The Astrophysical Journal, 867:137 (24pp), 2018 November 10 https://doi.org/10.3847/1538-4357/aae52b

Statistical Properties of Paired Fixed Fields

Francisco Villaescusa-Navarro[1], Sigurd Naess[1], Shy Genel[1,2], Andrew Pontzen[3], Benjamin Wandelt[1,4,5], Lauren Anderson[1], Andreu Font-Ribera[3], Nicholas Battaglia[1,6,7], and David N. Spergel[1,7]
[1] Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA; fvillaescusa@flatironinstitute.org
[2] Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, NY 10027, USA
[3] Department of Physics and Astronomy, University College London, 132 Hampstead Road, London, NW1 2PS, UK
[4] Institut d'Astrophysique de Paris 98bis, Boulevard Arago, F-75014 Paris, France
[5] Sorbonne Universites, Institut Lagrange de Paris 98 bis Boulevard Arago, F-75014 Paris, France
[6] Department of Astronomy, Cornell University, Ithaca, NY 14853, USA
[7] Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544-0010, USA
Received 2018 June 8; revised 2018 August 27; accepted 2018 August 29; published 2018 November 9

Abstract

The initial conditions of cosmological simulations are commonly drawn from a Gaussian ensemble. The limited number of modes inside simulations gives rise to *sample variance*: statistical fluctuations that limit the accuracy of the simulation predictions. Fixed fields offer an alternative initialization strategy; they have the same power spectrum as Gaussian fields but no intrinsic amplitude scatter. Paired fixed fields consist of two fixed fields with opposite phases that cancel phase correlations. We study the statistical properties of those fields for 19 different quantities at different redshifts through a large set of 600 *N*-body and 530 state-of-the-art magnetohydrodynamic simulations. We find that paired fixed simulations do not introduce a bias on any of the examined quantities. We quantify the statistical improvement brought by these simulations on different power spectra—matter, halos, cold dark matter, gas, stars, galaxies, and magnetic fields—finding that they can reduce their variance by factors as large as 10^6. We quantify the improvement achieved by fixing and by pairing, showing that sample variance can be highly suppressed by pairing after fixing. Paired fixed simulations do not change the scatter in quantities such as the probability distribution function or the halo, void, or stellar mass functions. We argue that procedures aiming at reducing the sample variance of those quantities are unlikely to work. Our results show that paired fixed simulations do not affect either mean relations or scatter of galaxy properties and suggest that the information embedded in one-point statistics is highly complementary to that in clustering.

Key words: large-scale structure of universe – methods: numerical – methods: statistical

1. Introduction

The standard model of cosmology is a well-established theoretical framework that explains with great success a large and diverse range of cosmological observables. The parameters of the model represent fundamental physics quantities such as the nature of dark energy, the density of dark matter, or the sum of the neutrino masses. The goal of current and upcoming cosmological surveys is to determine the value of those parameters with the highest accuracy possible, in order to improve our knowledge of fundamental physics.

The amount of information that can be extracted from cosmological surveys depends on the accuracy of the theoretical model. For instance, theoretical predictions are very accurate and fast to compute in the linear regime, but the amount of information that can be extracted with them is limited, since that regime can only accurately describe the largest scales. Perturbation theory (Bernardeau et al. 2002) is an ideal tool to make accurate theoretical predictions in the mildly nonlinear regime. However, theoretical predictions in the fully nonlinear regime require numerical simulations.

Ideally, the best way to extract cosmological information would be by evaluating the likelihood in every point of the parameter space by using the theoretical prediction from cosmological simulations. This procedure has been impractical so far (see, however, Palanque-Delabrouille et al. 2015, for similar efforts with the Lyα forest) as a result of several factors: (1) the volume of the parameter space can be very large, requiring many simulations for sampling it; (2) a very large number of simulations are needed to compute the covariance matrix in each point of the parameter space; (3) simulations covering representative survey volumes with the required mass resolution are computationally expensive; and (4) each simulation has an intrinsic variance, commonly called *sample variance*, arising from the limited number of modes it contains, such that many simulations are needed to compute the mean.

The first point can be addressed by running simulations on a subset of strategic locations in the parameter space (see, e.g., Heitmann et al. 2009). For the second and third points, a large amount of work has been carried out to speed up the running time of *N*-body simulations and to evaluate the covariance matrix, at the expense of accuracy (Monaco et al. 2002a, 2002b, 2013; Scoccimarro & Sheth 2002; Taffoni et al. 2002; Kitaura & Heß, 2013; Tassev et al. 2013, 2015; Kitaura et al. 2014; Chuang et al. 2015a, 2015b; Howlett et al. 2015; Feng et al. 2016a; Rizzo et al. 2017). We emphasize, however, that those methods, while aiming at capturing the effect of nonlinear matter evolution, do not include the nonlinear effects of baryons.

The scope of this paper is to investigate how to mitigate the fourth point, i.e., the intrinsic sample variance attached to each simulation. We focus our attention on paired fixed fields, introduced in Pontzen et al. (2016) and Angulo & Pontzen (2016). Those fields can be obtained from Gaussian density fields by performing certain operations on the amplitudes and/or the phases of their modes. Angulo & Pontzen (2016) showed that numerical simulations run with those fields as initial conditions lead to quantities, such as the matter power spectrum, with a much

lower variance than those obtained from traditional Gaussian fields.

The purpose of this work is to further investigate the properties of paired fixed fields and (1) identify the quantities for which paired fixed fields help in reducing the intrinsic statistical scatter, (2) quantify the statistical improvement, and (3) study whether a bias is introduced in any quantities.

We carry out our study using a large set of 600 *N*-body simulations with different box sizes and mass and spatial resolutions run specifically for this work using the GADGET-III code (Springel 2005). We use them to study the impact of paired fixed simulations on the matter, halo, and halo–matter power spectra; the halo bias; the probability distribution function (pdf) of matter density; the halo mass function; and the void radius function.

We then study the statistical properties of paired fixed simulations using a set of 530 state-of-the-art magnetohydrodynamic simulations also run specifically for this work through the AREPO code (Springel 2010). We investigate the properties of the above quantities, along with the power spectra of the other components: gas, cold dark matter (CDM), stars, galaxies, and magnetic fields. We also study the impact of paired fixed fields on the star formation rate history, on the stellar mass function, and on several internal galaxy properties such as radii or maximum circular velocity.

This paper is organized as follows. In Section 2 we define Gaussian, paired Gaussian, fixed, and paired fixed fields. The set of numerical simulations run for this project is described in Section 3, where we also explain the tools we use to carry out the statistical analysis. We present the results from our *N*-body and hydrodynamic simulations in Sections 4–6 for large, intermediate, and small scales, respectively. In Section 7 we investigate whether we can generate fields with reduced sample variance in both their the one-point and two-point statistics. Finally, we draw the main conclusions of this paper in Section 8.

2. Definitions

We now define what Gaussian, paired Gaussian, fixed, and paired fixed fields are. For a given density field $\rho(\boldsymbol{x})$, the density contrast is defined as

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

where $\bar{\rho} = \langle \rho(\boldsymbol{x}) \rangle$. We express its value in Fourier space as

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

where A is the mode's amplitude and θ is its phase. We note that the values of both A and θ depend on the particular wavelength, $\boldsymbol{k}$, considered. Since the density field is real, the modes satisfy $\delta(-\boldsymbol{k}) = \delta^*(\boldsymbol{k})$. The power spectrum of the field is defined as

$$\langle \delta(\boldsymbol{k}_1) \delta^*(\boldsymbol{k}_2) \rangle = \delta^D(\boldsymbol{k}_1 - \boldsymbol{k}_2) P(k_1), \tag{3}$$

and for a simulation of box size L and volume $V = L^3$ the above equation reads

$$P(k_1) = \frac{(2\pi)^3}{V} \delta_{\boldsymbol{k}_1,\boldsymbol{k}_2} \langle \delta(\boldsymbol{k}_1) \delta^*(\boldsymbol{k}_2) \rangle, \tag{4}$$

where $\delta_{\boldsymbol{k}_1,\boldsymbol{k}_2}$ is a Kronecker delta of the vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. In a Gaussian density field θ is a random variable distributed uniformly between 0 and 2π, whereas A follows a Rayleigh distribution

$$p(A)dA = \frac{A}{\sigma^2} e^{-A^2/2\sigma^2} dA, \tag{5}$$

with $\sigma^2 = VP(k)/(16\pi^3)$. The mean value of the mode amplitude is

$$\langle A \rangle = \int_0^\infty \frac{A^2}{\sigma^2} e^{-A^2/2\sigma^2} dA = \sqrt{\frac{VP(k)}{32\pi^2}}. \tag{6}$$

A density field built as above will satisfy

$$\langle \delta(\boldsymbol{k}) \delta^*(\boldsymbol{k}) \rangle = \langle A^2 \rangle = \int_0^\infty \frac{A^3}{\sigma^2} e^{-A^2/2\sigma^2} dA = \frac{VP(k)}{(2\pi)^3}. \tag{7}$$

A Gaussian field is completely described by its two-point correlation function, or power spectrum.

It is interesting to consider a different distribution for the amplitudes of the modes that fulfills two conditions: (1) the amplitude of the power spectrum is the same as in Gaussian fields, i.e., $\langle \delta(\boldsymbol{k}) \delta^*(\boldsymbol{k}) \rangle = VP(k)/(2\pi)^3$, and (2) it has no intrinsic scatter. The following distribution satisfies these two conditions:

$$p(A)dA = \delta^D\left(A - \sqrt{\frac{VP(k)}{(2\pi)^3}}\right) dA. \tag{8}$$

We note that in such fields the value we assign to each mode with wavenumber $\boldsymbol{k}$ is not the mean of the Rayleigh distribution (see Equation (6)). We emphasize that fields constructed with amplitudes drawn from the above distribution are not Gaussian.

We define Gaussian, paired Gaussian, fixed, and paired fixed fields as follows:

1. *Gaussian field*: a field with $\delta(\boldsymbol{k}) = A e^{i\theta}$, where A follows the Rayleigh distribution of Equation (5).
2. *Paired Gaussian field*: a pair of Gaussian fields, $\delta_1(\boldsymbol{k}) = A e^{i\theta}$ and $\delta_2(\boldsymbol{k}) = A e^{i(\theta+\pi)} = -\delta_1(\boldsymbol{k})$, where the values of A and θ are the same for the two fields and A follows the Rayleigh distribution of Equation (5).
3. *Fixed field*: a field with $\delta(\boldsymbol{k}) = A e^{i\theta}$, where A follows the distribution of Equation (8).
4. *Paired fixed field*: a pair of fields, $\delta_1(\boldsymbol{k}) = A e^{i\theta}$ and $\delta_2(\boldsymbol{k}) = A e^{i(\theta+\pi)} = -\delta_1(\boldsymbol{k})$, where the values of A and θ are the same for the two fields and A follows the distribution of Equation (8).

In all the above fields θ is a random variable distributed uniformly between 0 and 2π. Any of the above fields satisfies the Hermitian condition: $\delta(-\boldsymbol{k}) = \delta^*(\boldsymbol{k})$. In Figure 1 we show the 2D power spectrum, i.e., the amplitude of the power spectrum as a function of $k_\perp$ and $k_\parallel$ from a Gaussian and fixed field, and its comparison with the input power spectrum. Notice that our 2D power spectrum is in real space, so the symmetry between $k_\perp$ and $k_\parallel$ is not broken by redshift-space distortions (which are not included in this toy example). We also show schematically the effects of nonlinear evolution.

Paired Gaussian fields were introduced in Pontzen et al. (2016). Fixed fields have been relatively well known (see, e.g., Viel et al. 2010). Paired fixed fields were first studied in Angulo & Pontzen (2016).

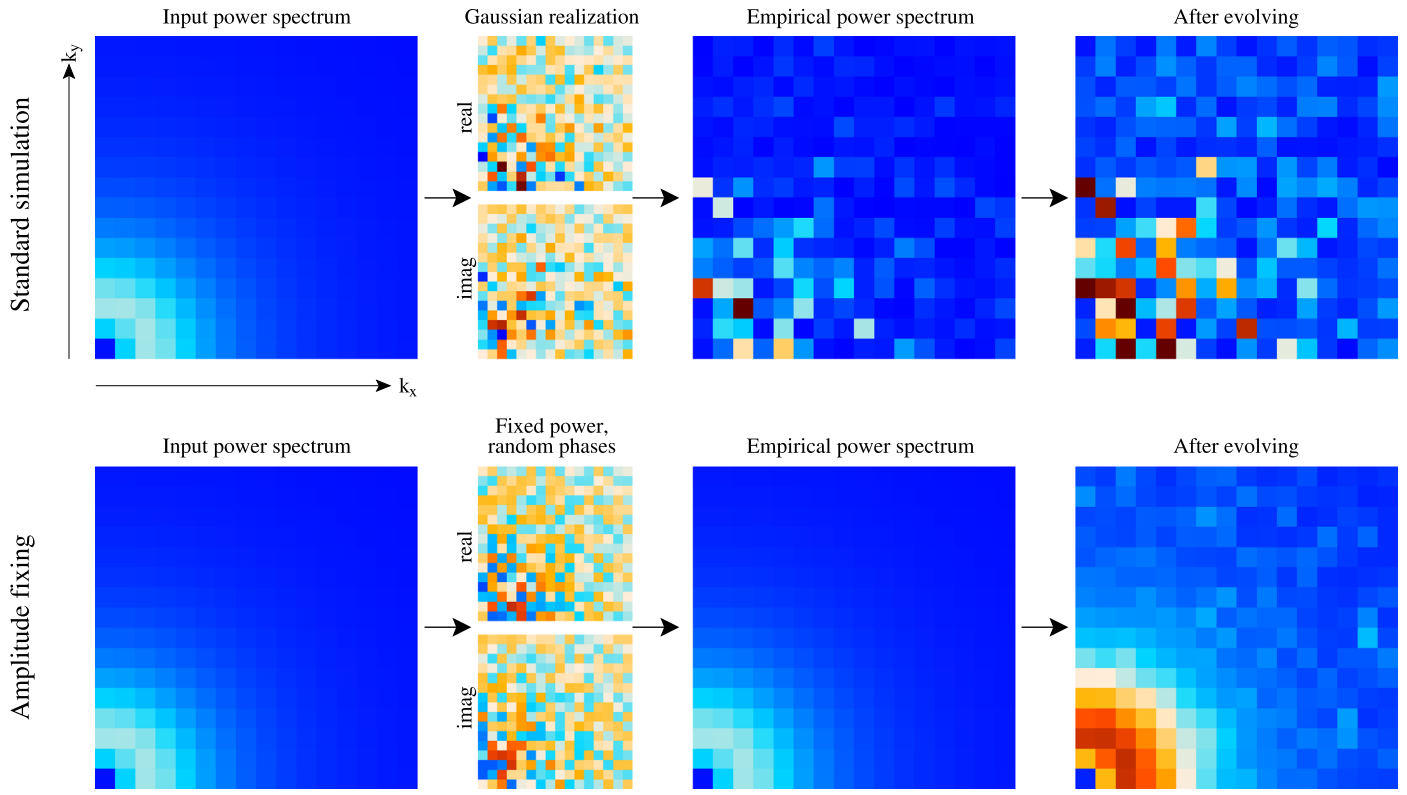

Figure 1. Toy example illustrating sample variance in standard (top row) and fixed (bottom row) simulations. In both cases, a random density field (second column, here shown as the real and imaginary Fourier components) is drawn from an input power spectrum (left column), but because the fixed simulations fix the complex amplitude, they avoid introducing sample variance into the empirical power spectrum measured from the random realization (third column). The rightmost panels show a sketch of the effect of further evolution: linear scales simply grow proportionally, preserving whatever sample variance was present, while on smaller scales, nonlinearities and mode mixing reintroduce some of the suppressed cosmic variance in the fixed simulation (see Figure 4).

In this paper we have run simulations where the initial conditions have been generated using the above fields. We refer to these simulations as standard, paired, fixed, and paired fixed simulations.

We note that although Gaussian and fixed fields share, by construction, the same power spectrum, they differ in higher-order correlations like the trispectrum. For instance, the variance of the power spectrum

$$\sigma^2(P(k)) = \langle (P(k) - \bar{P}(k))^2 \rangle = \frac{(2\pi)^6}{V^2} \langle \delta(\boldsymbol{k})\delta^*(\boldsymbol{k})\delta(\boldsymbol{k})\delta^*(\boldsymbol{k}) \rangle - \bar{P}^2(k) \quad (9)$$

is equal to $P(k)$ for Gaussian fields but is identically zero for fixed fields. For this reason, we expect that the scatter in the matter power spectrum of fixed (and paired fixed) fields will be lower than in Gaussian fields. One of the purposes of this paper is to study the reduction on the scatter of a considered quantity achieved by fixed and paired fixed simulations.

On the other hand, the value of some quantities, e.g., the pdf of the density field, will depend on the value of the n-point correlation function. Since the value of these functions may be different in Gaussian and fixed fields, we expect that fixed and paired fixed simulations may introduce a bias on the value of those quantities. Angulo & Pontzen (2016) argued using perturbation theory that observable quantities should be unbiased, but the accuracy of that statement in the nonlinear evolution remains to be tested. Thus, the other key point of this work is to quantify the magnitude of that bias.

3. Methods

In this section we describe the numerical simulations run for this work. We also explain the statistical analysis we carry out to quantify (1) whether paired fixed simulations introduce a bias on the considered quantity and (2) the statistical improvement achieved over standard simulations.

3.1. Numerical Simulations

A large number of realizations are needed to study the statistical properties of paired fixed simulations. Thus, in this work we have run an unusually large number, ∼1000, of standard and paired fixed simulations.

The purpose of this paper is to investigate the properties of those fields across a large range of scales, from linear to fully nonlinear scales. Doing so with a single set of simulations would require the simulations to have a large box size and a large number of particles. Having a sensible number of those simulations will be computationally expensive. Thus, we decided to run three different sets of simulations that encompass three different ranges of scales: (1) N-body simulations with box sizes of $1000\,h^{-1}$ Mpc at low mass resolution, (2) hydrodynamic simulations with boxes of $200\,h^{-1}$ Mpc at intermediate mass resolution, and (3) both N-body and hydrodynamic simulations with boxes of $20\,h^{-1}$ Mpc at high mass resolution.

All our simulations share the same value of the cosmological parameters, $\Omega_{\rm m} = 0.3175$, $\Omega_{\rm b} = 0.049$, $\Omega_\nu = 0$, $n_s = 0.96$, $h = 0.67$, $\sigma_8 = 0.834$, which are in agreement with the results by *Planck* (Planck Collaboration et al. 2016c). We have

Table 1
Specifications of the Simulations Run for This Paper

Name	Type	Code	No. Standard Realizations	No. Paired Fixed Realizations	$N_{\rm CDM}^{1/3}$	$N_{\rm gas}^{1/3}$	$m_{\rm CDM}$ ($h^{-1}\,M_\odot$)	$m_{\rm gas}$ ($h^{-1}M_\odot$)	ϵ (h^{-1} kpc)	Box Size (h^{-1} Mpc)
N1000	*N*-body	GADGET3	100	100	512	⋯	6.6×10^{11}	⋯	50	1000
N20	*N*-body	GADGET3	100	100	256	⋯	4.2×10^{7}	⋯	2	20
H200	Hydrodynamic	AREPO	30	25	640	640	2.3×10^{9}	4.2×10^{8}	8	200
H20	Hydrodynamic	AREPO	250	100	256	256	3.6×10^{7}	6.5×10^{6}	2	20

Note. The value of the cosmological parameters is the same for all simulations: $\Omega_{\rm m} = 0.3175$, $\Omega_{\rm b} = 0.049$, $\Omega_\nu = 0$, $n_s = 0.96$, $h = 0.67$, $\sigma_8 = 0.834$. We have four different sets of simulations, with different box sizes and numbers of particles. The first letter of the name set represents whether it is an *N*-body (N) or hydrodynamic (H) simulation, while the number thereafter is the box size in h^{-1} Mpc. Note that one paired fixed realization corresponds to two simulations with random phases flipped by π. The physics included in our magnetohydrodynamic simulations is radiative cooling, star formation, metal enrichment, galactic winds, black hole accretion, and feedback. The H200 simulations are similar to the illustrisTNG300-3 simulation.

generated the initial conditions by displacing and assigning peculiar velocities to particles initially laid down on a regular grid by using the Zel'dovich approximation at $z = 99$. The initial power spectrum and growth rates are computed by rescaling the $z = 0$ matter power spectrum and transfer functions according to the method described by Zennaro et al. (2017), i.e., we account for both the scale dependence of the growth factor and growth rate in simulations with two fluids.

The *N*-body simulations were run using the GADGET-III code, last described in Springel (2005). They consist of two different sets. One set follows the evolution of 512^3 CDM particles in a periodic box of 1000 comoving h^{-1} Mpc, while in the other 256^3 CDM particles are evolved in a box size of 20 comoving h^{-1} Mpc. The gravitational softening is set to 50 and 2 comoving h^{-1} kpc, respectively. We call these sets N1000 and N20, and we use them to study the statistical properties of paired fixed fields on large and small scales (and for very massive and low-mass objects), respectively. Each set contains 300 simulations: 100 standard simulations and 100 pairs of fixed simulations. We will show results obtained from the N1000 set, while the N20 set is mainly used to cross-check the results of the H20 simulation set that we describe below.

We also have two different sets of magnetohydrodynamic simulations, run with the AREPO code (Springel 2010). In one, we follow the evolution of 640^3 CDM plus 640^3 gas particles in a periodic box of 200 comoving h^{-1} Mpc, while in the other we have a box of 20 comoving h^{-1} Mpc with 256^3 CDM plus 256^3 gas particles. Both use the IllustrisTNG models of galaxy formation, which include gas radiative cooling, star formation, metal enrichment, galactic winds, and black hole accretion and feedback (Weinberger et al. 2017; Pillepich et al. 2018b). The numerical methods and subgrid physics models build on the Illustris simulation model (Vogelsberger et al. 2013, 2014a, 2014b; Genel et al. 2014; Torrey et al. 2014). The softening lengths are 8 and 2 comoving h^{-1}kpc, respectively. We name these sets H200 and H20, correspondingly. H200 has 80 simulations, 30 standard and 25 pairs, while H20 is made up of 450 simulations, 250 standard and 100 pairs. We use the H200 simulations, which are very close to the TNG300-3 simulation (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich et al. 2018a; Springel et al. 2018), to study the improvement on different power spectra (matter, CDM, gas) introduced by paired fixed simulations on intermediate scales. The H20 set is used to study the properties of paired fixed simulations on very small scales and to investigate the impact of those fields on galaxy properties. A summary of our simulation suite is shown in Table 1.

Snapshots are saved at different redshifts, from $z = 15$ down to $z = 0$. In this work we focus on redshifts 0, 1, and 5. Dark matter halos are identified using the friends-of-friends algorithm (Davis et al. 1985) with a value of the linking length parameter $b = 0.2$. In the hydrodynamic simulations we identify galaxies through the SUBFIND algorithm (Springel et al. 2001). We use the algorithm described in Banerjee & Dalal (2016) to identify voids in the matter distribution of our snapshots.

3.2. Formalism

Here we describe the formalism we use to carry out the statistical analysis for each quantity considered in this paper. The most important goals of this work are to (1) study whether paired fixed simulations introduce a bias with respect to standard simulations and (2) quantify the statistical improvement achieved by fixed and paired fixed simulations in comparison with standard simulations.

Throughout the paper we show plots that share the same structure and contain information on the above two statistical properties. An example of such a plot appears in the left panel of Figure 2.

We compute each quantity for each standard and paired fixed realization in the considered simulation set. For example, the left panel of Figure 2 considers the matter power spectrum at $z = 99$. We denote by $X_{\rm s,i}$ and $X_{\rm pf,i}$ the value of that quantity from the realization i of the standard and paired fixed simulations, respectively. We compute the value of $X_{\rm pf,i}$ as

$$X_{\rm pf,i} = \frac{1}{2}[X_{\rm pf,i,1} + X_{\rm pf,i,2}], \tag{10}$$

where $X_{\rm pf,i,1}$ and $X_{\rm pf,i,2}$ are the considered quantities in each simulation of a paired fixed realization.[8] From $X_{\rm s,i}$ and $X_{\rm pf,i}$ we estimate the mean and variance of each simulation type as

$$\bar{X}_\alpha \equiv \langle X_\alpha \rangle = \frac{1}{N}\sum_{i=1}^{N} X_{\alpha,{\rm i}} \tag{11}$$

[8] In the case of the nonlinear power spectrum this demonstrably cancels phase correlation errors at leading order. For other quantities, it may be possible to construct improved estimators using cross-correlations between simulations, but this is beyond the scope of the present work.

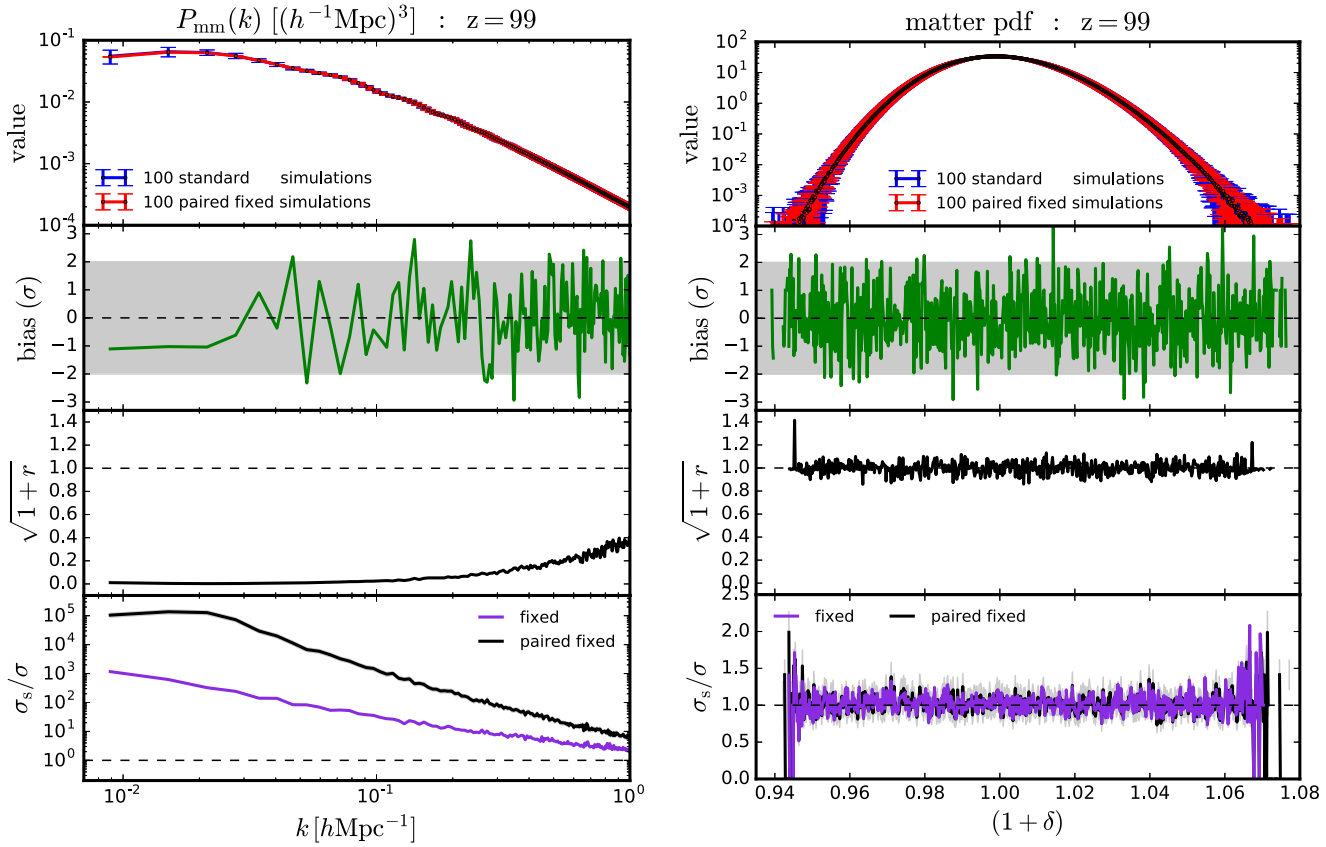

Figure 2. Most of the plots shown in this paper have the same structure, which we describe in detail here. First panel: mean and standard deviation of the considered quantity from the standard (blue) and paired fixed simulations (red). Second panel: an important aspect of our study is to investigate whether paired fixed simulations introduce a bias on the considered quantity; in this panel we show the difference between the means from the two simulation sets, divided by the expected error on the difference (see text for details). This panel quantifies thus the statistical agreement between both data sets. Any point within the gray region indicates that the agreement between means is within 2σ. Thus, points beyond the gray band indicate a bias on the means at more than 2σ. The black dashed line indicates a 0σ bias. Third panel: the other important quantify in paired fixed simulations is the statistical improvement they achieve with respect to standard simulations. For the considered quantity, the variance from standard and paired fixed simulations is given by $\sigma_{\rm s}^2$ and $\sigma_{\rm pf}^2$, respectively. We can express $\sigma_{\rm pf}$ as $\sigma_{\rm pf} = \sigma_{\rm f}\sqrt{(1+r)}$, where $\sigma_{\rm f}$ is the standard deviation of each individual pair and r is the cross-correlation coefficient between the pairs. Expressing the variance in that way is very helpful, as the improvement achieved by fixing the amplitude and flipping the phase is embedded in the value of $\sigma_{\rm f}$ and r, respectively. In this panel we show the value of $\sqrt{1+r}$. 0 values mean that the errors are completely anticorrelated and a large statistical improvement can be achieved. When the quantities from the two pairs are independent (dashed black line), $r = 0$ and $\sqrt{1+r} = 1$, and no statistical improvement is brought by pairing. If both simulations are completely correlated, the value of $\sqrt{1+r}$ is $\sqrt{2}$, and the normalized variance worsens. Fourth panel: this panel shows the ratios between the standard deviation from the standard and paired fixed simulations $\sigma_{\rm s}/\sigma_{\rm pf}$ (black line) and that from the standard and fixed simulations $\sigma_{\rm s}/\sigma_{\rm f}$ (purple line). That ratio indicates the statistical improvement achieved by the fixing and pair fixing over traditional simulations. The gray region around the black line in this panel represents the 1σ uncertainty on the ratio. The dashed line indicates 1 and can be interpreted as no statistical improvement with respect to standard simulations. We can see how fixed and paired fixed simulations largely reduce the scatter on the matter power spectrum, while they leave the variance on the matter density pdf unaffected.

$$\sigma_\alpha^2 \equiv \langle (X_\alpha - \bar{X}_\alpha)^2 \rangle = \frac{\sum_{i=1}^{N} (X_{\alpha,{\rm i}} - \bar{X}_\alpha)^2}{N-1}, \tag{12}$$

where $\alpha = \{{\rm s}, {\rm pf}\}$. The top panel always shows the mean and standard deviation from the standard and paired fixed simulations in blue and red, respectively.

In the second panel of each figure we quantify the bias introduced by the paired fixed simulations with respect to standard simulations. We calculate it by computing

$$\frac{\bar{X}_{\rm s} - \bar{X}_{\rm pf}}{\sigma_{\rm s-pf}}, \tag{13}$$

where $\sigma_{\rm s\text{-}pf}$ is the expected error on the difference between the means from the standard and paired fixed simulations. In this paper we have assumed that all the considered quantities are normally distributed. In that case, the expected error on the difference of the means is

$$\sigma^2_{\rm s\,-\,pf} = \frac{\sigma_{\rm s}^2}{N_{\rm s}} + \frac{\sigma_{\rm pf}^2}{N_{\rm pf}}, \tag{14}$$

where $N_{\rm s}$ and $N_{\rm pf}$ are the number of standard and paired fixed realizations. Notice that the above expression should also include a term to account for the covariance between the standard and paired fixed simulations. However, since the random seeds are different for the two setups, that covariance has an expected value equal to 0. We note that this is a reasonable assumption for power spectra, where the amplitude in a given k-bin receives contributions from many different independent modes. For halo mass function, void radius function, and pdf's, a more appropriate distribution will be a Poissonian. However, in this work we only show results for bins that contain many halos/voids/cells. In that case, the Poisson distribution is well approximated by a Gaussian.

The green line in the second panel thus measures the bias introduced by the paired fixed procedure, with respect to the standard simulations, in σ units. The gray band indicates where the bias is less than 2σ. Thus, any point beyond the gray region points toward a bias between the means larger than 2σ. It is important to emphasize that statistical fluctuations can give rise to deviations larger than 2σ. We note that this is a bias among the means, i.e., it is suppressed by the square root of the number of simulations. For a single realization, the absolute error is, in all the quantities studied in this work, negligible in practice.

In the third and fourth panels we quantify the statistical improvement achieved by the paired fixed simulations with respect to the standard simulations. The *normalized variance*[9] of the paired fixed simulations can be expressed as

$$\sigma^2_{\rm pf} = \sigma^2_{\rm f}(1 + r), \tag{15}$$

where $\sigma_{\rm f}$ is the variance of individual fixed simulations

$$\sigma^2_{\rm f} = \langle (X_{\rm pf,1} - \bar{X}_{\rm pf})^2 \rangle = \langle (X_{\rm pf,2} - \bar{X}_{\rm pf})^2 \rangle \tag{16}$$

and r is the cross-correlation coefficient between $X_{\rm pf,1}$ and $X_{\rm pf,2}$

$$r = \frac{1}{\sigma^2_{\rm f} N} \sum_{i=1}^{N} (X_{\rm pf,i,1} - \bar{X}_{\rm pf,1})(X_{\rm pf,i,2} - \bar{X}_{\rm pf,2}). \tag{17}$$

We note that $\sigma^2_{\rm f}$ can be interpreted as the variance obtained by fixing the amplitude without doing pairing (i.e., the variance of fixed simulations), while the value of r measures the correlation between the two sets of pairs and $1 + r$ can be interpreted as the statistical improvement on the variance achieved by pairing.

The third panel of Figure 2 shows the value of $\sqrt{1 + r}$. If the two pair quantities are independent, $r = 0$ and $\sqrt{1 + r} = 1$. In this case pairing does not bring any improvement, and the variance of paired fixed simulations will be just the variance of fixed simulations.[10] If the two pairs are completely correlated, $r = 1$, pairing does actually worsen the results. This happens because the second simulation adds no information and is therefore wasted. Finally, if both pair quantities are completely anticorrelated, $r = -1$, the variance of the paired fixed simulations reduces to 0. This can be understood taking into account that if both pair quantities are completely anticorrelated, as $X_{\rm pf,1}$ increases its value, $X_{\rm pf,2}$ shrinks such that $\frac{1}{2}(X_{\rm pf,1} + X_{\rm pf,1})$ remains constant. Thus, the lower the value of r, the larger the improvement brought by pairing. We emphasize that this is the improvement achieved by pairing once fixed. In other words, the value of r from just paired simulations that are not fixed can be different from that of paired fixed simulations (see Appendix A for further details). We provide explanations for the actual values in that figure in Section 4.

Finally, in the fourth panel of Figure 2 we show the ratios between the standard deviations of the standard and paired fixed simulations (solid black line), $\sigma_{\rm s}/\sigma_{\rm pf}$, and between the standard and fixed simulations (solid purple line), $\sigma_{\rm s}/\sigma_{\rm f}$. The purple line quantifies the statistical improvement achieved by fixing the amplitude, while the black line represents the gain obtained by fixing and pairing. We note that the black line can be obtained from the purple line and the line in the third panel through Equation (15). The dashed horizontal line in the fourth panel shows a value of 1, indicating the level where fixed and paired fixed simulations do not bring any statistical improvement over standard simulations. The black line is also surrounded by a gray shaded region (hard to see in Figure 2 owing to the large dynamic range), indicating the associated error on the standard deviation ratio, which we estimate as (see Appendix B)

$$\sigma\left(\frac{\sigma_{\rm s}}{\sigma_{\rm pf}}\right) = \frac{1}{2}\left(\frac{\sigma_{\rm s}}{\sigma_{\rm pf}}\right)\sqrt{\frac{2}{N_{\rm s}} + \frac{2}{N_{\rm pf}}}. \tag{18}$$

We only show it for the black line for clarity.

4. Large Scales: *N*-body

In this section we study the statistical properties of paired fixed fields on large scales using the N1000 N-body simulations. The halo catalogs are composed of all halos with masses above $\simeq 1.5 \times 10^{13}\, h^{-1}\, M_\odot$.

4.1. Initial Conditions

We start by quantifying the improvement achieved by paired fixed simulations at the level of initial conditions, as we naively expect that nonlinear evolution can, in general, only degrade it. We focus our analysis on the matter power spectrum and on the matter density pdf.

4.1.1. Clustering

For each realization of the standard and paired fixed simulations in the N1000 set we have computed the matter power spectrum at $z = 99$.

We show the results in the left panel of Figure 2. We find an excellent agreement between the results of both simulation sets, with paired fixed not introducing a bias on the results. Note that a few points show a bias larger than 2σ; this is expected under the assumption that the data are independent and normally distributed, which implies that $\simeq 5\%$ of the points should exhibit a bias larger than 2σ.

From the third panel we can see that the power spectra from the two simulations of each paired fixed realization are highly anticorrelated on almost all scales. We note that the deviation of $\sqrt{1 + r}$ from 0 is due primarily to aliasing. We have explicitly tested this by computing the power spectra using a grid with fewer cells. This anticorrelation is the origin of the large improvement that we obtain by pairing once we fix the amplitude, as we will see below.

From the fourth panel we can see how fixed simulations highly reduce the sample variance present in the standard simulations: from a factor of $\simeq 10^3$ at $k \simeq 10^{-2}\, h\,{\rm Mpc}^{-1}$ to a few at $k = 1\, h\,{\rm Mpc}^{-1}$. We find that the improvement worsens at smaller scales. This is an effect of the way the power spectrum is measured in an individual box; as we move to smaller scales, there are rapidly increasing numbers of modes per k-bin. Thus, the measured power spectrum asymptotes to the ensemble average at high k, and no initial improvement is achieved by fixing the power in this limit.

[9] While fixed simulations only contain one simulation, paired fixed simulations contain two. For quantities in which pairing and fixing do not help, we will still see an improvement when using paired fixed simulations simply because we are estimating the quantity through two simulations instead of one. We correct for that by computing the normalized variance, so that we can compare directly $\sigma_{\rm pf}$, $\sigma_{\rm f}$, and $\sigma_{\rm p}$ (see Appendix A for further details).

[10] Notice that we expect an improvement of $1/\sqrt{2}$ in the variance if we compute a quantity with two independent measurements instead of one (see Appendix A). However, in this work we are interested in the net gain, so we reabsorb that improvement in our definition.

Paired fixed simulations further reduce the sample variance amplitude with respect to fixed simulations, with ratios as large as 10^5 on the largest scales we probe. The improvement brought by pairing has its origin in the fact that the first-order nonlinear perturbations are canceled (Pontzen et al. 2016). Even at $z = 99$, the Zel'dovich approximation has introduced such nonlinearities.

4.1.2. Probability Distribution Function

We now investigate another key quantity to understand our results at lower redshifts: the pdf of the matter density field in real space.

For each initial condition realization of the standard and paired fixed simulations we have computed the matter density field by assigning particle positions to a grid with 128^3 cells using the cloud-in-cell (CIC) mass assignment scheme. We have then computed the pdf as the fraction of cells with matter overdensity, $1 + \delta = \rho/\bar{\rho}$, in the interval $[\rho, \rho + d\rho]/\bar{\rho}$. We show the results of our statistical analysis in the right panel of Figure 2.

As already pointed out in Angulo & Pontzen (2016), the pdf of paired fixed simulations shows a good agreement with that from standard simulations, as can be seen from the first panel. From the second panel we can see that paired fixed simulations do not introduce a bias on the matter density pdf with respect to the results from standard simulations.

In the third panel we show the cross-correlation coefficient between the pairs of the paired (orange) and paired fixed (black) simulations. We find that in both cases the value r is compatible with 0 ($\sqrt{1+r} = 1$), meaning that the results of both pairs are independent from each other. Thus, pairing does not help in reducing the variance on the matter density pdf from the standard simulations.

We show the statistical improvement achieved by fixed and paired fixed simulations, with respect to standard simulations, in the fourth panel. We find that all simulation types exhibit the same scatter as standard simulations. We do not find improvements on the variance amplitude for fixed or paired fixed simulations, meaning that fixing the amplitude does not reduce the pdf fluctuations either. In some ways this is a blessing: the local statistical properties of a fixed field are identical to the properties of its Gaussian counterpart, and therefore one can expect local physics such as galaxy formation to proceed correctly in a fixed universe.

Thus, we conclude that while paired fixed simulations can reduce the scatter on the power spectrum of the initial conditions by large factors, the pdf does not benefit from this and its scatter remains unchanged. We will see below that other quantities tightly related to the pdf, such as the halo mass function, the void radius function, the stellar mass function, and intrinsic galaxies properties, will not exhibit significant statistical improvement when estimated using paired fixed simulations.

We find only modest improvements on the variance of paired fixed simulations on the halo mass function, matter density pdf, and star formation rate history when analyzing the H20 simulations, as we will see in Section 6.

4.2. Clustering

For each simulation in N1000 we have computed the matter and halo auto-power spectrum and the halo–matter cross-power spectrum. The results of our statistical analysis are displayed in Figure 3.

The top row shows the results for the matter power spectrum at redshifts 0 (left), 1 (middle), and 5 (right). The bottom row displays the results for the halo–matter cross-power spectrum (left) and halo power spectrum (right). For those quantities we only show results at $z = 0$, since at $z = 5$ the number density of halos in our simulations is very low and results are very similar at $z = 1$.

From the first panels we can see that the agreement between the results of the standard and paired fixed simulations is very good in all cases. In the second panels we quantify the bias introduced by the paired fixed simulations with respect to standard simulations and find no evidence for a bias for any of the three quantities at the different redshifts considered. We emphasize that with a finite number of simulations, this kind of claim has to be considered as an upper bound. It may be that paired fixed simulations induce a bias on those quantities, but its magnitude is too small for detection with 100 realizations. We note that in some cases, e.g., the matter power spectrum at $z = 0$, there seems to be a systematic bias offset on small scales. This is, however, due to the fact that modes on those small scales are highly correlated, through nonlinear evolution, and therefore not fully independent.

In the third panels we show the value of the cross-correlation coefficient. We find that for all the considered quantities on small scales, $k \gtrsim 0.2$–$0.5\,h\,\mathrm{Mpc}^{-1}$, its value is compatible with $\sqrt{1+r} = 1$, indicating that the power spectra from the two pairs are independent. In that case, pairing does not help in reducing the statistical error due to sample variance. We find that the value of $\sqrt{1+r}$ is smaller than 1 on scales larger than $k \lesssim 0.2$–$0.5\,h\,\mathrm{Mpc}^{-1}$. The scale at which $\sqrt{1+r}$ equals 1 decreases with redshift, independently of the considered power spectrum, but the effect is more pronounced in the matter power spectrum.

The value of $\sqrt{1+r}$ for the matter power spectrum can be as low as $\simeq 4 \times 10^{-2}$, pointing out that pairing, once fixed, can reduce the scatter of the standard simulations by that factor. The value of $\sqrt{1+r}$ increases with scale, until reaching the value of 1. At $z = 5$, however, we find a dip around $3 \times 10^{-2}\,h\,\mathrm{Mpc}^{-1}$. It is interesting to note that the lowest values of $\sqrt{1+r}$ take place at $z = 0$ rather than $z = 5$. At present, we do not have an explanation for this.

We find much higher values of $\sqrt{1+r}$ for the matter–halo and the halo–halo power spectra than for the matter power spectrum. In those cases, we also find a dip around $4 \times 10^{-2}\,h\,\mathrm{Mpc}^{-1}$. On large scales, the value of $\sqrt{1+r}$ barely goes below 0.7, indicating that pairing can only reduce the variance by $\simeq 0.7$. We note that halos are the main driver of the increase in the value of $\sqrt{1+r}$, as on large scales the halo power spectrum barely deviates from 1.

From the fourth panels of Figure 3 we can see that for the matter power spectrum on large scales, reductions on the standard deviation of standard simulations can be as large as 10^3, at all redshifts considered. Since the standard deviation on the mean from standard simulations goes as $\propto 1/\sqrt{N_s}$, the above numbers can be interpreted as follows. A single paired fixed simulation can be used to evaluate the amplitude and shape of the matter power spectrum on large scales, with an error equal to that achieved by running $\sim 10^6$ standard simulations. On small scales, $k \sim 0.2$–$0.4\,h\,\mathrm{Mpc}^{-1}$, the ratios tend to 1, showing that no improvement is achieved by the

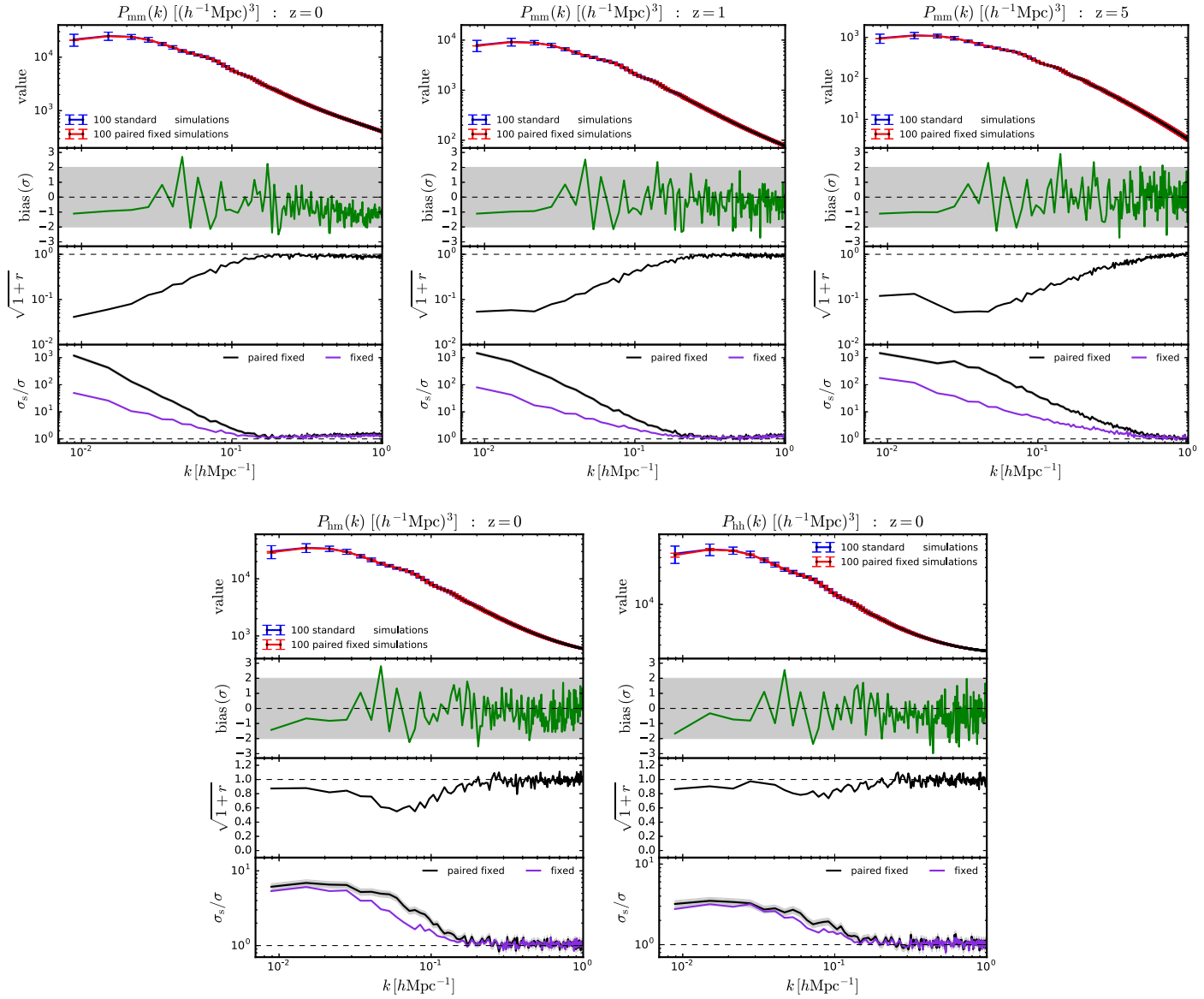

Figure 3. Impact of paired fixed simulations on the clustering of matter (top row), on halos (bottom right), and on the halo–matter cross-power spectrum (bottom left). We show results at redshifts 0 (left column), 1 (middle column), and 5 (right column) for matter and at $z = 0$ for the halo and halo–matter power spectra. Paired fixed simulations can reduce the sample variance scatter on these power spectra by large quantities without introducing a bias on them.

fixed or paired fixed simulations. The reason why fixed and paired fixed simulations do not improve the statistics of standard simulations on small scales is that in the nonlinear regime modes get mixed in a complicated manner that affects both the amplitudes and phases and gives rise to sample variance. We show this schematically in Figure 4, where a set of complex numbers with fixed amplitude end up with very different amplitudes after each mode mixes with its neighbors. This happens because whether complex numbers with the same amplitude add up or cancel depends on whether their phases align. In fixed and paired fixed simulations the phases are random.[11]

We note, however, that on very small scales and at $z = 0$, the $P_{\rm mm}(k)$ results for σ_s/σ are between 2 and 3. The improvement on those scales is mostly coming from fixing the amplitude rather than from pairing. We also observe this effect in the smaller box size simulations that we study in Section 5. Thus, the above argument can explain the behavior we find in simulations only qualitatively.

The statistical improvement on large scales is much smaller for the halo–matter and halo–halo power spectra. For the halo–matter cross-power spectrum, we reach values of $\sigma_{\rm s}/\sigma_{\rm pf} \simeq 6$ on large scales at both redshifts 0 and 1. For the halo auto-power spectrum those values shrink to $\sigma_{\rm s}/\sigma_{\rm pf} \simeq 3$. For those two power spectra no statistical improvement is achieved by fixed or paired fixed simulations on scales smaller than $k \gtrsim 0.3\,h\,{\rm Mpc}^{-1}$. Notice that we have not subtracted the shot-noise amplitude in the halo auto-power spectra. In the following subsections we will see how paired fixed simulations do not introduce a bias on the halo mass function. Thus, the halo mass functions from standard and paired fixed simulations will, on average, be identical. In that case, the mean abundance

[11] In paired fixed simulations there is a correlation between the phases in the initial conditions of the two simulations in a pair. The argument regarding random phases applies, however, to the mode mixing of each individual simulation in a pair.

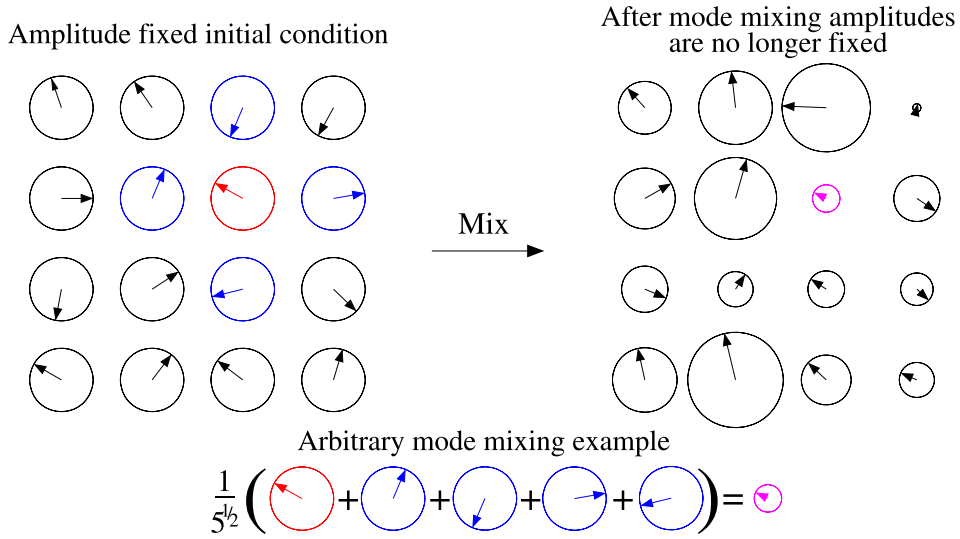

Figure 4. Illustration of how mode mixing from nonlinear evolution introduces sample variance in fixed simulations. Left: in the initial conditions each cell in Fourier space has a random phase (the arrow) but a fixed amplitude (the radius of the circle). Right: after each mode mixes with its neighbors, some amplitudes grow and others diminish, depending on how their phases align with those of their neighbors. Hence, the amplitudes are no longer fixed after mixing, reintroducing sample variance. The exact form of mixing shown here is just an example, while any kind of mixing will have similar effects.

of halos will be the same in both setups, and the shot-noise amplitude will be identical in the two situations.

We conclude that while paired fixed simulations can yield very large statistical improvements, $\sigma_s/\sigma_{pf} \simeq 1000$, for the matter power spectrum, for the halo–matter and halo power spectra the gain is much smaller, $\sigma_s/\sigma_{pf} \sim 5$, but still valuable.

4.3. Halo Bias

We now turn our attention to the halo bias. For each standard and paired fixed realization we have computed the halo bias using the estimator

$$b(k) = \frac{P_{hm}(k)}{P_{mm}(k)}. \tag{19}$$

We show the results of our statistical analysis in the top left panel of Figure 5. We only show the results at $z = 0$ since at $z = 1$ our conclusions are unchanged. From the first panel we see the very good agreement between the results of both simulations, while in the second panel we show that paired fixed simulations do not introduce a bias on this quantity. The value of the cross-correlation coefficient is, for almost all scales, compatible with 0 ($\sqrt{1+r} = 1$), implying that pairing does not help in reducing the scatter. Finally, in the fourth panel we can see how fixing the amplitude does not reduce the scatter either, and therefore paired fixed simulations exhibit the same scatter in the halo bias as standard simulations.

We have repeated the above analysis by computing the bias as $b(k) = \sqrt{P_{hh}(k)/P_{mm}(k)}$, reaching identical conclusions: fixed and paired fixed simulations exhibit the same scatter on the halo bias as standard simulations.

This result may appear surprising at first since, as we saw above, paired fixed simulations can reduce the scatter on the matter, halo–matter, and halo power spectra by factors as large as 10^3, 6, and 3, respectively. In order to understand the reason for this result, let us write the variance of the halo bias at linear order (see Appendix B for the derivation)

$$\sigma_b^2 = \frac{1}{2N(k)}\left(\frac{P_{hh}(k)P_{mm}(k) - P_{hm}^2}{P_{mm}^2}\right), \tag{20}$$

where $N(k)$ is the number of independent modes in the considered k-interval and the halo power spectrum includes both cosmological signal and the shot-noise term, i.e., $P_{hh}(k) = P_{hh}^{cosmo}(k) + 1/\bar{n}$. If the shot-noise amplitude were zero, $P_{hh} = b^2 P_{mm}$, $P_{hm} = bP_{mm}(k)$, and the linear-order variance of the bias would be zero too. This tells us that in the absence of shot noise, the halo–matter and matter–matter power spectra are perfectly correlated, and hence their ratio, the halo bias, has zero variance.

What breaks this perfect correlation and becomes the source of variance in our bias estimate is the presence of shot noise in the halo power spectrum. The amplitude of the shot noise on large scales is the same in paired fixed and standard simulations, as it only depends on the halo number density. Thus, it should not be surprising after all that the scatter in the halo bias from paired fixed and standard simulations is the same as well.

We leave for future work a formal derivation of this result on mildly nonlinear scales and a deeper understanding of why paired fixed simulations do not even reduce the scatter of the halo bias on nonlinear scales.

4.4. Matter Density pdf

We now focus our attention on the pdf of the matter density field. For each realization of the standard and paired fixed simulations we have computed the matter density field on a grid with 128^3 cells using the cloud-in-cell (CIC) mass assignment scheme. We show the results of our analysis in the top right panel of Figure 5. We only show results at $z = 0$ since results at higher redshift do not change our conclusions.

Unlike the matter power spectrum, where both pairing and amplitude fixing greatly reduced the variance, the matter density pdf is indifferent to these techniques, at least on this scale. There

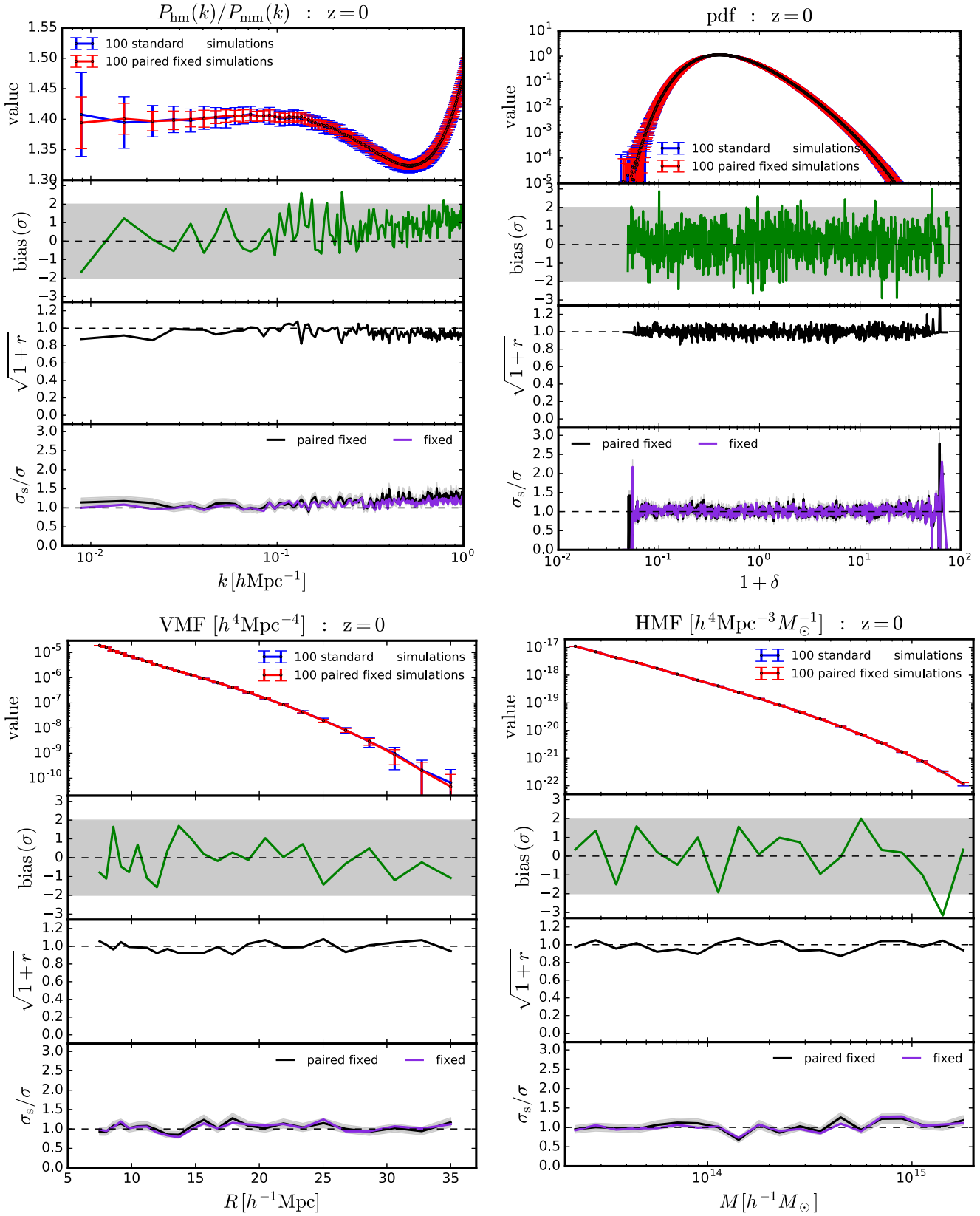

Figure 5. Impact of paired fixed simulations on halo bias (top left), matter density pdf (top right), void radius function (bottom left), and halo mass function (bottom right) at $z = 0$ from the N1000 simulation set. Paired fixed simulations do not introduce a bias on these quantities, but they do not reduce their scatter either.

is, however, no harm: the bias is consistent with zero, showing full agreement between standard and paired fixed simulations. But there is also no benefit: the value of $\sqrt{1+r}$ is consistent with 1, so pairing after fixing is of no help, and all the effects of amplitude fixing are washed out in this basis. See, however, Section 6.3 for how this changes on smaller scales.

We can interpret these results by taking into account that paired fixed simulations do not reduce the scatter on the pdf already at the starting redshift of the simulation (see Section 4.1). Thus, it is unlikely that nonlinear evolution would lead to different pdf's at low redshift.

4.5. Halo Mass Function and Void Radius Functions

Here we study the impact of paired fixed simulations on the halo and void radius functions. For each standard and paired fixed simulation we have computed the halo mass function, defined as the number density of halos per mass interval. We show the results in the bottom right panel of Figure 5. We have also computed the void radius function, defined as the number density of voids per radius interval, for each realization of the standard and paired fixed simulations. We show the results in the bottom left panel of Figure 5. For both cases we only show results at $z = 0$, as higher redshifts lead to identical conclusions.

From the top panels we find that the agreement between the standard and paired fixed simulations is very good for both the halo and void radius functions, and in the second panels we show that no bias is introduced on these quantities by the paired fixed simulations. In the third panels we show the cross-correlation coefficient from the results of each pair. Our results are compatible with the $\sqrt{1+r} = 1$, pointing out that the results of each pair are independent. From the fourth panels we find that there is no statistical improvement on these two quantities from fixed or paired fixed simulations.

We believe that paired fixed simulations do not improve the abundance of halos and voids statistics because the formation of those takes place on small scales, where the one-point properties are more relevant to determining the final outcome. A different way to see this is to take into account that the formation and evolution of halos and voids will be more sensitive to phases than to amplitudes. Besides, when pairing, a halo will become a void and vice versa (Pontzen et al. 2016), so it seems unlikely that paired fixed simulations can help in reducing the scatter of these quantities. As we saw in Section 4.1, these are not affected by the fixing and pairing procedure.

In Section 6 we will, however, see that paired fixed simulations slightly reduce the scatter of the halo mass function and matter density pdf when analyzing hydrodynamic simulations with small box sizes. This may be related to nonlinearities reaching the halo filtering scale, but further exploration is deferred to future work.

We thus conclude that large-scale box size paired fixed simulations reduce the scatter on clustering quantities like the matter, halo–matter, or halo power spectra. They do not, however, help in reducing the scatter of the halo bias or on one-point statistics like the halo or void radius functions, or the matter density pdf.

5. Intermediate Scales: Hydrodynamic

In this section we investigate the statistical properties of paired fixed simulations on intermediate scales using state-of-the-art magnetohydrodynamic simulations. We carry out the statistical analysis using the H200 simulations. Those simulations are computationally expensive, so we could only run 80 of them: 30 standard and 25 paired fixed realizations. This small number of simulations does not allow us to reach robust statistical conclusions for most of the quantities considered in this paper. For this reason we focus our analysis on clustering, where the effect is large enough to establish that paired fixed simulations do reduce the intrinsic scatter due to sample variance.

We have also computed the matter density pdf, the halo mass function, and the void radius function, and our results are in agreement with those from large scales, i.e., paired fixed simulations do not introduce a bias but also do not reduce the intrinsic scatter. However, the associated error bars are too large to rule out a small statistical improvement such as that we observe in the H20 simulations (see Section 6).

5.1. Clustering

For each standard and paired fixed simulation we have computed the power spectrum of matter, CDM, gas, stars, and galaxies at redshifts 0, 1, and 5. The relatively low resolution of the H200 simulation highly affects the power spectrum of stars and galaxies, due to the large amplitude of the shot noise, on all scales we probe. Hence, we focus our analysis on the matter ($P_{\rm mm}(k)$), CDM ($P_{\rm cc}(k)$) and gas ($P_{\rm gg}(k)$) power spectra.

We show our results in Figure 6. From the first panels we deduce that the agreement between the different power spectra from the different simulations is very good at all redshifts. From the second panels we see that no bias is introduced by paired fixed simulations, with respect to standard simulations, on these power spectra at $z = 5$. However, at lower redshifts, in some situations and for some scales, the bias deviates more than 2σ. In some cases a large set of points deviate continuously from 2σ. We emphasize that those scales are highly correlated, so it is expected that if one scale deviates, the others will exhibit the same behavior. Since the number of realizations we have in the H200 set is very small, it is not unreasonable to expect mean differences of $\simeq 2\sigma$. We find similar results at $z = 0$, where in some cases, e.g., gas power spectrum on very small scales, the difference between the mean of both data sets can be around 3σ, but again, on highly correlated scales. In order to verify that this bias is not statistically significant, we have repeated the above analysis but removing some random paired fixed or standard simulations. By doing so, we find that in some of the cases the bias between the two data sets decreases and remains below 2σ. This points out that our low number of realizations may be underestimating the intrinsic scatter. Furthermore, as we will see in the next section, with a much larger number of hydrodynamic simulations covering a range of scales similar to those we explore here, we do not find a bias on any of the power spectra studied here. We think that the bias we find in the H200 simulations may be due to statistical fluctuations. More simulations are, however, needed to clearly disentangle this issue. We emphasize again that this is the bias among the means. Even if the bias will be real and not a statistical fluctuation, for a single realization the absolute magnitude of this bias will be very small.

We find that the power spectra from the two pairs are strongly anticorrelated on large scales, for all the considered fields. This translates, as we shall see below, into large statistical improvements of the paired fixed simulations with respect to standard simulations. On smaller scales the value of the cross-correlation coefficient tends to zero, although it usually remains smaller than zero. We note that at $z = 0$ and for $k \simeq 1.5\,h\,{\rm Mpc}^{-1}$, the cross-correlation coefficient exhibits a

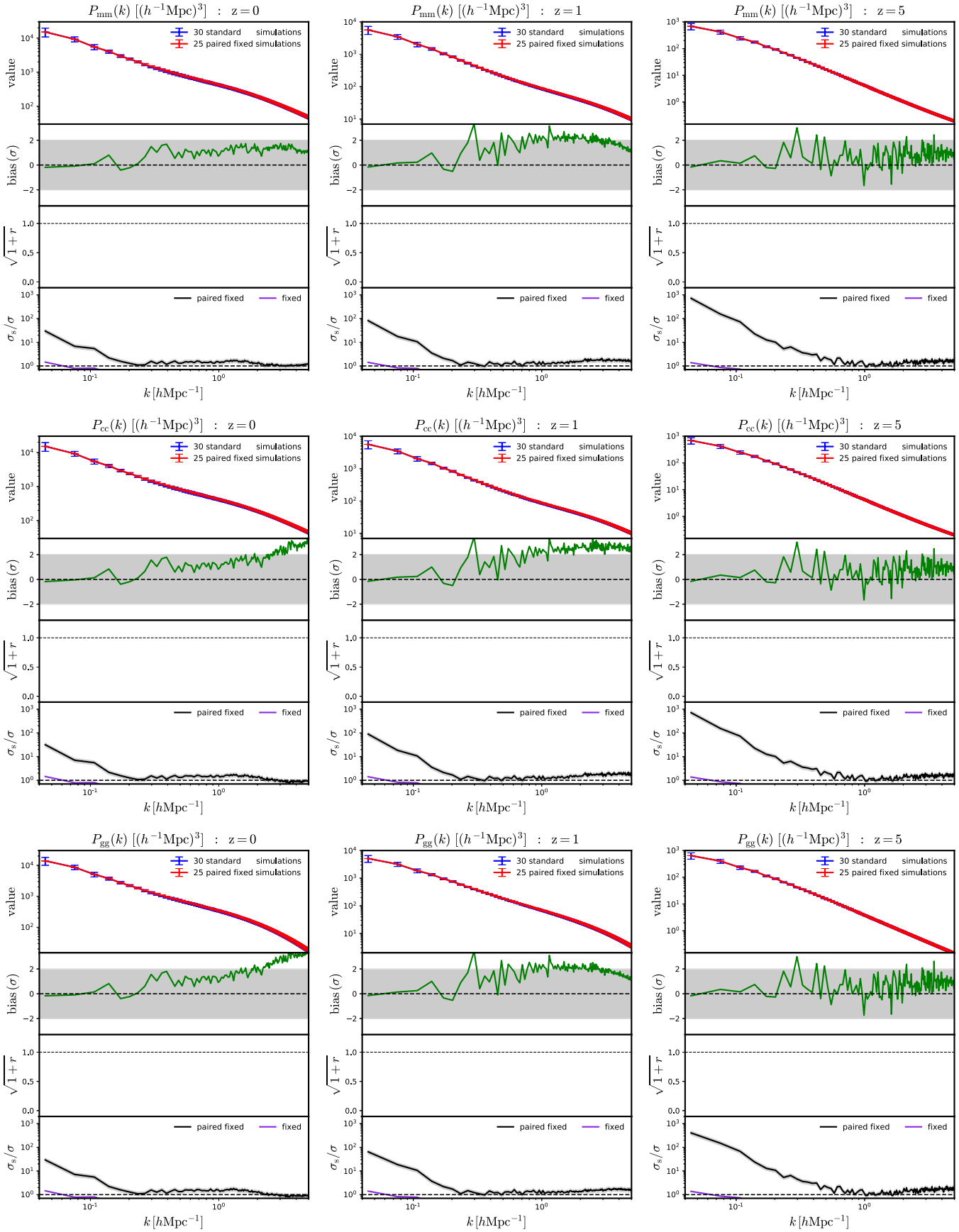

Figure 6. Impact of paired fixed simulations on the clustering of matter (top row), CDM (middle row), and gas (bottom row) from the magnetohydrodynamic simulations set H200 at redshifts 0 (left column), 1 (middle column), and 5 (right column). Paired fixed simulations largely reduce the sample variance errors associated with standard simulations on large scales. We believe that some of the rather large bias values we find are not statistically significant.

sizable dip. That dip also seems to take place at higher redshifts but on smaller scales.

On large scales and for fixed simulations we find an improvement on the standard deviation of standard simulations that ranges from $\simeq 7$ at $z = 0$ to $\simeq 30$ at $z = 5$. The improvement for paired fixed simulations is much higher, induced by the low values of the cross-correlation coefficient. It is worth pointing out that running one paired fixed realization can be used to determine the mean of the matter, CDM, or gas power spectrum with an error equal to that achieved by running $\simeq 900$ standard simulations for $k \simeq 0.1\,h\,\mathrm{Mpc}^{-1}$, a very important scale for baryonic acoustic oscillation (BAO) studies.[12] On smaller scales the statistical improvement vanishes, although we observe some residual improvement on scales where the value of the cross-correlation coefficient is below 0.

We thus conclude that paired fixed simulations bring large statistical improvements on the matter, CDM, and gas power spectra on large scales from full hydrodynamic simulations.

6. Small Scales: Hydrodynamic

We now push the limits of paired fixed simulations by studying their properties on small scales through the H20 hydrodynamic set. We focus our analysis on clustering, one-point statistics, and internal galaxy properties.

6.1. Initial Conditions

We have computed the matter, CDM, and gas power spectra of each realization of the H20 simulations. The result of our statistical analysis for these quantities is similar to what we found for the N1000 simulations, i.e., a very large improvement on the largest scales of the box, while on smaller scales the variance reduction is smaller. We thus do not show these results, as they do not add much to our discussion.

We have also computed the matter density pdf for each standard and paired fixed realization of the H20 simulations using a grid with 64^3 cells by employing the CIC mass assignment scheme. Figure 7 shows the result of our analysis. We find that paired fixed simulations do not introduce a bias on the matter density pdf of the standard simulations.

The value of $\sqrt{1+r}$ is compatible with 1 for almost all overdensities, with deviations being mostly statistical fluctuations. From the fourth panel we can see that both fixed and paired fixed simulations reduce the scatter of the matter density pdf of standard simulations in a nontrivial way. Those improvements, although small, are not statistical fluctuations. We obtain very similar results for the matter field when using the N20 simulations. We leave it for future work to understand the reason why paired fixed simulations reduce the scatter of the matter density pdf relative to standard simulations in the way they do.

6.2. Clustering

For each standard and paired fixed realization we have computed the power spectrum of matter, CDM, gas, magnetic fields, stars, galaxies ($P_{\rm gal}(k)$), halos, and halo–matter. In Figure 8 we show the results at redshifts 0, 1, and 5 for the total matter and gas power spectra (for gas only at redshifts 0 and 1) in the top and middle rows, respectively. We do not show the

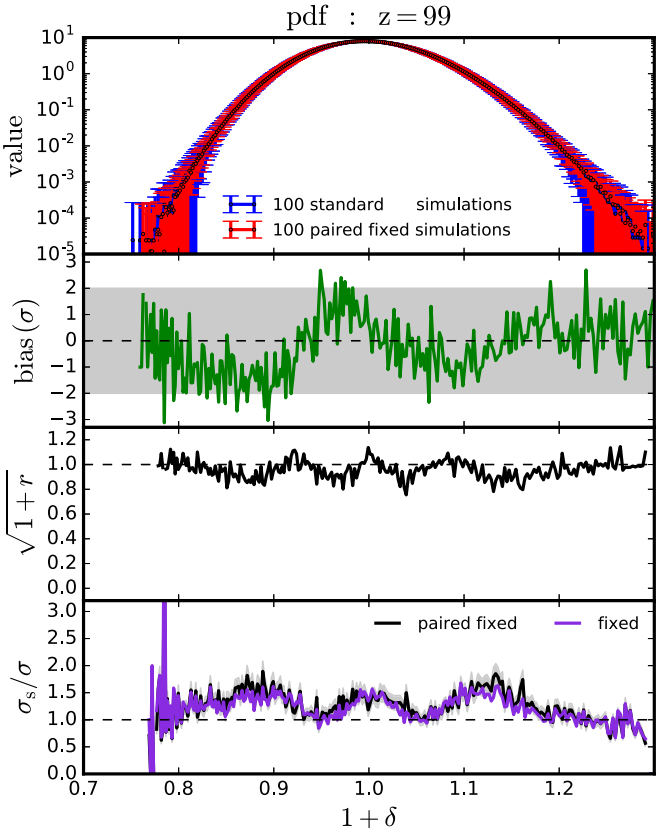

Figure 7. Impact of paired fixed simulations on the matter density pdf of the initial conditions of the H20 simulation set. Contrary to what we found for the N1000 simulations, for small smoothing scales we find that fixed and paired fixed simulations can slightly reduce the scatter in the matter density pdf without introducing a bias on it. The statistical improvement takes places over different overdensities in a complicated manner.

results for CDM since they are pretty similar to those from total matter and gas. The results for gas at $z = 5$ are also similar to those of matter at that redshift. In the middle right panel we show the results for the magnetic field power spectrum, while in the bottom row we display our findings for the stars, galaxies, and halo–matter power spectra. Since our conclusions for those components do not change significantly with redshift, we only show those at $z = 0$.

From the first panels we can see that the agreement between the results of the two simulation types is very good. In the second panels we demonstrate that for all power spectra and considered redshifts the bias introduced by paired fixed simulations, with respect to standard simulations, is compatible with 0. We emphasize that the scales we probe with the H20 simulations are highly nonlinear and correlated, as expected. This is why the green curves look so smooth in comparison with those of, e.g., Figure 3.

We find that on almost all scales, for all power spectra, and at all redshifts, the value of $\sqrt{1+r}$ is lower than 1, pointing out that the power spectra from the two pairs of the paired fixed simulations exhibit a degree of anticorrelation. At $z = 0$, and on the largest scales we can probe with the H20 simulations, the value of $\sqrt{1+r}$ is around 0.7. At higher redshift that value shrinks, reaching $\simeq 0.2$ for matter, CDM, and gas at $z = 5$. As we move to smaller scales, the value of the cross-correlation coefficient increases. At low redshift and for matter, CDM, gas, magnetic fields, and galaxies it tends to 1, while for stars it

[12] Notice that boxes with $200\,h^{-1}$ Mpc may be too small to properly capture the BAO feature.

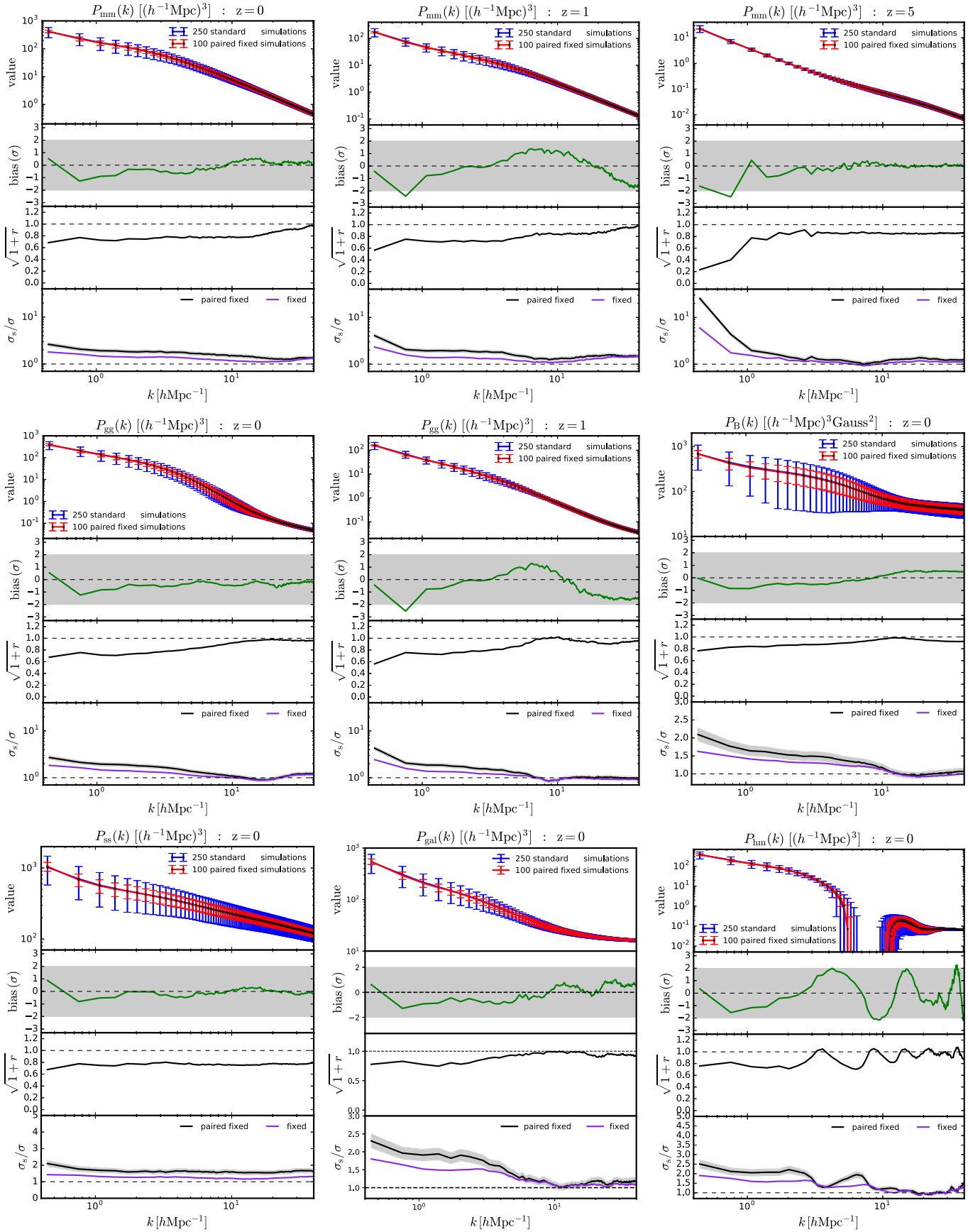

Figure 8. Impact of paired fixed simulations on the power spectrum of matter (top row), gas (left and middle panels of middle row), magnetic fields (right panel of middle row), stars (bottom left), galaxies (bottom middle), and halo–matter (bottom right) from the N20 simulations. Results for matter and gas are shown at redshifts 0 (left column), 1 (middle column), and 5 (right column, only for matter), while for magnetic fields, stars, galaxies, and the halo–matter we only show results at $z = 0$ since we observe little time evolution. Paired fixed simulations do not introduce a bias on any of these quantities and improve the statistics of standard simulations.

remains quite constant at $\sqrt{1+r} \simeq 0.8$. We observe a similar behavior at $z = 5$ for matter, CDM, and gas.

In the fourth panel we see that at low redshift the improvement on the sample variance reduction is moderate, with the standard deviation ratio reaching factors of 2 to 3 for matter, CDM, and gas on the largest scales. For the magnetic fields, stars, and galaxies the improvement is slightly lower. In the case of stars the improvement is surprisingly very scale independent. As we move to smaller scales, the improvement decreases, although showing a nonmonotonic dependence with redshift. The difference between the improvement from fixed and paired fixed is not large at low redshifts, while at $z = 5$ it can be a factor of almost 5 on the largest scales.

We note that the power spectrum of the magnetic field, stars, and galaxies is highly affected by shot noise (which we do not attempt to subtract). It is thus interesting to see that paired fixed simulations help to reduce the intrinsic error on it.

We find very interesting results for the halo–matter cross-power spectrum. First, notice that at $k \in 5$–$10\,h\,\mathrm{Mpc}^{-1}$ the value of the cross-power spectrum becomes negative. Second, the value of $\sqrt{1+r}$ exhibits an oscillatory behavior that is not due to statistical fluctuations and whose value is, in almost all scales, below 1. From the fourth panel we can see how on scales larger than $\simeq 2\,h\,\mathrm{Mpc}^{-1}$ fixed and paired fixed simulations slightly improve the statistics of the standard simulations. The oscillatory features we found in the value of the cross-correlation coefficient are reflected in the statistical improvement of paired fixed simulations, although fixed simulations also present that behavior, to a lesser extent.

We find similar oscillatory features in the halo auto-power spectrum and the halo bias. While the former are not due to the behavior of the cross-correlation coefficient, the latter exhibit the same features as the halo–matter power spectrum. We believe that the oscillations in the standard deviation ratio of the different halo power spectra are related to the features we observe in the matter density pdf of the initial conditions, which propagate to the matter density pdf and halo mass function at lower redshift (see next section). A more detailed study of this is beyond the scope of the present paper.

We thus conclude that even with small box size hydrodynamic simulations where all scales are nonlinear at low redshift, paired fixed simulations always produce power spectra with lower scatters than those from standard simulations. The statistical improvement can be pretty large at high redshift. Our results also point out that paired fixed simulations do not introduce a bias on any of the above power spectra.

6.3. One-point Statistics

We now study the impact of paired fixed simulations on one-point statistics. We focus our analysis on the halo mass function, the void radius function, the matter density pdf, the star formation rate history, and the stellar mass function. We only show results at $z = 0$,[13] since our conclusions are unchanged at higher redshifts.

6.3.1. Halo Mass Function

For each realization of the standard and paired fixed simulations we extracted halo catalogs by selecting all halos with masses above $\simeq 9 \times 10^8\,h^{-1}\,M_\odot$. We then computed the halo mass function for each realization, and we show the results in the top left panel of Figure 9. We find an excellent agreement between the results of both simulation types, and our results point out that paired fixed simulations do not introduce a bias. We can also see that the value of the $\sqrt{1+r}$ is compatible with 1 for all halo masses.

From the fourth panel we can see how fixed and paired fixed simulations slightly reduce the scatter on the halo mass function from standard simulations for some halo masses. This contrasts with our results of Section 4, where we found that paired fixed simulations do not reduce the scatter in the halo mass function. Note, however, that in Section 4 we only probed halos with masses above $\simeq 10^{13}\,h^{-1}\,M_\odot$; thus, for the halo mass range common to both simulations, our results are in agreement.

We note that the statistical improvement is not very significant, taking into account the error bars associated with the paired fixed simulations. In order to verify the robustness of these results, we have repeated the same analysis but using the N20 simulations, which are *N*-body and contain a different number of paired fixed realizations. By doing so, we find very similar results to what we find with the H20 simulations, implying that the improvement is not a statistical fluctuation but a physical effect.

Understanding the origin of this improvement on the halo mass function of small halos is beyond the scope of the current work.

6.3.2. Void Radius Function

For each realization of the standard and paired fixed simulations we have extracted voids in the matter field. In the top middle panel of Figure 9 we show the results for the void radius function. As always, we find a good agreement between the results of both simulations types and a bias between the mean of both simulations that is below $\simeq 2\sigma$. The value of the cross-correlation coefficient is compatible with 0 ($\sqrt{1+r} = 1$) for most of the void radii. We find that fixed and paired fixed simulations do not reduce the scatter on the void radius function. This is in agreement with our findings for larger voids in Section 4.

6.3.3. Matter Density pdf

We have computed the matter density field on a grid with 64^3 cells using the CIC interpolation scheme for each realization of the standard and paired fixed simulations. Our results for the pdf of the matter field are shown in the top right panel of Figure 9. We find good agreement among the results of both simulation types and that most of the points are below 2σ. The value of $\sqrt{1+r}$ is compatible with 1 for all overdensities with the exception of two dips for values of $1+\delta$ around 0.04 and 0.4. The origin of those dips is unclear to us, but we have verified that they are not statistical fluctuations. We obtain similar results by using the N20 simulations.

The fourth panel shows the statistical improvement achieved by fixed and paired fixed simulations with respect to standard simulations. We find that for overdensities larger than $\simeq 5$, fixed and paired fixed simulations do not reduce the intrinsic scatter of the standard simulations. For lower overdensities, we do, however, observe improvements. Those come from both the fixed and paired fixed simulations and manifest themselves as two bumps for overdensity values similar to those quoted above. In paired fixed simulations the improvement is more pronounced on those bumps owing to the anticorrelation of the

[13] For the star formation rate history we show results between redshifts 0 and 15.

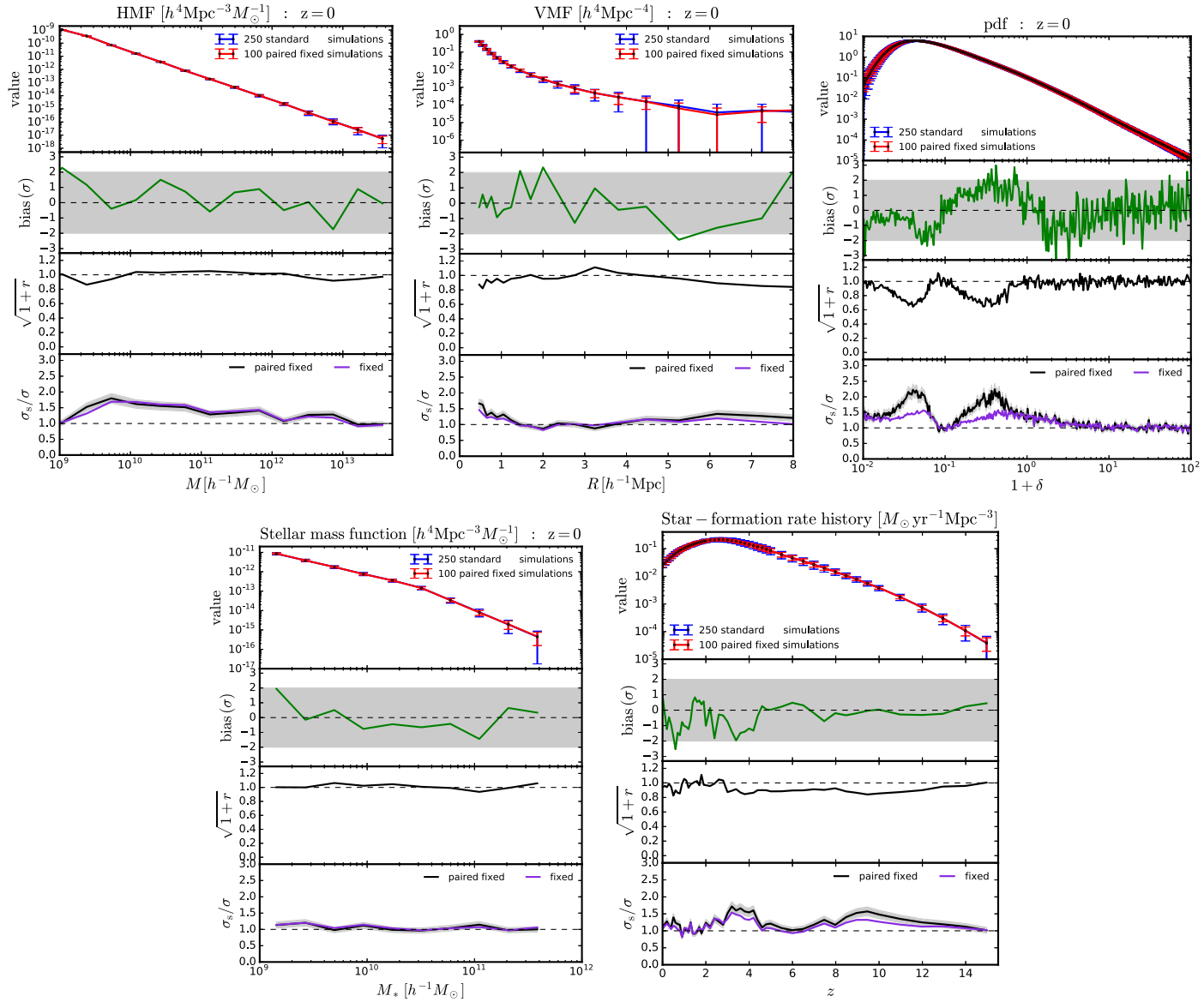

Figure 9. Impact of paired fixed simulations on the halo mass function (top left), void radius function (top middle), matter density pdf (top right), stellar mass function (bottom left), and star formation rate history (bottom right) from the N20 magnetohydrodynamic simulations at $z = 0$. We find similar results at higher redshifts. Paired fixed simulations do not introduce a bias on any of these quantities, and they slightly improve the statistics of some quantities in a complicated manner.

pdf's we find in the third panel. This result is different from what we found in Section 4, where we concluded that paired fixed simulations do not reduce the scatter of the matter density pdf. We note, however, that the scales we are probing in the two cases are very different. Besides, for these very small smoothing scales, we find that paired fixed simulations slightly improve the statistics of the matter density pdf already in the initial conditions (see Section 6.1). Future investigation of this effect will be required to disentangle whether the improvement propagates from the initial conditions or is brought by nonlinear evolution.

6.3.4. Stellar Mass Function

The results for the stellar mass function from the H20 simulation set are shown in the bottom left panel of Figure 9. As expected, the results from the two simulation types show a good agreement, and we find no bias between their means within 2σ (first and second panels). The third panel shows that the value of $\sqrt{1+r}$ is compatible with 1 for all stellar masses. Finally, we find no evidence for statistical improvement of fixed and paired fixed simulations over standard simulations for the stellar mass function (fourth panel).

6.3.5. Star Formation Rate History

We have computed the star formation rate history of each standard and paired fixed realization as the sum of the star formation rates of all gas particles divided by the simulation volume. That quantity informs us about the rate at which stars are being formed at a given redshift and therefore complements the stellar mass function when studying overall abundance. We show the results of our statistical analysis in the bottom right panel of Figure 9. The agreement between the results of both simulations is excellent, and we find no evidence that paired fixed simulations introduce a bias on that quantity. The value of the cross-correlation coefficient is compatible with 0 ($\sqrt{1+r}=1$), although between redshifts 4 and 12 it is less than 1.

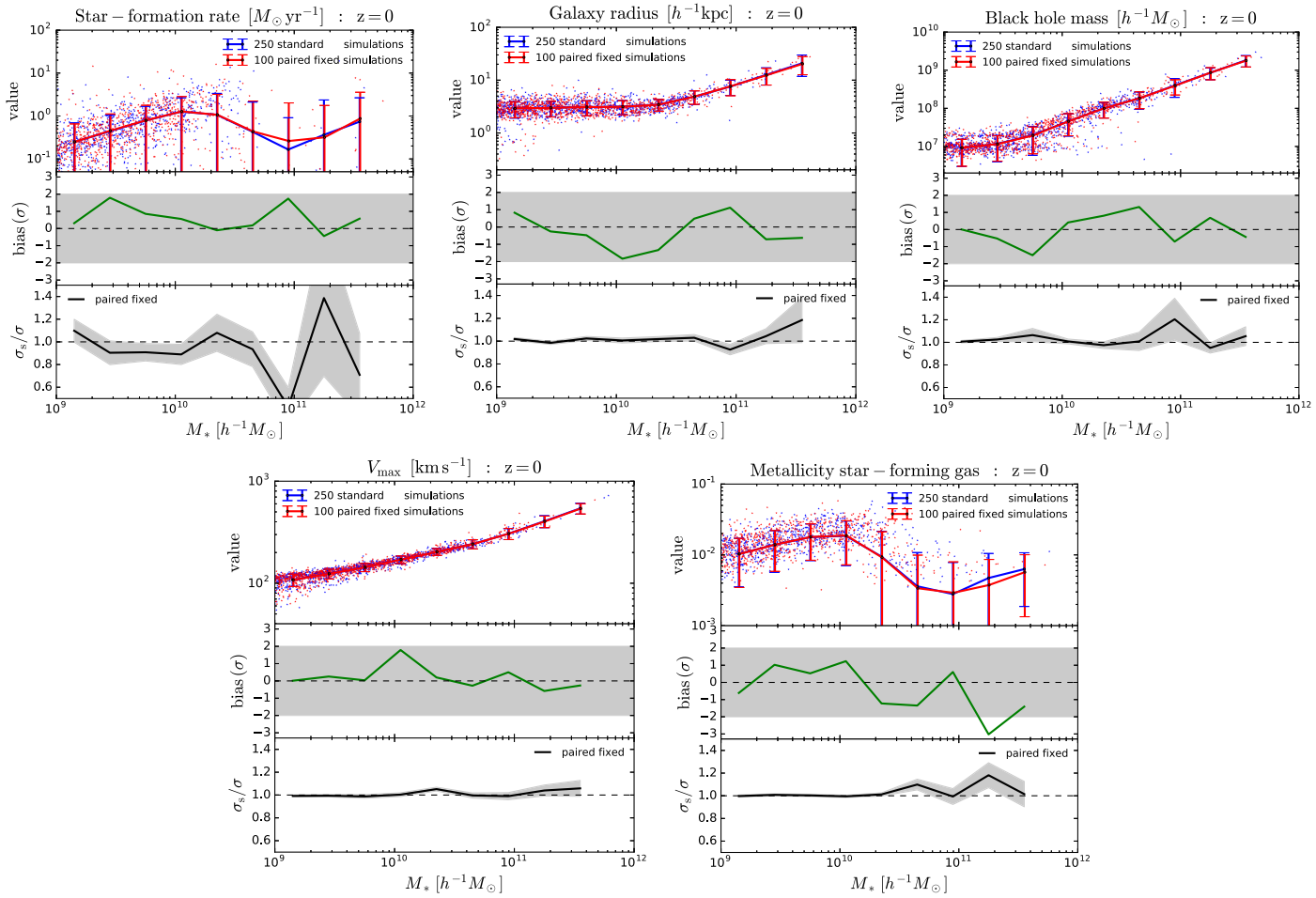

Figure 10. Impact of paired fixed simulations on internal galaxy properties. We show star formation rate (top left), radius (top middle), black hole mass (top right), maximum circular velocity (bottom left), and metallicity of star-forming gas (bottom right) as a function of stellar mass. Paired fixed simulations do not reduce the intrinsic physical scatter in these quantities, and they do not introduce a bias on them. Galaxies in paired fixed simulations look thus completely normal.

We find that fixed simulations barely improve the statistics of standard simulations, although there exist two significant bumps at redshifts $\simeq$4 and $\simeq$9. The improvement is slightly larger in paired fixed simulations, due to the values of the cross-correlation coefficient being less than 1 at some redshifts.

6.4. Galaxy Properties

The above results point out that, at least for clustering-related quantities, paired fixed or fixed simulations can improve the statistics of standard simulations without introducing a bias on the results. Thus, state-of-the-art cosmological hydrodynamic simulations such as ILLUSTRISTNG (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich et al. 2018a; Springel et al. 2018), EAGLE (Schaye et al. 2015), HORIZONAGN (Dubois et al. 2014), MAGNETICUM (Dolag et al. 2017), or BLUETIDES (Feng et al. 2016b) will highly benefit, for clustering analysis, by generating their initial conditions through fixed or paired fixed fields rather than standard Gaussian fields.

On the other hand, the main analysis scope of the above simulations is usually not clustering, but rather galaxy properties and evolution. It is thus very important to investigate (1) whether paired fixed simulations introduce a bias in internal galaxy properties and (2) whether the intrinsic physical scatter in their properties is changed in paired fixed simulations. The purpose of this section is to answer these two questions.

For each galaxy in each realization of the standard and paired fixed simulations, we have computed a number of different internal quantities: stellar mass, star formation rate, radius, black hole mass, maximum circular velocity, and metallicity of star-forming gas. We limit our analysis to well-resolved galaxies, which we define as those with a stellar mass above $10^9\,h^{-1}\,M_\odot$. We then make a scatter plot between the above quantities and stellar mass from the results of both simulation types. Finally, we take narrow bins in stellar mass and compute the mean and standard deviation of the results for the considered quantity.

The above procedure is slightly different from the treatment we have been using for the paired fixed simulations. For all the quantities considered so far in this work, we have estimated the value for the paired fixed realization as the average between the results within each pair. Here, for each paired fixed realization we just create the scatter plot and compute mean and standard deviation values for all galaxies (in a mass bin) together, without separating first between each simulation in the pair. This is because there is no way to pair individual objects for taking an average; indeed, individual halos become voids in their paired partner (Pontzen et al. 2016).

We show the results of this analysis in Figure 10. From the first panels we can see that the agreement between the results of both simulation types is very good, as in all the other quantities

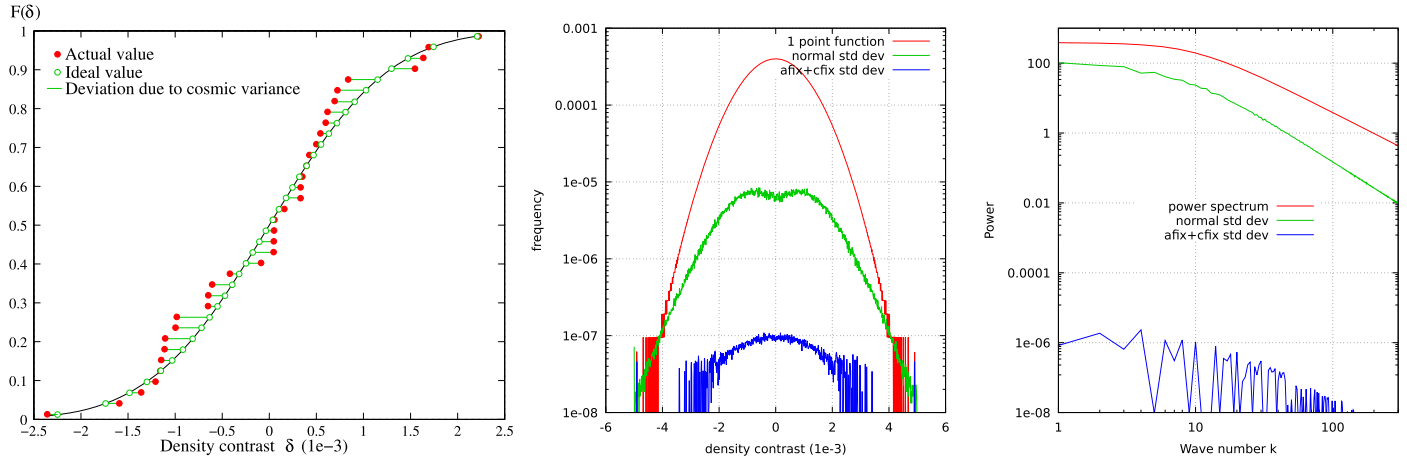

Figure 11. Middle and right: approximate CDF fixing and amplitude fixing can be applied at the same time by iteratively CDF fixing and amplitude fixing the same field. This is shown here for a simple Gaussian field with a truncated power-law power spectrum. The PDF (middle) and power spectrum (right) are shown in red, with corresponding standard deviations shown in green (before fixing) and blue (after fixing). Joint fixing results in a large reduction in the standard deviation of both statistics, with the largest improvement in the one that was last in the iteration scheme, in this case the power spectrum.

considered in this work. From the second panels we can see that paired fixed simulations do not introduce a bias on any of the studied internal galaxy properties. We have estimated the error on the difference of the means through Equation (14), but using the number of points in standard and paired fixed simulations in each bin as the value of N_s and N_{pf}, respectively.

The third panels show the ratio between the intrinsic scatter from each simulation type. Since the distribution of some of those properties is highly non-Gaussian, e.g., the distribution of star formation rates at fixed stellar mass, using Equation (18) with N_s and N_{pf} being the number of standard and paired fixed points in the scatter plot, respectively, will underestimate the errors on the ratio of the standard deviations. To avoid that, we have computed the errors on the ratio using bootstrap: for each studied quantity, we have created 15,000 bootstrap catalogs. For each catalog we have computed the ratio between the standard deviation of the standard and paired fixed simulations. Finally, we compute the standard deviation of the results from the previous step to get an estimate of the error on the standard deviation ratio from the whole sample. We create bootstrap catalogs by randomly subsampling, with replacement, the initial catalogs from the standard and paired fixed simulations.

The errors we obtain using this procedure are very similar to the ones we derive through Equation (18) for the radii, black hole mass, and maximum circular velocity versus stellar mass quantities, but very different for the star formation rate versus stellar mass.

We find that paired fixed simulations exhibit the same scatter on the considered quantities as standard simulations. In this case, this is precisely what we want, because the scatter on those quantities is due to internal physical processes and not to sample variance. We note, however, that we find a significantly lower scatter in the standard simulations for the star formation rate versus stellar mass of galaxies with stellar masses $\simeq 10^{11}\,h^{-1}\,M_\odot$. In that case, the ratio between the standard deviations is different from 1 at $\simeq 3.5\sigma$. Although the probability of having a point with such low standard deviation ratio is pretty low (under the assumption that the variance of standard and paired fixed is the same), we believe that it is not completely unreasonable given the large number of quantities considered. More simulations are, however, needed to disentangle whether this is a statistical fluctuation or pointing toward an increase in the scatter in paired fixed simulations.

We thus conclude that galaxies in paired fixed simulations look very much like those in standard simulations. We find no evidence that paired fixed simulations introduce a bias, and they do not reduce the internal physical scatter on their internal properties.

7. Improving One-point Statistics

So far we have seen that while paired fixed simulations can greatly reduce the sample variance in the power spectra, they have little to no effect on one-point statistics like the matter density pdf. The fact that amplitude fixing only works for the power spectra is not that surprising, since that procedure was designed to carefully tune complex amplitudes in Fourier space while letting the phases stay random. However, this clean separation between amplitudes and phases only exists in Fourier space. Other bases can be expressed as combinations of many different Fourier modes, and as we have seen (Figure 4), mixing the amplitudes of fixed modes undoes the fixing.

However, just like amplitude fixing is an operation designed to minimize sample variance in the power spectrum, we could construct different operations to minimize sample variance in other observables. For example, we could optimize for low sample variance in the initial one-point function of the density field by replacing the value in each grid cell with the value from the theoretical initial one-point function at that cell's quantile: if there are a total of n cells in the initial mass field, then the value of the cell with the kth largest value (counting from 0) would be replaced by the theoretical cumulative distribution's $[(2k+1)/(2n)]$th quantile (see the left panel of Figure 11).

We can call this operation *CDF fixing*, and it does eliminate the sample variance in the pdf of the density field in the initial conditions. Furthermore, since amplitude fixing and CDF fixing are defined in very different spaces, it turns out to be possible to perform both at the same time to high accuracy. A simple algorithm that achieves this is to iterate between fixing amplitudes in Fourier space and fixing the CDF in real space (see the middle and right panels of Figure 11).

In the same way that amplitude fixing for the power spectrum works as long as Fourier modes do not mix, CDF fixing works as long as real-space cells do not mix. Both

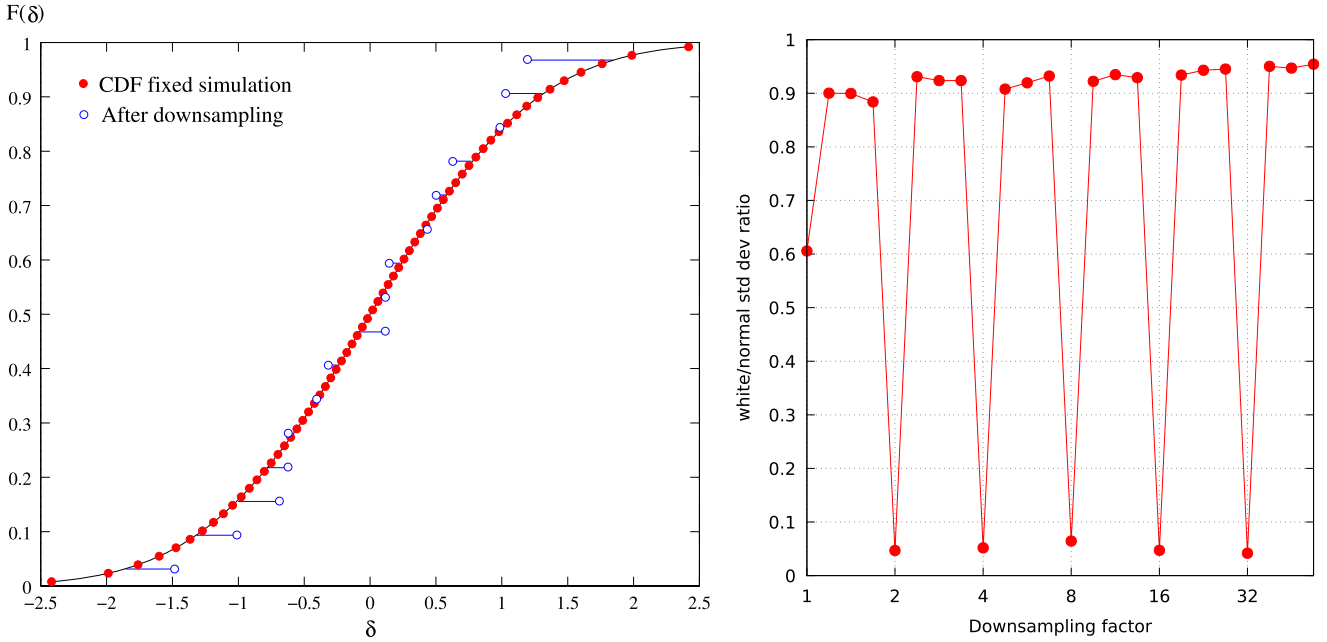

Figure 12. Left: CDF fixing is a nonlocal operation that only applies to a specific basis. This is an example of how a sample-variance-free CDF-fixed simulation completely loses its whiteness after being downsampled to half its original resolution. An 8×8 pixel white Gaussian field was simulated and CDF fixed, resulting in the sample-variance-free CDF shown in red. It was then downsampled to 4×4 via simple nearest-neighbor averaging. The blue points show the CDF of the downsampled field, and the horizontal blue lines show their deviation from the theoretical CDF. No trace of the fixing remains. Any operation or change in basis that mixes voxel values will have this effect, making CDF fixing very fragile. Right: a field can be approximately CDF fixed on multiple length scales at the same time by iteratively CDF fixing each scale. Here a 1024×1024 density field was CDF fixed at full resolution (downsampling factor of 1) and five power-of-two reductions in resolution (2, 4, 8, 16, 32). The variance was then measured on these scales, as well as several intermediate scales. Donwsampling was done using Fourier space truncation. The scales that were explicitly CDF fixed have reduced standard deviation, but almost no benefit is seen at any other length scale.

conditions are fulfilled under linear evolution, but once nonlinear effects appear, CDF fixing breaks down much more quickly than amplitude fixing. This happens because while nonlinear effects are relatively localized in Fourier space (they are most important at small scales), they occur practically everywhere in real space. Soon after nonlinear effects become important, all cells would start mixing, and the careful tuning of quantiles needed to cancel sample variance in the pdf would be lost.

Moreover, the reduction of the one-point function sample variance only happens for the exact set of cells it was defined for. Changing the resolution, or even just applying a noninteger displacement in position to the cells, will lead to destructive mixing. For example, if we optimize the pdf at a given grid size but measure it after downsampling to half resolution, the pixel mixing inherent in this operation completely destroys the sample variance cancellation. This is shown in Figure 12.

For the halo and void radius functions the problem is even worse, as the location and size of each halo and void are not known at the outset, preventing us from tuning the volumes that will end up as halos or voids to have reduced variance. And as we have seen, the tuning needs to exactly match the position and size of the objects we care about for there to be any effect. For example, simply tuning the pdf will not help, as each halo and void is a combination of multiple cells.

We conclude that while we can generate initial conditions with highly suppressed sample variance in the power spectrum, the corresponding operation for the matter density pdf is too fragile for practical use and would not survive even a small amount of nonlinearities and mode mixing. Even if such an operation were possible, we believe that it would not improve any other one-point statistics like the halo and void radius functions owing to the locality and nonlinearity involved in the formation of those objects and the highly nonlinear mode mixing involved thereby.

8. Discussion and Conclusions

Numerical simulations are an invaluable tool for understanding a large variety of processes such as the nonlinear growth of matter perturbations, the abundance of halos, and the formation and evolution of galaxies. The most powerful way to extract information from cosmological surveys will be to contrast observations with theoretical predictions from simulations.

The initial conditions of cosmological simulations are usually generated from Gaussian fields. The reason is that cosmic microwave background observations have shown that the temperature fluctuations in the early universe can be very accurately described by Gaussian fields (Planck Collaboration et al. 2016a, 2016b), whose properties are completely determined by their power spectra. The Fourier modes of a Gaussian field can be written as $\delta(\boldsymbol{k}) = Ae^{i\theta}$, where A follows the Rayleigh distribution of Equation (5) and θ is a random variable with a uniform distribution between 0 and 2π.

Running simulations with initial conditions generated from Gaussian fields gives rise to *sample variance*, i.e., statistical fluctuations arising from the fact that the mode distribution is not fully sampled. That problem is particularly important on scales approaching the box size, where only a few modes are sampled by simulations. To evaluate the likelihood and compute posteriors, the theoretical prediction should be free of statistical fluctuations. For this reason, many simulations are needed to beat down sample variance.

Fixed fields (see, e.g., Viel et al. 2010) are those with $\delta(\boldsymbol{k}) = Ae^{i\theta}$, where A takes a fixed value as specified by Equation (8) and θ is a random variable with a uniform distribution between 0 and 2π. The properties of those fields are that they share the same power spectrum of Gaussian fields but they do not exhibit any scatter around it.

Paired fixed fields consist of two fixed fields $\delta_1(\boldsymbol{k}) = Ae^{i\theta}$, $\delta_2(\boldsymbol{k}) = Ae^{i(\theta+\pi)} = -\delta_1(\boldsymbol{k})$, where the values of A and θ are the same in both. In Angulo & Pontzen (2016) it was shown that if simulations are run with initial conditions generated from those fields, large reductions on the sample variance amplitude of several important quantities can be achieved. The fixing serves to prevent sample variance in the linear amplitudes, while the pairing allows us to cancel some of the leading-order effects of phase correlations on nonlinear evolution in a finite box.

In this work we have further explored the properties of paired fixed fields by quantifying (1) the sample variance reduction achieved and (2) the bias introduced by paired fixed simulations with respect to standard simulations. We have carried out our analysis by using a large set of N-body (600) and state-of-the-art magnetohydrodynamic (506) simulations. Our simulations cover a wide range of scales, as well as mass and spatial resolutions, hence allowing us to investigate the statistical properties of paired fixed simulations in many different setups. We have analyzed the impact of paired fixed simulations in many different quantities: matter, CDM, gas, stars, galaxies, magnetic field, halo, and halo–matter power spectra, matter density field pdf's, void radius function, halo mass function, star formation rate history, stellar mass function, and internal galaxy properties such as black hole mass or galactic radii.

We now enumerate the main conclusions of this work. A summary the impact of paired fixed simulations on different statistics is shown in Table 2.

1. We find that paired fixed fields do not introduce a bias, with respect to standard Gaussian fields, on any of the quantities we have investigated in this paper. This is not an absolute statement. It may be that paired fixed simulations introduce a bias, but its magnitude has to be small since we do not find it with our rather large simulation set.
2. Paired fixed simulations reduce the scatter on the power spectrum of matter, halos, halo–matter, CDM, gas, stars, galaxies, and magnetic fields. The scatter reduction depends primarily on scale, with the variance on large scales being much more suppressed than on small scales.
3. Paired fixed simulations do not reduce the scatter on the halo bias. The linear-order explanation is that the variance on the halo bias is due to the amplitude of the shot noise, which is the same in standard and paired fixed simulations.
4. For large box sizes paired fixed simulations do not reduce the variance of the matter density pdf or the halo and void radius functions. For the matter density pdf, we find no improvement already in the initial conditions. Pairing has no effect because it simply mirrors the pdf around $\delta = 0$, and there is no special connection between points with values $\delta = +a$ and $\delta = -a$ in a simulation. Amplitude fixing has no effect either on the pdf since it is defined in real space, where the Fourier amplitudes and phases are scrambled.
5. For small boxes we find a small, statistically significant improvement on the matter density pdf and the halo mass function, but not on the void radius function. This may be due to a small reduction of the sample variance amplitude on the matter density pdf of those simulations that is already present in the initial conditions.
6. We find that paired fixed simulations do not reduce the scatter on the stellar mass function, while they seem to marginally improve it on the star formation rate history. We think that this follows as a result of the locality of the relevant physics.
7. Galaxies in paired fixed simulations look completely normal. We do not find any bias among the several intrinsic quantities, such as radii, black hole mass, star formation rate, metallicity, maximum circular velocity, and stellar mass, that we have investigated. The intrinsic, physical scatter on those quantities is not reduced by paired fixed simulations.
8. We have shown that procedures aiming at fixing the matter density pdf in the initial conditions are very fragile, and it seems almost impossible to fix the pdf on all possible scales. We thus conclude that it is unlikely that general operations performed in the initial conditions can be used to reduce the sample variance associated with statistics like the matter density pdf or the halo and void radius functions.

From the above results we can derive two further conclusions. First, let us recall that the values of parts of the trispectrum and higher-order moments are expected to be different in standard and fixed simulations. But we do not see any biases in any of our measured quantities. This suggests that the perturbation theory argument of Angulo & Pontzen (2016) —that the modifications do not propagate to observables except in a very specific, measure-zero subset—seems to hold even in highly nonlinear regimes.

The second conclusion is that, since paired fixed simulations help in reducing the scatter of clustering-related quantities while they do not improve the statistics of one-point quantities (or improve them marginally), we believe that the two quantities cannot be very correlated. If they were, we would have expected that as we reduce the scatter in one, the other should also be affected by it. We thus believe that the information embedded in clustering and one-point statistics should be highly complementary. While this is not surprising (see, e.g., Schaan et al. 2014), our conclusions arise from a completely different methodology than more traditional methods.

This paper constitutes an empirical confirmation of the benefits brought about by paired fixed simulations. We believe that upcoming large box size hydrodynamic simulations can highly benefit by being run with initial conditions from paired fixed fields. Since a paired fixed simulation requires running two simulations, the computational cost doubles with respect to a standard simulation. If that is computationally too expensive, one can run a single fixed simulation, which will bring a large fraction of the paired fixed simulation benefits at no extra cost.

We thank the referee, Pierluigi Monaco, for his very constructive report that has helped us to improve the quality of this work. This work has made extensive use of the python PYLIANS libraries, publicly available at https://github.com/franciscovillaescusa/Pylians. The simulations have been run in

Table 2
Summary of the Main Findings of This Work

Simulation Set	N_s	N_{pf}	Statistics	Redshift	Bias?	Max(σ_s/σ_f)	Max(σ_s/σ_{pf})	Corresponding Figures
N1000	100	100	$P_{mm}(k)$	99	no	1,169.0	136,821.0	2
				5	no	176.9	1,464.1	3
				1	no	80.0	1,472.4	3
				0	no	49.6	1,187.4	3
			$P_{hm}(k)$	0	no	6.3	7.3	3
			$P_{hh}(k)$	0	no	3.2	3.5	3
			$P_{hm}(k)/P_{mm}(k)$	0	no	1.3	1.6	5
			Matter pdf	99	no	1.4	1.4	2
				0	no	1.3	1.3	5
			Halo mass function	0	no	1.3	1.3	5
			Void radius function	0	no	1.2	1.3	5
H200	30	25	$P_{mm}(k)$	5	no	31.2	771.9	6
				1	no	10.8	87.1	6
				0	no	6.9	34.0	6
			$P_{cc}(k)$	5	no?	31.3	761.0	6
				1	no?	10.9	92.0	6
				0	no	7.0	34.3	6
			$P_{gg}(k)$	5	no?	37.2	426.3	6
				1	no	16.0	59.0	6
				0	no	8.8	29.3	6
H20	250	100	$P_{mm}(k)$	5	no	5.9	25.4	8
				1	no	2.3	4.1	8
				0	no	1.8	2.6	8
			$P_{gg}(k)$	5	no	6.3	23.9	8
				1	no	2.4	4.3	8
				0	no	1.8	2.7	8
			$P_{ss}(k)$	0	no	1.8	2.1	8
			$P_{gal}(k)$	0	no	1.8	2.3	8
			$P_B(k)$	0	no	1.6	2.1	8
			$P_{hm}(k)$	0	no	1.9	2.5	8
			Halo mass function	0	no	1.7	1.8	9
			Void radius function	0	no	1.4	1.6	9
			Matter pdf	99	no	1.5	1.6	7
				0	no	1.6	2.2	9
			Stellar mass function	0	no	1.2	1.2	9
			Star formation rate history	[0–15]	no	1.5	1.7	9
			Star formation rate versus stellar mass	0	no	⋯	1.4	10
			Radius versus stellar mass	0	no	⋯	1.2	10
			Black hole mass versus stellar mass	0	no	⋯	1.2	10

Table 2
(Continued)

Simulation Set	N_s	N_{pf}	Statistics	Redshift	Bias?	Max(σ_s/σ_f)	Max(σ_s/σ_{pf})	Corresponding Figures
			V_{max} versus stellar mass	0	no	⋯	1.1	10
			Metallicity versus stellar mass	0	no	⋯	1.2	10

Note. The first column indicates the simulation set used to carry out the analysis. The first letter indicates whether it is from *N*-body (N) or hydrodynamic (N) simulations, while the following number represents the simulation box size in h^{-1} Mpc. The second and third columns show the numbers of standard and pairs of fixed simulations composing each set, respectively. The fourth column represents the statistic considered, and its redshift is shown in the fifth column. The sixth column indicates whether we find that paired fixed simulations introduce a bias on the considered quantity with respect to standard simulations. The maximum reduction on the standard deviation from standard simulations achieved by fixed and paired fixed simulations is shown in the seventh and eighth columns, respectively. The relative error on those values is given by $0.5\sqrt{2/N_s + 2/N_{pf}}$. For galaxy properties (last five rows) we estimated the errors through bootstrap, finding that paired fixed simulations do not reduce the intrinsic scatter on galaxy properties. Finally, the ninth column shows the corresponding figure where we plot our results for the considered quantity. We note that for the matter density pdf we find larger reductions on the scatter of standard simulations, but we do not quote them, as they are due to statistical fluctuations.

the Gordon cluster at the San Diego Supercomputer Center. The work of F.V.-N., S.N., S.G., L.A., N.B., and D.N.S. is supported by the Simons Foundation. A.P. is funded by the Royal Society. This work was partially enabled by funding from the UCL Cosmoparticle Initiative.

Appendix A
Variance of Paired Fixed Simulations

In this appendix we derive Equation (15) and discuss the different origins of the statistical improvement of paired fixed simulations over traditional simulations for any generic quantity.

Suppose we are considering a quantity, X_s, e.g., the amplitude of the power spectrum at a given wavenumber k, or the halo mass function at mass M, from standard simulations, with a variance given by

$$\sigma_s^2 = \langle (X_s - \bar{X}_s)^2 \rangle = \langle X_s^2 \rangle - \bar{X}_s^2, \tag{21}$$

where $\bar{X}_s = \langle X_s \rangle$. Now consider the same quantity but estimated through the paired fixed simulations

$$X_{pf} = \frac{1}{2}(X_{pf,1} + X_{pf,2}), \tag{22}$$

where $X_{pf,1}$ and $X_{pf,2}$ are the values of X from the two pairs of a paired fixed simulation. The variance of X_{pf} is given by

$$\sigma_{pf}^2 = \langle (X_{pf} - \bar{X}_{pf})^2 \rangle = \langle X_{pf}^2 \rangle - \bar{X}_{pf}^2 \tag{23}$$

$$= \frac{1}{4}(\langle X_{pf,1}^2 \rangle + \langle X_{pf,2}^2 \rangle + 2\langle X_{pf,1} X_{pf,2} \rangle - \bar{X}_{pf,1}^2 - \bar{X}_{pf,2}^2 - 2\bar{X}_{pf,1}\bar{X}_{pf,2}) \tag{24}$$

$$= \frac{1}{4}(\sigma_{pf,1}^2 + \sigma_{pf,2}^2 + 2\mathrm{cov}_{12}), \tag{25}$$

where $\mathrm{cov}_{12} = \langle (X_{pf,1} - \bar{X}_{pf,1})(X_{pf,2} - \bar{X}_{pf,2}) \rangle$. Finally, since the variance of the two pairs from the paired fixed simulations is the same, $\sigma_{pf,1} = \sigma_{pf,2} = \sigma_f$, we obtain

$$\sigma_{pf}^2 = \sigma_f^2 \left(\frac{1 + r}{2} \right), \tag{26}$$

where the cross-correlation coefficient r is defined as $r = \mathrm{cov}_{12}/\sigma_f^2$, and it satisfies $-1 \leqslant r \leqslant 1$. We note that the variance of each individual pair within paired fixed simulations is, by definition, equivalent to the variance of individual fixed simulations. This is why we write σ_f above. It is interesting to consider some limiting situations:

1. The two sets of simulations of paired fixed simulations are independent, $r = 0$, and their variance is the same as in standard simulations, $\sigma_s = \sigma_f$. In this case, statistical improvement of the paired fixed Gaussian simulations will be just $\sigma_{pf} = \sigma_s/\sqrt{2}$. In this situation, the variance reduction arises simply because in the paired fixed simulations the quantity considered is estimated using two independent realizations instead of one.
2. The two sets of simulations of paired fixed simulations are completely correlated, $r = 1$, and the variance of each set is the same as in standard simulations, $\sigma_s = \sigma_f$. In this case no improvement is achieved by the paired fixed simulations: $\sigma_{pf} = \sigma_s$. This corresponds to a situation where the two paired fixed simulations are equivalent to one, e.g., the second is the same as the first, and therefore no improvement can be achieved.
3. The two sets of simulations of paired fixed simulations are completely anticorrelated, $r = -1$. In this case, the variance of the paired fixed simulations will be 0, independently of the variance of each pair, σ_f. The interpretation of this situation is that since the two simulations in each pair are completely anticorrelated, if $X_{pf,1}$ increases its value $X_{pf,2}$ will decrease, such that $X_{pf,1} + X_{pf,2}$ will be kept constant.
4. The variance of each set of paired fixed simulations is lower than the variance of the standard simulations, $\sigma_f^2 < \sigma_s^2$. In this case, even if the two paired fixed simulations are completely correlated, there will be a statistical improvement. This happens simply because even if the two pairs are completely correlated, i.e., only one independent realization is available, its variance is lower than that of a standard simulation. We note that this case applies to fixed simulations as well.

From the above arguments we see that, in most situations, paired fixed simulations will perform better than standard simulations by a factor of at least $1/\sqrt{2}$. This arises because each paired fixed realization contains two simulations while fixed or standard does only one. In order to avoid that artificial improvement, and to be able to carry out a fair comparison, in this paper we work with the *normalized variance*, defined as

the variance per number of simulations. In that case, we can express the normalized variance of paired fixed simulations as

$$\sigma^2_{\rm pf} = \sigma^2_{\rm f}(1+r). \tag{27}$$

This is the expression we have used along the text. It is interesting to relate the different pieces of Equation (27) to the properties of the paired fixed fields. On the one hand, a fixed field is expected to have different variance from a standard Gaussian field. Thus, the improvement of the fixed fields will arise from $\sigma_{\rm f}$ in Equation (27). On the other hand, the two simulations in a pair, independently of whether they are from pair simulations or paired fixed simulations, will contribute to the variance through r. We, however, emphasize that the value of r will, in general, be different for paired and paired fixed simulations. Thus, the correct interpretation of the $(1+r)$ factor is the statistical improvement brought by pairing (for paired simulations) or by pairing once the amplitude is fixed (for paired fixed simulations).

In other words, for paired simulations $\sigma_{\rm f} = \sigma_{\rm s}$ and any statistical improvement arises solely from r. For fixed simulations the statistical improvement comes through $\sigma_{\rm f}$, while for paired fixed simulations the improvement comes from both, by fixing the amplitude through $\sigma_{\rm f}$ and by pairing, once fixed, through the value of r.

Appendix B
Variance of the Ratio

Here we derive the expression we use to compute the variance of the ratio of two quantities. In general, given two random variables X and Y, the distribution of their ratio $Z = Y/X$ cannot be expressed analytically. We now derive a well-known expression for the variance of the ratio, making the assumption that the variances of both X and Y are smaller than their mean values. Given two random variables, X and Y, with means and variances given by

$$\bar{X} = \langle X\rangle \qquad \sigma^2_x = \langle (X-\bar{X})^2\rangle$$
$$\bar{Y} = \langle Y\rangle \qquad \sigma^2_y = \langle (Y-\bar{Y})^2\rangle,$$

we can Taylor-expand any function of them, $Z = f(X,Y)$, around the mean as

$$Z \simeq f(\bar{X},\bar{Y}) + \frac{\partial f}{\partial X}\bigg|_{\bar{X},\bar{Y}}(X-\bar{X}) + \frac{\partial f}{\partial Y}\bigg|_{\bar{X},\bar{Y}}(Y-\bar{Y}) + ... \tag{28}$$

At leading order, the mean of Z will be given by $\bar{Z} = f(\bar{X},\bar{Y})$, while its variance

$$\sigma^2_z = \langle (Z-\bar{Z})^2\rangle$$
$$\simeq \left\langle \left(\frac{\partial f}{\partial X}\bigg|_{\bar{X},\bar{Y}}(X-\bar{X}) + \frac{\partial f}{\partial Y}\bigg|_{\bar{X},\bar{Y}}(Y-\bar{Y})\right)^2\right\rangle + \ldots \tag{29}$$

$$= \left(\frac{\partial f}{\partial X}\bigg|_{\bar{X},\bar{Y}}\right)^2\sigma^2_x + \left(\frac{\partial f}{\partial Y}\bigg|_{\bar{X},\bar{Y}}\right)^2\sigma^2_y + 2\left(\frac{\partial f}{\partial X}\bigg|_{\bar{X},\bar{Y}}\right)\left(\frac{\partial f}{\partial Y}\bigg|_{\bar{X},\bar{Y}}\right)\sigma_x\sigma_y r + \ldots, \tag{30}$$

where r is the cross-correlation coefficient between X and Y. In the case where $Z = Y/X$ we obtain

$$\sigma^2_z \simeq \frac{\bar{Y}^2}{\bar{X}^4}\sigma^2_x + \frac{\sigma^2_y}{\bar{X}^2} - 2\frac{\bar{Y}}{\bar{X}^3}\sigma_x\sigma_y r. \tag{31}$$

We can finally express the above quantity as

$$\sigma^2_z \simeq \frac{\bar{Y}^2}{\bar{X}^2}\left(\frac{\sigma^2_x}{\bar{X}^2} + \frac{\sigma^2_y}{\bar{Y}^2} - 2r\frac{\sigma_x}{\bar{X}}\frac{\sigma_y}{\bar{Y}}\right). \tag{32}$$

For $b = P_{\rm hm}(k)/P_{\rm mm}(k)$ the above expression reduces to

$$\sigma^2_{\rm b} \simeq \frac{1}{2N(k)}\left(\frac{P_{\rm hh}(k)P_{\rm mm}(k) - P^2_{\rm hm}}{P^2_{\rm mm}}\right), \tag{33}$$

where we have used the fact that at linear order (see, e.g., Smith 2009)

$$\sigma^2(P_{\rm mm}(k)) = \frac{P^2_{\rm mm}(k)}{N(k)} \tag{34}$$

$$\sigma^2(P_{\rm hm}(k)) = \frac{P^2_{\rm hm}(k) + P_{\rm hh}(k)P_{\rm mm}(k)}{2N(k)} \tag{35}$$

$$r = P_{\rm hm}(k)P_{\rm mm}(k), \tag{36}$$

where $N(k)$ is the number of independent modes in the interval $[k, k+dk]$ where the different power spectra are measured and $P_{\rm hh}(k)$ is the halo power spectrum, which includes both the cosmological signal and the shot-noise term

$$P_{\rm hh}(k) = P^{\rm cosmo}_{\rm hh}(k) + \bar{n}^{-1}, \tag{37}$$

where $\bar{n}$ is the mean number density of halos.

We can also use Equation (32) to compute the error on the ratio between the standard deviation of standard and paired fixed simulations. Let us first compute the variance of $r^2 = \sigma^2_{\rm s}/\sigma^2_{\rm pf}$

$$\sigma^2_{r^2} = \frac{\sigma^4_{\rm s}}{\sigma^4_{\rm pf}}\left(\frac{\sigma^2_{\sigma^2_{\rm s}}}{\sigma^4_{\rm s}} + \frac{\sigma^2_{\sigma^2_{\rm pf}}}{\sigma^4_{\rm pf}}\right), \tag{38}$$

where $\sigma^2_{\sigma^2_{\rm s}}$ and $\sigma^2_{\sigma^2_{\rm pf}}$ denote the variance on the standard and paired fixed simulations, respectively. Under the assumption that data are Gaussian distributed, the quantity $\sum_{i=1}^{N}(X-\bar{X})^2$ follows a χ^2 distribution with N degrees of freedom. Thus, the variance of the variance is given by $\sigma^2_{\sigma^2} = 2\sigma^4/N$, and we obtain

$$\sigma^2_{r^2} = \frac{\sigma^4_{\rm s}}{\sigma^4_{\rm pf}}\left(\frac{2}{N_{\rm s}} + \frac{2}{N_{\rm pf}}\right). \tag{39}$$

We are, however, interested in the variance of the standard deviations, i.e., σ^2_r. By using the above Taylor expansion, we obtain $\sigma^2_{r^2} = 4r^2\sigma^2_r$; thus, the standard deviation of the standard deviation ratio is given by

$$\sigma_r = \frac{1}{2}\left(\frac{\sigma_{\rm s}}{\sigma_{\rm pf}}\right)\sqrt{\frac{2}{N_{\rm s}} + \frac{2}{N_{\rm pf}}}. \tag{40}$$

ORCID iDs

Shy Genel https://orcid.org/0000-0002-3185-1540
Andrew Pontzen https://orcid.org/0000-0001-9546-3849
Lauren Anderson https://orcid.org/0000-0001-5725-9329

References

Angulo, R. E., & Pontzen, A. 2016, MNRAS, 462, L1
Banerjee, A., & Dalal, N. 2016, JCAP, 11, 015
Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, PhR, 367, 1
Chuang, C.-H., Kitaura, F.-S., Prada, F., Zhao, C., & Yepes, G. 2015a, MNRAS, 446, 2621
Chuang, C.-H., Zhao, C., Prada, F., et al. 2015b, MNRAS, 452, 686
Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
Dolag, K., Mevius, E., & Remus, R.-S. 2017, Galax, 5, 35
Dubois, Y., Pichon, C., Welker, C., et al. 2014, MNRAS, 444, 1453
Feng, Y., Chu, M.-Y., & Seljak, U. 2016a, MNRAS, 463, 2273
Feng, Y., Di-Matteo, T., Croft, R. A., et al. 2016b, MNRAS, 455, 2778
Genel, S., Vogelsberger, M., Springel, V., et al. 2014, MNRAS, 445, 175
Heitmann, K., Higdon, D., White, M., et al. 2009, ApJ, 705, 156
Howlett, C., Manera, M., & Percival, W. J. 2015, A&C, 12, 109
Kitaura, F.-S., & Heß, S. 2013, MNRAS, 435, L78
Kitaura, F.-S., Yepes, G., & Prada, F. 2014, MNRAS, 439, L21
Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, MNRAS, 480, 5113
Monaco, P., Sefusatti, E., Borgani, S., et al. 2013, MNRAS, 433, 2389
Monaco, P., Theuns, T., & Taffoni, G. 2002a, MNRAS, 331, 587
Monaco, P., Theuns, T., Taffoni, G., et al. 2002b, ApJ, 564, 8
Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, MNRAS, 477, 1206
Nelson, D., Pillepich, A., Springel, V., et al. 2018, MNRAS, 475, 624
Palanque-Delabrouille, N., Yèche, C., Baur, J., et al. 2015, JCAP, 11, 011
Pillepich, A., Nelson, D., Hernquist, L., et al. 2018a, MNRAS, 475, 648
Pillepich, A., Springel, V., Nelson, D., et al. 2018b, MNRAS, 473, 4077
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016a, A&A, 594, A16
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, A&A, 594, A17
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016c, A&A, 594, A13
Pontzen, A., Slosar, A., Roth, N., & Peiris, H. V. 2016, PhRvD, 93, 103519
Rizzo, L. A., Villaescusa-Navarro, F., Monaco, P., et al. 2017, JCAP, 1, 008
Schaan, E., Takada, M., & Spergel, D. N. 2014, PhRvD, 90, 123523
Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
Scoccimarro, R., & Sheth, R. K. 2002, MNRAS, 329, 629
Smith, R. E. 2009, MNRAS, 400, 851
Springel, V. 2005, MNRAS, 364, 1105
Springel, V. 2010, MNRAS, 401, 791
Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676
Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726
Taffoni, G., Monaco, P., & Theuns, T. 2002, MNRAS, 333, 623
Tassev, S., Eisenstein, D. J., Wandelt, B. D., & Zaldarriaga, M. 2015, arXiv:1502.07751
Tassev, S., Zaldarriaga, M., & Eisenstein, D. J. 2013, JCAP, 6, 036
Torrey, P., Vogelsberger, M., Genel, S., et al. 2014, MNRAS, 438, 1985
Viel, M., Haehnelt, M. G., & Springel, V. 2010, JCAP, 6, 015
Vogelsberger, M., Genel, S., Sijacki, D., et al. 2013, MNRAS, 436, 3031
Vogelsberger, M., Genel, S., Springel, V., et al. 2014a, Natur, 509, 177
Vogelsberger, M., Genel, S., Springel, V., et al. 2014b, MNRAS, 444, 1518
Weinberger, R., Springel, V., Hernquist, L., et al. 2017, MNRAS, 465, 3291
Zennaro, M., Bel, J., Villaescusa-Navarro, F., et al. 2017, MNRAS, 466, 3244