Draft version September 12, 2019
Typeset using LaTeX **preprint** style in AASTeX62

Suppressing Cosmic Variance in the Bispectrum using Paired-Fixed Simulations

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

[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*

(Dated: `DRAFT --- 1a0a01d --- 2019-09-11 --- NOT READY FOR DISTRIBUTION`)

ABSTRACT

Keywords: cosmology: —

1. INTRODUCTION

Current LSS analyses rely on analytic models (i.e. perturbation theory) to calculate the likelihood combined with a Bayesian parameter inference framework to derive posterior probability distrubtions of cosmological parameters. While much progress has been made to extend perturbation theory to smaller more nonlinear scales, perturbation theory eventually breaks down. This is a shame because there's significant constraining power on small scales.

N-body simulations don't have the same issues on small scales. There's been a lot of progress in making fast simulations (e.g. fastpm, etc). In principle, we can exploit the constraining power on small scales, beyond perturbation theory limits, by calculating the likelihood directly using simulations. However, even with their progress, simulations are expensive and in standard MCMC analyses require evaluating the likelihood $> 100,000$ times. This makes a naive simulation based approach currently intractable.

Emulation based approaches have been put forth to address this issue. These approaches exploit the accuracy of simulations on small scales while reducing the compute time necessary. In this approach an emulator is built using simulations evaluated at a set of parameters spanning the parameter space to be explored. Most recent works have utilized space-filling sampling methods such as latin hypercubes (citecite) and build their emulators using Gaussian Processes (cite) or chaotic polynomial (cite). Yet even with such approaches the number of simulations is pretty large. Give numbers from Aemulus.

One reason why we still need a lot of sims even for emulators is sample variance. Sample variance is the statistical fluctations in our observables that come from using a finite size simulation. It directly contributes to the theoretical uncertainty of the emulator. Hence multiple realizations are necessary to beat down sample variance and get a less noisy estimate on the mean observable.

* hahn.changhoon@gmail.com

Recently, Pontzen et al. (2016) and Angulo & Pontzen (2016) proposed a method to suppress sample variance using *paired fixed simulations*. define in some concise way paired, fixed, and paired-fixed sims (see below).

$$\delta(\boldsymbol{x}) = \frac{\rho(\boldsymbol{x}) - \bar{\rho}}{\bar{\rho}} \tag{1}$$

$$\delta(\boldsymbol{k}) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3\boldsymbol{x} e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \delta(\boldsymbol{x}) = A e^{i\theta} \tag{2}$$

For Gaussian random field, θ is uniformly sampled from 0 to 2π and A is sampled from Rayleigh distribution:

$$p(A)\mathrm{d}A = \frac{A}{\sigma^2} e^{-A^2/2\sigma^2} \mathrm{d}A \tag{3}$$

where $\sigma^2 = V\ P(k)/(16\pi^3)$. The mean of this distribution is

$$\langle A \rangle = \int_0^\infty \frac{A^2}{\sigma^2} e^{-A^2/2\sigma^2} \mathrm{d}A = \sqrt{\frac{V\,P(k)}{32\pi^2}}. \tag{4}$$

Also,

$$\langle \delta(\boldsymbol{k}) \delta^*(\boldsymbol{k}) \rangle = \langle A^2 \rangle = \int_0^\infty \frac{A^3}{\sigma^2} e^{-A^2/2\sigma^2} \mathrm{d}A = \frac{V\,P(k)}{(2\pi)^3}. \tag{5}$$

A paired Gaussian field is where you have two fields δ_1 and δ_2 where $\delta_2(k) = Ae^{i(\theta+\pi)} = -\delta_1(k)$.

A fixed field is when the amplitude is fixed,

$$A = \sqrt{\frac{V\ P(k)}{(2\pi)^3}}, \tag{6}$$

such that the power spectrum is the same.

Paired fixed is when you do both.

? and Chuang et al. (2019) recently examined whether paired fixed simulations introduce any bias for a variety of observables. list observables. In this work, we focus on whether paired-fixed simulations introduce any bias for the full real-space and redshift-space bispectrum using over 23000 N-body simulations of the Quijote simulation suite. Furthermore, we examine whether these biases can propagate to parameter inference.

2. THE QUIJOTE SIMULATION SUITE

We use a subset of simulations from the Quijote suite, a set of 43,000 N-body simulations that spans over 7000 cosmological models and contains, at a single redshift, over 8.5 trillion particles. The Quijote suite was designed to quantify the information content of cosmological observables and also to train machine learning algorithms. Hence, the suite includes enough realizations to accurately estimate the covariance matrices of high-dimensional observables such as the bispectrum as well as the derivatives of these observables with respect to cosmological parameters. For the derivatives, the suite includes sets of simulations run at different cosmologies where only one parameter is varied from

Table 1. The QUIJOTE suite includes 15,000 standard N-body simulations at the fiducial cosmology to accurately estimate the covariance matrices. It also includes sets of 500 simulations at 13 different cosmologies, where only one parameter is varied from the fiducial value (underlined), to estimate derivatives of observables along the cosmological parameters. At every cosmology, the QUIJOTE suite also includes 250 pairs of paired-fixed simulations.

Name	M_ν	Ω_m	Ω_b	h	n_s	σ_8	ICs	standard realizations	paired-fixed pairs
Fiducial	0.0	0.3175	0.049	0.6711	0.9624	0.834	2LPT	15,000	250
Fiducial ZA	0.0	0.3175	0.049	0.6711	0.9624	0.834	Zel'dovich	500	250
M_ν^+	$\underline{0.1}$ eV	0.3175	0.049	0.6711	0.9624	0.834	Zel'dovich	500	250
M_ν^{++}	$\underline{0.2}$ eV	0.3175	0.049	0.6711	0.9624	0.834	Zel'dovich	500	250
M_ν^{+++}	$\underline{0.4}$ eV	0.3175	0.049	0.6711	0.9624	0.834	Zel'dovich	500	250
Ω_m^+	0.0	$\underline{0.3275}$	0.049	0.6711	0.9624	0.834	2LPT	500	250
Ω_m^-	0.0	$\underline{0.3075}$	0.049	0.6711	0.9624	0.834	2LPT	500	250
Ω_b^+	0.0	0.3175	$\underline{0.051}$	0.6711	0.9624	0.834	2LPT	500	250
Ω_b^-	0.0	0.3175	$\underline{0.047}$	0.6711	0.9624	0.834	2LPT	500	250
h^+	0.0	0.3175	0.049	$\underline{0.6911}$	0.9624	0.834	2LPT	500	250
h^-	0.0	0.3175	0.049	$\underline{0.6511}$	0.9624	0.834	2LPT	500	250
n_s^+	0.0	0.3175	0.049	0.6711	$\underline{0.9824}$	0.834	2LPT	500	250
n_s^-	0.0	0.3175	0.049	0.6711	$\underline{0.9424}$	0.834	2LPT	500	250
σ_8^+	0.0	0.3175	0.049	0.6711	0.9624	$\underline{0.849}$	2LPT	500	250
σ_8^-	0.0	0.3175	0.049	0.6711	0.9624	$\underline{0.819}$	2LPT	500	250

the 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). 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 simulations for $\{M_\nu^+, M_\nu^{++}, M_\nu^{+++}\} = \{0.1, 0.2, 0.4 \text{ eV}\}$. At each of these 14 cosmologies, QUIJOTE includes sets of both standard N-body and paired-fixed simulations.

The initial conditions for all the simulations were generated at $z = 127$ using second-order perturbation theory for simulations with massless neutrinos ($M_\nu = 0.0$ eV) and the Zel'dovich approximation for massive neutrinos ($M_\nu > 0.0$ eV). The initial conditions with massive neutrinos take their scale-dependent growth factors/rates into account using the Zennaro et al. (2017) method, while for the massless neutrino case we use the traditional scale-independent rescaling. For a standard N-body simulation, the amplitude A and phase θ of a Fourier mode $\delta(\boldsymbol{k}) = Ae^{i\theta}$ are drawn from a Rayleigh distribution and uniform distribution between 0 and π, respectively (see Section 1). For a pair of paired-fixed simulations, their amplitudes are fixed to the square root of the variance of the Rayleigh distribution and phases of the pair differ by π.

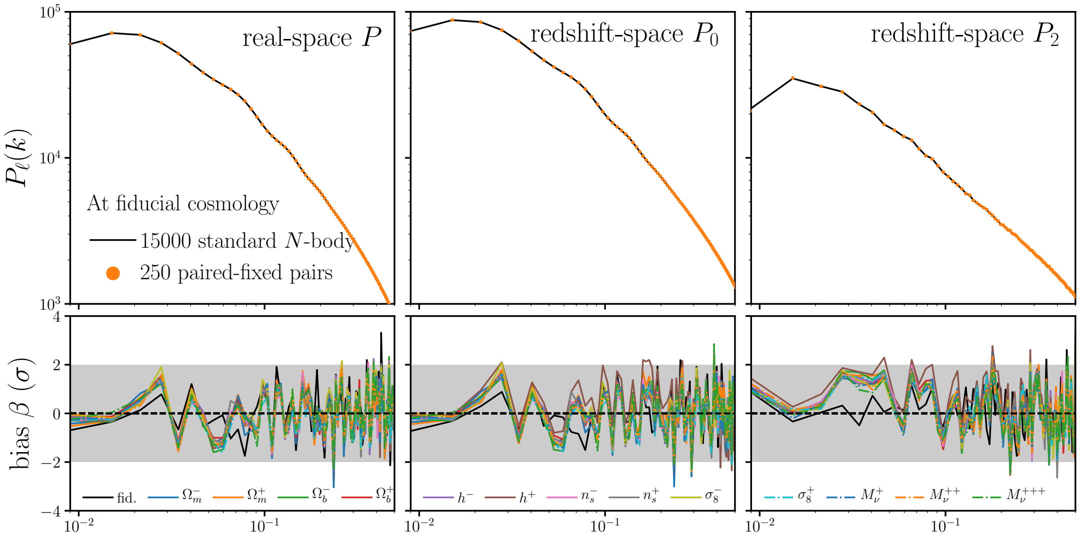

Figure 1. Comparison of the power spectra of standard N-body simulations and paired-fixed simulations. We compare the real-space power spectrum, redshift-space power spectrum monopole, and quadrupole in the left, center, and right columns, respectively. In the top panels, we compare the average power spectra of the standard simulations (black line) to the paired-fixed simulations (orange scatter) at the fiducial cosmology. In the bottom panels, we compare the bias, β of the paired-fixed simulations for the power spectra for all 14 cosmologies (Table 1). We mark bias within $\pm 2\sigma$ within the shaded region. At the fiducial cosmology we use 15,000 standard N-body and 250 pairs of paired-fixed simulations. For the other cosmologies, we use 500 standard N-body and 250 pairs of paired-fixed simulations. Consistent with previous results, *we find no significant bias in the real and redshift-space power spectra of paired-fixed simulations.*

From the initial conditions, the simulations follow the gravitational evolution of 512^3 dark matter particles, and 512^3 neutrino particles for massive neutrino models, to $z = 0$ using Gadget-III TreePM+SPH code (Springel 2005). Simulations with massive neutrinos are run using the "particle method", where neutrinos are described as a collisionless and pressureless fluid and therefore modeled as particles, same as CDM (Brandbyge et al. 2008; Viel et al. 2010). Halos are then identified using the Friends-of-Friends algorithm (FoF; Davis et al. 1985) with linking length $b = 0.2$ on the CDM + baryon distribution. We limit the halo catalogs to halos with masses above $M_{\rm lim} = 3.2 \times 10^{13} h^{-1} M_\odot$. For the fiducial cosmology, the halo catalogs have $\sim$156,000 halos ($\bar{n} \sim 1.56 \times 10^{-4}\ h^3{\rm Gpc}^{-3}$) with $\bar{n}P_0(k = 0.1) \sim 3.23$. We refer readers to Villaescusa-Navarro et al. (in preparation) and Hahn et al. (2019) for further details on the Quijote simulations.

3. RESULTS

We're interesting in determining whether paired-fixed simulations can be used, instead of standard N-body simulations, to reduce the statistical uncertainty from cosmic variance in applications where we want to accurately estimates of the mean of observables, such as in emulator-based approach. More specifically, we examine whether using paired-fixed simulations introduce bias for the full real and redshift-space bispectrum with the unprecedented statistical precision from over 22,000 Quijote simulatons. Furthermore, we exmaine whether such biases can propagate to constraints on

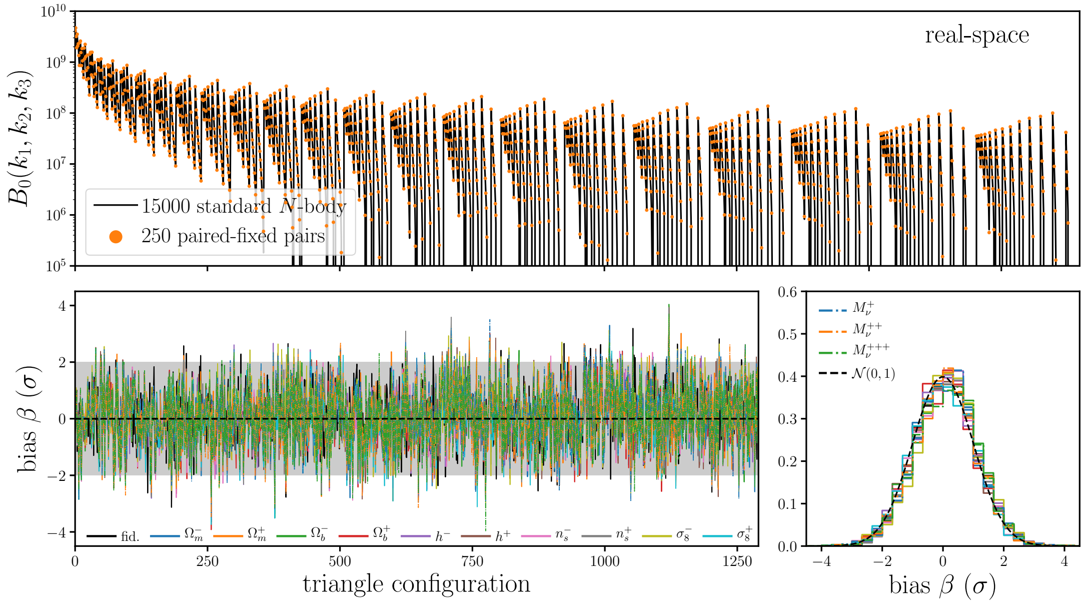

Figure 2. Comparison of the real-space bispectrum of standard N-body simulations and paired-fixed simulations. In the top panel, we compare the average bispectrum of the standard simulations (black line) to the paired-fixed simulations (orange scatter) at the fiducial cosmology. We include all 1898 triangles out to $k_{\rm max} = 0.5$. In the bottom panels, we compare the bias, β of the paired-fixed simulations for the bispectrum for all 14 cosmologies, as a function of triangle configuration (left) and their distribution (right). The β distributions are in good agreement with a Gaussian distribution, $\mathcal{N}(0, 1)$ (blackdashed). *Hence, we find no significant bias in the real-space bispectrum of paired-fixed simulations.*

inferred cosmological parameters. We also we reexamine whether paired-fixed simulations introduce bias for the power spectrum.

We quantify the bias of paired-fixed simulations for some observable X as

$$\beta_X = \frac{\overline{X}_{\rm std} - \overline{X}_{\rm pf}}{\sqrt{\sigma^2_{\rm std} + \sigma^2_{\rm pf}}} \tag{7}$$

where the mean observable of the standard N-body and paired-fixed simulations: $\overline{X}_{\rm std} = \frac{1}{N}\sum\limits_{i=1}^{N} X_{{\rm std},i}$ and $\overline{X}_{\rm pf} = \frac{1}{N}\sum\limits_{i=1}^{N} X_{{\rm pf},i}$. $X_{{\rm pf},i}$ is the mean observable of a pair of paired-fixed simulations:

$$X_{{\rm pf},i} = \frac{1}{2}\big[X^{(1)}_{{\rm pf},i} + X^{(2)}_{{\rm pf},i}\big]. \tag{8}$$

ACKNOWLEDGEMENTS

It's a pleasure to thank Arka Banerjee, Chia-Hsun Chuang, Joseph DeRose, Thomas McClintock, Jeremy L. Tinker ... for valuable discussions and comments.

APPENDIX

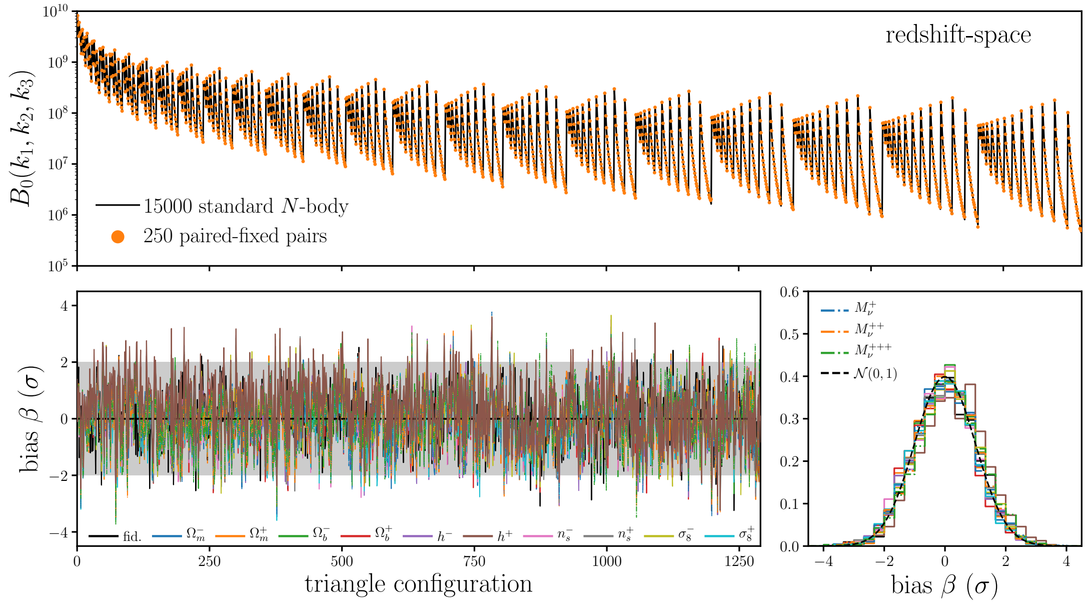

Figure 3. Same as Figure 2 but for the redshift-space bispectrum. The β distributions of the different cosmologies are mostly in good agreement with $\mathcal{N}(0,1)$ (lower right panel). A few cosmologies, h^+, M_ν^{+++}, have noticable discrepancies; however, these discrepancies are within 1σ. *Hence, we also find no significant bias in the redshift-space bispectrum of paired-fixed simulations.*

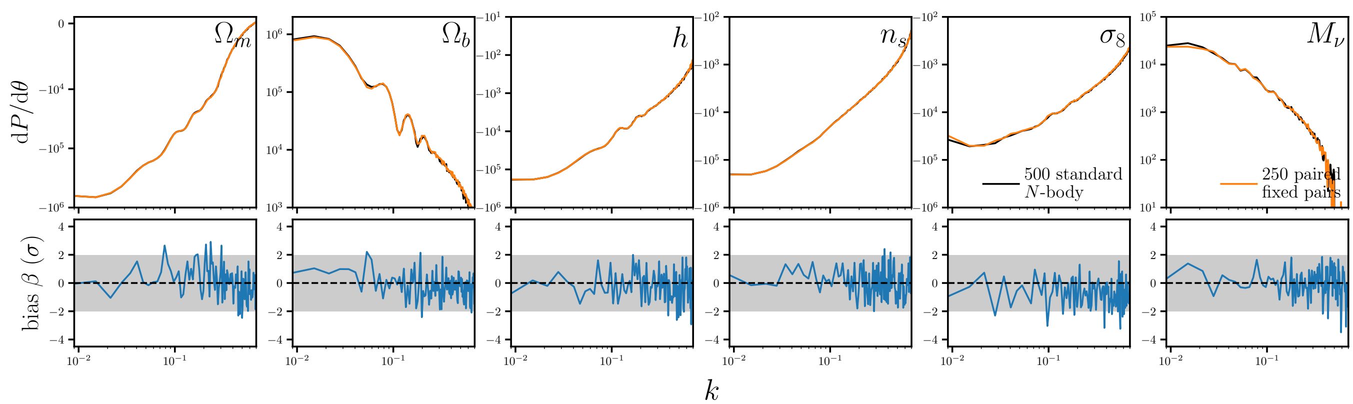

Figure 4. Comparison of the real-space power spectrum derivatives with respect to $\Omega_{\rm m}$, $\Omega_{\rm b}$, h, n_s, σ_8, and M_ν derived from standard N-body simulations versus paired-fixed simulations. In the top panels, we compare the derivatives from the standard simulations (black) to the paired-fixed simulations (orange). In the bottom panels, we compare the bias, β of each power spectrum derivative. The derivatives with respect to $\Omega_{\rm m}$ and σ_8 have noticably non-zero biases; however these deviations are within 1σ. Therefore, *we find no significant bias in the real-space power spectrum derivatives calculated from paired-fixed simulations.*

REFERENCES

Angulo, R. E., & Pontzen, A. 2016, Monthly Notices of the Royal Astronomical Society, 462, L1

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

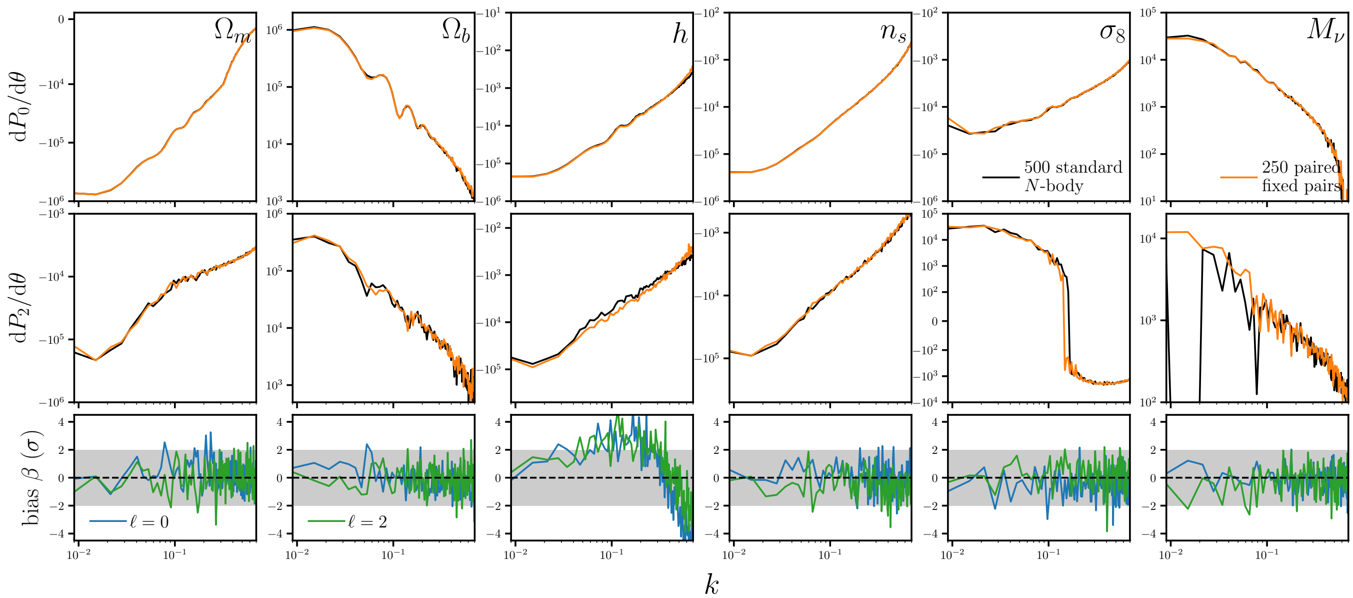

Figure 5. Same as Figure 4 but for the derivatives of the power spectrum monopole (top panels) and quadrupole (center panels) with respect to Ω_m, Ω_b, h, n_s, σ_8, and M_ν. The bottom panel reveals significantly larger biases for $\mathrm{d}P_0/\mathrm{d}\theta$ and $\mathrm{d}P_2/\mathrm{d}\theta$ than the real-space $\mathrm{d}P/\mathrm{d}\theta$, especially for h, σ_8 and M_ν. In fact, the bias for both $\mathrm{d}P_0/\mathrm{d}h$ and $\mathrm{d}P_2/\mathrm{d}h$ exceed 2σ in the range $k > 0.04\ h/\mathrm{Mpc}$. **CH:** some nuanced discussion about how this still is not a cause for concern because this is the bias for a 500 Gpc volume so it's well within the tolerance.

Chuang, C.-H., Yepes, G., Kitaura, F.-S., et al. 2019, Monthly Notices of the Royal Astronomical Society, stz1233

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, The Astrophysical Journal, 292, 371

Pontzen, A., Slosar, A., Roth, N., & Peiris, H. V. 2016, Physical Review D, 93, 103519

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

Viel, M., Haehnelt, M. G., & Springel, V. 2010, Journal of Cosmology and Astro-Particle Physics, 06, 015

Zennaro, M., Bel, J., Villaescusa-Navarro, F., et al. 2017, Monthly Notices of the Royal Astronomical Society, 466, 3244

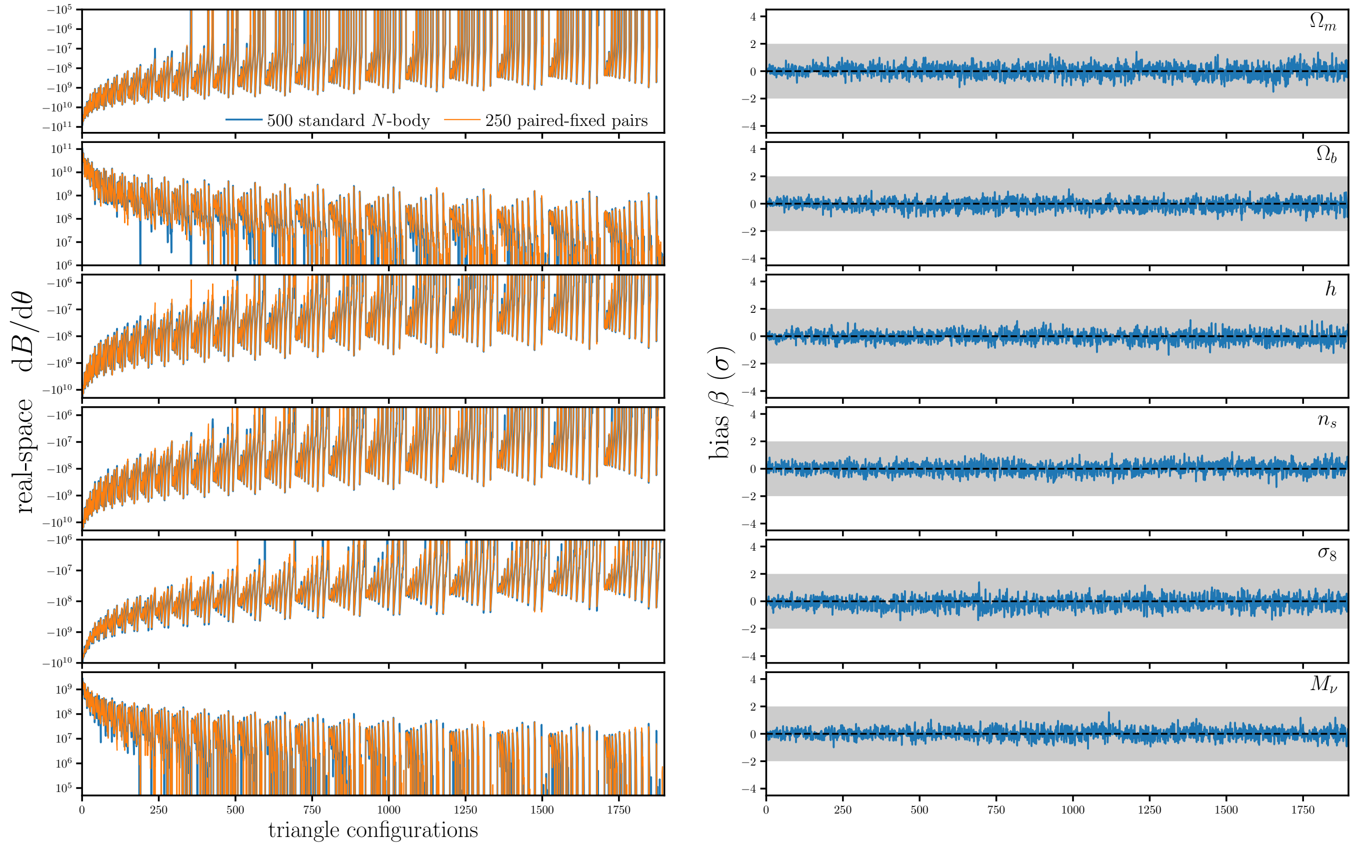

Figure 6. Comparison of the real-space bispectrum derivatives with respect to Ω_m, Ω_b, h, n_s, σ_8, and M_ν derived from standard N-body simulations versus paired-fixed simulations. In the left panels, we compare the derivatives from the standard simulations (black) to the paired-fixed simulations (orange). In the right panels, we compare the bias, β of the derivatives for each parameter. *We find no significant bias in the real-space bispectrum derivatives calculated from paired-fixed simulations.*

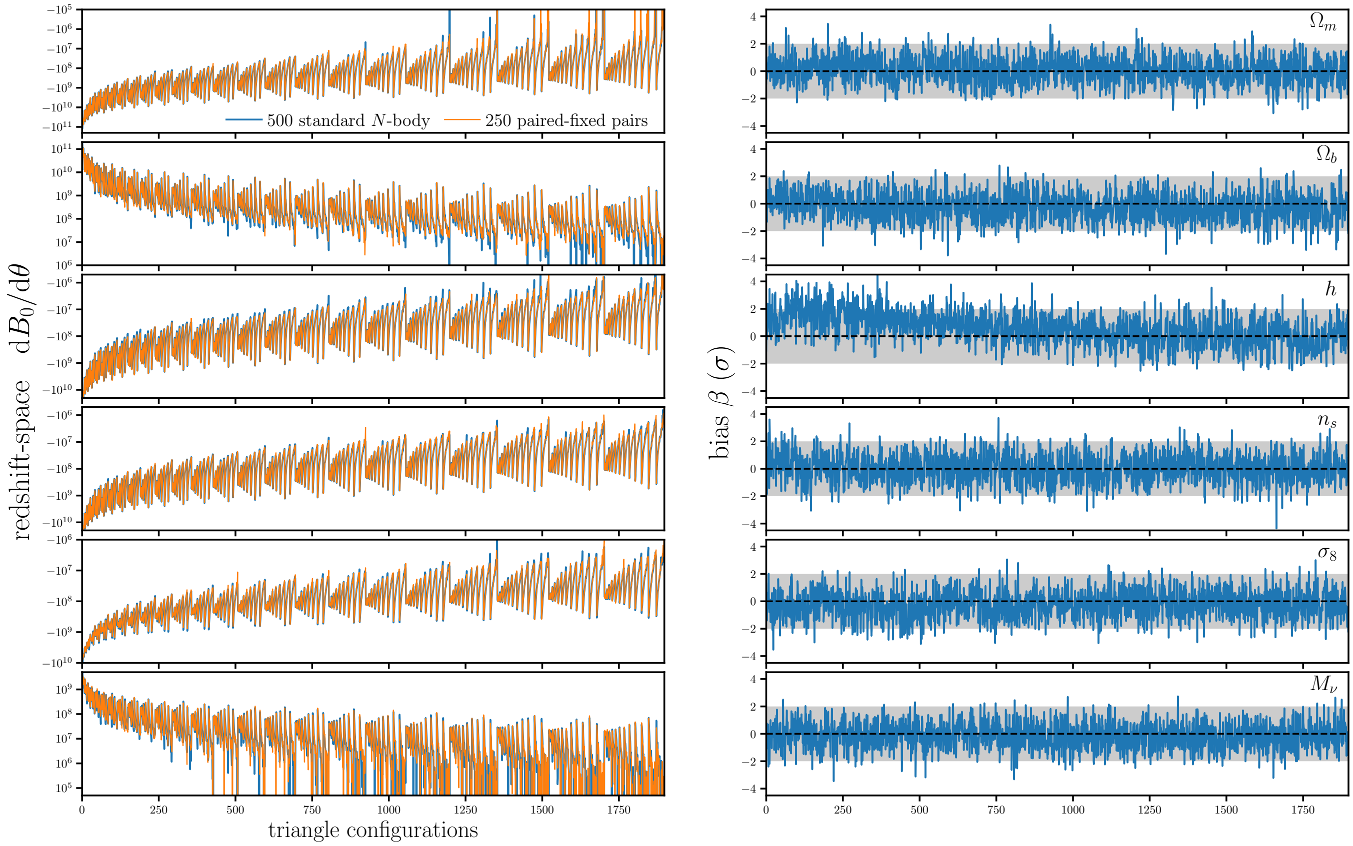

Figure 7. Same as Figure 6 but for the derivatives of the redshift-space bispectrum monopole with respect to the cosmological parameters. Similar to the power spectrum derivatives, we find significantly larger biases for redshift-space derivatives than in real-space. The bias for $\mathrm{d}B_0/\mathrm{d}h$ exceeds 2σ for triangle configurations with $0.1 \lesssim k_1, k_2, k_3 \lesssim 0.3\ h/\mathrm{Mpc}$. **CH:** again some nuanced discussion about how this still is not a cause for concern because this is the bias for a 500 Gpc volume so it's well within the tolerance.

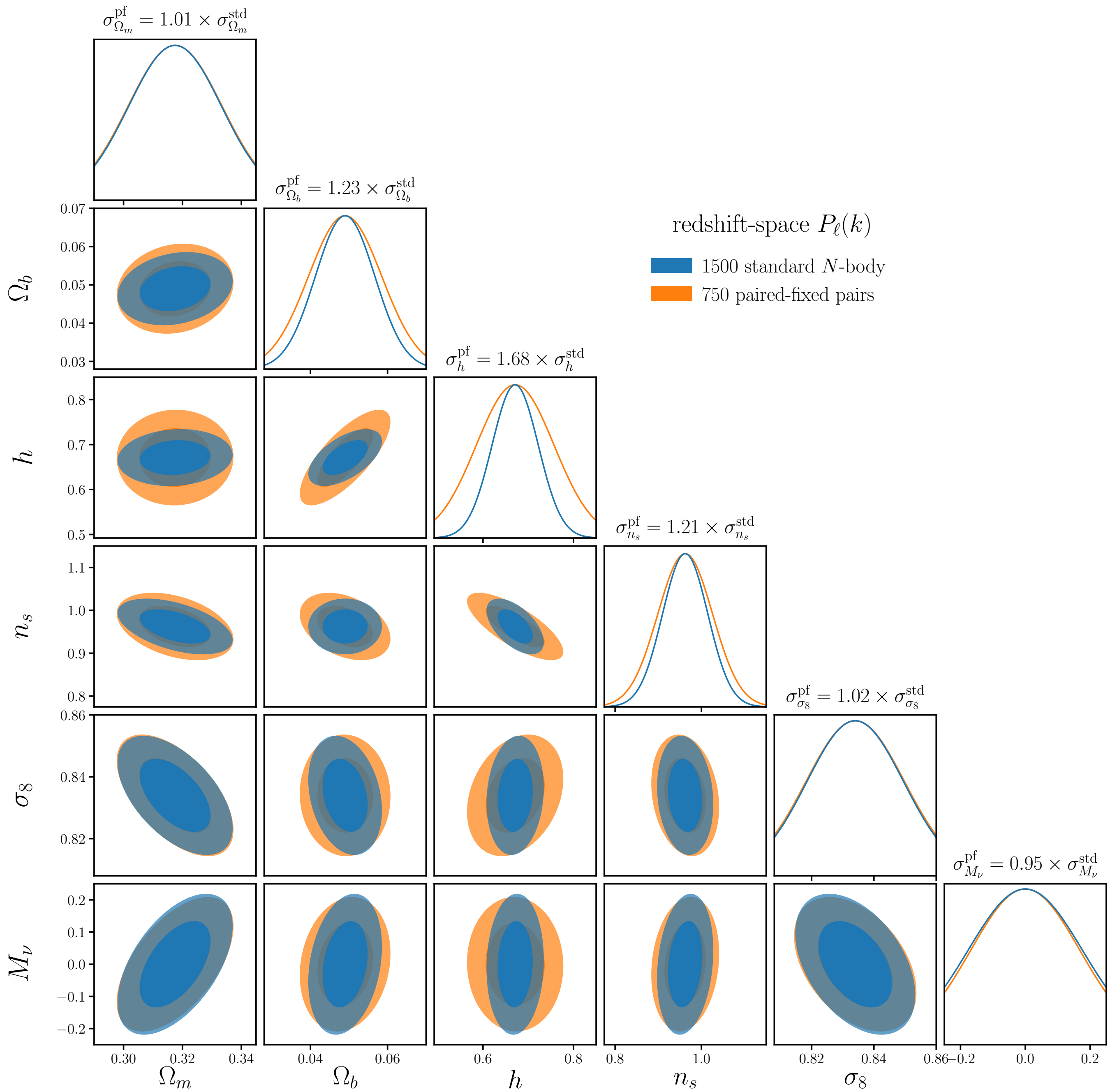

Figure 8. **CH:** to be updated

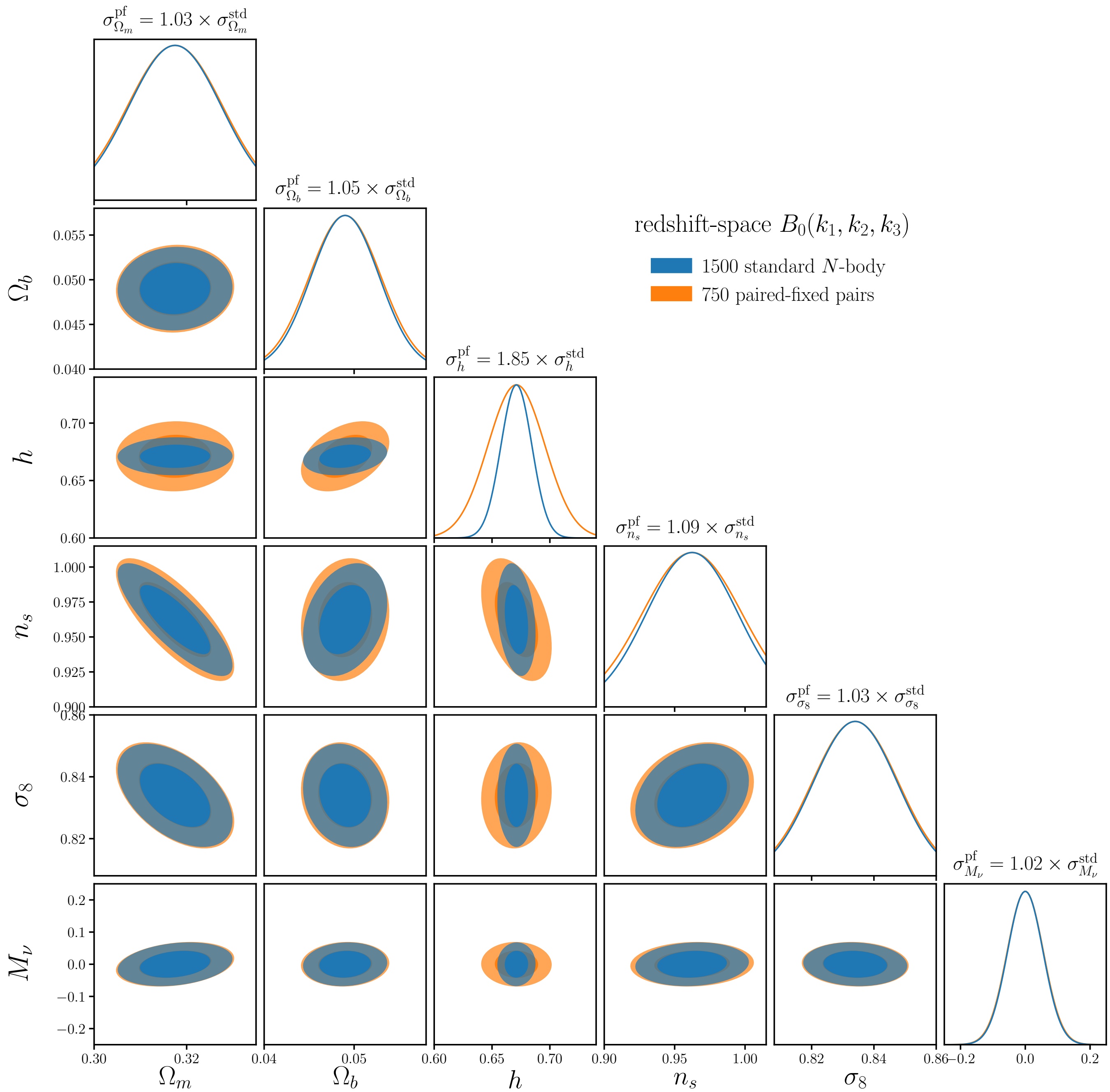

Figure 9. CH: to be updated