Constraining M_ν with the Bispectrum I: Breaking Parameter Degeneracies

ChangHoon Hahn,[1, 2, *] Francisco Villaescusa-Navarro,[3] Emanuele Castorina,[2, 1] and Roman Scoccimarro[4]

[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*

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 degerancies 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 over 23000 N-body simulations of the Quijote suite, we quantify for the first time the full information content of the redshift-space halo bispectrum using a Fisher matrix forecast of $\{\Omega_{\rm m}, \Omega_{\rm b}, h, n_s, \sigma_8, M_\nu\}$ down to nonlinear scales. For $k_{\rm max}$=0.5 h/Mpc, the bispectrum constrains $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, and σ_8 3.1, 4.1, 5.1, 5.9, and 3.3 times tighter than the power spectrum. For M_ν, the bispectrum improves the 1σ constraint from 0.1962 to 0.0342 eV — 6 times tigher 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 bias parameters to the next paper of the series, these constraints are derived for a $(1\ h^{-1}{\rm Gpc})^3$ 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: —

1. INTRODUCTION

* hahn.changhoon@gmail.com

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). These upper bound alone are not insufficient to distinguish between the mass hiearchies. 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); 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 (Collaboration et al. 2018b). Adding the Baryon Acoustic Oscillation (BAO) to the *Planck* likelihood breaks geometrical degeneracies (M_ν, h, $\Omega_{\rm m}$) and significantly tightens the upper bound to $M_\nu < 0.16$ eV. CMB lensing further tightens the bound fruther to $M_\nu < 0.13$ eV, though not as significantly. Combining these probes, however, introduce 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*: $\tau \sim 0.06$. However, 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). Meanwhile, proposed future space-based experiments such as LiteBIRD[1] and COrE[2], which have the 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

---

[1] http://litebird.jp/eng/
[2] http://www.core-mission.org/

neutrinos on 3D clustering of galaxies can be measured to constrain M_ν and with the sheer cosmic volumes mapped by upcoming surveys, *e.g.* DESI[3], PFS[4], EUCLID[5], and WFIRST[6], 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 conjuction, 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; Collaboration et al. 2018a; 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 only by $< 1\%$. The strong M_ν – σ_8 degeneracy poses a serious limitation of constraining M_ν with the power spectrum. However, information in the nonlinear regime cascade from the power spectrum to higher-order statistics— *e.g.* the bispectrum. In fact, the bispectrum has comparable a 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 analysis, Sefusatti et al. (2006) and Yankelevich & Porciani (2019) have shown that including the bispectrum in analyses improves constraints on cosmological parameters. Ruggeri et al. (2018) presented the first measurement of the bispectrum for N-body simulations with massive neutrions. Yet, no work to date has quantified the full information content and constraining power of the entire 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 more than 23000 N-body simulations with massive neutrinos. We first demonstrate that the bispectrum helps break the M_ν–σ_8 degeneracy found in the power spectrum. Then we present the full information content of the bispectrum for all triangle configurations down to $k_{\mathrm{max}} = 0.5\ h/\mathrm{Mpc}$ using a Fisher forecast where we estimate the covariance matrix and derivatives with the large set of simulations from the Quijote suite (Villaescusa-Navarro et al. in preparation). This paper is the first of a series of papers that aim to demonstrate the potential of the galaxy bispectrum analysis

---

[3] https://www.desi.lbl.gov/
[4] https://pfs.ipmu.jp/
[5] http://sci.esa.int/euclid/
[6] https://wfirst.gsfc.nasa.gov/

in constraining M_ν. In the subsequent paper, we will include HOD parameters 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 simulation suites, HADES and Quijote, and the halo catalogs cosntructed from them, used throughout the paper. We then describe in Section 3, how we measure the bispectrum using these simulations. We then use the redshift-space halo bispectrums to demonstrate the distinct imprint of M_ν on the bispectrum, which allows the bispectrum 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 dramatically improves the constraints on the cosmological parameters: $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, σ_8, and *especially M_ν*.

2. HADES AND QUIJOTE SIMULATION SUITES

The HADES[7] and Quijote suites are sets of, 23000 total, N-body simulations run on multiple cosmologies, include those with massive neutrinos ($M_\nu > 0$ eV). In this work, we use a subset of the HADES and Quijote simulations. Below, we briefly describe these simulations; a brief 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 $N_{\rm cdm} = 512^3$ CDM, plus $N_\nu = 512^3$ neutrino particles (for $M_\nu > 0$ eV cosmologies), to $z = 0$. They are run using the GADGET-III TreePM+SPH code (Springel 2005) in a periodic (1 h^{-1}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): $\Omega_{\rm m}$=0.3175, $\Omega_{\rm b}$=0.049, Ω_Λ=0.6825, n_s=0.9624, h=0.6711, and $k_{\rm pivot} = 0.05$ hMpc^{-1}.

The HADES suite includes N-body simulations with degenerate massive neutrinos of $M_\nu = 0.06, 0.10$, and 0.15 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 neutrino and different values of σ_8 to examine the $M_\nu - -\sigma_8$ degeneracy. The σ_8 values 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: only halos with masses above $3.2 \times 10^{13} h^{-1} M_\odot$ are included. We refer readers to Villaescusa-Navarro et al. (2018) for more details on the HADES simulations.
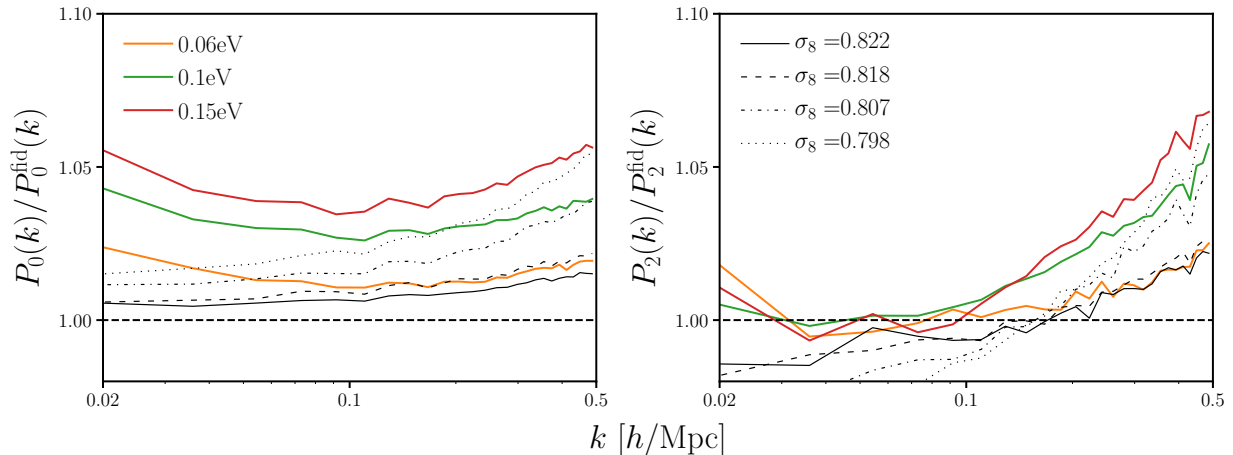
---

[7] https://franciscovillaescusa.github.io/hades.html

Figure 1. Impact of M_ν and σ_8 on the redshift-space halo power spectrum monopole, P_0, and quadrupole, P_2, measured using the HADES simulation suite. The P_ℓ measurements at each cosmology is averaged over 100 N-body realizations. M_ν and σ_8 produce almost identical effects on halo clustering on small scales ($k > 0.1\,h/\mathrm{Mpc}$). 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.*

In addition to HADES, we use simulations from the Quijote suite, a set of 23,000 N-body simulations that in total contain more than 3.3 trillion (3.3×10^{12}) particles over a volume of $23000(h^{-1}\mathrm{Gpc})^3$. These simulations were constructed to quantify the information content of different cosmological observables using Fisher matrix forecasting (Section 4.2). They are therefore designed to accurately calculate the covariance matrices of observables and the derivatives of observables with respect to cosmological parameters. The suite considers 6 cosmological parameters: Ω_m, Ω_b, h, n_s, σ_8, and M_ν.

To calculate covariance matrices, Quijote includes $N_\mathrm{cov} = 15,000$ N-body simulations run at a fiducial cosmology (Ω_m=0.3175, Ω_b=0.049, h=0.6711, n_s=0.9624, σ_8=0.834, and M_ν=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: $\{\Omega_\mathrm{m}^+, \Omega_\mathrm{m}^-, \Omega_\mathrm{b}^+, \Omega_\mathrm{b}^-, h^+, h^-, n_s^+, n_s^-, \sigma_8^+, \sigma_8^-\}$. 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 intial 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. 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). For further details on the Quijote simulations, we refer readers to Villaescusa-Navarro et al. (in preparation).

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 $(10^{10}h^{-1}M_\odot)$	realizations
				HADES suite				
Fiducial	0.0	0.3175	0.049	0.6711	0.9624	0.833	0.833	100
	0.06	0.3175	0.049	0.6711	0.9624	0.819	0.822	100
	0.10	0.3175	0.049	0.6711	0.9624	0.809	0.815	100
	0.15	0.3175	0.049	0.6711	0.9624	0.798	0.806	100
	0.0	0.3175	0.049	0.6711	0.9624	0.822	0.822	100
	0.0	0.3175	0.049	0.6711	0.9624	0.818	0.818	100
	0.0	0.3175	0.049	0.6711	0.9624	0.807	0.807	100
	0.0	0.3175	0.049	0.6711	0.9624	0.798	0.798	100
				Quijote suite				
Fiducial	0.0	0.3175	0.049	0.6711	0.9624	0.834	0.834	15,000
M_ν^+	<u>0.1</u>	0.3175	0.049	0.6711	0.9624	0.834	0.834	500
M_ν^{++}	<u>0.2</u>	0.3175	0.049	0.6711	0.9624	0.834	0.834	500
M_ν^{+++}	<u>0.4</u>	0.3175	0.049	0.6711	0.9624	0.834	0.834	500
Ω_m^+	0.0	<u>0.3275</u>	0.049	0.6711	0.9624	0.834	0.834	500
Ω_m^-	0.0	<u>0.3075</u>	0.049	0.6711	0.9624	0.834	0.834	500
Ω_b^+	0.0	0.3175	<u>0.051</u>	0.6711	0.9624	0.834	0.834	500
Ω_b^-	0.0	0.3175	<u>0.047</u>	0.6711	0.9624	0.834	0.834	500
h^+	0.0	0.3175	0.049	<u>0.6911</u>	0.9624	0.834	0.834	500
h^-	0.0	0.3175	0.049	<u>0.6511</u>	0.9624	0.834	0.834	500
n_s^+	0.0	0.3175	0.049	0.6711	<u>0.9824</u>	0.834	0.834	500
n_s^-	0.0	0.3175	0.049	0.6711	<u>0.9424</u>	0.834	0.834	500
σ_8^+	0.0	0.3175	0.049	0.6711	0.9624	<u>0.849</u>	<u>0.849</u>	500
σ_8^-	0.0	0.3175	0.049	0.6711	0.9624	<u>0.819</u>	<u>0.819</u>	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 massless neutrino 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 different cosmologies, where only one parameter is varied from the fiducial value (underlined), to estimate derivatives of observables along the cosmological parameters.

3. BISPECTRUM

We're 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
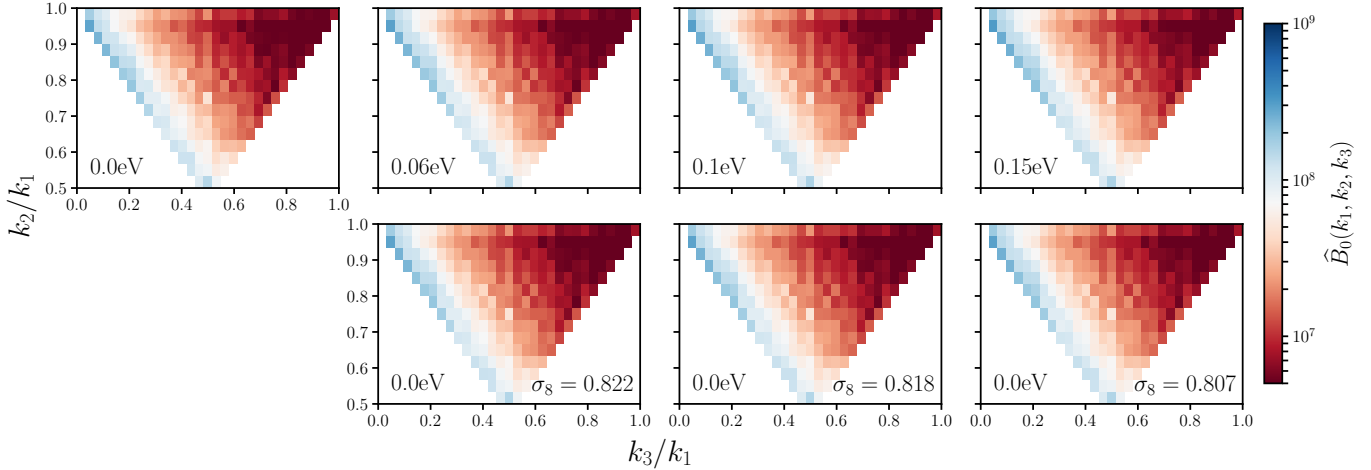
Figure 2. The redshift-space halo bispectrum, $\widehat{B}_0(k_1, k_2, k_3)$, as a function of triangle configuration shape for $M_\nu = 0.0, 0.06, 0.10$, and $0.15\,\mathrm{eV}$ (upper panels) and $\sigma_8 = 0.822, 0.818$, and 0.807 (lower panels). The $\widehat{B}_0$ at each cosmology (each panel) 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. 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 $k_1, k_2, k_3 \leq k_{\max} = 0.5\,h/\mathrm{Mpc}$ and use the $\widehat{B}_0$ estimator in Section 3.

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). Our estimator is similar to the estimators described in 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, $\delta(\boldsymbol{x})$, and Fourier transforming the grid to get $\delta(\boldsymbol{k})$. We use a fourth-order interpolation to interlaced grids, which has advantageous anti-aliasing properties that allow unbias measurements up to the Nyquist frequency (Hockney & Eastwood 1981; Sefusatti et al. 2016). Then using $\delta(\boldsymbol{k})$, we measure the bispectrum monopole as

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

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

$$\int_{k_i} \mathrm{d}^3 q \equiv \int_{k_i - \delta k/2}^{k_i + \delta k/2} \mathrm{d}q\, q^2 \int \mathrm{d}\Omega. \quad (2)$$
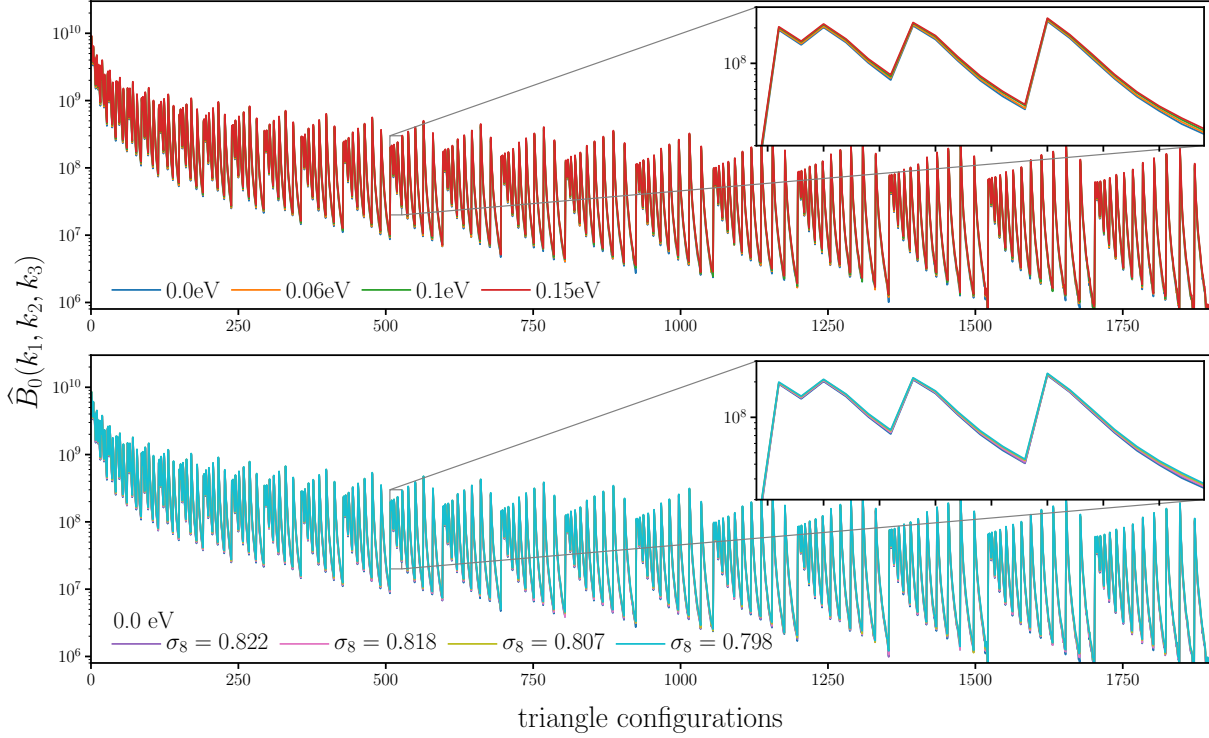
Figure 3. The redshift-space halo bispectrum, $\widehat{B}_0(k_1, k_2, k_3)$, as a function of triangle configurations for $M_\nu = 0.0, 0.06, 0.10$, and 0.15 eV (top panel) and $M_\nu = 0.0$ eV, $\sigma_8 = 0.822, 0.818, 0.807$, and 0.798 (lower panel). Each $\widehat{B}_0$ is averaged over 100 N-body 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 $k_1 \leq k_2 \leq k_3$. In the insets of the panels we zoom into triangle configurations with $k_1 = 0.113, 0.226 \leq k_2 \leq 0.283$, and $0.283 \leq k_3 \leq 0.377$ $h/$Mpc.

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

$$V_B = \int_{k_1} \mathrm{d}^3 q_1 \int_{k_2} \mathrm{d}^3 q_2 \int_{k_3} \mathrm{d}^3 q_3 \; \delta_{\mathrm{D}}(\boldsymbol{q}_{123}). \tag{3}$$

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

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

where $\bar{n}$ is the number density of objects (halos) and P_0 is the powerspectrum 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

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

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

where

$$I_{k_i}(\boldsymbol{x}) = \int_{k} \mathrm{d}^3 q \, \delta(\boldsymbol{q}) \, e^{i\boldsymbol{q}\cdot\boldsymbol{x}}. \tag{7}$$

At this point, we measure $\widehat{B}_{\ell=0}(k_1, k_2, k_3)$ by calcuating the I_{k_i}s with inverse FFTs and summing over in real space. For $\widehat{B}_{\ell=0}$ measurements throughout the paper, we use $\delta(\boldsymbol{x})$ grids with $N_{\mathrm{grid}} = 360$ and triangle configurations defined by k_1, k_2, k_3 bins of width $\Delta k = 3k_f = 0.01885 \, h/\mathrm{Mpc}$, three times the fundamental mode $k_f = 2\pi/(1000 \, h^{-1}\mathrm{Mpc})$ given the box size[8].

We present the redshift-space halo bispectrum of the HADE 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 $\widehat{B}_0(k_1, k_2, k_3)$ as a function of k_2/k_1 and k_3/k_1, which describe the triangle configuration shape. In each panel, the colormap of the $(k_2/k_1, \, k_3/k_1)$ bins represent the weighted average $\widehat{B}_0$ 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 ($k_1 = 2k_2 = 2k_3$). We include all possible 1898 triangle configurations with $k_1, k_2, k_3 < k_{\max} = 0.5 \, h/\mathrm{Mpc}$. The $\widehat{B}_0$ in the upper panels are HADES models with $M_\nu = 0.0$ (fiducial), $0.06, 0.10$, and $0.15 \, \mathrm{eV}$; $\widehat{B}_0$ in the lower panels are HADES models with $M_\nu = 0.0 \, \mathrm{eV}$ 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 $\widehat{B}_0(k_1, k_2, k_3)$ for all possible triangle configurations with $k_1, k_2, k_3 < k_{\max} = 0.5 \, h/\mathrm{Mpc}$ where we order the configurations by looping through k_3 in the inner most loop and k_1 in the outer most loop with $k_1 \le k_2 \le k_3$. In the top panel, we present $\widehat{B}_0$ of HADES models with $M_\nu = 0.0, 0.06, 0.10$, and $0.15 \, \mathrm{eV}$; in the lower panel, we present $\widehat{B}_0$ of HADES models with $M_\nu 0.0 \, \mathrm{eV}$ and $\sigma_8 = 0.822, 0.818$, and 0.807. We zoom into triangle configurations with $k_1 = 0.113$, $0.226 \le k_2 \le 0.283$, and $0.283 \le k_3 \le 0.377 \, h/\mathrm{Mpc}$ in the insets of the panels.

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 imprint 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.

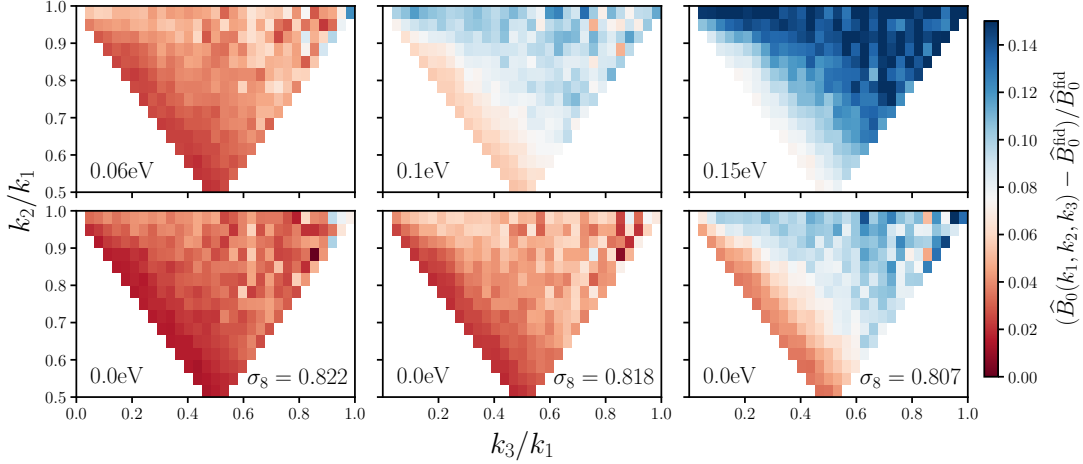[8] The code that we use to evaluate $\widehat{B}_{\ell=0}$ is publicly available at https://github.com/changhoonhahn/pySpectrum

Figure 4. The shape dependence of the M_ν and σ_8 imprint on the redshift-space halo bispectrum, $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$. 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. 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.*

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 separate 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.

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, $(\Delta\widehat{B}_0 = \widehat{B}_0 - \widehat{B}_0^{\rm fid})/\widehat{B}_0^{\rm fid}$, 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 bispectrum increases for all triangle shapes (top panels). This increase is due to halo bias (Villaescusa-Navarro et al. 2018, see also Figure 1). We impose a fixed $M_{\rm lim}$ on our halos so 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 $\widehat{B}_0^{\rm fid}$ for equilateral triangles while only $\sim 8\%$ higher for folded triangles (lower center). The noticeable difference in $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ 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
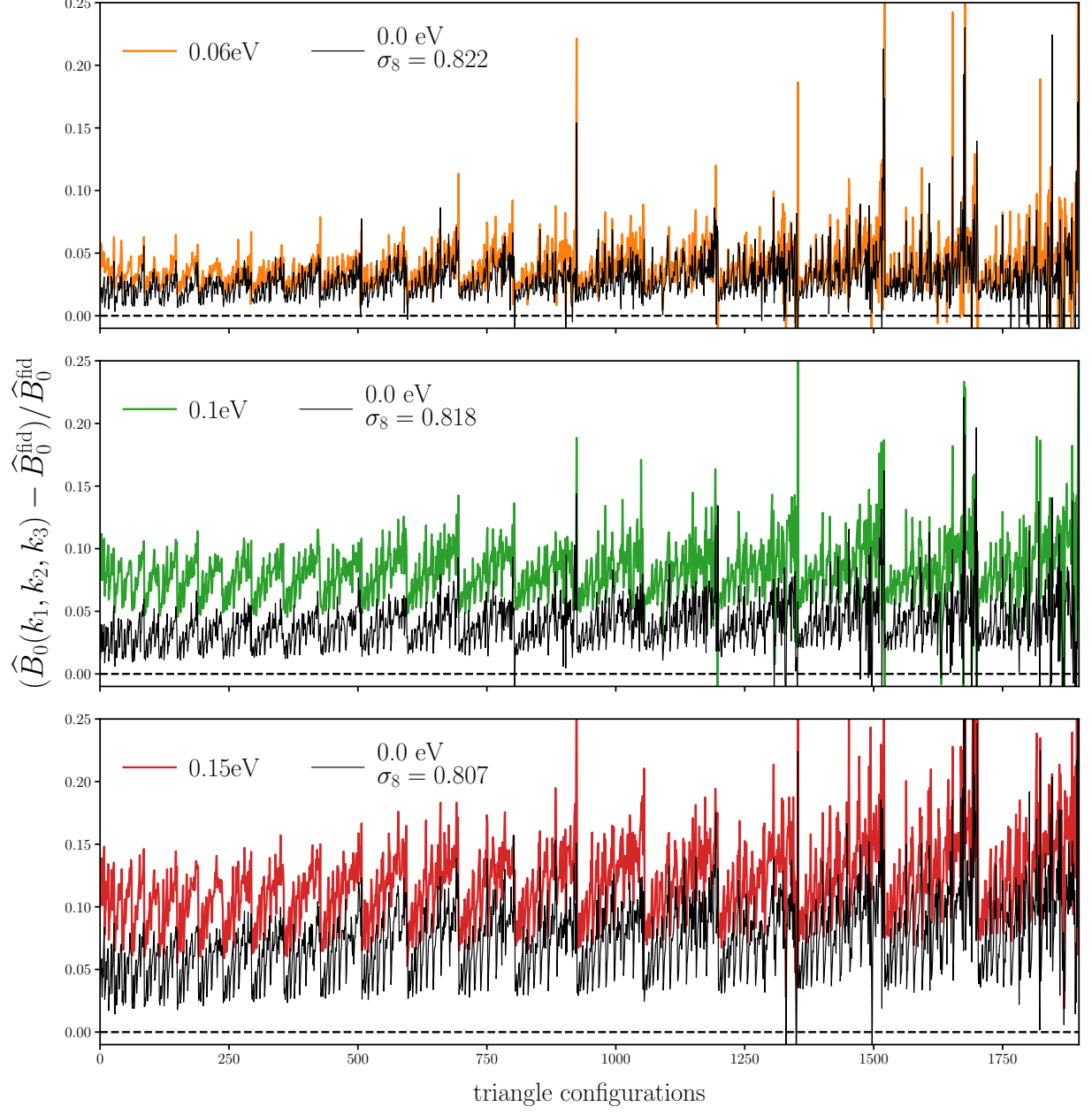
Figure 5. The impact of M_ν and σ_8 on the redshift-space halo bispectrum, $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$, for all 1898 triangle configurations with $k_1, k_2, k_3 \leq 0.5\ h/{\rm Mpc}$. We compare $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ of the $M_\nu = 0.06$ (top), 0.10 (middle), and 0.15 eV (bottom) HADES simulations to $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ 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 spectrums 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.*

significant differences in $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ for M_ν versus σ_8 alone. Between $M_\nu = 0.15$ eV and $\{0.0$ eV, $\sigma_8 = 0.807\}$ cosmologies, there is an overall $\gtrsim 5\%$ difference in the bispectrum. In addition, the shape dependence of the $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ increase is different for M_ν than σ_8. This is particularly clear in the differences between 0.1 eV (top center panel) and $\{0.0$ eV, $\sigma_8 = 0.807\}$ (bottom right panel): near equilateral triangles in the two panels have similar $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ while triangle shapes near the lower left edge from the squeezed to folded triangles have significantly different $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$. 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 for all triangle configurations. We present $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ for all 1898 possible triangle configurations with $k_1, k_2, k_3 < k_{\rm max} = 0.5$ h/Mpc in Figure 5. We compare $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ of the $M_\nu = 0.06, 0.10,$ and 0.15 eV HADES models to the $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ 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 spectrums differ by $< 1\%$ (Figure 1).

The comparison in the panels of Figure 5 confirms the difference in the configuration dependence in $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ 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 $k_1 \leq k_2 \leq k_3$. In this ordering, k_1 increases from left to right. $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ 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, $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ of M_ν has a smaller k_1 dependence than $\Delta\widehat{B}_0/\widehat{B}_0^{\rm fid}$ 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.

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. $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, and n_s. The Quijote simulation suite is *specifically* designed to answer this question through Fisher matrix forecast.

First, the Quijote suite includes $N_{\rm cov} = 15,000$ N-body realizations run at a fiducial cosmology: $M_\nu=0.0$eV, $\Omega_{\rm m}=0.3175, \Omega_{\rm b}=0.049, n_s=0.9624, h=0.6711,$ and $\sigma_8=0.834$ (see Table 1). This allows us to robustly estimate the 1898×1898 covariance matrix of the bispectrum, C (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
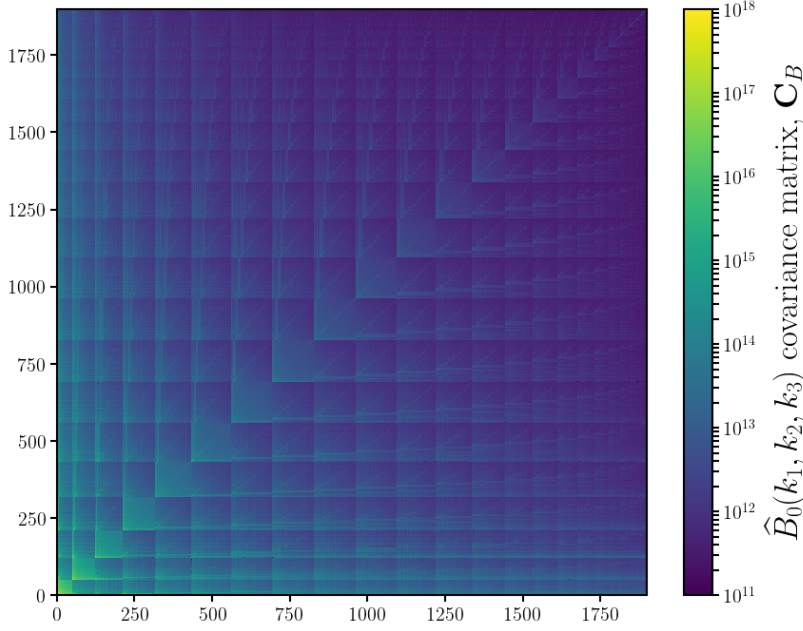
Figure 6. Covariance matrix of the redshift-space halo bispectrum estimated using $N_{\rm cov} = 15,000$ realizations of the Qujiote simulation suite at the fiducial cosmology: $\Omega_{\rm m}$=0.3175, $\Omega_{\rm b}$=0.049, h=0.6711, n_s=0.9624, σ_8=0.834, and M_ν=0.0 eV. We include all possible triangle configurations with $k_1, k_2, k_3 \leq k_{\rm max} = 0.5\ h/{\rm Mpc}$ 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.

effectively gives us $N_{\rm deriv.} = 1,500$ 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

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

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

$$F_{ij} = \frac{1}{2}\, {\rm Tr}\left[\boldsymbol{C}^{-1}\frac{\partial \boldsymbol{C}}{\partial \theta_i}\boldsymbol{C}^{-1}\frac{\partial \boldsymbol{C}}{\partial \theta_j} + \boldsymbol{C}^{-1}\left(\frac{\partial \overline{B}_0}{\partial \theta_i}\frac{\partial \overline{B}_0}{\partial \theta_j}^T + \frac{\partial \overline{B}_0}{\partial \theta_i}^T\frac{\partial \overline{B}_0}{\partial \theta_j} \right) \right]. \tag{9}$$

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

$$F_{ij} = \frac{1}{2}\, {\rm Tr}\left[\boldsymbol{C}^{-1}\left(\frac{\partial \overline{B}_0}{\partial \theta_i}\frac{\partial \overline{B}_0}{\partial \theta_j}^T + \frac{\partial \overline{B}_0}{\partial \theta_i}^T\frac{\partial \overline{B}_0}{\partial \theta_j} \right) \right], \tag{10}$$

directly with $\boldsymbol{C}$ and $\partial B_0/\partial \theta_i$ along each cosmological parameter from the Qujiote simulations.
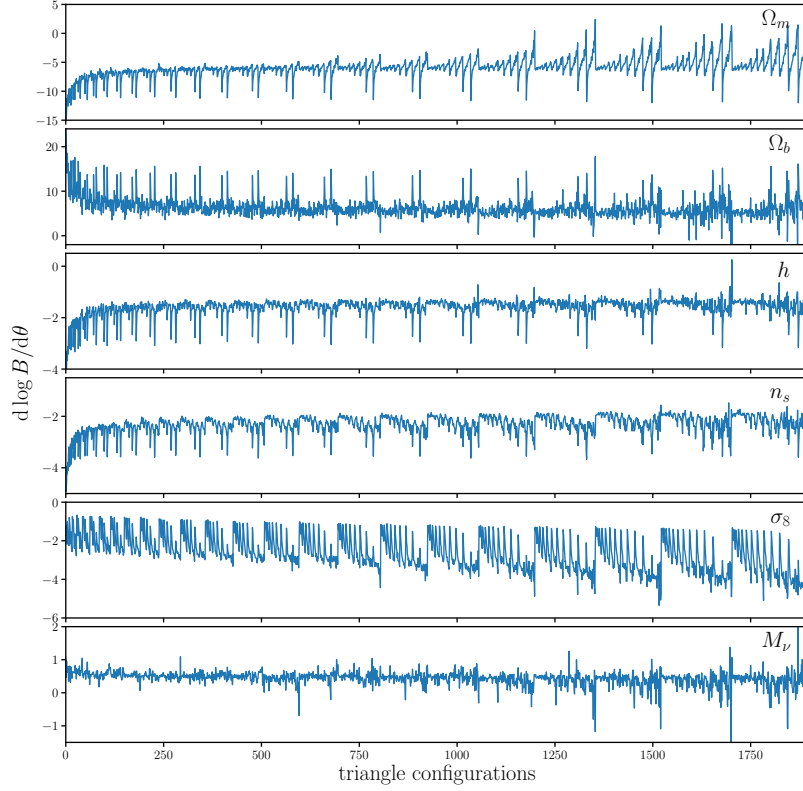
Figure 7. Derivatives of the redshift-space halo bispectrum, $\mathrm{d}\log B_0/\mathrm{d}\theta$ with respect to Ω_{m}, Ω_{b}, h, n_s, σ_8, and M_ν as a function of all 1898 triangle configurations with $k_1, k_2, k_3 \leq 0.5$ h/Mpc (top to bottom panels). We estimate the derivatives at the fiducial parameters using $N_{\mathrm{deriv}} = 1,500$ N-body realizations from the Quijote suite. The configurations above are ordered by looping through k_3 in the outer most loop and k_1 in the inner most loop satisfying $k_1 \leq k_2 \leq k_3$. By using N-body simulations for the derivatives, we exploit their accuracy in the nonlinear regime and rely on fewer assumptions and approximations than analytic methods like perturbation theory.

For Ω_{m}, Ω_{b}, h, n_s, and σ_8, we estimate

$$\frac{\partial \overline{B}_0}{\partial \theta_i} \approx \frac{\overline{B}_0(\theta_i^+) - \overline{B}_0(\theta_i^-)}{\theta_i^+ - \theta_i^-}, \tag{11}$$

where $\overline{B}_0(\theta_i^+)$ and $\overline{B}_0(\theta_i^-)$ are the average bispectrum of the $1,500$ realizations at θ^+ and θ^-, 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

$$\frac{\partial \overline{B}_0}{\partial M_\nu} \approx \frac{-21\overline{B}_0(M_\nu^{\mathrm{fid}}) + 32\overline{B}_0(M_\nu^+) - 12\overline{B}_0(M_\nu^{++}) + \overline{B}_0(M_\nu^{+++})}{1.2}, \tag{12}$$

which provides a $\mathcal{O}(\delta M_\nu^2)$ order approximation. By using these N-body simulations, instead of analytic methods (*e.g.* perturbation theory), we exploit the accuracy of the simulations in the nonlinear 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 full information content
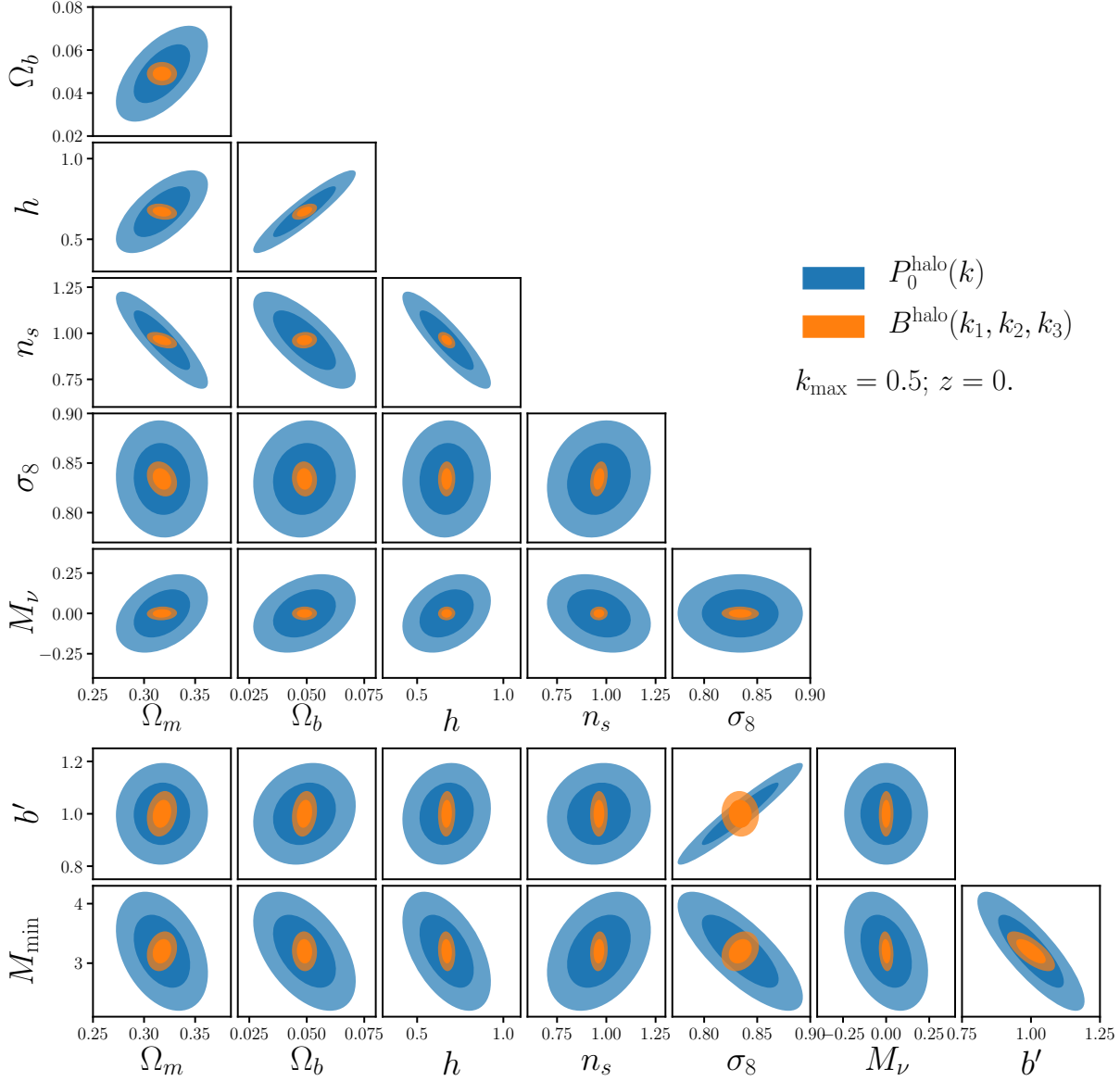
Figure 8. Fisher matrix constraints for M_ν and other cosmological parameters for the redshift-space halo bispectrum monopole (orange). We include Fisher parameter constraints for the redshift-space halo power-spectrum monopole in blue for comparison. The contours mark the 68% and 95% confidence interals. We set $k_{\rm max} = 0.5\ h/{\rm Mpc}$ for both power spectrum and bispectrum. We include in our forecasts b' and $M_{\rm min}$, 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 $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, and σ_8 improve by factors of 3.1, 4.1, 5.1, 5.9, and 3.3, respectively. *For M_ν, the bispectrum improves σ_{M_ν} from 0.1962 to 0.0342 eV — a factor of $\sim$6 improvement over the power spectrum.*

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 bispectrum derivative and tests of convegence and stability in Appendix B.

We present the constraints on M_ν and other cosmological parameters $\{\Omega_{\rm m}, \Omega_{\rm b}, h, n_s, \sigma_8\}$ derived from the redshift-space halo bispectrum Fisher matrix for $k_{\rm max} = 0.5\ h/{\rm Mpc}$ in Figure 8. We include Fisher constraints for the redshift-space halo power spectrum monopole with the same $k_{\rm max}$ for comparison (blue). The shaded contours mark the 68% and 95% confidence interals. We include in our Fisher constraints the following nuisance parameters: b' and $M_{\rm min}$. b' is a scaling factor on the bispectrum amplitude, analogous to linear bias, and $M_{\rm min}$ 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 $\partial \overline{B}_0 / \partial \theta_i$. For instance, the σ_8^+ and σ_8^- cosmologies have halo $\bar{n} = 1.586 \times 10^{-4}$ and $1.528 \times 10^{-4} (h^{-1}{\rm Mpc})^{-3}$. 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 $\partial \overline{B}_0 / \partial b' = \overline{B}_0$. $\partial \overline{B}_0 / \partial M_{\rm min}$, we estimate numerically using $\overline{B}_0$ evaluated at $M_{\rm min}^+ = 3.3 \times 10^{13} h^{-1} M_\odot$ and $M_{\rm min}^- = 3.1 \times 10^{13} h^{-1} M_\odot$ with all other parameters set to the fiducial value.

The bispectrum substantially improves constraints on all parameters over the power spectrum. For $k_{\rm max} = 0.5\ h/{\rm Mpc}$, the bispectrum tightens the marginalized 1σ constraints, σ_θ, of $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, and σ_8 by factors of $\sim$3, 4, 5, 6, and 3 over the power spectrum. *For M_ν, the bispectrum improves the constraint from $\sigma_{M_\nu} = 0.1962$ to $0.0342\,eV$ — a factor of 6 improvement over the power spectrum.* This $\sigma_{M_\nu} = 0.0342$eV constraint is for the *bispectrum alone* and only for a $1h^{-1}$Gpc box. For a larger volume, V, σ_θ scales as $\propto 1/\sqrt{V}$. We list the precise marginalized Fisher parameter constraints of both cosmological and nuisance parameters for P_0 and B_0 in Table 2.

Even below $k_{\rm max} < 0.5\ h/{\rm Mpc}$, the bispectrum significantly improves cosmological parameter constraints. We compare σ_θ of $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, σ_8, and M_ν as a function of $k_{\rm max}$ for B_0 (orange) and P_0 (blue) in Figure 9. We only include the $k_{\rm max}$ range where the Fisher forecast is well defined — *i.e.* more data bins than the number of parameters: $k_{\rm max} > 8\ k_{\rm f} \approx 0.05\ h/{\rm Mpc}$ for P_0 and $k_{\rm max} > 12\ k_{\rm f} \approx 0.075\ h/{\rm Mpc}$ for B_0. Figure 9 reveals that the improvement of the bispectrum σ_θ over the power spectrum σ_θ is larger at higher $k_{\rm max}$. Although limited by the $k_{\rm max}$ range, the figure suggests that on large scales ($k_{\rm max} \lesssim 0.1\ h/{\rm Mpc}$) σ_θ of P_0 crosses over σ_θ of B_0 so P_0 has more constraining power than B_0, as expected on linear scales. At slightly larger $k_{\rm max}$, $k_{\rm max} = 0.2\ h/{\rm Mpc}$, we find that the bispectrum substantially improves σ_θ by factors of $\sim 1.8, 2.1, 2.6, 2.5, 2.6$, and 3.1 for $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, σ_8, and M_ν respectively.

Our forecasts demonstrate that the bispectrum has significant constraining power beyond the powerspectrum in the weakly nonlinear regime ($k > 0.1\ h/{\rm Mpc}$). 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_0 and B_0: $1/\sqrt{F_{ii}}$ where F_{ii} is a diagonal element of the Fisher matrix. For $k < 0.4\ h/{\rm Mpc}$, the unmarginalized cosntraints from P_0 are tighter than those from B_0. Yet, once we marginalize the constraints over the other parameters, the B_0 constraints are tigher than P_0 for $k > 0.1\ h/{\rm Mpc}$. The derivatives, $\partial B_0 / \partial \theta_i$, also shed light on

Table 2. Marginalized Fisher parameter constraints from the redshift-space halo power spectrum (top) and bispectrum (bottom) for different values of $k_{\max}$. 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_0	0.2	±0.333	±0.052	±0.030	±0.372	±0.347	±0.128	±0.649	±5.045
	0.3	±0.277	±0.044	±0.023	±0.273	±0.276	±0.069	±0.383	±2.457
	0.4	±0.228	±0.040	±0.020	±0.235	±0.240	±0.059	±0.226	±1.270
	0.5	**±0.196**	±0.036	±0.018	±0.207	±0.213	±0.048	±0.157	±0.807
B_0	0.2	±0.107	±0.029	±0.014	±0.144	±0.140	±0.050	±0.265	±1.317
	0.3	±0.065	±0.020	±0.008	±0.077	±0.074	±0.023	±0.143	±0.657
	0.4	±0.043	±0.015	±0.006	±0.052	±0.047	±0.016	±0.088	±0.369
	0.5	**±0.034**	±0.012	±0.004	±0.040	±0.036	±0.014	±0.070	±0.269

how B_0 breaks parameter degeneracies. The parameter degeneracies in the P_0 forecasts of Figure 8 are consistent with similarities in the shape and scale dependence of P_0 derivatives $\partial P_0/\partial\theta$. On the other hand, the B_0 derivatives with respect to the parameters have significant different scale and triangle shape dependences, as seem in Figure 7.

By exploiting the massive number of N-body simulations of the Quijote suite, we present for the first time the full information content of the 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; Chan & Blot 2017). More recently, Chan & Blot (2017), used covariance matrices that include non-Gaussian contributions calibrated with 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 $k_{\max} \sim 0.1$ h/Mpc and increases to $\sim 40\%$ at $k_{\max} \sim 0.35$ h/Mpc. While these simple SNR measurements cannot be easily compared to Fisher analysis (), 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). Outside the k range explored by Chan & Blot (2017), $k_{\max} > 0.35$ h/Mpc, we find that the SNR of B_0 continues to increase at higher $k_{\max}$ in constract to the P_0 SNR, which saturates at $k_{\max} \sim 0.1$ h/Mpc. At $k_{\max} = 0.75$ h/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 with Fisher forecasts (Scoccimarro et al. 2004; Sefusatti et al. 2006; Sefusatti & Komatsu 2007; Song et al. 2015; Tellarini et al. 2016; Yamauchi et al. 2017; Karagiannis
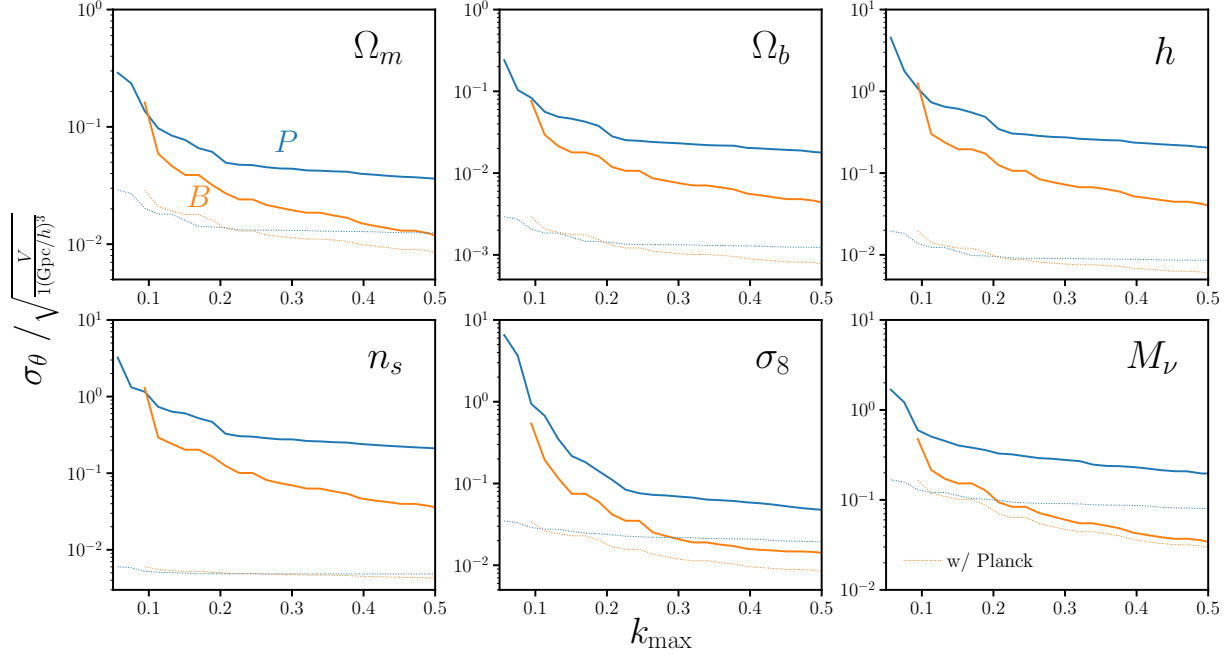
Figure 9. Marginalized 1σ constraints, σ_θ, of the cosmological parameters $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, σ_8, and M_ν as a function of $k_{\rm max}$ for the redshift-space halo bispectrum (orange) and power spectrum (blue). The constraints are marginalized over the nuisance parameters b' and $M_{\rm min}$ in our forecast (Section 4.2). We impose $k_{\rm max} > 8\,k_{\rm f}$ for P_0 and $k_{\rm max} > 12\,k_{\rm f}$ 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 $k_{\rm max} < 0.5\,h/{\rm Mpc}$, the bispectrum significantly improves cosmological parameter constraints. The improvement, however, is larger for higher $k_{\rm max}$. At $k_{\rm max} = 0.2\,h/{\rm Mpc}$, the bispectrum improves constraints on $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, σ_8, and M_ν by factors of ~ 1.8, 2.1, 2.6, 2.5, 2.6, and 3.1 over the power spectrum. Even with *Planck* priors, B_0 significantly improves σ_θ for $k_{\rm max} \gtrsim 0.2\,h/{\rm Mpc}$. While the constraining power of P_0 saturates at $k_{\rm max} = 0.2\,h/{\rm Mpc}$, the constraining power of B_0 continues to increase out to $k_{\rm max} = 0.5\,h/{\rm Mpc}$.

et al. 2018; Yankelevich & Porciani 2019). While most of these works fix most cosmological parameters and focus soley on forecasting constraints of primordial non-Gaussianity and bias parameters, Sefusatti et al. (2006) and Yankelevich & Porciani (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 $k_{\rm max} = 0.3\,h/{\rm Mpc}$, they find constraints on $\Omega_{\rm m}$, $\Omega_{\rm b}$, 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 $k_{\rm max} = 0.3\,h/{\rm Mpc}$ our B_0 constraints are tighter than P_0 constraints by factors of 2.1, 2.9, 3.5, 3.7, and 3.0. 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 break down on small scales. In a comparison to N-body simulations, Lazanu et al. (2016) find that perturbation theory models of the matter bispectrum deviate by $> 5\%$ at $k_{\rm max} \gtrsim 0.15\,h/{\rm Mpc}$.

Yankelevich & Porciani (2019) present Fisher forecasts for $\Omega_{\rm cdm}$, $\Omega_{\rm b}$, 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 $k_{max} = 0.15\ h$/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, their forecasts are at higher redshifts, $z > 0.7$, where we expect the constraining power of B to be weaker than at $z = 0$. 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 substantially impact the constraints especially on small scales (Chan & Blot 2017). Despite differences, Yankelevich & Porciani (2019) find that the constraining power of B relative to P increases for higher k_{max}, consistent with our forecasts as a function of k_{max} (Figure 9). Also, consistent with their results, for $k_{max} = 0.15\ h$/Mpc, we find similar posterior correlations between the P_0 and B_0 constraints. At $k_{max} = 0.5\ h$/Mpc, however, we find the posterior correlations are no longe rsimilar, which contribute to the constraining power of B_0 (Figure 8).

Various difference between our forecast and previous work prevent more thorough comparisons. Two crucial aspects, however, 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. This allows us to go beyond perturbation theory forecasts and quantify the full information content of the redshift-space bispectrum out to nonlinear regimes. Second, we present the first bispectrum forecast of cosmological parameters that includes neutrinos and demonstrates 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 P_0 and B_0 derivatives with respect to M_ν by comparing the derivatives computed using N-body simulations at three different sets of cosmologies: (1) {fiducial, M_ν^+, M_ν^{++}, and M_ν^{+++}} (Eq. 12), (2) {fiducial, M_ν^+, and M_ν^{++}}, and (3) {fiducial and M_ν^+ } (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_0 and B_0 derivatives are affected by the same factor, the relative improvement of the $B_0\ M_\nu$ constraint over the P_0 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}, repsectively. For N_{cov}, we find $< 5\%$ variation in the Fisher matrix elements, F_{ij}, for $N_{cov} > 5000$ and $< 1\%$ variation in σ_θ for $N_{cov} > 12000$. For N_{deriv}, we find $< 5\%$ variation in the F_{ij} elements and $< 5\%$ variation in σ_θ for $N_{deriv} > 1200$. 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 Apprendix 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 $\{\Omega_{\rm m}, \Omega_{\rm b}, h, n_s, \sigma_8\}$ tighter than the P_0 and B_0 alone (**CH: @paco: citation to planck prior**), the elements of the *Planck* prior matrix are larger than the elements of P_0 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_0 to $k_{\rm max} = 0.5\ h/{\rm Mpc}$, we derive the following constraints: $\sigma_{\Omega_{\rm m}} = 0.0125$, $\sigma_{\Omega_{\rm b}} = 0.0012$, $\sigma_h = 0.0086$, $\sigma_{n_s} = 0.0048$, and $\sigma_{\sigma_8} = 0.0194$. Including *Planck* priors expectedly tighten the constraints from P_0 alone. Meanwhile, with *Planck* priors and B_0 to $k_{\rm max} = 0.5\ h/{\rm Mpc}$, we get $\sigma_{\Omega_{\rm m}} = 0.0086$, $\sigma_{\Omega_{\rm b}} = 0.0008$, $\sigma_h = 0.0006$, $\sigma_{n_s} = 0.0043$, and $\sigma_{\sigma_8} = 0.0085$, 1.5, 1.6, 1.4, 1.1, and 2.3 times tighter constraints. For M_ν, $\sigma_{M_\nu} = 0.0802$ eV for P_0 and $\sigma_{M_\nu} = 0.0297$ eV, a factor of 2.7 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 $k_{\rm max}$ with *Planck* priors reveal that while the constraining power of P_0 saturates at $k_{\rm max} = 0.2\ h/{\rm Mpc}$, the constraining power of B_0 continues to increase out to $k_{\rm max} = 0.5\ h/{\rm Mpc}$ (dotted; Figure 9),

Our forecasts are derived using the power spectrum and bispectrum in a *periodic box*. 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. Survey geometry, 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 $M_{\rm min}$ in our forecasts to address the difference in halo bias and number densities among the Quijote cosmologies. Although we marginalize over $M_{\rm min}$, 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, $\}A_{\rm SN}, B_{\rm SN}, b_2, \gamma_2\}$, and examine their impact on our forecasts. $A_{\rm SN}$ and $B_{\rm SN}$ are multiplicative factors of the first and second terms of Eq. 4, which we include to account for any $\bar{n}$ dependence that may be introduced from the shotnoise correction (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', $M_{\rm min}$, $A_{\rm SN}$, $B_{\rm SN}$, b_2 and γ_2, we obtain the following constraints for B_0 with $k_{\rm max} = 0.5\ h/{\rm Mpc}$: $\sigma_{\Omega_{\rm m}} = 0.0129$, $\sigma_{\Omega_{\rm b}} = 0.0044$, $\sigma_h = 0.0404$, $\sigma_{n_s} = 0.0455$, $\sigma_{\sigma_8} = 0.0228$, and $\sigma_{M_\nu} = 0.0343$. 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; **?**; 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 constraing power of the galaxy bispectrum in the next paper of this series.

Marginalizing over galaxy bias parameters, will likely reduce the constraining power at high k. Hand et al. (2017), for instance, with their 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/Mpc to 0.4 h/Mpc. However, jointly analyzing power spectrum and bispectrum will help break parameter degeneracies and improve constraints on cosmological parameters (Sefusatti et al. 2006; Yankelevich & Porciani 2019). We again emphasize that the constraints we present is for a $1h^{-1}$Gpc box, a substantially smaller volume 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

CH: talk about DESI, PFS, WFIRST

ACKNOWLEDGEMENTS

APPENDIX

A. REDSHIFT-SPACE HALO BISPECTRUM

CH: do we want to keep this section?

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 ($\boldsymbol{C}$; 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
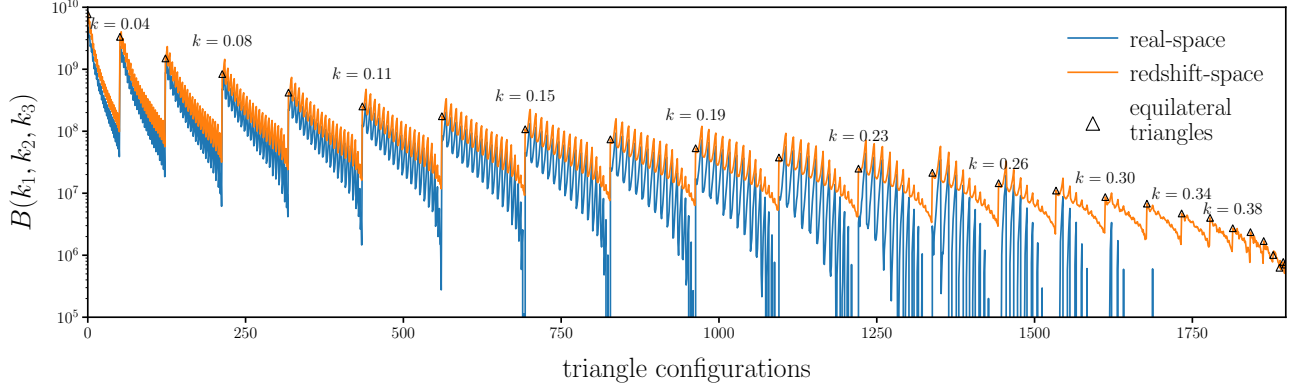
Figure 10. Comparison of the fiducial HADES simuluations real and redshift-space halo bispectrum for triangle configurations with $k_1, k_2, k_3 \leq k_{\max} = 0.5 h/\mathrm{Mpc}$ (blue and orange respectively). We mark equilateral triangle configurations (empty triangle marker) along with their side lengths k.

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 $\boldsymbol{C}$ and $\partial B_0/\partial\theta$ have converged and that numerical effects do not introduce any biases that impact our results. Below, we tests the convergence of $\boldsymbol{C}$ 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 $\boldsymbol{C}$. This is a *significantly* larger number of simulations than any previous bispectrum analyses; however, we also consider 1898 triangle configurations out to $k_{\max} = 0.5\ h/\mathrm{Mpc}$. 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 $\boldsymbol{C}$, N_{cov}, and see 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 $F_{ij}(N_{\mathrm{cov}})$, F_{ij} derived from $\boldsymbol{C}$ calculated with N_{cov} simulations, and $F_{ij}(N_{\mathrm{cov}} = 15,000)$ 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 $N_{\mathrm{cov}} > 5000$ and $\lesssim 1\%$ for $N_{\mathrm{cov}} > 10,000$. Next, we present ratio between $\sigma_\theta(N_{\mathrm{cov}})$, the marginalized 1σ constraints for $\{\Omega_{\mathrm{m}}, \Omega_{\mathrm{b}}, h, n_s, \sigma_8, M_\nu\}$ derived from $\boldsymbol{C}$ calculated with N_{cov} simulations, and $\sigma_\theta(N_{\mathrm{cov}} = 15,000)$ in the left panel of Figure 12. The constraints vary by $\lesssim 5\%$ for $N_{\mathrm{cov}} > 5000$ and $\lesssim 1\%$ for $N_{\mathrm{cov}} > 12000$. Hence, $N_{\mathrm{cov}} = 15,000$ is sufficient to accurately estimate $\boldsymbol{C}$ and its convergence does not impact our forecasts.

We estimate $\partial B_0/\partial\theta$ using $N_{\mathrm{deriv}} = 1,500$ 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 $F_{ij}(N_{\mathrm{deriv}})$, the Fisher matrix element derived from $\partial B_0/\partial\theta$ calculated with N_{deriv} simulations, and $F_{ij}(N_{\mathrm{deriv}} = 1,500)$ for all 36 elements of the Fisher matrix in the right panel of Figure 11. For $N_{\mathrm{deriv}} > 1000$, F_{ij} elements vary by $\lesssim 5\%$. Next, we present the ratio between $\sigma_\theta(N_{\mathrm{deriv}})$, the marginalized 1σ constraints for $\{\Omega_{\mathrm{m}}, \Omega_{\mathrm{b}}, h, n_s, \sigma_8, M_\nu\}$ 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 $\sigma_{\Omega_{\text{m}}}$ 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 Ω_{m}, Ω_{b}, h, n_s, σ_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,

$$\frac{\partial \overline{B}_0}{\partial M_\nu} \approx \frac{\overline{B}_0(M_\nu^{\text{fid}} + \delta M_\nu) - \overline{B}_0(M_\nu^{\text{fid}})}{\delta M_\nu}, \tag{B1}$$

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 $\partial \log B(k_1, k_2, k_3)/\partial M_\nu$ (right) and $\partial \log P(k)/\partial M_\nu$ (left) by comparing the derivatives computing using simulations at {fiducial, M_ν^+, M_ν^{++}, M_ν^{+++}} (blue), {fiducial, M_ν^+, M_ν^{++}} (orange), and {fiducial, M_ν^+} (green) in Figure 13. The three $\partial \log B/\partial M_\nu$ approximations differ from one another by $\sim 10\%$ with Eq. 12 producing the largest estimate for both P_0 and B_0. If we use the {fiducial, M_ν^+, M_ν^{++}} derivative and {fiducial, M_ν^+} derivative instead of Eq. 12 for our Fisher forecasts, the marginalized constraint on M_ν for $k_{\text{max}} = 0.5$ h/Mpc increases to 0.238 and 0.377eV for P_0 and 0.0416 and 0.0657eV for B_0. Compared to our $\sigma_{M_\nu} = 0.0342$ eV B_0 forecast, these correspond to a ~ 20 and 90% relative increase. While the derivative estimated from {fiducial, M_ν^+} 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 {fiducial, M_ν^+} derivatives are better $\mathcal{O}(\delta M_\nu^2)$ estimates for $\partial P/\partial M_\nu$ and $\partial B/\partial M_\nu$ 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 {fiducial, M_ν^+} derivatives and the other derivatives are not solely due to numerical stability. Moreover, because the discrepancies in the derivative propogate similarly to the P_0 and B_0 forecasts, the relative improvement of B_0 over P_0 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

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

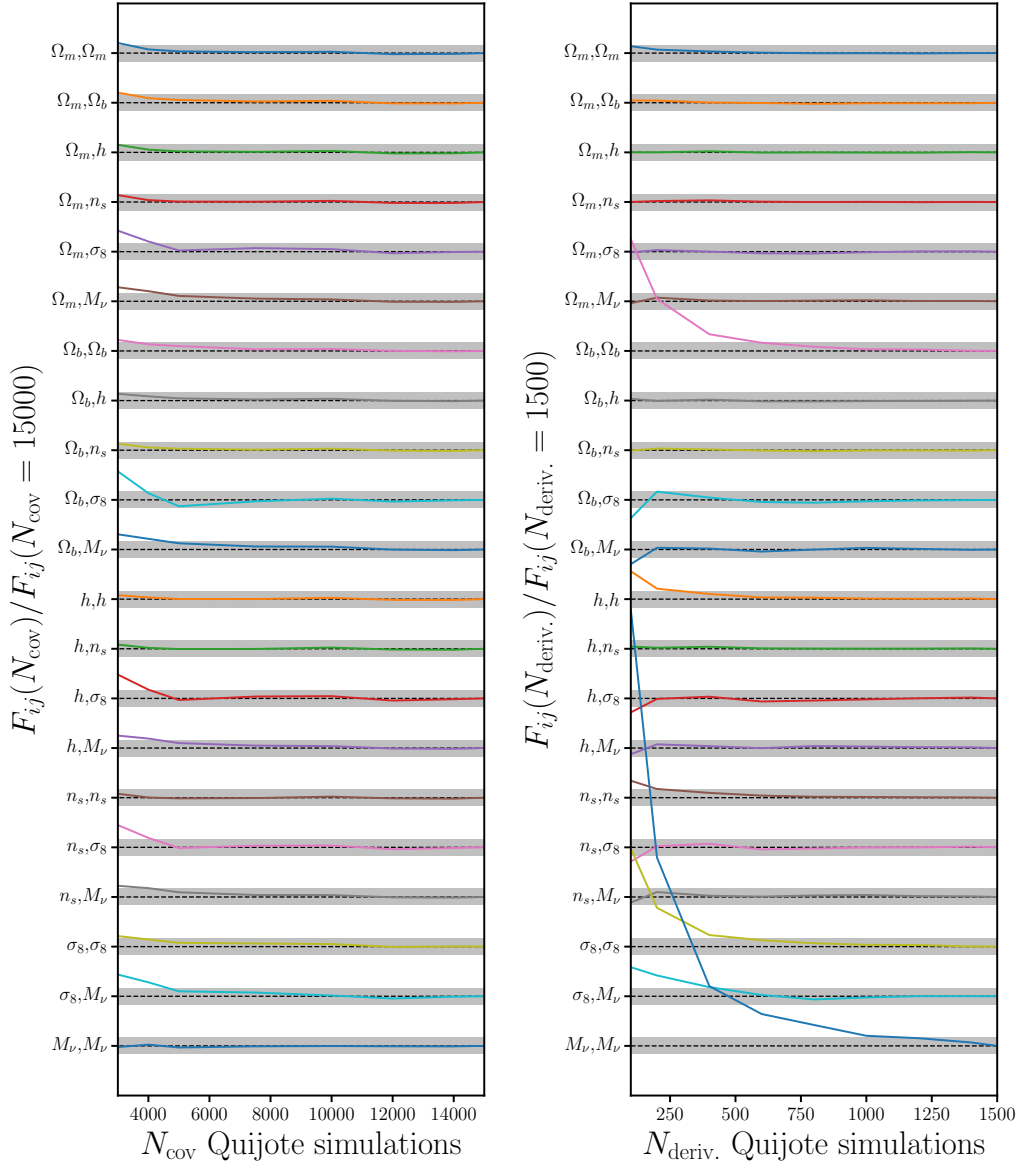**Figure 11.** *Left*: The convergence of all 36 Fisher matrix elements, F_{ij}, as as a function of $N_{\rm cov}$, the number of N-body simulations used to estimate the covariance matrix, $\boldsymbol{C}$. We present the ratio between $F_{ij}(N_{\rm cov})$, F_{ij} derived from $\boldsymbol{C}$ calculated with $N_{\rm cov}$ simulations, and $F_{ij}(N_{\rm cov} = 15,000)$. We mark $\pm 5\%$ deviations in the ratios with the shaded regions, for reference. All F_{ij} elements vary by $\lesssim 5\%$ for $N_{\rm cov} > 5000$ and $\lesssim 1\%$ for $N_{\rm cov} > 10,000$. *Right*: The convergence of F_{ij} as a function of $N_{\rm deriv}$, the number of N-body simulations used to estimate the derivatives, $\partial B_0/\partial \theta$. We plot the ratios between $F_{ij}(N_{\rm deriv})$ and $F_{ij}(N_{\rm deriv} = 1,500)$. All F_{ij} elements vary by $\lesssim 5\%$ for $N_{\rm deriv} > 1000$. Hence, $N_{\rm cov} = 15,000$ *and* $N_{\rm deriv} = 1,500$ *are sufficient and the convergence of* $\boldsymbol{C}$ *and* $\partial B_0/\partial \theta$ *does not impact* F_{ij}.
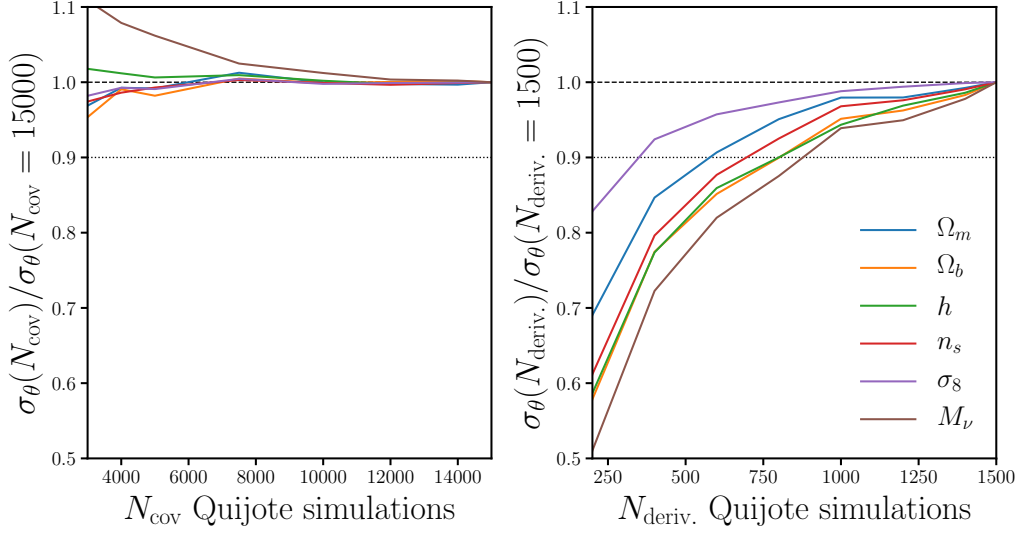
Figure 12. *Left*: The convergence of the marginalized 1σ constraints for $\{\Omega_m, \Omega_b, h, n_s, \sigma_8, M_\nu\}$, σ_θ, as a function of $N_{\rm cov}$, the number of Quijote simulations used to estimate $\boldsymbol{C}$. We plot the ratio between $\sigma_\theta(N_{\rm cov})$, derived from $\boldsymbol{C}$ with $N_{\rm cov}$ simulations, and $\sigma_\theta(N_{\rm cov} = 15,000)$, derived with all 15,000 simulations. σ_θ vary by $\lesssim 5$ and 1% for $N_{\rm cov} > 5000$ and $12,000$, respectively. *Right*: The convergence of σ_θ as a function of $N_{\rm deriv}$, the number of simulated used to estimate $\partial B/\partial M_\nu$. We plot the ratio between $\sigma_\theta(N_{\rm deriv})$, derived from $\partial B/\partial M_\nu$ with $N_{\rm deriv}$ simulations, and $\sigma_\theta(N_{\rm deriv} = 1,500)$, derived with all 1,500 simulations. Although $\sigma_\theta(N_{\rm deriv})/\sigma_\theta(N_{\rm deriv} = 1,500)$ vary among the parameters, the ratio vary by $\lesssim 10$ and 5% for $N_{\rm deriv} > 800$ and 1200, respectively. Hence, *we have a sufficient number of simulations to estimate $\boldsymbol{C}$ and the derivatives of the bipsectrum and our forecasts are robust to their convergence.*
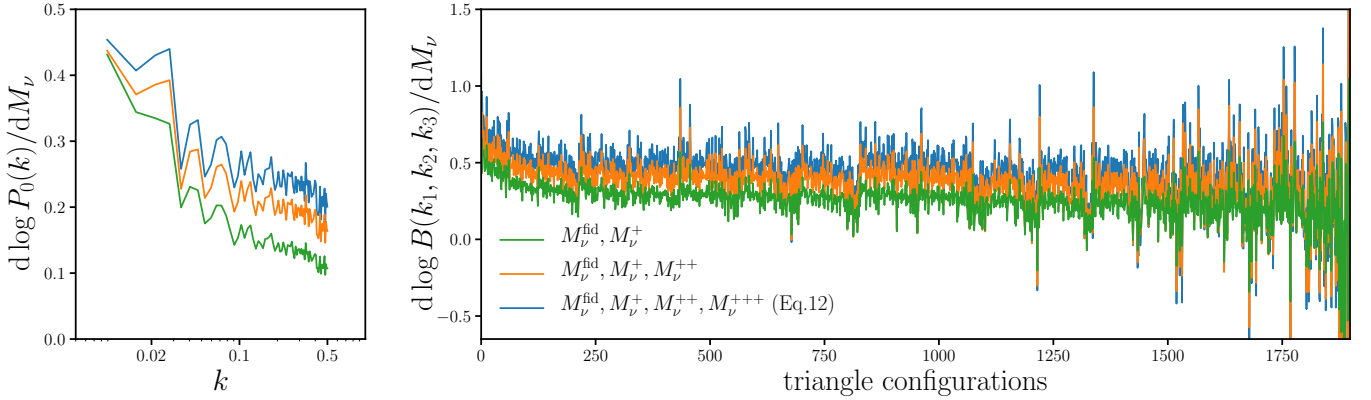

Figure 13. Comparison of $\partial \log B(k_1, k_2, k_3)/\partial M_\nu$ (right) and $\partial \log P(k)/\partial M_\nu$ (left), computed using Quijote simulations at $\{\text{fiducial}, M_\nu^+, M_\nu^{++}, M_\nu^{+++}\}$ (blue), $\{\text{fiducial}, M_\nu^+, M_\nu^{++}\}$ (orange), and $\{\text{fiducial}, M_\nu^+\}$ (green). The derivative approximations differ from one another by $\sim 10\%$ with Eq. 12 producing the largest estimate for both P_0 and B_0. Using the $\{\text{fiducial}, M_\nu^+, M_\nu^{++}\}$ derivatives instead of the Eq. 12 derivatives, increases the marginalized constraint on M_ν by $\sim 20\%$. However, the differences in the derivatives propogate similarly to the P_0 and B_0 forecasts so the relative improvement of B_0 over P_0 remains the same. Hence, we conclude that the derivatives with respect to M_ν are sufficiently stable and robust for our Fisher forecasts.

Audren, B., Lesgourgues, J., Bird, S., Haehnelt, M. G., & Viel, M. 2013, Journal of Cosmology and Astro-Particle Physics, 2013, 026

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

Bonn, J., Eitel, K., Glück, F., et al. 2011, Physics Letters B, 703, 310

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

Collaboration, E., Knabenhans, M., Stadel, J., et al. 2018a, arXiv:1809.04695 [astro-ph], arXiv:1809.04695 [astro-ph]

Collaboration, P., Aghanim, N., Akrami, Y., et al. 2018b, arXiv:1807.06209 [astro-ph], arXiv:1807.06209 [astro-ph]

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

Emberson, J. D., Yu, H.-R., Inman, D., et al. 2017, Research in Astronomy and Astrophysics, 17, 085

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

Gonzalez-Garcia, M. C., Maltoni, M., & Schwetz, T. 2016, Nuclear Physics B, 908, 199

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

Lesgourgues, J., & Pastor, S. 2012

—. 2014

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

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., Sefusatti, E., & Zaldarriaga, M. 2004, Physical Review D, 69, 103513

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

Sefusatti, E., 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

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

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