

## Lagrangian displacement field estimators in cosmology

Atsuhisa Ota<sup>1,2,\*</sup>, Hee-Jong Seo<sup>2,3</sup>, Shun Saito<sup>4,5</sup> and Florian Beutler<sup>6</sup>

<sup>1</sup>*HKUST Jockey Club Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, People's Republic of China*

<sup>2</sup>*Department of Physics and Astronomy, Ohio University, Athens, Ohio 45701, USA*

<sup>3</sup>*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA*

<sup>4</sup>*Institute for Multi-messenger Astrophysics and Cosmology, Department of Physics, Missouri University of Science and Technology, 1315 N. Pine Street, Rolla, Missouri 65409, USA*

<sup>5</sup>*Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, the University of Tokyo, Kashiwanoha, Kashiwa, Chiba 277-8583, Japan*

<sup>6</sup>*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, United Kingdom*



(Received 25 November 2022; accepted 19 May 2023; published 16 June 2023)

The late-time nonlinear Lagrangian displacement field is highly correlated with the initial field, so reconstructing it could enable us to extract primordial cosmological information. Our previous work [A. Ota *et al.*, *Phys. Rev. D* **104**, 123508 (2021)] carefully studied the displacement field reconstructed from the late-time density field using the iterative method proposed by Schmittfull *et al.* [*Phys. Rev. D* **96**, 023505 (2017)] and found that it does not fully converge to the true, underlying displacement field (e.g.,  $\sim 8\%$  offset at  $k \sim 0.2 h \text{ Mpc}^{-1}$  at  $z = 0.6$ ). We also constructed the Lagrangian perturbation theory model for the reconstructed field, but the model could not explain the discrepancy between the true and the reconstructed fields in the previous work. The main sources of the discrepancy were speculated to be a numerical artifact in the displacement estimator due to the discreteness of the sample. In this paper, we develop two new estimators of the displacement fields to reduce such a numerical discreteness effect, the normalized momentum estimator and the rescaled resumed estimator. We show that the discrepancy Ota *et al.* [*Phys. Rev. D* **104**, 123508 (2021)] reported is not due to the numerical artifacts. We conclude that the method from Schmittfull *et al.* [*Phys. Rev. D* **96**, 023505 (2017)] cannot fully reconstruct the shape of the nonlinear displacement field at the redshift we studied, while it is still an efficient baryon acoustic oscillation reconstruction method. In parallel, by properly accounting for the UV-sensitive term in a reconstruction procedure with an effective field theory approach, we improve the theoretical model for the reconstructed displacement field, by almost 5 times, from  $\sim 15\%$  to the level of a few percent at  $k \sim 0.2 h \text{ Mpc}^{-1}$  at the redshift  $z = 0.6$ .

DOI: [10.1103/PhysRevD.107.123523](https://doi.org/10.1103/PhysRevD.107.123523)

### I. INTRODUCTION

Baryon acoustic oscillation (BAO) imprints the sound horizon scale at recombination, which can be used to infer information about the nature of dark energy. The resulting precision critically depends on the strength of the BAO signal while the signal has been smeared as matter travels from the initial locations during nonlinear structure formation (e.g., [1–3]). However, the displacement of each mass tracer is mostly free from the degradation effect [4], as theoretically suggested by the Lagrangian resummation theory, where the exponential damping factor appears after resumming the one-loop density power spectrum by using the Lagrangian displacement [5]. Therefore, the Lagrangian displacement could be a useful degradation-free alternative,

for extracting the BAO information, to the traditional observables based on the Eulerian fluid dynamics [6].

Estimating the true displacement field from observed mass tracers is not straightforward for real surveys. First, we only measure the final locations of mass tracers. Also, there is a technical difficulty in estimating an unbiased displacement field from discrete, subsampled tracers as we will incorrectly measure the vanishing displacement field at the location where tracers do not exist.

Recently, there have been promising extensions of the standard density field reconstruction [7] suggested by various groups (e.g., [6,8–17]). Among these, Ref. [6] is one of the subset (for example, the method by Refs. [10–13] is also designed to derive the displacement field) that more directly focuses on the aspect of reconstructing the displacement field.

The method was demonstrated to return a superior BAO reconstruction performance, compared to the standard

\*iasota@ust.hk

method [7], particularly at a very low shot noise regime [18,19]. In detail, this method attempts to find the uniform Lagrangian frame by displacing each observed galaxy particle along the local density gradient, progressively reducing the smoothing scale. Once we achieve the almost homogenous mass distribution, we can estimate the displacement field by measuring the difference between the Eulerian and the estimated Lagrangian positions. If we can indeed recover the true nonlinear displacement field from such a method, the broadband shape of the resulting clustering could be modeled utilizing the perturbation theories of the displacement field (e.g., [4]), allowing a cosmological parameter extraction from the shape of the power spectrum in addition to the reconstructed BAO feature.

In Ref. [19], we constructed a theoretical model for the displacement reconstruction by Ref. [6]; there, we found that, while the method can reconstruct the BAO very well, it does not recover the shape of the true displacement field, and our theoretical model could not explain the deviation. Sources of the discrepancy were speculated, particularly including a numerical artifact due to the discreteness of the tracer sample for estimating the vector field.

This paper develops two new estimators of displacement fields that may reduce such discreteness effects in the tracers. The first estimator removes the effect of the tracer distribution, utilizing the fact that the Lagrangian positions of the (subsamped) particles are uncorrelated with the nonlinear displacement field. The second estimator asymptotically converts the displacement field to a density field.

We first test these estimators for measuring the true displacements from simulations (i.e., knowing the initial locations of the particles exactly) for various subsampling levels. While we perform this test as a sanity check before applying them to the reconstructed field, this application could be useful in estimating the nonlinear displacement field with sampling noise.

We then apply the estimators to the reconstructed displacement field and show that the discrepancy between the reconstructed and the true displacement fields remains even at  $k \sim 0.1 h \text{Mpc}^{-1}$  and it is physical, not due to a numerical artifact. We, therefore, show that the method from Ref. [6] cannot fully recover the true nonlinear displacement field near the redshift we studied ( $z = 0.6$ ), while it is still an effective BAO reconstruction method. By better securing the results against the discreteness effect, we improve the theoretical model for the reconstructed displacement field by properly accounting for the UV-sensitive term in a reconstruction procedure with an effective field theory (EFT) approach.

We organize this paper as follows. In Sec. II, we review the discreteness effect in galaxy surveys and pose a question about the displacement field estimators. Sections III and IV introduce two new displacement field estimators and give a theoretical background for them. A comparison of those new estimators and the previous

mass-weighted one is presented in Sec. V. Then we apply the estimators for the iterative reconstruction in Sec. VI, and we discuss the remaining inconsistencies in the post-reconstruction estimators, the true displacement field, and the one-loop perturbation theory modeling. The final section is devoted to the conclusions.

## II. DISCRETENESS EFFECT IN GALAXY SURVEYS

This section investigates the theoretical aspects of the discreteness effect in galaxy surveys. We first briefly review the mathematics for the discreteness effect in the galaxy field. Then, we illustrate the issue in the displacement field measurement. This is a problem with interpreting a vector field from measuring mass tracers.

### A. Galaxy number density field

Let  $\bar{n}_g$  be the average galaxy number density and  $V$  be the volume of a given three-dimensional pixel. Then the probability that the number of galaxies  $N$  found in the pixel follows a Poisson distribution whose average is  $\bar{n}_g V$ , i.e.,

$$e^{-\bar{n}_g V} \frac{(\bar{n}_g V)^N}{N!}. \quad (1)$$

Using the Poissonian random variable at the position  $\mathbf{x}$ , i.e.,  $N(\mathbf{x})$ , the galaxy number density field is given as

$$n_g(\mathbf{x}) = \frac{N(\mathbf{x})}{V(\mathbf{x})}. \quad (2)$$

The two-point correlation function of the number density is given by

$$\langle n_g(\mathbf{x}) n_g(\mathbf{y}) \rangle_{\text{Po}} = \left\langle \frac{N(\mathbf{x})}{V(\mathbf{x})} \frac{N(\mathbf{y})}{V(\mathbf{y})} \right\rangle_{\text{Po}}, \quad (3)$$

where the subscript ‘‘Po’’ means that the average is taken by the locally defined Poisson distribution of Eq. (1). For  $\mathbf{x} = \mathbf{y}$ , we get

$$\langle n_g(\mathbf{x})^2 \rangle_{\text{Po}} = \frac{\bar{n}_g}{V(\mathbf{x})}. \quad (4)$$

For  $\mathbf{x} \neq \mathbf{y}$ , the distributions are uncorrelated, and we find

$$\langle n_g(\mathbf{x}) n_g(\mathbf{y}) \rangle_{\text{Po}} = \langle n_g(\mathbf{x}) \rangle_{\text{Po}} \langle n_g(\mathbf{y}) \rangle_{\text{Po}} = \bar{n}_g^2. \quad (5)$$

To summarize, we derive [20]

$$\langle n_g(\mathbf{x}) n_g(\mathbf{y}) \rangle_{\text{Po}} = \frac{\delta_{\mathbf{x},\mathbf{y}}}{V(\mathbf{x})} (\bar{n}_g - V(\mathbf{x}) \bar{n}_g^2) + \bar{n}_g^2, \quad (6)$$

where  $\delta_{\mathbf{x},\mathbf{y}} = 1$  for  $\mathbf{x} = \mathbf{y}$ , and otherwise zero. In the small pixel limit [ $V(\mathbf{x}) \bar{n}_g \ll 1$ ], defining the Poisson noise  $\delta_g \equiv (n_g - \bar{n}_g) / \bar{n}_g$ , we get

$$\langle \delta_g(\mathbf{x})\delta_g(\mathbf{y}) \rangle_{\text{Po}} = \frac{1}{\bar{n}_g} \delta_D^{(3)}(\mathbf{x} - \mathbf{y}), \quad (7)$$

where  $\delta_{\mathbf{x},\mathbf{y}}/V(\mathbf{x}) \approx \delta_D^{(3)}(\mathbf{x} - \mathbf{y})$  is the three-dimensional Dirac's delta function. For simplicity, we assume that the galaxy distribution is linearly related to the underlying dark matter density fluctuation  $\delta_m$ . Then the number fluctuation due to the primordial density field is written as  $b\delta_m$ . We replace  $\bar{n}_g$  with the local number  $\bar{n}_g(1 + b\delta_m)$  when we normalize the Poisson distribution of Eq. (1), and we derive

$$\langle \delta_g(\mathbf{x})\delta_g(\mathbf{y}) \rangle_{\text{Po,G}} = \frac{\delta_D^{(3)}(\mathbf{x} - \mathbf{y})}{\bar{n}_g} + b^2 \langle \delta_m(\mathbf{x})\delta_m(\mathbf{y}) \rangle_G, \quad (8)$$

where ‘‘G’’ implies the Gaussian average of the primordial density perturbations. The first term comes from the Poisson shot noise, while the second term is the cosmological signal.

## B. Displacement field

Next, we consider the discreteness effect for the displacement field. For a given Eulerian coordinate  $\mathbf{x}$  and its associated Lagrangian coordinate  $\mathbf{q}$ , the displacement field is defined as

$$\mathbf{x} = \mathbf{q} + \Psi(\mathbf{q}). \quad (9)$$

This  $\Psi(\mathbf{q})$  in Eq. (9) is a continuous field, and in practice, we may interpret the displacement field by measuring the position of a test particle on the field. We assign the individual particle displacements to the grid in data analysis. Let  $V(\mathbf{q})$  be a pixel volume at  $\mathbf{q}$  and  $N(\mathbf{q})$  be the number of particles found in the pixel, which is a Poissonian variable defined at a Lagrangian position  $\mathbf{q}$ . One may evaluate the center of the mass displacement field

$$\tilde{\Psi}(\mathbf{q}) = \frac{\sum_{i=1}^{N(\mathbf{q})} \Psi(\mathbf{q}_i)}{N(\mathbf{q})}, \quad (10)$$

for the cell at  $\mathbf{q}$  where  $i$  is the particle label in the pixel. The issue is that Eq. (10) is ill-defined for empty pixels, i.e., for  $N(\mathbf{q}) = 0$ . One may consider interpolating the value at the empty pixels, e.g., based on the values at the neighboring pixels, but any *ad hoc* prescription to  $N = 0$  will introduce additional complexity in estimating the effect of the Poisson fluctuations on the displacement estimator. More importantly, the interpolation becomes quickly inefficient because most pixels (assuming a reasonable pixel resolution, e.g., 5 Mpc/h for  $\bar{n}_g = 0.001h^3/\text{Mpc}^3$ ) would be empty in a sparse system such as the galaxy field. In Refs. [18,19], we selectively set  $\tilde{\Psi} = 0$  for the empty pixel and attempt to correct the resulting large-scale effect by rescaling the clustering amplitude by a constant factor. However, we indeed found that there still is a residual discrepancy in the small-scale power that depends on the

sampling fraction (e.g., at the level of 6%–7% at  $k \sim 0.3 h \text{ Mpc}^{-1}$  in Fig. 2 between L500 and subL500). Thus, it was nontrivial to extract the unbiased displacement fields on small scales using the mass-weighted estimator (MWE) we adopted.

The MWE fails because the displacement field is irrespective of the occupation of the pixel. An empty pixel does not mean the displacement of the pixel is zero. Eulerian velocity estimators also suffer from a similar issue to the mass-weighted estimators. To our knowledge, Ref. [21] was the first to point out that mass-weighted velocity estimators return a biased estimator of the actual velocity field since the measurements relying on the mass tracer counting give the momentum rather than the velocity. They proposed two volume-weighted assignments, the Voronoi tessellation method and the Delaunay tessellation method for velocity measurements, which were applied for describing shell crossing in Ref. [22]. Furthermore, various volume weighted assignments are proposed [23–28]. On the other hand, giving up the displacement/velocity and considering the momentum field can also be an option [29–32]. An advantage of the momentum field is that the sampling issue is solved as momentum is, correctly, zero without a mass tracer. However, higher-order effects such as the galaxy bias would complicate estimating the true momentum field. While the displacement field discreteness effects are similar to the velocity, we focus on the properties specific to the Lagrangian perspective to find a solution to our issue. The following sections introduce new displacement estimators that can reduce the discreteness effect.

As a caveat, while we first test our new estimators with the true displacement fields in a simulation, we note that the displacement field is not directly observable in real surveys since we only measure the final Eulerian position of each galaxy. Therefore, we are interested in ‘‘reconstructing’’ the Lagrangian position from the observed Eulerian position, i.e., the displacement field. Our new estimators are developed to interpret the reconstructed displacement field properly. The reconstruction scheme to find the Lagrangian frame itself was also discussed in Refs. [6,18,19], and we review the idea in Sec. VI.

## III. NORMALIZED MOMENTUM ESTIMATOR

This section introduces a new momentum estimator and shows how one can normalize it to obtain a volume-weighted displacement field estimator to a good approximation.

### A. Definition

We propose to compute a normalized momentum estimator (NME)

$$\tilde{\xi}_{ij}(\mathbf{q} - \mathbf{r}) \equiv \frac{\left\langle \sum_{a=1}^{N(\mathbf{q})} \sum_{b=1}^{N(\mathbf{r})} \Psi_i(\mathbf{q}_a) \Psi_j(\mathbf{r}_b) \right\rangle}{\langle N(\mathbf{q})N(\mathbf{r}) \rangle}, \quad (11)$$

and we will show that  $\tilde{\xi}_{ij} \rightarrow \xi_{ij} \equiv \langle \Psi_i(\mathbf{q}) \Psi_j(\mathbf{r}) \rangle$  for the small pixel limit below. With this approximation, we can avoid the ill-defined mass-weighted displacement in Eq. (10). In Eq. (11), one first computes the numerator, which is the correlation function of a momentumlike quantity, and then we normalize the correlation function by the density correlation function. This would allow us to construct an approximately volume-weighted estimator without, e.g., the Delaunay tessellation method.

Let us prove  $\tilde{\xi}_{ij} \rightarrow \xi_{ij}$  for the  $V \rightarrow 0$  limit. For simplicity, the particles are Poisson sampled in the uniform Lagrangian frame. The bracket of Eq. (11) means that we take both the Poisson and ensemble averages.  $N(\mathbf{q})$  and  $N(\mathbf{r})$  in Eq. (11) are independent of the ensemble average in Lagrangian space so that we may exclusively take the ensemble average for the displacement field.  $\xi$  is a function of the distance due to the statistical isotropy and homogeneity. Then Eq. (11) is expanded into

$$\tilde{\xi}_{ij}(\mathbf{q} - \mathbf{r}) = \sum_{n=0}^{\infty} \mathcal{F}_n(\mathbf{q}, \mathbf{r}) \partial_x^n \xi_{ij}(x) \Big|_{x=|\mathbf{q}-\mathbf{r}|}, \quad (12)$$

where we used the Taylor expansion

$$\xi_{ij}(\mathbf{q}_a - \mathbf{r}_b) = \sum_{n=0}^{\infty} \frac{\epsilon_{ab}^n}{n!} \partial_x^n \xi_{ij}(x) \Big|_{x=|\mathbf{q}-\mathbf{r}|}, \quad (13)$$

$$\epsilon_{ab} \equiv |\mathbf{q}_a - \mathbf{r}_b| - |\mathbf{q} - \mathbf{r}|, \quad (14)$$

and then defined

$$\mathcal{F}_n(\mathbf{q}, \mathbf{r}) \equiv \left\langle \sum_{a=1}^{N(\mathbf{q})} \sum_{b=1}^{N(\mathbf{r})} \right\rangle_{\text{Po}}^{-1} \left\langle \sum_{a=1}^{N(\mathbf{q})} \sum_{b=1}^{N(\mathbf{r})} \frac{\epsilon_{ab}^n}{n!} \right\rangle_{\text{Po}}. \quad (15)$$

We cannot explicitly evaluate  $\mathcal{F}_n$  because we never know the implicit Poissonian dependence of  $\epsilon_{ab}$ , i.e., the Poissonian dependence of the particle positions. However, we can find the upper bound easily. An example configuration for  $\mathbf{q}$ ,  $\mathbf{q}_a$ ,  $\mathbf{r}$ , and  $\mathbf{r}_b$  is illustrated in Fig. 1. Let  $V(\mathbf{q})$  be a cube and  $L = V^{1/3}$ . The maximum distance between particles in the same box is  $\sqrt{3}L$ , so we get

$$|\mathcal{F}_n| \leq \frac{(\sqrt{3}L)^n}{n!}. \quad (16)$$

Then one finds

$$|\tilde{\xi}_{ij}(|\mathbf{q} - \mathbf{r}|) - \xi_{ij}(|\mathbf{q} - \mathbf{r}|)| \leq \sum_{n=1}^{\infty} \frac{(\sqrt{3}L)^n}{n!} |\partial_x^n \xi_{ij}(x)|_{x=|\mathbf{q}-\mathbf{r}|}. \quad (17)$$

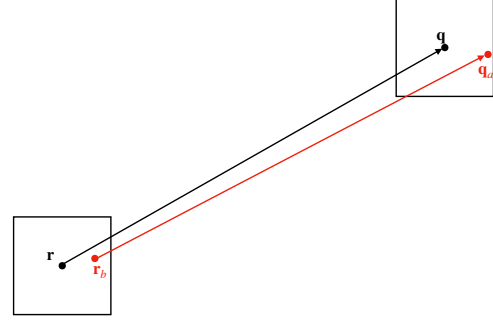


FIG. 1. An example configuration for  $\mathbf{q}$ ,  $\mathbf{q}_a$ ,  $\mathbf{r}$ , and  $\mathbf{r}_b$ .  $\epsilon_{ab} \equiv |\mathbf{q}_a - \mathbf{r}_b| - |\mathbf{q} - \mathbf{r}|$  is within the size of a box.

We can take a sufficiently small  $V$  such that  $L^n \partial^n \xi_{ij} \rightarrow 0$ , as long as  $\partial^n \xi_{ij}$  is finite.

The above proof is applicable as particle sampling is independent of the primordial fluctuations, so the distribution of particles does not have to be Poissonian as long as it is independent of the ensemble average. Interestingly, the mass window is canceled in Eq. (11) as we have  $\mathcal{F}_0 = 1$  so that we obtain the volume-weighted correlation function, the shot noise is zero, and the final result is independent of the number density  $\bar{n}_g$ . We would not claim that the covariance vanishes, but the discreteness effect vanishes. The above proof is not valid for a singular  $\xi$ .

Equation (11) is different from the nominal momentum estimator in Ref. [32] that corresponds to

$$\left\langle \sum_{a=1}^{N(\mathbf{q})} \sum_{b=1}^{N(\mathbf{r})} \Psi_i(\mathbf{q}_a) \Psi_j(\mathbf{r}_b) \right\rangle. \quad (18)$$

Equation (18) is a mass-weighted quantity without the normalization in the denominator of Eq. (11) and therefore does not correspond to a volume-weighted estimator.

In Eulerian space, e.g., for the late-time velocity field, we cannot exclusively take the ensemble average in Eq. (11). This is because the Poisson distribution depends on the local stochastic variable, as discussed in Sec. II A. Hence, the bispectrum or higher-order cumulant appears and cannot be canceled. Therefore, the above proof only applies to the Eulerian velocity field at leading order perturbations.

Another crucial remark is that because the new estimator is effectively volume weighted, we expect that Eq. (11) is a promising estimator for very sparse samples, such as a biased galaxy tracer. In our future work, we plan to investigate how well such an estimator of the galaxy field would relate to the matter displacement field as a function of scale.

## B. Numerical implementation

Below, we compare the mass-weighted estimator used in Ref. [19] with the operation conducted for the NME in this paper.

The former mass-weighted displacement was defined as

$$\Psi_p^{\text{obs}} = \frac{\sum_i W_{\text{CIC}}(\mathbf{x}_p, \mathbf{x}_i) \Psi^{\text{obs}}(\mathbf{x}_i)}{\sum_i W_{\text{CIC}}(\mathbf{x}_p, \mathbf{x}_i)}, \quad (19)$$

which corresponds to Eq. (10). Here  $\Psi^{\text{obs}}(\mathbf{x}_i)$  can be either the true displacement or the iteratively reconstructed displacement of each mass tracer labeled by  $\mathbf{x}_i$ .  $W_{\text{CIC}}$  is the pixel window function indicating that we are using the cloud-in-cell assignment, and  $\Psi_p^{\text{obs}}$  is the estimator assigned at a pixel centered at  $\mathbf{x}_p$ .

For NME, we have the same catalog of  $\Psi^{\text{obs}}(\mathbf{x}_i)$ , but we directly evaluate its power spectrum without explicitly introducing the displacement field estimators like Eq. (19). First, we estimate the momentum field

$$\mathbf{P}_p^{\text{obs}} = \sum_i W_{\text{CIC}}(\mathbf{x}_p, \mathbf{x}_i) \Psi^{\text{obs}}(\mathbf{x}_i). \quad (20)$$

Then, we consider fast Fourier transformation (FFT) of the momentum field  $\hat{\mathbf{P}}_p^{\text{obs}}$  and evaluate the momentum power spectrum  $\langle \hat{\mathbf{P}}_p^{\text{obs}}(\mathbf{k}_1) \cdot \hat{\mathbf{P}}_p^{\text{obs}}(\mathbf{k}_2) \rangle$ . Then, we compute the correlation function of the momentum field  $\xi_p$  from the momentum power spectrum. We also calculate the correlation function  $\xi_n$  of the number density field (i.e., the discreteness selection function):

$$n_p^{\text{obs}} = \sum_i W_{\text{CIC}}(\mathbf{x}_p, \mathbf{x}_i). \quad (21)$$

Then, we compute  $\xi_p/\xi_n$  to convert the momentum field two-point statistics to the displacement field statistics, and finally, Fourier transform it to find the displacement field divergence spectrum. We ignored the curl component for simplicity so we have  $(\mathbf{k} \cdot \Psi)^2 = k^2 \Psi \cdot \Psi$ . As emphasized in the previous section, because  $\mathbf{x}_i$  is in the Lagrangian space for the true displacement field or presumably close to the true Lagrangian space in the reconstructed case, we can ignore the higher order cumulants in  $\xi_n$  and  $\xi_p$ . Again, the same operation would not be justified in calculating a late-time field where  $\mathbf{x}_i$  is not close to the Lagrangian space, e.g., the Eulerian velocity field. To compare the density power spectrum, we use  $k^2 P_\Psi$  rather than  $P_\Psi$ . In Sec. V, we compare the NME and the exact displacement field.

#### IV. RESCALED RESUMMATION ESTIMATOR

This section describes another displacement field estimator for a sparse sample we propose in this paper, which we call the rescaled resummation estimator (RRE). We measure the displacement field divergence power spectrum as a rescaled density power spectrum without explicitly evaluating the vector field on the grid; therefore, we can avoid the issue of the displacement field assignment with sparse tracers.

#### A. Lagrangian resummation theory

Let us consider the density fluctuation exponentiated by the rescaled displacement field  $\Psi_{\text{NL}}/\Lambda$ :

$$\delta_{\text{NL}}^\Lambda \equiv \int d^3q e^{-i\mathbf{k} \cdot \mathbf{q}} (e^{-i\mathbf{k} \cdot \frac{\Psi_{\text{NL}}(\mathbf{q})}{\Lambda}} - 1). \quad (22)$$

For  $\Lambda = 1$ , this equation recovers a known integral representation of the density perturbation in Lagrangian perturbation theory (LPT). In simulations, Eq. (22) corresponds to measure the density fluctuation after moving the particles from their initial/Lagrangian locations by  $\Psi_{\text{NL}}/\Lambda$ . Taylor expanding Eq. (22), we get

$$\Lambda \delta_{\text{NL}}^\Lambda = i\mathbf{k} \cdot \Psi_{\text{NL}} + \mathcal{O}(\Lambda^{-1}), \quad (23)$$

and we get the exact nonlinear displacement field in the  $\Lambda \rightarrow \infty$  limit. From Lagrangian resummation theory [5], the resummed power spectrum of Eq. (22) is given as

$$\Lambda^2 P_{\delta_{\text{NL}}^\Lambda} = \exp\left(-\frac{k^2 \int p^2 dp P_{\Psi_{\text{NL}}} + \mathcal{O}(\Lambda^{-3})}{6\pi^2 \Lambda^2}\right) \times (k^2 P_{\Psi_{\text{NL}}} + \mathcal{O}(\Lambda^{-1})). \quad (24)$$

The correction terms are higher-order cumulants in  $\Psi_{\text{NL}}$ , which are relatively suppressed to the leading order power spectrum since they carry additional negative powers of  $\Lambda$ . The exponential damping is the degradation effect due to the dark matter displacement from the initial BAO configurations. After rescaling (22), the damping effect is reduced thanks to the  $\Lambda^2$  in the denominator of the exponential and the estimator asymptotes to the displacement field power spectrum. We obtain the displacement field divergence power spectrum without explicitly evaluating vector fields and the power spectrum as the density power spectrum by taking a large  $\Lambda$  limit.

#### B. Noise modeling

A large  $\Lambda$  will improve the recovery of the displacement field power spectrum (and the BAO feature in it), but multiplying large  $\Lambda$  may also amplify the noise in the density field. This subsection illustrates the potential issue and how to mitigate some effects. A possible error comes from the uncertainty  $\chi$  in the estimated Lagrangian position during reconstruction; i.e., the estimated Lagrangian position should be written as  $\mathbf{q} + \chi$  rather than  $\mathbf{q}$ . Because of this error, the estimated displacement field is also shifted by the same amount. Therefore, the rescaled particle location will be replaced as

$$\mathbf{q} + \frac{\Psi_{\text{NL}}}{\Lambda} \rightarrow \mathbf{q} + \chi + \frac{\Psi_{\text{NL}} - \chi}{\Lambda}. \quad (25)$$

We cannot isolate  $\chi$  from the observed data, so the rescaling happens only for  $\Psi_{\text{NL}} - \chi$ . Nonvanishing  $\chi$  is generally

inevitable when reconstructing the displacement field from actual data. Equation (22) is generalized to

$$\delta_{\text{NL}}^{\Lambda} \rightarrow \int d^3q e^{-i\mathbf{k}\cdot\mathbf{q}} \left[ e^{-i\mathbf{k}\cdot\left(\frac{(\Lambda-1)\chi(\mathbf{q})}{\Lambda} + \frac{\Psi_{\text{NL}}(\mathbf{q})}{\Lambda}\right)} - 1 \right]. \quad (26)$$

The  $\chi$  term in the above equation is not scaled for a large  $\Lambda$ , so the noise term is relatively amplified after multiplying  $\Lambda$  in Eq. (23). The resummed power spectrum is given as

$$\Lambda^2 P_{\delta_{\text{NL}}^{\Lambda}} \rightarrow \exp \left[ -\frac{k^2 \int p^2 dp}{6\pi^2} \left( \frac{P_{\Psi_{\text{NL}}}}{\Lambda^2} + \frac{(\Lambda-1)^2}{\Lambda^2} P_{\chi} \right) \right] \times k^2 [P_{\Psi_{\text{NL}}} + 2(\Lambda-1)P_{\chi\Psi_{\text{NL}}} + (\Lambda-1)^2 P_{\chi}], \quad (27)$$

where we ignored  $\mathcal{O}(\chi^3, \Lambda^{-1})$  terms.  $\chi$  and  $\Psi_{\text{NL}}$  are generally correlated. Equation (27) implies that the RRE works well without amplifying noise when we have the following relation:

$$(\Lambda-1)^2 \ll \frac{P_{\Psi_{\text{NL}}}}{P_{\chi}}, \quad \Lambda-1 \ll \frac{P_{\Psi_{\text{NL}}}}{P_{\chi\Psi_{\text{NL}}}}, \quad (28)$$

which will not be satisfied for large  $\Lambda$  or large inhomogeneities in  $\chi$ . How can we then mitigate the amplification of the  $\chi$  contribution? We provide a simple method as follows. By setting  $\Lambda = \infty$  we can estimate the noise field as

$$\delta_{\text{NL}}^{\infty} = \int d^3q e^{-i\mathbf{k}\cdot\mathbf{q}} (e^{-i\mathbf{k}\cdot\chi} - 1). \quad (29)$$

This is the nonvanishing inhomogeneity in the estimated Lagrangian space, as reconstruction is not perfect. In our method, the error field can be approximately estimated from the final density field of the galaxies after a series of iterative reconstructions. We can therefore subtract the noise field to define

$$\hat{\delta}_{\text{NL}}^{\Lambda} \equiv \delta_{\text{NL}}^{\Lambda} - \delta_{\text{NL}}^{\infty}. \quad (30)$$

The corrected RRE is expanded into

$$\Lambda \hat{\delta}_{\text{NL}}^{\Lambda} = i\mathbf{k} \cdot (\Psi_{\text{NL}} - \chi) + \mathcal{O}(\Lambda^{-1}). \quad (31)$$

Thus, the positive power of  $\Lambda$  is removed so we can avoid amplifying the noise term, whereas the noise  $\chi$  is inevitable.<sup>1</sup> The resummed power spectrum has the following form:

<sup>1</sup>We could have chosen to subtract by  $(1-1/\Lambda)\delta_{\text{NL}}^{\infty}$  in Eq. (30) to cancel out the effect of  $\chi$  more efficiently. We find that with that option, by boosting the small-scale contribution slightly due to the factor  $(1-1/\Lambda)$ , the method reaches a better reconstruction at  $n < 6$  while the agreement with respect to the theory model and the NME is worse. For  $n > 8$  we find this choice also converges to NME.

$$\Lambda^2 P_{\delta_{\text{NL}}^{\Lambda}} = \exp \left[ -\frac{k^2 \int p^2 dp}{6\pi^2} P_{\chi} \right] \times k^2 [P_{\Psi_{\text{NL}}} - 2P_{\chi\Psi_{\text{NL}}} + P_{\chi} + \mathcal{O}(\chi, \Lambda^{-1})]. \quad (32)$$

Now, the condition that  $\Lambda \hat{\delta}_{\text{NL}}^{\Lambda}$  converges to the displacement is independent from  $\Lambda$ , and Eq. (28) reduces to

$$1 \ll \frac{P_{\Psi_{\text{NL}}}}{P_{\chi}}, \quad 1 \ll \frac{P_{\Psi_{\text{NL}}}}{P_{\chi\Psi_{\text{NL}}}}. \quad (33)$$

The subtraction in Eq. (30) bears a resemblance to the operation in the standard BAO reconstruction estimator as well as the standard iterative scheme introduced in Ref. [33], which is written as the difference of the shifted reference density perturbation (denoted as  $\delta_s$  in Ref. [33]) and the displaced galaxy density perturbation ( $\delta_d$  in Ref. [33]).  $\delta_d$  in standard/iterative reconstruction adds reconstructed small-scale information; likewise, we find that in RRE, subtracting with  $\delta_{\text{NL}}^{\infty}$  (i.e.,  $\delta_d$ ) recovers the small-scale clustering by removing most of the effect of  $\chi$ . We assume this procedure also subtracts the shot noise contribution to a good extent and therefore does not apply a separate shot-noise subtraction.

## V. NEW ESTIMATORS IN $N$ -BODY SIMULATIONS

We proposed the two new displacement field estimators in the previous sections and will now assess their performance in numerical simulations. As mentioned earlier, we want to reconstruct the displacement field from actual galaxy surveys and apply the estimators for the post-reconstruction data. In this section, we first test the methods by deriving the true displacement field (i.e., the difference between the initial Lagrangian positions and the final Eulerian positions in the simulations); we check if these estimators have advantages in mitigating the discreteness and/or subsampling effect. We will then consider the post-reconstruction data in the next section.

This paper uses two different simulations from Ref. [19] for different purposes. We briefly summarize the parameters of these simulations. First, we focus on dark matter simulation rather than galaxies or halos for simplicity. Both simulations are based on the flat  $\Lambda$ CDM cosmology in Ref. [34] with  $\Omega_m = 0.3075$ ,  $\Omega_b h^2 = 0.0223$ ,  $h = 0.6774$ , and  $\sigma_8 = 0.8159$ . Full  $N$ -body simulations were produced using the MP-Gadget [35–37] with the box size of 500 Mpc/ $h$  and 1500 Mpc/ $h$ , and the simulations evolve  $1536^3$  particles from  $z = 99$  by computing forces in a grid of  $1536^3$ . For the former 500 Mpc/ $h$  simulation, we average five realizations and use only 4% of the dark matter particles at  $z = 0.6$  for the dataset named L500 and 0.15% for subL500 in Table I. We use a grid of  $512^3$  to Fourier transform and reconstruct this nonlinear field for L500 and subL500. These datasets are prepared to test the robustness of the estimators: L500 has almost no empty FT grids, but the subsampled particles are not necessarily in the

TABLE I. Simulations and sampling parameters used in this paper. The simulations assume a flat  $\Lambda$ CDM cosmology in Ref. [34] ( $\Omega_m = 0.3075$ ,  $\Omega_b h^2 = 0.0223$ ,  $h = 0.6774$ , and  $\sigma_8 = 0.8159$ ).

Name	Subsampling %	Number of meshes	Box size [Mpc/h] <sup>3</sup>	Number of original particles	Number of simulations
fullL1500	100	1536 <sup>3</sup>	1500 <sup>3</sup>	1536 <sup>3</sup>	1
L500	4	512 <sup>3</sup>	500 <sup>3</sup>	1536 <sup>3</sup>	5
subL500	0.15	512 <sup>3</sup>	500 <sup>3</sup>	1536 <sup>3</sup>	5
subsubL500	0.015	512 <sup>3</sup>	500 <sup>3</sup>	1536 <sup>3</sup>	5

center of each mesh, while subL500 represents a sparse sample with 96% of grids/meshes being empty. We prepare the latter 1500 Mpc/h box simulation to generate the reference displacement field that we called the “true displacement field.” We name the set fullL1500, which samples all particles in the center of each grid in Lagrangian space. Therefore, there is no empty pixel in the set, and the displacement measured in each grid is straightforwardly interpreted as the displacement field without sampling noise as a volume-weighted displacement.

In this work we compute the auto power spectrum of the displacement field divergence  $P_{ik,\Psi}$ , and the cross-power spectrum of the displacement field divergence and the linear matter fluctuation,  $P_{\delta_L, ik,\Psi}$ . Both are normalized by the linear matter power spectrum  $P_L$ . We also refer  $P_{\delta_L, \nabla, \Psi} / P_L$  to the propagator  $C_{ik,\Psi}$ . As we have only 1 or 5 realizations for simulations, it is difficult to derive reliable error bars. We therefore maximally utilize the variance cancellation by comparing propagators and power spectra relative to their corresponding initial conditions when presenting the results.

In Figs. 2 and 3, we show the displacement field power spectrum measured in L500, subL500, and fullL1500, normalized by the linear matter power spectrum for various estimators. The NME and MWE converge for  $k \lesssim 0.2 h \text{ Mpc}^{-1}$  within 1%, but the MWE starts to depart from the true displacement field at higher  $k$ . The NME (Fig. 2) is more robust than the MWE. A few percent convergences are extended to  $k < 0.5 h \text{ Mpc}^{-1}$ ; the solid green (L500) and dotted green (subL500) show much better consistency with respect to the true displacement field (solid black) even though subL500 has 96% of its grid cells empty, compared to MWE. We also add subsubL500 for the NME in the same figure (faint dashed), which corresponds to  $n_{\text{particle}} = 0.0043 h^3 \text{ Mpc}^{-3}$  (i.e., 99.6% of the meshes we used were empty). Despite the high sampling noise, it still shows the convergence to the true displacement field in both the power spectrum and the propagator.

For the RREs (Fig. 3),  $\Lambda = 1$  shown in the purple line returns the true nonlinear density field as expected from Eq. (22), as a sanity check. With increasing  $\Lambda$ , RRE approaches the true displacement field with decreasing BAO damping. Note that with  $\Lambda = 10$ , the propagator already recovered the propagator of the true displacement.

In Fig. 3, the convergence of  $\Lambda = 100$  and  $\Lambda = 200$  implies that the correction of  $\mathcal{O}(\Lambda^{-1})$  is negligible for  $\Lambda \gtrsim 100$  at least. However, both curves do not converge to the true displacement field or NME for  $k \gtrsim 0.2 h \text{ Mpc}^{-1}$ . As a caveat, we know the initial Lagrangian locations of the particles exactly in this test, and therefore  $\chi = 0$  in Eq. (25). However, the RREs and true displacement field still disagree. The disagreement is manifest even for large  $\Lambda$ , which implies that errors are described by neither  $\Lambda$  nor  $\chi$ . The discrepancies in L500 should be different from the sampling noise since almost one tracer per mesh is observed in these data. The solid pink curve in Fig. 3 shows RRE with  $\Lambda = 100$  when we assign the rescaled displacement at the *center* of the FFT grid

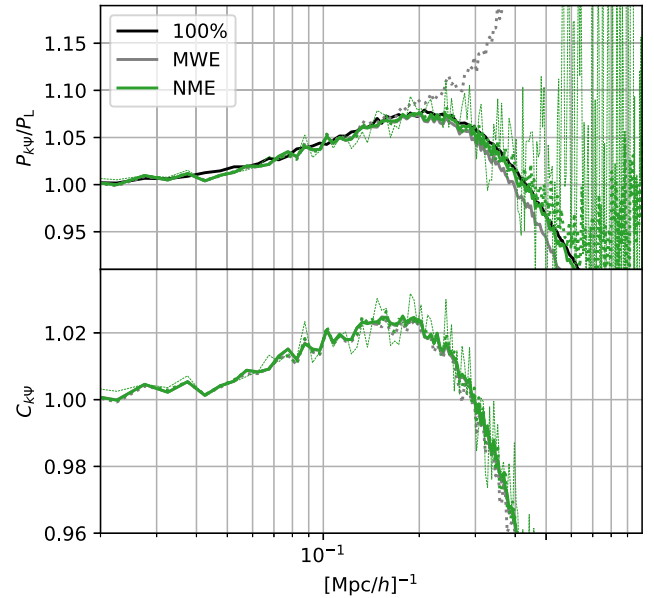


FIG. 2. The displacement field divergence power spectra (top) and propagators (bottom) of the MWE and NME for various sampling fractions. The linear matter power spectrum normalizes the power spectrum, and the propagator is defined as the cross-power spectrum normalized by the linear matter power spectrum. The figure is based on the  $N$ -body simulation L500 (solid line) and subL500 (dotted line) summarized in Table I. We also add subsubL500 for the NME (faint dashed line) that corresponds to  $n_{\text{particle}} = 0.0043 h^3 \text{ Mpc}^{-3}$ , which still shows the convergence to the true displacement field in both panels despite the high sampling noise. The black solid curve corresponds to the fullL1500.

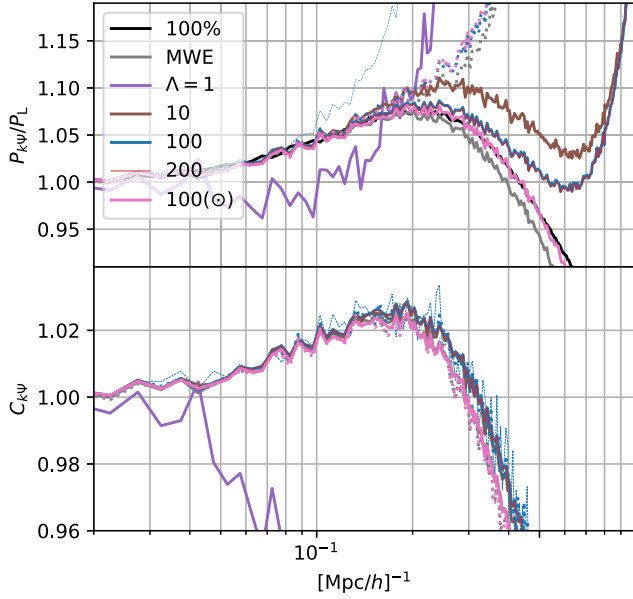


FIG. 3. The displacement field divergence power spectra (top) and propagators (bottom) of the MWE and RRE for various sampling fractions. The linear matter power spectrum normalizes the power spectrum, and the propagator is defined as the cross-power spectrum normalized by the linear matter power spectrum. The figure is based on the  $N$ -body simulation L500 (solid line) and subL500 (dotted line) summarized in Table I. The black solid curve corresponds to the fullL1500. The violet line reproduces the nonlinear field with  $\Lambda = 1$ .  $\odot$  implies that we assigned the displacement fields at the center of each pixel. We again add subsubL500 for the RRE  $\Lambda = 100$  (faint blue dashed line) that corresponds to  $n_{\text{particle}} = 0.0043 h^3 \text{Mpc}^{-3}$ . Contrary to NME, the top panel shows a residual contribution from the sampling noise, while the propagator fairly well converged to the true displacement field.

where the initial position of each particle falls, rather than at the actual initial particle position. This manipulation introduces  $\chi \neq 0$ , but the deviation from the true displacement is removed for L500. Thus, a part of the unknown error is reduced by this procedure. In more subsampled cases, the pink and blue dotted curves (subL500) show a similar noise, which can now be understood as the shot noise. Thus, centering seems to reduce some of the unknown errors but cannot reduce the sampling noise. We note that this empirical observation is missing the theoretical explanation. It is not obvious how to apply this “centering” on a highly subsampled case, so the centering operation is not necessarily useful. Thus, we conclude that the NME is more robust on scales  $k > 0.2 h \text{Mpc}^{-1}$  than the MWE approach used in Refs. [6, 19], while the RRE requires further investigation for the error correction.

## VI. POST-RECONSTRUCTION ESTIMATORS

So far, we have discussed how to assign given particle displacements to the grid and constructed the displacement

field estimators robust to the discreteness effect. In Figs. 2 and 3, we tested our proposed methods for estimating the true displacement field of particles (i.e., in simulations, we knew all initial particle positions *a priori* and compute the particle displacements by subtracting the initial positions from the final positions). However, this is not the case in real data. We only know the final locations of galaxies in the actual surveys, so we need to reconstruct the displacement field without prior knowledge of the initial mass locations.

### A. Displacement field reconstruction

A displacement field reconstruction method from the observed mass locations is proposed in Ref. [6]. They proposed to move the mass tracers along the smoothed local density gradient iteratively by progressively reducing the smoothing scale. After several iterations, they found that the mass configuration becomes almost uniform, and the final locations can be interpreted as estimated Lagrangian positions for each mass. In this section, we review the mathematical aspect of the algorithm and apply the new estimators NME and RRE on the iteratively reconstructed field.

To reconstruct the displacement field from the observed density field and the mass tracer locations, in principle, we have to solve

$$\delta_{\text{NL}}(\mathbf{k}) = \int d^3q e^{-i\mathbf{k}\cdot\mathbf{q}} (e^{-i\mathbf{k}\cdot\boldsymbol{\Psi}_{\text{NL}}(\mathbf{q})} - 1), \quad (34)$$

for  $\boldsymbol{\Psi}_{\text{NL}}$ . However, this equation is a complicated nonlinear integral equation, which we cannot solve exactly.<sup>2</sup> Let us consider the following eigenfunction decomposition of the perturbation on the flat Friedmann Lemaître Robertson Walker background:

$$\boldsymbol{\Psi}_{\text{NL}} = i\mathbf{k}\phi_{\text{NL}} + \boldsymbol{\beta}, \quad (35)$$

where  $\nabla \cdot \boldsymbol{\beta} = 0$ . Expanding Eq. (34) to the leading order in  $\phi_{\text{NL}}$ , we get

$$\delta_{\text{NL}} \simeq k^2 \phi_{\text{NL}} + R, \quad (36)$$

where the residual  $R$  is the higher order terms in  $\phi_{\text{NL}}$  and  $\boldsymbol{\beta}$  is shown to be small in Ref. [4]. The inverse for Eq. (36) is easy if we can ignore  $R$ . However, the Zel’dovich approximation  $\delta_{\text{NL}} \simeq k^2 \phi_{\text{NL}}$  is valid only for low  $k$ . The iterative reconstruction assumes that there exists a cutoff scale  $k_{\text{cut}}$  such that  $\delta_{\text{NL}} \simeq k^2 \phi_{\text{NL}}$ , for all  $k < k_{\text{cut}}$ , and then we introduce the smoothed negative displacement

$$\mathbf{s} = -\frac{i\mathbf{k}}{k^2} S\delta_{\text{NL}}, \quad (37)$$

<sup>2</sup>Note that the displacement field divergence is not the log normal field:  $\ln[1 + \delta_{\text{NL}}(\mathbf{q})] \neq -\nabla_{\mathbf{q}} \cdot \boldsymbol{\Psi}_{\text{NL}}(\mathbf{q})$ .

where  $S = \exp(-k^2/2k_c^2)$  such that  $SR \approx 0$ . This is equivalent to solving the smoothed linearized continuity equation, as we may identify the time derivative of the displacement field with the velocity. Then we shift the particles from  $\mathbf{x}$  to  $\mathbf{x} + \mathbf{s}(\mathbf{x})$ . Thus, we partly estimated the Lagrangian position for  $k < k_{\text{cut}}$  modes. The new shifted frame would be closer to the uniform Lagrangian frame, so the nonlinearity would be reduced there. The new cutoff scale can be bigger;  $k_{\text{cut,new}} > k_{\text{cut}}$ . Then we reestimate the negative displacement for the reduced density perturbations by reducing the smoothing scale as  $k_{\text{cut}} \rightarrow k_{\text{cut,new}} = \epsilon k_{\text{cut}}$  with  $\epsilon > 1$  in Eq. (37). We repeat this cycle until we derive almost zero  $n$ th step displacement. The shifted frame is given recursively as

$$\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} + \mathbf{s}^{(n)}(\mathbf{x}^{(n)}), \quad (38)$$

and the final location  $\mathbf{x}^{(\infty)}$  will be close to the uniform density frame, which is the estimated Lagrangian position. In this way, we find a ‘‘coordinate transformation’’ from the Eulerian to Lagrangian coordinate in the simulation. In the iteration, Eq. (36) is generalized to

$$\delta_{\text{NL}}^{(n)} = k^2 \phi_{\text{NL}}^{(n)} + R^{(n)}, \quad (39)$$

where the superscript  $n$  implies a step of iteration; that is,  $\delta_{\text{NL}}^{(n)}$  is the density perturbation measured in the  $n$ th step particle distribution, and  $\phi_{\text{NL}}^{(n)}$  is the  $n$ th step displacement field. We attempt to reduce the nonlinearity in the input mass distribution by solving the linear algebra iteratively. References [6,18,19] confirmed that the BAO damping is significantly reduced by reconstructing the displacement field in this way.

Our new estimators reduce the potential discreteness effect in evaluating the reconstructed displacement field. We follow the same iterative reconstruction process as in Ref. [19]. The location of the mass tracers after each iteration is identical among the estimators we are testing in this paper.

## B. Discrepancy between true and reconstructed displacement fields

In our earlier work, we also investigated to what extent the iterative procedure reconstructs the true displacement field in the broadband shape of the power spectrum [19]. In that work, we saw about 8% discrepancies between the true displacement field measured in the same simulation and the reconstructed displacement field measured with the MWE estimator for  $k > 0.1 h \text{ Mpc}^{-1}$ . We suspected the inconsistency was partly due to the discreteness effect in the MWE estimator. In Fig. 4, we compare the MWE, NME, and RRE estimators of the reconstructed displacement field to revisit this inconsistency since we expect the new estimators are more robust against the discreteness effect.

The solid lines are for L500, and the dotted lines are for subL500. Although we expect some sampling noise of the observed field [38], MWE shows a drastic difference between the two subsamplings rather even though the shot noise of subL500 is still quite negligible, as pointed out in Ref. [19]. Our new estimators, NME and RRE are more stable for subsampling than MWE is. As we progress to  $n > 6$ , we begin to see differences in the power spectrum on small scales due to the noise of the field affecting reconstruction and also the difference between RRE and NME; because the power spectrum is divided by the linear matter power spectrum, a small residual shot noise will appear significant in this plot. On the other hand, the propagator shows the expected consistency between these two subsamplings when using NME and RRE. The dashed lines show subsubL500, which corresponds to  $n_{\text{particle}} = 0.0043 h^3 \text{ Mpc}^{-3}$ , i.e., closer to a realistic dense galaxy sample such as from the Bright Galaxy Survey in Dark Energy Spectroscopic Instrument [39]. The reconstruction efficiency of this subsample is expected to be noticeably lower than L500 and subL500, and the small scale will be dominated by shot noise (a typical particle separation of  $6.15 h^{-1} \text{ Mpc}$ ). Indeed, the maximum efficiency for this sample happens at  $n \sim 5-6$ , where the smoothing scale of the step ( $3.5 - 6 h^{-1} \text{ Mpc}$ ) is close to this average particle spacing, and a further iteration with a smaller smoothing scale decreases the efficiency.

While robustness against sampling/discreteness is improved, we still found inconsistency with respect to the true displacement field (gray), even on the quasilinear scale,  $k \sim 0.1 h \text{ Mpc}^{-1}$ , where the estimators are converging toward each other. The difference between NME and RRE is small at that scale, so we can now interpret the discrepancy with respect to the true displacement field as an indication that the iterative displacement field reconstruction cannot recover the true displacement field even on the quasilinear scale perfectly, contrary to our theory expectation in Ref. [19]. We still recover the initial density information. The propagator in the lower panel of the figure shows that iterative reconstruction is an efficient density field reconstruction method, but it does not fully recover the nonlinear displacement field in the broadband.

Based on this result we argue that the iterative reconstruction method in Ref. [6] does not reproduce the true displacement field. Introducing 2LPT or higher-order corrections for Eq. (36) could improve the agreement with respect to the true displacement. Reference [3] implemented 2LPT as an extension to the standard reconstruction, but found a minor improvement in the reconstructed density field. In the current case, the combination of the iterative steps and the 2LPT focusing on the displacement field is worth investigating, since the iterative process extends to the higher  $k$  than the standard BAO reconstruction. We leave this investigation for future work.

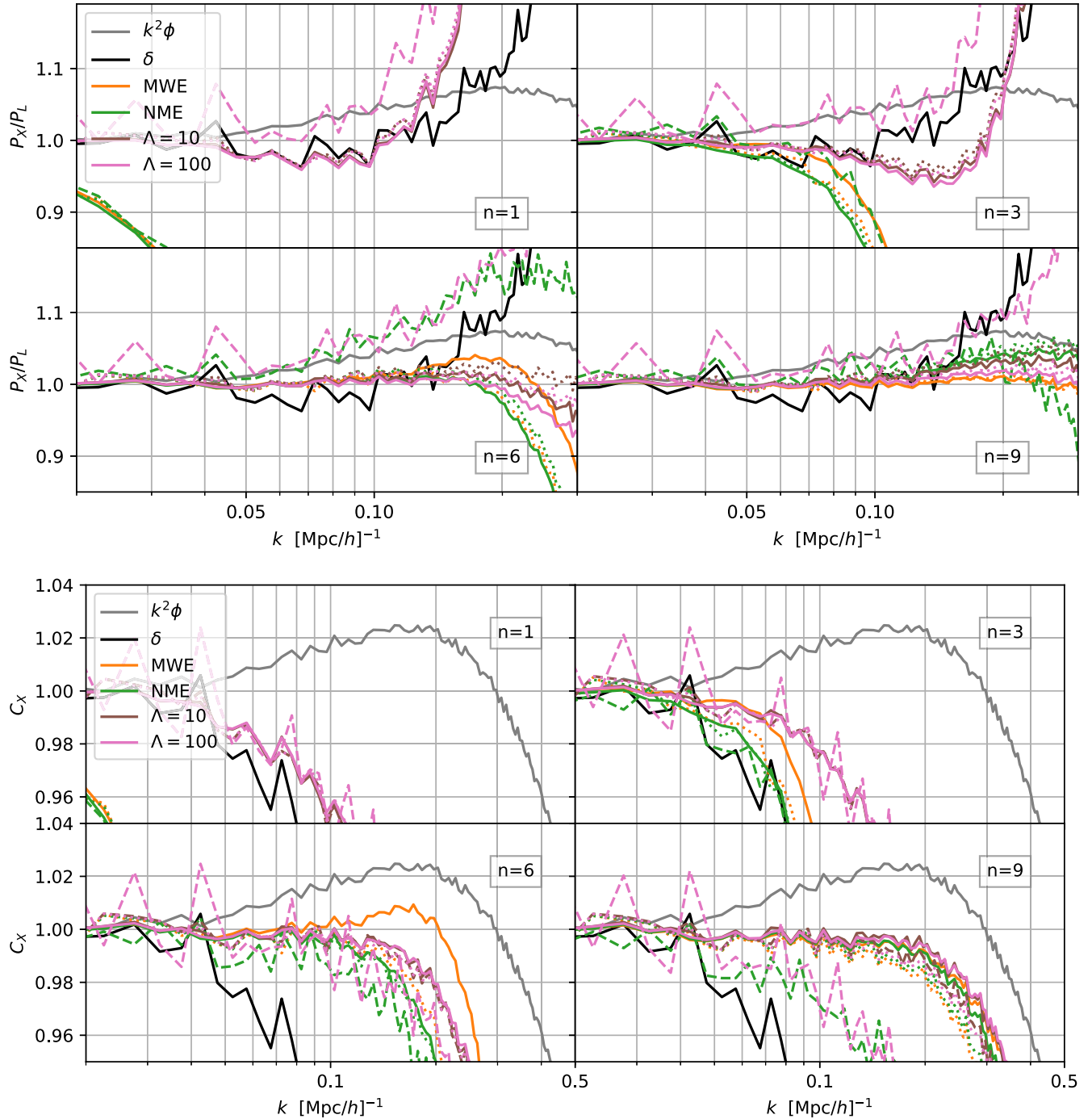


FIG. 4. A comparison of the power spectra (top) and propagators (bottom) of the various estimators for iterative reconstruction for  $n = 1, 3, 6,$  and  $9$ , in comparison to the true displacement (gray line). The black curve means the correlation functions with the nonlinear density perturbations. A solid (dotted) curve represents L500 (subL500). Various estimators agree with each other for  $n \leq 6$  and  $k < 0.2 h \text{Mpc}^{-1}$ , except for the traditional mass weighted estimators (orange line). The behavior of NME and RRE as a function of  $n$  shows a stable trend compared to MWE. The dashed lines show subsubL500, which corresponds to  $n_{\text{particle}} = 0.0043 h^3 \text{Mpc}^{-3}$ , i.e., closer to a realistic galaxy sample; here, the difference is mainly due to the effect of the sampling noise on the reconstruction efficiency.

### C. Discrepancy between reconstructed displacement fields and the model

So far, we have developed methods to reduce the discreteness effect in estimating the displacement fields,

which was considered an obstacle for comparing the reconstructed field and the true displacement field in Ref. [19]. Then we found that the discreteness effect cannot explain the discrepancy of 8% at  $k \sim 0.2 h \text{Mpc}^{-1}$ .

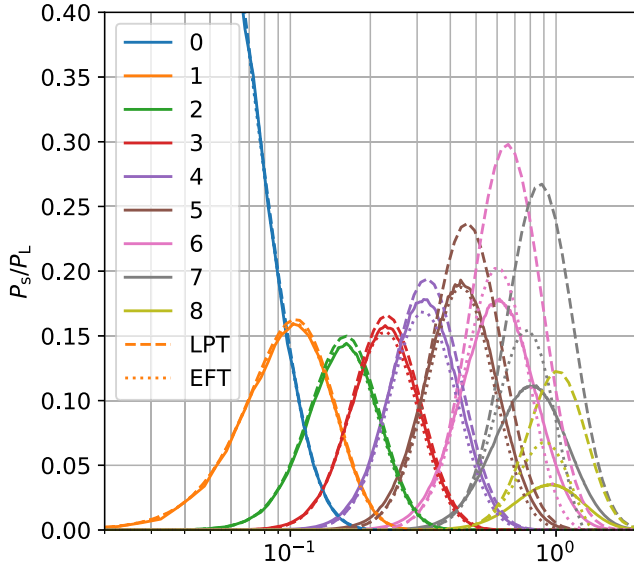


FIG. 5. A comparison of the shift vector [i.e., the input to reconstruction as in Eq. (40) or (41)] power spectra normalized by the linear power spectrum for the simulation (solid line), in comparison to our theoretical model based on LPT (dashed line) and EFT (dotted line) at each iteration. The y axis is the power spectrum of the shift vector normalized by the linear power spectrum, which is the smoothed density fluctuation power spectrum normalized by the linear spectrum. The shift vector of the simulation is better approximated by introducing the EFT terms [in Eq. (41)] up to  $n \leq 6$  for  $k \leq 0.4 h \text{Mpc}^{-1}$ .

Note that the discrepancy is apparent even from the very first iteration. Therefore, we revisit and improve the theoretical modeling to address this discrepancy better. In Ref. [19], we modeled the reconstruction procedure up to one-loop order in LPT (see the Appendix for a summary). A possible cause of discrepancy is that the standard perturbation theory (SPT) prediction fails at the quasi-nonlinear scale; i.e., we have  $\delta_{\text{NL,SPT}}^{(n)} \gg \delta_{\text{NL,sim}}^{(n)}$  for  $k \gtrsim 0.1 h \text{Mpc}^{-1}$  at  $z = 0.6$ , and therefore theory overestimates the displacement field (see Fig. 5). As a result, the estimation of the shift vector within SPT,

$$\mathbf{s}^{(n)} \approx -\frac{i\mathbf{k}}{k^2} S^{(n)} \delta_{\text{NL,SPT}}^{(n)}, \quad (40)$$

may not be very accurate, depending on the smoothing kernel and the redshift. Indeed, Fig. 5 shows the difference between Eq. (40) (dashed curves) and the actual shift vector (solid curves) from the simulations. The shortcoming of the one-loop calculation is mainly because of the UV-sensitive loop integrals. We can try to correct this term with effective field theory [40]. At one-loop order, a possible EFT correction is written as

$$\mathbf{s}^{(n)} \approx -\frac{i\mathbf{k}}{k^2} S^{(n)} [\delta_{\text{NL,SPT}}^{(n)} + \alpha k^2 \delta_{\text{L}}^{(n)}]. \quad (41)$$

We observe the nonlinear density field at every step and find a single parameter  $\alpha$  by fitting all steps simultaneously. With the least-squares method, we found  $\alpha = -0.999 \pm 0.011$  by fitting simulations in  $0.2 < k \text{Mpc}/h < 2$ .<sup>3</sup> In Fig. 5, we show a comparison of the power spectra of the shift vector for the simulation (solid line), LPT (dashed line), and EFT (dotted lines). The figure shows that this single EFT term can reduce the discrepancy in  $\mathbf{s}^{(n)}$  between the theory and the simulation, especially for  $n \leq 5$ .

We now propagate this EFT correction to the reconstructed displacement potential, which is written as

$$\phi_{\text{rec,EFT}}^{(n)} = \phi_{\text{rec,LPT}}^{(n)} - \alpha \delta_{\text{L}} \sum_{i=0}^{n-1} S^{(i)} B^{(i)}. \quad (42)$$

$B^{(n)}$  is the reconstruction kernel defined in the Appendix. The corresponding post-reconstruction power spectrum is written as

$$P_{\text{rec,EFT}}^{(n)} = P_{\text{rec,LPT}}^{(n)} - 2k^2 \alpha \bar{B}^{(n)} P_{\text{L}} \sum_{i=0}^{n-1} S^{(i)} B^{(i)}, \quad (43)$$

where  $\bar{B}^{(n)} \equiv 1 - B^{(n)}$ . Figure 6 shows a comparison of one-loop LPT and EFT for iterative reconstruction modeling, and we show our EFT helps to reach a few percent agreements for the displacement field up to  $k < 0.5 h \text{Mpc}^{-1}$  for  $z = 0.6$ , particularly for NME and for  $n < 6$  while the agreement is less for RRE. The EFT term can improve the agreement between the assumption in the theoretical model and the simulation and, therefore, the theoretical model of the resulting reconstructed displacement field within 1% for NME at  $k < 0.2 h \text{Mpc}^{-1}$ . Therefore our new theory model explains that we do not recover the true displacement field with the current reconstruction method. While we fixed  $\alpha$  in the above consideration, we also varied  $\alpha$  for each iteration and considered independent fits; however, it did not improve the EFT correction any further.

In Ref. [19], we tested a different option of adopting EFT; i.e., the EFT fit directly to the post-reconstruction spectrum, by introducing the one-loop EFT term for the displacement field from the LPT perspective. That is, we considered

$$\bar{P}_{\text{rec,EFT}}^{(n)} = P_{\text{rec,LPT}}^{(n)} + 2\bar{\alpha} k^2 P_{\text{rec,LPT}}^{(n),\text{lin}}, \quad (44)$$

with the linear power spectrum of  $n$ th step displacement field  $\phi_{\text{rec,LPT}}^{(n)}$ , that is,  $P_{\text{rec,LPT}}^{(n),\text{lin}}$ . This prescription is based on the EFT for the displacement field in Ref. [4]. Such a

<sup>3</sup> $\alpha \sim -1$  is just by accident as  $\alpha$  has dimension.

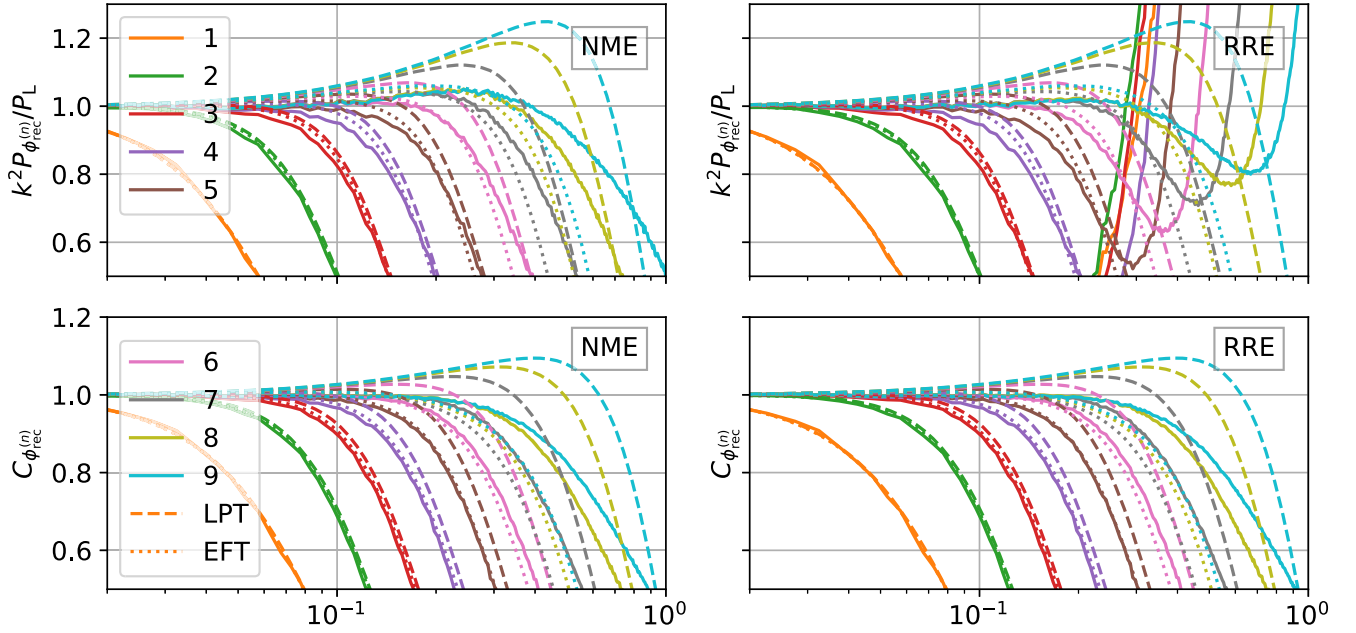


FIG. 6. The post-reconstruction power spectrum normalized by the linear power spectrum (top) and correlation function (bottom) for the simulation, one-loop LPT, and one-loop EFT. The EFT parameter is obtained by fitting the shift vector power spectra with a one-loop EFT correction. The simulation estimators are the NME (left) and the corrected RRE (right) for L500.

EFT choice was good at reproducing the true displacement field, but it could not model the power spectrum of the post-reconstruction field. In detail, the offset between the model and the reconstructed simulation happened mainly in the propagator (approximately  $P_{13}$ ); i.e., the simulation returned  $P_{13}$  more negative than the theory. In contrast, the  $P_{22}$  contribution seemed in a good agreement.

In this paper, we fit the post-reconstruction spectrum at a few percent precision with the proper choice of the EFT terms. The corrections in the shift vectors at each iteration with a fitting parameter  $\alpha$  are considered. The new model predicts the  $P_{13}$  contribution to 1% for  $n \leq 9$  in the lower panel of Fig. 6. The accuracy of the improved theory model is almost at the same level as the one-loop EFT fit for the pre-reconstruction matter power spectrum.

To summarize, once we correct Eq. (40) with the EFT, our model explains that Ref. [6] could not reconstruct the nonlinear displacement field. This implies that truncation of  $R^{(n)}$  in Eq. (39) causes the discrepancy in Sec. VI B. Thus, at least partly, the nonlinear physics prevents us from recovering the nonlinear displacement field.

## VII. CONCLUSIONS

The Lagrangian displacement can be useful in large-scale structure analysis since it contains the BAO feature such as the linear field. Reference [6] proposed a reconstruction method of the displacement field and confirmed the post-reconstruction field is highly correlated with the linear field. In Ref. [19], we compared the

reconstructed displacement field and the true one. We also constructed a theoretical model for the displacement reconstruction for broadband analysis in that work. Then we found two discrepancies: the difference between the reconstructed displacement and the true one; and the disagreement between the simulations and the model. The former is the level of 8%, and the latter is 15% at  $k \sim 0.2 h \text{Mpc}^{-1}$  for  $n = 9$  at  $z = 0.6$ . Sources of the discrepancies were speculated, particularly including a numerical artifact due to the discreteness of the sample.

This paper worked on mitigating such numerical artifacts by developing new estimators. The new estimators are robust to sampling, but we still observed the difference between the true displacement and the reconstructed one. Therefore, we conclude that the method in Ref. [6] does not reproduce the true displacement field. Based on the new estimators, we identified the source of the discrepancy between the theoretical model and the reconstruction method. Using the EFT approach, we improved our theoretical model and decreased the discrepancy between the model and simulation to a few percent at  $k \sim 0.2 h \text{Mpc}^{-1}$  for  $n = 9$  at  $z = 0.6$ . We summarize our results below.

First, the mass-weighted displacement estimator we used in Ref. [19] is subject to the error particularly sensitive to the number of empty pixels. We proposed two new displacement field estimators to overcome the discreteness effect: the NME and the RRE. The NME is the momentum-like estimator that does not suffer significantly from the empty pixel. Then, utilizing the fact that the Lagrangian positions of the tracers are independent of their displacement field, we showed that the momentum correlation

function normalized by the mass correlation function could closely return the displacement field. This method can avoid the complexity of direct, volume-weighted estimators using tessellation [21]. Another new estimator, RRE, is a density field-based estimator of the displacement field we devised. Based on the Lagrangian resummation theory [5], we showed that a rescaled density field power spectrum asymptotes to the displacement divergence power spectrum. That is, the RRE reduces to the displacement field divergence spectrum without explicit evaluation of the vector field, and thus we could avoid the empty pixel issue. We investigated the convergence of two estimators for various sampling cases. We concluded that NME performs much better than the RRE and MWE for dealing with the sampling artifacts. We confirmed that NME is stable for  $k \lesssim 0.2 h \text{Mpc}^{-1}$  at  $z \sim 0.6$  even if 99.6% of pixels are empty.

We applied our new methods for the post-reconstruction displacement fields. Then we identified the residual difference between the theoretical model in Ref. [19] and the reconstructed field in our simulation. With a physically motivated implementation of effective field theory for iterative reconstruction, we could produce a more accurate model for the post-reconstructed field. The precision of our model for the broadband of the reconstructed power spectrum is a few percent at  $k < 0.2 h \text{Mpc}^{-1}$  at  $z = 0.6$ .

This work only considered dark matter  $N$ -body simulation in real space. In terms of the numerical operation, we expect that NME and RRE can be straightforwardly extended to galaxy samples in redshift space, which we plan for future investigation. Also, we plan to extend our

perturbation theory model approach for a more realistic setup with galaxy bias and redshift-space distortion, in addition to the shot noise effect.

## ACKNOWLEDGMENTS

The authors thank Marcel Schmittfull for providing simulations and valuable discussions. A. O. is supported in part by the National Key R&D Program of China (No. 2021YFC2203100). A. O. and H.-J. S. are supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics under DE-SC0019091. S. S. acknowledges support for this work from NSF-2219212. S. S. is also supported by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 853291). F. B. is a University Research Fellow.

## APPENDIX: LPT MODELING OF ITERATIVE RECONSTRUCTION

The main text considers the effective field theory modeling of post-reconstruction displacement based on the one-loop LPT modeling discussed in Ref. [19]. In this appendix, we summarize the derivation in the reference. In that work, we started from a general perturbative expansion ansatz of the  $n$ th step nonlinear displacement  $\phi^{(n)}$ :

$$k^2 \phi^{(n)}(\mathbf{k}) = B^{(n)}(\mathbf{k}) \delta_L(\mathbf{k}) + \frac{1}{2!} \int \frac{d^3 k_1 d^3 k_2}{(2\pi)^6} (2\pi)^3 \delta_D^{(3)}(\mathbf{k} - \mathbf{k}_1 - \mathbf{k}_2) B^{(n)}(\mathbf{k}_1, \mathbf{k}_2) \delta_L(\mathbf{k}_1) \delta_L(\mathbf{k}_2) + \frac{1}{3!} \int \frac{d^3 k_1 d^3 k_2 d^3 k_3}{(2\pi)^9} (2\pi)^3 \delta_D^{(3)}(\mathbf{k} - \mathbf{k}_1 - \mathbf{k}_2 - \mathbf{k}_3) B^{(n)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \delta_L(\mathbf{k}_1) \delta_L(\mathbf{k}_2) \delta_L(\mathbf{k}_3). \quad (\text{A1})$$

We determine the LPT kernels  $B^{(n)}$  by solving the recurrence relations, which are derived by modeling the steps summarized in Sec. VI. In this appendix, we consider the expansion with respect to  $\delta_L$ , while the original expansion was for  $\phi^{(0)}$  since we expanded  $\phi^{(0)}$  into  $\delta_L$  in the end. We confirmed that the final results are unchanged for both conventions. Then, the estimated displacement field is the sum of the total negative displacement

$$\phi_{\text{rec.LPT}}^{(n)} \equiv \sum_{i=0}^{n-1} (\phi^{(i)} - \phi^{(i+1)}) = \phi^{(0)} - \phi^{(n)}. \quad (\text{A2})$$

The LPT post-reconstruction power spectrum of Eq. (A2) up to one-loop order is written as follows:

$$P_{\phi_{\text{rec}}^{(n)}}^{\text{LPT}} = P_{\phi_{\text{rec}}^{(n)}11} + P_{\phi_{\text{rec}}^{(n)}22} + P_{\phi_{\text{rec}}^{(n)}13}, \quad (\text{A3})$$

where we defined

$$k^4 P_{\phi_{\text{rec}}^{(n)}11} = (B^{(0)} - B^{(n)})^2 P_L, \quad (\text{A4})$$

$$k^4 P_{\phi_{\text{rec}}^{(n)} 22} = \frac{k^3}{4\pi^2} \int_0^\infty x^2 dx \int_{-1}^1 d\mu P_L(ky) P_L(kx) \frac{(X^{(0)} - X^{(n)})^2}{2}, \quad (\text{A5})$$

$$k^4 P_{\phi_{\text{rec}}^{(n)} 13} = \frac{k^3}{4\pi^2} \bar{B}^{(n)} P_L \int_0^\infty x^2 dx \int_{-1}^1 d\mu P_L(kx) (Y^{(0)} - Y^{(n)}). \quad (\text{A6})$$

$X^{(n)}$ ,  $Y^{(n)}$ , and  $Z^{(n)}$  are defined as

$$X^{(n)}(k, kx, \mu) \equiv B^{(n)}(\mathbf{k} - \mathbf{k}', \mathbf{k}'), \quad (\text{A7})$$

$$Y^{(n)}(k, kx, \mu) \equiv B^{(n)}(-\mathbf{k}, \mathbf{k}', -\mathbf{k}'), \quad (\text{A8})$$

$$Z^{(n)}(k, kx, \mu) \equiv B^{(n)}(-\mathbf{k}, \mathbf{k}'), \quad (\text{A9})$$

with  $|\mathbf{k}'| = kx$ ,  $(\mathbf{k} \times \mathbf{k}')^2 / (kk')^2 = 1 - \mu^2$ , and  $y^2 = 1 - 2x\mu + x^2$ . The recurrence relations for  $X$ ,  $Y$ , and  $Z$  are

$$\begin{aligned} X^{(n+1)} = & X^{(n)} - S^{(n)}(k)X^{(n)} - x^{-1}y^{-2}\mu(1 - x\mu)S^{(n)}(k)B^{(n)}(kx)B^{(n)}(ky) \\ & - x^{-1}y^{-2}(x - \mu)(1 - x\mu)S^{(n)}(ky)B^{(n)}(kx)B^{(n)}(ky) - y^{-2}\mu(x - \mu)S^{(n)}(kx)B^{(n)}(kx)B^{(n)}(ky), \end{aligned} \quad (\text{A10})$$

$$\begin{aligned} Y^{(n+1)} = & Y^{(n)} - S^{(n)}(k)Y^{(n)} + 2x^{-2}\mu^2 S^{(n)}(k)B^{(n)}(k)B^{(n)}(kx)^2 + 2\mu^2 S^{(n)}(kx)B^{(n)}(k)B^{(n)}(kx)^2 \\ & - 2x^{-1}y^{-2}\mu(1 - x\mu)S^{(n)}(k)B^{(n)}(kx)Z^{(n)} - 2\mu(x - \mu)y^{-2}S^{(n)}(kx)B^{(n)}(kx)Z^{(n)} \\ & - 2x^{-1}(x - \mu)(1 - x\mu)y^{-2}S^{(n)}(ky)B^{(n)}(kx)Z^{(n)} \\ & - 2y^{-2}x^{-2}(x - \mu)^2(1 - x\mu)^2 S^{(n)}(ky)B^{(n)}(k)B^{(n)}(kx)^2, \end{aligned} \quad (\text{A11})$$

$$\begin{aligned} Z^{(n+1)} = & Z^{(n)} - S^{(n)}(ky)Z^{(n)} - x^{-1}(x - \mu)(1 - x\mu)S^{(n)}(ky)B^{(n)}(k)B^{(n)}(kx) \\ & - \mu(x - \mu)S^{(n)}(kx)B^{(n)}(k)B^{(n)}(kx) - x^{-1}\mu(1 - x\mu)S^{(n)}(k)B^{(n)}(k)B^{(n)}(kx). \end{aligned} \quad (\text{A12})$$

Note that the factors of 2 in Eq. (A11) come from the  $\mu \rightarrow -\mu$  symmetry. The initial conditions for the recurrence relations are given as

$$X^{(0)} = \frac{3(1 - \mu^2)}{7y^2}, \quad (\text{A13})$$

$$Y^{(0)} = \frac{10(1 - \mu^2)^2}{21y^2}, \quad (\text{A14})$$

$$Z^{(0)} = \frac{3(1 - \mu^2)}{7}. \quad (\text{A15})$$

Similarly, we find that the power spectra of the shift vectors Eq. (40) are written as

$$P_{s^{(n)}} = P_{s^{(n)} 11} + P_{s^{(n)} 22} + P_{s^{(n)} 13}, \quad (\text{A16})$$

where we defined

$$k^4 P_{s^{(n)} 11} = B^{(n)}(k)^2 P_L(k), \quad (\text{A17})$$

$$k^4 P_{s^{(n)} 22} = \frac{k^3}{4\pi^2} \int x^2 dx \int d\mu P_L(kx) P_L(ky) \frac{1}{2} \left[ X^{(n)} + \frac{\mu(1 - x\mu)}{xy^2} B^{(n)}(kx) B^{(n)}(ky) \right]^2, \quad (\text{A18})$$

$$k^4 P_{s^{(n)} 13} = \frac{k^3}{4\pi^2} B^{(n)} P_L \int x^2 dx \int d\mu P_L(kx) \left[ \frac{2(1 - x\mu)\mu}{xy^2} Z^{(n)} B^{(n)}(kx) + Y^{(n)} - \frac{\mu^2}{x^2} B^{(n)}(kx)^2 B^{(n)}(k) \right]. \quad (\text{A19})$$

We confirmed that the above equation coincides with the one-loop SPT spectrum for  $n = 0$ .

- [1] Daniel J. Eisenstein, Hee-jong Seo, and Martin J. White, On the robustness of the acoustic scale in the low-redshift clustering of matter, *Astrophys. J.* **664**, 660 (2007).
- [2] Martin Crocce and Roman Scoccimarro, Nonlinear evolution of baryon acoustic oscillations, *Phys. Rev. D* **77**, 023533 (2008).
- [3] Hee-Jong Seo, Jonathan Eckel, Daniel J. Eisenstein, Kushal Mehta, Marc Metchnik, Nikhil Padmanabhan, Phillip Pinto, Ryuichi Takahashi, Martin White, and Xiaoying Xu, High-precision predictions for the acoustic scale in the non-linear regime, *Astrophys. J.* **720**, 1650 (2010).
- [4] Tobias Baldauf, Emmanuel Schaun, and Matias Zaldarriaga, On the reach of perturbative descriptions for dark matter displacement fields, *J. Cosmol. Astropart. Phys.* **03** (2016) 017.
- [5] Takahiko Matsubara, Resumming cosmological perturbations via the Lagrangian picture: One-loop results in real space and in redshift space, *Phys. Rev. D* **77**, 063530 (2008).
- [6] Marcel Schmittfull, Tobias Baldauf, and Matias Zaldarriaga, Iterative initial condition reconstruction, *Phys. Rev. D* **96**, 023505 (2017).
- [7] Daniel J. Eisenstein, Hee-jong Seo, Edwin Sirko, and David Spergel, Improving cosmological distance measurements by reconstruction of the baryon acoustic peak, *Astrophys. J.* **664**, 675 (2007).
- [8] Svetlin Tassev and Matias Zaldarriaga, Towards an optimal reconstruction of baryon oscillations, *J. Cosmol. Astropart. Phys.* **10** (2012) 006.
- [9] Xin Wang, Hao-Ran Yu, Hong-Ming Zhu, Yu Yu, Qiaoyin Pan, and Ue-Li Pen, Isobaric reconstruction of the baryonic acoustic oscillation, *Astrophys. J.* **841**, L29 (2017).
- [10] Hong-Ming Zhu, Ue-Li Pen, and Xuelei Chen, Primordial density and BAO reconstruction, *arXiv:1609.07041*,
- [11] Hong-Ming Zhu, Yu Yu, Ue-Li Pen, Xuelei Chen, and Hao-Ran Yu, Nonlinear reconstruction, *Phys. Rev. D* **96**, 123502 (2017).
- [12] Yu Yu, Hong-Ming Zhu, and Ue-Li Pen, Halo nonlinear reconstruction, *Astrophys. J.* **847**, 110 (2017).
- [13] Yanlong Shi, Marius Cautun, and Baojiu Li, New method for initial density reconstruction, *Phys. Rev. D* **97**, 023505 (2018).
- [14] Ryuichiro Hada and Daniel J. Eisenstein, An iterative reconstruction of cosmological initial density fields, *Mon. Not. R. Astron. Soc.* **478**, 1866 (2018).
- [15] Ryuichiro Hada and Daniel J. Eisenstein, Application of the iterative reconstruction to simulated galaxy fields, *Mon. Not. R. Astron. Soc.* **482**, 5685 (2019).
- [16] Tian-Xiang Mao, Jie Wang, Baojiu Li, Yan-Chuan Cai, Bridget Falck, Mark Neyrinck, and Alex Szalay, Baryon acoustic oscillations reconstruction using convolutional neural networks, *Mon. Not. R. Astron. Soc.* **501**, 1499 (2021).
- [17] Christopher J. Shallue and Daniel J. Eisenstein, Reconstructing cosmological initial conditions from late-time structure with convolutional neural networks, *Mon. Not. R. Astron. Soc.* **520**, 6256 (2023).
- [18] Hee-Jong Seo, Atsuhisa Ota, Marcel Schmittfull, Shun Saito, and Florian Beutler, Iterative reconstruction excursions for baryon acoustic oscillations and beyond, *Mon. Not. R. Astron. Soc.* **511**, 1557 (2022).
- [19] Atsuhisa Ota, Hee-Jong Seo, Shun Saito, and Florian Beutler, Modeling iterative reconstruction and displacement field in the large scale structure, *Phys. Rev. D* **104**, 123508 (2021).
- [20] P. J. E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, 1980).
- [21] Francis Bernardeau and Rien van de Weygaert, A new method for accurate velocity statistics estimation, *Mon. Not. R. Astron. Soc.* **279**, 693 (1996).
- [22] Sebastian Pueblas and Roman Scoccimarro, Generation of vorticity and velocity dispersion by orbit crossing, *Phys. Rev. D* **80**, 043504 (2009).
- [23] Stephane Colombi, Michal Chodorowski, and Romain Teyssier, Cosmic velocity–gravity relation in redshift space, *Mon. Not. R. Astron. Soc.* **375**, 348 (2007).
- [24] Yi Zheng, Pengjie Zhang, Yipeng Jing, Weipeng Lin, and Jun Pan, Peculiar velocity decomposition, redshift space distortion and velocity reconstruction in redshift surveys. II. Dark matter velocity statistics, *Phys. Rev. D* **88**, 103510 (2013).
- [25] Pengjie Zhang, Yi Zheng, and Yipeng Jing, Sampling artifact in volume weighted velocity measurement. I, Theoretical modeling, *Phys. Rev. D* **91**, 043522 (2015).
- [26] Yi Zheng, Pengjie Zhang, and Yipeng Jing, Sampling artifact in volume weighted velocity measurement. II, Detection in simulations and comparison with theoretical modeling, *Phys. Rev. D* **91**, 043523 (2015).
- [27] Yu Yu, Jun Zhang, Yipeng Jing, and Pengjie Zhang, Kriging interpolating cosmic velocity field, *Phys. Rev. D* **92**, 083527 (2015).
- [28] Yu Yu, Jun Zhang, Yipeng Jing, and Pengjie Zhang, Kriging interpolating cosmic velocity field. II. Taking anisotropies and multistreaming into account, *Phys. Rev. D* **95**, 043536 (2017).
- [29] Changbom Park, Cosmic momentum field and mass fluctuation power spectrum, *Mon. Not. R. Astron. Soc.* **319**, 573 (2000).
- [30] Chan-Gyung Park and Changbom Park, Power spectrum of cosmic momentum field measured from the SFI galaxy sample, *Astrophys. J.* **637**, 1 (2006).
- [31] Cullan Howlett, The redshift-space momentum power spectrum—I. Optimal estimation from peculiar velocity surveys, *Mon. Not. R. Astron. Soc.* **487**, 5209 (2019).
- [32] Jun Pan, Estimating power spectrum of discrete cosmic momentum field with fast Fourier transform, *Res. Astron. Astrophys.* **20**, 146 (2020).
- [33] Hee-Jong Seo, Atsuhisa Ota, Shun Saito, Florian Beutler, and Marcel Schmittfull, Iterative reconstruction excursions for Baryon Acoustic Oscillations and beyond, *Mon. Not. R. Astron. Soc.* **511**, 1557 (2022).
- [34] P. A. R. Ade *et al.*, Planck 2015 results. XIII. Cosmological parameters, *Astron. Astrophys.* **594**, A13 (2016).
- [35] Volker Springel, Naoki Yoshida, and Simon D. M. White, Gadget: A code for collisionless and gasdynamical cosmological simulations, *New Astron.* **6**, 79 (2001).
- [36] Volker Springel, The cosmological simulation code Gadget-2, *Mon. Not. R. Astron. Soc.* **364**, 1105 (2005).

- [37] Yu Feng, Simeon Bird, Lauren Anderson, Andreu Font-Ribera, and Chris Pedersen, MP-Gadget/MP-Gadget: A tag for getting a DOI, 10.5281/zenodo.1451799 (2018).
- [38] Martin White, Shot noise and reconstruction of the acoustic peak, [arXiv:1004.0250v1](https://arxiv.org/abs/1004.0250v1).
- [39] A. Aghamousa, The DESI experiment part I: Science, targeting, and survey design (2016), [arXiv:1611.00036](https://arxiv.org/abs/1611.00036).
- [40] John Joseph M. Carrasco, Mark P. Hertzberg, and Leonardo Senatore, The effective field theory of cosmological large scale structures, *J. High Energy Phys.* **09** (2012) 082.