

# Increasing the Fisher Information Content in the Matter Power Spectrum through Lagrangian Space Reconstruction

Qiaoyin Pan,<sup>1,\*</sup> Ue-Li Pen,<sup>2,3,4,5,†</sup> Derek Inman,<sup>2,6</sup> and Hao-Ran Yu<sup>2,7</sup>

<sup>1</sup>*School of Physics, Nankai University, 94 Weijin Rd, Nankai, Tianjin, 300071, China*

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

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

<sup>3</sup>*Dunlap Institute for Astronomy and Astrophysics,*

*University of Toronto, Toronto, ON M5S 3H4, Canada*

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

<sup>5</sup>*Perimeter Institute for Theoretical Physics, Waterloo, ON, N2L 2Y5, Canada*

<sup>6</sup>*Department of Physics, University of Toronto, 60 St. George, Toronto, ON M5S 1A7, Canada*

<sup>7</sup>*Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China*

(Dated: November 24, 2016)

A new reconstruction method using *Moving-mesh* (MM) algorithm from *Adaptive Particle-mesh* (APM) algorithm originally aiming at improving the resolution of N-body simulation is recently introduced into cosmology matter field, which is expected to give a better reconstruction from non-linear density fields to linear ones in many cases. We are motivated to adapt this method to test the Fisher information difference, checking the effect of the reconstruction method. We reconstruct 136 non-linear density fields given by independent N-body simulations, in order to recover some of the information lost in the non-linear regime of large-scale structure. Through analyzing the power spectra of both density fields from simulations and deformation potentials from reconstructions, we find that after reconstruction, the non-linear regime of correlation matrix shrinks to  $k \sim 0.6$ . We also find that the cumulative Fisher information has a increase by a factor up to 40 at the translinear plateau,  $k \sim 0.8$ .

PACS numbers:

## I. INTRODUCTION

Power spectrum is widely used in modern cosmology to measure the matter fluctuations. In the early universe, initial Gaussian density fields can be completely described by the power spectrum, or the two-point statistics. However, gravitational instability and nonlinear large scale structure (LSS) formation drive the matter distribution highly non-Gaussian, and the galaxy distribution also follows this non-Gaussian distribution. In these cases one needs to compute higher statistics which are computationally more expensive and more difficult to interpret into initial cosmological parameters. Fisher information is usually used to quantify the amount of independent information that is contained in the power spectrum estimation.

Rimes and Hamilton [1] first studied the Fisher information as a function of scale contained in the matter power spectrum given by  $N$ -body simulations, and find that there is a plateau on trans-linear scales ( $k \simeq 0.2 - 0.8 h/\text{Mpc}$ ), which shows that on these scales, there is a strong coupling of Fourier modes and thus the power spectrum on smaller scales, gives little additional independent information.

There are many approaches to recover the lost information in the power spectrum of matter density field, by reversing the final density field into a more Gaussian, early stage density field. For example, Gaussianization

transforms are commonly used [2, 3] to make the logarithmic distribution more Gaussian. Nonlinear Wiener filters are used in wavelet space to Gaussianize the fields and can also improve the Fisher information [4–6]. It is shown in [6] that, although these methods or their combinations may have different abilities to recover the Fisher information, by means of reducing the mode coupling and variances in the auto power spectrum of Gaussianized density fields, they do not necessarily improve the cross correlation (propagator in their context) between the initial density field and the final density field, and thus result in a smearing out of the baryonic acoustic oscillations (BAO) peak in the two-point correlation function. If one concerns about mapping the initial conditions to final conditions (e.g. measurement of BAO) these methods are unable to extract valid information from initial conditions, at least in the cross power spectrum between initial and final conditions.

Reconstruction techniques (including the one described in [6]) are able to increase the Fisher information while also improve the cross correlation to the initial conditions and sharpen the BAO peak. It is based on the coupling of linear density field  $\delta_L(\mathbf{q}, t_0)$  to the displacement field  $\Psi(\mathbf{q})$ , first derived by [7], and the displacement field is estimated by a smoothed final density field. [8] shows a new method in the estimation of displacement field in 1-dimensional (1D), according to which the 1D linear density field is reconstructed in Lagrangian space and successfully improve the BAO measurement. In 3D cases, it is nontrivial to estimate the displacement field, but [9] show that the displacement field given by  $N$ -body simulations can be used to recover the  $\delta_L$ .

In this paper we generalize the displacement field esti-

\*Electronic address: panda@mail.nankai.edu.cn

†Electronic address: pen@cita.utoronto.ca

mation method from 1D [8] to 3D, reconstruct  $\delta_L$  and study the Fisher information recovery in  $\delta_L$ . Here, the displacement field estimation is done by a *Moving-mesh* (MM) algorithm, which is based on the *Adaptive Particle-mesh* (APM) simulation algorithm [10, 11].

This paper is organized as follows. In Section II, we present the main steps of the N-body simulation code that was used to simulate the dark matter density fields. In Section III, we briefly describe the reconstruction algorithm. In Section IV, we calculate and compare the power spectra, correlation matrix and Fisher information given by simulation and reconstruction. Conclusion and discussion are presented in Section V

## II. N-BODY SIMULATION OF DARK MATTER DENSITY FIELDS

We run 136 simulations with a box size of  $300 h^{-1}\text{Mpc}$ , resolution of  $1024^3$  cells and  $512^3$  particles, using the cosmological simulation code CubeP3M[12]. The initial condition is given by reading the transfer function computed by CAMB [13] and then evolving the power linearly to  $z = 100$ . Then Zel'dovich approximation is used to calculate the displacement field and velocity field, which are assigned to the particles. The cosmological parameters used are  $\Omega_M = 0.32$ ,  $\Omega_\Lambda = 0.679$ ,  $h = 0.67$ ,  $\sigma_8 = 0.83$ , and  $n_s = 0.96$ . Then the initial densities are evolved up to  $z = 0$ . Different seeds are used to produce the initial conditions for different simulations so that those simulations are independent to each other.

## III. RECONSTRUCTION ALGORITHM

The basic idea of MM algorithm is to build a PM scheme on a curvilinear coordinate system, in which the number of the particles per grid cell is set approximately constant. Consider a numerical grid of curvilinear coordinates  $\xi = (\xi_1, \xi_2, \xi_3)$ . In order to determine the physical position of each grid point, one needs to specify the Euclidean coordinate  $\mathbf{x}(\xi, t)$  as a function of grid position. In the Euclidean coordinate, the flat metric is Kronecker delta function  $\delta_{ij}$ , while the curvilinear metric is given by

$$g_{\mu\nu} = \frac{\partial x^i}{\partial \xi^\mu} \frac{\partial x^j}{\partial \xi^\nu} \delta_{ij}. \quad (1)$$

We use the convention that Latin indices denote Cartesian coordinate, while Greek indices denote the curvilinear grid coordinate. In principle, there are many different methods to connect the Cartesian coordinate and curvilinear coordinate of each grid cell. In APM method, the connection is described by an irrotational deformation,

$$x^i = \xi^\mu \delta_\mu^i + \Delta x^i, \quad (2)$$

where

$$\Delta x^i = \frac{\partial \phi}{\partial \xi^\nu} \delta_\nu^i. \quad (3)$$

This choice of the deformation can minimize mesh distortion and twisting.  $\phi$  is called the deformation potential, and  $\Delta x^i$  the lattice displacement. The deformation potential can be given in terms of the continuity equation in curvilinear coordinate,

$$\frac{\partial(\sqrt{g}\rho)}{\partial t} + \partial_\mu [\rho \sqrt{g} e_\mu^i (v^i - \Delta \dot{x}^i)] = 0 \quad (4)$$

where  $\sqrt{g} \equiv (\partial x^i / \partial \xi^\alpha)$  is the volume element and thus  $\sqrt{g}\rho$  represents the particle mass in the volume element under the curvilinear coordinate system.  $e_\mu^i = \partial \xi^\mu / \partial x^i$  is the triad.  $\Delta \dot{x}^i = \delta^{i\nu} \partial_\nu \dot{\phi}$  is chosen such that the first term in equation (4) is zero, resulting in a constant mass per volume element. And the velocity field divergence is replaced by the deviation density field  $\Delta \rho = \bar{\rho} - \rho \sqrt{g}$ , which ideally should be zero. Then the deformation potential is described in the elliptic equation,

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

The equation (5) can be solved using multigrid algorithm described in Ref. [10, 11]. Then the displacement is given by the gradient of the deformation potential as in (3). A 2-D projection of one layer of the deformed grids and the original density field on the selected deformed grids are given in Fig. II. As expected, there's no grid crossing after reconstruction.

## IV. POWER SPECTRA AND INFORMATION CONTENT

The power spectrum is the Fourier transform of the correlation function and measures the amount of clustering in the matter distribution in terms of the wavenumber  $k$  in unit of  $h/\text{Mpc}$ ,

$$\langle \delta(\mathbf{k}) \delta(\mathbf{k}') \rangle = (2\pi)^3 P(\mathbf{k}) \hat{\delta}(\mathbf{k} - \mathbf{k}'), \quad (6)$$

where  $\delta(\mathbf{k})$  is the density fluctuation in wave space, while  $\hat{\delta}$  is the delta function. Of equal interest is  $\Delta_k^2$ , the power spectrum in its dimensionless form, defined as

$$\Delta_k^2 \equiv \frac{k^3 P(k)}{2\pi^2} \quad (7)$$

The power spectra of the mass distributions are calculated using the "Nearest Grid Point" (NGP) mass assignment scheme, which calculates the position of each particle based on which grid point it is nearest. In Fig. 2(a) we plot the mean cross correlation function,  $r = P_{\delta\delta_L} / \sqrt{P_\delta P_{\delta_L}}$  of non-linear power spectrum and linear power spectrum, reconstructed power spectrum and linear power spectrum respectively. One can see that after reconstruction, the damping scale of cross correlation shrinks greatly. The wave number where cross correlation drops to a half increase from  $k \sim 0.2$  to  $k \sim 0.6$ . In Fig. 2(b) we plot the linear power spectrum which is the transfer function, and mean power spectrum (with error bars) of 136 non-linear density fields and reconstructed density fields simply given by  $\delta_r = \mathbf{k} \cdot \mathbf{k}\xi$ . The

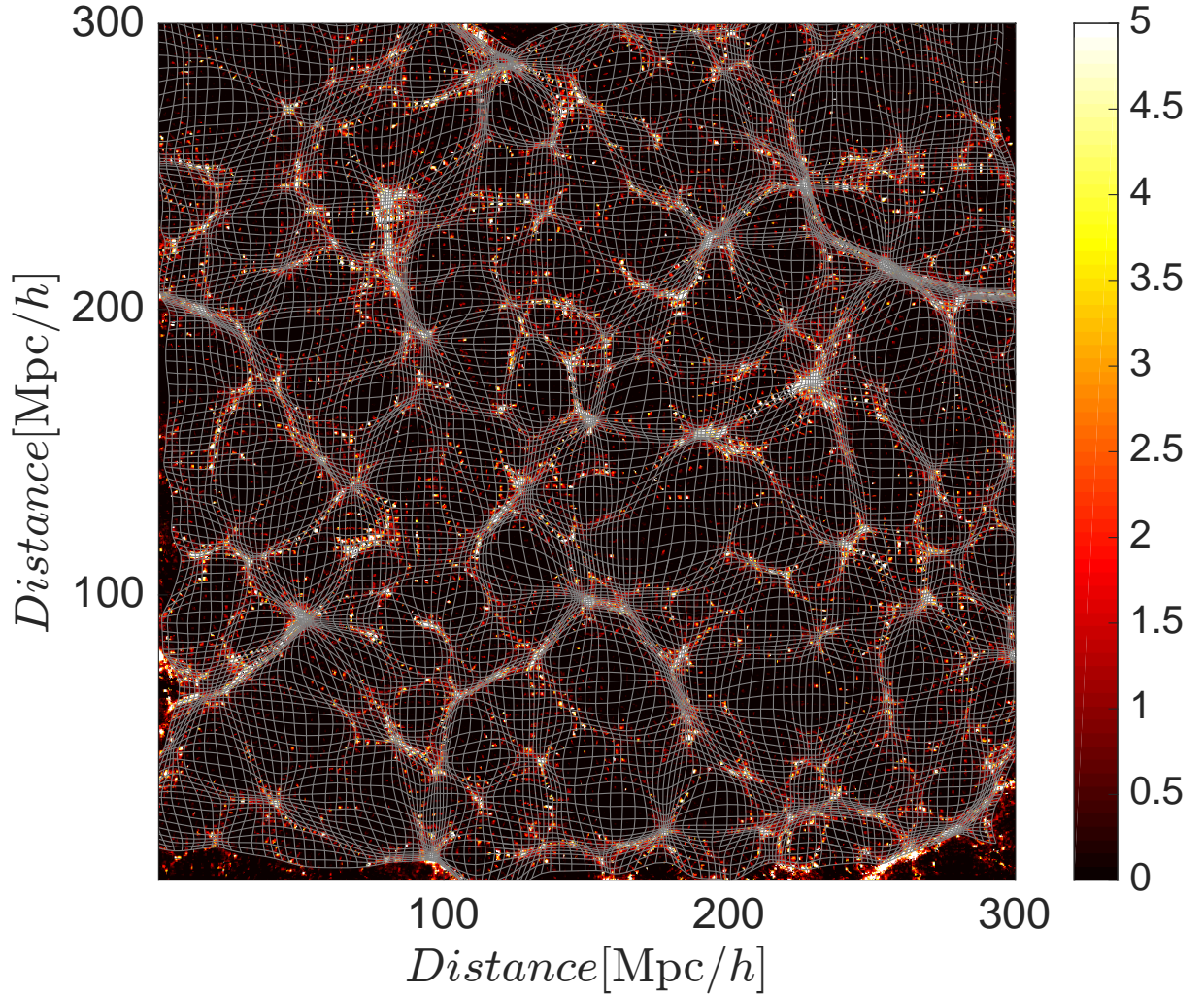


FIG. 1: 2-D projection of one layer of the deformed grids of one out of 136 N-body simulations, and the density field on the selected deformed grids, with the magnitude as the number of particles per cell. The simulations have  $300 h^{-1}\text{Mpc}$  width box and  $1024^3$  pixels.

reconstructed power spectrum drops at non-linear scale ( $k \gtrsim 0.3$ ) since the reconstructed density fields are totally irrotational. The result is similar to that of E-mode displacement reconstruction described in [9], in which the reconstructed power spectrum drops, but in a different scale and a different speed.

Mathematically, Fisher information [14]  $I$  in the log of amplitude  $A$  of the initial matter power spectrum is defined as

$$I_A \equiv - \left\langle \frac{\partial^2 \ln \mathcal{L}}{\partial A^2} \right\rangle, \quad (8)$$

in which  $\mathcal{L}$  denotes the likelihood. For Gaussian fluctuations, the likelihood depends on parameters only through the power spectrum  $P(k)$ , and the information  $I$  in  $A$  defined by (8) can be written as [1]

$$I_A = - \left\langle \sum_{k,k'} \frac{\partial \ln P(k)}{\partial \ln A} \frac{\partial^2 \ln \mathcal{L}}{\partial \ln P(k) \partial \ln P(k')} \frac{\partial \ln P(k')}{\partial \ln A} \right\rangle \quad (9)$$

in which the angle bracket denotes the average of all the power spectra samples.

For any density field  $\delta$ , we can conveniently decompose it as

$$\delta(k) = b(k)\delta_L(k) + n(k), \quad (10)$$

in which  $\delta_L$  denotes the linear density field. Safely select  $b(k)$  and  $n(k)$  such that the correlation  $\langle \delta_L(k)n(k) \rangle$  is zero. Correlating the density field  $\delta$  and linear density field  $\delta_L$ ,

$$\langle \delta(k)\delta_L(k) \rangle = b(k)\langle \delta_L(k)\delta_L(k) \rangle, \quad (11)$$

we obtain

$$b(k) = \frac{P_{\delta\delta_L}(k)}{P_{\delta_L\delta_L}(k)}. \quad (12)$$

Nonlinear evolution drives  $b(k)$  to drop from unity, and generates the noise  $n(k)$ . Correlating the density field  $\delta$  and itself, we can write its power spectrum as

$$P_\delta(k) = \mathcal{D}(k)P_{\delta_L}(k) + P_n(k), \quad (13)$$

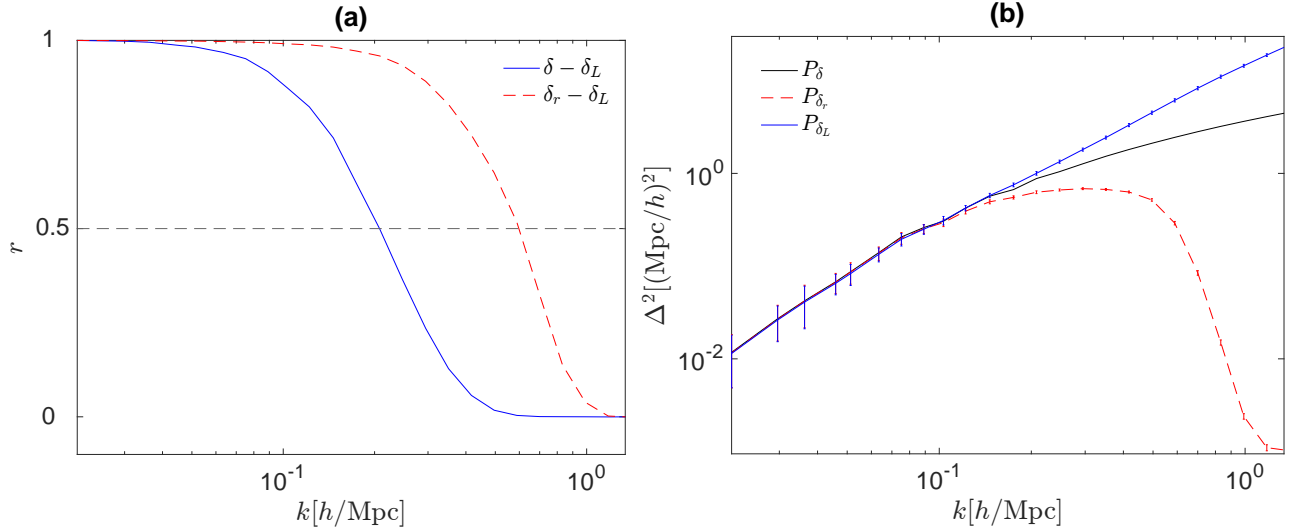


FIG. 2: (a) Cross correlation function  $r(\delta, \delta_L)$  (blue line) and  $r(\delta_r, \delta_L)$  (red dash line). (b) Linear power spectrum given by transfer function (black line), mean power spectrum with error bar of 136 non-linear density fields (blue line) and reconstructed density fields (red dash line)

where  $\mathcal{D}(k) \equiv b^2(k)$  is the non-linear damping factor, which is  $\exp(-k^2 \Sigma^2/2)$  in Gaussian BAO damping model.  $P_n$  is the mode-coupling term.

With the help of (12) and (13), we can replace the partial derivatives  $\partial \ln p(k)/\partial \ln A$  in (9) with the square of crosscorrelation function  $r^2(k)$  of  $\delta$  and  $\delta_L$ . the second partial derivative terms in 9, the Hessian of the vector  $\ln P(k)$ , has the expectation value of the Fisher matrix with respect to the log powers. For linear density fields, the Fisher matrix is approximately equal to the inverse of the covariance matrix of power spectrum estimates, which should be diagonal, with diagonal elements equal to the number of modes in each wavenumber bin (when considering  $\mathbf{k}$  and  $-\mathbf{k}$  as the same mode). Thus we can write down a simpler form of cumulative Fisher information in (9),

$$I_A(< k_n) = r^2(k)^T [C_{\text{norm}}^{-1}(k_i, k_j)] r^2(k')(i, j \leq n), \quad (14)$$

where  $C_{\text{norm}}$  is the normalized covariance matrix with size per dimension up to  $k_n$ , defined as

$$C_{\text{norm}}(k, k') = \frac{\text{Cov}(k, k')}{\langle P(k) \rangle \langle P(k') \rangle}, \quad (15)$$

and  $r$  is the mean cross correlation of a given density field and linear one as a function of  $k$  up to  $k_n$ . It's reliable to define samely as (14) for non-linear density fields, since the Fisher matrix is approximately the same as that of linear density fields in linear scale. The covariance matrix is defined as

$$\text{Cov}(k, k') \equiv \frac{\sum_{i=1}^N [P_i(k) - \langle P(k) \rangle] [P_j(k') - \langle P(k') \rangle]}{N-1}, \quad (16)$$

where angle brackets mean the expected values, and  $N$  is the total number of simulations. The correlation matrix,

normalized version of the covariance matrix,

$$\text{Corr}(k, k') = \frac{\text{Cov}(k, k')}{\sqrt{\text{Cov}(k, k) \text{Cov}(k', k')}}, \quad (17)$$

represents the correlation between different  $k$  modes. The correlation matrices for non-linear and reconstructed power spectra from reconstructions are shown in Fig. 3. For the non-linear power spectra, the correlation matrix in linear regime,  $k \lesssim 0.2$ , is almost diagonal. The off-diagonal elements are produced by strong mode coupling in non-linear scale, and super-survey tidal effect which is small in linear scale but dominates in the weakly non-linear regime [15]. The correlation matrix for non-linear power spectra has few negative elements, smallest  $\text{Corr} \simeq -0.1$ , which are produced by the unbiased error and thus will vanish with more simulations [16]. For the reconstructed correlation matrix, however, the linear regime expand up to  $k \simeq 0.6$ , but the number and magnitude of negative off-diagonal elements increase, smallest  $\text{Corr} \simeq -0.8$ .

We plot the cumulative Fisher information of the non-linear, linear and reconstructed power spectra in Fig.4(a). The Fisher information of the non-linear power spectra drops from the linear one at  $k \sim 0.05$ , and has a flat plateau in the translinear regime,  $k \sim 0.3$ . It indicates that there's nearly no independent information in the translinear regime of the power spectrum. But the information curve of the reconstructed power spectra keeps increasing roughly the same as the linear information until  $k \sim 0.3$ , and reaches it's plateau at  $k \sim 0.8$  up to a factor of 40. It indicates that MM reconstructed method can strongly recover the lost information within this scale. We compare the Fisher information given by MM reconstruction method with logarithmic density mapping method [3] as an example to illustrate their strength. We find that MM reconstruction gives more than 10 times more information than logarithmic mapping. In



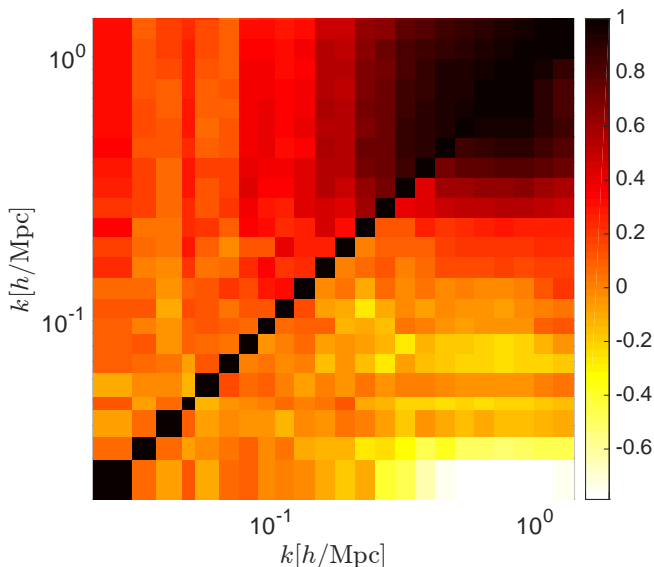


FIG. 3: Correlation coefficient matrix as found from 136 non-linear power spectra (the upper-left elements) and the reconstructed power spectra (the lower-right off-diagonal elements).

some papers, the cross-correlation  $r^2$  terms are set to be unity in (14), which apparently increases the non-linear information. We also plot those in Fig. 4(b) for better comparison. But we find that in this case, the MM reconstructed and logarithmic mapping information in the scale  $k \sim 0.2 - 0.5$  is higher than the linear one, which is not expected.

## V. CONCLUSION AND DISCUSSION

We use the code "CubeP3M" to generate 136 independent dark matter density fields, then give the reconstructed deformation potentials which are pure divergent using MM algorithm. We analyze the power spectra of both the non-linear density fields and reconstructed density fields, after which we give the cross correlation matrix. We find that the power spectra are highly correlated on small scales, since these scales are in non-linear regime. But after reconstruction, the strongly correlated regime sinks from  $k \sim 0.2$  to  $k \sim 0.6$ . We also calculate the cumulative information, and find that the plateau of the reconstructed information curve keeps matching the Gaussian information curve until  $k \sim 0.3$  and reach a plateau at  $k \sim 0.8$ , rising by a factor of 40 compared to the non-linear information curve.

We argue that the reconstructed Fisher information will still increasing to a greater magnitude in smaller scale since the cross correlation of the reconstructed and linear power spectrum increases in a higher resolution analysis cite(bib:ZhuH2016).

The new reconstruction method successfully recovers the lost linear information on the mildly non-linear scale, better than previous methods [3–5, 17, 18] and pushes the non-linear scale to a smaller scale in our case. The result in dark matter density fields gives a strong motivation to adapt MM reconstruction in halo fields, neutrino fields, etc, so that we have access to know more clearly about the physics in smaller scale. Some efforts were made to improve cosmology measurements to BAO scale (e.g. [19, 20]). MM reconstruction gives the reconstructed displacement given on the Lagrangian coordinates transformed from the Eulerian coordinates. It's successful try on BAO reconstruction in 1-D [8] and 3-D cite(bib:ZhuH2016) cosmology provided an intuitive view of the algorithm to push forward the BAO research.

MM reconstruction effectively decomposes the irrotational part and the curl part of the displacement field of particles. The decomposition is not clean, in other words, the difference between the reconstructed displacement and the real displacement in N-body simulation includes part of irrotational displacement as well. On the other hand, the reconstructed displacement might be greatly different from real displacement in N-body simulation, since it's sensitive to the late stage shell-crossing and non-linear process so that the original position of some specific particles can be replaced by each other. It's meaningful to compare the irrotational displacement field through MM reconstruction and that from E-mode displacement reconstruction [9] which decomposes cleanly the irrotational and curl part of the real displacement in N-body simulation. Since MM reconstruction only needs the density field input and gives a large recovering of lost information, it's expected to have a good effect on reconstructing the matter density field from observation.

## Acknowledgments

We thank Hong-Ming Zhu, Yu-Yu and Xin Wang for friendly and helpful discussions. Computations were performed on the General Purpose Cluster supercomputer at the SciNet HPC Consortium. SciNet is funded by: the Canadian Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto.

- 
- [1] C. D. Rimes and A. J. S. Hamilton, in *American Astronomical Society Meeting Abstracts #207* (2006), vol. 207 of *American Astronomical Society Meeting Abstracts*, p. 206.01.
  - [2] D. H. Weinberg, *MNRAS* **254**, 315 (1992).
  - [3] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, *ApJ* **698**,

- L90 (2009), 0903.4693.
- [4] T.-J. Zhang, H.-R. Yu, J. Harnois-Déraps, I. MacDonald, and U.-L. Pen, *ApJ* **728**, 35 (2011), 1008.3506.
- [5] H.-R. Yu, J. Harnois-Déraps, T.-J. Zhang, and U.-L. Pen, *MNRAS* **421**, 832 (2012), 1012.0444.
- [6] J. Harnois-Déraps, H.-R. Yu, T.-J. Zhang, and U.-L. Pen,

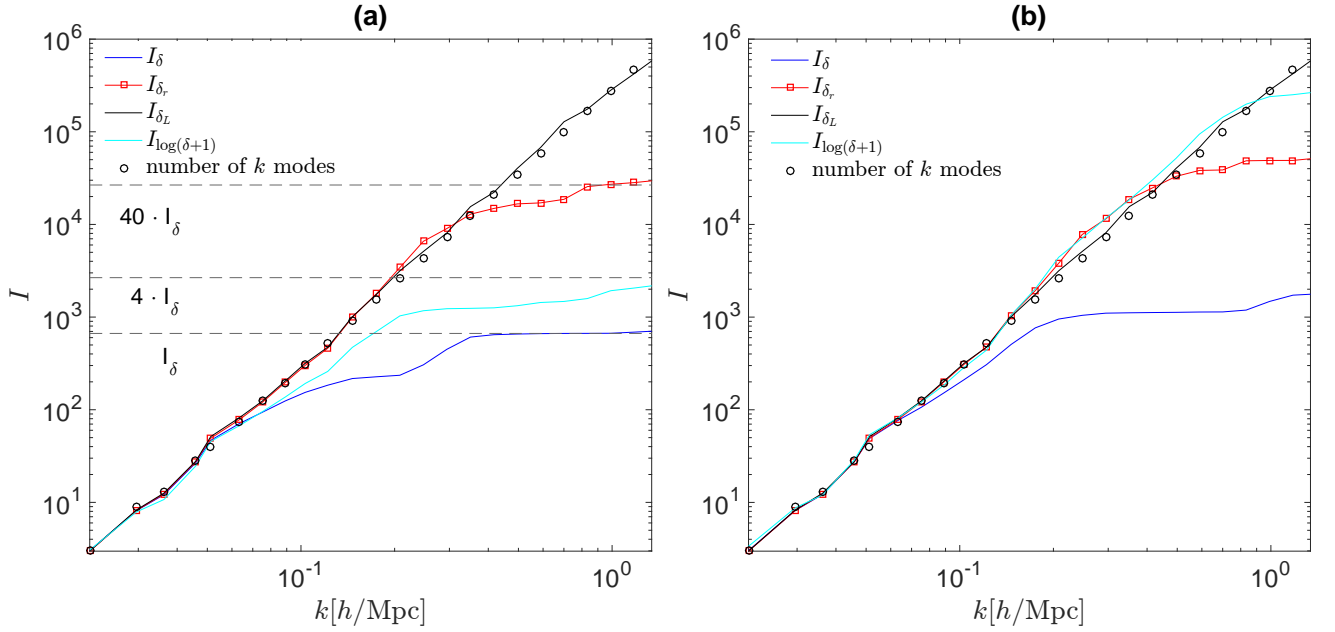


FIG. 4: (a) Cumulative information in the power spectra as a function of wavenumber. The blue line correspond to the non-linear density field by simulation; the red line with squares correspond to the reconstructed deformation potential; the dark line corresponds to linear density field; the cycles correspond to number of  $k$  modes up to that wave bin. (b) Cumulative information given by setting the cross-correlation to be unity.

- MNRAS **436**, 759 (2013), 1205.4989.
- [7] Y. B. Zel'dovich, A&A **5**, 84 (1970).
- [8] H.-M. Zhu, U.-L. Pen, and X. Chen, ArXiv e-prints (2016), 1609.07041.
- [9] H.-R. Yu, U.-L. Pen, and H.-M. Zhu, ArXiv e-prints (2016), 1610.07112.
- [10] U.-L. Pen, ApJS **100**, 269 (1995).
- [11] U.-L. Pen, ApJS **115**, 19 (1998), astro-ph/9704258.
- [12] J. Harnois-Déraps, U.-L. Pen, I. T. Iliev, H. Merz, J. D. Emberson, and V. Desjacques, MNRAS **436**, 540 (2013), 1208.5098.
- [13] A. Lewis, A. Challinor, and A. Lasenby, ApJ **538**, 473 (2000), astro-ph/9911177.
- [14] M. Tegmark, A. N. Taylor, and A. F. Heavens, ApJ **480**, 22 (1997), astro-ph/9603021.
- [15] K. Akitsu, M. Takada, and Y. Li, ArXiv e-prints (2016), 1611.04723.
- [16] R. Takahashi, N. Yoshida, M. Takada, T. Matsubara, N. Sugiyama, I. Kayo, A. J. Nishizawa, T. Nishimichi, S. Saito, and A. Taruya, ApJ **700**, 479 (2009), 0902.0371.
- [17] M. C. Neyrinck, I. Szapudi, and C. D. Rimes, MNRAS **370**, L66 (2006), astro-ph/0604282.
- [18] M. C. Neyrinck, in *Statistical Challenges in 21st Century Cosmology*, edited by A. Heavens, J.-L. Starck, and A. Krone-Martins (2014), vol. 306 of *IAU Symposium*, pp. 251–254, 1407.4815.
- [19] D. J. Eisenstein, H.-J. Seo, E. Sirko, and D. N. Spergel, ApJ **664**, 675 (2007), astro-ph/0604362.
- [20] M. White, MNRAS **450**, 3822 (2015), 1504.03677.