

Sparsity Weak Lensing 3-D Density Map Reconstruction: Reducing the line of sight smearing.

ABSTRACT

A new method is developed to reconstruct 3-D density contrast maps from photometric weak-lensing shear measurements. The 3-D density contrast maps are modeled as a summation of the NFW basis atoms which have 2-D multi-scale NFW surface density profiles on the transverse plane and 1-D Dirac delta functions in the line of sight direction. With the prior assumption that the density fields can be sparsely represented by the NFW basis atoms, it is reconstructed using an oracle algorithm: adaptive lasso. Our method is tested with realistic simulations using the HSC-like shape noise and photo- z uncertainty. Our findings are summarized as follows: 1) The global minima solution of the standard lasso algorithm suffers from a smear of structure in the line of sight direction even in the absence of shape noise. 2) The global minima solution of adaptive lasso algorithm significantly removes the line of sight smear of structure.

1. INTRODUCTION

Light from distant galaxies is distorted by the intervening inhomogeneous density distribution along the line of sight due to the influence of gravity. As a result of the light distortion, the shapes of the background galaxies are coherently distorted. Such effect, which is known as weak-lensing, imprints the information of foreground mass density distribution to the background galaxy images and offers a direct probe into the mass density distribution in our universe (see [Kilbinger 2015](#); [Mandelbaum 2018](#), for recent reviews).

The expected shear measurements (γ) on distant galaxies are related to the foreground density contrast field (δ) through a linear transformation

$$\gamma = \mathbf{T}\delta, \quad (1)$$

where \mathbf{T} is used to denote the linear transformation operator which includes not only physical lensing effect but also systematic effects from observations (e.g. pixelization and smoothing of shear field in transverse plane, photo- z uncertainty).

Several large scale surveys target to study the weak-lensing effect at high precision level (e.g. HSC ([Aihara et al. 2018](#)), KIDS ([de Jong et al. 2013](#)), DES ([The Dark Energy Survey Collaboration 2005](#)), LSST ([LSST Science Collaboration et al. 2009](#)), Euclid ([Laureijs et al. 2011](#)), WFIRST ([Spergel et al. 2015](#))).

The primary goal of most weak-lensing surveys is to constrain the cosmology model through 2-point correlations. The studies include galaxy-galaxy lensing which cross correlating the shear field (γ) with the positions of foreground galaxies ([Han et al. 2015](#); [More et al. 2015](#); [Prat et al. 2018](#)), and cosmic shear which auto-correlates the shear measurements ([Morrison et al. 2016](#); [Troxel](#)

[et al. 2018](#); [Hikage et al. 2019](#); [Hamana et al. 2020](#)). Since shear is directly related to the matter distribution as shown in eq. (1), Galaxy-galaxy lensing probes into the correlation between the matter field and galaxy field, on the other hand, cosmic shear probes into the auto-correlation of matter field.

The reconstruction of density map from shear measurements also receive considerable interest as it reaches the nonlinear scales. 2-D density map reconstruction which recover an integration of projected mass along the line of sight has been well studied within the community ([Kaiser & Squires 1993](#); [Lanusse et al. 2016](#); [Price et al. 2020](#)) and applied to large scale surveys ([Oguri et al. 2018](#); [Chang et al. 2018](#); [Jeffrey et al. 2018](#)). However, the reconstruction of 3-D mass map is still a challenging task.

In order to fully reconstruct the 3-D mass density distribution (δ) from the photometric shear observations (γ), the density contrast field is modeled as a summation of basis atoms in a model dictionary

$$\delta = \Phi x, \quad (2)$$

where Φ is the transformation operator from the parameters in the dictionary space to the density contrast and x denotes the parameters. [Simon et al. \(2009\)](#) reconstruct the density field in Fourier space, which is equivalent to model the mass field with sinusoidal functions. On the other hand, [Leonard et al. \(2014\)](#) models the mass field with Starlets ([Starck et al. 2015](#)).

The projection coefficients are estimated through optimizing a regularized loss function. The estimator is generally defined as

$$\hat{x} = \arg \min_x \left\{ \frac{1}{2} \left\| \Sigma^{-\frac{1}{2}} (\gamma - \mathbf{T}\Phi x) \right\|_2^2 + \lambda C(x) \right\}, \quad (3)$$

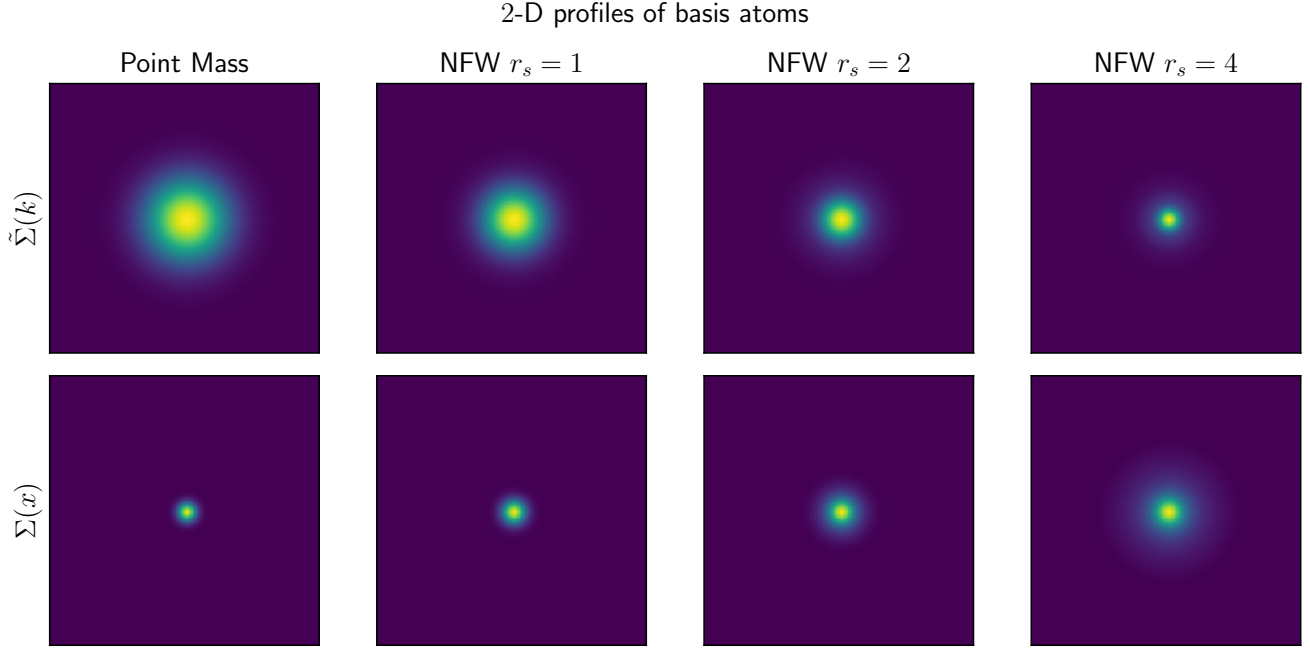


Figure 1. The smoothed basis atoms. The first row shows the smoothed basis atoms in Fourier space and the second row shows the smoothed basis atoms in Real space. The first column is point mass atom and the other columns are multi-scale NFW atoms. The smoothing kernel is Gaussian with scale of 1.5 pixels.

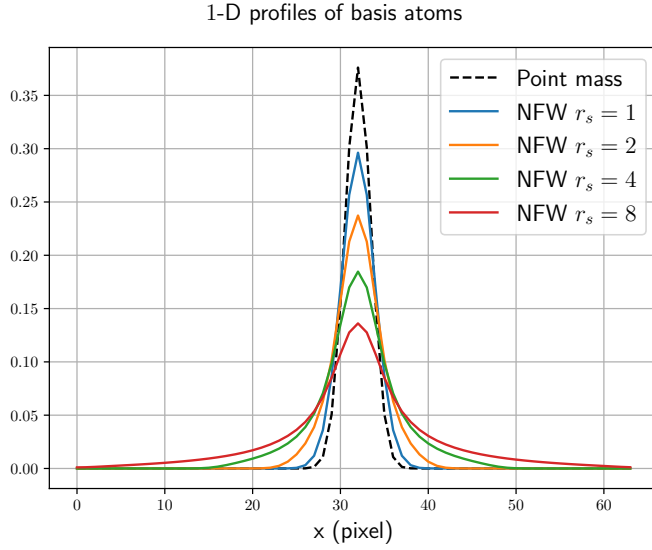


Figure 2. The 1-D slices for smoothed basis atoms at $x = 0$. The corresponding 2-D profiles are shown in Figure 1.

where $\left\| \Sigma^{-\frac{1}{2}}(\gamma - \mathbf{T}\Phi x) \right\|_2^2$ is the l^2 chi-square term¹ measuring the residuals between the prediction and the data, while $C(x)$ is the regularization term measuring the deviation of the estimation of the parameter (x) from the

¹ weighted by the inverse of the diagonal covariance matrix of error on the shear measurements (Σ).

prior assumptions. The l^p norm is defined as

$$\|x\|_p = \left(\sum_i |x_i|^p \right)^{\frac{1}{p}}. \quad (4)$$

Such estimation prefers the parameters that are able to describe the observations and also align with the prior assumptions. The regularization parameter λ adjusts the relative weight between the observations and prior assumptions in the optimization process.

Simon et al. (2009) propose to use the Wiener filter, which is also known as l^2 ridge regulation ($C = \|x\|_2^2$), to find a regularized solution in Fourier space. Oguri et al. (2018) apply the method of Simon et al. (2009) to the first-year data of the Hyper Suprime-Cam Survey (Aihara et al. 2018). However, the density maps reconstructed by this method suffer from a smearing along the line of sight direction with a standard deviation of $\sigma_z = 0.2 \sim 0.3$.

Leonard et al. (2014) propose to use a derivative version of l^1 lasso regulation ($C = \|x\|_1^1$) to find a sparse solution in the Starlets dictionary space (Starck et al. 2015). Leonard et al. (2014) apply a greedy coordinate descent algorithm, which selects the steepest coordinate in each iteration, to find the minima of a non-convex loss function penalized with the firm thresholding function. Leonard et al. (2014) significantly reduce the smearing along the line of sight, however, the stability of the non-convex optimization and greedy coordinate descent algorithm has not been fully justified. Moreover, the Starlets

functions are not designed to model the profile of clumpy mass in the universe.

N -body simulations have shown that the dark matter is distributed in halos connected by filaments, and the density profile of a single halo follows the NFW function (Navarro et al. 1997). We construct a dictionary with multi-scale NFW atoms. The atoms follow multi-scale surface density profiles of NFW functions (Takada & Jain 2003) on the transverse plane. We follow Leonard et al. (2014) to neglect the depth of halos since the resolution scale of the reconstruction in the line of sight direction is much larger than the scales of halos. Therefore, we set the profile of the NFW atoms in the line of sight as the Dirac delta function. The multi-scale NFW atoms are used to model the profiles of halos which is assumed to be sparsely distributed in the universe. With the sparsity prior, the adaptive lasso regularization (Zou 2006) is used to reconstruct the density field. We find that, in contrast to the standard lasso estimator which smears the structure along the line of sight, the adaptive lasso is able to significantly reduce the smearing effect.

Comparing with Leonard et al. (2014), our dictionary is built up to describe the clumpy mass in the universe which has a clear physical motivation. Furthermore, the adaptive lasso algorithm is strictly convex and can be directly optimized with the FISTA algorithm (Beck & Teboulle 2009) without relying on any greedy coordinate descent approaches. The stability of this convex optimization has been well studied.

This paper is organized as follows. In Section 2, we propose the new method for 3-D density map reconstruction. In Section 3, we test the novel algorithm on HSC-like simulations. In Section 4, we summarize and discuss the future development of the method.

2. METHODOLOGY

We first review the lensing process in section 2.1. Then, we introduce the dictionary used to model the foreground density maps in section 2.2.

Subsequently, in section 2.3, we discuss several systematic effects from observations which include photo- z uncertainty (section 2.3.1), smoothing (section 2.3.2), masking (section 2.3.3), and pixelization (section 2.3.4).

Finally, we solve the mass reconstruction problem in section 2.4 using the adaptive lasso algorithm (Zou 2006) optimized with the FISTA algorithm (Beck & Teboulle 2009).

2.1. Lensing

The lensing convergence map at the comoving distance χ_s caused by the foreground inhomogeneous density distribution at the comoving distance χ_l ($\chi_l < \chi_s$) along

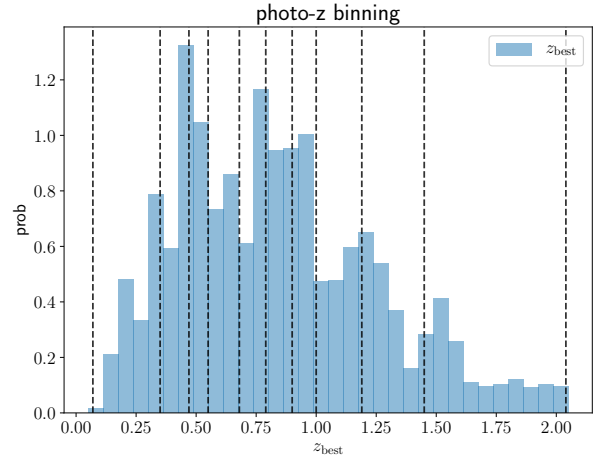


Figure 3. The source galaxies are binned into 10 redshift bins according to their MLZ best photo- z estimation. The blue histogram is the number distribution of the best photo- z estimation. The vertical dashed lines are the bounds of bins. The galaxies are evenly distributed in each bins.

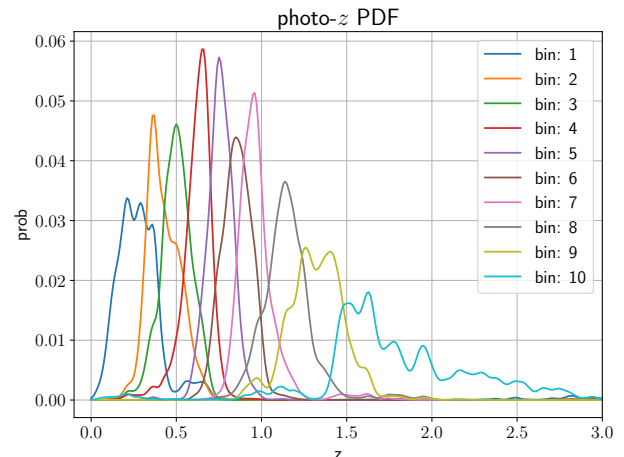


Figure 4. The average PDF of MLZ photo- z error for 10 source redshift bins.

the line of sight is

$$\kappa(\vec{\theta}, \chi_s) = \frac{3H_0^2\Omega_M}{2c^2} \int_0^{\chi_s} d\chi_l \frac{\chi_l \chi_{sl}}{\chi_s} \frac{\delta(\vec{\theta}, \chi_l)}{a(\chi_l)}, \quad (5)$$

where $\delta = \rho(\vec{\theta}, \chi_l)/\bar{\rho} - 1$ is the density contrast at the position of lens, H_0 is the Hubble parameter, Ω_M is the matter density parameter, c is the speed of light, and $a(\chi_l)$ is the scale parameter at the lens position.

Substitute comoving distance (χ) with redshift (z), we have

$$\kappa(\vec{\theta}, z_s) = \int_0^{z_s} dz_l K(z_l, z_s) \delta(\vec{\theta}, z_l). \quad (6)$$

where $K(z_l, z_s)$ is the lensing kernel defined as

$$K(z_l, z_s) = \begin{cases} \frac{3H_0\Omega_M}{2c} \frac{\chi_l \chi_{sl}(1+z_l)}{\chi_s E(z_l)} & (z_s > z_l), \\ 0 & (z_s \leq z_l), \end{cases} \quad (7)$$

where $E(z)$ is the Hubble parameter as a function of redshift, in units of H_0 .

As shown in Kaiser & Squires (1993), the shear field is related to the kappa field at the same redshift plane via

$$\gamma_L(\vec{\theta}, z_s) = \int d^2\theta' D(\vec{\theta} - \vec{\theta}') \kappa(\vec{\theta}', z_s), \quad (8)$$

where

$$D(\vec{\theta}) = -\frac{1}{\pi} (\theta_1 - i\theta_2)^{-2}. \quad (9)$$

Here we denote the physical shear distortion field as γ_L and we note that the final observed shear measurements are also physical shear distortions influenced by systematic errors from observations. The systematic errors will be discussed in Section 2.3.

Combining equation (6) with equation (8), the expectation of lensing shear signal is

$$\gamma_L(\vec{\theta}, z_s) = \int_0^{z_s} dz_l K(z_l, z_s) \int d^2\theta' \vec{D}(\vec{\theta} - \vec{\theta}') \delta(\vec{\theta}', z_l). \quad (10)$$

To simplify the expression, we define the lensing transform operator as

$$\mathbf{Q} = \int_0^{z_s} dz_l K(z_l, z_s) \int d^2\theta' \vec{D}(\vec{\theta} - \vec{\theta}'), \quad (11)$$

and eq. (10) is simplified to

$$\gamma_L = \mathbf{Q}\delta. \quad (12)$$

2.2. Dictionary

The density contrast field is modeled as a summation of basis atoms in the dictionary

$$\delta(\vec{r}) = \sum_{s=1}^N \int d^3r' \phi_s(\vec{r} - \vec{r}') x_s(\vec{r}'), \quad (13)$$

where $\phi_s(\vec{r})$ are the basis atoms of the dictionary. The basis atoms have 'N' different scale frames and the atoms in each scale frame are shifted by \vec{r}' to form basis at different positions. $x_s(\vec{r}')$ is the projection coefficients of the density contrast field onto the basis atoms.

We propose to use multi-scale NFW atoms, denoted as $\{\phi_1, \dots, \phi_N\}$, as our dictionary basis. On the transverse plane, the NFW atoms follow surface density profiles of NFW halos (Takada & Jain 2003) with scale radius θ_α and truncation radius $c\theta_\alpha$, where c is the concentration of NFW halos. As the scales of halos are much less than

the reachable redshift resolution, we neglect the depth of halo on the line of sight direction and set the profiles of NFW atoms in the line of sight direction to 1-D Dirac delta functions as suggested by (Leonard et al. 2014). The multi-scale NFW atoms are defined as

$$\phi_\alpha(\vec{r}) = \frac{f}{2\pi\theta_\alpha^2} F(|\vec{\theta}|/\theta_\alpha) \delta_D(z), \quad (14)$$

($s = 1..N$)

where

$$F(x) = \begin{cases} -\frac{\sqrt{c^2-x^2}}{(1-x^2)(1+c)} + \frac{\text{arccosh}\left(\frac{x^2+c}{x(1+c)}\right)}{(1-x^2)^{3/2}} & (x < 1), \\ \frac{\sqrt{c^2-1}}{3(1+c)} \left(1 + \frac{1}{c+1}\right) & (x = 1), \\ -\frac{\sqrt{c^2-x^2}}{(1-x^2)(1+c)} + \frac{\arccos\left(\frac{x^2+c}{x(1+c)}\right)}{(x^2-1)^{3/2}} & (1 < x \leq c), \\ 0 & (x > c). \end{cases} \quad (15)$$

$f = 1/[\ln(1+c) - c/(1+c)]$. In this work, we fix $c = 4$ for NFW atoms in different scale frames.

To simplify the notation, we define the projection pa-

rameters as a column vector: $x = \begin{pmatrix} x_0 \\ x_1 \\ \dots \\ x_N \end{pmatrix}$, and the dic-

tionary transform operator as a row vector:

$$\Phi = \left(\int d^3r \phi_0(\vec{r}) \int d^3r \phi_1(\vec{r}) \dots \int d^3r \phi_N(\vec{r}) \right). \quad (16)$$

We substitute eq. (13) into eq. (10)

$$\gamma_L = \mathbf{Q}\Phi x. \quad (17)$$

In this paper, dictionary constructed with point mass atoms is used to compare with the dictionary of multi-scale atoms. The point mass atoms is a 3-D Dirac function defined as follows

$$\phi_{\text{PM}}(\vec{r}) = \delta_D(\theta_1) \delta_D(\theta_2) \delta_D(z). \quad (18)$$

The 2-D profiles of the point mass atom and the multi-scale NFW atoms on the transverse plane are shown in Figure 1. The 1-D slices of the profiles are demonstrated in Figure 2. Note that these profiles are smoothed with a Gaussian kernel and pixelized into evenly spaced grids. The smoothing operation is discussed in Section 2.3.2 and the pixelization operation is discussed in Section 2.3.4.

2.3. Systematics

The observed shear measurements are deviated from the physical shear prediction due to the systematic errors from observations. The influence of systematics is carefully studied and incorporated into the forward modeling in this section.

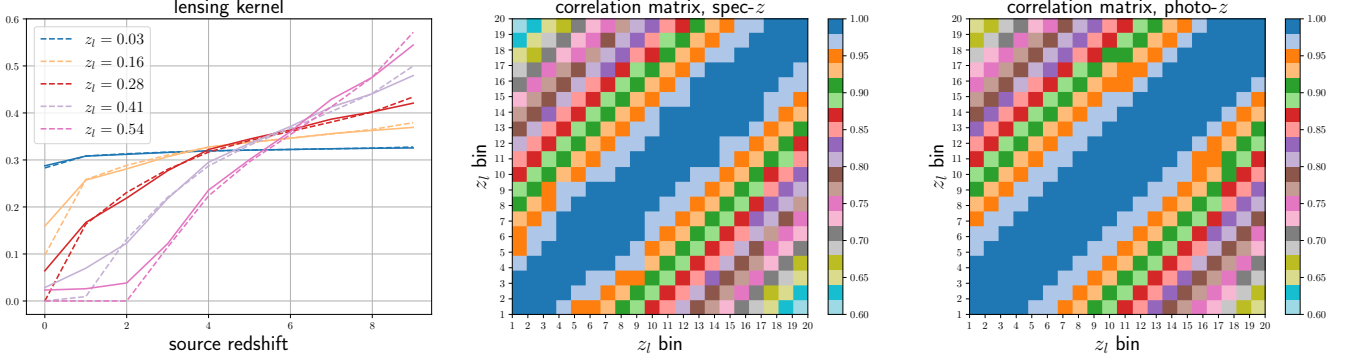


Figure 5. The left panel shows the lensing kernels for five different lens redshifts. The dashed lines are the kernels for spectroscopic redshift which assumes that the redshifts of source galaxies are precisely estimated. The solid lines are for photometric redshifts which accounts for the influence of photometric redshift uncertainty. The other two panels show the correlation between lensing kernels of different lens redshifts. The middle panel is for spectroscopic redshift and the right panel is for photometric redshifts. The lensing kernels are normalized so that the diagonal elements of the correlation matrices equal one.

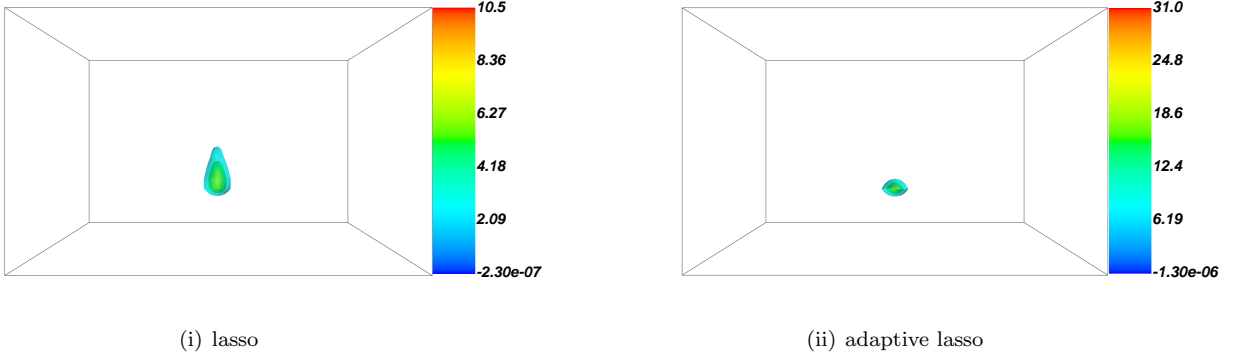


Figure 6. The result of density map reconstruction with standard lasso (left) and adaptive lasso (right). The input density map is from a NFW halo with mass $M_{200} = 10^{15} h^{-1} M_{\odot}$ at redshift 0.35. The lasso reconstruction smear the density field along the line of sight direction (vertical direction in the plots). Noises on shear measurements are neglected in this simulation.

2.3.1. Photo- z Uncertainty

The photometric redshifts of source galaxies in the current large scale survey are estimated with a limited number of board photometric bands (e.g. 9 bands for KIDS+VIKING survey (Hildebrandt et al. 2020), 5 bands for DES survey and HSC survey). As a result, the estimated redshifts of galaxies suffer from much larger uncertainties, compared with redshifts estimated with spectroscopic observations. Such photo- z uncertainty smears the lensing kernels statistically since a galaxy with a best fit photo- z estimation of z_s has possibilities of being actually located at different redshifts (z). The probability function for the photo- z uncertainty is denoted as $P(z|z_s)$ and the expected shear distortion on the galaxy is

$$\int dz_s P(z|z_s) \gamma_L(\vec{\theta}, z_s). \quad (19)$$

With the definition of photo- z smearing operator

$$\mathbf{P} = \int dz_s P(z|z_s), \quad (20)$$

the photo- z uncertainty changes the shear as

$$\gamma_L \rightarrow \mathbf{Q} \gamma_L, \quad (21)$$

where we use \rightarrow to denote the changes of shear by the systematic operator.

Figure 3 shows the histogram of the best MLZ photo- z estimation Tanaka et al. (2018) for galaxies in tract 9347 of HSC S16A data release (Aihara et al. 2018). These Galaxies are divided into ten source galaxy bins according to the photo- z best estimation and the boundaries of the bins are shown as vertical dash lines in Figure 3. Figure 4 shows the average probability density function (PDF) in for galaxies in each redshift bin.

The left panel of Figure 5 shows the lensing kernels for lenses at five different redshifts as functions of source

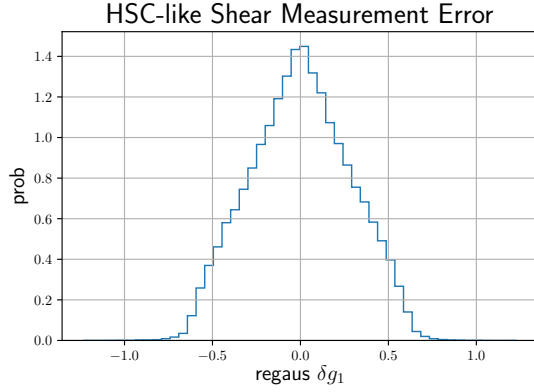


Figure 7. The Histogram of HSC-like shear measurement error (including errors from shape noise and photon noise) on the first component of shear (g_1).

galaxy redshift bins. The dashed lines are the lensing kernel for spectroscopic redshifts with neglectable redshift uncertainty. The solid lines are the lensing kernel for photometric redshifts with redshift uncertainties shown in Figure 4. As shown by the dashed lines in the left panel of Figure 5, the lensing kernels converge to zero for source redshifts lower than the lens redshift, if the uncertainties on the source galaxy redshift estimations are neglectable. However, as demonstrated by the solid lines in the same panel, for the source redshifts with large photo- z uncertainties, the lensing kernels do not converge to zero at redshifts lower than the lens redshift. This is because the galaxies with photo- z estimations lower than the lens redshifts may be actually located at higher redshifts due to the photo- z uncertainties.

The middle and right panels show the correlation between lensing kernels for lenses at different redshifts. The middle panel is for spect- z and the right panel is for photo- z . As demonstrated in the middle panel, the lensing kernels are highly correlated even though the redshift estimation is precise. Comparing the correlation matrices shown in the middle panel and the right panel of Figure 5, we conclude that the photo- z uncertainties further increase the correlations between lensing kernels at different lens plane.

2.3.2. Smoothing

The observed galaxies have random irregular (unequally-spaced) spatial distribution. In order to boost the computational speed, we smooth the shear measurements from galaxy shapes and pixelize the smoothed measurements onto regular grids. After the pixelization, the fast Fourier transform (FFT) can be directly conducted on the transverse plane in each source redshift bin.

The smoothing is conducted by convolving the shear measurements with a smoothing kernel

$$\gamma_{\text{sm}}(\vec{\theta}) = \frac{\sum_i W(\vec{\theta} - \vec{\theta}_i, z - z_i) \gamma_i}{\sum_i W(\vec{\theta} - \vec{\theta}_i, z - z_i)}, \quad (22)$$

where $W(\vec{\theta}, z)$ is a 3-D smoothing kernel. γ_i , z_i and θ_i are the shear, photometric redshift, and transverse position of the ‘ i -th’ galaxy in the catalog.

$W(\vec{\theta}, z)$ can be decomposed into a transverse component $W_T(\vec{\theta})$ and a line of sight component $W_\times(z)$

$$W(\vec{\theta}, z) = W_T(\vec{\theta}) W_\times(z). \quad (23)$$

In this paper, we use an isotropic 2-D Gaussian kernel and a 1-D top-hat kernel to smooth the measurements in the transverse plane and the line of sight direction, the smoothing kernel is

$$W_T(\vec{\theta}) = \frac{1}{2\pi\beta^2} \exp\left(-\frac{|\vec{\theta}|^2}{2\beta^2}\right), \quad (24)$$

$$W_\times(z) = \begin{cases} 1/\Delta z & (|z| < \Delta z/2), \\ 0 & \text{else.} \end{cases}$$

With the approximation that the density of galaxy number ($n(\vec{r})$) varies slowly on the smoothing scale, as $\int d^3r W(\vec{r}) = 1$, the smoothed galaxy number density which is defined as

$$n_{\text{sm}}(\vec{r}) = \sum_i W(\vec{r} - \vec{r}_i, z - z_i), \quad (25)$$

equals the galaxy number density: $n_{\text{sm}}(\vec{r}) = n(\vec{r})$. We note that the galaxy number density experience a steep drop on the boundary of the survey, therefore the smoothed galaxy number density does not equal the galaxy number density close to the boundary of the survey.

The smoothing operator is defined as

$$\mathbf{W} = \int d^3r' W(\vec{r} - \vec{r}'), \quad (26)$$

and the smoothing procedure influence the shear signal through

$$\gamma_L \rightarrow \mathbf{W} \gamma_L. \quad (27)$$

As we will discuss in Section 2.3.4, the smoothed shear field is pixelized into equally spaced grids. We note that another widely used scheme is to directly average the shear measurements located in each pixel. Such scheme is equivalent to an equal-space sampling of the shear field smoothed with a 3-D top-hat kernel with the same scale as the pixels.

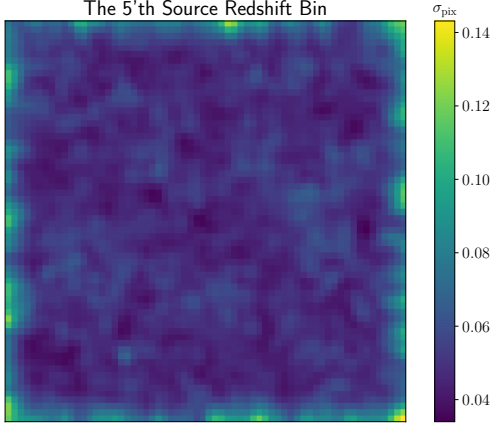


Figure 8. The standard deviation map of shear measurement error for the fifth source bin ($0.69 \leq z < 0.80$).

2.3.3. Masking

In real observations, shear measurements are available in a finite region of the sky and the boundary of the region is always irregular. Moreover, there are also many isolated sub-regions near the bright stars are masked out since the light from bright stars tends to influence the shear measurements on neighboring galaxies.

We define the masking window function according to the smoothed number density (defined in eq. (25)) of the galaxies as follows

$$M(\vec{r}) = \begin{cases} 0 & n_{\text{sm}} > 1, \\ 1 & \text{else.} \end{cases} \quad (28)$$

The mask changes the shear measurements though

$$\gamma_L(\vec{\theta}, z) \rightarrow M(\vec{\theta}, z) \gamma_L(\vec{\theta}, z), \quad (29)$$

We define the masking operator as

$$\mathbf{M} = \int d^3 r' M(\vec{r}') \delta_D(\vec{r} - \vec{r}'), \quad (30)$$

where $\delta_D(\vec{r})$ is 3-D Dirac delta function. The shear is influenced by the masking through $\gamma_L \rightarrow \mathbf{M} \gamma_L$.

The final observed shear field, taking into account all of the aforementioned systematics from observations, is

$$\gamma = \mathbf{MWPQ}\Phi x. \quad (31)$$

For simplicity, we denote $\mathbf{A} = \mathbf{MWPQ}\Phi$ and eq. (31) is written as

$$\gamma = \mathbf{A}x. \quad (32)$$

2.3.4. Pixelization

We pixelize the smoothed shear field into a $N_\theta \times N_\theta \times N_s$ grid, where N_θ is the number of pixels of two orthogonal axes of the transverse plane and N_s is the number

of pixels of the line of sight axis. γ_α is used to denote the value recorded on pixel with index α , where $\alpha = 1 \dots N_\theta \times N_\theta \times N_s$. The grids on the transverse plans are equally spaced, therefore fast Fourier transform (FFT) can be used to boost the speed of linear operation on the transverse plan, whereas the grids in the line of sight direction follow equal number binning as shown in Figure 3.

Similarly, we pixelize each scale frame of the parameter x into a $N_\theta \times N_\theta \times N_l$ grid. The pixelization of the parameter field x on the transverse plan for each scale frame is exactly the same as the pixelization of the smoothed shear field on the transverse plane. Whereas the parameter field is pixelized into equal spaced grids in the line of sight direction. Here, we use N_l to denote the number of the lens planes and x_β to denote the element of the parameter with index β , where $\beta = 1 \dots N_\theta \times N_\theta \times N_l \times N$. The corresponding pixelized elements of the forward transform matrix \mathbf{A} is denoted as $A_{\alpha\beta}$.

We term the column vectors of the forward transform matrix \mathbf{A} as the effective basis atoms. We note that the effective basis atoms have different l^2 norm. The l^2 norm of the i 'th column vectors of the effective basis atoms are $\mathcal{N}_i = \sum_\alpha A_{i\alpha} A_{i\alpha}$. Before solving the density map reconstruction problem, we normalize the column vectors of the transform matrix through the following rescaling

$$\begin{aligned} A'_{\alpha\beta} &= A_{\alpha\beta} / \mathcal{N}_\alpha^{\frac{1}{2}}, \\ x'_\beta &= x_\beta \mathcal{N}_\beta^{\frac{1}{2}}. \end{aligned} \quad (33)$$

2.4. Density map reconstruction

2.4.1. Adaptive lasso

The lasso algorithm uses l^1 norm of parameters as the regularization term and the estimator is defined as

$$\hat{x}'^{\text{lasso}} = \arg \min_x \left\{ \frac{1}{2} \left\| \Sigma^{-\frac{1}{2}} (\gamma - \mathbf{A}' x') \right\|_2^2 + \lambda_{ls} \|x'\|_1 \right\}, \quad (34)$$

where $\|\cdot\|_1$ and $\|\cdot\|_2$ refer to the l^1 norm and l^2 norm, respectively. λ_{ls} denotes the penalization parameter for the lasso estimation.

The lasso algorithm selects the parameters which are relevant to the measurements and simultaneously estimate the value of the selected parameters. However, it has been shown by Zou (2006) that when the column vectors of the projection matrix \mathbf{A}' are highly correlated, lasso cannot select the relevant parameters from the parameter space consistently. Moreover, the estimated parameters are biased due to the shrinkage process in the lasso regression. We note that, for the density

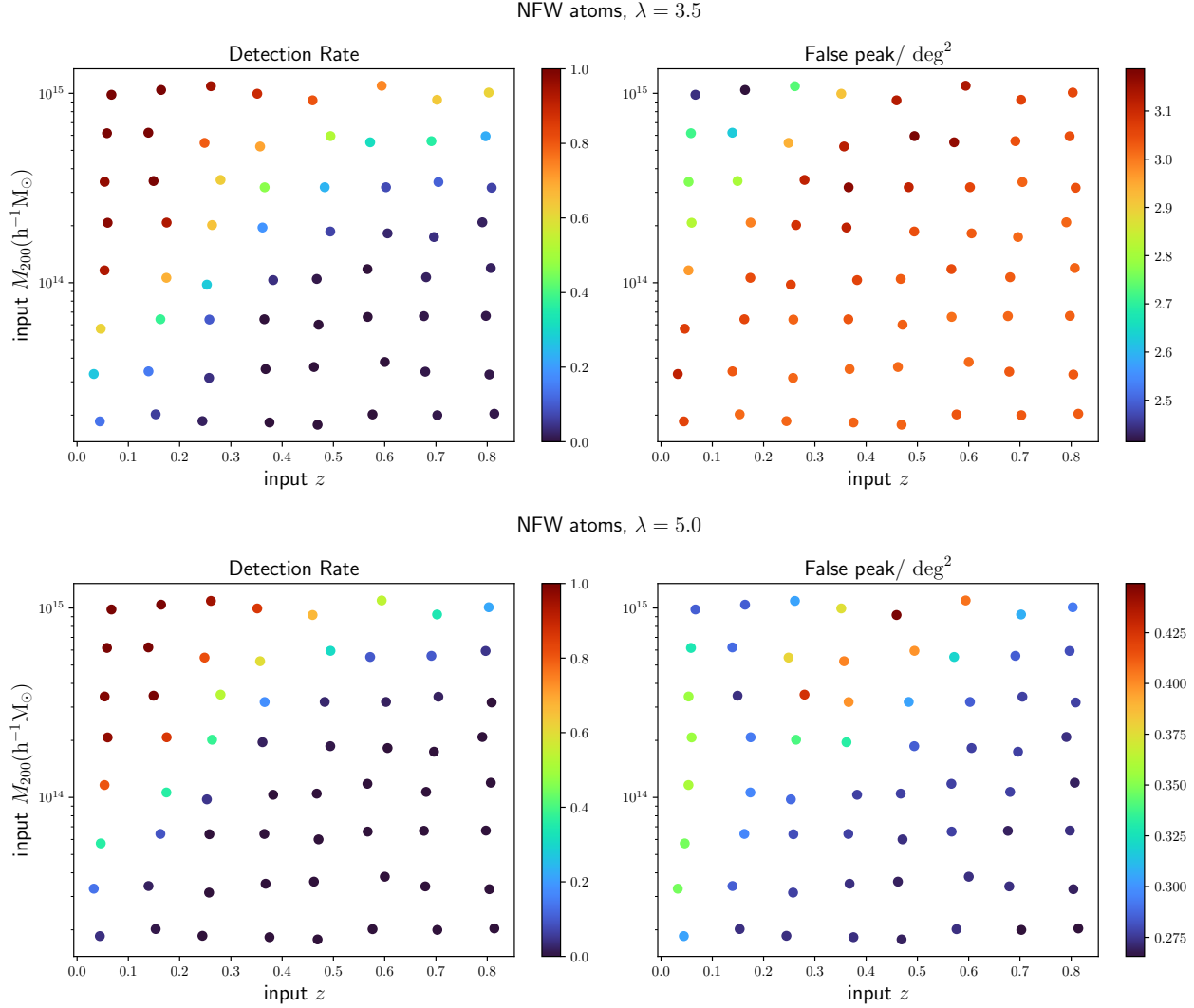


Figure 9. The upper panels show the detection rate and false peak per square degree for mass map reconstruction with NFW atoms and $\lambda = 3.5$. The low panels show the results for $\lambda = 5.0$.

map reconstruction problem, the effective basis vectors are highly correlated since, as shown in Figure 5, the lensing kernels for lenses at different redshifts are highly correlated. Therefore, the lasso algorithm cannot select the consistent mass in the line of sight direction and the reconstructed mass suffers from smears in the line of sight direction even in the absence of noise on shear measurements. Figure 13 shows the reconstructions of the mass map of a single halo with halo mass equals $M_{200} = 10^{15} h^{-1} M_{\odot}$ at redshift 0.35. The noises on shear measurements are not included. The left panel of Figure 13 is the reconstruction with lasso algorithm which shows a significant smear along the line of sight.

Zou (2006) proposes the adaptive lasso algorithm which uses adaptive weights to penalize different parameters in the l^1 penalty. The adaptive lasso algorithm is a two-steps process. In the first step, lasso is used to

estimate the parameters and the preliminary lasso estimation is denoted as \hat{x}'^{lasso} . In the second step, the preliminary lasso estimation is used to weight the penalization. The weight on penalty is defined as

$$\hat{w} = \frac{1}{\left| \hat{x}'^{\text{lasso}} \right|^{\tau}}, \quad (35)$$

here we set the hyper-parameter τ to 2. The adaptive lasso estimator is expressed as

$$\hat{x}' = \arg \min_{x'} \left\{ \frac{1}{2} \left\| \Sigma^{-\frac{1}{2}} (\gamma - \mathbf{A}' x') \right\|_2^2 + \hat{w} \lambda_{\text{als}} \|x'\|_1 \right\}. \quad (36)$$

Here λ_{als} is the penalization parameter for adaptive lasso which does not need to be the same as the penalization parameter for the preliminary lasso estimation (λ_{ls}).

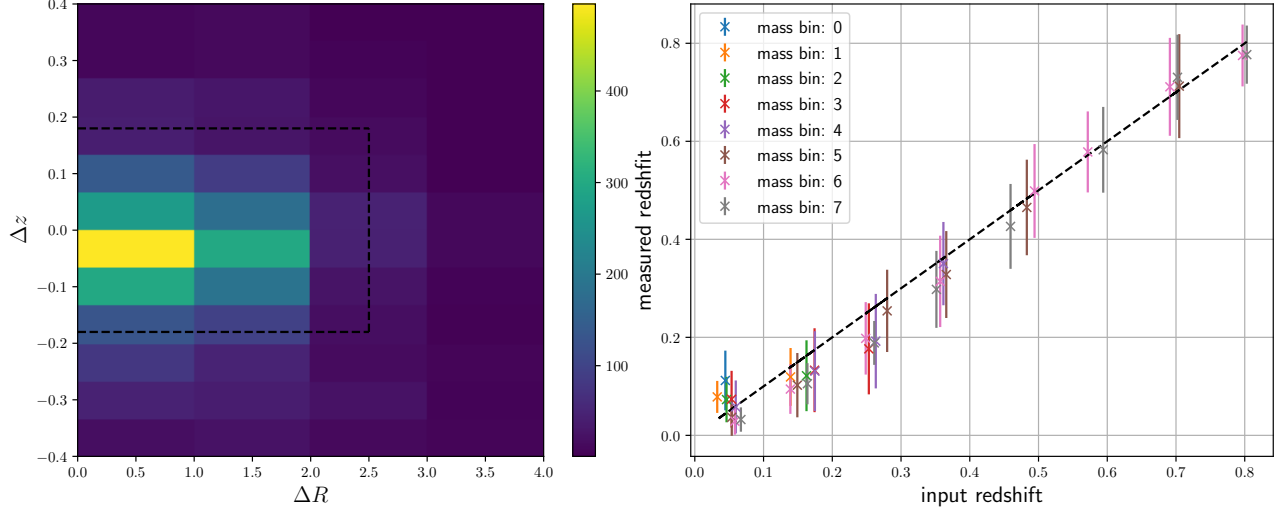


Figure 10. The left panel shows the stacked distribution of deviations of detected peak positions from the centers of the corresponding input halos. The x -axis is for the deviated distance in the transverse plane and the y -axis is for the deviation of the redshift. The peaks inside the dashed black box are regarded as true detections. The right panel focus on the deviation of detected peaks in the line of sight direction. The x -axis shows the input redshifts and the y -axis is the redshift of the detected peak. The cross symbols show the average redshift of detected peak for halos with different input redshift and mass and the error-bars are the scatter of the corresponding peak redshifts. All of the results in this figure are based on the NFW atoms.

The loss function can be rewritten with Einstein notation

$$L(x') = \frac{1}{2}(\Sigma^{-1})_{\alpha\beta}(\gamma_{\alpha}^* - A'_{\alpha i}x'_i)(\gamma_{\beta} - A'_{\beta j}x'_j) + \lambda_{\text{als}}\hat{w}_{\beta}|x'_{\beta}|. \quad (37)$$

To simplify the notation in future, we define the quadruple term in the loss function as $G(x')$:

$$G(x') = \frac{1}{2}\Sigma_{\alpha\beta}^{-1}(\gamma_{\alpha}^* - A'_{\alpha i}x'_i)(\gamma_{\beta} - A'_{\beta j}x'_j). \quad (38)$$

2.4.2. FISTA

In this work, we apply the Fast Iterative Soft Thresholding Algorithm (FISTA) of [Beck & Teboulle \(2009\)](#) to solve the adaptive lasso estimator.

The parameters are initialized as $x_i^{(1)} = 0$. According to the FISTA algorithm, we iteratively update the elements of the parameter vector (x). In the n 'th iteration, a temporary update is first calculated as

$$x_i'^{(n+1)} = \text{ST}_{\lambda} \left(x_i'^{(n)} - \mu \partial_i G(x'^{(n)}) \right), \quad (39)$$

where ST is the soft thresholding function defined as

$$\text{ST}_{\lambda}(x') = \text{sign}(x') \max(|x'| - \lambda, 0). \quad (40)$$

μ is the step size of the gradient descent iteration. $\partial_i G(x'^{(n)})$ refers to the i 'th element of the gradient vector of G at point $x'^{(n)}$

$$\partial_i G(x'^{(n)}) = \Sigma_{\alpha\beta}^{-1} \text{Re} \left(A'_{\alpha i} (\gamma_{\beta} - A'_{\beta j} x'_j) \right), \quad (41)$$

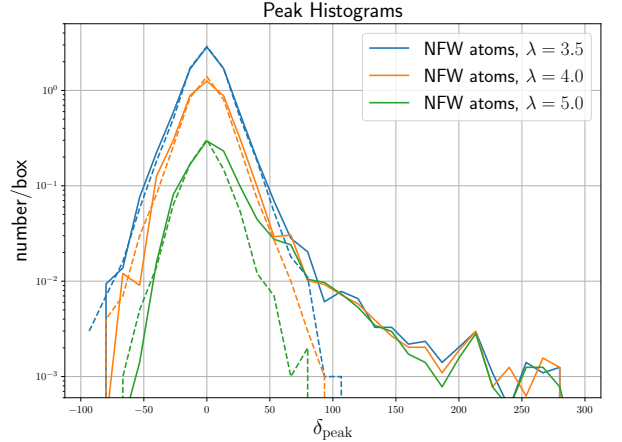


Figure 11. The histograms of detected peak values for all of the halo and error realizations. The dotted lines with different colors are for the results of reconstruction with NFW atoms dictionary penalized with different regularization parameter λ .

where $\text{Re}(\bullet)$ is the function returns the real part of the input function. The FISTA algorithm requires an additional step amounting to a weighted average between $x'^{(n+1)}$ and $x'^{(n)}$:

$$t_{n+1} = \frac{1 + \sqrt{1 + 4(t^{(n)})^2}}{2}, \quad (42)$$

$$x'_{n+1} \leftarrow x'^{(n+1)} + \frac{t^{(n)} - 1}{t^{(n+1)}}(x'^{(n+1)} - x'^{(n)}),$$

where the relative weight is initialized as $t^{(1)} = 1$.

Note that the FISTA algorithm converges as long as the gradient descent step size μ satisfies

$$0 < \mu < \frac{1}{\|\mathbf{A}^\dagger \boldsymbol{\Sigma}^{-1} \mathbf{A}\|}, \quad (43)$$

where $\|\mathbf{A}^\dagger \boldsymbol{\Sigma}^{-1} \mathbf{A}\|$ refers to the spectrum norm of the matrix $\mathbf{A}^\dagger \boldsymbol{\Sigma}^{-1} \mathbf{A}$. The spectral norm is estimated by simulating a large number of random vectors with l^2 norms equal one with different realizations. Then, the matrix operator $\mathbf{A}^\dagger \boldsymbol{\Sigma}^{-1} \mathbf{A}$ is applied to each random vector and get a corresponding transformed random vector. The spectral norm of the matrix $\mathbf{A}^\dagger \boldsymbol{\Sigma}^{-1} \mathbf{A}$ is the maximum l^2 norm of the transformed vectors.

2.4.3. The Algorithm

The algorithm is described in Algorithm 2.4.3.

Algorithm	Our Algorithm
Input:	γ : Pixelized complex 3-D array of shear
Output:	δ : 3-D array of density contrast
1:	Normalize column vectors of \mathbf{A}
2:	Estimate step size μ and $\boldsymbol{\Sigma}$
3:	Initialization:
4:	$x^{(1)} = 0$
5:	$\hat{w} = 1, \lambda = \lambda_{\text{ls}}$
6:	$t^{(1)} = 1, i = 1, j = 1$
7:	while $j \leq 2$ do
8:	while $i \leq N_{\text{iter}}$ do
9:	$x_i^{(n+1)} = \text{ST}_{\hat{w}\lambda} \left(x_i^{(n)} - \mu \partial_i G(x^{(n)}) \right)$
10:	$t^{(n+1)} = \frac{1 + \sqrt{1 + 4(t^{(n)})^2}}{2}$
11:	$x'_{n+1} \leftarrow x^{(n+1)} + \frac{t^{(n)} - 1}{t^{(n+1)}} (x^{(n+1)} - x^{(n)})$
12:	$i = i + 1$
13:	end while
14:	Reinitialization:
15:	$\hat{w} = \left \hat{x}'^{\text{lasso}} \right ^{-2}, \lambda = \lambda_{\text{als}}$
16:	$\hat{x}^{(1)} = x'^{(N_{\text{iter}})}$
17:	$t^{(1)} = 1, i = 1$
18:	$j = j + 1$
19:	end while
20:	$\delta = \Phi \mathcal{N}^{-\frac{1}{2}} x'^{(N_{\text{iter}})}$

3. TESTS

In this section, we simulate weak-lensing shear fields induced by a group of halos with various halo masses and redshifts. The shear distortion fields are applied to HSC mock catalogs with different realizations of HSC-like shape noises and photo- z uncertainties (Section 3.1).

Then, we test our algorithm on the simulations with two different setups of the algorithm. The first setup uses dictionary constructed with NFW atoms (Section 3.2) and the second setup uses dictionary constructed with point mass atoms (Section 3.3).

The Λ CDM cosmology used for the simulations is from the best fitting result of the final full-mission Planck observation of the cosmic microwave background (CMB) with $H_0 = 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.315$, $\Omega_\Lambda = 0.685$ (Planck Collaboration et al. 2020).

3.1. Simulations

We sample halos in two dimensional redshift-mass plane. The redshift-mass plane is evenly divided into eight redshift bins and eight mass bins. The input redshifts and masses are shifted by random values for each halo from the centers of the bins. The concentration of the NFW halo is a function of the halo's mass and redshift according to Ragagnin et al. (2019)

$$c_h = 6.02 \times \left(\frac{M_{200}}{10^{13} M_\odot} \right)^{-0.12} \left(\frac{1.47}{1 + z_h} \right)^{0.16}. \quad (44)$$

The weak-lensing shear fields of these NFW halos are simulated according to Takada & Jain (2003). The shear distortions are applied to galaxy catalogs with HSC-like shape noise and photo- z uncertainty.

The galaxy catalogs are simulated with the HSC S16A shape catalog (Mandelbaum et al. 2018). We use galaxies in a one square degree region at the center of tract 9347 (Aihara et al. 2018). By randomly rotating the galaxies in the shape catalog, we simulate HSC-like shear estimation errors with different realizations. The positions of the galaxies are randomized so that the galaxies are homogeneously distributed in a one square degree stamp. For each galaxy, we randomly assign its redshift following the MLZ photo- z probability distribution function (Tanaka et al. 2018) of the galaxy.

3.2. NFW atoms

In this subsection, we test the performance of our algorithm with the default setup which model the matter density field with multi-scale NFW atoms. The dictionary space is constructed with NFW atom with NFW scale radius set to three different comoving distances which are 0.12 Mpc/h, 0.24 Mpc/h, 0.36 Mpc/h. The truncation radius is set to four times of the scale radius. We test the algorithm with different regularization parameters for the preliminary lasso set to 3.5 and 5.0. The regularization parameter for the adaptive lasso is set to $\lambda_{\text{als}} = \lambda_{\text{ls}}^{\gamma+1}$.

3.3. Point mass atoms

In this subsection, we test the performance of our algorithm with another setup which model the matter density field with point mass atoms. The dictionary space is constructed with point mass atoms in each lens redshift plane. We also test the algorithm with regularization

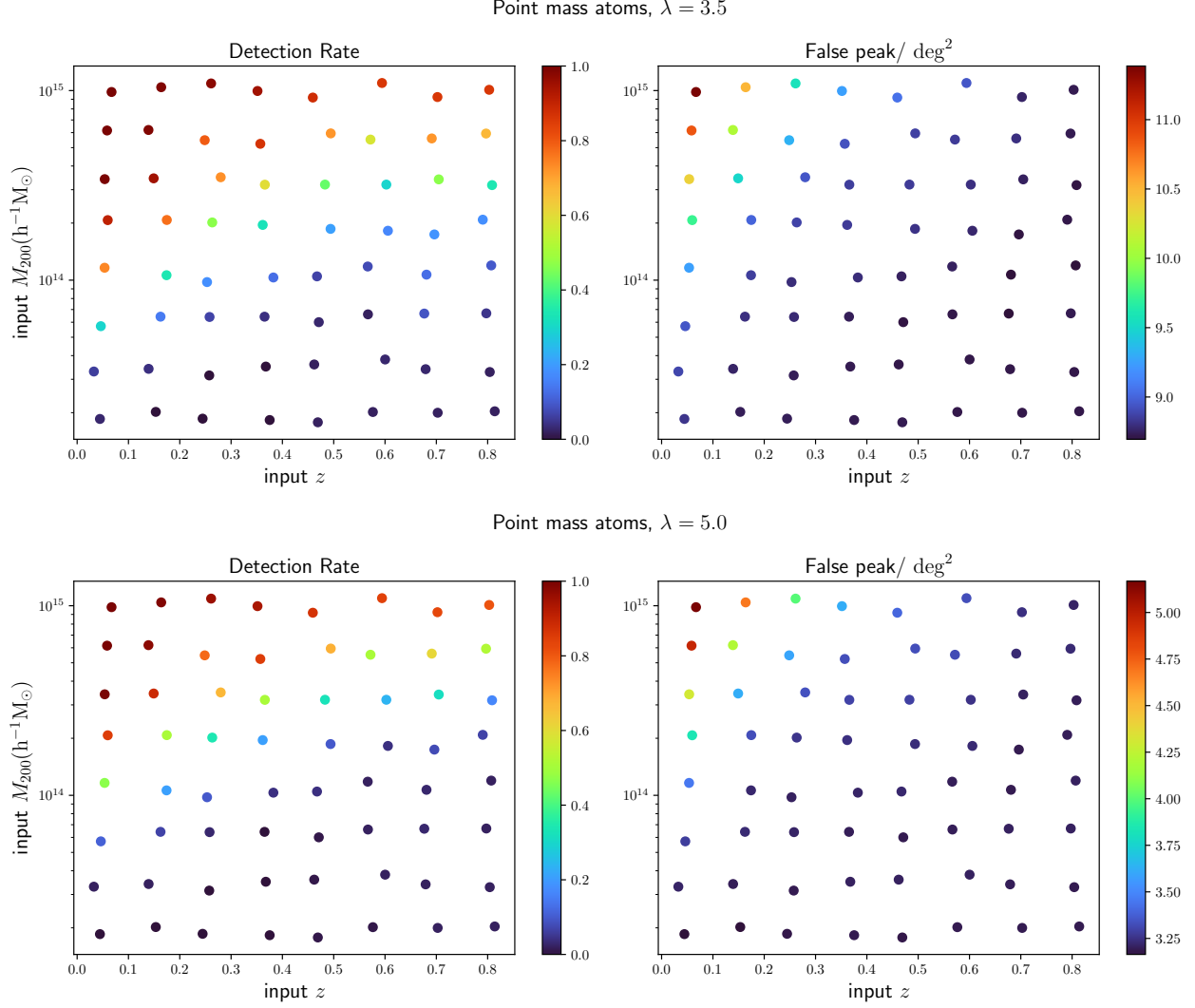


Figure 12. The upper panels show the detection rate and false peak per square degree for mass map reconstruction with point mass atoms and $\lambda = 3.5$. The low panels show the results for $\lambda = 5.0$.

parameter for the preliminary lasso is set to 3.5 and 5.0. The regularization parameter for the adaptive lasso is set to $\lambda_{\text{als}} = \lambda_{\text{ls}}$.

Inspired by [Pramanik & Zhang \(2020\)](#), which propose to incorporate external group information into different penalization weights for the regression coefficients by setting the penalization weights for adaptive lasso for model coefficients in the same group to the same value, we use a smoothed weight for adaptive lasso. The preliminary lasso estimation is smoothed with a top-hat filter of comoving scale 0.25 Mpc/h, which is denoted as $\hat{x}_{\text{sm}}^{\text{ls}}$ and the penalization weights are set to $1/|\hat{x}_{\text{sm}}^{\text{ls}}|^\gamma$.

4. SUMMARY

We develop a novel method to reconstruct 3-D density contrast maps from photometric weak-lensing shear

measurements. Our method models 3-D density contrast maps as a summation of NFW atoms with multiple comoving radius and point mass atoms in the 3-D space binned with photometric redshift. The NFW atoms are used to model the mass in isolated halos and the point mass atoms are used to model the structures close to the resolution limit of the reconstruction.

With the prior assumption that the basis atoms sparsely distributes in the 3-D space, the density field is reconstructed using the adaptive lasso algorithm ([Zou 2006](#)) which has the oracle properties.

The method is tested with realistic simulations using HSC-like shape noise and photo- z uncertainties.

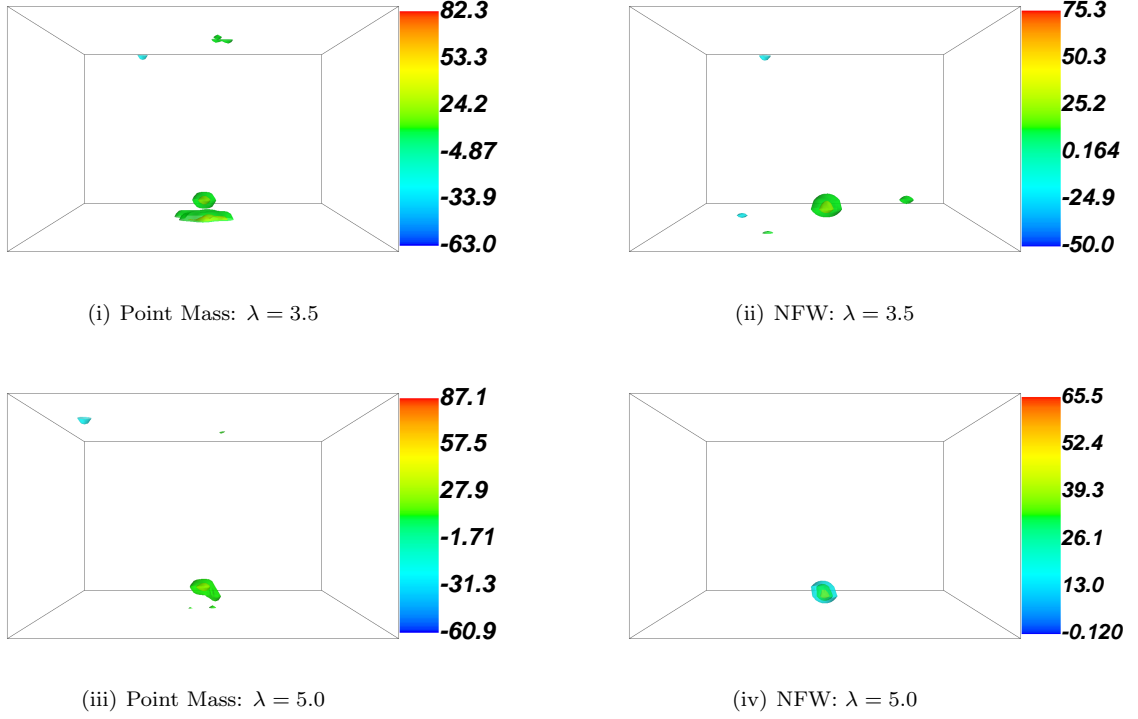


Figure 13. The result of density map reconstruction with Point Mass atoms (left) and NFW atoms (right). The input density map is from a NFW halo with mass $M_{200} = 10^{15.02} h^{-1} M_{\odot}$ at redshift 0.164.

REFERENCES

- Aihara, H., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, S8, doi: [10.1093/pasj/psx081](https://doi.org/10.1093/pasj/psx081)
- Beck, A., & Teboulle, M. 2009, SIAM Journal on Imaging Sciences, 2, 183
- Chang, C., Pujol, A., Mawdsley, B., et al. 2018, MNRAS, 475, 3165, doi: [10.1093/mnras/stx3363](https://doi.org/10.1093/mnras/stx3363)
- de Jong, J. T. A., Kuijken, K., Applegate, D., et al. 2013, The Messenger, 154, 44
- Hamana, T., Shirasaki, M., Miyazaki, S., et al. 2020, PASJ, 72, 16, doi: [10.1093/pasj/psz138](https://doi.org/10.1093/pasj/psz138)
- Han, J., Eke, V. R., Frenk, C. S., et al. 2015, MNRAS, 446, 1356, doi: [10.1093/mnras/stu2178](https://doi.org/10.1093/mnras/stu2178)
- Hikage, C., Oguri, M., Hamana, T., et al. 2019, PASJ, 71, 43, doi: [10.1093/pasj/psx010](https://doi.org/10.1093/pasj/psx010)
- Hildebrandt, H., Köhlinger, F., van den Busch, J. L., et al. 2020, ap, 633, A69, doi: [10.1051/0004-6361/201834878](https://doi.org/10.1051/0004-6361/201834878)
- Jeffrey, N., Abdalla, F. B., Lahav, O., et al. 2018, MNRAS, 479, 2871, doi: [10.1093/mnras/sty1252](https://doi.org/10.1093/mnras/sty1252)
- Kaiser, N., & Squires, G. 1993, pj, 404, 441, doi: [10.1086/172297](https://doi.org/10.1086/172297)
- Kilbinger, M. 2015, Reports on Progress in Physics, 78, 086901, doi: [10.1088/0034-4885/78/8/086901](https://doi.org/10.1088/0034-4885/78/8/086901)
- Lanusse, F., Starck, J. L., Leonard, A., & Pires, S. 2016, ap, 591, A2, doi: [10.1051/0004-6361/201628278](https://doi.org/10.1051/0004-6361/201628278)
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints. <https://arxiv.org/abs/1110.3193>
- Leonard, A., Lanusse, F., & Starck, J.-L. 2014, MNRAS, 440, 1281, doi: [10.1093/mnras/stu273](https://doi.org/10.1093/mnras/stu273)
- LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, ArXiv e-prints. <https://arxiv.org/abs/0912.0201>
- Mandelbaum, R. 2018, ARA&A, 56, 393, doi: [10.1146/annurev-astro-081817-051928](https://doi.org/10.1146/annurev-astro-081817-051928)
- Mandelbaum, R., Miyatake, H., Hamana, T., et al. 2018, PASJ, 70, S25, doi: [10.1093/pasj/psx130](https://doi.org/10.1093/pasj/psx130)
- More, S., Miyatake, H., Mandelbaum, R., et al. 2015, The Astrophysical Journal, 806, 2, doi: [10.1088/0004-637x/806/1/2](https://doi.org/10.1088/0004-637x/806/1/2)
- Morrison, C. B., Klaes, D., van den Busch, J. L., et al. 2016, Monthly Notices of the Royal Astronomical Society, 465, 1454, doi: [10.1093/mnras/stw2805](https://doi.org/10.1093/mnras/stw2805)
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, pj, 490, 493, doi: [10.1086/304888](https://doi.org/10.1086/304888)
- Oguri, M., Miyazaki, S., Hikage, C., et al. 2018, PASJ, 70, S26, doi: [10.1093/pasj/psx070](https://doi.org/10.1093/pasj/psx070)

- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, ap, 641, A6, doi: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910)
- Pramanik, S., & Zhang, X. 2020, arXiv e-prints, arXiv:2006.02041. <https://arxiv.org/abs/2006.02041>
- Prat, J., Sánchez, C., Fang, Y., et al. 2018, Phys. Rev. D, 98, 042005, doi: [10.1103/PhysRevD.98.042005](https://doi.org/10.1103/PhysRevD.98.042005)
- Price, M. A., Cai, X., McEwen, J. D., et al. 2020, MNRAS, 492, 394, doi: [10.1093/mnras/stz3453](https://doi.org/10.1093/mnras/stz3453)
- Ragagnin, A., Dolag, K., Moscardini, L., Biviano, A., & D’Onofrio, M. 2019, MNRAS, 486, 4001, doi: [10.1093/mnras/stz1103](https://doi.org/10.1093/mnras/stz1103)
- Simon, P., Taylor, A. N., & Hartlap, J. 2009, MNRAS, 399, 48, doi: [10.1111/j.1365-2966.2009.15246.x](https://doi.org/10.1111/j.1365-2966.2009.15246.x)
- Spiegel, D., Gehrels, N., Baltay, C., et al. 2015, ArXiv e-prints. <https://arxiv.org/abs/1503.03757>
- Starck, J., Murtagh, F., & Bertero, M. 2015, Starlet transform in astronomical data processing, ed. O. Scherzer, Vol. 1 (United States: Springer New York), 2053–2098
- Takada, M., & Jain, B. 2003, MNRAS, 340, 580, doi: [10.1046/j.1365-8711.2003.06321.x](https://doi.org/10.1046/j.1365-8711.2003.06321.x)
- Tanaka, M., Coupon, J., Hsieh, B.-C., et al. 2018, Publications of the Astronomical Society of Japan, 70, S9, doi: [10.1093/pasj/psx077](https://doi.org/10.1093/pasj/psx077)
- The Dark Energy Survey Collaboration. 2005, ArXiv Astrophysics e-prints
- Troxel, M. A., MacCrann, N., Zuntz, J., et al. 2018, Phys. Rev. D, 98, 043528, doi: [10.1103/PhysRevD.98.043528](https://doi.org/10.1103/PhysRevD.98.043528)
- Zou, H. 2006, Journal of the American Statistical Association, 101, 1418, doi: [10.1198/016214506000000735](https://doi.org/10.1198/016214506000000735)

APPENDIX