# Nonlinear *E*-mode clustering in Lagrangian space

Hao-Ran Yu,<sup>1,2,\*</sup> Ue-Li Pen,<sup>2,3,4,5,†</sup> and Hong-Ming Zhu<sup>6,7,‡</sup>

<sup>1</sup>Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China

<sup>2</sup>Canadian Institute for Theoretical Astrophysics, University of Toronto,

60 St. George Street, Toronto, Ontario M5S 3H8, Canada

<sup>3</sup>Dunlap Institute for Astronomy and Astrophysics, University of Toronto,

50 St. George Street, Toronto, Ontario M5S 3H4, Canada

<sup>4</sup>Canadian Institute for Advanced Research, CIFAR Program in Gravitation and Cosmology,

Toronto, Ontario M5G 1Z8, Canada

<sup>5</sup>Perimeter Institute for Theoretical Physics,

31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada

<sup>6</sup>Key Laboratory for Computational Astrophysics, National Astronomical Observatories,

Chinese Academy of Sciences, 20A Datun Road, Beijing 100012, China

<sup>7</sup>University of Chinese Academy of Sciences, Beijing 100049, China

(Received 8 November 2016; published 2 February 2017)

We study the nonlinear *E*-mode clustering in Lagrangian space by using large scale structure *N*-body simulations and use the displacement field information in Lagrangian space to recover the primordial linear density field. We find that, compared to Eulerian nonlinear density fields, the *E*-mode displacement fields in Lagrangian space improves the cross-correlation scale k with initial density field by a factor of 6–7, containing 2 orders of magnitude more primordial information. This illustrates ability of potential density reconstruction algorithms, to improve the baryonic acoustic oscillation measurements from current and future large scale structure surveys.

DOI: 10.1103/PhysRevD.95.043501

# I. INTRODUCTION

Our Universe starts from primordial Gaussian perturbations at a very early stage, and from those fluctuations, the gravitational instability drives the formation of the large scale structure (LSS) distribution of matter [1,2]. These structures grow linearly until the perturbations are large enough so that the first order perturbation theories are unable to analytically describe the LSS distributions [3]. As a result, the final nonlinear LSS distribution contains higher order statistics, and thus makes it more challenging to be interpreted into basic cosmological parameters. One such example is that the baryonic acoustic oscillation (BAO) scale can be used as a "standard ruler" to constrain the cosmic expansion history and thus probes the dark energy properties [4], but nonlinear evolution smears the BAO features and lowers the measurement accuracy [5,6]. There are various attempts to recover earlier stages of LSS, in which statistics are closer to Gaussian [7,8]. Because Gaussian fields can be adequately described by two-point statistics, ideally after some recovery algorithms, more information can be extracted, more straightforwardly, by power spectra or two-point correlation functions [9,10].

Standard BAO reconstruction algorithms apply in Eulerian space. They usually smooth the nonlinear density

field on linear scale (~10 Mpc/h) and reverse the large scale bulk flows by a negative Zel'dovich linear displacement [11-13]. In the linear Lagrangian perturbation theory (LPT), the negative divergence of the displacement field  $-\nabla_q \cdot \Psi(q)$  with respect to Lagrangian coordinates q gives the linear density field [14], and in many literatures [15–17]  $\nabla_a \cdot \Psi(q)$  is compared with various standard LPT and corrected LPT initial conditions. Their motivation was trying to correct or modify the LPT displacement fields in order that the final positions of particles are brought closer to N-body results. Because of the absence of Lagrangian space information in observations, few density reconstruction algorithms are developed according to the relation between displacement and linear density. However, there are techniques to estimate the displacement field from a final distribution of matter. For example, when a homogeneous initial matter distribution is assumed, there is a unique solution of curl-less displacement field to relate the initial and final distributions without shell crossing. This solution can be solved by a metric transformation equation [18,19]. In the one-dimensional (1D) case, this solution simplifies to an ordering of mass elements by their final, Eulerian coordinates. Zhu et al. [20] apply this algorithm to the result of a 1D simulation [3] and obtain an estimated displacement field  $\tilde{\Psi}(q)$ , and find that this new method well recovers the linear information and reconstructs the 1D BAO peak in the correlation function. In 3D and more realistic cases, one needs to carefully consider

haoran@cita.utoronto.ca

pen@cita.utoronto.ca

hmzhu@nao.cas.cn

## HAO-RAN YU, UE-LI PEN, and HONG-MING ZHU

effects of curl, shell crossing, complicated baryonic physics and biased tracers (galaxies).

Before these steps, we need to quantify the amount of linear information that can be recovered<sup>1</sup> from the full nonlinear displacement field  $\Psi(q)$ . This field can be decomposed into a curl-less "E-mode" component and a divergenceless "B-mode" component,  $\Psi = \Psi_E + \Psi_B$ , where the *B*-mode is raised by late stage non-Gaussianities and is not dominant in  $\Psi$  [16]. In  $\Psi(q)$  from N-body simulations, (1) E-mode displacement field is used for recovering the linear information (although *B*-mode is also present); (2) shell-crossing effects are fully considered; (3) the setup is clean in the absence of baryons and realistic observables. Further reconstructions by  $\Psi(q)$  can be compared with this result. Furthermore, we study the scale dependent cross correlation between  $-\nabla_q \cdot \Psi(q)$  and linear density field and construct optimal Wiener filters to get the optimal filtered recovered linear density field and recovered linear power spectrum.

In the rest of the paper, we describe the simulation, density recovery algorithm and Wiener filter setups in Sec. II. We show results in Sec. III. Discussion and the conclusion are in Sec. IV.

### **II. METHOD**

We show the LSS simulation and displacement field setups in Sec. II A. In Sec. II B, we recover the linear density field from the displacement field  $\Psi(q)$  from simulations. Note that potential reconstruction algorithms are based on an estimated displacement field  $\tilde{\Psi}(q)$  instead of  $\Psi(q)$ . In the following sections we use  $\delta_R$  to label the recovered linear density field from  $\Psi(q)$ .

#### A. Simulation

We use the open source cosmological simulation code CUBE [21]. Cosmological parameters are set as  $\Omega_m = 0.27$ ,  $\Omega_{\Lambda} = 0.73$ ,  $h_0 = 0.68$ ,  $n_s = 0.96$  and  $\sigma_8 = 0.83$ . Initial conditions are generated at redshift z = 50 using Zel'dovich approximation, and using a CLASS transfer function [22].  $N_p = 512^3$  N-body particles are evolved via their mutual gravitational interactions to z = 0, in a periodic box with L = 400 Mpc/h per side. The code is set to use a standard particle-mesh algorithm [23] on two-level mesh grids (for details, see [24]) and cloud-in-cell (CIC) is used in particle interpolations in force calculation and obtaining the density field  $\rho(\mathbf{x})$  in Eulerian coordinates  $\mathbf{x}$  at late stages. We use density contrast  $\delta \equiv \rho/\langle \rho \rangle - 1$  to describe the density fluctuations. The primordial linear density field  $\delta_L$  is given by the initial stage and scaled to

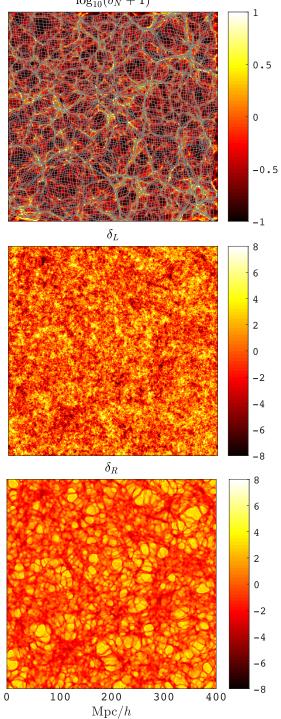


FIG. 1. Visualization of the nonlinear density field  $\delta_N$  (top), linear density field  $\delta_L$  (middle) and the raw recovered density field  $\delta_R$  (bottom). These projections have 9.375 Mpc/*h* thickness and 400 Mpc/*h* per side. The top panel shows the nonlinear displacement field  $\Psi$  by the deformed mesh, which traces the LSS of  $\delta_N$ .

z = 0 by the linear growth factor. In the top and middle panels of Fig. 1 we show projections of the nonlinear density field  $\delta_N$  given by the simulation and the linear density field  $\delta_L$ .  $\delta_N$  is obtained by the particle distribution

<sup>&</sup>lt;sup>1</sup>To avoid ambiguity, we call  $\delta_R \equiv -\nabla_q \cdot \Psi(q)$  recovered linear density field, which requires  $\Psi(q)$  from *N*-body simulations, while  $-\nabla_q \cdot \tilde{\Psi}(q)$  is called the reconstructed density field, which uses the estimated displacement field.

at redshift z = 0, and the particles are interpolated by using the cloud-in-cell (CIC) algorithm. Because  $\delta_N$  is highly nonlinear and follows an approximate log-normal distribution, we plot  $\log_{10}(\delta_N + 1)$  instead, and show the color scale  $\log_{10}(\delta_N + 1) \in [-1, 1]$  (or  $\rho_N / \langle \rho_N \rangle \in [0.1, 10]$ ) only, for a better visualization. The nonlinear evolution of  $\delta_N$  makes it very different from  $\delta_L$  in appearance.

The two-point statistics of these density fields are quantified by the cross power spectrum  $P_{ij}(k) \equiv (2\pi)^{-3} \langle |\delta_i(k)|| \delta_j(k)| \rangle$ , where subscripts *i*, *j* may refer to linear (*L*), nonlinear (*N*), recovered (*R*), or noise (*n*) density fields. When i = j it reduces to the auto power spectrum  $P_{ii}(k)$  or P(k). We usually plot the dimensionless power spectrum  $\Delta^2(k) \equiv k^3 P(k)/2\pi^2$ .

### **B.** Density recovery

In the simulation, we use particle-ID (PID) to record the initial (Lagrangian) location q of particles, and the information is tracked until the z = 0 and we can get the Lagrangian displacement vector  $\Psi \equiv x - q$  for every particle. Then these vectors are interpolated onto the initial Lagrangian coordinates q of particles and we get the displacement field  $\Psi(q)$ . To visualize the  $\Psi$  field, we draw a 3D uniform mesh over the volume, and use the given  $\Psi$ field to deform the mesh according to the direction and physical amplitude of  $\Psi$ . In the top panel of Fig. 1, The resulting mesh illustrates a "pseudocurvilinear coordinate" similar to [18], however the mesh can be overlapped due to shell crossing. The densest mesh grids trace the densest structures of  $\delta_N$ , whereas the undeformed grid positions are the Lagrangian coordinates in which we do the density recovery. The raw recovered density field is given by the differential motion of matter elements,

$$\delta_R = -\nabla \cdot \Psi(q). \tag{1}$$

Because the recovery processes are implemented on Lagrangian coordinates,  $\delta_R$  takes the coordinates of q instead of x. We just write q's Fourier wave number  $k_q$  as k to simplify the expression.

To quantify the linear information in the recovered density field  $\delta_R$ , we decompose  $\delta_L$  in Fourier space into two uncorrelated parts,

$$\delta_L(k) = r'\delta_R + \delta_n,\tag{2}$$

where the first term  $r'\delta_R$  is completely correlated with  $\delta_R$ , meaning the linear information we can extract from  $\delta_R$ . The noise term  $\delta_n$  is uncorrelated with  $\delta_R$ , since the rest of the linear information is not contained in  $\delta_R$ . Correlating Eq. (2) with  $\delta_R$  gives

$$P_{LR} = r' P_{RR} + P_{nR}, \tag{3}$$

where  $P_{ij} \equiv \langle \delta_i \delta_j \rangle$  denotes the cross power spectrum. Since  $\delta_n$  is uncorrelated with  $\delta_R$ ,  $P_{nR} = 0$ . With the definition of cross correlation coefficient  $r(\delta_L, \delta_R) \equiv P_{LR}/\sqrt{P_{LL}P_{RR}}$  and bias  $b^2 = P_{RR}/P_{LL}$ , we solve  $r' = P_{LR}/P_{RR} = rb^{-1}$ . Note that these computations above and below can also be applied on  $\delta_N$  for comparison, by replacing " $_R$ " with " $_N$ " in the equations, while we do not rewrite them explicitly in the paper for simplicity. From these, we plot the cross correlation coefficient  $r_{LN} = r(\delta_L, \delta_N)$  and  $r_{LR} = r(\delta_L, \delta_R)$  in Fig. 2.  $r_{LN}$  shows no correlation starting from  $k \approx 0.4 \ h/Mpc$  [25]. Clearly,  $\delta_R$  contains much more linear information on smaller scales.

According to Eq. (2), the auto power spectrum of  $\delta_L$  is decomposed as

$$P_{LL} = r^2 b^{-2} P_{RR} + P_{nn}, (4)$$

and  $P_{nn} = (1 - r^2)P_{LL}$ . We also explicitly compute the cross power spectrum between  $\delta_R$  and  $\delta_n = \delta_L - rb^{-1}\delta_R$ , and found that  $rb^{-1}P_{nR}$  is about 2 orders of magnitude lower than  $P_{LL}$ , being consistent with zero. This confirms that the signal term  $rb^{-1}\delta_R$  and the noise term  $\delta_n$  is indeed uncorrelated and validates Eq. (4). According to these two terms, we construct a Wiener filter to filter out the uncorrelated part in  $\delta_R$ :

$$W(k) = \frac{r^2 b^{-2} P_{RR}}{r^2 b^{-2} P_{RR} + P_{nn}} = r^2.$$
 (5)

The optimal recovered density is given by

$$\tilde{\delta}_R = W b^{-1} \delta_R,\tag{6}$$

and the optimal recovered power spectrum is given by

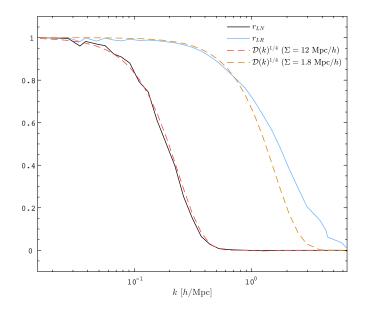


FIG. 2. Correlation functions  $r(\delta_L, \delta_N)$  and  $r(\delta_L, \delta_R)$  (solid lines) and their scaled BAO damping models (dotted lines).

HAO-RAN YU, UE-LI PEN, and HONG-MING ZHU

$$\tilde{P} = W^2 b^{-2} P_{RR} = W^2 P_{LL} + W^2 b^{-2} P_{NN}.$$
 (7)

Here  $W^2$  describes the damping of the linear power spectrum.

## **III. RESULTS**

To visualize the above algorithms, a projection of  $\delta_R$  is plotted in the bottom panel of Fig. 1, which looks closer to  $\delta_L$  compared to  $\delta_N$ . However the smallest scale structures are unable to be recovered.

As discussed in Sec. II B, Fig. 2 shows the cross correlation functions  $r_{LN}$  and  $r_{LR}$ . The latter extends the correlation with  $\delta_L$  to smaller scales by nearly an order of magnitude. The extra correlation scales well cover the BAO scales of our interest.

In Fig. 3, we show the auto power spectra of  $\delta_L$  and  $\delta_N$  in black dashed and blue solid curves. Their difference shows the nonlinear evolution of LSS on small scales. Their cross power (not shown for clarity of the figure) drops to a very low value on scales  $k \gtrsim 0.1 h/Mpc$ , indicating a loss of linear information in the nonlinear power spectrum  $P_{NN}$ . This scenario directly leads to how  $P_{LL}$  is decomposed according to Eq. (4). In the nonlinear case [with <sub>R</sub> replaced by <sub>N</sub> Eq. (4)], on small scales  $k \gtrsim 0.1 h/Mpc$ ,  $P_{LL}$  is dominated by uncorrelated, nonlinear noise, shown in the green dotted line. In the case of  $\delta_R$ , however,  $P_{LL}$  is decomposed into the orange dash-dotted correlated part and the purple dotted uncorrelated part according to Eq. (4). The correlated power spectrum is dominated on BAO scales of our interest.

To quantify the improvement of cross correlation in the power spectrum, we compute the damping factors  $W^2(k)$  respectively for the optimal filtered nonlinear and

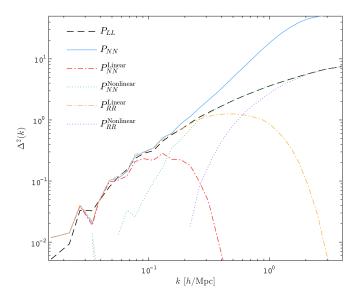


FIG. 3. Power spectra of  $\delta_L$ ,  $\delta_N$ , and the decomposition of  $P_{LL}$  into correlated parts and noise terms according to Eq. (4).

recovered density fields  $\tilde{\delta}_N$  and  $\tilde{\delta}_R$ . We fit Gaussian BAO damping models  $\mathcal{D}(k) = \exp(-k^2\Sigma^2/2)$  to these  $W^2(k)$ 's and give  $\Sigma = 1.8 \text{ Mpc}/h$  and  $\Sigma = 12 \text{ Mpc}/h$  for  $\delta_R$  and  $\delta_N$ . Since  $\mathcal{D}(k) = W^2 = r^4$ , we plot  $\mathcal{D}_N^{1/4}$  and  $\mathcal{D}_R^{1/4}$  over  $r_{LN}$  and  $r_{LR}$  in Fig. 2. The analyses are repeated with various box sizes (100, 300, 800 Mpc/*h* per side) and give consistent results.

To further illustrate the improvement in real space one point function correlations, in Fig. 4 we use the probability distribution as functions (PDFs) of  $(\delta_L, \delta_N)$  and  $(\delta_L, \delta_R)$  to show the point-point correlation between these two pairs of density fields. Since  $\delta_n$  in Eq. (2) is uncorrelated, we use Wiener filtered fields. To keep the consistency over  $\delta_L$ ,  $\delta_N$ and  $\delta_R$ , we use the  $W(k) = r_{LR}^2$  as the Wiener filter. The grey-scaled plots in the center of both panels show

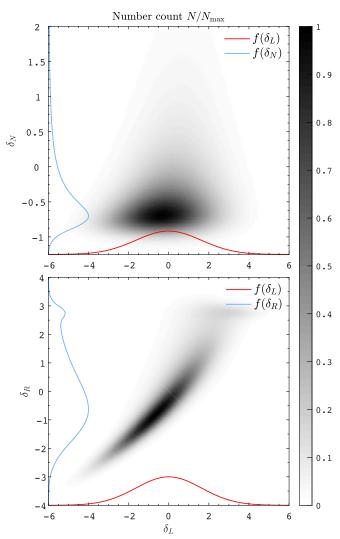


FIG. 4. Probability distribution functions (PDFs)  $f(\delta_L, \delta_N)$  and  $f(\delta_L, \delta_R)$ , showing in the upper and lower panels respectively. Both red curves on the *x* axes show  $f(\delta_L)$ , following a Gaussian distribution. The two curves on the *y* axes in the upper and the lower panel show respectively  $f(\delta_N)$  and  $f(\delta_R)$ .

the two-variable PDFs, whereas their projections onto each variable are just one-variable PDFs— $f(\delta_L)$ ,  $f(\delta_N)$ and  $f(\delta_R)$ , shown as red/blue curves on the axes of Fig. 4. In the top panel,  $\delta_N$  shows an approximate log-normal distribution (blue curve) and  $\delta_L$  follows an expected Gaussian distribution. They show tiny positive correlation in the 2D PDF. Because in Fourier space,  $\delta_L$  and  $\delta_N$  have correlations on only very large, linear scales (Fig. 2), they result in little correlation in real space-initial density fluctuations in Lagrangian coordinates are evolved/ transformed to Eulerian coordinates. As the recovery is done in Lagrangian space, it recovers a certain amount of correlation, as shown in the 2D PDF of the bottom panel of Fig. 4. One can also see that  $\delta_R$  follows a much closer Gaussian distribution (blue curve of the bottom panel). In denser regions of  $\delta_L$ ,  $\delta_R$  is saturated at  $\delta_R = 3$ , signifying the extreme collapsing of matter [15]:  $\delta_R = -\nabla \cdot \Psi = \nabla \cdot q = 3$ . Shell crossing makes  $\delta_R$  oscillate around 3. These second uncorrelated peaks damp out as we go to higher redshifts.

# **IV. DISCUSSION AND CONCLUSION**

We extract the actual displacement field of matter elements in cosmological N-body simulations, and use this displacement field to study the LSS nonlinear clustering in Lagrangian space. The displacement information is used to recover the primordial linear perturbations. The result shows a prominent improvement from  $r_{LN}$  to  $r_{LR}$  in Fig. 2—recovering the lost linear information on nearly an order of magnitude smaller scales. This is achieved by implementing differential movement information of matter elements on Lagrangian coordinates, rather than on Eulerian coordinates. This result illustrates the feasibility of using estimated displacement field  $\Psi(q)$  to reconstruct the primordial linear density field. A straightforward example of an estimation of  $\Psi(q)$  is given by [18,19]. In reality, one needs to consider all aspects including vorticity, shell crossing, bias, noise and data complexities. The impact of these factors can be quantitatively compared with the impact of different estimation methods, and with the exact solution by *N*-body simulations.

The advantage of using displacement field in reconstruction is its insensitive response from high nonlinearities. Nonlinear densities  $\delta_N$  can be arbitrarily large one expects virialized regions to be observable, where nonlinear density is given by the inverse determinant of the Jacobian  $\delta_N = |\mathbf{J}|^{-1} \gtrsim 200$ . However, reconstructed densities  $\delta_R$  are given by the trace tr(**J**) and saturates at 3. Actually, the displacement fields are dominated by early stage linear processes, which is the Lagrangian-Eulerian coordinate transform, while late stage shell crossing, nonlinear and baryonic processes only fine-tune the final position x [16]. Compared with estimated displacement fields  $\tilde{\Psi}(q)$  by [18], which do not have shell crossing, the additional shell-crossing information in  $\Psi(q)$  is uncorrelated with  $\delta_L$ . This insensitive response from nonlinearities enables the stability of reconstruction algorithms, which are expected to give similar results of this paper. In contrast, traditional treatments in reconstruction deals directly on density fields which sensitively relies on nonlinear processes-density values can vary by orders of magnitude due to nonlinear/baryonic physics and many sources of errors.

# ACKNOWLEDGMENTS

We thank Pengjie Zhang and Kwan Chuen Chan for helpful discussions and comments. We acknowledge funding from NSERC. H. R. Y. acknowledges General Financial Grant No. 2015M570884 and Special Financial Grant No. 2016T90009 from the China Postdoctoral Science Foundation. H. M. Z. acknowledges the support of the Chinese MoST 863 program under Grant No. 2012AA121701, the CAS Science Strategic Priority Research Program XDB09000000 and the NSFC under Grant No. 11373030. Research at the Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research & Innovation.

- [1] Y. B. Zel'dovich, Astron. Astrophys. 5, 84 (1970).
- [2] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White, Astrophys. J. 292, 371 (1985).
- [3] M. McQuinn and M. White, J. Cosmol. Astropart. Phys. 01 (2016) 043.
- [4] D. J. Eisenstein, New Astron. Rev. 49, 360 (2005).
- [5] D. J. Eisenstein et al., Astrophys. J. 633, 560 (2005).
- [6] W. Ngan, J. Harnois-Déraps, U.-L. Pen, P. McDonald, and I. MacDonald, Mon. Not. R. Astron. Soc. 419, 2949 (2012).
- [7] D. H. Weinberg, Mon. Not. R. Astron. Soc. 254, 315 (1992).

- [8] J. Harnois-Déraps, H.-R. Yu, T.-J. Zhang, and U.-L. Pen, Mon. Not. R. Astron. Soc. 436, 759 (2013).
- [9] C. D. Rimes and A. J. S. Hamilton, Mon. Not. R. Astron. Soc. 360, L82 (2005).
- [10] H.-R. Yu, J. Harnois-Déraps, T.-J. Zhang, and U.-L. Pen, Mon. Not. R. Astron. Soc. 421, 832 (2012).
- [11] D. J. Eisenstein, H.-J. Seo, E. Sirko, and D. N. Spergel, Astrophys. J. 664, 675 (2007).
- [12] Y. Noh, M. White, and N. Padmanabhan, Phys. Rev. D 80, 123501 (2009).

HAO-RAN YU, UE-LI PEN, and HONG-MING ZHU

- [13] N. Padmanabhan, M. White, and J. D. Cohn, Phys. Rev. D 79, 063523 (2009).
- [14] D. Jeong, Cosmology with high (z > 1) redshift galaxy surveys, Ph.D. thesis, University of Texas at Austin, 2010.
- [15] M. C. Neyrinck, Mon. Not. R. Astron. Soc. 428, 141 (2013).
- [16] K. C. Chan, Phys. Rev. D 89, 083515 (2014).
- [17] T. Baldauf, E. Schaan, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 03 (2016) 017.
- [18] U.-L. Pen, Astrophys. J. Suppl. Ser. 100, 269 (1995).
- [19] U.-L. Pen, Astrophys. J. Suppl. Ser. 115, 19 (1998).

- [20] H.-M. Zhu, U.-L. Pen, and X. Chen, arXiv:1609.07041.
- [21] H. R. Yu and U.-L. Pen (to be published).
- [22] D. Blas, J. Lesgourgues, and T. Tram, J. Cosmol. Astropart. Phys. 07 (2011) 034.
- [23] R. W. Hockney and J. W. Eastwood, Computer simulation using particles (1988).
- [24] J. Harnois-Deraps, U.-L. Pen, I.T. Iliev, H. Merz, J.D. Emberson, and V. Desjacques, Mon. Not. R. Astron. Soc. 436, 540 (2013).
- [25] U. Seljak and M. S. Warren, Mon. Not. R. Astron. Soc. 355, 129 (2004).