

HALO NONLINEAR RECONSTRUCTION

YU YU^{1,*}, HONG-MING ZHU^{2,3}, AND UE-LI PEN^{4,5,6,7}

submitted to ApJ

ABSTRACT

We apply the nonlinear density reconstruction method (Zhu et al. (2016b)) to simulated halo fields. For halo number density $2.77 \times 10^{-2} (h^{-1}\text{Mpc})^{-3}$ at $z = 0$, corresponding to the SDSS main sample density, we find the scale where the noise saturates the linear signal is improved to $k \gtrsim 0.36 h\text{Mpc}^{-1}$, a factor of 2.29 improvement in scale, or 12 in number of linear modes. The improvement is less for higher redshift or lower halo density. We expect this to substantially improve the BAO accuracy of dense, low redshift surveys, including the SDSS main sample, 6dFGS and 21cm intensity mapping initiatives.

Keywords: cosmology: large-scale structure of universe

1. INTRODUCTION

Measuring the expansion history and structure growth has been a major focus of modern observational cosmology. A lot of information about our Universe is being decoded from the current surveys, like BOSS (Dawson et al. (2013)), 6dFGS (Jones et al. (2009)), and etc. Ambitious on-going and future surveys, like eBOSS⁹, DESI (DESI Collaboration et al. (2016)), PFS (Takada et al. (2014)), CHIME (Bandura et al. (2014)), HIRAX (Newburgh et al. (2016)), BINGO (Battye et al. (2016)), Tianlai¹⁰, HETDEX¹¹, Euclid¹², WFIRST¹³, and etc, will further extend the exploration in both width and depth. These surveys are expected to bring us significant improvement in the understanding of our Universe.

Most of the future surveys are aiming at the high redshift universe. These surveys will observe matter distributions with less nonlinear evolution in a huge cosmic volume. However, only very bright objects/features could be detected in these surveys. On the contrary, the low redshift universe is relatively easy to observe, but suffers from significant nonlinear effects and limited volume. The late-time nonlinear evolution is a complicated pro-

cess and the statistics is difficult to model. This induces systematics in the statistics of cosmic probes (e.g., the broadening and shifting of the baryon acoustic oscillation (BAO) peak in galaxy 2-point correlation function). It also prevents a simple mapping from the final state to the initial conditions that are predicted by theories. As a result, a large portion of the cosmic information is encoded into complicated high-order statistics, and even worse, some of them is lost (e.g., Rimes & Hamilton (2005, 2006); Neyrinck et al. (2006); Carron & Neyrinck (2012)).

A part of the information loss could be recovered by a process known as “reconstruction” (e.g., Eisenstein et al. (2007)). The nonlinear density field is smoothed on the linear scale ($\sim 10h^{-1}\text{Mpc}$) to make the Zel’dovich approximation valid. The linear displacement is estimated and used to move both the galaxy and a random sample back. The reconstructed field has sharper BAO peak, leading to more stringent cosmological parameter constraints. In the literatures, this reconstruction method was theoretically understood and modeled (e.g., Padmanabhan et al. (2009), Noh et al. (2009), White (2015), Seo et al. (2016)), tested against numerical simulations (e.g., Seo et al. (2008), Seo et al. (2010), Mehta et al. (2011), Achitouv & Blake (2015), Schmittfull et al. (2015), Obuljen et al. (2016)), and applied on observations (e.g., Padmanabhan et al. (2012), Xu et al. (2013), Anderson et al. (2014), Kazin et al. (2014), Ross et al. (2015), Beutler et al. (2016), Hinton et al. (2017), Beutler et al. (2017)) extensively.

Zhu et al. (2016b) presented a direct approach to non-parametrically reconstruct the linear density field, by solving for a unique displacement potential consistent with the nonlinear density map and positive definite coordinate transformation. Different from Eisenstein et al. (2007), the reconstructed displacement is not used to move the galaxy position. The reconstructed density field is directly derived by the negative divergence of the reconstructed displacement. Hereafter, we call this reconstruction process as *nonlinear reconstruction*. Zhu et al. (2016b) found that the reconstructed density field recovers the coherence with linear initial conditions up to $k \sim 1 h\text{Mpc}^{-1}$, a factor of ~ 6 improvement in scale compared to the nonlinear density field. The informa-

¹ Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai, 200030, China

* yuyu22@shao.ac.cn

² Key Laboratory for Computational Astrophysics, National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Beijing 100012, China

³ University of Chinese Academy of Sciences, Beijing 100049, China

⁴ Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, Ontario M5S 3H8, Canada

⁵ Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, Ontario M5S 3H4, Canada

⁶ Canadian Institute for Advanced Research, CIFAR Program in Gravitation and Cosmology, Toronto, Ontario M5G 1Z8, Canada

⁷ Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, N2L 2Y5, Canada

⁹ eBOSS, <http://www.sdss.org/surveys/eboss/>

¹⁰ Tianlai, <http://tianlai.bao.ac.cn>

¹¹ HETDEX, <http://www.hetdex.org/>

¹² EUCLID, <http://sci.esa.int/euclid/>

¹³ WFIRST, <http://wfirst.gsfc.nasa.gov/>

tion content is also increased by a factor of ~ 50 after the nonlinear reconstruction (see Pan et al. (2016)). Yu et al. (2016) quantified the limits of all Lagrangian reconstruction methods by measuring the correlations in Lagrangian space (i.e., the correlations between the initial displacement and final displacement) in simulation.

Nonlinear reconstruction generalizes the application of linear displacement theory to fully non-linear fields, potentially substantially expanding the BAO and redshift space distortions (RSD) information content of dense large-scale structure surveys. As shown in Zhu et al. (2016b), the noise part of the reconstructed field dominates over the linear signal at $k \gtrsim 0.6 \text{ hMpc}^{-1}$, suggesting that all BAO peaks may be recovered from the present day density field, opening up the potential of recovering cosmic information including BAO down to the Poisson noise limit. To apply this novel reconstruction method to observations, we need to consider the reconstruction from galaxy/halo density fields. This paper presents the performance test on the simulated halo fields with three number densities and three redshifts.

This article is organized as follows. In Sect. 2, we briefly introduce the reconstruction method. The performance tests on various situations are presented in Sect. 3. We summarize the result and discuss future directions in Sect. 4.

2. RECONSTRUCTION

2.1. Displacement

We are aiming at reconstructing a density field with more linear information. Actually, the reconstructed density field is a byproduct of the reconstruction method proposed by Zhu et al. (2016b), which non-parametrically reconstructs the unobservable nonlinear displacement.

In Lagrangian scenario of structure formation, the displacement field $\Psi(\mathbf{q}, t)$ describes the total motion at time t for each mass element starting from the initial Lagrangian position \mathbf{q} .

The displacement field, as a vector field, could be uniquely decomposed into a gradient part and a curl part.

$$\Psi(\mathbf{q}) = \Psi_E(\mathbf{q}) + \Psi_B(\mathbf{q}) , \quad (1)$$

where $\nabla \times \Psi_E = \mathbf{0}$ and $\nabla \cdot \Psi_B = 0$. One could further decompose the E -mode component into a linear part and a nonlinear part, in which the linear part is completely correlated with the linear displacement and the nonlinear part is the remaining. This physically motivated decomposition helps to understand the displacement field.

In the nonlinear regime, shell crossing happens. Mass elements lost their memory on the initial positions. This process generates the nonlinear and B -mode component in the displacement field. However, the final displacement is still dominated by the linear part. The B -mode and nonlinear part only fine-tune the final position (Zhu et al. (2017), in prep.). Thus, a unique displacement potential solution consistent with the nonlinear density field and positive definite coordinate transformation is expected to be a good approximation for the nonlinear displacement field (Zhu et al. (2016b)).

The reconstructed displacement field is potentially very useful in cosmology and the reconstruction of the linear density field is the first application. Further appli-

cations include RSD modelling, peculiar velocity reconstruction, and etc (Zhu et al. (2016b)).

2.2. Nonlinear reconstruction

The basic idea of the reconstruction is to build a bijective curvilinear coordinate system $\xi \equiv (\xi_1, \xi_2, \xi_3) = \xi(\mathbf{x})$, where the mass per volume element is constant,

$$\rho(\xi) d^3\xi = \text{constant} . \quad (2)$$

Here, $\sqrt{g} = \det(e_\mu^i)$ is the Jacobian of the transformation matrix $e_\mu^i = \partial x^i / \partial \xi^\mu$. Latin indices denote Cartesian coordinate labels x^i , while Greek indices denote the curvilinear coordinates ξ^μ . We call this curvilinear coordinate system as *potential mass gauge (coordinate)*. It becomes analogous to “synchronous gauge” and “Lagrangian coordinates” before shell crossing, but allows a unique mapping even after shell crossing. In this scenario, the mass element at final physical position \mathbf{x} comes from the estimated initial Lagrangian position given by the potential mass coordinate ξ .

Since we attempt to follow the potential flow instead of the vorticity, we define a coordinate transformation that is a pure gradient,

$$x^i = \xi^\mu \delta_\mu^i + \Delta x^i , \quad (3)$$

where the displacement

$$\Delta x^i \equiv \frac{\partial \phi}{\partial \xi^\nu} \delta^{i\nu} . \quad (4)$$

We call ϕ as the displacement potential.

Since the displacement from the initial Lagrangian position to the final Eulerian position can be large, it's difficult to obtain the solution directly. One efficient and robust algorithm is the moving mesh approach, which is originally introduced for the adaptive particle-mesh N-body algorithm and the moving mesh hydrodynamics algorithm (See Pen (1995, 1998)). These algorithms attempt to evolve the curvilinear coordinate along with the matter/energy density field to maintain constant mass/energy-resolution. In our case, this approach solves for the displacement potential perturbatively and iteratively. The evolution of the curvilinear coordinate system is determined by a linear elliptic evolution equation

$$\partial_\mu (\rho \sqrt{g} e_i^\mu \delta^{i\nu} \partial_\nu \phi) = \Delta \rho , \quad (5)$$

where $\Delta \rho = \bar{\rho} - \rho \sqrt{g}$. The time derivative here is relative to the iteration step. We obtain the change of the displacement potential $\Delta \phi = \dot{\phi} \Delta t$ at each iteration step and then update the density field in the new coordinate frame. The final solution is given by

$$\phi = \Delta \phi^{(1)} + \Delta \phi^{(2)} + \Delta \phi^{(3)} + \dots , \quad (6)$$

where $\Delta \phi^{(i)}$ is the result from the i th iteration. We also implement the smoothing and limiting schemes to guarantee the triad e_i^μ is positive definite at each step. See Zhu et al. (2016b) for a simple physical interpretation of Equation (5). The elliptic equation can be solved using the multigrid algorithm described also in Pen (1995).

The above process ensures that the coordinate transform is positive definite. The coordinate lines will not cross and the eigenvalues for this coordinate transform

n_h	2.77×10^{-2}	2.77×10^{-3}	2.77×10^{-4}	M_{\max}
$z = 0.0$	2.15×10^{10}	1.84×10^{12}	2.10×10^{13}	2.11×10^{15}
$z = 0.5$	6.44×10^{10}	1.70×10^{12}	1.53×10^{13}	1.16×10^{15}
$z = 1.0$	5.58×10^{10}	1.44×10^{12}	1.05×10^{13}	6.12×10^{14}

Table 1

Detailed information for various halo samples used in the performance test. The first line indicates the sample number density, in unit of $(h^{-1}\text{Mpc})^{-3}$. The minimum halo mass for three number densities and three redshifts is listed, in unit of M_{\odot}/h . In the last column, the maximum halo mass is listed for three redshifts.

are always positive. In the case that particles follow a irrotational potential flow and no shell crossing happens, the reconstructed displacement is exact up to a global spatial translation. However, shell crossing happens in the nonlinear regime. This reconstruction algorithm gives an approximate solution.

We define the negative Laplacian of the reconstructed displacement potential as the *reconstructed density field*,

$$\delta_r(\xi) \equiv -\nabla_{\xi} \cdot \Delta \mathbf{x}(\xi) = -\nabla_{\xi}^2 \phi. \quad (7)$$

3. IMPLEMENTATION AND RESULT

3.1. Halo density field

To test the performance of the reconstruction algorithm, we use a simulation run with the CUBEP³M code (See Harnois-Déraps et al. (2013)), involving 2048^3 dark matter particles in a box with side length of $600 h^{-1}\text{Mpc}$. The particle-mesh process adopts 4096^3 grid and the particle-particle process is involved. We output the snapshot at $z = 0, 0.5$, and 1 to cover the typical redshift for current and on-going galaxy surveys. The reconstruction and analysis are performed on 512^3 curvilinear and uniform grids. This grid size ensures that the following analysis is reliable and the reconstruction is not computationally expensive. We adopt a spherical overdensity halo finder. The smallest halo contains 10 particles and has a halo mass of $2.15 \times 10^{10} M_{\odot}/h$. We construct three halo samples with number density of 2.77×10^{-2} , 2.77×10^{-3} , $2.77 \times 10^{-4} (h^{-1}\text{Mpc})^{-3}$ by setting the lower limit in halo mass. The detailed halo sample information is listed in Table 1.

The halo field is a discrete, highly non-uniform field. Traditional mass assignment methods (e.g., NGP, CIC, TSC) will leave many grids empty. As the reconstruction algorithm aims to build a potential mass coordinate in which the mass per volume element is constant, a large area with no matter will cause the algorithm unstable. We need a new mass assignment method to avoid this situation.

The new mass assignment we adopt is similar to the NGP method. First, we find the nearest halo for each grid. If N grids sharing the same halo as the nearest halo, these grids obtain $1/N$ mass. For those halos not being the nearest halo for any grid, NGP mass assignment is adopted. It is a simplified version of Voronoi tessellation method, without the calculation for the cross-section between each Voronoi polyhedral and the grid volume. This assignment will leave no empty grid and has negligible effect in the high density region or for a high number density sample. For further discussion on the new mass assignment, we refer the readers to the Appendix.

Figure 1 shows the simulated dark matter slice at $z = 0$

and the corresponding three halo fields, with downgraded halo number density. The color blocks observed in the bottom-right plot is a feature for such a low density sample. Next, we apply the reconstruction algorithm on these halo fields to quantify the performance.

3.2. Number of iteration steps

To maintain positive coordinate transform for each step, smoothing and limiting scheme is adopted during the reconstruction. This prevents reaching a final status of exactly constant mass per volume element. In practice, we use the power spectrum of the reconstructed density field and its cross-correlation coefficient with the initial condition as the convergence criteria.

Comparing results at different iteration steps, we found that we only need ~ 600 steps for the reconstruction algorithm to reach convergence for these halo samples, much less than ~ 1500 steps for the reconstruction from a nonlinear dark matter density field. By observing the power spectrum of the reconstructed density field and its cross-correlation coefficient with the initial condition at different iteration steps, we found that the results reach convergence fast on large scales. For the reconstruction from a nonlinear dark matter density field, most of the computing time is spent on the reconstruction from nonlinear small scales. This reconstruction method outperforms others by taking use of these information residing in nonlinear regime (See Zhu et al. (2016b)). However, the halo field is a discrete sample, in which a part of small scale information is missing. This comes out as the fast convergence for reconstruction from halo field.

For a halo sample with lower number density, it contains less usable small scale information. The number of iteration steps to reach convergence decreases toward lower halo density. For convenience, we just fix the number of iteration steps to be 600 for all the halo samples.

3.3. Dependence on halo number density

In the left panel of Figure 2, we present the power spectrum for input halo fields and the reconstructed density fields. The cross-correlation coefficients with the linear density field are presented in the right panel. As reference, we also plot the result for the reconstruction from the nonlinear dark matter density field at the top row.

Different from the method proposed by Eisenstein et al. (2007), the reconstruction process in this work is purely a mathematical approach, which does not involve any cosmological dynamics. Without the assumptions on the fiducial cosmology, galaxy bias, growth rate, and etc, the algorithm always find the potential mass coordinate on which each grid element has approximately constant mass. This process ensures that the reconstructed density field is same with the input field at sufficiently large scale. Halo is a biased tracer of the underlying dark matter density field. Thus, the reconstructed fields also show the same bias to linear density field at small k . The main purpose of this work is to valid the reconstruction method on the halo field, by investigating the cross-correlation with the initial conditions. The bias presented in the reconstructed field will not influence the results.

The reconstruction performance is influenced by the number density. For DM fields, the cross-correlation coefficient curve is shifted toward small scale by a factor

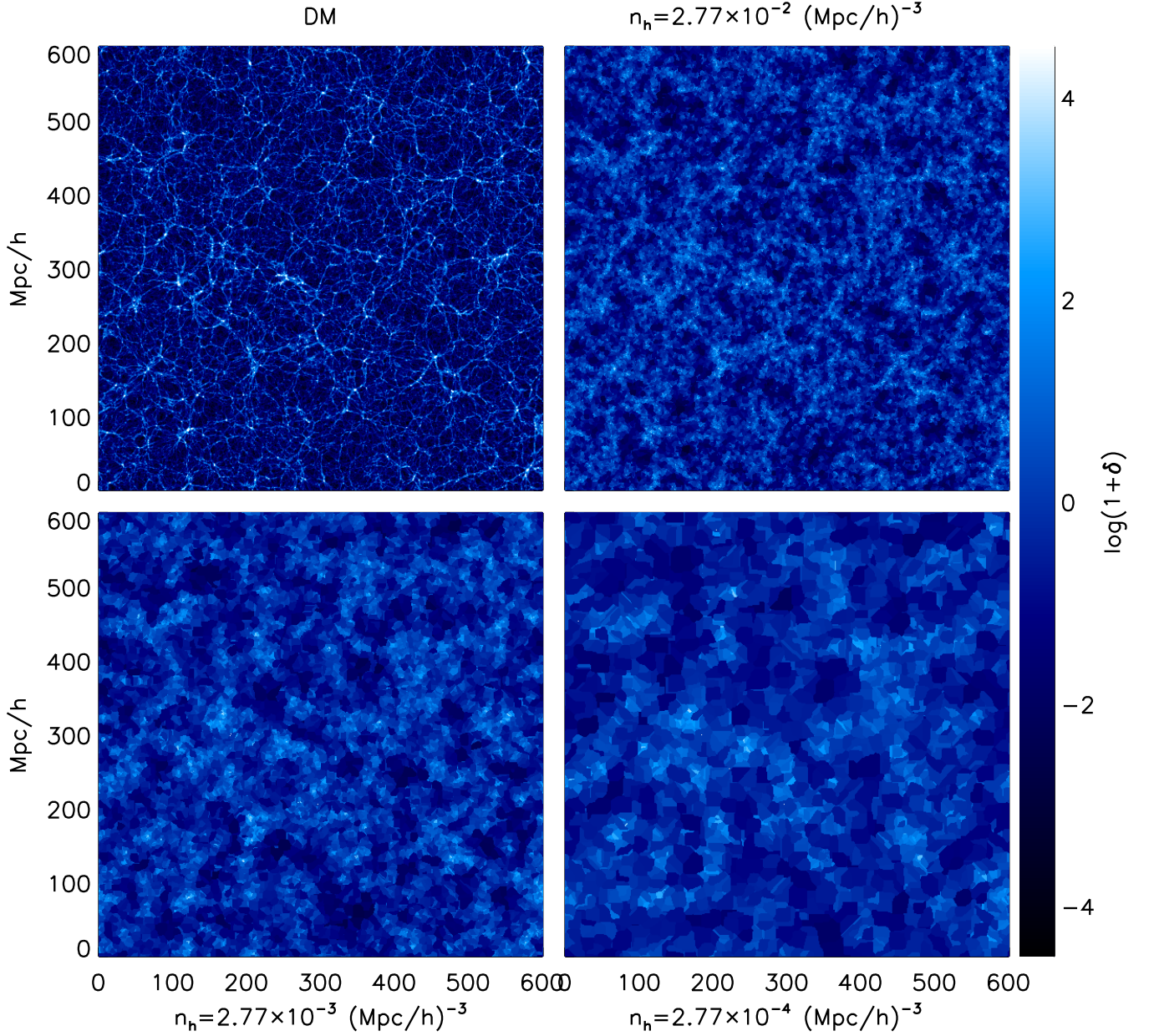


Figure 1. The input DM density field (top-left) and halo fields with number density of 2.77×10^{-2} , 2.77×10^{-3} , $2.77 \times 10^{-4} (h^{-1}\text{Mpc})^{-3}$ for $z = 0$. These halo fields adopt new assignment method to avoid empty grid and improve the stability of the reconstruction algorithm.

~ 6 by the reconstruction. We observe smaller amount of shift (less improvement) toward decreasing halo number density. For the highest density case, $n_h = 2.77 \times 10^{-2} (h^{-1}\text{Mpc})^{-3}$ sample, the shift is ~ 2.5 . For the intermediate sample, $n_h = 2.77 \times 10^{-3} (h^{-1}\text{Mpc})^{-3}$, the shift is $\lesssim 2$. For the lowest density case, only a very small improvement is observed.

Similar as the analysis in Seo et al. (2016), to clearly show how much linear signal is recovered by the reconstruction process, we decompose the density field into two terms,

$$\delta(k) = C(k)\delta_L(k) + n(k), \quad (8)$$

in which $C(k)\delta_L(k)$ is completely correlated with the linear density field. The pre-factor $C(k)$, often dubbed as the “propagator” (e.g., Crocce & Scoccimarro (2006, 2008); Matsubara (2008); Taruya et al. (2009)), could be obtained by

$$C(k) = \frac{P_{\delta_L\delta}(k)}{P_{\delta_L}(k)}. \quad (9)$$

The remaining noise term is just $n(k) = \delta(k) - C(k)\delta_L(k)$. This decomposition of the density field directly leads to the decomposition of the power spectrum into a linear signal term plus a noise term (also called

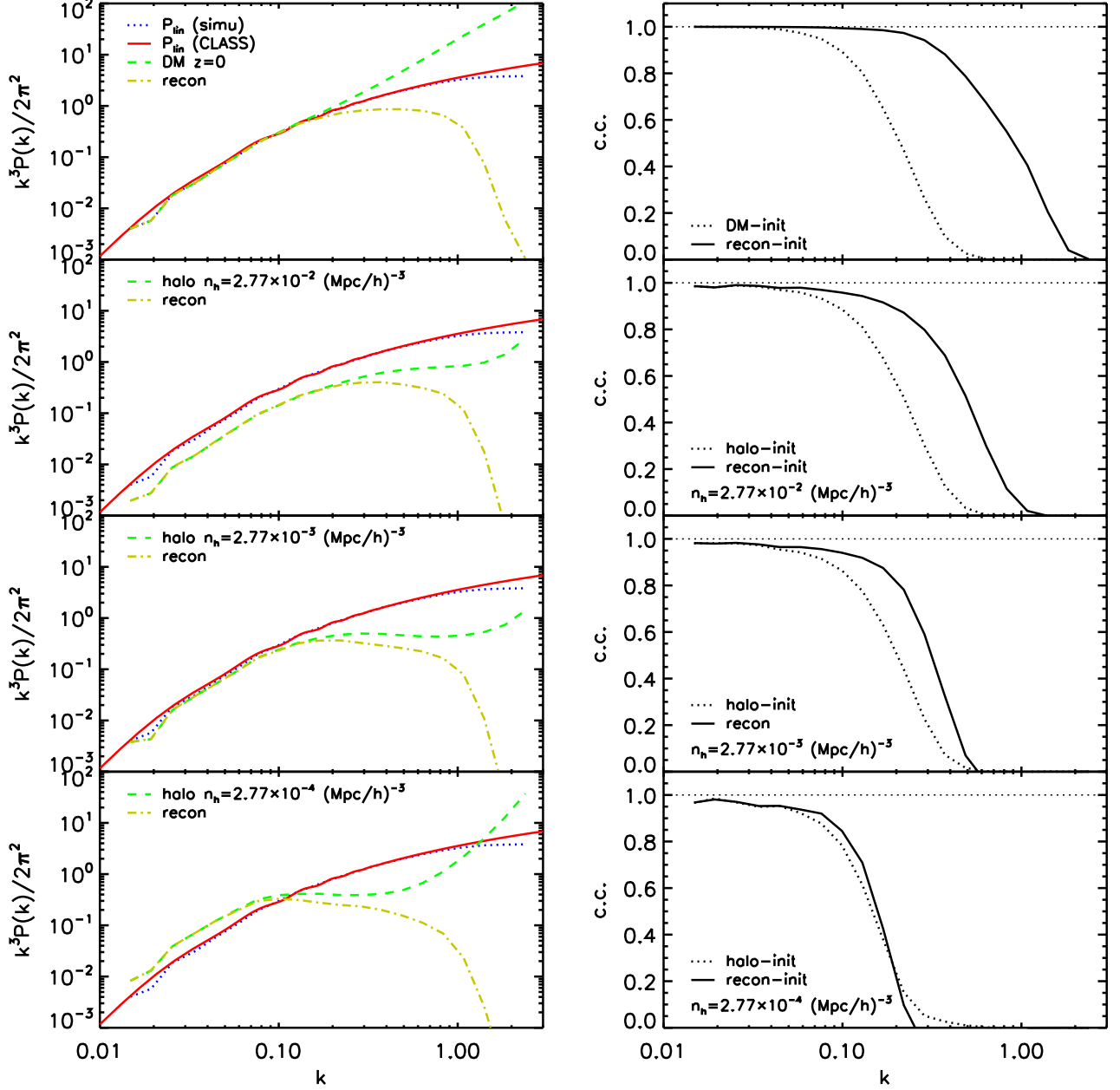


Figure 2. The left panel presents the power spectrum of the input fields and the reconstructed fields for $z = 0$. As reference, linear power spectrum from theory and simulation initial condition are also plotted. Right panel presents the cross-correlation coefficients between the reconstructed fields and linear density field in solid line. For comparison, the cross-correlation coefficients between the input fields and linear density field are plotted in dotted line. From top to bottom, the input field is DM density field, halo fields with number density of 2.77×10^{-2} , 2.77×10^{-3} , $2.77 \times 10^{-4} (h^{-1}\text{Mpc})^{-3}$, respectively.

“mode-coupling” term),

$$P_{\delta}(k) = P_s(k) + P_n(k), \quad (10)$$

in which $P_s(k) = C^2(k)P_{\delta_L}(k)$ is the linear signal term. The lower limit in scale where the BAO signal could be robustly measured is quantified by the scale where the linear signal term equals the noise term.

We present the power spectrum decomposition Equation (10) for both the input halo fields and the reconstructed fields in Figure 3. The results are scaled so that the total power $P_s(k) + P_n(k)$ equals the halo power spectrum with familiar NGP mass assignment, to better explain the results. The reconstructed field not only

has larger linear signal term, but also has lower noise term than the input field. The downward arrows indicate the scale where the linear signal term equals to the noise term. We clearly see the decrease of the scale (increase in k) by the reconstruction process. The improvement reaches a factor of 2.29, 1.69 and 1.15 for $n_h = 2.77 \times 10^{-2}$, 2.77×10^{-3} , $2.77 \times 10^{-4} (h^{-1}\text{Mpc})^{-3}$, respectively. Specifically, for the most concern case, this scale is decreased from $k = 0.16$ to $0.36 h\text{Mpc}^{-1}$.

Note that the noise term has two contributions, one from the nonlinear clustering and another from the shot noise. We present the expected shot noise power spectrum for NGP mass assignment in Figure 3. For the

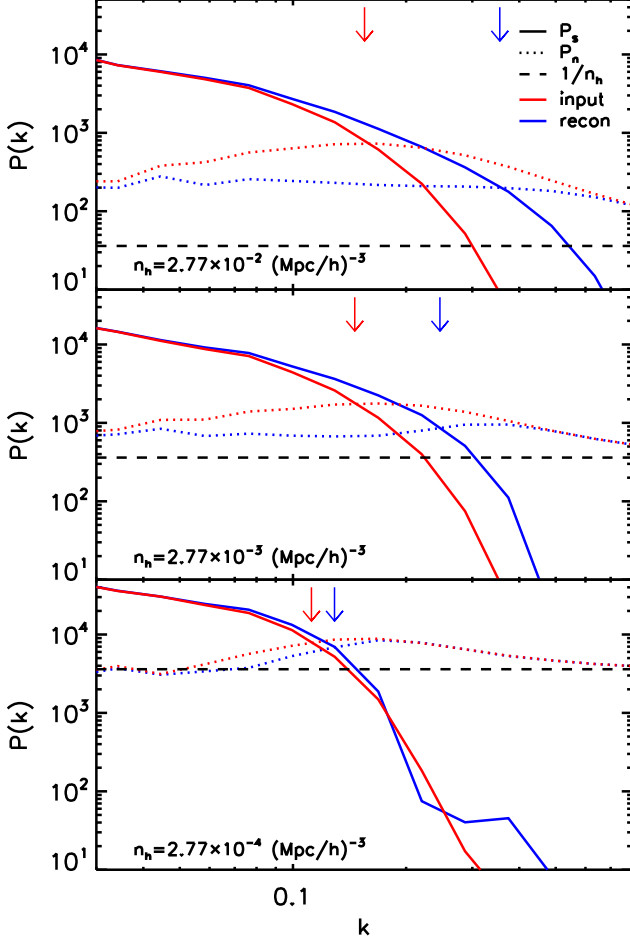


Figure 3. The power spectrum decomposition for the input (red) and reconstructed (blue) field. The linear signal term is presented in solid line, while the noise term in dotted line. All the results are scaled so that the total power equals to the halo power spectrum with familiar NGP mass assignment. The expected shot noise for NGP mass assignment, $1/n_h$, is also plotted in dashed line. From top to bottom, the input halo field has number density of 2.77×10^{-2} , 2.77×10^{-3} , 2.77×10^{-4} ($h^{-1}\text{Mpc}$) $^{-3}$, respectively. The downward arrows indicate the scale where the signal term equals the noise term.

highest number density case (top panel), the noise term of the input field exceeds the expected shot noise contribution a lot. This implies that the nonlinear clustering dominates over the shot noise. After the reconstruction, linear signal is recovered and meanwhile the nonlinear clustering is suppressed. Thus, the nonlinear reconstruction is powerful for this case. For the lowest number density case (bottom panel), the noise term of the input field is dominated by the shot noise. The halo field with such a low number density is almostly the combination of the linear signal plus the Poisson noise. In this case, there is no usable nonlinear information and the reconstruction performance is limited.

In Figure 4, we present the comparison of cross-correlation coefficient for various input and reconstructed fields, including the DM field and halo samples with $n_h = 2.77 \times 10^{-2}$ and 2.77×10^{-3} ($h^{-1}\text{Mpc}$) $^{-3}$. These three input fields have similar coherence with the initial condition. This implies that $n_h \gtrsim 2.77 \times 10^{-3}$ ($h^{-1}\text{Mpc}$) $^{-3}$ is roughly sufficient for extracting BAO in-

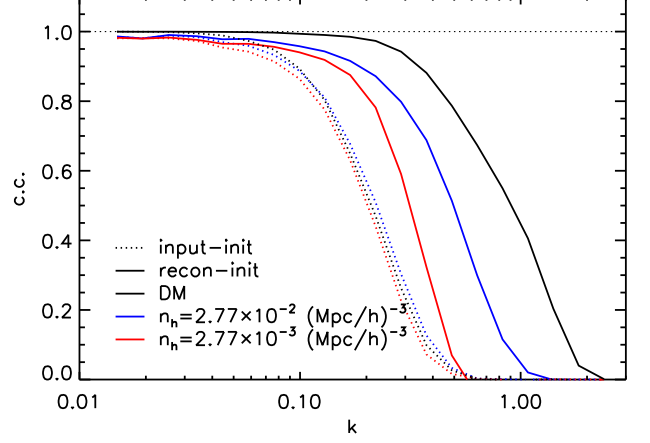


Figure 4. The comparison of cross-correlation coefficient for various input and reconstructed fields, including the DM field and two halo samples. The three input fields have similar coherence with the initial condition. But, the reconstructed fields benefit from higher halo number density.

formation without any reconstruction process. However, the reconstruction performance indeed benefits from high halo number density. This strongly motivates us to go deeper for low redshift (local) surveys.

3.4. Dependence on redshift

We present the redshift dependence of the reconstruction performance for each halo sample in Figure 5. Note that these halo samples at different redshifts are not the same halo population at different cosmic epoch, since we just choose a specific halo number density for all three redshifts.

The top panel shows the result for the input halo sample with number density of 2.77×10^{-2} ($h^{-1}\text{Mpc}$) $^{-3}$. The higher redshift halo field experiences less nonlinear evolution, and thus contains more linear information. The reconstruction increases the cross-correlation coefficients similarly for all three redshifts. The performance is only mildly degraded for $z = 1$.

For the input halo field with a number density of 2.77×10^{-3} ($h^{-1}\text{Mpc}$) $^{-3}$, we can still see the improvements for all three redshifts. However, the improvement is less obvious toward higher redshift. This could be explained by the fact that the halo sample with such a low number density (large mean separation) is much less sensitive on the nonlinear effects. Thus, the power of the nonlinear reconstruction is limited. As a result, the reconstruction recovers the cross-correlation to comparable level for all the three redshifts.

For the number density of 2.77×10^{-4} ($h^{-1}\text{Mpc}$) $^{-3}$ case, the input halo fields at three redshifts contain similar coherence with the initial condition. After the reconstruction, we only observe very small improvement or no improvement for these redshifts.

In the two lowest density case, we observe the convergence for the recovered cross-correlation coefficients for different redshifts. If we define a characteristic scale to quantify the performance, namely where the cross-correlation coefficients drops to 0.7 (corresponds to $S/N = 1$ for the power spectrum decomposition Equation (10), see Zhu et al. (2016a)), we find this roughly scales as $n_h^{-1/3}$. This implies that the reconstruction per-

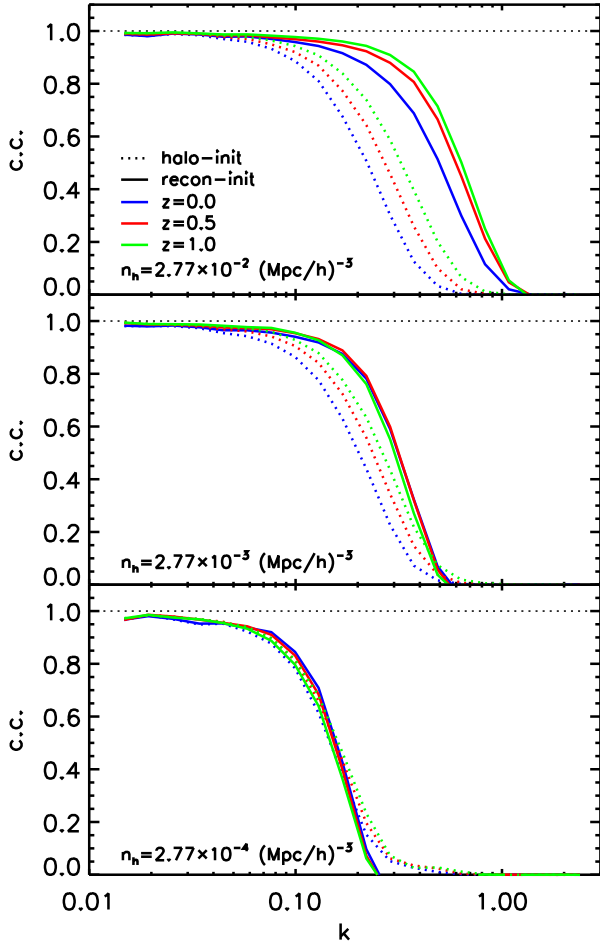


Figure 5. The cross-correlation coefficient between the reconstructed fields and the linear density field is presented in solid line. For comparison the result for the input halo fields is presented in dotted line. For a given halo number density, the result for $z = 0$, 0.5, and 1 is presented in blue, red, and green, respectively.

formance is already limited by the information possible to be reconstructed for a given halo number density. It is also this limitation that mildly degrades the reconstruction performance for $n_h = 2.77 \times 10^{-2} (h^{-1}\text{Mpc})^{-3}$ at $z = 1$.

4. CONCLUSION AND DISCUSSION

We tested the reconstruction method proposed in Zhu et al. (2016b) on the simulated halo fields with three different number densities and at three redshifts. The reconstruction performance is quantified by the extension in scale where the BAO signal could be robustly measured. We decomposed the power spectrum of both the input field and the reconstructed field into a linear signal term plus a noise term. This scale is defined at where the linear term equals the noise term. For the most concerned case, $n_h = 2.77 \times 10^{-2} (h^{-1}\text{Mpc})^{-3}$ at $z = 0$ which is close to the condition of SDSS main sample, we found the improvement reaches a factor of 2.29 in scale (from $k = 0.16$ to $0.36 h\text{Mpc}^{-1}$), or equivalently, a factor of 12 for available modes for BAO measurement.

We found better reconstruction performance for higher halo number density case. At $z = 0$, the improvements reach a factor of 2.29, 1.69 and 1.15 for halo sample with

number density of $n_h = 2.77 \times 10^{-2}$, 2.77×10^{-3} , $2.77 \times 10^{-4} (h^{-1}\text{Mpc})^{-3}$, respectively. We also reported the limited reconstruction power for the low number density cases, which roughly correspond to the targets of the ongoing high redshift surveys.

To apply the new reconstruction method on discrete halo fields, we need a suitable mass assignment method, which does not induce extra degree-of-freedom in the process. The simplified Voronoi tessellation method we adopted in this work is a straightforward attempt. There may exist better mass-assignment worth a try to deal with this issue (e.g., Delaunay tessellation method, Cautun & van de Weygaert (2011)). However, for the most concerned high density case, the choice of mass assignment is less important. For low density case, none of the mass assignment could present the underlying DM field well without introducing extra parameters.

As discussed in Zhu et al. (2016b), the Lagrangian BAO reconstruction algorithm involves displacing individual objects according to the linear displacement that is computed from the observed galaxy distribution under certain model assumptions (the fiducial cosmology, galaxy bias, growth rate, etc (See Eisenstein et al. (2007))). The results depend on the assumed fiducial model and must be tested against different parameter choices, which is computationally expensive (See Padmanabhan et al. (2012)). The reconstruction method used in this work is a purely mathematical approach with no cosmological dynamics involved.

The reconstruction from the nonlinear DM field gives an estimate for displacement field (See Zhu et al. (2016b)). This is not the case for the reconstruction from a biased tracer, i.e. the halo field. The reconstructed displacement from a biased halo sample does not responds to the halo displacement in the Lagrangian halo formation scenario. However, one could involve the above mentioned model assumptions (the fiducial cosmology, galaxy bias, growth rate, etc) to correct the bias and RSD effect prior to the reconstruction process. In this way, the reconstructed displacement field is more physically motivated and might have further applications. We leave this to future investigation.

We focused on investigating the dependence on the halo number density and redshifts. One immediate and urgent future work is the performance test on the halo samples with the RSD effect. The RSD effect is due to the structure growth, and thus contains important cosmological information. The observed position of an object is shifted by its peculiar velocity along the line-of-sight. This simply adds an extra offset on the real displacement field. Thus, the displacement field reconstructed from the observed density field automatically includes this additive offset. Since much less nonlinearities involved in the displacement potential, the modeling of RSD will be improved significantly. This also helps to simultaneously model the BAO and RSD signal.

To apply the proposed reconstruction method in observations, there are many observational issues need investigation, including the selection function, the survey geometry, and etc. We will investigate these issues in near future.

ACKNOWLEDGMENT

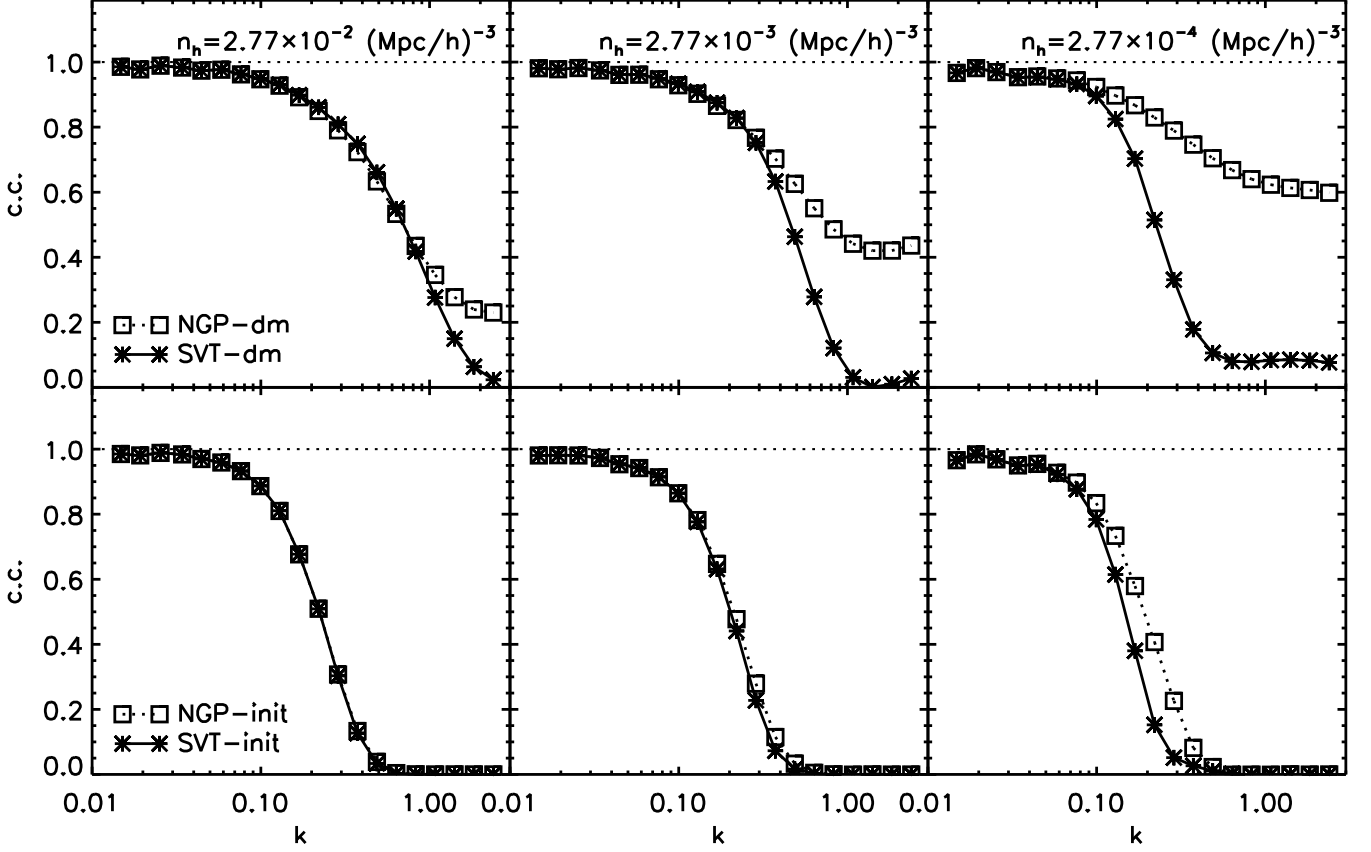


Figure 6. The upper row shows the cross-correlation coefficients between the halo fields with the simulated DM field. The solid lines with star present the result for the halo fields produces by the new mass assignment (simplified voronoi tessellation method). For comparison, the result for the NGP halo fields is plotted in dotted lines with square. From left to right, the halo number density is 2.77×10^{-2} , 2.77×10^{-3} , $2.77 \times 10^{-4} (h^{-1}\text{Mpc})^{-3}$, respectively. The bottom row shows the cross-correlation coefficients with the initial condition. In the calculation, shot noise is not subtracted and window functions are not deconvolved, since we want to see how different the two halo fields are.

We thank Pengjie Zhang, Xin Wang, Xuelei Chen, Kwan Chuen Chan, Haoran Yu, Morgan Bennett and Qiaoyin Pan for useful discussions. This work was supported by the national science foundation of China (Grant No. 11403071, 11373030), the Strategic Priority Research Program “The Emergence of Cosmological

Structures” of the Chinese Academy of Sciences (grant No. XDB09000000). The simulations are performed on the BGQ super-computer at the SciNet HPC Consortium. SciNet is funded by the Canada Foundation for Innovation under the auspices of Compute Canada, the Government of Ontario, the Ontario Research Fund Research Excellence, and the University of Toronto.

APPENDIX

HALO MASS ASSIGNMENT

We want to check the difference between the new mass assignment (simplified Voronoi tessellation method (SVT)) and the NGP method. We calculate the cross-correlation coefficients between the halo fields produced by these two methods with the DM field, for each halo number density. Note that the comparison here is not aiming at calculating the clustering statistics (e.g., the power spectrum) precisely. Thus, we do not subtract the shot noise and not deconvolve the window functions in all the power spectrum and cross-correlation coefficient calculation. The result is presented in upper row of Figure 6.

For the highest halo number density, the difference only appears at small scale $k > 1 h\text{Mpc}^{-1}$, much smaller than the scale where the recovered BAO information locates. Thus, the reconstruction performance test is reliable for this case. The difference is larger toward decreasing halo number density, and begins to show up at smaller k . For $n_h = 2.77 \times 10^{-3} (h^{-1}\text{Mpc})^{-3}$ case, the difference shows up at $k \sim 0.3 h\text{Mpc}^{-1}$, and for $n_h = 2.77 \times 10^{-4} (h^{-1}\text{Mpc})^{-3}$ it shows up at $k \sim 0.1 h\text{Mpc}^{-1}$.

In the bottom row of Figure 6, we present the cross-correlation coefficients with the initial condition. We find that, only for $n_h = 2.77 \times 10^{-4} (h^{-1}\text{Mpc})^{-3}$, the two halo fields from different mass assignment method contain slightly different linear information. Furthermore, the comparison between the upper row and bottom row implies that the halo field produced by the new mass assignment traces the DM field better than initial condition (i.e., to a smaller scale). This implies that there exists information exceeding the linear information and residing in the final density

field. This part of information could be used by the nonlinear reconstruction. Thus, the reconstruction performance is not limited by the mass assignment and the unsatisfying performance for $n_h = 2.77 \times 10^{-4} (h^{-1} \text{Mpc})^{-3}$ is due to the low number density instead of the effect from the new mass assignment.

REFERENCES

- Achitouv, I., & Blake, C. 2015, PRD , 92, 083523
- Anderson, L., Aubourg, É., Bailey, S., et al. 2014, MNRAS , 441, 24
- Bandura, K., Addison, G. E., Amiri, M., et al. 2014, in Proc. SPIE, Vol. 9145, Ground-based and Airborne Telescopes V, 914522
- Battye, R., Browne, I., Chen, T., et al. 2016, ArXiv e-prints, arXiv:1610.06826
- Beutler, F., Blake, C., Koda, J., et al. 2016, MNRAS , 455, 3230
- Beutler, F., Seo, H.-J., Ross, A. J., et al. 2017, MNRAS , 464, 3409
- Carron, J., & Neyrinck, M. C. 2012, ApJ , 750, 28
- Cautun, M. C., & van de Weygaert, R. 2011, The DTFE public software: The Delaunay Tessellation Field Estimator code, Astrophysics Source Code Library, arXiv:1105.0370
- Crocce, M., & Scoccimarro, R. 2006, PRD , 73, 063520
- . 2008, PRD , 77, 023533
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ , 145, 10
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, ArXiv e-prints, arXiv:1611.00036
- Eisenstein, D. J., Seo, H.-J., Sirko, E., & Spergel, D. N. 2007, ApJ , 664, 675
- Harnois-Déraps, J., Pen, U.-L., Iliev, I. T., et al. 2013, MNRAS , 436, 540
- Hinton, S. R., Kazin, E., Davis, T. M., et al. 2017, MNRAS , 464, 4807
- Jones, D. H., Read, M. A., Saunders, W., et al. 2009, MNRAS , 399, 683
- Kazin, E. A., Koda, J., Blake, C., et al. 2014, MNRAS , 441, 3524
- Matsubara, T. 2008, PRD , 77, 063530
- Mehta, K. T., Seo, H.-J., Eckel, J., et al. 2011, ApJ , 734, 94
- Newburgh, L. B., Bandura, K., Bucher, M. A., et al. 2016, in Proc. SPIE, Vol. 9906, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 99065X
- Neyrinck, M. C., Szapudi, I., & Rimes, C. D. 2006, MNRAS , 370, L66
- Noh, Y., White, M., & Padmanabhan, N. 2009, PRD , 80, 123501
- Obuljen, A., Villaescusa-Navarro, F., Castorina, E., & Viel, M. 2016, ArXiv e-prints, arXiv:1610.05768
- Padmanabhan, N., White, M., & Cohn, J. D. 2009, PRD , 79, 063523
- Padmanabhan, N., Xu, X., Eisenstein, D. J., et al. 2012, MNRAS , 427, 2132
- Pan, Q., Pen, U.-L., Inman, D., & Yu, H.-R. 2016, ArXiv e-prints, arXiv:1611.10013
- Pen, U.-L. 1995, ApJS , 100, 269
- . 1998, ApJS , 115, 19
- Rimes, C. D., & Hamilton, A. J. S. 2005, MNRAS , 360, L82
- . 2006, MNRAS , 371, 1205
- Ross, A. J., Samushia, L., Howlett, C., et al. 2015, MNRAS , 449, 835
- Schmittfull, M., Feng, Y., Beutler, F., Sherwin, B., & Chu, M. Y. 2015, PRD , 92, 123522
- Seo, H.-J., Beutler, F., Ross, A. J., & Saito, S. 2016, MNRAS , 460, 2453
- Seo, H.-J., Siegel, E. R., Eisenstein, D. J., & White, M. 2008, ApJ , 686, 13
- Seo, H.-J., Eckel, J., Eisenstein, D. J., et al. 2010, ApJ , 720, 1650
- Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ , 66, R1
- Taruya, A., Nishimichi, T., Saito, S., & Hiramatsu, T. 2009, PRD , 80, 123503
- White, M. 2015, MNRAS , 450, 3822
- Xu, X., Cuesta, A. J., Padmanabhan, N., Eisenstein, D. J., & McBride, C. K. 2013, MNRAS , 431, 2834
- Yu, H.-R., Pen, U.-L., & Zhu, H.-M. 2016, ArXiv e-prints, arXiv:1610.07112
- Zhu, H.-M., Pen, U.-L., & Chen, X. 2016a, ArXiv e-prints, arXiv:1609.07041
- Zhu, H.-M., Yu, Y., Pen, U.-L., Chen, X., & Yu, H.-R. 2016b, ArXiv e-prints, arXiv:1611.09638
- Zhu et al. 2017, in prep.