ν as a function of k_{max} for B_0 (orange) and $P_{\ell=0,2}$ (blue) in Figure 9. We only include the k_{max} range where the Fisher forecast is well defined — *i.e.* more data bins than the number of parameters: $k_{\text{max}} > 5 k_f \approx 0.03 h/\text{Mpc}$ for P_ℓ and $k_{\text{max}} > 12 k_f \approx 0.075 h/\text{Mpc}$ for B_0 . Figure 9 reveals that the improvement of the bispectrum σ_θ over the power spectrum σ_θ is larger at higher k_{max} . Although limited by the k_{max} range, the figure suggests that on large scales ($k_{\text{max}} \lesssim 0.1 h/\text{Mpc}$) σ_θ of P_ℓ crosses over σ_θ of B_0 so P_ℓ has more constraining power than B_0 , as expected on linear scales. On slightly smaller scales, $k_{\text{max}} = 0.2 h/\text{Mpc}$, we find that the bispectrum improves σ_θ by factors of ~ 1.3 , 1.1 , 1.3 , 1.5 , 1.7 , and 2.8 for Ω_m , Ω_b , h , n_s , σ_8 , and M_ν respectively.

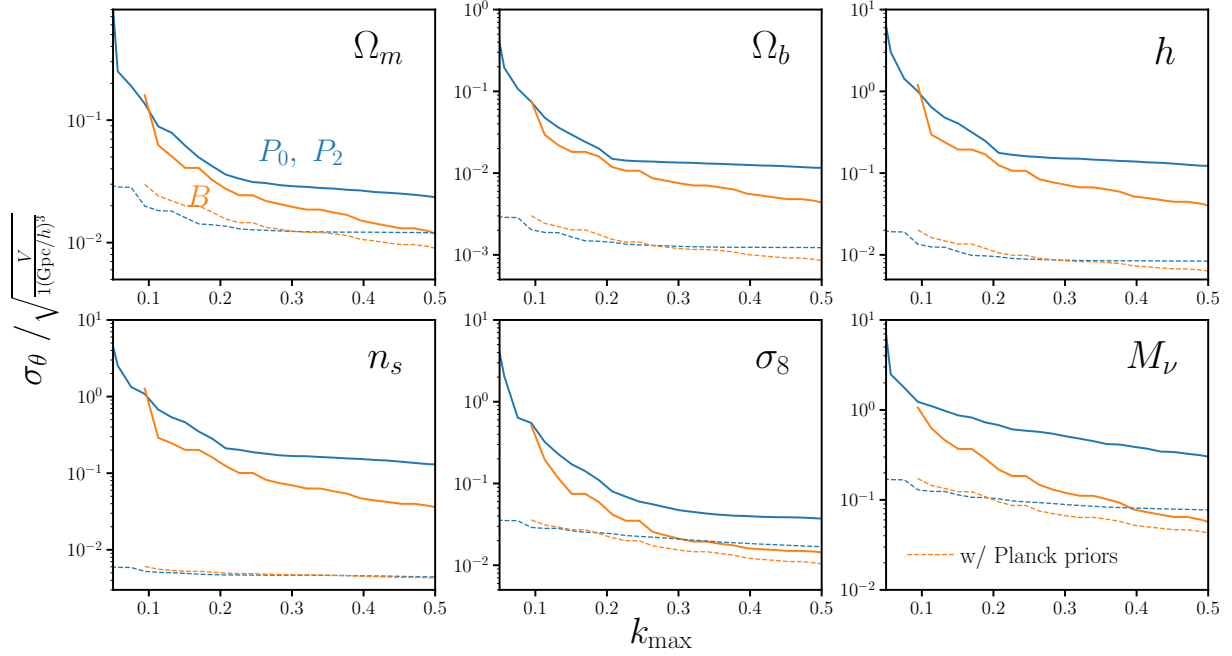


Figure 9. Marginalized 1σ constraints, σ_θ , of the cosmological parameters Ω_m , Ω_b , h , n_s , σ_8 , and M_ν as a function of k_{\max} for the redshift-space halo bispectrum (orange) and power spectrum ($\ell = 0, 2$; blue). The constraints are marginalized over the nuisance parameters b' and M_{\min} in our forecast (Section 4.2). We impose $k_{\max} > 5 k_f$ for P_ℓ and $k_{\max} > 12 k_f$ for B_0 , the k ranges where we have more data bins than number of parameters. We also include σ_θ constraints with *Planck* priors (dotted). Even below $k_{\max} < 0.5 h/\text{Mpc}$, the bispectrum significantly improves cosmological parameter constraints. The improvement, however, is larger for higher k_{\max} . At $k_{\max} = 0.2 h/\text{Mpc}$, the bispectrum improves constraints on Ω_m , Ω_b , h , n_s , σ_8 , and M_ν by factors of ~ 1.3 , 1.1 , 1.3 , 1.5 , 1.7 , and 2.8 over the power spectrum. Even with *Planck* priors, B_0 significantly improves σ_θ for $k_{\max} \gtrsim 0.2 h/\text{Mpc}$. While the constraining power of P_ℓ saturates at $k_{\max} = 0.2 h/\text{Mpc}$, the constraining power of B_0 continues to increase out to $k_{\max} = 0.5 h/\text{Mpc}$.

Our forecasts demonstrate that the bispectrum has significant constraining power beyond the power spectrum in the weakly nonlinear regime ($k > 0.1 h/\text{Mpc}$). This constraining power comes from the bispectrum breaking degeneracies among the cosmological and nuisance parameters. This is evident when we compare the unmarginalized constraints from P_ℓ and B_0 : $1/\sqrt{F_{ii}}$ where F_{ii} is a diagonal element of the Fisher matrix. For $k < 0.4 h/\text{Mpc}$, the unmarginalized constraints from P_ℓ are tighter than those from B_0 . Yet, once we marginalize the constraints over the other parameters, the P_ℓ constraints are degraded and the B_0 constraints are tighter for $k > 0.1 h/\text{Mpc}$. The derivatives, $\partial B_0/\partial\theta_i$, also shed light on how B_0 breaks parameter degeneracies. The parameter degeneracies in the P_ℓ forecasts of Figure 8 are consistent with similarities in the shape and scale dependence of P_ℓ derivatives $\partial P_0/\partial\theta$ and $\partial P_2/\partial\theta$. On the other hand, the B_0 derivatives with respect to the parameters have significant different scale and triangle shape dependencies (Figure 7).

In Figure 7, we mark the triangles with $k_1 = 72k_f$ in the shaded region and triangles with $k_1 = 72k_f$ and $k_2 = 69k_f$ in the darker shaded region. In the shaded region, k_2 increases for triangles to the right while k_3 and θ_{12} , the angle between k_1 and k_2 , increase for the triangles to the right in the darker

shaded region. $\partial \log B_0 / \partial M_\nu$ has little scale dependence for triangles with small θ_{12} ($k_1, k_2 \gg k_3$; folded or squeezed triangles). This is in contrast to the Ω_m , Ω_b , h , and n_s derivatives, which have significant scale dependence at $k < 0.2 h/\text{Mpc}$. For a given k_1 and k_2 , $\partial \log B_0 / \partial M_\nu$ decreases as θ_{12} increases. The Ω_m , h , and n_s derivatives have the opposite θ_{12} dependence. $\partial \log B_0 / \partial \sigma_8$ decreases as θ_{12} increases and also has little scale dependence for folded or squeezed triangles. However, the θ_{12} dependence of $\partial \log B_0 / \partial \sigma_8$ is scale independent unlike $\partial \log B_0 / \partial M_\nu$, which has little θ_{12} dependence on large scales and stronger θ_{12} dependence on small scales. The detailed differences in the shape dependence of $\partial \log B_0 / \partial M_\nu$, $\partial \log B_0 / \partial \sigma_8$, other the other derivatives, highlighted in the shaded regions, ultimately allow B_0 to break parameter degeneracies.

By exploiting the massive number of N -body simulations of the Quijote suite, we present for the first time the total information content of the full redshift-space bispectrum beyond the linear regime. The information content of the bispectrum has previously been examined using perturbation theory. Previous works, for instance, measure the signal-to-noise ratio (SNR) of the bispectrum derived from covariance matrices estimated using perturbation theory (Sefusatti & Scoccimarro 2005; Sefusatti et al. 2006; Chan & Blot 2017). More recently, Chan & Blot (2017), used covariance matrices from ~ 4700 N -body simulations to find that the cumulative SNR of the halo bispectrum is $\sim 30\%$ of the SNR of the halo power spectrum at $k_{\text{max}} \sim 0.1 h/\text{Mpc}$ and increases to $\sim 40\%$ at $k_{\text{max}} \sim 0.35 h/\text{Mpc}$. While these simple SNR measurements cannot be easily compared to Fisher analysis (Repp et al. 2015; Blot et al. 2016), we note that they are roughly consistent with the unmarginalized constraints, which loosely represent the SNRs of the derivatives. Also, when we measure the the halo power spectrum and bispectrum SNRs using our covariance matrices (Figure 6), we find a relation between the SNRs consistent with Chan & Blot (2017). Beyond the k range explored by Chan & Blot (2017), $k_{\text{max}} > 0.35 h/\text{Mpc}$, we find that the SNR of B_0 continues to increase at higher k_{max} in contrast to the P_0 SNR, which saturates at $k_{\text{max}} \sim 0.1 h/\text{Mpc}$. At $k_{\text{max}} = 0.75 h/\text{Mpc}$, the largest k we measure the B_0 , the SNR of B_0 is $\sim 75\%$ of the SNR of P_0 .

Beyond these signal-to-noise calculations, a number of previous works have quantified the information content of the bispectrum (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). While most of these works fix most cosmological parameters and focus solely on forecasting constraints of primordial non-Gaussianity and bias parameters, Sefusatti et al. (2006), Yankelevich & Porciani (2019), and Chudaykin & Ivanov (2019) provide bispectrum forecasts for full sets of cosmological parameters. In Sefusatti et al. (2006), they present likelihood analysis forecasts for ω_d , ω_b , Ω_Λ , n_s , A_s , w , τ . For ΛCDM , with fixed bias parameters, and $k_{\text{max}} = 0.3 h/\text{Mpc}$, they find constraints on Ω_m , Ω_b , h , n_s , and σ_8 from WMAP, P_0 , and B_0 is ~ 1.5 times tighter than constraints from WMAP and P_0 . In comparison, for $k_{\text{max}} = 0.3 h/\text{Mpc}$ our B_0 , constraints are tighter than P_0 constraints by factors of 1.4, 1.7, 2.0, 2.3, 2.0, and 3.9. Both Sefusatti et al. (2006) and our analysis find significantly tighter constraints with the bispectrum. They however include the WMAP likelihood in their forecast and use perturbation theory models, which is limited to larger scales than used here (Scoccimarro et al. 1998, 1999; Sefusatti et al. 2010; Pollack et al. 2012; Gil-Marín et al. 2014; Lazanu et al. 2016; Eggemeier et al. 2019).

Yankelevich & Porciani (2019) present Fisher forecasts for Ω_{cdm} , Ω_{b} , h , n_s , A_s , w_0 , and w_0 for a Euclid-like survey (Laureijs et al. 2011) in 14 non-overlapping redshift bins over $0.65 < z < 2.05$. They use the full redshift-space bispectrum, rather than just the monopole, and a more sophisticated bias expansion than Sefusatti et al. (2006) but use a perturbation theory bispectrum model, which consequently limit their forecast to $k_{\text{max}} = 0.15 \text{ h/Mpc}$. They find similar constraining power on cosmological parameters from B alone as P . They also find that combining the bispectrum with the power spectrum only moderately improves parameter constraints because posterior correlations are similar for P and B . While this seemingly conflicts with the results we present, there are significant differences between our forecasts. For instance, they forecast the Euclid survey (*i.e.* $z > 0.7$), while our forecasts are for $z = 0$. They also forecast the *galaxy* P and B and marginalizes over 56 nuisance parameters (14 z bins each with 3 bias parameters and 1 RSD parameter). They also neglect non-Gaussian contributions to the B covariance matrix, which play a significant role on small scales (Chan & Blot 2017) and may impact the constraints. Despite differences, Yankelevich & Porciani (2019) find that the constraining power of B relative to P_ℓ increases for higher k_{max} , consistent with our forecasts as a function of k_{max} (Figure 9). Also, consistent with their results, for $k_{\text{max}} = 0.15 \text{ h/Mpc}$, we find similar posterior correlations between the P_ℓ and B_0 constraints. At $k_{\text{max}} = 0.5 \text{ h/Mpc}$, however, we find the posterior correlations are no longer similar, which contribute to the constraining power of B_0 (Figure 8).

Finally, Chudaykin & Ivanov (2019) present power spectrum and bispectrum forecasts for ω_{cdm} , ω_{b} , h , n_s , A_s , n_s , and M_ν of a Euclid-like survey in 8 non-overlapping redshift bins over $0.5 < z < 2.1$. They use a one-loop perturbation theory model for the redshift-space power spectrum multipoles ($\ell = 0, 2, 4$) and a tree-level bispectrum monopole model. Also, rather than imposing a k_{max} cutoff to restrict their forecasts to scales where their perturbation theory model can be trusted, they use a theoretical error covariance model approach from Baldauf et al. (2016). They find ~ 1.4 , 1.5 , 1.2 , 1.5 , and 1.3 times tighter constraints from P_ℓ and B_0 than from P_ℓ alone. For M_ν , they find a factor of 1.4 improvement, from 0.038 eV to 0.028 eV . Overall, Chudaykin & Ivanov (2019) find more modest improvements from including B_0 than the improvements we find in our $k_{\text{max}} = 0.5 \text{ h/Mpc}$ B_0 constraints over the P_ℓ constraints.

Among the differences from our forecast, most notably, Chudaykin & Ivanov (2019) marginalize over 64 parameters of their bias model (4 bias and 4 counterterm normalization parameters at each redshift bin). They also include the Alcock-Paczynski (AP) effect for P_ℓ , which significantly improve the P_ℓ parameter constraints (*e.g.* tightens M_ν constraints by $\sim 30\%$). They however, do not include AP effects in B_0 . Furthermore, Chudaykin & Ivanov (2019) use theoretical error covariance to quantify the uncertainty of their perturbation theory models. Because they use the tree-level perturbation theory model for B_0 , theoretical errors for B_0 quickly dominate at $k \gtrsim 0.1 \text{ h/Mpc}$, where the one- and two-loop contribute significantly (*e.g.* Lazanu & Liguori 2018). Hence, their forecasts do not include the constraining power on nonlinear scales. Also different from our analysis, Chudaykin & Ivanov (2019) use an Markov-Chain Monte-Carlo (MCMC) approach, which derives more accurate parameter constraints than our Fisher approach. They, however, neglect non-Gaussian contributions to both the P_ℓ and B_0 covariance matrices and do not include the covariance between P_ℓ and B_0 for

their joint constraints. While their theoretical error covariance, which couples different k -modes, may partly take this into account, we find that neglecting the covariance between P_ℓ and B_0 overestimates constraints, tighter by $\sim 20\%$ for $k_{\max} = 0.2 \ h/\text{Mpc}$. Nonetheless, Chudaykin & Ivanov (2019) and our results both find significant improves in cosmological parameter constraints from including B_0 . In fact, taking the theoretical errors into account, the improvements from B_0 they find are loosely consistent with our results at $k_{\max} \sim 0.2 \ h/\text{Mpc}$.

Various differences between our forecast and previous work prevent more thorough comparisons. However, crucial aspects of our simulation based approach distinguish our forecasts from other works. We present the first bispectrum forecasts for a full set of cosmological parameters using bispectrum measured entirely from N -body simulations. By using the simulations, we go beyond perturbation theory models and accurately model the redshift-space bispectrum to the nonlinear regime. Furthermore, by exploiting the immense number of simulations, we accurately estimate the full high-dimensional covariance matrix of the bispectrum. With these advantages, we present the first forecast of cosmological parameters from the bispectrum down to nonlinear scales and demonstrate the constraining power of the bispectrum for M_ν . Below, we underline a few caveats of our forecasts.

Our forecasts are derived from Fisher matrices. Such forecasts make the assumption that the posterior is approximately Gaussian and, as a result, they underestimate the constraints for posteriors that are highly non-elliptical or asymmetric (Wolz et al. 2012). Fisher matrices also rely on the stability, and in our case also convergence, of numerical derivatives. We examine the stability of the B_0 derivatives with respect to M_ν by comparing the derivatives computed using N -body simulations at three different sets of cosmologies: (1) $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+, M_\nu^{++}, M_\nu^{+++}\}$ (Eq. 12), (2) $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+, M_\nu^{++}\}$, and (3) $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+\}$ (see Appendix B; Figure 13). The derivatives computed using the different set of cosmologies, do not impact the Ω_m , Ω_b , h , n_s , and σ_8 constraints. They do however affect the M_ν constraints; but because P_ℓ and B_0 derivatives are affected by the same factor, the relative improvement of the B_0 M_ν constraint over the P_ℓ constraints is *not* impacted. In addition to the stability, because we use N -body simulations, we test whether the convergence of our covariance matrix and derivatives impact our forecasts by varying the number of simulations used to estimate them: N_{cov} and N_{deriv} , respectively. For N_{cov} , we find $< 5\%$ variation in the Fisher matrix elements, F_{ij} , for $N_{\text{cov}} > 5000$ and $< 1\%$ variation in σ_θ for $N_{\text{cov}} > 12000$. For N_{deriv} , we find $< 5\%$ variation in the F_{ij} elements and $< 5\%$ variation in σ_θ for $N_{\text{deriv}} > 1200$. Since our constraints vary by $< 10\%$ for sufficient N_{cov} and N_{deriv} , the convergence of the covariance matrix and derivatives do not impact our forecasts to the accuracy level of Fisher forecasting. We refer readers to Appendix B for a more details on the robustness of our results to the stability of the derivatives and convergence.

We argue that the constraining power of the bispectrum and its improvement over the power spectrum come from breaking degeneracies among the cosmological parameters. However, numerical noise can impact our forecasts when we invert the Fisher matrix. Since *Planck* constrain $\{\Omega_m, \Omega_b, h, n_s, \sigma_8\}$ tighter than either P_ℓ or B_0 alone, the elements of the *Planck* prior matrix are larger than the elements of P_ℓ and B_0 Fisher matrices. Including *Planck* priors (*i.e.* adding the prior matrix to the Fisher matrix) increases the numerical stability of the matrix inversion. It also reveals whether the bispectrum still improves parameter constraints once we include CMB constraints. With *Planck*

priors and P_ℓ to $k_{\text{max}} = 0.5 \text{ h/Mpc}$, we derive the following constraints: $\sigma_{\Omega_m} = 0.0120$, $\sigma_{\Omega_b} = 0.0012$, $\sigma_h = 0.0084$, $\sigma_{n_s} = 0.0044$, and $\sigma_{\sigma_8} = 0.0169$. Including *Planck* priors expectedly tighten the constraints from P_ℓ . Meanwhile, with *Planck* priors and B_0 to $k_{\text{max}} = 0.5 \text{ h/Mpc}$, we get $\sigma_{\Omega_m} = 0.0090$, $\sigma_{\Omega_b} = 0.0009$, $\sigma_h = 0.0064$, $\sigma_{n_s} = 0.0043$, and $\sigma_{\sigma_8} = 0.0104$, 1.3, 1.4, 1.3, 1.0, and 1.6 times tighter constraints. For M_ν , $\sigma_{M_\nu} = 0.0773 \text{ eV}$ for P_ℓ and $\sigma_{M_\nu} = 0.0430 \text{ eV}$ for B_0 , a factor of 1.8 improvement. Since we find substantial improvements in parameter constraints with the *Planck* prior, the improvement from B_0 are numerically robust. Furthermore, σ_θ as a function of k_{max} with *Planck* priors reveal that while the constraining power of P_ℓ saturates at $k_{\text{max}} = 0.2 \text{ h/Mpc}$, the constraining power of B_0 continues to increase out to $k_{\text{max}} = 0.5 \text{ h/Mpc}$ (dotted; Figure 9).

Our forecasts are derived using the power spectrum and bispectrum in *periodic boxes*. We do not consider a realistic geometry or radial selection function of actual observations from galaxy surveys. A realistic selection function will smooth out the triangle configuration dependence and consequently degrade the constraining power of the bispectrum. In Sefusatti & Scoccimarro (2005), for instance, they find that the signal-to-noise of the bispectrum is significantly reduced once survey geometry is included in their forecast. We also do not account for super-sample covariance (*e.g.* Hamilton et al. 2006; Sefusatti et al. 2006; Takada & Hu 2013; Li et al. 2018), which may also degrade constraints. Survey geometry and super-sample covariance, however, also degrades the signal-to-noise of their power spectrum forecasts. Hence, with the substantial improvement in the M_ν constraints of the bispectrum, even with survey geometry we expect the bispectrum will significantly improve M_ν constraints over the power spectrum.

We include the nuisance parameter M_{min} in our forecasts to address the difference in halo bias and number densities among the Quijote cosmologies. Although we marginalize over M_{min} , this may not fully account for the extra information from \bar{n} and nonlinear bias leaking into the derivatives. To test this, we include extra nuisance parameters, $\{A_{\text{SN}}, B_{\text{SN}}, b_2, \gamma_2\}$, and examine their impact on our forecasts. A_{SN} and B_{SN} account for any \bar{n} dependence that may be introduced from the shot noise correction. They are multiplicative factors of the first and second terms of Eq. 4. b_2 and γ_2 are the quadratic bias and nonlocal bias parameters (Chan et al. 2012; Sheth et al. 2013) to account for information from nonlinear bias. Marginalizing over $\{b', M_{\text{min}}, A_{\text{SN}}, B_{\text{SN}}, b_2, \gamma_2\}$, we obtain the following constraints for B_0 with $k_{\text{max}} = 0.5 \text{ h/Mpc}$: $\sigma_{\Omega_m} = 0.0129$, $\sigma_{\Omega_b} = 0.0044$, $\sigma_h = 0.0404$, $\sigma_{n_s} = 0.0456$, $\sigma_{\sigma_8} = 0.0228$, and $\sigma_{M_\nu} = 0.0575$. While constraints on n_s and σ_8 are broadened from our fiducial forecasts, by 27% and 60%, the other parameters, especially M_ν , are not significantly impacted by marginalizing over the extra nuisance parameters. As another test, we calculate derivatives using halo catalogs from Quijote θ^- and θ^+ cosmologies with fixed \bar{n} . We similarly find no significant impact on the B_0 parameter constraints. Forecasts using additional nuisance parameters and with fixed \bar{n} derivatives, both support the robustness of our forecast. Yet these tests do not ensure that our forecast entirely marginalizes over halo bias.

In this paper, we focus on the halo bispectrum and power spectrum. However, constraints on M_ν will ultimately be derived from the distribution of galaxies. Besides the cosmological parameters, bias and nuisance parameters that allow us to marginalize over galaxy bias need to be incorporated to forecast M_ν and other cosmological parameter constraints for the galaxy bispectrum. Although we

include a *naive* bias model through b' and M_{\min} , and even b_2 and γ_2 in our tests, this is insufficient to describe how galaxies trace matter. A more realistic bias model such as a halo occupation distribution (HOD) model involve extra parameters that describe the distribution of central and satellite galaxies in halos (*e.g.* Zheng et al. 2005; Leauthaud et al. 2012; Tinker et al. 2013; Zentner et al. 2016; Vakili & Hahn 2019). We, therefore, refrain from a more exhaustive investigation of the impact of halo bias on our results and focus quantifying the constraining power of the galaxy bispectrum in the next paper of this series: Hahn et al. (in preparation).

Marginalizing over galaxy bias parameters, will likely reduce the constraining power at high k . Improvements that come from extending to smaller scales will also be diminished by the extra parameters needed to accurately model those scales. Hand et al. (2017), for instance, using a 13 parameter model only find a 15-30% improvement in $f\sigma_8$ when they extend their power spectrum multipole analysis from $0.2 h/\text{Mpc}$ to $0.4 h/\text{Mpc}$. Although we focus on parameter constraints from the bispectrum alone in this work, jointly analyzing the power spectrum and bispectrum will improve matters in this regard. A combined analysis will help constrain bias parameters, further break parameter degeneracies, and improve constraints on cosmological parameters (Sefusatti et al. 2006; Yankelevich & Porciani 2019; Chudaykin & Ivanov 2019; Coulton et al. 2019). Furthermore, we emphasize that the constraints we present is for a $1h^{-1}\text{Gpc}$ box and $\bar{n} \sim 1.56 \times 10^{-4} h^3\text{Mpc}^{-3}$, a substantially smaller volume and lower number density than upcoming surveys. Thus, even if the constraining power at high k is reduced, our forecasts suggest that the bispectrum offers significant improvements over the power spectrum, especially for constraining M_ν .

5. SUMMARY

A precise measurement of M_ν can distinguish between the ‘normal’ and ‘inverted’ neutrino mass hierarchy scenarios and reveal physics beyond the Standard Model of particle physics. The total neutrino mass, through impact the expansion history and the growth of cosmic structure in the Universe, can be measured using cosmological observables (*e.g.* CMB and large-scale structure). In fact, cosmological probes have the potential to more precisely measure M_ν than current and upcoming laboratory experiments. The current tightest cosmological M_ν constraints come from combining CMB data with other cosmological probes. The degeneracy between M_ν and τ , the optical depth of reionization, however, is a major bottleneck for M_ν constraints from CMB data that will *not* be addressed by upcoming ground-based CMB experiments. Measuring the imprint of neutrinos on the 3D clustering of galaxies provides a promising alternative to tightly constrain M_ν , especially with the unprecedented cosmic volumes that will be mapped by upcoming surveys.

Recent developments in simulation based emulation methods have improved the accuracy of theoretical predictions beyond linear scales and address the challenges of unlocking the information content in nonlinear clustering to constrain M_ν . Yet, for power spectrum analyses, the strong degeneracy between the imprints of M_ν and σ_8 poses a serious limitation for constraining M_ν . Information in the nonlinear regime, however, cascade from the power spectrum to higher-order statistics such as the bispectrum. Previous works have demonstrated that the bispectrum has comparable signal-to-noise as the power spectrum on nonlinear scales (Sefusatti & Scoccimarro 2005; Chan & Blot 2017) and that it improves constraints on cosmological parameters (Sefusatti et al. 2006; Yankelevich &

Porciani 2019; Chudaykin & Ivanov 2019). No work to date has quantified the total information content and constraining power of the full bispectrum down to nonlinear scales.

In this work, we examined the effect of massive neutrinos on the redshift-space halo bispectrum using $\sim 23,000$ N -body simulations with massive neutrinos from the HADES and Quijote suite. Using N -body simulations and with such a massive number of them, we directly address key challenges of accurately modeling the bispectrum in the nonlinear regime and estimating its high dimensional covariance matrix. More specifically,

- We examine the imprint of M_ν and σ_8 on the redshift-space halo bispectrum, B_0 , using two sets of HADES simulations. One with massive neutrinos, $M_\nu = 0.0, 0.06, 0.1, \text{ and } 0.15 \text{ eV}$; the other with $M_\nu = 0.0 \text{ eV}$ and matching σ_8^c . M_ν and σ_8 leave distinct imprints on B_0 with significantly different scale and triangle shape dependencies. Thus, we demonstrate that B_0 helps break the M_ν - σ_8 degeneracy found in the power spectrum.
- We quantify the total information content of B_0 using a Fisher matrix forecast of $\{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\}$. With bispectrum measured from 22,000 N -body simulations from the Quijote suite, we are able to derive stable and converged covariance matrix and derivatives for our forecast. Furthermore, our simulation based approach allows us to extend beyond perturbation theory and use all 1898 triangle configurations of the bispectrum down to $k_{\text{max}} = 0.5 \text{ h/Mpc}$.
- For $k_{\text{max}} = 0.5 \text{ h/Mpc}$, the bispectrum produces $\Omega_m, \Omega_b, h, n_s$, and σ_8 constraints 1.9, 2.6, 3.1, 3.6, and 2.6 times tighter than the power spectrum. For M_ν , we derive 1σ constraint of 0.0572 eV — over 5 times tighter than the power spectrum. Even with priors from *Planck*, the bispectrum improves M_ν constraints by a factor of 1.8. These constraints are derived for a $(1 \text{ h}^{-1}\text{Gpc})^3$ box, a substantially smaller volume than upcoming surveys.

While our results clearly showcase the advantages of the bispectrum for more precisely constraining M_ν , as well as the cosmological parameters, a number of assumptions go into our forecast. Fisher matrix forecasts assume that the posterior is approximately Gaussian and, as a result, they overestimate the constraints for posteriors that are highly non-elliptical or asymmetric. Furthermore, our forecasts are derived using the power spectrum and bispectrum in periodic boxes, instead of a realistic geometry or radial selection function of galaxy surveys, which degrades the constraining power. Since we use periodic boundary conditions, we also do not account for super-sample covariance, which in practice results from a non-trivial survey geometry. Lastly, our forecast focuses on the halo bispectrum, while the *galaxy* bispectrum is what is actually measured from observations. We marginalize over a simplistic bias model through b' and M_{min} ; however, such a model is insufficient to flexibly describe how galaxies trace matter. A more realistic bias model involve extra parameters that describe the distribution of central and satellite galaxies. In the next paper of the series, we include HOD parameters in our forecasts and quantify the full information content of the redshift-space galaxy bispectrum.

The constraints from our forecasts are derived for a $(1 \text{ h}^{-1}\text{Gpc})^3$ volume. These constraints roughly scale as $\propto 1/\sqrt{V}$ with volume. Upcoming spectroscopic galaxy surveys will map out vastly larger cosmic volumes: PFS $\sim 9 \text{ h}^{-3}\text{Gpc}^3$ and DESI $\sim 50 \text{ h}^{-3}\text{Gpc}^3$ (Takada et al. 2014; Collaboration et al.

2016). Euclid and WFIRST, space-based surveys, will expand these volumes to higher redshifts. A key science goal in these surveys will be constraining M_ν . Scaling our 0.0572 eV M_ν constraint from the bispectrum to the survey volumes gives 0.0191 eV and 0.0081 eV 1σ constraints for PFS and DESI. Galaxy samples in these surveys will also have significantly higher number densities than the halos used in our forecast: *e.g.* the DESI BGS at $z \sim 0.3$, DESI LRG, and PFS at $z \sim 1.3$ will have $\sim 20, 3$, and $5\times$ higher number densities, respectively. With such precision, we would detect M_ν by $> 3\sigma$ and distinguish between the normal and inverted neutrino mass hierarchy scenarios by $> 2\sigma$ from the bispectrum alone. Such naive projections, however, should be taken with more than a grain of salt. Nevertheless, the substantial improvement we find in constraints with the bispectrum over the power spectrum, strongly advocate for analyzing future surveys with more than the power spectrum. Our results demonstrate the potential of the bispectrum to tightly constrain M_ν with unprecedented precision.

ACKNOWLEDGEMENTS

It's a pleasure to thank Michael R. Blanton, William Coulton, Enea Di Dio, Daniel Eisenstein, Simone Ferraro, Simon Foreman, Patrick McDonald, Shirley Ho, Emmanuel Schaan, Uroš Seljak, Zachary Slepian, David N. Spergel, Licia Verde, and Benjamin D. Wandelt for valuable discussions and comments. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under contract No. DE-AC02-05CH11231. This project 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.

APPENDIX

A. REDSHIFT-SPACE BISPECTRUM

In this work, we present how the redshift-space bispectrum helps break degeneracies among cosmological parameters, especially between M_ν and σ_8 , and improve constraints on M_ν as well as the other cosmological parameters. We measure the redshift-space bispectrum monopole, \hat{B}_0 , of the HADES and Quijote simulations (Section 2) using an FFT based estimator similar to the ones in Sefusatti (2005), Scoccimarro (2015), and Sefusatti et al. (2016) (see Section 3 for details). We use triangle configurations defined by k_1 , k_2 , and k_3 bins of width $\Delta k = 3k_f = 0.01885 h/\text{Mpc}$. For $k_{\text{max}} = 0.5 h/\text{Mpc}$ we have \hat{B}_0 for 1898 triangle configurations. For $k_{\text{max}} = 0.4, 0.3, 0.2$, and $0.1 h/\text{Mpc}$, we have 1056, 428, 150, and 28 triangle configurations respectively. The number of triangle configurations sharply increases for higher k_{max} . This contributes to the significant increase in constraining power as we go to increasingly nonlinear scales (Figure 9). We highlight the triangle configurations at different k_{max} for the \hat{B}_0 of the Quijote simulation at the fiducial cosmology in the top panel of Figure 10. We mark the configurations within $k_{\text{max}} = 0.1, 0.2, 0.3, 0.4$, and $0.5 h/\text{Mpc}$ in purple, red, green, orange, and blue respectively.

In Figure 10, as well as in Figures 3, 5, 7, we order the triangle configurations by looping through k_3 in the inner most loop and k_1 in the outer most loop such that $k_1 \geq k_2 \geq k_3$. This ordering is different from the ordering in Gil-Marín et al. (2017), which loops through k_3 in the inner most loop

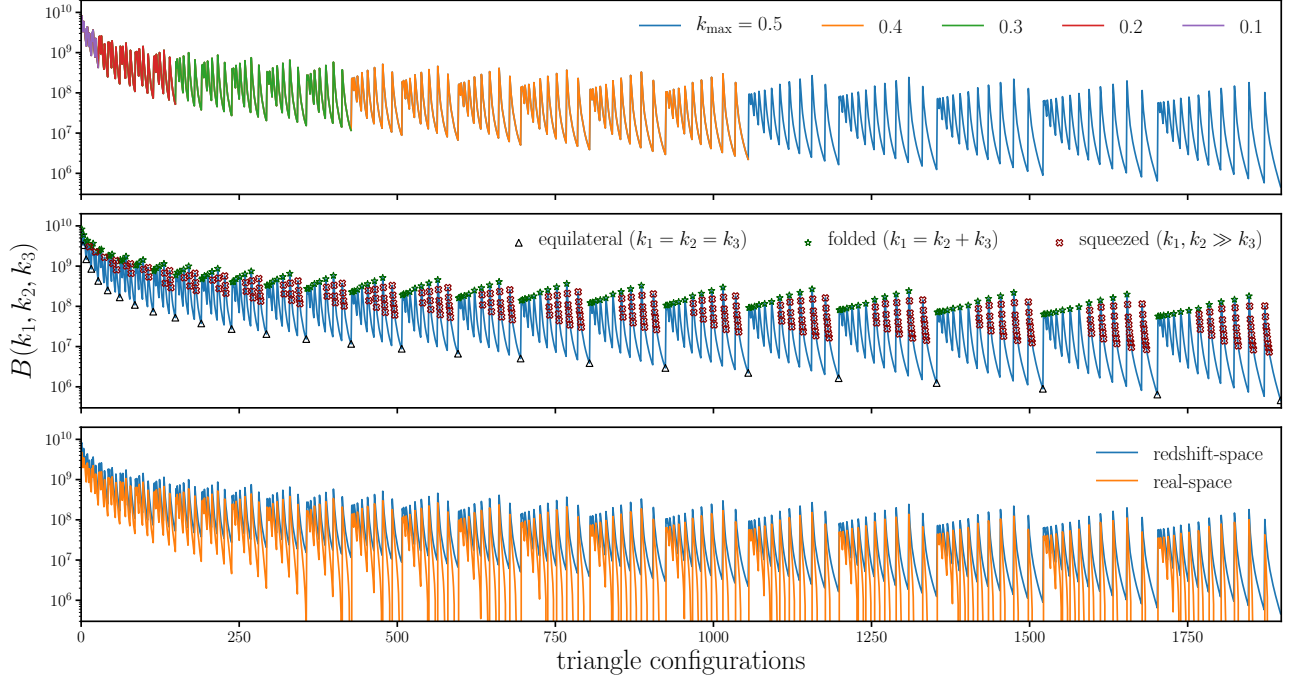


Figure 10. *Top:* We highlight the triangle configurations of the bispectrum that fall within: $k_{\max} < 0.1$ (purple), 0.2 (red), 0.3 (green), 0.4 (orange), and 0.5 h/Mpc (blue). The number of triangle configurations increase significantly at higher k_{\max} , which contributes to the significant constraining power on nonlinear scales. *Center:* We mark triangle configurations with three different shapes: equilateral (black triangle), folded (green star), and squeezed (red cross) triangles. Equilateral triangles have $k_1 = k_2 = k_3$; folded triangles have $k_1 = k_2 + k_3$; squeezed triangles have $k_1, k_2 \gg k_3$, above we mark configurations with $k_1, k_2 > 2.4k_3$ as squeezed. *Bottom:* Comparison of the real (orange) and redshift-space (blue) halo bispectrum for fiducial Quijote simulations. The redshift-space \hat{B}_0 has a significantly higher overall amplitude than the real-space \hat{B}_0 , with a shape and k_{\max} dependence. As a result, the redshift-space \hat{B}_0 has a higher signal-to-noise and more constraining power than the real-space \hat{B}_0 .

and k_1 in the outer most increasing loop but with $k_1 \leq k_2 \leq k_3$. As the top panel demonstrates, our ordering clearly reflects the k_{\max} range of the configurations. Furthermore, the repeated configuration sequences in our ordering cycles through all available triangle configurations for a given k_1 . In the center panel of Figure 10, we mark equilateral (black triangle), folded (green star), and squeezed (red cross) triangle configurations. Equilateral triangles have $k_1 = k_2 = k_3$, folded triangles have $k_1 = k_2 + k_3$, and squeezed triangles have $k_1, k_2 \gg k_3$ — $k_1, k_2 > 2.4k_3$ in Figure 10. These shapes correspond to the three vertices of the bispectrum shape plots of Figures 2 and 4.

Lastly, in the bottom panel of Figure 10 we compare the redshift-space bispectrum (blue) to the real-space bispectrum of the same Quijote simulations at the fiducial cosmology (orange). Overall, the redshift-space \hat{B}_0 has a higher amplitude than the real-space \hat{B}_0 with a significant triangle shape dependence: equilateral triangles have the largest difference in amplitude while folded triangles have the smallest. The relative amplitude difference also depends significantly on k_{\max} where the difference

is larger for configurations with higher k_1 . With its higher amplitude, the redshift-space \hat{B}_0 has a higher signal-to-noise and more constraining power than the real-space \hat{B}_0 .

B. FISHER FORECASTS USING N -BODY SIMULATIONS

The two key elements in calculating the Fisher matrices in our bispectrum forecasts are the bispectrum covariance matrix (\mathbf{C} ; Figure 6) and the derivatives of the bispectrum along the cosmological and nuisance parameters ($\partial B_0/\partial\theta$; Figure 7). We compute both these elements directly using the N -body simulations of the Quijote suite (Section 2). This exploits the accuracy of N -body simulations in the nonlinear regime and allows us to accurately quantify the constraining power of the bispectrum beyond perturbation theory models. However, we must ensure that both \mathbf{C} and $\partial B_0/\partial\theta$ have converged and that numerical effects do not introduce any biases that impact our results. Below, we test the convergence of \mathbf{C} and $\partial B_0/\partial\theta$ and discuss some of the subtleties and caveats of our $\partial B_0/\partial\theta$ calculations.

We use 15,000 Quijote N -body simulations at the fiducial cosmology to estimate \mathbf{C} . This is a *significantly* larger number of simulations than any previous bispectrum analyses; however, we also consider 1898 triangle configurations out to $k_{\text{max}} = 0.5 \ h/\text{Mpc}$. For reference, Gil-Marín et al. (2017) recently used 2048 simulations to estimate the covariance matrix of the bispectrum with 825 configurations. To check the convergence of covariance matrix, we vary the number of simulations used to estimate \mathbf{C} , N_{cov} , and determine whether this significantly impacts the elements of the Fisher matrix, F_{ij} , or the final marginalized Fisher parameter constraints, σ_θ . In the left panel of Figure 11, we present the ratio between $F_{ij}(N_{\text{cov}})$, F_{ij} derived from \mathbf{C} calculated with N_{cov} simulations, and $F_{ij}(N_{\text{cov}} = 15,000)$ for all 36 elements of the Fisher matrix. We shade $\pm 5\%$ deviations in the ratios for reference. The F_{ij} elements vary by $\lesssim 5\%$ for $N_{\text{cov}} > 5000$ and $\lesssim 1\%$ for $N_{\text{cov}} > 10,000$. Next, we present ratio between $\sigma_\theta(N_{\text{cov}})$, the marginalized 1σ constraints for $\{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\}$ derived from \mathbf{C} calculated with N_{cov} simulations, and $\sigma_\theta(N_{\text{cov}} = 15,000)$ in the left panel of Figure 12. The constraints vary by $\lesssim 5\%$ for $N_{\text{cov}} > 5000$ and $\lesssim 1\%$ for $N_{\text{cov}} > 12000$. Hence, $N_{\text{cov}} = 15,000$ is sufficient to accurately estimate \mathbf{C} and its convergence does not impact our forecasts.

We estimate $\partial B_0/\partial\theta$ using $N_{\text{deriv}} = 1,500$ N -body simulations at 13 different cosmologies listed in Table 1. To check the convergence of $\partial B_0/\partial\theta$ and its impact on our results, we examine the ratio between $F_{ij}(N_{\text{deriv}})$, the Fisher matrix element derived from $\partial B_0/\partial\theta$ calculated with N_{deriv} simulations, and $F_{ij}(N_{\text{deriv}} = 1,500)$ for all 36 elements of the Fisher matrix in the right panel of Figure 11. For $N_{\text{deriv}} > 1000$, F_{ij} elements vary by $\lesssim 5\%$. Next, we present the ratio between $\sigma_\theta(N_{\text{deriv}})$, the marginalized 1σ constraints for $\{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\}$ derived from $\partial B_0/\partial\theta$ calculated with N_{deriv} simulations, and $\sigma_\theta(N_{\text{deriv}} = 1,500)$ in the right panel of Figure 12. Unlike $\sigma_\theta(N_{\text{cov}})$, $\sigma_\theta(N_{\text{deriv}})$ depend significantly on θ . For instance, σ_{σ_8} and σ_{Ω_m} vary by $\lesssim 10\%$ for $N_{\text{deriv}} > 600$ and $\lesssim 1\%$ for $N_{\text{deriv}} > 1200$. σ_θ for the other parameter vary significantly more. Nonetheless, for $N_{\text{deriv}} > 800$ and 1200 they vary by $\lesssim 10$ and 5%, respectively.

For $\Omega_m, \Omega_b, h, n_s, \sigma_8$, and also the nuisance parameter M_{lim} we estimate $\partial B_0/\partial\theta$ using a centered difference approximation (Eq. 11). However, for M_ν we cannot have values below 0.0 eV and, thus, cannot estimate the derivative with the same method. If we use the forward difference approximation,

$$\frac{\partial \bar{B}_0}{\partial M_\nu} \approx \frac{\bar{B}_0(\theta_{\text{fid}}^{\text{ZA}} + \delta M_\nu) - \bar{B}_0(M_\nu^{\text{fid}})}{\delta M_\nu}, \quad (\text{B1})$$

the error goes as $\mathcal{O}(\delta M_\nu)$. Instead, we use a finite difference approximation with the Quijote simulations at M_ν^+ , M_ν^{++} , M_ν^{+++} , and the fiducial cosmology, a $\mathcal{O}(\delta M_\nu^2)$ order approximation (Eq. 12). We can also use a finite difference approximation with simulations at M_ν^+ , M_ν^{++} , and the fiducial cosmology. We examine the stability of $\partial \log B_0 / \partial M_\nu$ (right) and $\partial \log P_\ell / \partial M_\nu$ (left) by comparing the derivatives computing using simulations at $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+, M_\nu^{++}, M_\nu^{+++}\}$ (blue), $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+, M_\nu^{++}\}$ (orange), and $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+\}$ (green) in Figure 13. The three $\partial \log B_0 / \partial M_\nu$ approximations differ from one another by $\sim 10\%$ with Eq. 12 producing the largest estimate for both P_ℓ and B_0 . If we use the $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+, M_\nu^{++}\}$ derivative and $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+\}$ derivative instead of Eq. 12 for our Fisher forecasts, the marginalized constraint on M_ν for $k_{\text{max}} = 0.5 \text{ h/Mpc}$ increases to 0.390 and 0.682 eV for P_ℓ and 0.0754 and 0.1354 eV for B_0 . Compared to our $\sigma_{M_\nu} = 0.0572 \text{ eV}$ B_0 forecast, these correspond to a ~ 30 and 130% relative increase. While the derivative estimated from $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+\}$ significantly impact the forecasts, we emphasize that this is a $\mathcal{O}(\delta M_\nu)$ approximation, unlike the other $\mathcal{O}(\delta M_\nu^2)$ approximations. In fact, the $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+\}$ derivatives are better $\mathcal{O}(\delta M_\nu^2)$ estimates for $\partial P_\ell / \partial M_\nu$ and $\partial B_0 / \partial M_\nu$ at 0.05 eV. When we compare the derivative of the linear theory power spectrum, $P^{(\text{LT})}$, we find that $\partial P^{(\text{LT})} / \partial M_\nu$ at 0.0 is larger than at 0.05 eV. Hence, the differences between the $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+\}$ derivatives and the other derivatives are not solely due to numerical stability. Moreover, because the discrepancies in the derivative propagate similarly to both P_ℓ and B_0 forecasts, the relative improvement of B_0 over P_ℓ remains roughly the same. Hence, we conclude that the derivatives with respect to M_ν are sufficiently stable and robust for our Fisher forecasts.

REFERENCES

- Abazajian, K. N., Adshead, P., Ahmed, Z., et al. 2016, [arXiv:1610.02743 \[astro-ph, physics:gr-qc, physics:hep-ph, physics:hep-th\]](#), [arXiv:1610.02743 \[astro-ph, physics:gr-qc, physics:hep-ph, physics:hep-th\]](#)
- Adamek, J., Durrer, R., & Kunz, M. 2017, [arXiv:1707.06938 \[astro-ph, physics:gr-qc\]](#), [arXiv:1707.06938 \[astro-ph, physics:gr-qc\]](#)
- Ade, P. a. R., Aghanim, N., Arnaud, M., et al. 2016, *Astronomy & Astrophysics*, 594, A13
- Agarwal, S., & Feldman, H. A. 2011, *Monthly Notices of the Royal Astronomical Society*, 410, 1647
- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *The Astrophysical Journal Supplement Series*, 219, 12
- Allison, R., Caucal, P., Calabrese, E., Dunkley, J., & Louis, T. 2015, *Physical Review D*, 92, 123535
- Archidiacono, M., Brinckmann, T., Lesgourgues, J., & Poulin, V. 2017, *Journal of Cosmology and Astro-Particle Physics*, 2017, 052
- Audren, B., Lesgourgues, J., Bird, S., Haehnelt, M. G., & Viel, M. 2013, *Journal of Cosmology and Astro-Particle Physics*, 2013, 026
- Baldauf, T., Mirbabayi, M., Simonović, M., & Zaldarriaga, M. 2016
- Banerjee, A., & Dalal, N. 2016, *Journal of Cosmology and Astro-Particle Physics*, 2016, 015
- Bird, S., Viel, M., & Haehnelt, M. G. 2012, *Monthly Notices of the Royal Astronomical Society*, 420, 2551
- Blot, L., Corasaniti, P. S., Amendola, L., & Kitching, T. D. 2016, *Monthly Notices of the Royal Astronomical Society*, 458, 4462
- Bonn, J., Eitel, K., Glück, F., et al. 2011, *Physics Letters B*, 703, 310

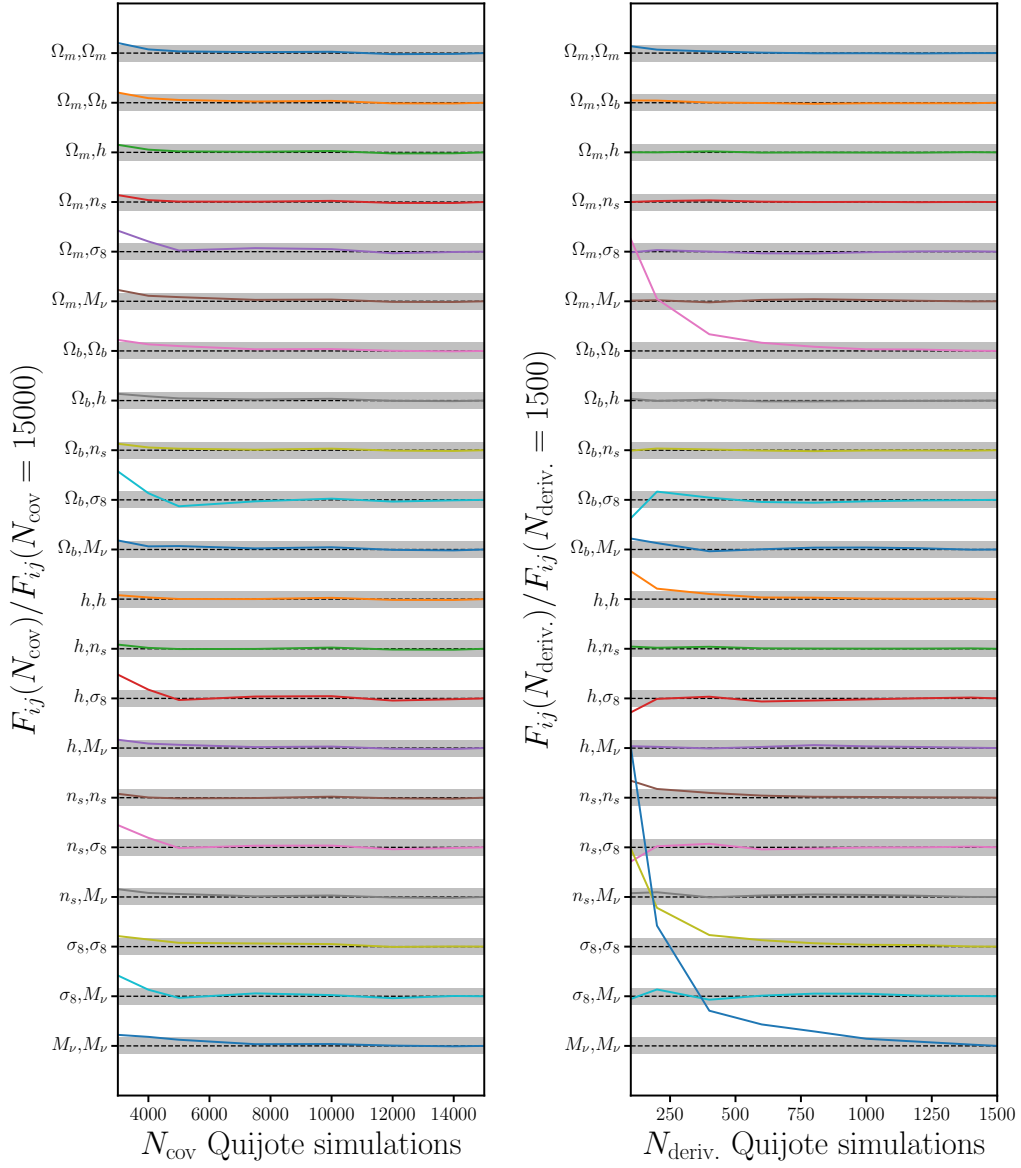


Figure 11. *Left:* The convergence of all 36 Fisher matrix elements, F_{ij} , as a function of N_{cov} , the number of N -body simulations used to estimate the covariance matrix, \mathbf{C} . We present the ratio between $F_{ij}(N_{\text{cov}})$, F_{ij} derived from \mathbf{C} calculated with N_{cov} simulations, and $F_{ij}(N_{\text{cov}} = 15,000)$. We mark $\pm 5\%$ deviations in the ratios with the shaded regions for reference. All F_{ij} elements vary by $\lesssim 5\%$ for $N_{\text{cov}} > 5,000$ and $\lesssim 1\%$ for $N_{\text{cov}} > 10,000$. *Right:* The convergence of F_{ij} as a function of N_{deriv} , the number of N -body simulations used to estimate the derivatives, $\partial B_0/\partial\theta$. We plot the ratios between $F_{ij}(N_{\text{deriv}})$ and $F_{ij}(N_{\text{deriv}} = 1,500)$. All F_{ij} elements vary by $\lesssim 5\%$ for $N_{\text{deriv}} > 1,000$. Hence, $N_{\text{cov}} = 15,000$ and $N_{\text{deriv}} = 1,500$ are sufficient and the convergence of \mathbf{C} and $\partial B_0/\partial\theta$ does not impact F_{ij} .

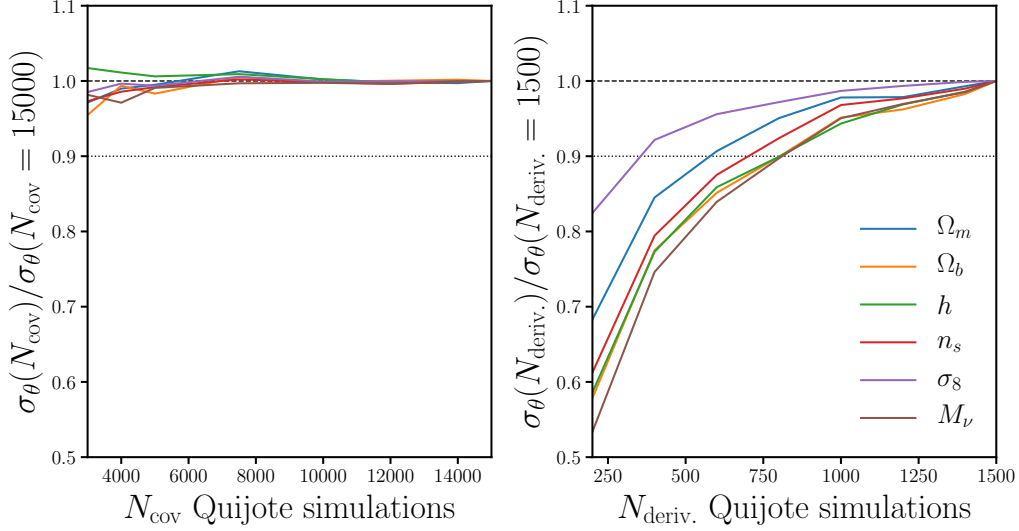


Figure 12. *Left:* The convergence of the marginalized 1σ constraints for $\{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\}, \sigma_\theta$, as a function of N_{cov} , the number of Quijote simulations used to estimate \mathbf{C} . We plot the ratio between $\sigma_\theta(N_{\text{cov}})$, derived from \mathbf{C} with N_{cov} simulations, and $\sigma_\theta(N_{\text{cov}} = 15,000)$. σ_θ vary by $\lesssim 5$ and 1% for $N_{\text{cov}} > 5000$ and 12,000, respectively. *Right:* The convergence of σ_θ as a function of $N_{\text{deriv.}}$, the number of simulated used to estimate $\partial B_0/\partial M_\nu$. We plot the ratio between $\sigma_\theta(N_{\text{deriv.}})$, derived from $\partial B/\partial M_\nu$ with $N_{\text{deriv.}}$ simulations, and $\sigma_\theta(N_{\text{deriv.}} = 1,500)$. Although $\sigma_\theta(N_{\text{deriv.}})/\sigma_\theta(N_{\text{deriv.}} = 1,500)$ vary among the parameters, the ratio vary by $\lesssim 10$ and 5% for $N_{\text{deriv.}} > 800$ and 1200, respectively. Hence, *we have a sufficient number of simulations to estimate \mathbf{C} and the derivatives of the bipsectrum and our forecasts are robust to their convergence.*

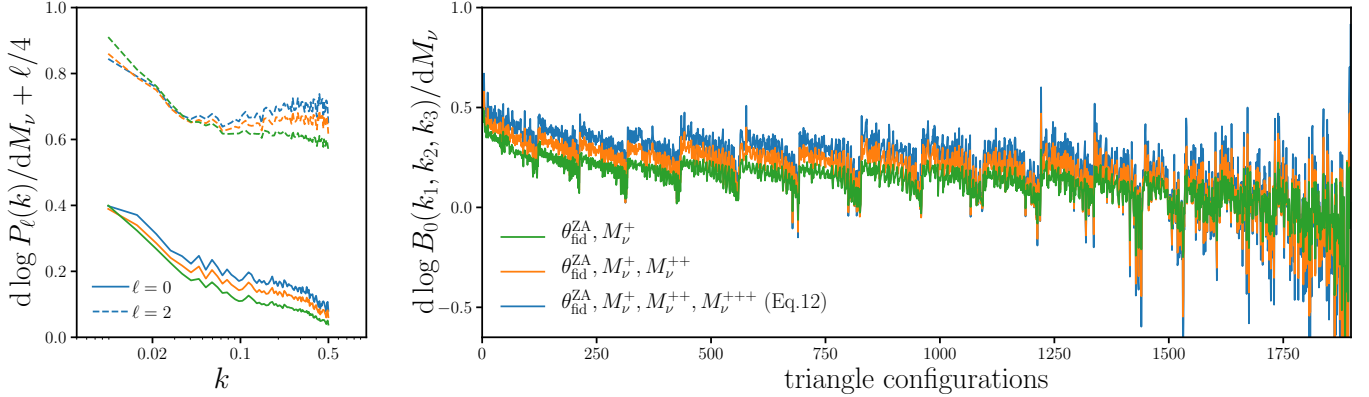


Figure 13. Comparison of $\partial \log B_0(k_1, k_2, k_3)/\partial M_\nu$ (right) and $\partial \log P_\ell(k)/\partial M_\nu$ (left), computed using Quijote simulations at $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+, M_\nu^{++}, M_\nu^{+++}\}$ (blue), $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+, M_\nu^{++}\}$ (orange), and $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+\}$ (green). The derivative approximations differ from one another by $\sim 10\%$ with Eq. 12 producing the largest estimate for both P_ℓ and B_0 . Using the $\{\theta_{\text{fid}}^{\text{ZA}}, M_\nu^+, M_\nu^{++}\}$ derivatives instead of the Eq. 12 derivatives, increases the marginalized constraint on M_ν by $\sim 30\%$. However, the differences in the derivatives propagate similarly to the P_ℓ and B_0 forecasts so the relative improvement of B_0 over P_ℓ remains the same. Hence, we conclude that *the derivatives with respect to M_ν are sufficiently stable and robust for our Fisher forecasts.*

- Boyle, A., & Komatsu, E. 2018, *Journal of Cosmology and Astro-Particle Physics*, 2018, 035
- Brandbyge, J., Hannestad, S., Haugbølle, T., & Thomsen, B. 2008, *Journal of Cosmology and Astro-Particle Physics*, 08, 020
- Carron, J. 2013, *Astronomy & Astrophysics*, 551, A88
- Castorina, E., Carbone, C., Bel, J., Sefusatti, E., & Dolag, K. 2015, *Journal of Cosmology and Astro-Particle Physics*, 2015, 043
- Castorina, E., Sefusatti, E., Sheth, R. K., Villaescusa-Navarro, F., & Viel, M. 2014, *Journal of Cosmology and Astro-Particle Physics*, 02, 049
- Chan, K. C., & Blot, L. 2017, *Physical Review D*, 96, arXiv:1610.06585
- Chan, K. C., Scoccimarro, R., & Sheth, R. K. 2012, *Physical Review D*, 85, 083509
- Chudaykin, A., & Ivanov, M. M. 2019, arXiv:1907.06666 [astro-ph, physics:hep-ph], arXiv:1907.06666 [astro-ph, physics:hep-ph]
- Collaboration, D., Aghamousa, A., Aguilar, J., et al. 2016, arXiv:1611.00036 [astro-ph], arXiv:1611.00036 [astro-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
- Dodelson, S. 2003, *Modern Cosmology*
- Drexlin, G., Hannen, V., Mertens, S., & Weinheimer, C. 2013, *Advances in High Energy Physics*
- Eggemeier, A., Scoccimarro, R., & Smith, R. E. 2019, *Physical Review D*, 99, 123514
- Emberson, J. D., Yu, H.-R., Inman, D., et al. 2017, *Research in Astronomy and Astrophysics*, 17, 085
- Euclid Collaboration, Knabenhans, M., Stadel, J., et al. 2018, arXiv:1809.04695 [astro-ph], arXiv:1809.04695 [astro-ph]
- Font-Ribera, A., McDonald, P., Mostek, N., et al. 2014, *Journal of Cosmology and Astro-Particle Physics*, 05, 023
- Forero, D. V., Tórtola, M., & Valle, J. W. F. 2014, *Physical Review D*, 90, 093006
- Gerbino, M. 2018, arXiv e-prints, arXiv:1803.11545
- Gil-Marín, H., Percival, W. J., Verde, L., et al. 2017, *Monthly Notices of the Royal Astronomical Society*, 465, 1757
- Gil-Marín, H., Wagner, C., Noreña, J., Verde, L., & Percival, W. 2014, *Journal of Cosmology and Astro-Particle Physics*, 12, 029
- Gonzalez-Garcia, M. C., Maltoni, M., & Schwetz, T. 2016, *Nuclear Physics B*, 908, 199
- Hamilton, A. J. S., Rimes, C. D., & Scoccimarro, R. 2006, *Monthly Notices of the Royal Astronomical Society*, 371, 1188
- Hand, N., Seljak, U., Beutler, F., & Vlah, Z. 2017, arXiv:1706.02362 [astro-ph], arXiv:1706.02362 [astro-ph], rEAD
- Heavens, A. 2009, arXiv:0906.0664 [astro-ph], arXiv:0906.0664 [astro-ph]
- Heitmann, K., Higdon, D., White, M., et al. 2009, *The Astrophysical Journal*, 705, 156
- Hockney, R. W., & Eastwood, J. W. 1981, *Computer Simulation Using Particles*
- Ichiki, K., & Takada, M. 2012, *Physical Review D*, 85, 063521
- Jungman, G., Kamionkowski, M., Kosowsky, A., & Spergel, D. N. 1996, *Physical Review D*, 54, 1332
- Karagiannis, D., Lazanu, A., Liguori, M., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 478, 1341
- Kwan, J., Heitmann, K., Habib, S., et al. 2015, *The Astrophysical Journal*, 810, 35
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints, arXiv:1110.3193
- Lazanu, A., Giannantonio, T., Schmittfull, M., & Shellard, E. P. S. 2016, *Physical Review D*, 93, 083517
- 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
- Lesgourgues, J., & Pastor, S. 2012
- . 2014
- Li, Y., Schmittfull, M., & Seljak, U. 2018, *Journal of Cosmology and Astro-Particle Physics*, 2018, 022
- Liu, A., Pritchard, J. R., Allison, R., et al. 2016, *Physical Review D*, 93, 043013
- LoVerde, M. 2014, *Physical Review D*, 90, 083518

- Marulli, F., Carbone, C., Viel, M., Moscardini, L., & Cimatti, A. 2011, *Monthly Notices of the Royal Astronomical Society*, 418, 346
- McClintock, T., Rozo, E., Becker, M. R., et al. 2018, arXiv:1804.05866 [astro-ph], arXiv:1804.05866 [astro-ph]
- Petracca, F., Marulli, F., Moscardini, L., et al. 2016, *Monthly Notices of the Royal Astronomical Society*, 462, 4208
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv:1807.06209 [astro-ph], arXiv:1807.06209 [astro-ph]
- Pollack, J. E., Smith, R. E., & Porciani, C. 2012, *Monthly Notices of the Royal Astronomical Society*, 420, 3469
- Repp, A., Szapudi, I., Carron, J., & Wolk, M. 2015, *Monthly Notices of the Royal Astronomical Society*, 454, 3533
- Ruggeri, R., Castorina, E., Carbone, C., & Sefusatti, E. 2018, *Journal of Cosmology and Astroparticle Physics*, 2018, 003
- Saito, S., Takada, M., & Taruya, A. 2008, *Physical Review Letters*, 100, 191301
- . 2009, *Physical Review D*, 80, 083528
- Sartoris, B., Biviano, A., Fedeli, C., et al. 2016, *Monthly Notices of the Royal Astronomical Society*, 459, 1764
- Scoccimarro, R. 2015, *Physical Review D*, 92, arXiv:1506.02729
- Scoccimarro, R., Colombi, S., Fry, J. N., et al. 1998, *The Astrophysical Journal*, 496, 586
- Scoccimarro, R., Couchman, H. M. P., & Frieman, J. A. 1999, *The Astrophysical Journal*, 517, 531
- Scoccimarro, R., Sefusatti, E., & Zaldarriaga, M. 2004, *Physical Review D*, 69, 103513
- Sefusatti, E. 2005, Ph.D. Thesis, 2114
- Sefusatti, E., Crocce, M., & Desjacques, V. 2010, *Monthly Notices of the Royal Astronomical Society*, 406, 1014
- Sefusatti, E., Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, *Physical Review D*, 74, arXiv:astro-ph/0604505
- Sefusatti, E., Crocce, M., Scoccimarro, R., & Couchman, H. M. P. 2016, *Monthly Notices of the Royal Astronomical Society*, 460, 3624
- Sefusatti, E., & Komatsu, E. 2007, *Physical Review D*, 76, 083004
- Sefusatti, E., & Scoccimarro, R. 2005, *Physical Review D*, 71, arXiv:astro-ph/0412626
- Sheth, R. K., Chan, K. C., & Scoccimarro, R. 2013, *Physical Review D*, 87, 083002
- 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
- Takada, M., & Hu, W. 2013, *Physical Review D*, 87, 123504
- Takada, M., Ellis, R. S., Chiba, M., et al. 2014, *Publications of the Astronomical Society of Japan*, 66, R1
- Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, *The Astrophysical Journal*, 480, 22
- 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
- Upadhye, A., Kwan, J., Pope, A., et al. 2016, *Physical Review D*, 93, 063515
- Vakili, M., & Hahn, C. 2019, *The Astrophysical Journal*, 872, 115
- Verde, L. 2010, arXiv:0911.3105 [astro-ph], 800, 147
- Viel, M., Haehnelt, M. G., & Springel, V. 2010, *Journal of Cosmology and Astro-Particle Physics*, 06, 015
- Villaescusa-Navarro, F., Banerjee, A., Dalal, N., et al. 2018, *The Astrophysical Journal*, 861, 53
- Villaescusa-Navarro, F., Bird, S., Peña-Garay, C., & Viel, M. 2013, *Journal of Cosmology and Astro-Particle Physics*, 2013, 019
- Villaescusa-Navarro, F., Marulli, F., Viel, M., et al. 2014, *Journal of Cosmology and Astro-Particle Physics*, 03, 011
- Villaescusa-Navarro, F., Hahn, C., Massara, E., et al. 2019, arXiv:1909.05273 [astro-ph], arXiv:1909.05273 [astro-ph]
- Watts, D. J., Wang, B., Ali, A., et al. 2018, *The Astrophysical Journal*, 863, 121
- Wibking, B. D., Salcedo, A. N., Weinberg, D. H., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 484, 989
- Wolz, L., Kilbinger, M., Weller, J., & Giannantonio, T. 2012, *Journal of Cosmology and Astroparticle Physics*, 2012, 009
- Wong, Y. Y. Y. 2008, *Journal of Cosmology and Astroparticle Physics*, 2008, 035

- 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. 2018, [arXiv:1804.05867 \[astro-ph\]](#), [arXiv:1804.05867 \[astro-ph\]](#)
- Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, [The Astrophysical Journal](#), **633**, 791