

## Constraining $M_\nu$ with the Bispectrum I: Breaking Parameter Degeneracies

CHANGHOON HAHN,<sup>1,2,\*</sup> FRANCISCO VILLAESCUSA-NAVARRO,<sup>3</sup> EMANUELE CASTORINA,<sup>2,1</sup> AND  
ROMAN SCOCCIMARRO<sup>4</sup>

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

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

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

<sup>4</sup>*Center for Cosmology and Particle Physics, Department of Physics, New York University, NY 10003, New York, USA*

(Dated: DRAFT --- fa692f0 --- 2019-06-13 --- 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,  $M_\nu$ , a key ingredient for particle physics beyond the Standard Model. However, degeneracies with cosmological parameters limit the constraining power of standard two-point clustering statistics. In this work, we investigate whether we can break these parameter degeneracies and constrain  $M_\nu$  with the next higher-order correlation function, the bispectrum. We first examine the degeneracy between  $M_\nu$  and  $\sigma_8$  using 800  $N$ -body simulations from the HADES suite and demonstrate that the redshift-space halo bispectrum helps break this degeneracy. Next, we quantify the information content of the redshift-space halo bispectrum in a Fisher matrix forecast that includes  $M_\nu$  and cosmological parameters  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$ . This is the first  $N$ -body simulation based Fisher matrix forecast of cosmological parameters using the bispectrum. More importantly, for  $k_{\text{max}} = 0.2$  and  $0.5$  we find

*Keywords:* cosmology: —

### 1. INTRODUCTION

The lower bound on the sum of neutrino masses ( $M_\nu \gtrsim 0.06$  eV), discovered by neutrino oscillation experiments, provides strong evidence of physics beyond the Standard Model of particle physics (Forero et al. 2014; Gonzalez-Garcia et al. 2016). A more precise measurement of  $M_\nu$  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  $M_\nu$  (*e.g.* double beta decay and tritium beta decay experiments) have the potential to place upper bounds of  $M_\nu < 0.2$  eV

\* hahn.changhoon@gmail.com

in upcoming experiments (Drexlin et al. 2013, CH: hmm, who do I cite). These upper bound alone are not insufficient 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  $M_\nu$ .

Neutrinos, in the early universe, are relativistic and contribute to the energy density of radiation. Later as they become non-relativistic, they contribute to the energy density of matter. This transition affects the expansion history of the universe and leaves imprints observable in the cosmic microwave background (CMB) anisotropy spectrum (citecitecite). 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 the clustering and thereby reduce 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 also leaves measurable imprints on the CMB as well as large-scale structure. For more details the effect of neutrinos in cosmological observables, we refer readers to Lesgourgues & Pastor (2012, 2014); Gerbino (2018).

The tightest cosmological constraints on  $M_\nu$  currently come from combining CMB data with other cosmological probes. Temperature and large angle polarization data from the *Planck* satellite places an upper bound of  $M_\nu < 0.54$  eV with 95% confidence level (Collaboration et al. 2018b). Adding the Baryon Acoustic Oscillation (BAO) to the *Planck* likelihood breaks geometrical degeneracies ( $M_\nu$ ,  $h$ ,  $\Omega_m$ ) and significantly tightens the upper bound to  $M_\nu < 0.16$  eV. CMB lensing further tightens the bound further to  $M_\nu < 0.13$  eV, though not as significantly. Combining these probes, however, introduce a strong correlation between  $M_\nu$  and  $A_s$ , the scale amplitude. Since CMB experiments measure the combined quantity  $A_s e^{-2\tau}$ , where  $\tau$  is the optical depth of reionization, this leads to a strong degeneracy between  $M_\nu$  and  $\tau$  (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  $M_\nu$  constraints from CMB data. The best constraints on  $\tau$  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  $\tau$  (Abazajian et al. 2016). Meanwhile, proposed future space-based experiments such as LiteBIRD<sup>1</sup> and CoRE<sup>2</sup>, which have the potential to precisely measure  $\tau$ , have yet to be confirmed. CMB data, however, is not the only way to improve  $M_\nu$  constraints. The imprint of neutrinos on 3D clustering can be measured to constrain  $M_\nu$  and with the sheer cosmic volumes mapped by upcoming surveys, *e.g.* DESI<sup>3</sup>, PFS<sup>4</sup>, EUCLID<sup>5</sup>, and WFIRST<sup>6</sup>, will be able tightly constrain  $M_\nu$  (Audren et al. 2013; Font-Ribera et al. 2014; Petracca et al. 2016; Sartoris et al. 2016; Boyle & Komatsu 2018).

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

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

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

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

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

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

A major limitation of using 3D clustering is modeling 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; 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  $M_\nu$ .

Various works have examined the impact of neutrino masses on nonlinear clustering of matter in real-space (*e.g.* Saito et al. 2008; Wong 2008; Saito et al. 2009; Viel et al. 2010; Agarwal & Feldman 2011, **CH: add more citations**) 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  $M_\nu$  on the redshift-space matter and halo power spectrum to find the imprint of  $M_\nu$  and  $\sigma_8$  on the power spectrum are degenerate and differ only by  $< 1\%$ . The strong  $M_\nu - \sigma_8$  degeneracy poses a serious limitation of constraining  $M_\nu$  with the power spectrum. However, information in the nonlinear regime extend beyond 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  $M_\nu$  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 neutrinos. Yet, no work to date has quantified the full information content and constraining power of the bispectrum down to nonlinear scales — especially for  $M_\nu$ .

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  $M_\nu$ - $\sigma_8$  degeneracy found in the power spectrum. Then we present the full information content of the bispectrum for all triangle configurations down to  $k_{\text{max}} = 0.5 \, h/\text{Mpc}$  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 in constraining  $M_\nu$ . 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, that we use throughout the paper. Then in Section 3, we describe how we measure the bispectrum of these simulations. We then present in Section 4.1, using the redshift-space halo bispectrum of HADES simulations, how  $M_\nu$  leaves a distinct imprint on the bispectrum which allows the bispectrum to break the degeneracy between  $M_\nu$  and  $\sigma_8$ . Finally, in Section 4.2 we present the full information content of the halo bispectrum

with a Fisher forecast of cosmological parameters using the Quijote simulations and demonstrate the how the bispectrum drastically improves the constraints on the cosmological parameters:  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$ , and *especially*  $M_\nu$ .

## 2. HADES AND QUIJOTE SIMULATION SUITES

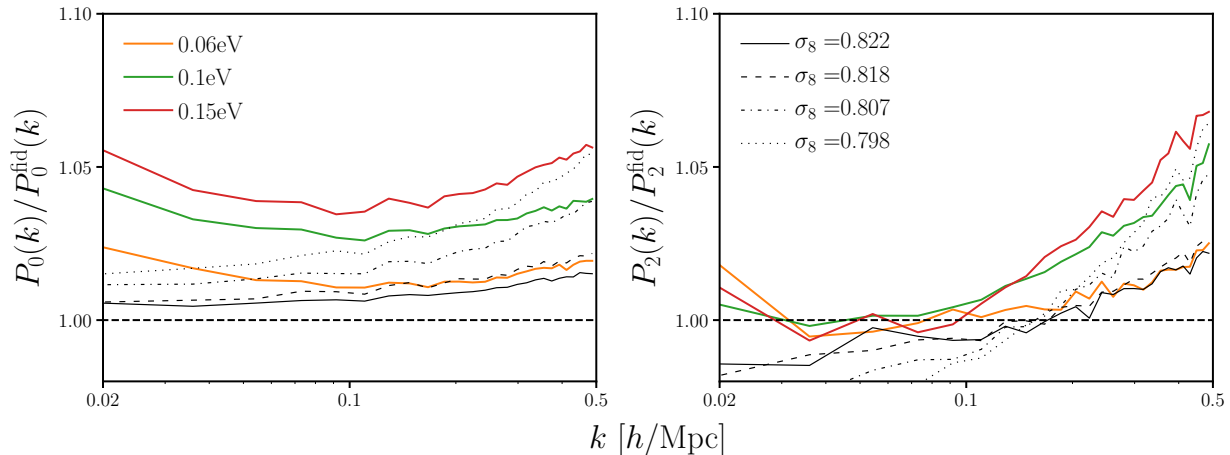
We use a subset of the HADES<sup>7</sup> and Quijote simulation suites. Below, we briefly describe these simulations; a brief summary of the simulations can be found in Table 1. The HADES simulations start from Zel’dovich approximated initial conditions generated at  $z = 99$  using the Zennaro et al. (2017) rescaling method and follow the gravitational evolution of  $N_{\text{cdm}} = 512^3$  CDM, plus  $N_\nu = 512^3$  neutrino particles (for massive neutrino models), to  $z = 0$ . They are run using the GADGET-III TreePM+SPH code (Springel 2005) in a periodic  $(1h^{-1}\text{Gpc})^3$  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 models with degenerate massive neutrinos of different masses:  $M_\nu = 0.06$ ,  $0.10$ , and  $0.15$  eV. These massive neutrino models 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 models with massless neutrino and different values of  $\sigma_8$  to examine the  $M_\nu - \sigma_8$  degeneracy. The  $\sigma_8$  values were chosen to match either  $\sigma_8^m$  or  $\sigma_8^c - \sigma_8$  computed with respect to total matter (CDM + baryons +  $\nu$ ) or CDM + baryons — of the massive neutrino models:  $\sigma_8 = 0.822, 0.818, 0.807$ , 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  $> 3.2 \times 10^{13} h^{-1} M_\odot$  are included. For further details on the HADES simulations, we refer readers to Villaescusa-Navarro et al. (2018).

In addition to HADES, we use simulations from the Quijote simulation suite, a set of 23,000  $N$ -body simulations that in total contain more than 3.3 trillion ( $3.3 \times 10^{12}$ ) particles over a volume of  $23000(h^{-1}\text{Gpc})^3$ . These simulations were constructed to quantify the information content of different cosmological observables using Fisher matrix forecasting (*e.g.* 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:  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$ , and  $M_\nu$ .

To calculate covariance matrices, Quijote includes  $N_{\text{cov}} = 15,000$   $N$ -body simulations run 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  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$ , the

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



**Figure 1.** Impact of  $M_\nu$  and  $\sigma_8$  on the redshift-space halo power spectrum monopole and quadrupole measured using the HADES simulation suite.  $M_\nu$  and  $\sigma_8$  produce almost identical effects on halo clustering on small scales ( $k > 0.1 h/\text{Mpc}$ ). This degeneracy can be partially broken through the quadrupole; however,  $M_\nu$  and  $\sigma_8$  produce almost the same effect on two-point clustering — within a few percent.

fiducial cosmology is adjusted by either a small step above and below the fiducial value. Along  $M_\nu$ , because the derivative of certain observable with respect to  $M_\nu$  is noisy, Quijote includes sets of 500 simulations for  $M_\nu = 0.1, 0.2$ , and  $0.4$  eV. In Table 1, we list 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. 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). The simulations run at the fiducial cosmology for covariance matrix estimation are standard  $N$ -body simulations. However, the rest are paired fixed simulations, which greatly reduce cosmic variance without introducing bias for a large set of statistics (Angulo & Pontzen 2016; Pontzen et al. 2016; Villaescusa-Navarro et al. 2018). We confirm that the paired fixed simulations do not introduce any bias for the redshift-space halo bispectrum (the observable we consider in this paper). For further details on the Quijote simulations, we refer readers to Villaescusa-Navarro et al. (in preparation).

### 3. BISPECTRUM

We’re interested in breaking parameter degeneracies that limit the constraining power on  $M_\nu$  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 Fast Fourier Transforms (FFTs). Our estimator is similar to the estimators described in Scoccimarro (2015); Sefusatti et al. (2016); we also follow their formalism in our description below. Although Sefusatti et al. (2016) and Scoccimarro

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

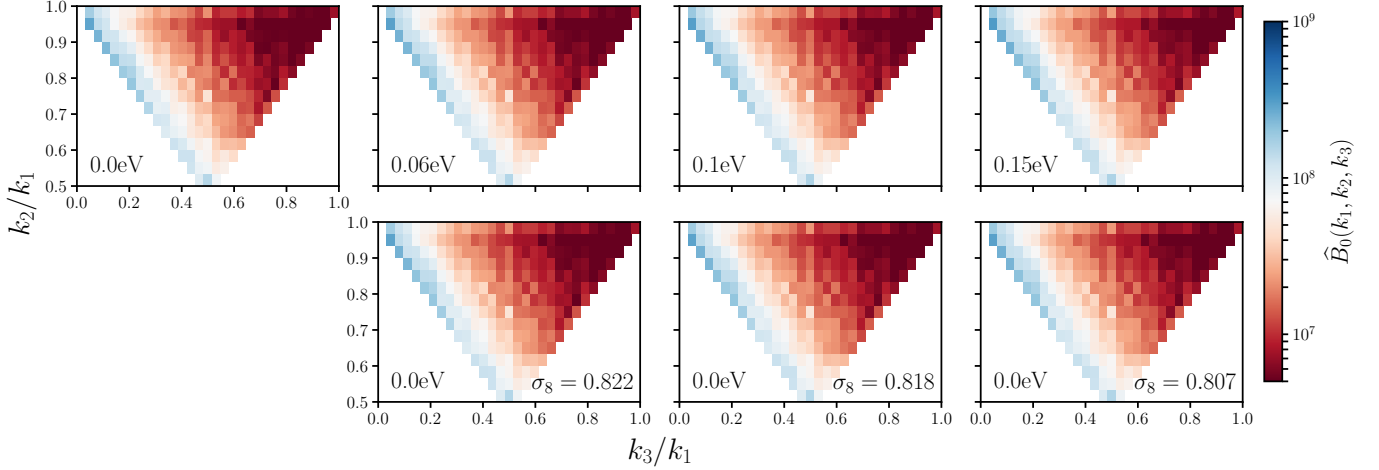
| Name          | $M_\nu$<br>(eV) | $\Omega_m$    | $\Omega_b$   | $h$           | $n_s$         | $\sigma_8^m$<br>( $10^{10}h^{-1}M_\odot$ ) | $\sigma_8^c$<br>( $10^{10}h^{-1}M_\odot$ ) | realizations |
|---------------|-----------------|---------------|--------------|---------------|---------------|--|--|--------------|
| 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       |
| $M_\nu^+$     | <u>0.1</u>      | 0.3175        | 0.049        | 0.6711        | 0.9624        | 0.834                                      | 0.834                                      | 500          |
| $M_\nu^{++}$  | <u>0.2</u>      | 0.3175        | 0.049        | 0.6711        | 0.9624        | 0.834                                      | 0.834                                      | 500          |
| $M_\nu^{+++}$ | <u>0.4</u>      | 0.3175        | 0.049        | 0.6711        | 0.9624        | 0.834                                      | 0.834                                      | 500          |
| $\Omega_m^+$  | 0.0             | <u>0.3275</u> | 0.049        | 0.6711        | 0.9624        | 0.834                                      | 0.834                                      | 500          |
| $\Omega_m^-$  | 0.0             | <u>0.3075</u> | 0.049        | 0.6711        | 0.9624        | 0.834                                      | 0.834                                      | 500          |
| $\Omega_b^+$  | 0.0             | 0.3175        | <u>0.051</u> | 0.6711        | 0.9624        | 0.834                                      | 0.834                                      | 500          |
| $\Omega_b^-$  | 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          |
| $\sigma_8^+$  | 0.0             | 0.3175        | 0.049        | 0.6711        | 0.9624        | <u>0.849</u>                               | <u>0.849</u>                               | 500          |
| $\sigma_8^-$  | 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 paired fixed simulations at different cosmologies, where only one parameter is varied from the fiducial value (underlined), to estimate derivatives of observables along the cosmological parameters.

(2015) respectively describe estimators in redshift- and real-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





**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  $\sigma_8 = 0.822, 0.818$ , and  $0.807$  (lower panels). The HADES simulations of the top and bottom panels in the three right-most columns, have matching  $\sigma_8$  values (Section 2). We describe the triangle configuration shape by the ratio of the triangle sides:  $k_3/k_1$  and  $k_2/k_1$ . 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 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.

to interlaced grids, which has advantageous anti-aliasing properties (Hockney & Eastwood 1981; Sefusatti et al. 2016) that allow unbiased measurements up to the Nyquist frequency. Then using  $\delta(\mathbf{k})$ , we measure the bispectrum monopole as

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

$\delta_D$  above is a 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 form a closed triangle. Each of the integrals above represent an integral over a spherical shell in  $k$ -space with radius  $\delta k$  centered at  $\mathbf{k}_i$  — *i.e.*

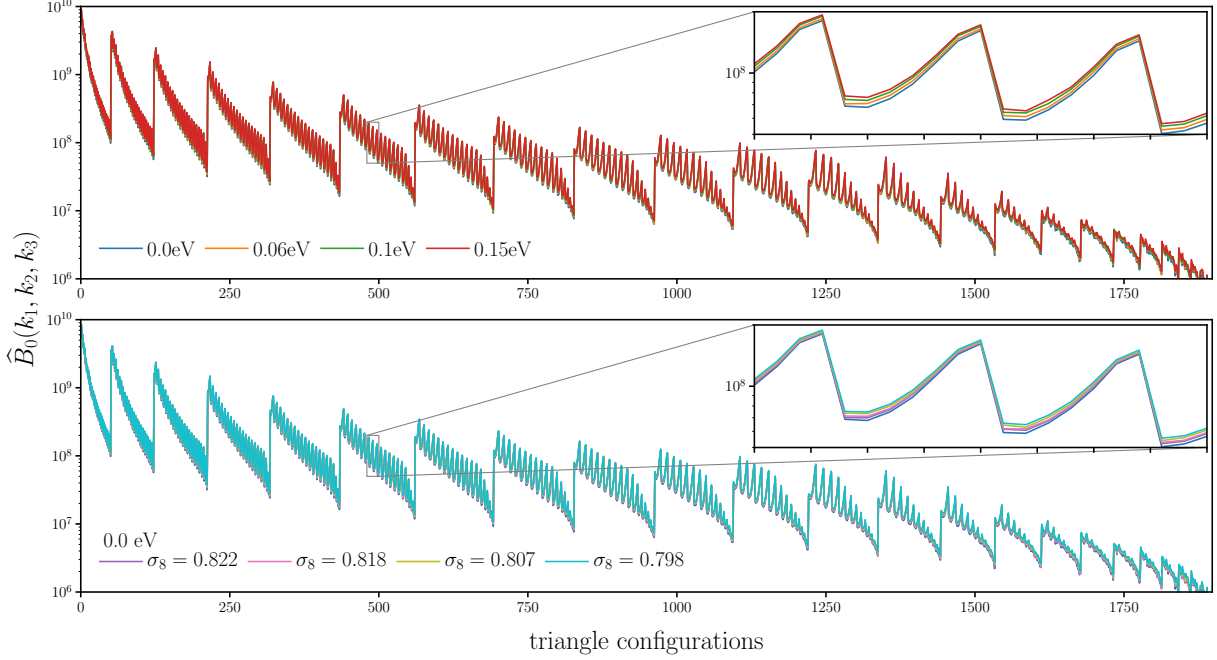
$$\int_{k_i} d^3 q \equiv \int_{k_i - \delta k/2}^{k_i + \delta k/2} dq q^2 \int d\Omega. \quad (2)$$

$V_B$  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  $\delta k$ :

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

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

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



**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). We include all possible triangle configurations with  $k_1, k_2, k_3 \leq k_{\text{max}} = 0.5$   $h/\text{Mpc}$  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  $0.283 \leq k_3 \leq 0.377$   $h/\text{Mpc}$ .

$\bar{n}$  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

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

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

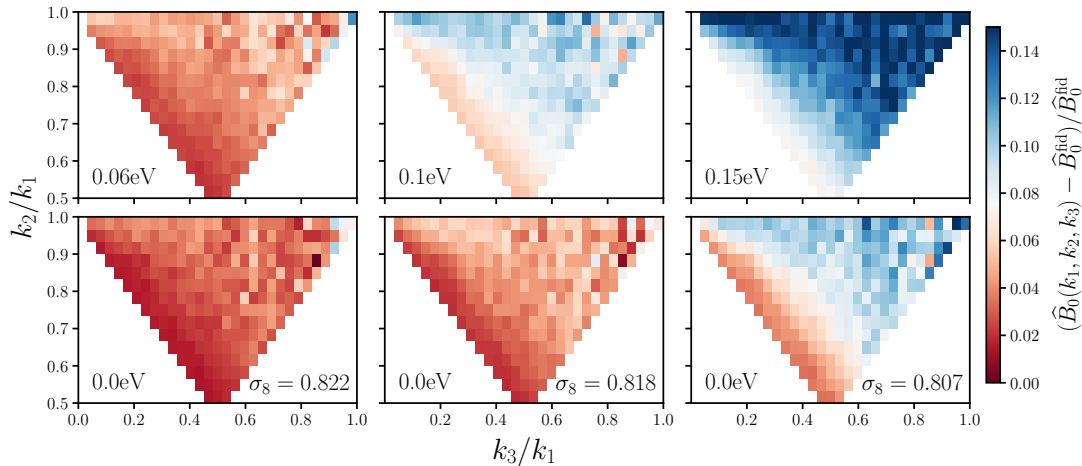
where

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

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.<sup>8</sup> 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  $\Delta k = 3k_f = 0.01885$   $h/\text{Mpc}$ , three times the fundamental mode  $k_f = 2\pi/(1000$   $h/\text{Mpc})$  given the box size.

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





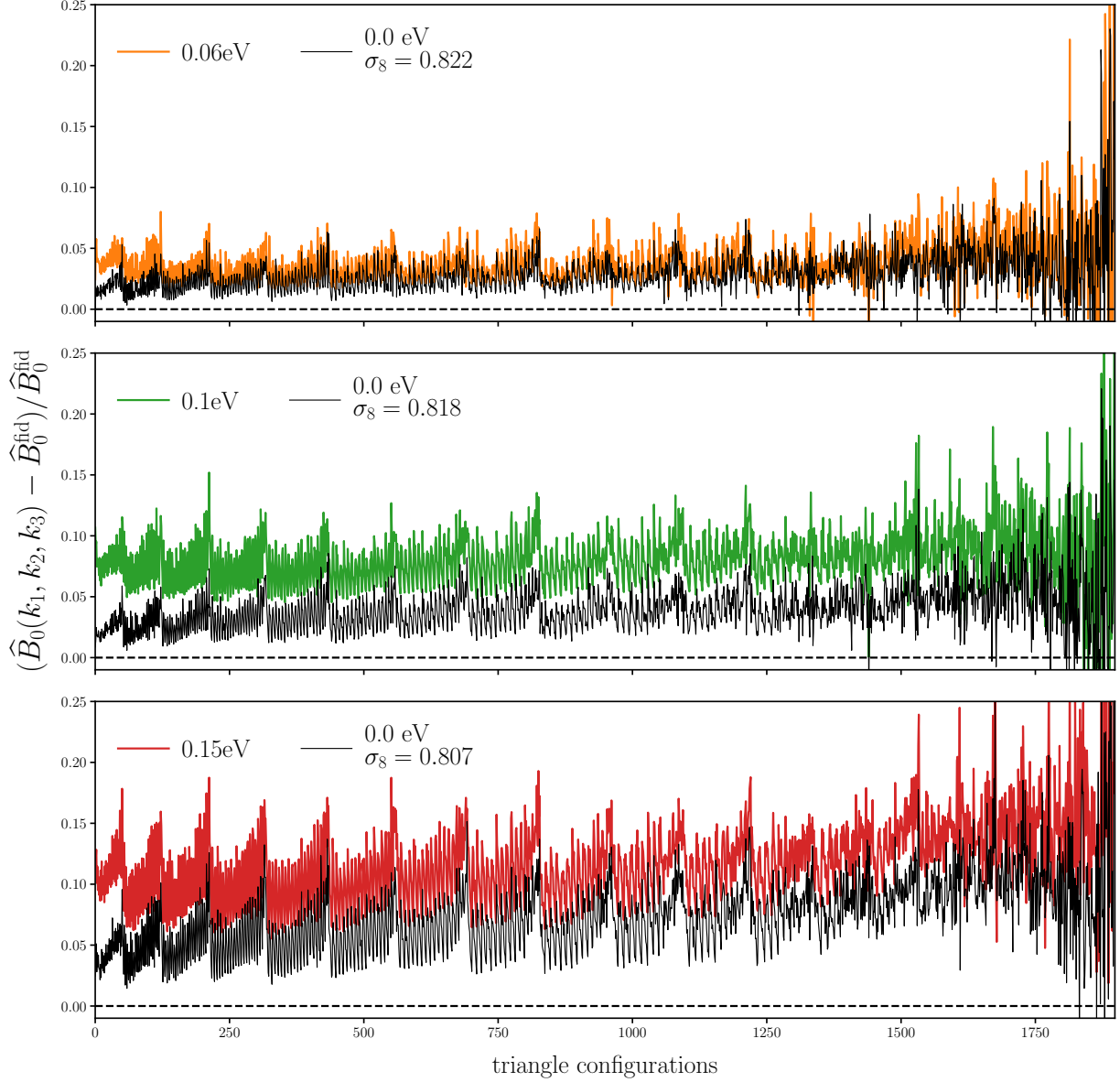
**Figure 4.** The shape dependence of the  $M_\nu$  and  $\sigma_8$  imprint on the redshift-space halo bispectrum,  $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ . We align the  $M_\nu = 0.06, 0.10$ , and  $0.15$  eV HADES models in the upper panels with  $M_\nu = 0.0$  eV  $\sigma_8 = 0.822, 0.818$ , and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

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  $\hat{B}_0(k_1, k_2, k_3)$  as a function of  $k_2/k_1$  and  $k_3/k_1$ , which describe the triangle configuration shape. In each panel, the colormap in each  $(k_2/k_1, k_3/k_1)$  bin is the weighted average  $\hat{B}_0$  amplitude of all triangle configurations in the bin. 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 triangle configurations with  $k_1, k_2, k_3 < k_{\text{max}} = 0.5$  h/Mpc. The  $\hat{B}_0$  in the upper panels are HADES models with  $M_\nu = 0.0$  (fiducial),  $0.06, 0.10$ , and  $0.15$  eV;  $\hat{B}_0$  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  $\sigma_8$  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  $k_1 \leq k_2 \leq k_3$ . 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  $k_1 = 0.113$ ,  $0.226 \leq k_2 \leq 0.283$ , and  $0.283 \leq k_3 \leq 0.377$  h/Mpc in the insets of the panels.

## 4. RESULTS

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



**Figure 5.** The impact of  $M_\nu$  and  $\sigma_8$  on the redshift-space halo bispectrum,  $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$ , for all 1898 triangle configurations with  $k_1, k_2, k_3 \leq 0.5h/\text{Mpc}$ . 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 models to  $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$  of  $M_\nu = 0.0$  eV  $\sigma_8 = 0.822, 0.818$ , and  $0.807$  models. The impact of  $M_\nu$  on the bispectrum has a significantly different amplitude than the impact of  $\sigma_8$ . 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 powerspectrums only differ by  $< 1\%$  (Figure 1). Combined with the different shape-dependence (Figure 4), the distinct imprint of  $M_\nu$  on the bispectrum illustrate that the bispectrum can break the degeneracy between  $M_\nu$  and  $\sigma_8$  that degrade constraints from two-point analyses.

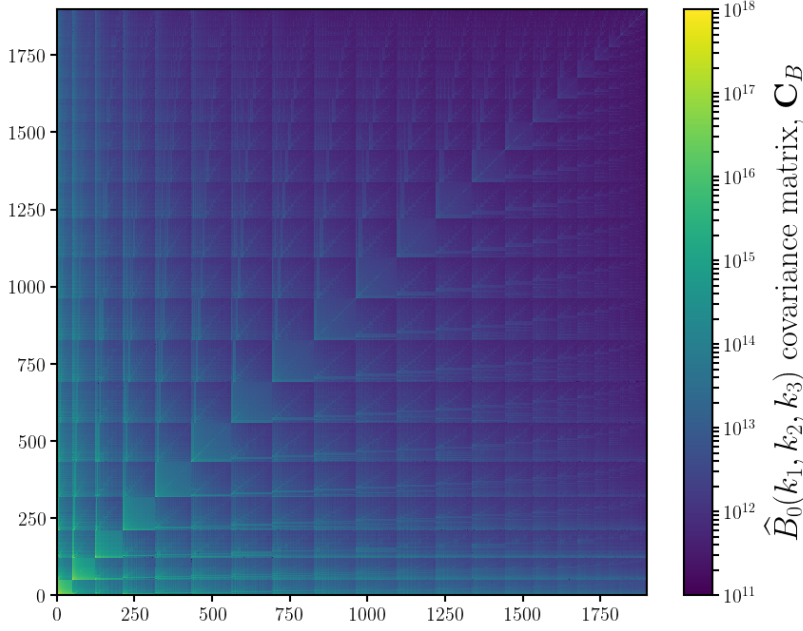
One major bottleneck of constraining  $M_\nu$  with the power spectrum alone is the strong  $M_\nu - \sigma_8$  degeneracy. The imprint of  $M_\nu$  and  $\sigma_8$  on the power spectrum are degenerate and for models with the same  $\sigma_8^c$ , the power spectrum only differ by  $< 1\%$  (see Figure 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  $\sigma_8^c$  —  $\sigma_8 = 0.822, 0.818$ , and  $0.807$ , provide an ideal set of simulations to separate the impact of  $M_\nu > 0.0$  eV and examine the degeneracy between  $M_\nu$  and  $\sigma_8$  (Section 2 and Table 1). Hence, by measuring bispectrum of these simulations (Figure 2 and 3), we can determine whether the bispectrum helps break the  $M_\nu - \sigma_8$  degeneracy. Below, we present our comparison of the HADES bispectrum and illustrate that the bispectrum can significantly improve  $M_\nu$  constraints by breaking the  $M_\nu - \sigma_8$  degeneracy.

We begin by examining the triangle shape dependent imprint of  $M_\nu$  on the redshift-space halo bispectrum versus  $\sigma_8$  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  $\sigma_8^c$ . Overall as  $M_\nu$  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  $M_{\text{lim}}$  on our halos so lower values of  $\sigma_8$  translate to a larger halo bias, which boosts the amplitude of the bispectrum. Within the overall increase in amplitude, however, equilateral triangles (upper left) have the largest increase. For  $M_\nu = 0.15$  eV, the bispectrum is  $\sim 15\%$  higher than  $\hat{B}_0^{\text{fid}}$  for equilateral triangles. Meanwhile, the bispectrum increases by  $\sim 8\%$  for folded triangles for  $0.15$  eV (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  $\sigma_8$  increases, with  $M_\nu = 0.0$  eV fixed, the bispectrum 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  $M_\nu$  versus  $\sigma_8$  alone. Between  $M_\nu = 0.15$  eV and  $0.0$  eV  $\sigma_8 = 0.807$ , there is an overall  $\gtrsim 5\%$  difference. In addition, the shape dependence of the  $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$  increase is different for  $M_\nu$  than  $\sigma_8$ . This is particularly clear in the differences between  $0.1$  eV (top center panel) and  $0.0$  eV and  $\sigma_8 = 0.807$  (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,  $M_\nu$  leaves an imprint on the bispectrum with a distinct triangle shape dependence than  $\sigma_8$  alone. In other words, unlike the power spectrum, the triangle shape dependent impact of  $M_\nu$  on the bispectrum cannot be replicated by varying  $\sigma_8$ .

We next examine the amplitude of the  $M_\nu$  imprint on the redshift-space halo bispectrum versus  $\sigma_8$  alone for all triangle configurations. We present  $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$  for all 1898 possible triangle configurations with  $k_1, k_2, k_3 < k_{\text{max}} = 0.5 h/\text{Mpc}$  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  $M_\nu$  and  $\sigma_8$  (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).



**Figure 6.** Covariance matrix of the redshift-space halo bispectrum estimated using  $N_{\text{cov}} = 15,000$  realizations of the Quijote simulation suite at the fiducial cosmology:  $\Omega_{\text{m}}=0.3175$ ,  $\Omega_{\text{b}}=0.049$ ,  $h=0.6711$ ,  $n_{\text{s}}=0.9624$ ,  $\sigma_8=0.834$ , 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.

The comparison in the panels of Figure 5 also reveal a difference in the configuration dependence in  $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$  between  $M_{\nu}$  versus  $\sigma_8$ . 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.  $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$  of  $M_{\nu}$  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  $M_{\nu}$  has a smaller  $k_1$  dependence than  $\Delta\hat{B}_0/\hat{B}_0^{\text{fid}}$  of  $\sigma_8$ . Combined with the different shape-dependence (Figure 4), the distinct imprint of  $M_{\nu}$  on the redshift-space halo bispectrum illustrates that the bispectrum can break the degeneracy between  $M_{\nu}$  and  $\sigma_8$ . Moreover it illustrates that by including the bispectrum, we can more precisely constrain  $M_{\nu}$  than with the power spectrum alone.

#### 4.2. $M_{\nu}$ and other Cosmological Parameter Forecasts

We demonstrate in the previous section with the HADES simulations, that the bispectrum helps break the  $M_{\nu}$ – $\sigma_8$  degeneracy, a major challenge in precisely constraining  $M_{\nu}$  with the power spectrum. This establishes the bispectrum as a promising probe for  $M_{\nu}$ . However, we are ultimately interested in determining the constraining power of the bispectrum for an analysis that include cosmological parameters beyond  $M_{\nu}$  and  $\sigma_8$ — *i.e.*  $\Omega_{\text{m}}$ ,  $\Omega_{\text{b}}$ ,  $h$ , and  $n_{\text{s}}$ . The Quijote suite of simulations 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,  $\Omega_{\text{m}}=0.3175$ ,  $\Omega_{\text{b}}=0.049$ ,  $n_{\text{s}}=0.9624$ ,  $h=0.6711$ , and  $\sigma_8 = 0.834$  (see Table 1. This allows

us to robustly estimate the covariance matrix of the bispectrum,  $\mathbf{C}$ , which has  $\sim 1,800$  triangle configurations (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). On top of this, we apply redshift-space distortions along 3 different directions for these 500 realizations, which then gives us  $N_{\text{deriv.}} = 1,500$  realizations. These realizations 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

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

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

$$F_{ij} = \frac{1}{2} \text{Tr} \left[ \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_i} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_j} + \mathbf{C}^{-1} \left( \frac{\partial \bar{B}_0}{\partial \theta_i} \frac{\partial \bar{B}_0}{\partial \theta_j}^T + \frac{\partial \bar{B}_0}{\partial \theta_i}^T \frac{\partial \bar{B}_0}{\partial \theta_j} \right) \right]. \quad (9)$$

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

$$F_{ij} = \frac{1}{2} \text{Tr} \left[ \mathbf{C}^{-1} \left( \frac{\partial \bar{B}_0}{\partial \theta_i} \frac{\partial \bar{B}_0}{\partial \theta_j}^T + \frac{\partial \bar{B}_0}{\partial \theta_i}^T \frac{\partial \bar{B}_0}{\partial \theta_j} \right) \right], \quad (10)$$

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

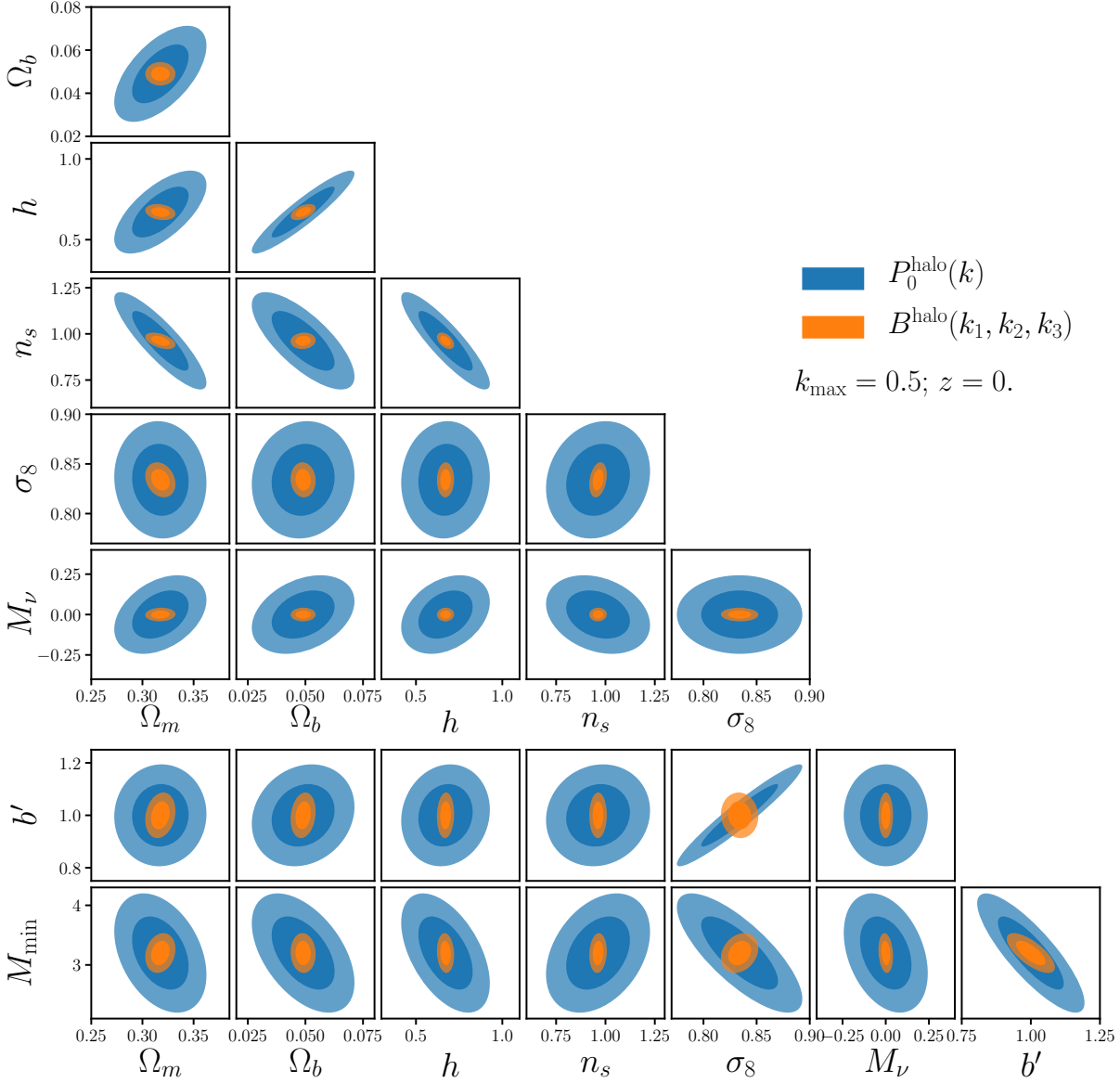
For  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$ , we estimate

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

where  $\bar{B}_0(\theta_i^+)$  and  $\bar{B}_0(\theta_i^-)$  are the average bispectrum of the 1,500 realizations at  $\theta^+$  and  $\theta^-$ . Meanwhile, for  $M_\nu$ , where the fiducial value is 0.0 eV and we cannot have negative  $M_\nu$ , we use the Quijote simulations at  $M_\nu^+$ ,  $M_\nu^{++}$ ,  $M_\nu^{+++} = 0.1, 0.2, 0.4$  eV (Table 1) to estimate

$$\frac{\partial \bar{B}_0}{\partial M_\nu} \approx \frac{-21\bar{B}_0(M_\nu^{\text{fid}}) + 32\bar{B}_0(M_\nu^+) - 12\bar{B}_0(M_\nu^{++}) + \bar{B}_0(M_\nu^{+++})}{1.2}, \quad (12)$$

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 numerical simulations in the nonlinear regime and rely on fewer assumptions and approximations. In fact, *these  $N$ -body simulation estimated derivatives are the key ingredients that enables us to quantify, for the first time, the full information content of the redshift-space bispectrum in the non-linear regime.* We discuss subtleties of Eq. 12 bispectrum derivative and tests of convergence in Appendix B.



**Figure 7.** Fisher matrix constraints for  $M_\nu$  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 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  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$  improve by factors of 3.1, 4.1, 5.1, 5.9, and 3.3, respectively. For  $M_\nu$ , the bispectrum improves  $\sigma_{M_\nu}$  from 0.1962 to 0.0342 eV — a factor of  $\sim 6$  improvement over the power spectrum.



We present the constraints on  $M_\nu$  and other cosmological parameters  $\{\Omega_m, \Omega_b, h, n_s, \sigma_8\}$  derived from the redshift-space halo bispectrum Fisher matrix (Eq. 10) for  $k_{\max} = 0.5 \text{ h/Mpc}$  in Figure 7, respectively. We include Fisher constraints for the redshift-space halo power spectrum monopole with the same  $k_{\max}$  for comparison (blue). We mark the 68% and 95% confidence intervals with the contours. We include in our Fisher constraints the following nuisance parameters:  $b'$ , a scaling factor on the bispectrum amplitude, and  $M_{\min}$ , the halo mass limit.  $b'$  is analogous to linear bias. Meanwhile, we choose  $M_{\min}$  as a nuisance parameter to address the difference in the number densities among the Quijote cosmologies, which impacts the derivatives  $\partial \bar{B}_0 / \partial \theta_i$ . For instance, the  $\sigma_8^+$  and  $\sigma_8^-$  cosmologies have halo  $\bar{n} = 1.586 \times 10^{-4}$  and  $1.528 \times 10^{-4} (h^{-1} \text{Mpc})^3$ . 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  $\partial \bar{B}_0 / \partial b' = \bar{B}_0$ . Meanwhile, we numerically estimate  $\partial \bar{B}_0 / \partial M_{\min}$  using  $\bar{B}_0$  evaluated at  $M_{\min}^+ = 3.3 \times 10^{13} h^{-1} M_\odot$  and  $M_{\min}^- = 3.1 \times 10^{13} h^{-1} M_\odot$ , with all other parameters set to the fiducial value.

The bispectrum *substantially* improves constraints on all parameters over the power spectrum. For  $k_{\max} = 0.5 \text{ h/Mpc}$ , the bispectrum improves the marginalized constraints,  $\sigma_\theta$  of  $\Omega_m, \Omega_b, h, n_s$ , and  $\sigma_8$  by factors of  $\sim 3, 4, 5, 6$ , and  $3$  over the power spectrum. *For  $M_\nu$ , the bispectrum improves the constraint from  $\sigma_{M_\nu} = 0.1962$  to  $0.0342 \text{ eV}$  — a factor of 6 improvement over the power spectrum.* We emphasize that this  $\sigma_{M_\nu} = 0.0342 \text{ eV}$  constraint is for the *bispectrum alone* and only for a  $1 h^{-1} \text{Gpc}$  box. 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  $k_{\max} < 0.5 \text{ h/Mpc}$ , the bispectrum significantly improves cosmological parameter constraints. We compare  $\sigma_\theta$ , the marginalized  $1\sigma$  constraints of  $\Omega_m, \Omega_b, h, n_s, \sigma_8$ , and  $M_\nu$ , as a function of  $k_{\max}$  for  $B_0$  (orange) and  $P_0$  (blue) in Figure 8. We focus only on the  $k_{\max}$  range where the Fisher forecast is well defined — *i.e.* more data bins than the number of parameters:  $k_{\max} > 8k_f \approx 0.05 \text{ h/Mpc}$  for  $P_0$  and  $k_{\max} > 12k_f \approx 0.075 \text{ h/Mpc}$  for  $B_0$ . Figure 8 reveals that the improvement of the bispectrum  $\sigma_\theta$  over the power spectrum  $\sigma_\theta$  is larger at higher  $k_{\max}$ . Although limited by the  $k_{\max}$  range, the figure suggests that on large scales ( $k_{\max} \lesssim 0.1 \text{ h/Mpc}$ )  $P_0 \sigma_\theta$  crosses over  $B_0 \sigma_\theta$  so  $P_0$  has more constraining power than  $B_0$  — as expected on linear scales. On slightly larger  $k_{\max}$ , even at  $k_{\max} = 0.2 \text{ h/Mpc}$ , we find that the bispectrum substantially improves  $\sigma_\theta$  by factors of  $\sim 1.8, 2.1, 2.6, 2.5, 2.6$ , and  $3.1$  for  $\Omega_m, \Omega_b, h, n_s, \sigma_8$ , and  $M_\nu$  respectively.

Our forecasts demonstrate that the bispectrum has significant constraining power in the weakly non-linear regime ( $k > 0.1 \text{ h/Mpc}$ ) beyond the power spectrum. 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  $k < 0.4 \text{ h/Mpc}$ , the unmarginalized constraints 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 tighter than  $P_0$  for  $k > 0.1 \text{ h/Mpc}$ . The derivatives,  $\partial B_0 / \partial \theta_i$ , also shed light on how  $B_0$  breaks parameter degeneracies. The parameter degeneracies in the  $P_0$  forecasts of Figure 7 are consistent with similarities among the shape and scale dependence of  $P_0$  derivatives  $\partial P_0 / \partial \theta$ . On

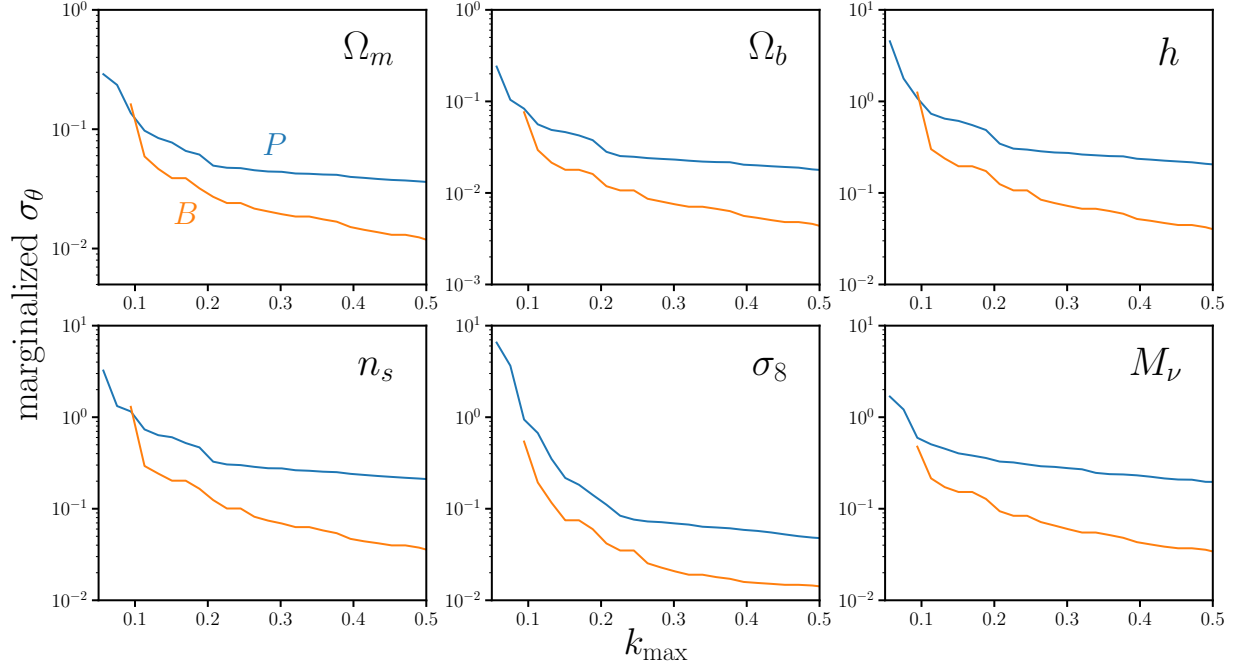
**Table 2.** Marginalized Fisher parameter constraints from the redshift-space halo power spectrum (top) and bispectrum (bottom) for different values of  $k_{\max}$ . We list constraints for cosmological parameters  $M_\nu$ ,  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$  as well as nuisance parameters  $b'$  and  $M_{\min}$

|       | $k_{\max}$<br>( $h/\text{Mpc}$ ) | $M_\nu$<br>(eV)      | $\Omega_m$  | $\Omega_b$  | $h$         | $n_s$       | $\sigma_8$  | $b'$        | $M_{\min}$<br>( $10^{13}h^{-1}M_\odot$ ) |
|-------|----------------------------------|----------------------|-------------|-------------|-------------|-------------|-------------|-------------|--|
|       |                                  | 0.0                  | 0.3175      | 0.049       | 0.6711      | 0.9624      | 0.834       | 1.          | 3.2                                      |
| $P_0$ | 0.2                              | $\pm 0.333$          | $\pm 0.052$ | $\pm 0.030$ | $\pm 0.372$ | $\pm 0.347$ | $\pm 0.128$ | $\pm 0.649$ | $\pm 5.045$                              |
|       | 0.3                              | $\pm 0.277$          | $\pm 0.044$ | $\pm 0.023$ | $\pm 0.273$ | $\pm 0.276$ | $\pm 0.069$ | $\pm 0.383$ | $\pm 2.457$                              |
|       | 0.4                              | $\pm 0.228$          | $\pm 0.040$ | $\pm 0.020$ | $\pm 0.235$ | $\pm 0.240$ | $\pm 0.059$ | $\pm 0.226$ | $\pm 1.270$                              |
|       | 0.5                              | $\pm \mathbf{0.196}$ | $\pm 0.036$ | $\pm 0.018$ | $\pm 0.207$ | $\pm 0.213$ | $\pm 0.048$ | $\pm 0.157$ | $\pm 0.807$                              |
| $B_0$ | 0.2                              | $\pm 0.107$          | $\pm 0.029$ | $\pm 0.014$ | $\pm 0.144$ | $\pm 0.140$ | $\pm 0.050$ | $\pm 0.265$ | $\pm 1.317$                              |
|       | 0.3                              | $\pm 0.065$          | $\pm 0.020$ | $\pm 0.008$ | $\pm 0.077$ | $\pm 0.074$ | $\pm 0.023$ | $\pm 0.143$ | $\pm 0.657$                              |
|       | 0.4                              | $\pm 0.043$          | $\pm 0.015$ | $\pm 0.006$ | $\pm 0.052$ | $\pm 0.047$ | $\pm 0.016$ | $\pm 0.088$ | $\pm 0.369$                              |
|       | 0.5                              | $\pm \mathbf{0.034}$ | $\pm 0.012$ | $\pm 0.004$ | $\pm 0.040$ | $\pm 0.036$ | $\pm 0.014$ | $\pm 0.070$ | $\pm 0.269$                              |

the other hand, the  $B_0$  derivatives with respect to the parameters have significant different scale and triangle shape dependences. **CH: derivative figure?**

By exploiting the unprecedented 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, however, 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 (*e.g.* Sefusatti & Scoccimarro 2005; Chan & Blot 2017). Most recently, Chan & Blot (2017), using covariance matrices that include non-Gaussian contributions calibrated with  $N$ -body simulations, find that the cumulative SNR of the halo bispectrum is  $\sim 30\%$  of the SNR of the halo power spectrum at  $k_{\max} \sim 0.1 h/\text{Mpc}$  and increases to  $\sim 40\%$  at  $k_{\max} \sim 0.35 h/\text{Mpc}$ . While these simple SNR measurements cannot be easily compared to our Fisher analysis, we note that they are loosely consistent with the unmarginalized constraints, where we find tighter unmarginalized constraints from  $P_0$  than  $B_0$  at  $k_{\max} < 0.4 h/\text{Mpc}$ . Also, when we measure 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 of Chan & Blot (2017) ( $k_{\max} > 0.35 h/\text{Mpc}$ ), 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  $k_{\max} \sim 0.1 h/\text{Mpc}$ . At  $k_{\max} = 0.75 h/\text{Mpc}$ , the largest  $k$  we probe, the  $B_0$  SNR is  $\sim 75\%$  of the  $P_0$  SNR.

Beyond these signal-to-noise calculations, a number of previous works quantify 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 et al. 2018; Yankelevich & Porciani 2019). While most of these works fix most cosmological parameters and focus solely on forecasting constraints of primordial non-Gaussianity ( $f_{\text{NL}}$ ) and bias parameters,



**Figure 8.** Marginalized  $1\sigma$  constraints of the cosmological parameters  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$ , and  $M_\nu$  ( $\sigma_\theta$ ) as a function of  $k_{\max}$  for the redshift-space halo bispectrum (orange) and power spectrum (blue). Though not included in the figure, we marginalize over the nuisance parameters  $b'$  and  $M_{\min}$  in our forecast (Section 4.2). We only include  $k_{\max} > 8k_f$  for  $P_0$  and  $k_{\max} > 12k_f$  for  $B_0$  —  $k_{\max}$  ranges where we have more data bins than number of parameters. Even at  $k_{\max} < 0.5$   $h/\text{Mpc}$ , 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  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$ , and  $M_\nu$  by factors of  $\sim 1.8$ ,  $2.1$ ,  $2.6$ ,  $2.5$ ,  $2.6$ , and  $3.1$  over the power spectrum.

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  $\omega_d$ ,  $\omega_b$ ,  $\Omega_\Lambda$ ,  $n_s$ ,  $A_s$ ,  $w$ ,  $\tau$ . For  $\Lambda\text{CDM}$ , with fixed bias parameters, and  $k_{\max} = 0.3$   $h/\text{Mpc}$ , they find constraints on  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$  from WMAP,  $P_0$ , and  $B_0$  is a factor of 1.71, 1.3, 1.4, 1.2, and 1.5 tighter than constraints from WMAP and  $P_0$  alone. In comparison, for  $k_{\max} = 0.3$   $h/\text{Mpc}$  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 only present constraints with both  $P_0$  and  $B_0$ . We also emphasize that they use perturbation theory bispectrum models, which break down on small scales. In a comparison with  $N$ -body simulations, Lazanu et al. (2016) find that perturbation theory models of the matter bispectrum deviate by  $> 5\%$  at  $k_{\max} \gtrsim 0.15$   $h/\text{Mpc}$ .

Meanwhile, in Yankelevich & Porciani (2019), they present Fisher forecasts for  $\Omega_{\text{cdm}}$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $A_s$ ,  $w_0$ , 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) that depend on the tidal

field. They use a perturbative model for the bispectrum, which consequently limit their forecast to  $k_{\text{max}} = 0.15 \text{ h/Mpc}$ . From their forecasts, 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  $k_{\text{max}}$ , consistent with our forecasts as a function of  $k_{\text{max}}$  (Figure 8). Also, similar to their forecast, for  $k_{\text{max}} = 0.15 \text{ h/Mpc}$ , we also find similar posterior correlations between the  $P_0$  and  $B_0$  constraints. At  $k_{\text{max}} = 0.5 \text{ h/Mpc}$ , however, we find significantly less correlations (Figure 7). Based on our forecasts, at higher  $k_{\text{max}}$  adding  $B_0$  to the analysis improve parameter constraints beyond the improvement found in Yankelevich & Porciani (2019).

While the various difference between our forecast and previous work prevent more thorough comparisons, the main difference is that we present the first bispectrum forecasts for a full set of cosmological parameters using bispectrum measurements entirely from  $N$ -body simulations. This allows us to go beyond previous perturbation theory based forecasts and quantify the full information content of the redshift-space bispectrum all the way out to the non-linear regime. Moreover, we also present the first bispectrum forecast of cosmological parameters including neutrinos and demonstrate the potential of the bispectrum in constraining  $M_\nu$ . 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 convergence, of numerical derivatives. We examine the stability of the  $P_0$  and  $B_0$  derivatives with respect to  $M_\nu$  by comparing the derivatives computed using  $N$ -body simulations at three different sets of cosmologies: (1) {fiducial,  $M_\nu^+$ ,  $M_\nu^{++}$ , and  $M_\nu^{+++}$ } (Eq. 12), (2) {fiducial,  $M_\nu^+$ , and  $M_\nu^{++}$ }, and (3) {fiducial and  $M_\nu^+$ } (Appendix B; Figure 11). Compared to the derivatives computed with the (1) cosmologies, the derivatives with (2) and (3) differ by  $\sim 20$  and  $\sim 50\%$  for both  $P_0$  and  $B_0$ . These differences do not impact the  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$  constraints. They do however affect the  $M_\nu$  constraints; but because  $P_0$  and  $B_0$  derivatives are scaled by the same factor, *the relative improvement of the  $B_0$  constraint over the  $P_0$  constraints is not impacted*.

In addition to the stability of the derivatives, 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:  $N_{\text{cov}}$  and  $N_{\text{deriv}}$ , respectively. For  $N_{\text{cov}}$ , we find  $< 5\%$  variation in the Fisher matrix elements,  $F_{i,j}$  for  $N_{\text{cov}} > 5000$ . Meanwhile, the marginalized parameter constraints,  $\sigma_\theta$ , vary by  $< 1\%$  for  $N_{\text{cov}} > 12000$ . When we vary  $N_{\text{deriv}}$ , we find  $< 5\%$  variation in the  $F_{i,j}$  elements

and  $\sigma_\theta$  for  $N_{\text{deriv}} > 1200$ . Since our constraints vary by  $< 10\%$  for sufficient  $N_{\text{cov}}$  and  $N_{\text{deriv}}$ , we conclude that the convergence of the covariance matrix and derivatives do not impact our forecasts. 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  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ , and  $\sigma_8$  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. By including Planck priors (*i.e.* adding the prior matrix to the Fisher matrix) increases the numerical stability of our forecasts. It also reveals whether the bispectrum still improves parameter constraints once we include CMB constraints. With Planck priors and  $P_0$  to  $k_{\text{max}} = 0.5 \text{ h/Mpc}$ , we derive the following constraints:  $\sigma_{\Omega_m} = 0.0125$ ,  $\sigma_{\Omega_b} = 0.0012$ ,  $\sigma_h = 0.0086$ ,  $\sigma_{n_s} = 0.0048$ , 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  $k_{\text{max}} = 0.5 \text{ h/Mpc}$ , we get  $\sigma_{\Omega_m} = 0.0086$ ,  $\sigma_{\Omega_b} = 0.0008$ ,  $\sigma_h = 0.0006$ ,  $\sigma_{n_s} = 0.0043$ , and  $\sigma_{\sigma_8} = 0.0085$ , 1.5, 1.6, 1.4, 1.1, and 2.3 times tighter constraints. For  $M_\nu$ ,  $\sigma_{M_\nu} = 0.0802 \text{ eV}$  for  $P_0$  and  $\sigma_{M_\nu} = 0.0297 \text{ eV}$ , a factor of 2.7 improvement. Hence, even with Planck priors,  $B_0$  improves substantially parameter constraints, especially for  $M_\nu$ . This also suggests that the improvement from  $B_0$  we find in our forecasts are numerically robust.

So far in our analysis, 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  $M_\nu$  constraints of the bispectrum, even with survey geometry we expect the bispectrum will significantly improve  $M_\nu$  constraints over the power spectrum.

We include the nuisance parameter  $M_{\text{min}}$  in our forecasts to address the difference in halo number densities among the Quijote cosmologies. Although we marginalize over  $M_{\text{min}}$ , 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,  $A_{\text{SN}}$ ,  $B_{\text{SN}}$ ,  $b_2$ , and  $\gamma_2$ , and examine their impact on our forecasts.  $A_{\text{SN}}$  and  $B_{\text{SN}}$  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  $B_0$  shotnoise correction.  $b_2$  and  $\gamma_2$  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  $b'$ ,  $M_{\text{min}}$ ,  $A_{\text{SN}}$ ,  $B_{\text{SN}}$ ,  $b_2$  and  $\gamma_2$ , we obtain the following constraints from  $B_0$  with  $k_{\text{max}} = 0.5 \text{ h/Mpc}$ :  $\sigma_{\Omega_m} = 0.0129$ ,  $\sigma_{\Omega_b} = 0.0044$ ,  $\sigma_h = 0.0404$ ,  $\sigma_{n_s} = 0.0455$ ,  $\sigma_{\sigma_8} = 0.0228$ , and  $\sigma_{M_\nu} = 0.0343$ . Only constraints on  $n_s$  and  $\sigma_8$  are broadened from our fiducial forecasts, by 27% and 60%. Other parameters, especially  $M_\nu$ , are not impacted by marginalizing over the extra nuisance parameters. As another test to further isolate

the impact of the differing  $\bar{n}$ , we calculate derivatives using halo catalogs from Quijote  $\theta^-$  and  $\theta^+$  cosmologies—with fixed  $\bar{n}$ . We similarly find no significant impact on the parameter constraints from  $B_0$ . 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  $M_\nu$  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  $M_\nu$  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  $\gamma_2$  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.

Marginalizing over galaxy bias parameters, will likely reduce the constraining power at high  $k$ . Hand et al. (2017), for instance, find for their 13 parameter model only find a 15-30% improvement in  $f\sigma_8$  when they extend their power spectrum multipole analysis from  $0.2 h/\text{Mpc}$  to  $0.4 h/\text{Mpc}$ . Furthermore, 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  $1h^{-1}\text{Gpc}$  box, a substantially volume than upcoming surveys. Thus, even if the constraining power at high  $k$  is reduced, our forecasts suggest that the bispectrum still offers significant improvements over the power spectrum, especially for constraining  $M_\nu$ .

## 5. SUMMARY

**CH:** talk about DESI, PFS, WFIRST

## ACKNOWLEDGEMENTS

It's a pleasure to thank Enea Di Dio, Daniel Eisenstein, Simone Ferraro, Shirley Ho, Emmaneul Schaan, Zachary Slepian, David N. Spergel, and Benjamin D. Wandelt for valuable discussions and comments.

## APPENDIX

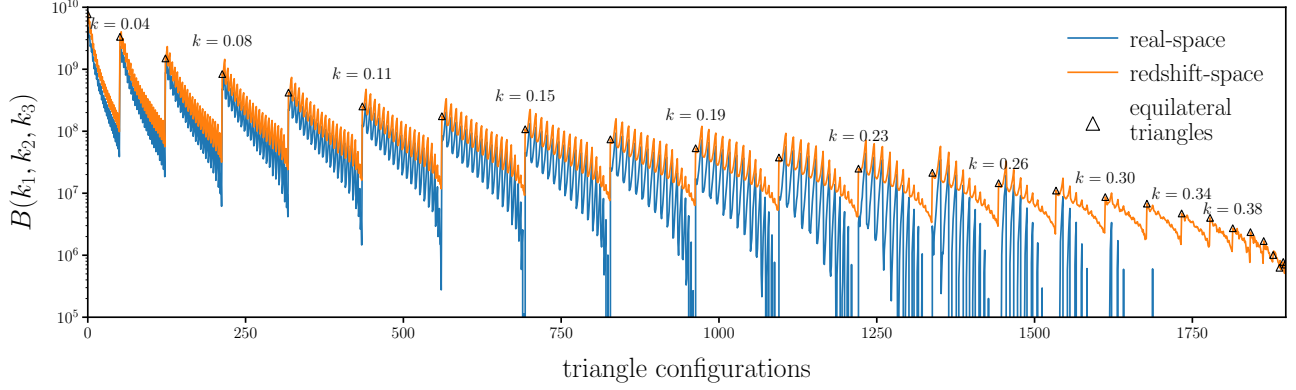
### A. REDSHIFT-SPACE HALO BISPECTRUM

**CH:** do we want to keep this section?

### B. FISHER FORECASTS USING $N$ -BODY SIMULATIONS

The two key elements in calculating the Fisher matrices we use in our forecasts are the bispectrum covariance matrix ( $\mathbf{C}$ ; Figure 6) and the derivatives of the bispectrum along the cosmological and





**Figure 9.** Comparison of the fiducial HADES simulations real and redshift-space halo bispectrum for triangle configurations with  $k_1, k_2, k_3 \leq k_{\max} = 0.5h/\text{Mpc}$  (blue and orange respectively). We mark equilateral triangle configurations (empty triangle marker) along with their side lengths  $k$ .

nuisance parameters,  $\partial B_0/\partial\theta_i$  (Section 4.2). We compute both these elements directly using the  $N$ -body simulations of the Quijote suite (Section 2). This takes advantage of the  $N$ -body simulations and allows us to accurately quantify the constraining power of the bispectrum that come from the nonlinear regime. However, to trust our forecast, we must ensure that both  $\mathbf{C}$  has converged and that the numerically calculated  $\partial B_0/\partial\theta_i$  do not introduce any biases. Below, we test the convergence of  $\mathbf{C}$  and  $\partial B_0/\partial\theta_i$  and discuss some of the subtleties and caveats of numerically calculating  $\partial B_0/\partial\theta_i$  from the Quijote simulations. **CH: mention somewhere the P convergence looks good**

To estimate  $\mathbf{C}$ , we use 15,000 Quijote  $N$ -body simulations at the fiducial cosmology. This is a *significantly* larger number of simulations than previous bispectrum analyses; however, we also consider a larger number of triangle configurations — 1898 triangles out to  $k_{\max} = 0.5 h/\text{Mpc}$ . For reference, the recent Gil-Marín et al. (2017) analysis used 2048 simulations to estimate the covariance matrix of the bispectrum with 825 configurations. We, therefore, check the convergence of  $\mathbf{C}$  by varying  $N_{\text{fid}}$ , the number of simulations used to estimate  $\mathbf{C}$ , and examining whether this significantly impacts the Fisher parameter constraints. We present  $\sigma_\theta(N_{\text{fid}})/\sigma_\theta(N_{\text{fid}} = 15,000)$ , the ratio of the  $1\sigma$  Fisher constraint for  $\theta = \Omega_m, \Omega_b, h, n_s, \sigma_8$ , and  $M_\nu$  calculated with  $N_{\text{fid}}$  over the constraint calculated with  $N_{\text{fid}} = 15,000$ , as a function of  $N_{\text{fid}}$  (Figure 10 left panel). The  $1\sigma$  Fisher constraints on the parameters vary by  $< 10\%$  for  $N_{\text{fid}} > 7000$ ; in fact, the constraints vary by  $< 1\%$  for  $N_{\text{fid}} > 14,000$ . Hence, we conclude that we have a sufficient number of simulations to estimate the bispectrum  $\mathbf{C}$  and our forecasts are robust to the convergence of  $\mathbf{C}$ .

We estimate  $\partial B_0/\partial\theta_i$  numerically using 13 sets of  $N_{\text{fp}} = 500$  fixed paired simulations (Table 1). To check the convergence of  $\partial B_0/\partial\theta_i$  and its impact on our forecast we present the ratio of the  $1\sigma$  Fisher constraint for  $\theta$  calculated using  $N_{\text{fp}}$  simulations over the constraint calculated with  $N_{\text{fp}} = 500$ ,  $\sigma_\theta(N_{\text{fp}})/\sigma_\theta(N_{\text{fp}} = 500)$ , as a function of  $N_{\text{fp}}$  (Figure 10 right panel). Unlike  $\sigma_\theta(N_{\text{fid}})$ ,  $\sigma_\theta(N_{\text{fp}})$  depend significantly on  $\theta$ . For instance,  $\sigma_\theta$  for  $\sigma_8$  and  $\Omega_m$  vary by  $< 10\%$  for  $N_{\text{fp}} > 300$  and  $< 2\%$  for  $N_{\text{fp}} > 450$ .  $\sigma_\theta$  for the other parameters vary significantly more. Nonetheless, for  $N_{\text{fp}} > 400$  and  $450$  they vary by  $< 10$  and  $5\%$ , respectively.

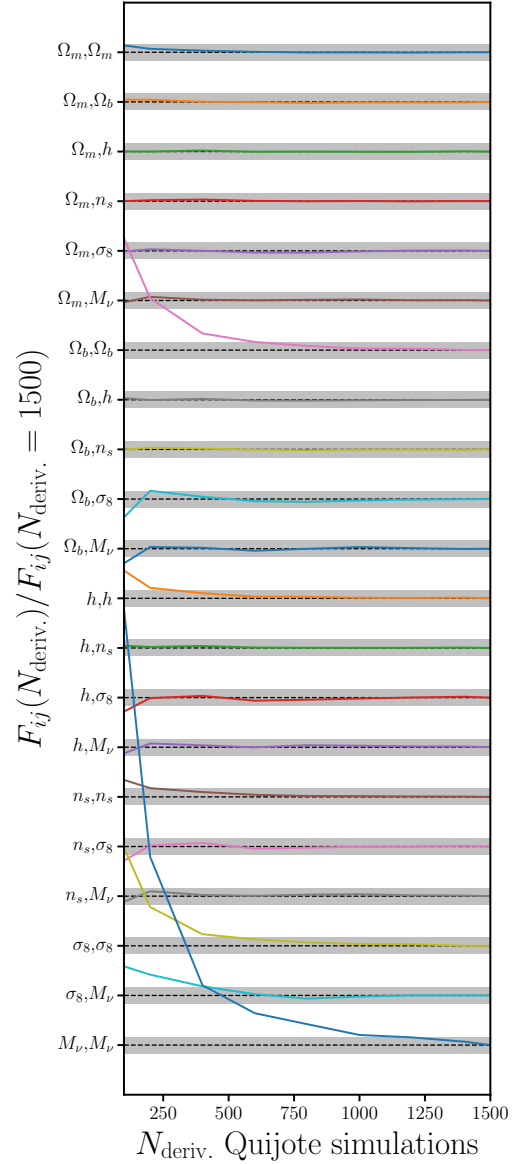
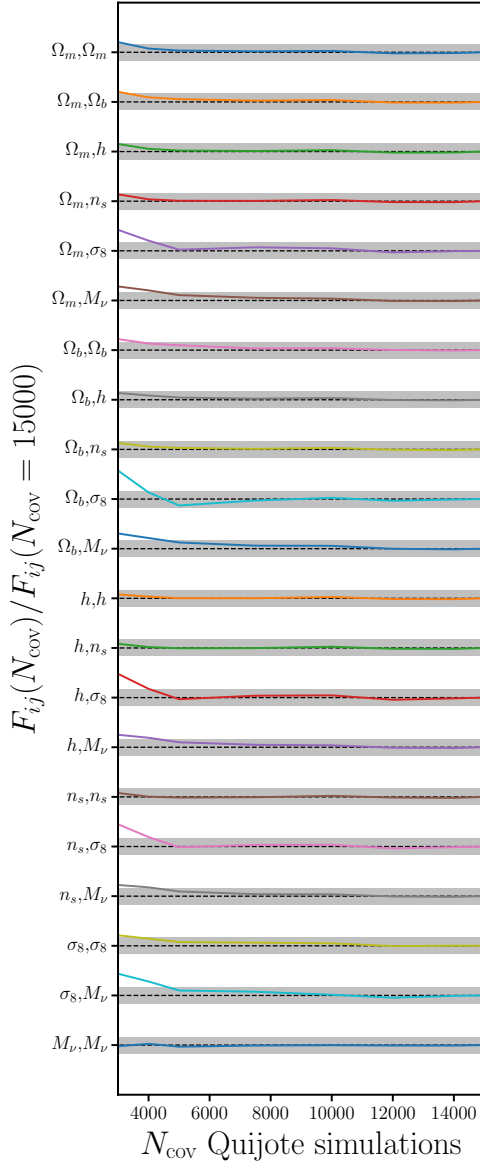
For  $\Omega_m$ ,  $\Omega_b$ ,  $h$ ,  $n_s$ ,  $\sigma_8$ , and  $M_{\text{lim}}$  we estimate  $\partial B_0/\partial\theta_i$  using a centered difference approximation (Eq. 11). However, for  $M_\nu$ , where we cannot have values below 0.0 eV, we cannot estimate the derivative with the same method.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

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

the error goes as  $\mathcal{O}(\delta M_\nu)$ . Instead, we use Eq. 12, which provides a  $\mathcal{O}(\delta M_\nu^2)$  order approximation. In our  $\partial B_0/\partial M_\nu$  approximation, we use the Quijote simulations at  $M_\nu^+$ ,  $M_\nu^{++}$ , and  $M_\nu^{+++}$ . We compare  $\partial \log B(k_1, k_2, k_3)/\partial M_\nu$  (right) and  $\partial \log P(k)/\partial M_\nu$  (left), computed using Eq. 12 (blue) and the forward difference approximation (green) in Figure 11. We also include  $P$  and  $B$  derivatives approximated using only the  $M_\nu^+$  and  $M_\nu^{++}$  simulations in orange. **CH: update numbers:** The three approximations for the derivatives differ from one another by roughly 10% with Eq. 12 producing the largest estimate for both  $P_0$  and  $B_0$ . If we use the  $M_\nu^+$  and  $M_\nu^{++}$  and forward difference derivatives instead of the Eq. 12 for our Fisher forecasts, we find the following marginalized  $M_\nu$  constraints for  $k_{\text{max}} = 0.5 h/\text{Mpc}$ : 0.196 and 0.294eV for  $P_0$  and 0.0308 and 0.0483eV for  $B_0$ . These correspond to a  $\sim 20$  and 80% relative increase from our forecasts in Section 4.2. The forward difference derivatives have a significant impact on our forecasts; however, we emphasize that this is a  $\mathcal{O}(\delta M_\nu)$  approximation, unlike the other  $\mathcal{O}(\delta M_\nu^2)$  approximations. Moreover, because the discrepancies in the derivative propagate similarly to the  $P_0$  and  $B_0$  constraints, the relative improvement of  $B_0$  over  $P_0$  remains roughly the same. Hence we conclude that the derivatives have sufficiently converged 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
- Angulo, R. E., & Pontzen, A. 2016, *Monthly Notices of the Royal Astronomical Society*, 462, L1
- 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
- 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

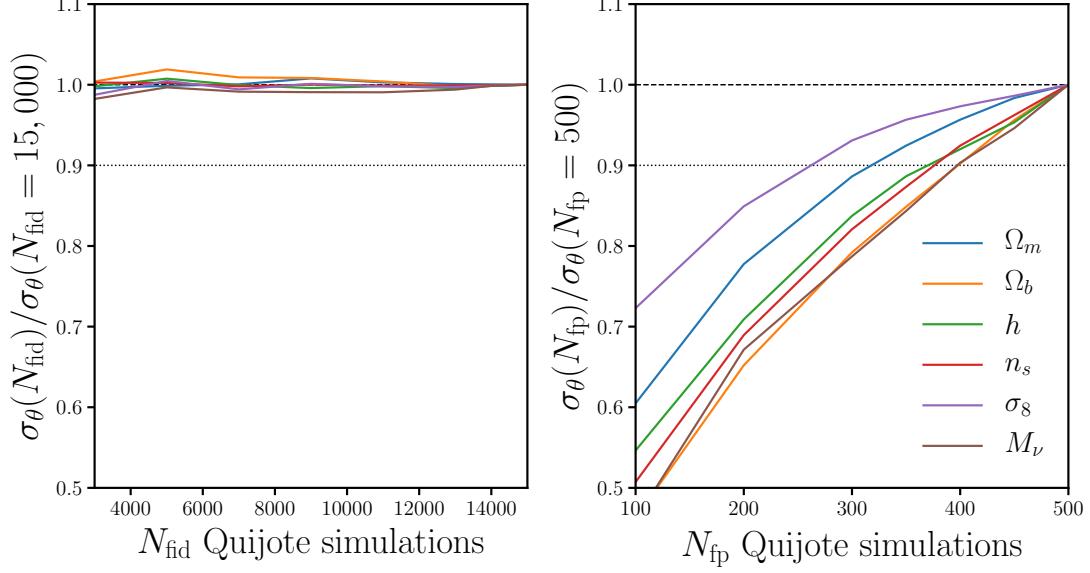
Dodelson, S. 2003, *Modern Cosmology*

Drexlin, G., Hannen, V., Mertens, S., & Weinheimer, C. 2013

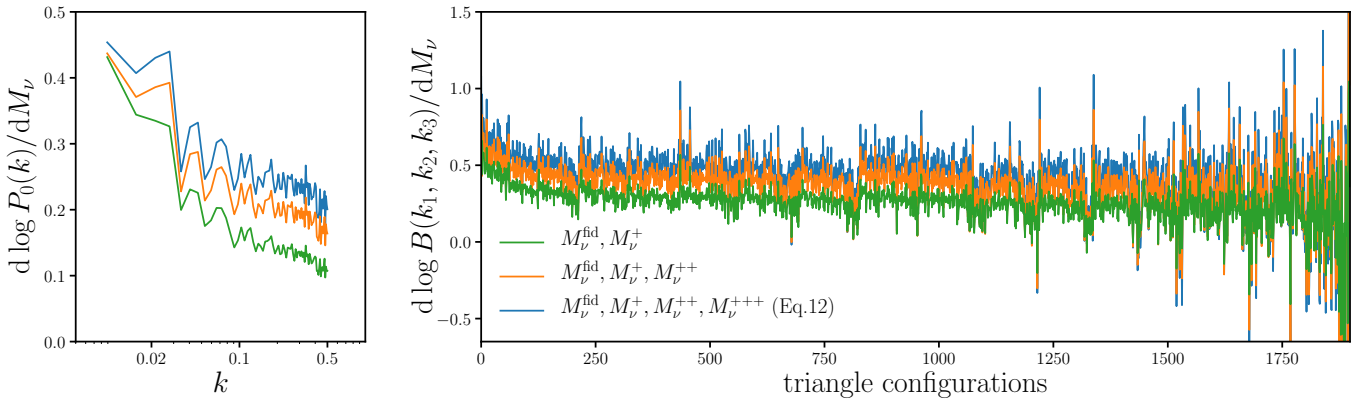
Emberson, J. D., Yu, H.-R., Inman, D., et al. 2017, *Research in Astronomy and Astrophysics*, 17, 085

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



**Figure 10. Left:** The ratio of the  $1\sigma$  Fisher constraint for  $\theta = \Omega_m, \Omega_b, h, n_s, \sigma_8$ , and  $M_\nu$  calculated using  $N_{\text{fid}}$  Quijote simulations over the constraint calculated with all 15,000 simulations,  $\sigma_\theta(N_{\text{fid}})/\sigma_\theta(N_{\text{fid}} = 15,000)$ , as a function of  $N_{\text{fid}}$ . The  $N_{\text{fid}}$  simulations are used to estimate  $\mathbf{C}$  used to calculate the Fisher matrix (Eq. 10). The Fisher parameter constraints vary by  $< 10$  and  $1\%$  for  $N_{\text{fid}} > 7000$  and  $14,000$ , respectively. **Right:** The ratio of the  $1\sigma$  Fisher constraint for  $\theta$  calculated using  $N_{\text{fp}}$  simulations over the constraint calculated with all 500 fixed paired simulations,  $\sigma_\theta(N_{\text{fp}})/\sigma_\theta(N_{\text{fp}} = 500)$ , as a function of  $N_{\text{fp}}$ . The  $N_{\text{fp}}$  fixed paired simulations are used to numerically estimate  $\partial B_0/\partial\theta_i$  in Eq. 10. Although  $\sigma_\theta(N_{\text{fp}})/\sigma_\theta(N_{\text{fp}} = 500)$  vary among the parameters, for  $N_{\text{fp}} > 400$  and  $450$  they vary by  $< 10$  and  $5\%$ , 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 11.** 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 Eq.12 (blue), excluding  $M_\nu^{+++}$  (orange), and the forward difference approximation (green).

- 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
- Leauthaud, A., Tinker, J., Bundy, K., et al. 2012, *The Astrophysical Journal*, 744, 159
- 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
- Pontzen, A., Slosar, A., Roth, N., & Peiris, H. V. 2016, *Physical Review D*, 93, 103519
- 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