



Cosmological Hydrodynamic Simulations with Suppressed Variance in the Ly α Forest Power Spectrum

Lauren Anderson¹ , Andrew Pontzen² , Andreu Font-Ribera² , Francisco Villaescusa-Navarro¹ , Keir K. Rogers³, and Shy Genel¹ 

¹ Flatiron Institute, 162 5th Ave., New York, NY 10010, USA; landerson@flatironinstitute.org

² University College London, Gower St., Kings Cross, London WC1E 6BT, UK

³ Oskar Klein Centre for Cosmoparticle Physics, Stockholm University, Stockholm SE-106 91, Sweden

Received 2018 October 31; accepted 2018 December 1; published 2019 January 28

Abstract

We test a method to reduce unwanted sample variance when predicting Ly α forest power spectra from cosmological hydrodynamical simulations. Sample variance arises due to sparse sampling of modes on large scales and propagates to small scales through nonlinear gravitational evolution. To tackle this, we generate initial conditions in which the density perturbation amplitudes are fixed to the ensemble average power spectrum—and are generated in pairs with exactly opposite phases. We run 50 such simulations (25 pairs) and compare their performance against 50 standard simulations by measuring the Ly α 1D and 3D power spectra at redshifts $z = 2, 3$, and 4. Both ensembles use periodic boxes of $40 h^{-1}\text{Mpc}$ containing 512^3 particles each of dark matter and gas. As a typical example of improvement, for wavenumbers $k = 0.25 h\text{Mpc}^{-1}$ at $z = 3$, we find estimates of the 1D and 3D power spectra converge 34 and 12 times faster in a paired–fixed ensemble compared with a standard ensemble. We conclude that, by reducing the computational time required to achieve fixed accuracy on predicted power spectra, the method frees up resources for exploration of varying thermal and cosmological parameters—ultimately allowing the improved precision and accuracy of statistical inference.

Key words: early universe – intergalactic medium – methods: numerical

1. Introduction

Precision cosmology with the Ly α forest is an area of increasing interest, both for its ability to constrain the evolution of the universe (McDonald & Eisenstein 2007; Busca et al. 2013; Slosar et al. 2013; Font-Ribera et al. 2014; Bautista et al. 2017; du Mas des Bourboux et al. 2017) and for its unique insight into the small-scale power spectrum that is a potential diagnostic of warm dark matter, neutrinos, and other particle physics phenomena (Seljak et al. 2005, 2006a, 2006b; Viel et al. 2006, 2013; Palanque-Delabrouille et al. 2015; Iršič et al. 2017). For the latter studies it is essential to explore the effect of both particle physics and astrophysical parameters on the measured power spectrum, since there are important degeneracies with the thermal and ionization history of the intergalactic medium that need to be considered (McDonald et al. 2005b; McQuinn 2016; Walther et al. 2018).

Ideally one runs a series of cosmological simulations to forward model the parameters into mock observables that can be compared with data (McDonald et al. 2005a; Viel & Haehnelt 2006; Borde et al. 2014). However, the mock observations suffer from intrinsic sample variance due to the finite simulation box size; sparse sampling of random phases and amplitudes of large-scale modes in the initial conditions (ICs) propagate to an uncertainty in the final mock power spectrum. This uncertainty can affect even small-scale modes through nonlinear gravitational coupling. In practice, model uncertainty is the limiting factor on large scales where it exceeds the observational uncertainty (Palanque-Delabrouille et al. 2015). Suppressing the uncertainties can be achieved either by expanding the box size, or by averaging many simulated realizations of the same volume size; either way a more precise mean power spectrum (or other observable) can be achieved. But each simulation is expensive, especially as the

box size is increased. To extract the best possible constraints, one has to tension this need for large sample sizes or volumes for a single point in parameter space against the requirement to adequately sample the parameter space itself.

This paper examines an approach to lessen the tension, by initializing simulations with carefully chosen amplitudes and phases for each mode, thus reducing the total volume or number of simulations required to achieve a given statistical accuracy on the mean mock power spectrum. This approach will free up computing time per parameter instantiation to increase the density of samples in parameter space. Methods, such as emulators (S. Bird et al. 2018, in preparation; K. Rogers et al. 2018, in preparation), which interpolate between hydrodynamic simulation predictions, can then be more accurate. Interpolation is necessary because only a limited number of simulations can feasibly be run.

We minimize the number of simulation realizations by running paired–fixed simulations (Angulo & Pontzen 2016; Pontzen et al. 2016). The difference between the standard simulations and the paired–fixed simulations is in the input density field. For a Gaussian random field, the amplitude of each mode is drawn from a Rayleigh distribution with zero mean and variance equal to the power spectrum at that scale, and the phase drawn with uniform probability between 0 and 2π . Instead, for paired–fixed simulations, we fix the amplitudes of the Fourier modes to the square root of the power spectrum, and invert the phase of one simulation in a pair relative to the other. The fixing reduces the variance of the power spectrum on the largest scales, which remain in or near the linear regime, while the pairing cancels leading-order errors due to nonlinear evolution of chance correlations between large-scale modes.

On its own, pairing is a benign manipulation in the sense that both realizations in the pair are legitimate, equally likely draws from a Gaussian random field (Pontzen et al. 2016). Fixing, on

the other hand, generates an ensemble that is quantitatively different from Gaussian; Angulo & Pontzen (2016) argued that, to any order in perturbation theory, the differences affect only the variance on the power spectrum and cannot propagate to other observables such as the power spectrum or observable measures of non-Gaussianity. Combined, pairing, and fixing thus formally removes the leading order and next-to-leading order uncertainties that generate scatter in predicted mock observations from simulations. In Villaescusa-Navarro et al. (2018), we thoroughly explored the statistics of paired–fixed simulations for both dark matter evolution and galaxy formation, finding that there are no statistically significant biases in the approach, and in many settings it does yield significantly enhanced statistical accuracy.

In this paper, we apply these techniques for the first time to the study of the power spectrum of fluctuations in the Ly α forest, as measured from hydrodynamical simulations. The Ly α forest is the collection of absorption features present in distant quasar spectra due to intervening neutral hydrogen absorbing Ly α photons as the quasar light redshifts. Physically, the forest comprises a relatively smooth, low-density environment of hydrogen that traces the underlying matter density of the universe well in the redshift range $2 < z < 5$ (Hernquist et al. 1996; Croft et al. 1998). At higher redshifts $z > 5$ and especially beyond reionization the neutral hydrogen is too optically thick (Gunn & Peterson 1965); at lower redshift $z < 2$ the forest cannot be observed with optical instruments. Although use of the Ly α forest as a tracer of the matter density is limited to a finite redshift interval, it is a unique probe of a wide range of otherwise inaccessible redshifts and scales. It has the unique advantage of tracing the matter density at small scales, down to 10s of kiloparsecs, in regions where the evolution is still close to linear.

This paper is structured as follows. In Section 2 we describe the simulations used in this analysis, before presenting results in Section 3. We conclude in Section 4 with a brief summary and discussion.

2. Simulations

In this section we describe the simulations used in this study. We start by describing the different sets of ICs (Gaussian, fixed and paired–fixed); we will then discuss the hydrodynamical simulations and the method to simulate Ly α forest sightlines from the simulations.

2.1. Gaussian, Fixed, and Paired–Fixed Random Fields

In this paper we will follow the naming convention of Villaescusa-Navarro et al. (2018), which contains more detailed definitions including power spectrum conventions. Here we present a brief summary of the required definitions to interpret our results.

Given a density field $\rho(\mathbf{x})$, the density contrast is defined as

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}, \quad (1)$$

where $\bar{\rho} = \langle \rho(\mathbf{x}) \rangle$. In a simulation we take discretized values of these quantities at the center of a cell \mathbf{x}_i , $\delta_i = \delta(\mathbf{x}_i)$, and define

discrete Fourier modes $\tilde{\delta}_n$ as

$$\tilde{\delta}_n = \sum_i \delta_i e^{i\mathbf{x}_i \cdot \mathbf{k}_n} = A_n e^{i\theta_n}, \quad (2)$$

where \mathbf{k}_n identifies the wavenumber, A_n is the amplitude of the discrete Fourier mode, and θ_n is the phase.

For a Gaussian density field, each θ_n is a random variable distributed uniformly between 0 and 2π and each A_n follows a Rayleigh distribution

$$p(A_n) dA_n = \frac{A_n}{\sigma_n^2} e^{-A_n^2/2\sigma_n^2} dA_n, \quad (3)$$

where $\langle A_n \rangle = \sqrt{\frac{\pi}{2}} \sigma_n$, and σ_n^2 is proportional to the power spectrum $P(|\mathbf{k}_n|)$ (with the constant of proportionality depending on the particular conventions adopted; see Villaescusa-Navarro et al. 2018).

For generating ICs representing a discretized Gaussian density field for a simulation of the universe, the mode amplitudes (A_n) and phases (θ_n) are chosen randomly from their distributions above. This randomness generates a source of variance in the final mock power spectrum from the simulation; the amplitudes A_n propagate directly, while the phases θ_n become significant during nonlinear evolution because their values determine how modes of different scales interact with each other.

In this paper, we explore alternative approaches to generating ICs such that these variance effects are minimized. Similar to Villaescusa-Navarro et al., we define Gaussian, paired Gaussian, fixed, and paired–fixed fields as follows:

1. *Gaussian field*: A field with $\tilde{\delta}_n = A_n e^{i\theta_n}$, where A_n follows the Rayleigh distribution of Equation (3) and θ_n is a random variable distributed uniformly between 0 and 2π .
2. *Paired Gaussian field*: A pair of Gaussian fields δ and δ' , where $\tilde{\delta}_n = A_n e^{i\theta_n}$ and $\tilde{\delta}'_n = A_n e^{i(\theta_n + \pi)} = -\tilde{\delta}_n$; the values of A_n and θ_n are the same for the two fields and A_n is drawn from the Rayleigh distribution of Equation (3). Because the leading-order uncertainties are typically from the amplitudes rather than the phases, we do not use paired fields without fixing the amplitudes in this work.
3. *Fixed field*: A field with $\tilde{\delta}_n = A_n e^{i\theta_n}$, where we fix $A_n = \sqrt{2} \sigma_n$. The phase θ_n remains a uniform random variable between 0 and 2π .
4. *Paired–fixed field*: A pair of fields, $\tilde{\delta}_n = A_n e^{i\theta_n}$ and $\tilde{\delta}'_n = A_n e^{i(\theta_n + \pi)} = -\tilde{\delta}_n$, where the values of A_n are the same for the two fields and fixed to $A_n = \sqrt{2} \sigma_n$, and the values of θ_n are also the same for the two fields, as defined, rendering $\tilde{\delta}_n$ and $\tilde{\delta}'_n$ exactly out of phase with respect to each other. Here θ_n remains a uniform random variable between 0 and 2π .

The purpose of this paper is to compare the statistical accuracy of paired–fixed simulations with standard simulations (based on Gaussian fields), in the context of Ly α forest power spectrum analyses.

2.2. Hydrodynamical Simulations

We perform hydrodynamical simulations with the GADGET-3 code, a modified version of the publicly available code GADGET-2 (Springel 2005). The code incorporates radiative cooling by primordial hydrogen and helium, alongside heating by the UV background following the model of Viel et al. (2013). The `quick-lya` flag is enabled, so gas particles with overdensities larger than 1000 and temperatures below 10^5 K are converted into collisionless particles (Viel et al. 2004). This approach focuses the computational effort on low-density gas associated with the Ly α forest, and is much more efficient than either allowing dense gas to accumulate or adopting a realistic galaxy formation and feedback model, to which the forest is largely insensitive. Thus, this computational efficiency has no significant effect on the accuracy of estimated Ly α forest observables.

The ICs are generated at $z = 99$ by displacing cold dark matter (CDM) and gas particles from their initial positions in a regular grid, according to the Zel'dovich approximation. We account for the different power spectra and growth factors of CDM and baryons by rescaling the $z = 0$ linear results according to the procedure described in Zennaro et al. (2017).

Our simulations evolve 512^3 CDM and 512^3 gas particles from $z = 99$ down to $z = 2$ within a periodic box of $40 h^{-1}$ Mpc, storing snapshots at redshifts 4, 3, and 2. We run a total of 100 simulations (25 fixed pairs, and 50 standard simulations). In order to study the dependence of our results on volume we also run an identical set of simulations as those described above but with 256^3 CDM and 256^3 gas particles in a box of size $20 h^{-1}$ Mpc.

The values of the cosmological parameters, the same for all simulations, are in good agreement with results from Planck Collaboration et al. (2016): $\Omega_m = 0.3175$, $\Omega_b = 0.049$, $\Omega_\nu = 0$, $\Omega_\Lambda = 0.6825$, $h = 0.67$, $n_s = 0.9624$, and $\sigma_8 = 0.834$.

2.3. Modeling Ly α Absorption

In order to study Ly α forest flux statistics we generate a set of mock spectra containing only the Ly α absorption line from neutral hydrogen for each snapshot in our simulation suite. The spectra are calculated on a square grid of 160,000 spectra for the 40 Mpc box. Each spectrum extends the full length of the box, and maintains the periodic boundary conditions of the underlying simulation, with sizes in velocity space of [3793, 4267, 4698] km s $^{-1}$ respectively at $z = [2, 3, 4]$. In each spectrum, the optical depth τ is calculated by taking the product of the column density of neutral hydrogen and the atomic absorption coefficient for the Ly α line (Humlicek 1979) appropriately redshifted for both its cosmological and peculiar velocities. Measurements are made in bins of velocity width 10 km s^{-1} using the code `fake_spectra` (Bird 2017).

Once we have computed the Ly α optical depth, we derive the transmitted flux fraction (or flux for short):

$$F = e^{-\tau}. \quad (4)$$

We then calculate \bar{F} , the mean flux over all pixels in our box. Next, we adopt the common practice of rescaling the optical depth of each simulation, $\tau \rightarrow \alpha\tau$ where α is chosen to fix \bar{F} to a target value. This is normally adopted when making comparisons to data because it reduces the impact of uncertainties in the UV background. In our case, we set α in each simulation such that \bar{F} is fixed to the mean over all simulations. We verified that, if we do not renormalize τ , the

benefits of paired and fixed simulations remain (in fact they even increase further on small scales). After renormalization, we calculate the fluctuations around the mean, δ_F :

$$\delta_F = \frac{F}{\bar{F}} - 1. \quad (5)$$

Finally, we calculate flux power spectra using the code `lyman_alpha` (Rogers et al. 2018a, 2018b). Studies of the Ly α forest either adopt 1D or 3D power spectra depending on the density of available observational skewers; the former approach is computationally simpler and captures small-scale line-of-sight information, while the latter approach is required to capture larger-scale correlations that do not fit within a single quasar spectrum. In this work we will study the effects of paired–fixed simulations in both settings.

To create a mock power spectrum, the first step is to generate the Fourier transform ($\tilde{\delta}_F$) of the fluctuation field δ_F . For 1D power spectra, the transformation is taken separately for each skewer along the line-of-sight direction; otherwise, we apply the 3D transformation to the entire box. In both cases we are able to exploit fast Fourier transforms since all velocity bins are of equal width and spectra are evenly sampled in the transverse direction. We then estimate power spectra by averaging $|\tilde{\delta}_F|^2$ over all skewers for the 1D power spectrum, or in $|k|$ shells for the 3D power spectrum.

When working with paired simulations, we average the power spectra to form a single estimate from each pair. This leaves us with 25 independent estimates from our ensemble of 50 paired–fixed simulations, a point that will be discussed further below.

3. Results

In this section we will present the statistical performance of our different simulations in predicting Ly α power spectra. To quantify the performance of any given ensemble, we calculate the mean of the measured power across all realizations, $\bar{P}(k)$

$$\bar{P}(k) = \frac{1}{N_{\text{est}}} \sum_{i=1}^{N_{\text{est}}} P_i(k), \quad (6)$$

where N_{est} is the number of independent estimates of $\bar{P}(k)$ and $P_i(k)$ are the individual estimates. We also calculate the single-estimate variance in $P(k)$:

$$\sigma^2(P(k)) = \frac{1}{N_{\text{est}} - 1} \sum_{i=1}^{N_{\text{est}}} (P_i(k) - \bar{P}(k))^2, \quad (7)$$

As noted above, in the case of paired (or paired–fixed) simulations $N_{\text{est}} = N_{\text{sim}}/2$, while in the case of a standard ensemble, $N_{\text{est}} = N_{\text{sim}}$ where in both cases $N_{\text{sim}} = 50$, the number of simulations. This difference arises because paired power spectrum estimates are not independent and instead the average of each pair is regarded as providing a single sample $P_i(k)$.

We are actually interested in the expected uncertainty on the mean power spectrum, $(\Delta\bar{P})^2$, estimated as

$$(\Delta\bar{P})^2 \equiv \frac{\sigma^2(P(k))}{N_{\text{est}}}. \quad (8)$$

The ratio of the standard to the paired–fixed uncertainties gives the most useful performance metric; in our case where

N_{sim} is the same between the two ensembles, this reduces to

$$\frac{(\Delta\bar{P}_S)^2}{(\Delta\bar{P}_{\text{PF}})^2} = \frac{\sigma_S^2(P(k))}{2\sigma_{\text{PF}}^2(P(k))}, \quad (9)$$

where S and PF subscripts refer to the standard (Gaussian) and paired–fixed ensembles respectively. This ratio can also be thought of as the computational speed-up for achieving a fixed target accuracy (assuming the uncertainty on the mean power spectrum scales proportional to $\sqrt{N_{\text{sim}}}$). It limits to unity when the paired–fixed approach is not providing any improvement.

From our paired–fixed simulations we can also form a fixed ensemble simply by discarding one of each pair. Thus we will also show results for a fixed ensemble with $N_{\text{est}} = N_{\text{sim}} = 25$. Once again the estimated accuracy of the mean power is given by Equation (9), including the factor 2 renormalization for the relative number of estimates.

To ensure that the manipulations of the ICs do not introduce any bias, we also calculate $\bar{P}_S - \bar{P}_{\text{PF}}$. For finite N_{sim} this difference contains a residual statistical uncertainty; one should expect

$$\langle(\bar{P}_S - \bar{P}_{\text{PF}})^2\rangle = (\Delta\bar{P}_S)^2 + (\Delta\bar{P}_{\text{PF}})^2, \quad (10)$$

given the independence of the two ensembles. Therefore, we check that $\bar{P}_S - \bar{P}_{\text{PF}}$ remains comparable in magnitude to this expected residual.

3.1. Matter Power Spectrum

We start by discussing the matter power spectrum (i.e., the 3D power spectrum of the overdensity field without any transformation to Ly α flux); Villaescusa-Navarro et al. (2018) also discussed this quantity, but not at the redshifts in which we are interested in this work. Our simulations also have different baryonic physics implementations.

In Figure 1 we show the matter power spectrum measured at $z = 2, 3,$ and 4 . In the top panel we show the average power spectrum in the two ensembles: the set of 50 standard simulations \bar{P}_S and the set of 25 paired–fixed simulations \bar{P}_{PF} . Error bars show the uncertainties $\Delta\bar{P}_S$ in the mean of the standard simulations, computed using Equation (8). The uncertainties are small due to the large number of simulations. The uncertainties in the paired–fixed simulation are too small to be plotted in the top panel, as we will shortly discuss.

The middle panel shows the difference between the mean power measured in the two sets of simulations, as a fraction of the total power in the standard simulations. Additionally, we show, as a gray band, the $\pm 1\sigma$ uncertainty at $z = 3^4$ defined by Equation (10). The two approaches to estimating the nonlinear power spectrum thus agree well, consistent with Angulo & Pontzen (2016) and Villaescusa-Navarro et al. (2018). In the bottom panel we quantify the statistical improvement achieved by the paired–fixed simulations with respect to the standard simulations. We show the ratios of the uncertainty on the means, comparing the standard and paired–fixed simulations, $(\Delta\bar{P}_S)^2/(\Delta\bar{P}_{\text{PF}})^2$, Equation (9).⁵ This ratio represents how quickly the paired–fixed simulations converge on the true mean relative to the standard simulations. We also show the ratio of

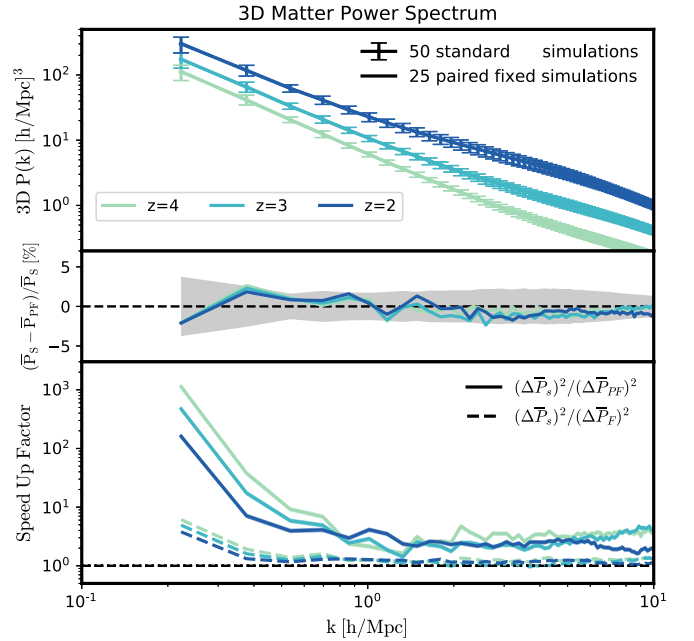


Figure 1. Matter power spectrum measured at $z = 2, 3,$ and 4 . The top panel shows the mean power spectrum in the two sets of simulations, the standard simulations, \bar{P}_S , with the 1σ uncertainty on the mean, and the paired–fixed simulations, \bar{P}_{PF} ; the means are indistinguishable. The middle panel shows the fractional difference between the mean power measured in the two sets of simulations, with a shaded region showing the expected 1σ uncertainty on that difference which is given by Equation (10). Across all scales, there is no evidence for bias in the estimate of the paired–fixed simulations relative to the traditional simulations. The bottom panel shows the ratio between the uncertainty on the mean of the standard and paired–fixed simulations, $(\Delta\bar{P}_S)^2/(\Delta\bar{P}_{\text{PF}})^2$, with dashed lines showing instead the effect of fixing without pairs $(\Delta\bar{P}_S)^2/(\Delta\bar{P}_F)^2$. These ratios summarize how fast the paired–fixed (or fixed) simulations converge on the mean power spectrum relative to the standard simulations. As a typical example of improvement, for wavenumbers $k = 0.25 h\text{Mpc}^{-1}$ and $2 h\text{Mpc}^{-1}$ at $z = 3$, we find estimates of the matter power spectra converge 390 and 2.0 times faster in the paired–fixed simulations compared with the standard simulations.

the uncertainty on the means of the standard simulations and fixed (but not paired) simulations, shown as the dashed lines. The dotted horizontal line shows a value of 1, indicating the level where fixed and paired–fixed simulations do not bring any statistical improvement over standard simulations. Comparing the solid and dashed lines shows that both fixing and pairing the simulations contributes to reducing uncertainty on the mean power spectrum at low k . As a typical example of improvement, for wavenumbers $k = 0.25 h\text{Mpc}^{-1}$ and $2 h\text{Mpc}^{-1}$ at $z = 3$, we find estimates of the matter power spectra converge 390 and 2 times faster in paired–fixed simulations compared with standard simulations.

Broadly this agrees with results from Villaescusa-Navarro et al. (2018) but, unlike in the earlier work, we find that the factor of $\simeq 2$ improvement in a paired–fixed simulation continues to $k = 10 h\text{Mpc}^{-1}$. To isolate the cause of this difference, we first considered the effect of box size. Our simulations are of boxes with side length $40 h\text{Mpc}^{-1}$, which is intermediate between boxes considered by Villaescusa-Navarro et al. (2018). We therefore ran a suite of smaller $20 h\text{Mpc}^{-1}$ simulations for direct comparison to the smallest boxes of that earlier work. However, we found similar results at large k in this additional suite, and conclude that box size is not the driver for the different behavior. The only remaining difference between our present simulations and the hydrodynamic

⁴ We find that our calculations of this expected scatter as a fraction of the mean power is almost independent of redshift, as the fractional uncertainties depend primarily just on the number of modes in each power spectrum bin.

⁵ Note that this is the square of the quantity shown in the bottom panels in Villaescusa-Navarro et al. (2018).

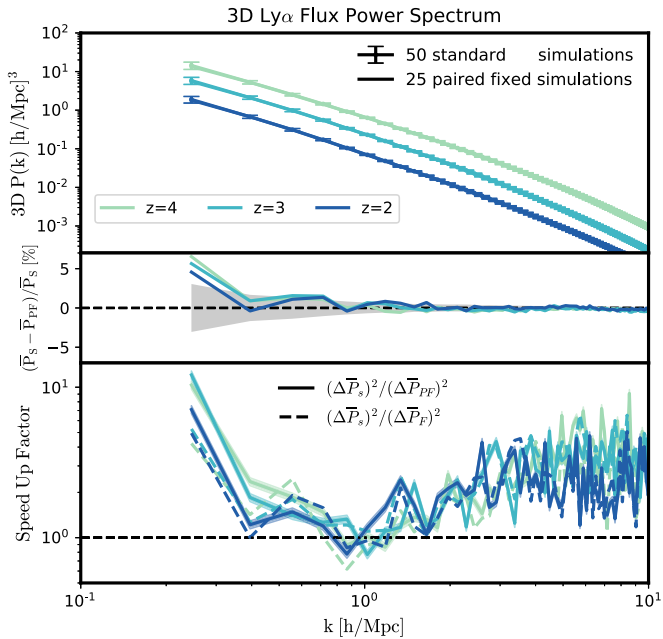


Figure 2. Same as Figure 1, but for the 3D Ly α flux power spectrum. As a typical example of improvement, for wavenumbers $k = 0.25 \text{ hMpc}^{-1}$ and 2 hMpc^{-1} at $z = 3$, we find estimates of the 3D Ly α flux power spectra converge 12 and 2.0 times faster in a paired-fixed ensemble compared with a standard ensemble.

simulations of Villaescusa-Navarro et al. (2018) lies in the drastically different baryonic physics implementation. As discussed in Section 2.2, we do not implement feedback but instead use an efficient prescription to convert high density gas into collisionless particles. By contrast, Villaescusa-Navarro et al. (2018) adopts a state-of-the-art star formation and feedback prescription that results in significant mass transport due to galactic outflows. We verified that the $20 \text{ h}^{-1}\text{Mpc}$ dark-matter-only runs of Villaescusa-Navarro et al. (2018), like our Ly α runs, also generate substantial improvements from pairing and fixing at high k . It therefore seems that potential efficiency gains at $k \gg 1 \text{ hMpc}^{-1}$ can be erased by stochastic noise from sufficiently energetic feedback and winds.

Our practical conclusions will be insensitive to this regime. Uncertainties at high k are dominated by observational rather than theoretical uncertainty, due to the rapidly shrinking absolute magnitude of simulation sample variance. Because the sample variance is intrinsically small, it is only important that the power spectrum remains unbiased at these scales.

3.2. 3D Ly α Power Spectrum

We next turn our attention to the 3D Ly α power spectrum. On large, quasi-linear scales, the 3D Ly α power spectrum is proportional to the matter power spectrum, with an amplitude set by the scale-independent bias parameter, and an angular dependence described by the Kaiser (1987) model of linear redshift-space distortions. In this work, we include these distortions but consider the angle-averaged (i.e., monopole) power spectrum rather than divide the power spectrum into angular bins.

We present the measured 3D Ly α forest power spectrum in Figure 2. The panels are computed and arranged as in Figure 1, but starting from the flux δ_F instead of the overdensity δ . In the top panel the uncertainties on the mean power spectrum of the

standard simulations are again small due to the large number of simulations. The middle panel shows the fractional difference between the paired-fixed and standard simulation power spectra, with the gray shaded region again indicating the expected 1σ scatter. Several measured points scatter outside this envelope (most notably at $k < 0.3 \text{ hMpc}^{-1}$) but this is consistent with a statistical fluctuation driven by scatter in the standard ensemble. At high k the difference in the mean is subpercent. Comparing the central panels of Figures 1 and 2 reveals that the shape of the 1σ contour is different between the Ly α and matter power cases. At high k , the matter power enters a highly nonlinear regime where modes are strongly coupled, whereas the flux arises from low-density clouds that are still quasi-linear. In the case of flux, modes are near-decoupled and the uncertainty decays proportionally to the increasing number of modes per k shell.

The bottom panel of Figure 2 quantifies the statistical improvement achieved by the paired-fixed simulations with respect to the standard simulations, shown as the solid line. The dashed line represents simulations that are just fixed. The dotted horizontal line shows a value of 1, indicating the level where fixed and paired-fixed simulations do not bring any statistical improvement over standard simulations. As a typical example of improvement, for wavenumbers $k = 0.25 \text{ hMpc}^{-1}$ and 2 hMpc^{-1} at $z = 3$, we find estimates of the 3D Ly α flux power spectra converge 12 and 2 times faster in a paired-fixed ensemble compared with a standard ensemble.

There is a minimum level of improvement at $k \simeq 1 \text{ hMpc}^{-1}$ where the paired-fixed approach does not outperform the standard ensemble at $z = 3$. The relative performance improvements then increase with increasing k beyond this point, which is surprising. Previously it has been argued that the improvements generated by the paired-fixed approach can be understood within the framework of standard perturbation theory (Angulo & Pontzen 2016; Pontzen et al. 2016), which more obviously applies to intermediate quasi-linear scales than the $k > 1 \text{ hMpc}^{-1}$ regime.

At present we cannot fully explain why pairing and fixing improves the Ly α power spectrum accuracy in this regime. It might be that super-sample covariance terms couple high k to low k in a unique way given that the forest is dominated by underdense regions (unlike the matter power spectrum; Takada & Hu 2013). However, we did not investigate further for this work because the improvements are not particularly relevant for observational studies; the absolute magnitude of the model uncertainties in the $k > 1 \text{ hMpc}^{-1}$ rapidly become smaller than observational uncertainties (see the top panel of Figure 2 for a visualization of the magnitude of the model uncertainties). Thus the improvements in accuracy at low k are of most practical benefit to future work.

3.3. 1D Ly α Power Spectrum

Most hydrodynamical simulations of the Ly α forest have been used to predict the 1D power spectrum (McDonald et al. 2005a; Viel & Haehnelt 2006; Borde et al. 2014; Lukić et al. 2015; Walther et al. 2018). As described in Section 2.3, the 1D spectrum captures small-scale information along the line of sight without the computational complexity of cross-correlating between different skewers.

We present the measured 1D Ly α power spectrum in Figure 3, using again the same panel layout as Figure 1. In the top panel the uncertainties on the mean power spectrum of the

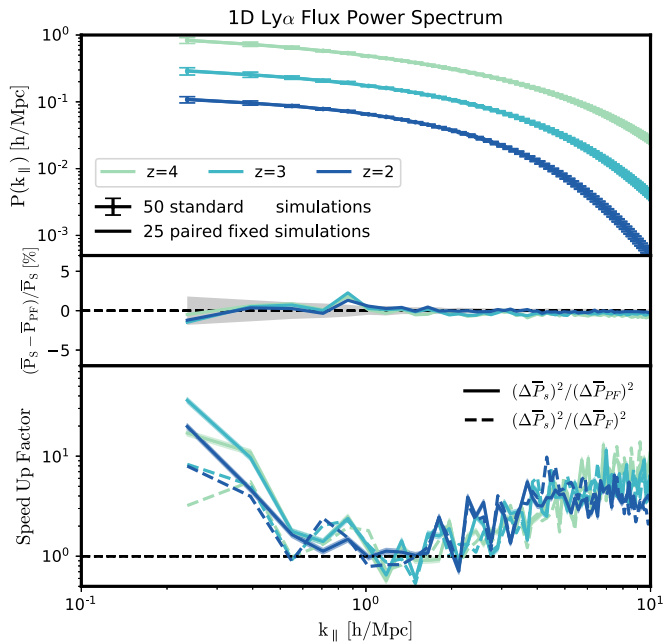


Figure 3. Same as Figure 1, but for the 1D Ly α flux power spectrum. The line-of-sight wavenumber k_{\parallel} corresponds to an integral over many 3D wavenumbers k , so there is not a direct correspondence between the scales. As a typical example of improvement, for wavenumbers $k_{\parallel} = 0.25 \text{ hMpc}^{-1}$ and 2 hMpc^{-1} at $z = 3$, we find estimates of the 1D Ly α flux power spectra converge 34 and 1.7 times faster in a paired-fixed ensemble compared with a standard ensemble.

standard simulations are, as usual, small due to the large number of simulations. The middle panel shows the fractional uncertainty in the paired-fixed simulations relative to the standard simulations, with the gray shaded region indicating the 1σ uncertainty on this difference. Once again there is overall good agreement, with the strongest deviation arising at $k = 0.9 \text{ hMpc}^{-1}$ at a significance of almost 2σ . However, the difference between our standard and paired-fixed ensemble remains less than 2%, and (given the large number of independent k bins) we believe the difference to be consistent with the expected level of statistical fluctuations.

The bottom panel quantifies the statistical improvement achieved by the paired-fixed simulations (solid line) with respect to the standard simulations. The dashed line represents simulations that are fixed but not paired. As a typical example of improvement, for wavenumbers $k_{\parallel} = 0.25 \text{ hMpc}^{-1}$ and 2 hMpc^{-1} at $z = 3$, we find estimates of the 1D Ly α flux power spectra converge 34 and 1.7 times faster in a paired-fixed ensemble compared with a standard ensemble. As with the 3D Ly α flux power (Section 3.2) there is a steadily increasing improvement at high k , forming a local minimum at $k_{\parallel} \simeq 1 \text{ hMpc}^{-1}$; since the 1D power spectrum mixes multiple modes from the 3D power spectrum, this is consistent with projecting the improvement discussed in Section 3.2.

4. Discussion and Conclusions

We have shown that using the method of paired-fixed simulations reduces the uncertainty of the mean mock Ly α forest power spectrum measured in hydrodynamical simulations, and therefore requires less computing time to estimate the expected value of the considered quantity, which is needed to evaluate the likelihood of the data. As a typical example of

improvement, for wavenumbers $k = 0.25 \text{ hMpc}^{-1}$ at $z = 3$, we find estimates of the 1D and 3D power spectra converge 34 and 12 times faster in a paired-fixed ensemble compared with a standard ensemble. The largest improvements are at small k , where model uncertainties are larger than observational uncertainties. The improvements are minimal at $k = 1 \text{ hMpc}^{-1}$, but at these scales the observational uncertainties are larger than the model uncertainties, so reduced uncertainty on the mean mock power spectrum is unnecessary. This suggests that running a large simulation box size with paired-fixed ICs will provide the most accurate mock Ly α forest power spectrum over the largest range of scales.

By reducing the computational time required to achieve a target accuracy for mock power spectra, the method frees up resources for a more thorough exploration of astrophysical and cosmological parameters. It is essential to be able to sample efficiently over this space, with at least three parameters describing the shape and amplitude of the input linear power spectrum and at least two for astrophysical effects to span the reionization redshift and level of heat injection. In the future, likelihoods for data sets such as eBOSS and DESI will likely take advantage of emulators, which interpolate predicted power spectra within these high-dimensional parameter spaces (Heitmann et al. 2016; Walther et al. 2018, K. Rogers et al. 2018, in preparation, S. Bird et al. 2018, in preparation). In this context, freeing up CPU time by using the paired-fixed approach will allow for a denser sampling of training points for the emulator. That in turn will beat down interpolation errors and thus lead to more accurate inferences from forthcoming data.

It is a pleasure to thank Anže Slosar for useful discussions. The work of L.A., F.V.N., and S.G. is supported by the Simons Foundation. A.P. was supported by the Royal Society. A.F.R. was supported by an STFC Ernest Rutherford Fellowship, grant reference ST/N003853/1. A.P. and A.F.R. were further supported by STFC Consolidated Grant No. ST/R000476/1. This work was partially enabled by funding from the University College London (UCL) Cosmoparticle Initiative. This work was supported by collaborative visits funded by the Cosmology and Astroparticle Student and Postdoc Exchange Network (CASPEN). The simulations were run on the Gordon cluster at the San Diego Supercomputer Center supported by the Simons Foundation.

Software: The code used in this project is available from <https://github.com/andersdot/LyA-InvertPhase>. This research utilized the following open-source Python packages: *Astropy* (Astropy Collaboration et al. 2013), *scipy* (Jones et al. 2001), *matplotlib* (Hunter 2007), *fake_spectra* (Bird 2017), and *numpy* (Van der Walt et al. 2011). It also utilized the python pylans libraries, publicly available at <https://github.com/franciscovillaescusa/PyLans>, and *lyman_alpha* (Rogers et al. 2018a, 2018b).

ORCID iDs

Lauren Anderson <https://orcid.org/0000-0001-5725-9329>
 Andrew Pontzen <https://orcid.org/0000-0001-9546-3849>
 Andreu Font-Ribera <https://orcid.org/0000-0002-3033-7312>
 Francisco Villaescusa-Navarro <https://orcid.org/0000-0002-4816-0455>
 Shy Genel <https://orcid.org/0000-0002-3185-1540>

References

- Angulo, R. E., & Pontzen, A. 2016, *MNRAS*, **462**, L1
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
- Bautista, J. E., Busca, N. G., Guy, J., et al. 2017, *A&A*, **603**, A12
- Bird, S. 2017, FSFE: Fake Spectra Flux Extractor, Astrophysics Source Code Library, ascl:1710.012
- Borde, A., Palanque-Delabrouille, N., Rossi, G., et al. 2014, *JCAP*, **7**, 005
- Busca, N. G., Delubac, T., Rich, J., et al. 2013, *A&A*, **552**, A96
- Croft, R. A. C., Weinberg, D. H., Katz, N., & Hernquist, L. 1998, *ApJ*, **495**, 44
- du Mas des Bourboux, H., Le Goff, J.-M., Blomqvist, M., et al. 2017, *A&A*, **608**, A130
- Font-Ribera, A., Kirkby, D., Busca, N., et al. 2014, *JCAP*, **5**, 27
- Gunn, J. E., & Peterson, B. A. 1965, *ApJ*, **142**, 1633
- Heitmann, K., Bingham, D., Lawrence, E., et al. 2016, *ApJ*, **820**, 108
- Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, *ApJL*, **457**, L51
- Humlicek, J. 1979, *JQSRT*, **21**, 309
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Iršič, V., Viel, M., Berg, T. A. M., et al. 2017, *MNRAS*, **466**, 4332
- Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open source scientific tools for Python, <http://www.scipy.org/>
- Kaiser, N. 1987, *MNRAS*, **227**, 1
- Lukić, Z., Stark, C. W., Nugent, P., et al. 2015, *MNRAS*, **446**, 3697
- McDonald, P., & Eisenstein, D. J. 2007, *PhRvD*, **76**, 063009
- McDonald, P., Seljak, U., Cen, R., et al. 2005a, *ApJ*, **635**, 761
- McDonald, P., Seljak, U., Cen, R., Bode, P., & Ostriker, J. P. 2005b, *MNRAS*, **360**, 1471
- McQuinn, M. 2016, *ARA&A*, **54**, 313
- Palanque-Delabrouille, N., Yèche, C., Lesgourgues, J., et al. 2015, *JCAP*, **2**, 45
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, **594**, A13
- Pontzen, A., Slosar, A., Roth, N., & Peiris, H. V. 2016, *PhRvD*, **93**, 103519
- Rogers, K. K., Bird, S., Peiris, H. V., et al. 2018a, *MNRAS*, **474**, 3032
- Rogers, K. K., Bird, S., Peiris, H. V., et al. 2018b, *MNRAS*, **476**, 3716
- Seljak, U., Makarov, A., McDonald, P., et al. 2005, *PhRvD*, **71**, 103515
- Seljak, U., Makarov, A., McDonald, P., & Trac, H. 2006a, *PhRvL*, **97**, 191303
- Seljak, U., Slosar, A., & McDonald, P. 2006b, *JCAP*, **10**, 14
- Slosar, A., Iršič, V., Kirkby, D., et al. 2013, *JCAP*, **4**, 26
- Springel, V. 2005, *MNRAS*, **364**, 1105
- Takada, M., & Hu, W. 2013, *PhRvD*, **87**, 123504
- Van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, **13**, 22
- Viel, M., Becker, G. D., Bolton, J. S., & Haehnelt, M. G. 2013, *PhRvD*, **88**, 043502
- Viel, M., & Haehnelt, M. G. 2006, *MNRAS*, **365**, 231
- Viel, M., Haehnelt, M. G., & Lewis, A. 2006, *MNRAS*, **370**, L51
- Viel, M., Haehnelt, M. G., & Springel, V. 2004, *MNRAS*, **354**, 684
- Villaescusa-Navarro, F., Naess, S., Genel, S., et al. 2018, arXiv:1806.01871
- Walther, M., Oñorbe, J., Hennawi, J. F., & Lukić, Z. 2018, arXiv:1808.04367
- Zennaro, M., Bel, J., Villaescusa-Navarro, F., et al. 2017, *MNRAS*, **466**, 3244