# Inverted initial conditions: Exploring the growth of cosmic structure and voids

Andrew Pontzen,[1,*] Anže Slosar,[2] Nina Roth,[1] and Hiranya V. Peiris[1]

[1]*Department of Physics and Astronomy, University College London, Gower Street,*
*London WC1E 6BT, United Kingdom*
[2]*Brookhaven National Laboratory, Upton, New York 11973, USA*
(Received 28 November 2015; published 18 May 2016)

We introduce and explore "paired" cosmological simulations. A pair consists of an A and B simulation with initial conditions related by the inversion $\delta_A(\mathbf{x}, t_{\text{initial}}) = -\delta_B(\mathbf{x}, t_{\text{initial}})$ (underdensities substituted for overdensities and vice versa). We argue that the technique is valuable for improving our understanding of cosmic structure formation. The A and B fields are by definition equally likely draws from $\Lambda$CDM initial conditions, and in the linear regime evolve identically up to the overall sign. As nonlinear evolution takes hold, a region that collapses to form a halo in simulation A will tend to expand to create a void in simulation B. Applications include (i) contrasting the growth of A-halos and B-voids to test excursion-set theories of structure formation, (ii) cross-correlating the density field of the A and B universes as a novel test for perturbation theory, and (iii) canceling error terms by averaging power spectra between the two boxes. Generalizations of the method to more elaborate field transformations are suggested.

## I. INTRODUCTION

The interpretation of cosmological observations increasingly requires a precise understanding of nonlinear structure formation. In addition to the power spectrum of the matter distribution, the properties and abundances of structures such as clusters [1] or voids [2] also have the potential to constrain cosmological parameters.

Excursion-set theories [3–5] show that the formation of voids from initial underdensities is nearly but not precisely analogous to the formation of halos from overdensities [6–9]. The imperfect symmetry suggests that directly contrasting void and halo formation could be informative. In this work we take a first step in this direction by comparing results from two simulations with precisely opposite initial conditions (underdensities substituted for overdensities and vice versa). We refer to these simulations as being "paired."

The paired simulations can also be used to improve both practical estimation and theoretical understanding of the matter power spectrum (and higher-order correlations). There are presently two approaches to calculating the nonlinear power spectrum: analytic perturbation theory, or computational $N$-body simulations. The former comes in a wide variety of flavors, because the simplest perturbative treatment of gravitational instability (standard perturbation theory, SPT [10]) suffers from divergences at increasing comoving wave number $k$. These can be brought under control by partially resumming some of the SPT series [11] or writing down an

effective theory [12]. The resulting theories can be tested or calibrated on simulations [13–18].

The most familiar example of a nonstandard perturbation theory is the Zel'dovich approximation, a linear expansion in Lagrangian space which leads to a regrouping of terms. While the raw Zel'dovich predictions for the auto-power spectrum are inaccurate, in many respects it behaves better than Eulerian perturbation theory [11,19,20]. In particular, it correctly predicts the decay of the cross-correlation between initial conditions and the final nonlinear field [11,14,20–22]. In the present work we cross-correlate the nonlinear density fields of the paired simulations, providing an alternative performance comparison of different perturbative schemes from a physical perspective. We find that the Zel'dovich approximation continues to offer insight in this new regime.

From a purely practical perspective the science case for forthcoming large-scale structure surveys requires percent-level accuracy on computations even on strongly nonlinear, megaparsec scales [23]. Our third application for paired simulations shows how they can be used to cancel a large class of finite-volume errors that can compromise this requirement. The same cancellation can be approximately achieved by averaging over a large ensemble of uncorrelated simulations, but the paired approach is more computationally efficient.

After describing the simulation setup (Sec. II) we discuss the asymmetry in the evolution of halos and voids (Sec. II A) and then show how the technique generates new insights into perturbation theory (Sec. II B) and improves the accuracy of power spectrum estimates (Sec. II C). Possible extensions to the technique are discussed in Sec. III. We conclude in Sec. IV.

---

[*]a.pontzen@ucl.ac.uk

103519-1

## II. RESULTS

In this paper we present results from paired cosmological simulations drawn from an ensemble described by the WMAP 7-year recommended cosmological parameters [24] ("WMAP $+$ BAO $+$ $H_0$ ML"). While these are no longer the most precise parameters available [25], they are sufficiently close for our present purposes where we do not compare to observational data; adopting WMAP7 parameters allowed us to make use of an existing simulation which we refer to as "A." We perform dark-matter-only simulations, adding the baryon density to that of the dark matter.

We used CAMB [26] to generate the initial power spectrum of fluctuations from which we drew a random realization $\delta_S(x, t_{\text{initial}})$ on a uniform $512^3$ grid in a $(200h^{-1}\,\text{Mpc})^3$ volume, so probing wave numbers $0.031 < k/(h\,\text{Mpc}^{-1}) < 16$. The initial particle displacements and velocities were generated using the Zel'dovich approximation at redshift $z = 99$, deep in the linear regime for the relevant scales. All simulations were run to $z = 0$ using the P-GADGET-3 code [27,28]. The particle softening was set to $\epsilon = 5h^{-1}$ kpc. Halos were identified with the SUBFIND algorithm [29].

We generated the initial conditions for the first simulation (denoted "A") using the same code as Ref. [30] and flipped the sign of the overdensity to generate the "B" initial conditions, $\delta_B(x, t_{\text{initial}}) = -\delta_A(x, t_{\text{initial}})$. Both simulations are on an equal footing in the sense that they are equally probable draws from the underlying statistical description of the initial conditions. However, their relationship with each other allows for systematic investigations into various aspects of structure formation as we describe in the following sections.

### A. Evolution of antihalos

To begin our study of the relationship between the A and B simulations, we show that halos map reliably onto voids (and vice versa). This situation is illustrated in Fig. 1 which shows, from left to right, a $20h^{-1}$ Mpc slice through the matter density field of the $z = 0$ A simulation, the $z = 99$ A simulation, the $z = 99$ B simulation and the $z = 0$ B simulation. The brightness represents projected mass density while colors track the fraction of particles identified as halos in the A simulation. Note that the color hue of each particle is therefore fully determined by the A simulation, even in the B-simulation panels. The voids (light shades) in the B simulation are identified as A-halo-dominated regions (colored blue) interspersed by filaments, which are A-halo-free regions (colored orange). Although our presentation privileges the A simulation halo catalog, the overall relationship is symmetric: a similar figure can be made starting from the halos in the B simulation.

Figure 1 suggests that voids in the B simulation can be identified with "antihalos," i.e., the Lagrangian region defined by the particles making up A-halos. However, theories of structure formation using the excursion-set formalism emphasize that there is an asymmetry between the evolution of halos and voids [6–8]: voids can be crushed by a large-scale overdensity that collapses at late times, whereas halos are not erased by living in a large-scale void.

The A/B comparison allows us to search for direct evidence of this asymmetry. We take the $z = 0$ A-halos in three mass bins: $10^{14} < M/M_\odot < 10^{15}$, $10^{13} < M/M_\odot < 10^{14}$ and $10^{12} < M/M_\odot < 10^{13}$; there are respectively 353, 4699 and 38 537 in each range. For each A-halo at $z = 0$, we track the constituent particles through
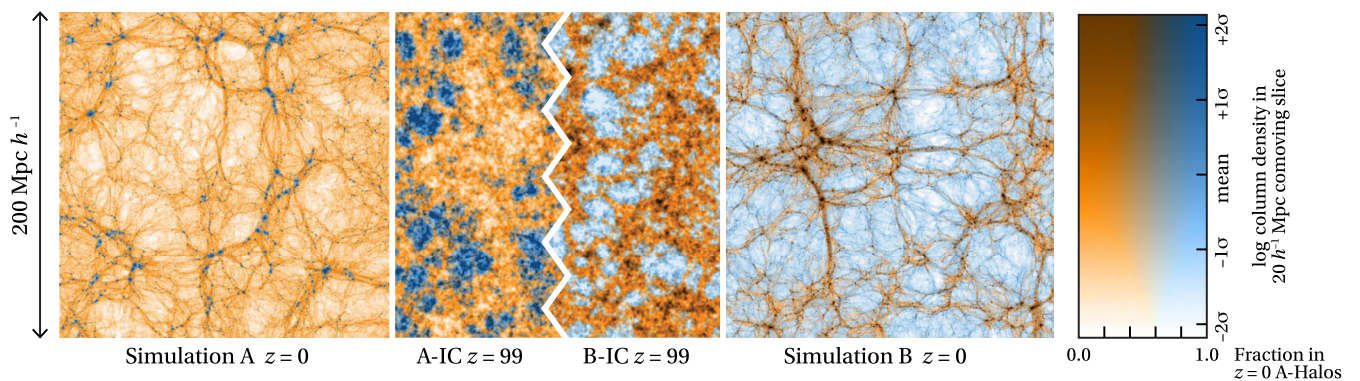


FIG. 1. An illustration of the paired simulation technique. A standard ΛCDM simulation is performed as described in the text. The left panel shows the present-day ($z = 0$) projected density field in a $20h^{-1}$ Mpc slice through the simulation. Collapsed dark matter halos have been identified using a friends-of-friends algorithm; the fraction of the column density contributed by particles in such structures is color coded from orange (no contribution) to blue (100% contribution). The center-left panel shows the initial conditions, color coded according to the same scheme. The initial conditions for the B simulation are obtained by reversing the sign of the overdensity field (center-right panel). While the statistical properties of the linear field are unchanged by this transformation, the blue "A-halo" particles are now associated with underdensities. Evolving the B simulation to $z = 0$ gives rise to the right panel. The B-voids are seen to be associated with the same particles (i.e., the same Lagrangian regions) as the A-halos.
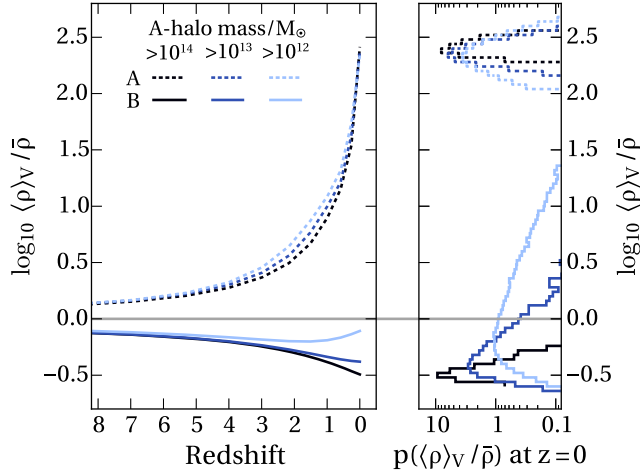
FIG. 2. The volume-averaged density of Lagrangian regions corresponding to $z = 0$ A-halos of different mass ranges (from darkest to lightest: $10^{14} < M/M_\odot < 10^{15}$, $10^{13} < M/M_\odot < 10^{14}$ and $10^{12} < M/M_\odot < 10^{13}$ respectively). The left panel shows the evolution of each mass bin's mean density; the right panel shows the spread of halo-averaged densities within the bin at $z = 0$. Dashed lines show the regions in the A simulation (so at $z = 0$ these correspond to halos); solid lines show the corresponding "antihalo" regions in the B simulation. The most massive antihalos can be identified as voids. At lower masses there is a tail of crushed, high-density antihalos, reflecting the known void-in-cloud evolution asymmetry.

time in both A and B simulations to follow the collapse of the halo (A) or expansion of the antihalo (B). At each output time step, we record the volume-weighted[1] mean density of each Lagrangian region,

$$\langle \rho \rangle_V \equiv \frac{\sum_i \rho_i h_i^3}{\sum_i h_i^3}, \qquad (1)$$

where the sum is over all particles $i$ associated with a particular region, $\rho_i$ is a local density estimate computed by PYNBODY [31] using the 64 nearest-neighbor particles, and $h_i$ is the physical distance to the furthest of these neighbors. We divide by the cosmic mean density $\bar\rho$ to remove the effects of the background expansion.

The results of the density calculation are shown in Fig. 2. Over time, the Lagrangian region corresponding to the final $z = 0$ halos grows in density (dashed lines, left panel). The differences between the three mass bins are relatively small, with a slight trend for lower-mass regions to reach higher densities at earlier times. The right panel shows a histogram of the densities of the individual halos making up each mass bin at $z = 0$; once again, the A-simulation results are shown by dashed lines. The variance in the mean density is

---

[1]The volume weighting is crucial because much of the mass within voids is contained inside rare but dense halos [6] which contaminate the mass-weighted mean.

small, which is to be expected given that the halos are identified based on a friends-of-friends algorithm which specifies a fixed density for their boundary [29].

The solid lines show the corresponding quantities for the antihalo regions in the B simulation. The left panel shows that, at early times, the selected regions are underdense, as demanded by the antisymmetry in the initial conditions. Over time the largest antihalos become progressively less dense, as expected for voids. The histogram (right panel) confirms that the most massive antihalos are all well below the cosmic mean density and can be robustly identified as voids, confirming the more qualitative picture painted by Fig. 1.

In the lowest mass bin, the average density of the antihalos turns around and starts to grow (relative to the cosmic mean) at low redshift. This is consistent with the expected "void-crushing" process [6]. The right panel shows that the majority of antihalos remain underdense, but the mean is dragged up by a few regions. Inspection of these high-density cases confirms that they are being crushed by larger-scale collapse. The effect is only evident at low mass; otherwise, even if antihalos are contained within a B overdensity, there has not been time for gravitational collapse to crush them. We can describe the antihalos above a density threshold of $\langle \rho \rangle_V / \bar\rho > 200$ as "fully crushed," since they have achieved a mean density comparable to that of a halo. Even at $10^{12} \, M_\odot$ (the minimum mass we can reliably resolve), only 0.1% of antihalos at $z = 0$ exceed the threshold. It is far more common to find antihalos that have been crushed only along two dimensions, and now form the diffuse mass in a cosmic filament.

In summary, antihalos correspond closely to voids, especially on large scales. There is presently significant interest in formulating reliable ways of defining voids so that such structures can be identified and used for cosmological inference in large-scale structure surveys [32]. By selecting antihalos that have not been crushed, one could arrive at a clean definition of voids. We will explore this further in a future paper.

### B. Perturbation theory

In this section we show how our paired simulation approach can be used to study cosmological perturbation theory. Given a density field for a given simulation labeled $X$ we define

$$\delta_X(\mathbf{k}) \equiv \int d^3\mathbf{x} \left( \frac{\rho_X(\mathbf{x})}{\bar\rho} - 1 \right) e^{-i\mathbf{k}\cdot\mathbf{x}}, \qquad (2)$$

where $\mathbf{k}$ is a comoving wave vector, $\delta_X(\mathbf{k})$ is the Fourier-space overdensity, $\mathbf{x}$ is the comoving position, $\rho_X(\mathbf{x})$ is the density, and $\bar\rho$ is the mean density. The cross-power spectrum between fields $X$ and $Y$, $P_{XY}(k)$, is defined by

$$\frac{1}{2} \langle \delta_X^\star(\mathbf{k})\delta_Y(\mathbf{k}') + \delta_Y^\star(\mathbf{k})\delta_X(\mathbf{k}') \rangle$$
$$= (2\pi)^3 P_{XY}(k)\delta^D(\mathbf{k} - \mathbf{k}'), \tag{3}$$

where angular brackets denote the ensemble average, and $\delta^D$ represents the Dirac delta function. We will make use of the simulated density fields A and B, but also the linearly evolved field L which is defined as

$$\delta_L(\mathbf{k}, t) = \delta_A(\mathbf{k}, t_{\text{initial}})D(t, t_{\text{initial}}), \tag{4}$$

where $D(t, t_{\text{initial}})$ is the linear growth factor. From these three fields, there are six power spectra that can be constructed; however, in the true ensemble average, two of these ($P_{AA} = P_{BB}$ and $P_{AL} = -P_{BL}$) contain identical information. (In practice, since the volume of simulations is finite, there is residual information in $P_{AA} - P_{BB}$ and $P_{AL} + P_{BL}$ that we will discuss in Sec. II C.)

We modified the GENPK code[2] [33,34] to calculate cross-correlations between two GADGET outputs. For the purposes of the present discussion we construct four power spectrum estimates: $P(k) = (P_{AA}(k) + P_{BB}(k))/2$, $P_{\times L}(k) = (P_{AL}(k) - P_{BL}(k))/2$, $P_{AB}$ and $P_{LL}(k)$. We assume that particle shot noise is uncorrelated between A and B simulations and therefore do not subtract its contribution to the cross-spectra; the validity of this assumption does not affect our results, since shot noise is highly subdominant over the scales of interest.

We start by focussing on the cross-correlations $P_{\times L}$ and $P_{AB}$; these are plotted for a range of redshifts in Fig. 3 (upper and lower panels respectively), normalized by $P_{LL}$. The quantity $P_{\times L}/P_{LL}$ is sometimes called the propagator; it expresses the degree of coherence between the nonlinear and linear fields and has been studied extensively [11,14,20,35]. These studies have revealed that the connection between the initial overdensity and final nonlinear structure is poorly described by SPT, but (as we rederive in the Appendix) accurately predicted by resumming the Zel'dovich approximation which gives

$$P_{\times L}^{\text{zel}}(k) = e^{-(k/k_{\text{NL}})^2}P_{LL}(k) \tag{5}$$

where $k_{\text{NL}}$ is the wave vector corresponding to the scale at which the linear and nonlinear fields decohere,

$$k_{\text{NL}}^{-2} = \frac{1}{12\pi^2} \int_0^\infty P_L(k')\mathrm{d}k'. \tag{6}$$

The qualitative reason for this decoherence is straightforward: in the nonlinear evolution, the particles have moved (from their initial positions) an rms distance $\langle \Delta x^2 \rangle^{1/2}$ that is proportional to $k_{\text{NL}}^{-1}$; for scales below this limit, the original information has been erased by the displacements.[3]
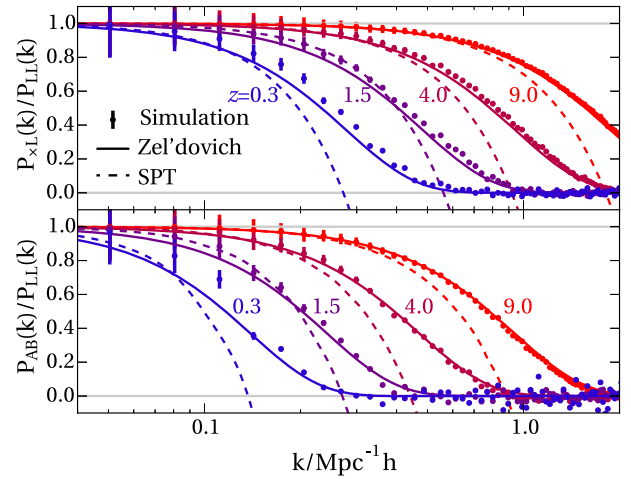


FIG. 3. The cross-power spectra, for four redshifts $0.3 < z < 9.0$, of simulated and linearly evolved fields ($P_{\times L}$, upper panel) and A and B simulations ($P_{AB}$, lower panel), each normalized by the linear power $P_{LL}$. The Zel'dovich resummation from the Appendix is shown as a solid line, and gives excellent agreement with the simulations even at low redshift. Standard perturbation theory at one-loop order is shown by the dashed line and is in poor agreement with the simulations at all redshifts. These results are known for the linear cross-correlation (upper panel) but continue to hold for the new $P_{AB}$ cross-correlation (lower panel).

The quantitative corrections to the Zel'dovich description arising from higher-order effects can be estimated in a formal resummation scheme and are seen to be small in the case of $P_{\times L}$ [36]. By contrast, finite-order standard perturbation theory—an expansion in the Eulerian density contrast, shown by dashed lines in Fig. 3—performs worse in describing the decorrelation. For that reason, the Zel'dovich result has been used as an inspiration for partially resumming perturbation theory to combine the best of both worlds. For example, both resummed standard perturbation theory (RPT) [11,35] and resummed Lagrangian perturbation theory (LPT) [20] are designed to reproduce the Zel'dovich cross-correlation result at tree level [14].

We will now show that the cross-correlation between the A and B simulations provides a new test bed for perturbation theory schemes. The $P_{AB}$ measurements from the simulations are shown in the lower panel of Fig. 3 along with one-loop SPT (dashed lines), showing that the perturbation theory gives a reasonable prediction at sufficiently low $k$ but once more diverges in the high-$k$ limit.[4] As in the AL case, it is possible to use the Zel'dovich approximation to find a far better description of the AB decorrelation. The solid line in the lower panel of Fig. 3 shows the result, derived in the Appendix:

---

[2]http://github.com/sbird/GenPK.
[3]Even better agreement could be found by fitting the scale of the Gaussian suppression, $k_{\text{NL}}$, at each redshift.

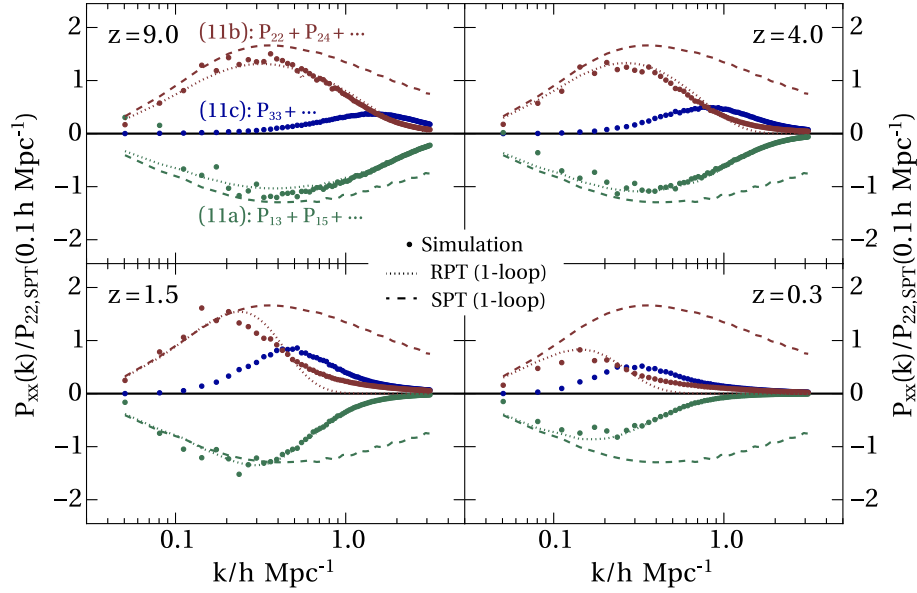[4]We used the COPTER code [14] to calculate perturbation theory results for this paper.

FIG. 4. The power spectrum empirically split into different SPT series according to Eq. (11), normalized for convenience to $P_{22,\mathrm{SPT}}(0.1h\,\mathrm{Mpc}^{-1})$. The three equations are labeled by their SPT contributions. Dashed lines show the SPT calculation for these terms (although we can only compute for one-loop order). Dotted lines show the equivalent results for RPT which incorporates high-order effects even in the one-loop truncation. Points show measurements from the paired simulations, from which we can verify the effectiveness of resummation, and also see directly that the magnitude of two-loop $P_{33}$ terms is small at sufficiently low $k$.

$$P_{AB}^{\mathrm{zel}}(k) = e^{-(2k/k_{\mathrm{NL}})^2} P_{LL}(k), \qquad (7)$$

giving an excellent fit to the simulation results. One justification for this form is to imagine that the particle displacements are doubled in magnitude relative to the AL displacements, so that the relevant wave number is halved. However the more formal derivation of Eq. (7) as given in the Appendix requires an unconventional choice of resummation. The reason why this particular choice is appropriate for the $P_{AB}^{\mathrm{zel}}$ calculation is discussed further in the Appendix.

Instead of directly studying cross-correlations, we can use the new information to empirically constrain the terms within SPT. In this approach, the nonlinear overdensity field $\delta_A$ is written as the sum of terms of increasing powers of the linear field, $\delta_A = \delta_1 + \delta_2 + \cdots$. The power spectrum is then expanded as a series in the auto- and cross-power of these individual terms; schematically

$$P_{AA} = P_{BB} = P_{11} + P_{13} + P_{22} + P_{15} + P_{24} + P_{33} + \cdots \qquad (8)$$

to two-loop order, where $P_{ij}$ denotes the parts of the power spectrum formed from contracting a term which is $i$th order in $\delta_L$ with one which is $j$th order.[5] Note that there are no terms for which $i+j$ is odd, since the ensemble

---

[5]By convention the $P_{ij}$ term absorbs the $P_{ji}$ term if $i \neq j$; this leads to a potentially confusing factor of 2 notational discrepancy between the formal definition of symmetric ($i = j$) and asymmetric ($i \neq j$) terms [37]. We nevertheless adopt this convention for compatibility with the existing literature.

expectation value is identically zero in such cases. We have suppressed the $k$ parameter for brevity.

By cross-correlating the A and B simulated densities with the linear field, we obtain the series

$$P_{AL} = -P_{BL} = P_{11} + \frac{1}{2}(P_{13} + P_{15} + \cdots), \qquad (9)$$

where the factor $1/2$ arises because, unlike in the auto-correlation case, there are no $P_{ji}$ terms to absorb into the $P_{ij}$ terms. Cross-correlating with the B field gives

$$P_{AB} = -P_{11} - P_{13} + P_{22} - P_{15} + P_{24} - P_{33} + \cdots, \qquad (10)$$

where we have picked up a minus sign in front of each $P_{ij}$ term for which $i$ and $j$ are odd (so that an odd number of B linear fields appears).

Truncating at the two-loop order, the relationships (8), (9) and (10) can be partially inverted to obtain

$$P_{13} + P_{15} + \cdots = \frac{1}{2}(P_{LA} - P_{LB}) - P_{LL}, \qquad (11a)$$

$$P_{22} + P_{24} + \cdots = \frac{1}{2}(P_{AA} + P_{BB} + P_{AB}), \qquad (11b)$$

$$P_{33} + \cdots = P_{LL} - P_{LA} + P_{LB} - \frac{1}{2}P_{AB} + \frac{1}{4}(P_{AA} + P_{BB}). \qquad (11c)$$

The leading-order lhs of Eqs. (11a) and (11b) can be computed using one-loop SPT. The rhs can be measured

from our paired simulations, and predicted by RPT (or other resummed theories). The comparison is shown in Fig. 4: the simulation results are shown as points, SPT by dashed lines and RPT by dotted lines.

The one-loop SPT predictions for the $P_{13}$ and $P_{22}$ class terms are poor for $k > 0.2h\,\mathrm{Mpc}^{-1}$, although the errors are opposite in sign and so significantly cancel to produce reasonable predictions of the autocorrelation [38]. The strength of RPT (dotted lines) in predicting the $P_{13}$ series derives directly from its exact agreement with the Zel'dovich approximation in the $P_{AL}$ cross-correlation (Fig. 3).

By adding the AB cross-correlation we have been able to construct an expression that, at leading order, returns $P_{33}$ (a two-loop term) directly from simulations for the first time. Such terms must be small at low $k$ for perturbation theory to be valid. Our results demonstrate that this requirement does in fact hold in numerical simulations.

Being able to extract different perturbation theory terms empirically also gives the opportunity for testing resummation schemes. The addition of the "B" simulation gives access to a distinctive higher-order test that is not available from existing methods. At present we do not have code to calculate two-loop predictions so this comparison is left for future work.

### C. Improving the accuracy of structure formation simulations

As a final example application of paired simulations, we turn to a more immediately practical question. Since numerical simulations of nonlinear structure formation can probe only a finite dynamic range, practitioners need to balance the box size against the ability to resolve small scales. The finite volume has two effects. First, it removes all power below $k_{\min} = 2\pi/L_{\mathrm{box}}$, where $L_{\mathrm{box}}$ is the comoving size of the box [39]. This could be tackled in a computationally efficient way by assuming a separate-universe approximation and rescaling the background cosmology in each "patch" [40]; we will not consider this further here. The second effect of the small box is that it samples only a small number of modes for wave numbers reaching $k = k_{\min}$, leading to variance effects that vanish in a true ensemble mean. Paired simulations can be helpful in tackling this problem.

In this section we will need to make a clear distinction between the theoretical power spectrum $P(k)$, as defined by Eq. (3), and the measured power spectrum $\tilde{P}(k)$ which is defined with reference to the discretized density field components in a simulation

$$\tilde{P}(k) = \frac{1}{N_k} \sum_{i \in S_k} \delta_i^* \delta_i, \qquad (12)$$

where $\delta_i$ is the density field Fourier component with index $i$, $S_k$ is the set of such components that are used in the power spectrum estimate for wave number $k$, and $N_k$ is the size of that set. Shot-noise corrections [41] can be applied to Eq. (12) without changing the discussion; we omit it for simplicity. Expanding $\delta_i$ to second order in perturbation theory, we have

$$\tilde{P}(k) \simeq \frac{1}{N_k} \sum_{i \in S_k} (\delta_{i,L}^* \delta_{i,L} + G_{ijk}(\delta_{j,L}^* \delta_{k,L}^* \delta_{i,L}) + \mathrm{c.c.} + \cdots)$$

$$(13)$$

where $\delta_{i,L}$ is the linear amplitude for component $i$, c.c. indicates the complex conjugate of the preceding term, $D$ is the linear growth factor, and $G_{ijk}$ describes how mode $i$ grows in response to the amplitude of modes $j$ and $k$, evaluated at a given redshift. The linear growth function has been absorbed into the definition of the linear field according to Eq. (4). There is an assumed summation over all modes $j$ and $k$.

Given simulations of a specified box size, the ideal quantity to calculate is $\langle \tilde{P}(k) \rangle$, i.e., an average over all possible realizations of the initial field $\delta_{i,L}$. This is typically attempted by computing tens or even hundreds of realizations [42,43], which is computationally costly. The leading-order correction that this generates compared to a single realization can actually be applied by hand, since $\langle \tilde{P}(k) \rangle = P_L(k) + \cdots$. The required "first-order corrected" power spectrum estimate is

$$\tilde{P}(k)_{\mathrm{corr},1} = \tilde{P}_{AA}(k) + P_L(k) - \tilde{P}_{LL}(k), \qquad (14)$$

so that $\tilde{P}(k)_{\mathrm{corr},1} - \langle \tilde{P}(k) \rangle$ is third order in $\delta$. Note that this is different from the usual approach of "canceling" sample variance [44] by estimating

$$\tilde{P}(k)_{\mathrm{corr,std}} = \tilde{P}_{AA}(k) \frac{P_L(k)}{\tilde{P}_{LL}(k)}, \qquad (15)$$

which is hard to justify from a theoretical point of view (though one gets the right answer on linear scales by construction).

With a paired simulation in hand, we can go further and apply the next-to-leading-order correction because, inspecting Eq. (13), the error has odd parity in $\delta_L$ and so reverses sign in $P_{BB}$. Thus,

$$\tilde{P}(k)_{\mathrm{corr},2} = \frac{1}{2}(\tilde{P}_{AA}(k) + \tilde{P}_{BB}(k)) + P_L(k) - \tilde{P}_{LL}(k). \qquad (16)$$

Residual errors in $\tilde{P}(k)_{\mathrm{corr},2}$ compared to $\langle \tilde{P}(k) \rangle$ are then fourth order in $\delta$.

The corrections arising from this change are highly significant in the case of a $200h^{-1}$ Mpc box. Figure 5 shows the correction $\Delta\tilde{P}(k) = (\tilde{P}_{BB}(k) - \tilde{P}_{AA}(k))/2$, as a fraction of $\tilde{P}(k)_{\mathrm{corr},2}$. The corrections reach $\sim 5\%$ even at small scales, $k = 1.0h^{-1}$ Mpc, and modest redshifts, $z = 1$, where one might hope that the box-size effects are minimal.
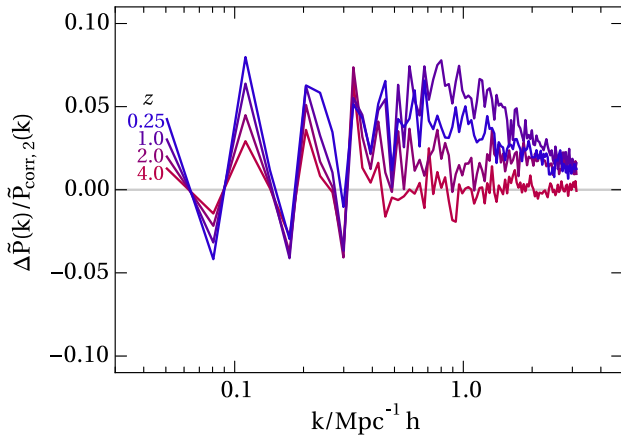
FIG. 5.    Finite-box errors in power spectrum estimates can be vastly reduced using a paired simulation. Here we show our correction for the next-to-leading-order error term $\Delta \tilde{P}(k) = \tilde{P}(k)_{\text{corr},2} - \tilde{P}(k)_{\text{corr},1}$ as a fraction of $\tilde{P}(k)_{\text{corr},2}$. Despite having a reasonable $200h^{-1}$ Mpc box size, correlated artifacts from the small number of large-scale modes propagate down to create an apparent bias in the power spectrum, reaching $\simeq 5\%$ on scales as small as $k = 1$ Mpc$^{-1}h$ at $z = 1$. This cannot be corrected by older techniques that divide out cosmic variance in the linear power, as it is an inherently nonlinear effect.

This is consistent with what is found by averaging over hundreds of realizations [43] or comparing to increased box sizes which better sample the large-scale modes [23].

Moreover, unlike the second-order cosmic variance, the third-order error is strongly correlated over different $k$'s, presumably because it arises from the coupling to a small number of low-$k$ modes [45]. In other words, the sample variance is not determined solely by the number of modes in the initial conditions at the same wave number, but can instead be dominated by the $G_{ijk}$ coupling to the poorly sampled low-$k$ modes.

By performing just one additional simulation, it is possible to remove this bias to third-order accuracy. The fourth-order term is left invariant by the averaging. Overall, our paired technique enables significant gains in computational efficiency when generating nonlinear power spectra for comparison with large cosmological surveys. A more comprehensive exploration of how pairing can improve the accuracy of structure formation simulations, including on measurement correlations beyond the power spectrum, is presented in Ref. [46].

## III. EXTENSIONS

As well as fleshing out the three applications in Sec. II, future work could examine wider classes of statistics-preserving transformations; for example, anything of the form

$$\delta(\mathbf{k}) \to T(\mathbf{k})\delta(\mathbf{k}) \qquad (17)$$

with $|T(\mathbf{k})|^2 = 1$ is suitable. For the overdensity field to remain real, one additionally requires $T(\mathbf{k}) = T(-\mathbf{k})^{\star}$ but otherwise there are no restrictions. In particular there is no requirement for $T$ to be isotropic or homogeneous.

Condition (17) ensures that the power spectrum is unchanged; our method of cross-correlation will then allow the study of structure growth in the presence of the correct cosmological background. The form of $T(\mathbf{k})$ is dictated by the specific aspect of structure growth under study.

Section II's A-B simulations correspond to the simplest case of $T(\mathbf{k}) = -1$; as another example, a translation $\mathbf{x} \to \mathbf{x} + \Delta\mathbf{x}$ corresponds to the case $T(\mathbf{k}) = e^{i\mathbf{k}\cdot\Delta\mathbf{x}}$. Let us briefly consider one further illustrative extension, given by

$$T(\mathbf{k}) = 1 - 2\Theta(k_0 - |\mathbf{k}|), \qquad (18)$$

where $\Theta$ is the Heaviside step function. The resulting transformation flips the sign of $\delta$ for wave numbers below a critical $k_0$. We can refer to the simulation resulting from the new initial conditions as "spliced" (abbreviated to S) since the initial conditions are identical to A for $k > k_0$ and to B for $k < k_0$. In terms of perturbation theory, this splicing operation is more complex than a $k$-independent transformation because it breaks loop terms into an infrared and ultraviolet part with different signs. Being able to segregate parts of loop integrals fully within a numerical simulation in principle allows a very detailed comparison with perturbation theory. Here we will consider only the qualitative results.

Because the S and A simulations are anticorrelated on large scales, the low-$k$ modes destroy the high-$k$ correlations over time (Fig. 6, top panel) just as with the A-B cross correlation (Fig. 3, lower panel). On the other hand, cross-correlating S with B reveals that the low-$k$ modes remain positively correlated at all times, showing that the anticorrelation on small scales does not affect the larger scales. Furthermore, as nonlinear power grows in the late-time universe, the ratio $P_{SB}/P$ ultimately becomes positive at large $k$: structure growth is coherent between the S and B universes because it is regulated by the largest scale modes. A full understanding of the coupling of large- and small-scale modes is necessary for distinguishing the bispectrum due to nonlinear evolution [47] from any primordial contribution. Using our technique this behavior is exposed to quantitative study without ever changing the power spectrum away from $\Lambda$CDM, and with just three simulations rather than expensive averages over large numbers [45].

One could expand to an even broader class of transformations where two independent, uncorrelated initial realizations ($\delta_{\text{I}}$ and $\delta_{\text{II}}$, say) are available:

$$\delta(\mathbf{k}) \to T(\mathbf{k})\delta_{\text{I}}(\mathbf{k}) + S(\mathbf{k})\delta_{\text{II}}(\mathbf{k}) \qquad (19)$$

where $|T(\mathbf{k})|^2 + |S(\mathbf{k})|^2 = 1$. This includes another interesting special case where the $k < k_0$ modes are kept
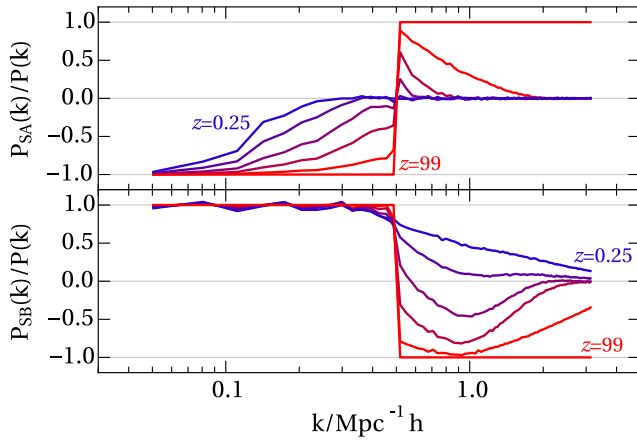
FIG. 6. As an example extension, we show the cross-correlation between the spliced ("S") simulation and the original ("A" and "B") pair. Modes in the S initial conditions are equal to the "A" modes on small scales ($k > 0.5$ Mpc$^{-1}h$), but to the "B" modes on large scales ($k < 0.5$ Mpc$^{-1}h$). In cross-correlation with the A simulation, one sees the effect of large-scale streaming decorrelating the small scales. Conversely, the lower panel shows a trend towards coherence between S and B simulations on small scales, showing that the final power at large $k$ is being sourced by low-$k$ density fluctuations. All three simulations have a $\Lambda$CDM power spectrum, so the effects are being directly measured within a realistic cosmological setting rather than a toy model.

fixed while $k > k_0$ modes are randomized (rather than anticorrelated). This specialization has been studied elsewhere [48,49] using large numbers of runs with independent $\delta_{\mathrm{II}}$ fields. Applications include studying the insensitivity of the evolved universe to early small-scale fluctuations [48], and averaging away stochastic fluctuations in halo spin alignments [50] or local bias measurements [51]. Our approach of pulling out information from a single additional simulation with transformed initial conditions could also be applied to this specialization, for example as another way to isolate the contribution of modes in a specific $k$ range to the perturbation theory loop terms.

## IV. CONCLUSIONS

We have introduced the technique of "paired" simulations. We run two simulations ("A" and "B") that are identical except for having inverted initial linear overdensities ($\delta_A = -\delta_B$). Since by definition the linear field is symmetric about zero, the two simulations have identical statistical properties. We illustrated how this can be used to better understand the evolution of voids, extract information on the physical basis of perturbation theories, and eliminate a class of finite-volume effects from power spectrum estimates with greater efficiency than existing techniques. Extensions to a broader class of transformations of the initial density field could further enhance the power of this technique.

## APPENDIX: AB CROSS-CORRELATIONS IN THE ZEL'DOVICH APPROXIMATION

In this appendix, we derive the result quoted in the text for the cross-power spectrum of the A and B simulations in the Zel'dovich approximation. Our approach closely follows that of previous works [20,52], but extends to the cross-power spectrum with an alternative resummation that we will describe in due course.

The Zel'dovich approximation is the linear-order solution to Lagrangian perturbation theory. The central quantity is the displacement field $\mathbf{\Psi}(\mathbf{q})$ which describes the movement of particles from their initial positions $\mathbf{q}$ to their final position $\mathbf{x} = \mathbf{q} + \mathbf{\Psi}(\mathbf{q})$. All information about the system is then expressed in terms of $\mathbf{\Psi}(\mathbf{q})$. For example, the local density is

$$\rho(\mathbf{x}) = \bar{\rho} \left| \frac{\mathrm{d}^3 \mathbf{q}}{\mathrm{d}^3 \mathbf{x}} \right|, \tag{A1}$$

where $|\mathrm{d}^3\mathbf{q}/\mathrm{d}^3\mathbf{x}|$ denotes the Jacobian determinant of the transformation between Lagrangian and Eulerian coordinates $\mathbf{q}$ and $\mathbf{x}$, and $\bar{\rho}$ is the volume-averaged density. This allows the fractional overdensity in Fourier space, $\delta(\mathbf{k})$, to be written

$$\delta(\mathbf{k}) \equiv \int \mathrm{d}^3\mathbf{x} \left( \frac{\rho(\mathbf{x})}{\bar{\rho}} - 1 \right) e^{-i\mathbf{k}\cdot\mathbf{x}}$$

$$= \int \mathrm{d}^3\mathbf{q} \, e^{-i\mathbf{k}\cdot\mathbf{q}} [e^{-i\mathbf{k}\cdot\mathbf{\Psi}(\mathbf{q})} - 1], \tag{A2}$$

where Eq. (A1) has been used to transform the integration variable to $\mathbf{q}$ for the first term, whereas the second term has been rewritten with a relabeling of the integration coordinate from $\mathbf{x}$ to $\mathbf{q}$. Mass conservation demands that $\langle\delta(\mathbf{k})\rangle = 0$ which, combined with Eq. (A2), implies

$$\int d^3 q e^{-i\mathbf{k}\cdot\mathbf{q}} \langle e^{-i\mathbf{k}\cdot\mathbf{\Psi}(\mathbf{q})} \rangle = \int d^3 q e^{-i\mathbf{k}\cdot\mathbf{q}}, \qquad (A3)$$

a result that we will use momentarily.

The cross-power spectrum $P_{XY}(k)$ between fields $X$ and $Y$ is defined by

$$\langle \delta_X(\mathbf{k})\delta_Y(\mathbf{k}') \rangle = (2\pi)^3 \delta^D(\mathbf{k}+\mathbf{k}') P_{XY}(k), \qquad (A4)$$

where $\delta^D$ is the Dirac delta function. Substituting two copies of Eq. (A2) into this definition gives the following expression for the cross-power in terms of the displacement fields:

$$P_{XY}(k) = \int d^3 r e^{-i\mathbf{k}\cdot\mathbf{r}} [\langle e^{-i\mathbf{k}\cdot\Delta\mathbf{\Psi}} \rangle - 1], \qquad (A5)$$

where we have used Eq. (A3) and defined $\Delta\mathbf{\Psi} \equiv \mathbf{\Psi}_X(\mathbf{q}) - \mathbf{\Psi}_Y(\mathbf{q}')$ with $\mathbf{r} \equiv \mathbf{q} - \mathbf{q}'$.

The treatment to this point has been exact (up to shell crossing). We now introduce the perturbative element by employing the Zel'dovich approximation, in which the displacement is related to the linear-theory density field $\delta_L(\mathbf{k})$ by

$$\mathbf{\Psi}_X^{\text{zel}}(\mathbf{k}) = i\alpha_X \frac{\mathbf{k}}{k^2} \delta_L(\mathbf{k}) \qquad (A6)$$

for a constant $\alpha_X$ depending on which field $X$ we consider. For the A field, $\alpha_A = 1$; for the B field, $\alpha_B = -1$.

We additionally want to be able to calculate the cross-correlation $P_{AL}$ between the true and linearly evolved fields. Since the Zel'dovich approximation and linear theory have to agree in the limit of small density variations, we can represent the linear theory by multiplying the displacements by some small number $\alpha_L \ll 1$, but then dividing the output density field (A2) by the same small number $\alpha_L$. One can think of this procedure as rescaling the input linear field by a growth factor appropriate for some very early time, then undoing the scaling in the final expression. As a verification, we can recalculate the linear field from Eq. (A2), substituting Eq. (A6) and then Taylor expanding in $\alpha_L$:

$$\delta_L^{\text{zel}}(\mathbf{k}) \simeq \frac{1}{\alpha_L} \int d^3 q e^{-i\mathbf{k}\cdot\mathbf{q}} \int \frac{d^3 k'}{(2\pi)^3} \alpha_L \frac{\mathbf{k}\cdot\mathbf{k}'}{\mathbf{k}'^2} \delta_L(\mathbf{k}') e^{i\mathbf{k}'\cdot\mathbf{q}}$$
$$= \delta_L(\mathbf{k}), \qquad (A7)$$

confirming the recovery of the linear density field.

We now return to Eq. (A5) and insert Eq. (A6) for the displacement fields. In the Zel'dovich picture, $\mathbf{\Psi}_X$ is always Gaussian, so we can apply the cumulant expansion $\langle e^{-G} \rangle = \exp[-\frac{1}{2}\langle G^2 \rangle]$. We divide the final expression by $|\alpha_X \alpha_Y|$; for $\alpha_A = 1$ and $\alpha_B = -1$ this has no effect,

whereas for $\alpha_L \ll 1$ this captures the shift to linear theory described above. Put together, we obtain the following expression for auto- and cross-spectra:

$$P_{XY}^{\text{zel}}(k) = \frac{1}{|\alpha_X \alpha_Y|} \int d^3 r e^{-i\mathbf{k}\cdot\mathbf{r}} [e^{-(\alpha_X - \alpha_Y)^2 I(\mathbf{k},0)/2 + \alpha_X \alpha_Y J(\mathbf{k},\mathbf{r})} - 1]$$
$$(A8)$$

where $I(\mathbf{k},\mathbf{r})$ captures the effect of the displacement field for two points at distance $\mathbf{r}$ on a Fourier mode with wave vector $\mathbf{k}$ and is given by

$$I(\mathbf{k},\mathbf{r}) = \int \frac{d^3 k'}{(2\pi)^3} \frac{(\mathbf{k}\cdot\mathbf{k}')^2}{k'^4} \cos(\mathbf{k}'\cdot\mathbf{r}) P_L(k'), \qquad (A9)$$

and $J(\mathbf{k},\mathbf{r}) = I(\mathbf{k},\mathbf{r}) - I(\mathbf{k},0)$. The exponent in the integral above has two pieces, one that depends on $\mathbf{r}$ (proportional to $J(\mathbf{k},\mathbf{r})$) and the one that does not [proportional to $I(\mathbf{k},0)$]. The piece that does not depend on $\mathbf{r}$ can be pulled out of the integral to generate a $k$-dependent exponential suppression. The exponential for the other piece is expanded to first order and integrated to generate a linear power spectrum term and a harmless $k = 0$ correction

$$P_{XY}^{\text{zel}}(k) = e^{-(\alpha_X - \alpha_Y)^2 (k/k_{\text{NL}})^2} [P_L(k) - I(\mathbf{k},0)\delta^D(\mathbf{k})], \qquad (A10)$$

where $k_{\text{NL}}$ is the wave number corresponding to a non-linearity scale, defined by

$$k_{\text{NL}}^{-2} = \frac{1}{2} I(k,0)k^{-2} = \frac{1}{12\pi^2} \int_0^\infty P_L(k') dk'. \qquad (A11)$$

It follows that

$$P_{AL}^{\text{zel}}(k) = e^{-(k/k_{\text{NL}})^2} P_L(k), \quad \text{and} \qquad (A12)$$

$$P_{AB}^{\text{zel}}(k) = e^{-4(k/k_{\text{NL}})^2} P_L(k), \qquad (A13)$$

which are the results quoted in the main text. The decorrelation scale between the A and B fields is half that of the decorrelation scale between A and L fields, a result confirmed in our measurements from the simulations.

This derivation looks deceptively close to the resummed LPT approach of Ref. [20], which we denote by "M." However, there is an important conceptual difference in the detail. In Ref. [20], the equivalent of our Eq. (A8) is given by

$$P_{XY}^{\text{zel,M}}(k) = \frac{1}{|\alpha_X \alpha_Y|} \int d^3 r e^{-i\mathbf{k}\cdot\mathbf{r}} [e^{-(\alpha_X^2 + \alpha_Y^2)I(\mathbf{k},0)/2 + \alpha_X \alpha_Y I(\mathbf{k},\mathbf{r})} - 1].$$
$$(A14)$$

Equations (A8) and (A14) are mathematically equivalent: we have just moved a constant from the $\mathbf{r}$-independent piece back into the $\mathbf{r}$-dependent piece (which we later

expand). At infinite order in the expansion this does not matter, but following the same reasoning as above we obtain to the first order in expanded integral,

$$P_{AL}^{\mathrm{zel,M}}(k) = e^{-(k/k_{\mathrm{NL}})^2} P_L(k), \qquad (A15)$$

$$P_{AB}^{\mathrm{zel,M}}(k) = e^{-2(k/k_{\mathrm{NL}})^2} P_L(k). \qquad (A16)$$

The cross-correlation between the initial and final fields is the same, but the AB decorrelation scale differs by a factor of 2. The difference can be attributed to resummation of a different set of operators, so one needs to justify why our choice is appropriate for the problem in hand.

For the case of constructing the auto-power spectrum, any large-scale displacements appear (at sufficiently short wavelengths) to be a uniform translation, and so become

irrelevant. However this is no longer true in the AB case where such displacements—occurring in opposite directions in the two fields—strongly suppress power. For this reason, we want our tree-level solution for the cross-correlation to include the effect of coherent displacements; in other words we construct Eq. (A8) such that the perturbative terms vanish for small separations, $r \to 0$, an approach also taken by Ref. [53]. This construction implicitly underlies the intuitive description given in Sec. II B that the AB particle displacement is doubled relative to the AL displacement: one has to assume that the displacements are coherent, which only becomes exactly true in the $r \to 0$ ("eikonal") limit. Conversely, Eq. (A14) is well suited to expanding the auto-power ($\alpha_X = \alpha_Y = 1$) precisely because it is insensitive to large-wavelength displacements.

[1] P. A. R. Ade *et al.* (Planck Collaboration), arXiv:1502.01597.
[2] G. Lavaux and B. D. Wandelt, Astrophys. J. **754**, 109 (2012).
[3] W. H. Press and P. Schechter, Astrophys. J. **187**, 425 (1974).
[4] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser, Astrophys. J. **379**, 440 (1991).
[5] A. R. Zentner, Int. J. Mod. Phys. D **16**, 763 (2007).
[6] R. K. Sheth and R. van de Weygaert, Mon. Not. R. Astron. Soc. **350**, 517 (2004).
[7] S. R. Furlanetto and T. Piran, Mon. Not. R. Astron. Soc. **366**, 467 (2006).
[8] A. Paranjape, T. Y. Lam, and R. K. Sheth, Mon. Not. R. Astron. Soc. **420**, 1648 (2012).
[9] E. Jennings, Y. Li, and W. Hu, Mon. Not. R. Astron. Soc. **434**, 2167 (2013).
[10] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Phys. Rep. **367**, 1 (2002).
[11] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063519 (2006).
[12] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, J. High Energy Phys. 09 (2012) 082.
[13] T. Nishimichi *et al.*, Publ. Astron. Soc. Jpn. **61**, 321 (2009).
[14] J. Carlson, M. White, and N. Padmanabhan, Phys. Rev. D **80**, 043531 (2009).
[15] N. Roth and C. Porciani, Mon. Not. R. Astron. Soc. **415**, 829 (2011).
[16] H. Gil-Marín, C. Wagner, L. Verde, C. Porciani, and R. Jimenez, J. Cosmol. Astropart. Phys. 11 (2012) 029.
[17] Z. Vlah, U. Seljak, and T. Baldauf, Phys. Rev. D **91**, 023508 (2015).
[18] S. Foreman, H. Perrier, and L. Senatore, arXiv:1507.05326.
[19] T. Buchert, Mon. Not. R. Astron. Soc. **254**, 729 (1992).
[20] T. Matsubara, Phys. Rev. D **77**, 063530 (2008).
[21] F. Bernardeau, M. Crocce, and R. Scoccimarro, Phys. Rev. D **78**, 103521 (2008).
[22] M. White, Mon. Not. R. Astron. Soc. **439**, 3630 (2014).
[23] A. Schneider, R. Teyssier, D. Potter, J. Stadel, J. Onions, D. S. Reed, R. E. Smith, V. Springel, and F. R. Pearce, arXiv:1503.05920.
[24] E. Komatsu *et al.*, Astrophys. J. Suppl. Ser. **192**, 18 (2011).
[25] P. A. R. Ade *et al.* (Planck Collaboration), arXiv:1502.01589.
[26] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000).
[27] V. Springel, Mon. Not. R. Astron. Soc. **364**, 1105 (2005).
[28] V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J. F. Navarro, C. S. Frenk, and S. D. M. White, Mon. Not. R. Astron. Soc. **391**, 1685 (2008).
[29] V. Springel, S. D. M. White, G. Tormen, and G. Kauffmann, Mon. Not. R. Astron. Soc. **328**, 726 (2001).
[30] N. Roth, A. Pontzen, and H. V. Peiris, Mon. Not. R. Astron. Soc. **455**, 974 (2016).
[31] A. Pontzen, R. Roskar, G. Stinson, and R. Woods, PYNBODY: *N*-body/SPH analysis for python, Astrophysics Source Code Library ascl:1305.002 (2013).
[32] P. M. Sutter, P. Elahi, B. Falck, J. Onions, N. Hamaus, A. Knebe, C. Srisawat, and A. Schneider, Mon. Not. R. Astron. Soc. **445**, 1235 (2014).
[33] S. Bird, H. V. Peiris, M. Viel, and L. Verde, Mon. Not. R. Astron. Soc. **413**, 1717 (2011).
[34] S. Bird, M. Viel, and M. G. Haehnelt, Mon. Not. R. Astron. Soc. **420**, 2551 (2012).
[35] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063520 (2006).
[36] M. Crocce, R. Scoccimarro, and F. Bernardeau, Mon. Not. R. Astron. Soc. **427**, 2537 (2012).
[37] N. Makino, M. Sasaki, and Y. Suto, Phys. Rev. D **46**, 585 (1992).

[38] D. Jeong and E. Komatsu, Astrophys. J. **651**, 619 (2006).

[39] Y. Li, W. Hu, and M. Takada, Phys. Rev. D **89**, 083519 (2014).

[40] R. E. Angulo and S. D. M. White, Mon. Not. R. Astron. Soc. **405**, 143 (2010).

[41] Y. P. Jing, Astrophys. J. **620**, 559 (2005).

[42] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon, Astrophys. J. **715**, 104 (2010).

[43] K. Heitmann, E. Lawrence, J. Kwan, S. Habib, and D. Higdon, Astrophys. J. **780**, 111 (2014).

[44] V. Springel *et al.*, Nature (London) **435**, 629 (2005).

[45] A. Meiksin and M. White, Mon. Not. R. Astron. Soc. **308**, 1179 (1999).

[46] R. E. Angulo and A. Pontzen, arXiv:1603.05253.

[47] C. Wagner, F. Schmidt, C.-T. Chiang, and E. Komatsu, Mon. Not. R. Astron. Soc. **448**, L11 (2015).

[48] B. Little, D. H. Weinberg, and C. Park, Mon. Not. R. Astron. Soc. **253**, 295 (1991).

[49] M. A. Aragon-Calvo, Mon. Not. R. Astron. Soc. **455**, 438 (2016).

[50] M. A. Aragon-Calvo and L. F. Yang, Mon. Not. R. Astron. Soc. **440**, L46 (2014).

[51] M. C. Neyrinck, M. A. Aragón-Calvo, D. Jeong, and X. Wang, Mon. Not. R. Astron. Soc. **441**, 646 (2014).

[52] A. N. Taylor and A. J. S. Hamilton, Mon. Not. R. Astron. Soc. **282**, 767 (1996).

[53] J. Carlson, B. Reid, and M. White, Mon. Not. R. Astron. Soc. **429**, 1674 (2013).