Sparsity Weak Lensing 3- Mass Reconstruction

ABSTRACT

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 and the shapes of the background galaxies are coherently sheared due to the lensing effect. 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 the universe (see Kilbinger 2015; Mandelbaum 2018, for recent reviews).

Mathematically, the shear field (γ) measured distant galaxies is a linear mapping of the density contrast (δ)

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

where $\mathbf{T}$ is the transformation operator. The transformation operator ($\mathbf{T}$) includes not only physical lensing effect but it also includes observational systematic effects.

In order to fully reconstruct the 3-D mass density distribution (δ) from the 3-D shear field (γ) observed from distorted galaxy images, a premise is take that the density field can be decomposed into a set of models

$$\delta = \mathbf{\Phi}x, \qquad (2)$$

where $\mathbf{\Phi}$ is the transformation operator from the model space to the mass density and x is the projector. In the previous works, Starlets (wavelets transform) and sinusoidial functions (Fourier transform) have been used to modeled the density field .

Subsequently, an ill-posed inverse problem

$$\min_{\mathrm{x}}\{\|\gamma - \mathbf{T}\mathbf{\Phi}\mathrm{x}\|_2^2 + \mathrm{C(x)}\}, \qquad (3)$$

need to be solved to reconstruct the mass density map, where $C(x)$ is the regulation term which makes the ill-posed inversion solvable. The l^1 sparsity regulation and the l^2 ridge regulation have been applied to the mass density map reversion problem. In their application, the l^1 sparsity regulation ($\|x\|_1^1$) is combined with Starlets modeling (Leonard et al. 2014) and the l^2 ridge regulation ($\|x\|_2^2$) is combined with sinusoidial modeling[1]

---

[1] It is also known as Wiener filter.

(Simon et al. 2009). See Oguri et al. (2018) for a recent application of Simon et al. (2009) in the Hyper Suprime-Cam Survey (Aihara et al. 2018).

We propose to use NFW halos (Navarro et al. 1997) to model the density contrast field and construct the density contrast field with l^1 sparsity regulation.

This paper is organized as follows. Section 2 proposes the new method for 3-D mass map reconstruction. Section 3 introduces the realistic simulations we use to test the new method. Section 4 presents the results of our method on the simulations. Section 5 summarizes and discusses the future development of the method.

2. METHODOLOGY

In this section, we first review how the foreground density contrast induces the weak lensing shear distortion on background galaxies in section 2.1. After that, we describe several systematic effects which exist in real observation mathematically. The systematic effects include photo-z uncertainty (section 2.2), smoothing (section 2.3), mask and noise (section 2.4) and pixelation (section 2.6).

Subsequently, we propose a novel method to reconstruct the 3-D density contrast field from the weak lensing shear distortion field. With the premise that the density field can be decomposed into NFW halos and point masses, we build up a dictionary using NFW models and point mass models in section 2.5. The models in our dictionary are used to fit the shear field and the optimal fit is determined by minimizing a loss function. As shown in section 2.7 the loss function is composed of a normal chi-2 term, a l_1 sparsity constrain and a quadratic total square variance constrain. The minimum of the loss function is achieved with the pathwise coordinate descent algorithm in section 2.8.

2.1. *Density Contrast to Shear*

The lensing convergence κ at the comoving distance χ_s caused by the foreground inhomogeneous density distribution at the comoving distance χ_l ($\chi_l < \chi_s$) along the line-of-sight is

$$\kappa(\vec{\theta}, \chi_s) = \frac{3H_0^2\Omega_M}{2c^2} \int_0^{\chi_s} d\chi_l \frac{\chi_l\chi_{sl}}{\chi_s} \frac{\delta(\vec{\theta}, \chi_l)}{a(\chi_l)}, \qquad (4)$$

(Leonard et al. 2014). where χ refers to the comoving distance, $\delta = \rho(\vec{\theta}, \chi_l)/\bar{\rho} - 1$ is the density contrast at the
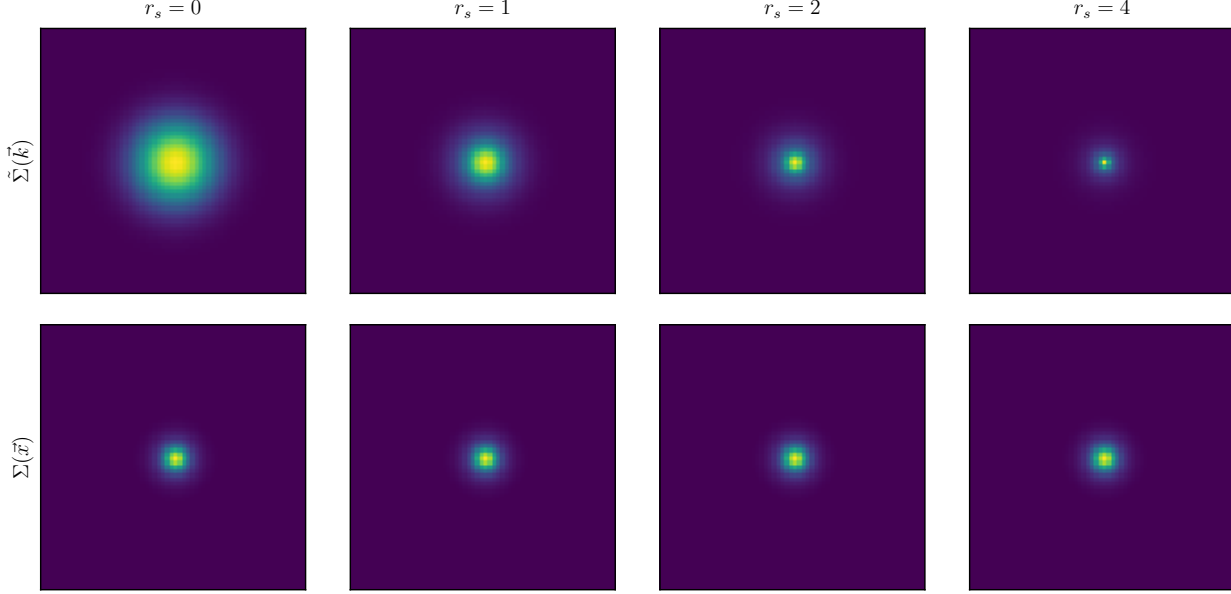
Figure 1. The NFW models with different scale radius (r_s). The first row shows the NFW models in Fourier space and the second row shows the NFW model in Real space.

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) z and we have

$$\kappa(\vec{\theta}, z_s) = \int_0^{z_s} dz_l \delta_c^{-1}(z_l, z_s)\delta(\vec{\theta}, z_l). \qquad (5)$$

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

$$K(z_l, z_s) = \begin{cases} \frac{3H_0\Omega_M}{2c}\frac{\chi_l\chi_{sl}(1+z_l)}{\chi_s E(z_l)} & (z_s > z_l), \\ 0 & (z_s \leq z_l). \end{cases} \qquad (6)$$

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

$$\gamma^t(\vec{\theta}, z_s) = \int d^2\theta' D(\vec{\theta} - \vec{\theta}')\kappa(\vec{\theta}', z_s), \qquad (7)$$

where

$$D(\vec{\theta}) = -\frac{1}{\pi}(\theta_1 - i\theta_2)^{-2}. \qquad (8)$$

Here we denote the true shear distortion as γ^t to distinguish the observed shear signal with systematics which will be introduced in the following subsections.

Combining equation (5) with equation (7), the true shear signal is derived as

$$\gamma^t(\vec{\theta}, z_s) = \int_0^{z_s} \frac{dz_l}{\delta_c(z_l, z_s)} \int d^2\theta' \vec{D}(\vec{\theta} - \vec{\theta}')\delta(\vec{\theta}', z_l). \qquad (9)$$

Here we define the lensing functional as

$$\mathbf{Q} = \int_0^{z_s} \frac{dz_l}{\delta_c(z_l, z_s)} \int d^2\theta' \vec{D}(\vec{\theta} - \vec{\theta}'), \qquad (10)$$
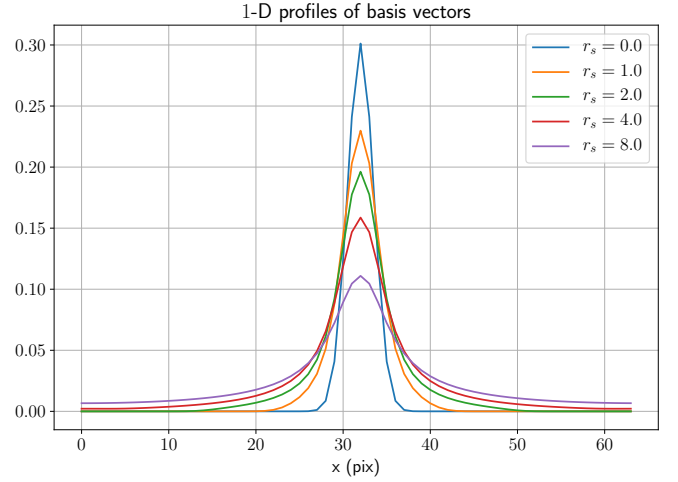

Figure 2. The 1-D slices of NFW models with different scale radius ($r_s = 0, 1, 2, 4, 8$).

then eq. (9) is simplified to $\gamma^t = \mathbf{Q}\delta$.

2.2. *Photo-z Uncertainty*

Since the photometric redshifts of source galaxies in the current large scale survey are estimated with a limited number of photometric bands, the estimated redshift of a single galaxy suffers from large uncertainty. We denote the probability of a galaxies, with photo-z estimated as z_s, being actually located at redshift z as $P(z|z_s)$. Note that, in order to simplify the future cal-

culation, we assume the variation of the PDF across the transverse plane is small and neglect such variation.

Taking the uncertainty of redshift into account, the shear signal changes to

$$\gamma^p(\vec{\theta}, z) = \int dz_s P(z|z_s)\gamma^t(\vec{\theta}, z_s). \qquad (11)$$

With the definition of photo-z functional

$$\mathbf{P} = \int dz_s P(z|z_s), \qquad (12)$$

the relation between shear signal and density contrast is $\gamma^p = \mathbf{PQ}\delta$.

2.3. Smoothing

Since the observed galaxies have random irregular (unequally-spaced) distribution, it is necessary to smooth the shear signal in the observation. The smoothing is expressed as follows

$$\hat{\gamma} = \frac{\sum_i W(\vec{\theta} - \vec{\theta}_i, z - z_i)e_i}{\sum_i R_i W(\vec{\theta} - \vec{\theta}_i, z - z_i)}, \qquad (13)$$

where $W(\vec{\theta}, z)$ is a 3-D smoothing kernel. e_i, R_i, z_i and θ_i are the ellipticity, response, reshift, and transverse position of the 'i-th' galaxy in the galaxy catalog.

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

$$W(\vec{\theta}, z) = W_T(\vec{\theta})W_\times(z). \qquad (14)$$

In this paper, we set

$$W_T(\vec{\theta}) = \frac{1}{2\pi\beta^2}\exp(-\frac{|\vec{\theta}|}{2\beta^2}),$$
$$W_\times(z) = \begin{cases} 1/\Delta z & (|z| < \Delta z/2), \\ 0 & else. \end{cases} \qquad (15)$$

With the approximation that the density of response R and the density of galaxy number vary slowly on the smoothing scale and the fact that $\int d^3 r W(\vec{r}) = 1$, the smoothed shear is

$$\gamma^s(\vec{r}) = \int d^3 r' W(\vec{r} - \vec{r'})\gamma^p(\vec{r'}) \qquad (16)$$

We define the smoothing functional as

$$\mathbf{W} = \int d^3 r' W(\vec{r} - \vec{r'}), \qquad (17)$$

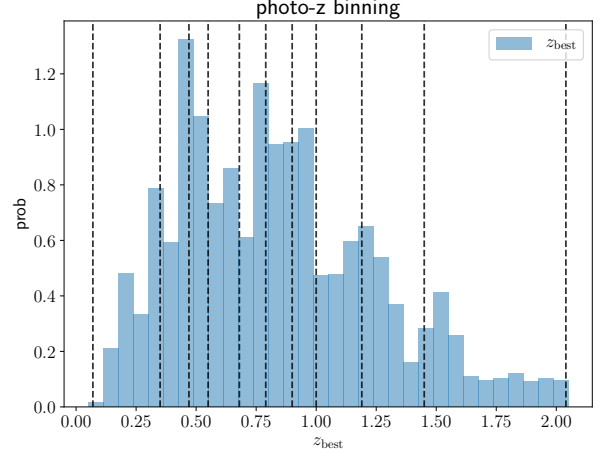and the relation between shear signal and density contrast is $\gamma^s = \mathbf{WPQ}\delta$.

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.

2.4. Mask and Noise

In real observations, the influence of mask and noise should also be taken into account. The final observed shear is

$$\gamma(\vec{\theta}, z) = M(\vec{\theta}, z)\gamma^s(\vec{\theta}, z) + \epsilon(\vec{\theta}, z), \qquad (18)$$

where ϵ represents noise typically caused by random shape of intrinsic galaxies and such noise is modeled as uncorrelated Gaussian noise. $M(\vec{\theta}, z_s)$ is the mask function.

With the definition of masking functional,

$$\mathbf{M} = \int d^3 r' M(\vec{r'})\delta_D(\vec{r} - \vec{r'}), \qquad (19)$$

where $\delta_D(\vec{r})$ is 3-D Dirac delta function, we can write equation (18) into

$$\gamma = \mathbf{MWPQ}\delta + \epsilon. \qquad (20)$$

2.5. Dictionary

N-body simulations show that the dark matter to be largely distributed in halos connected by filaments, therefore, we assume that the density contrast field can be decomposed into multi-scaled NFW halo (Navarro et al. 1997) and point mass at different positions

$$\delta(\vec{r}) = \sum_{s=0}^{N} \int d^3 r' \phi_s(\vec{r} - \vec{r'})x_s(\vec{r'}) \qquad (21)$$

where ϕ_0 is 3-D Dirac delta function to model point masses

$$\phi_0(\vec{r}) = \delta_D(\theta_1)\delta_D(\theta_2)\delta_D(z). \qquad (22)$$

Since the scale of halo is much less than the reachable redshift resolution, we neglect the depth of halo on the line-of-sight direction following (Leonard et al. 2014). On the other hand, in the transverse directions, the surface density profiles of NFW halos with scale θ_α and truncation radius $c\theta_\alpha$ (Takada & Jain 2003) are used to model halos, where c is generally known as concentration of NFW halo.

$$\phi_\alpha(\vec{r}) = \frac{f}{2\pi\theta_\alpha^2} F(|\vec{\theta}|/\theta_\alpha)\delta_D(z), \qquad (23)$$
$$(s = 1..N)$$

where

$$F(x) = \begin{cases} -\frac{\sqrt{c^2-x^2}}{(1-x^2)(1+c)} + \frac{\mathrm{arccosh}\left(\frac{x^2+c}{x(1+c)}\right)}{(1-x^2)^{3/2}} & (x < 1), \\ \frac{\sqrt{c^2-1}}{3(1+c)}\left(1+\frac{1}{c+1}\right) & (x = 1), \\ -\frac{\sqrt{c^2-x^2}}{(1-x^2)(1+c)} + \frac{\arccos\left(\frac{x^2+c}{x(1+c)}\right)}{(x^2-1)^{3/2}} & (1 < x \le c), \\ 0 & (x > c). \end{cases} \qquad (24)$$

$f = 1/[\ln(1+c)-c/(1+c)]$, and x_α is the corresponding projection coefficient of the density contrast field onto our basis NFW model.

We define the projection parameters set as $x = \begin{pmatrix} x_{[0]} \\ x_{[1]} \\ ... \\ x_{[N]} \end{pmatrix}$, and the dictionary transforming functional as

$$\boldsymbol{\Phi} = \left(\int d^3r\phi_0(\vec{r}) \ \int d^3r\phi_1(\vec{r}) \ ... \ \int d^3r\phi_N(\vec{r}) \right). \qquad (25)$$

Then we write equation (20) into

$$\gamma = \mathbf{MPQ\Phi}x + \epsilon, \qquad (26)$$

where $\epsilon(\vec{r})$ is the error from the shear measurement which includes contributions from both random galaxy orientations and photon noise.

For simplicity, we denote $\mathbf{A} = \mathbf{MPQ\Phi}$ and eq. (26) reduces to

$$\gamma = \mathbf{A}x + \epsilon. \qquad (27)$$

2.6. *Pixelization*

We pixelize the smoothed shear field (γ) 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 $\alpha = 1..N_\theta \times N_\theta \times N_s$. The grid on the transverse plan is
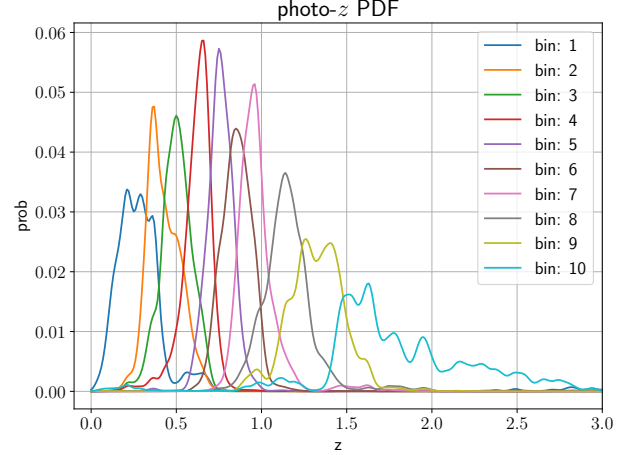

Figure 4. The PDF of photo-z error for 10 source redshift bins.

equally spaced so that we can perform the Fast Fourier Transform (FFT) on the transverse plan.

Similarly, we pixelize each frame [2] of the projector into a $N_\theta \times N_\theta \times N_l$ grid. The pixelization on the transverse plan of each parameter frame is exactly the same as that of the observational frame. N_l is the number of pixels on the line-of-sight axis. x_β is used to denote the value recorded on pixel with index β, where $\beta = 1..N_\theta \times N_\theta \times N_l \times N$.

With Einstein notation, eq. (27) reduces to

$$\gamma_\alpha = A_{\alpha\beta}x_\beta + \epsilon_\alpha, \qquad (28)$$

where ϵ_α refers the noise (ϵ) on the pixel of index α.

2.7. *Loss Function with Constrains*

We define the loss function as

$$L(x) = \frac{1}{2}\|\gamma - \mathbf{A}x\|_2^2 + \frac{1}{2}\tau\mathrm{TSV}(x) + \lambda\sigma\|x\|_1, \qquad (29)$$

where $\|\cdot\|_1$ and $\|\cdot\|_2$ refer to the l^1 norm and l^2 norm, respectively. $\mathrm{TSV}(x)$ refers to the total square variance of the first dictionary frame, which is defined as

$$\mathrm{TSV}(x) = \int d^2\theta dz(\frac{\partial x}{\partial \theta_1})^2 + \int d^2\theta dz(\frac{\partial x}{\partial \theta_2})^2. \qquad (30)$$

The total square variance can be expressed in a quadratic form as

$$\mathrm{TSV}(x) = \left\|\frac{\partial x}{\partial \theta_1}\right\|_2^2 + \left\|\frac{\partial x}{\partial \theta_2}\right\|_2^2. \qquad (31)$$

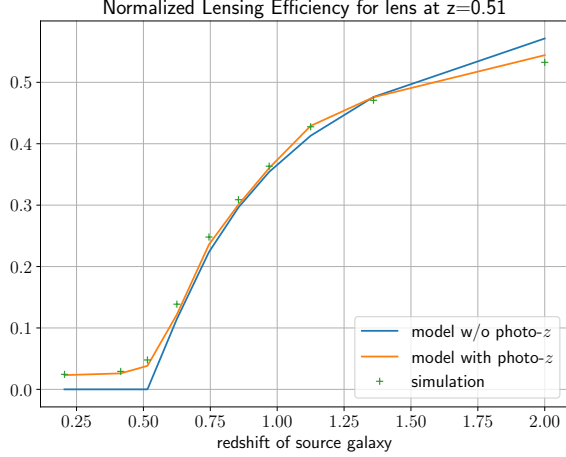[2] Here we have N dictionary frames in total

Figure 5. The blue line shows the normalized lensing efficiency for lens bin at $z_l = 0.51$. The orange line is the normalized lensing efficiency while taking into account the photo-z uncertainty. The green data points show the averaged κ field for different redshift bins in the simulation.

The l^1 term is the sparsity constrain. The TSV term constrains on the smoothness of the mass distribution in the point mass dictionary frame.

The loss function defined on the pixelized data is

$$
\begin{aligned}
L(x) = &\frac{1}{2}(\gamma_\alpha - A^*_{\alpha\beta}x_\beta)(\gamma_\alpha - A_{\alpha\mu}x_\mu) \\
&+ \frac{\tau}{2}[(D^1_{\alpha\beta}x_\beta)(D^1_{\alpha\mu}x_\mu) + (D^2_{\alpha\beta}x_\beta)(D^2_{\alpha\mu}x_\mu)] \\
&+ \lambda\sigma_\beta\|x_\beta\|_1,
\end{aligned}
\tag{32}
$$

where $\mathbf{D^1}$ and $\mathbf{D^2}$ and difference operator on θ_1 and θ_2 directions, respectively.

2.8. *Pathwise Coordinate Descent Algorithm*

The loss function can be separated into a summation of a quadratic term and a l^1 term as

$$
L(x) = G(x) + \lambda\sigma_\beta\|x_\beta\|_1.
\tag{33}
$$

The quadratic term is defined as

$$
\begin{aligned}
G(x) = &\frac{1}{2}(\gamma_\alpha - A^*_{\alpha\beta}x_\beta)(\gamma_\alpha - A_{\alpha\mu}x_\mu) \\
&+ \frac{\tau}{2}[(D^1_{\alpha\beta}x_\beta)(D^1_{\alpha\mu}x_\mu) + (D^2_{\alpha\beta}x_\beta)(D^2_{\alpha\mu}x_\mu)],
\end{aligned}
\tag{34}
$$

Many algorithms have been proposed to find the minima of this kind of loss function. We base our method on the coordinate descent algorithm (Wright 2015) which is described as follows. Firstly, we initialize the projector as $x^{(1)} = 0$. According to the coordinate descent algorithm, we subsequently update the projector (x) along

the direction of one specific coordinate (i) as

$$
x_i^{(n+1)} = \text{ST}_{\lambda\sigma_i}(x_i^{(n)} - \frac{\partial_i G(x^{(n)})}{A_{\alpha i}A_{\alpha i} + 4\tau}),
\tag{35}
$$

where ST is the soft thresholding function defined as

$$
\text{ST}_\lambda(x) = \text{sign}(x)\max(|x| - \lambda, 0),
\tag{36}
$$

and the projector along the other coordinates are kept the same. The minima is finally approached by iteratively updating along each coordinate for multiple times.

To simplify the notation, here we define the projector difference for the n-th iteration as

$$
\Delta x^{(n)} = -\frac{\bigtriangledown G(x^{(n)})}{A_{\alpha i}A_{\alpha i} + 4\tau},
\tag{37}
$$

where $\bigtriangledown G(x^{(n)})$ refers to the gradient of quadratic function G at $x^{(n)}$. The signal to noise ratio (SNR) of the projector difference is defined as $s^{(n)} = \Delta x^{(n)}/\sigma$.

In our algorithm, the coordinate descent algorithm in conducted in a pathwise manner (Friedman et al. 2007). We begin with a regulation parameter ($\lambda^{(1)}$) which is slightly smaller than the maximum coordinate of $s^{(1)}$ (denoted as $s_{\max}^{(1)}$) so that in the first iteration, we only update on the coordinate with the maximum SNR. Then, we find the maximum projector difference SNR of the second iteration ($s_{\max}^{(2)}$) and update the regulation parameter to $\lambda^{(2)}$, where $\lambda^{(2)}$ is slightly smaller than $s_{\max}^{(2)}$.

The algorithm is fully described in Algorithm 5.

Algorithm 1 Our Algorithm

Input: γ: Complex 3-D array of shear
Output: δ: 3-D array of density contrast
1: **Initialization** $x^{(1)} = 0$, $\lambda^{(1)} = 100$
2: **while** $i \leq N_{\text{iter}}$ **do**
3:
4:
5:
6:
7: **end while**

---

3. SIMULATION

This section simulates lensing shear fields induced by a group of dark matter halos with various halo mass and redshifts. HSC-like shape measurement error and photo-z error are added to the shear field.

3.1. *Weak Lensing Fields*

We simulate weak lensing shear fields of NFW halos according to Takada & Jain (2003) and sample halos in the mass-redshift plane as shown in Figure 7. We

assume a dependency of the concentration on the mass and the redshift of a halo

$$c_h = 6.02 \times \left(\frac{M_{200}}{10^{13}M_\odot}\right)^{-0.12}\left(\frac{1.47}{1. + z_h}\right)^{0.16}. \qquad (38)$$
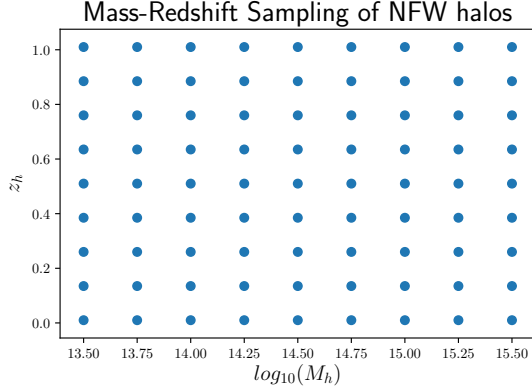
Mass-Redshift Sampling of NFW halos

Figure 6. The sampling points of halos in the mass-redshift plane.

3.2. *HSC-like Errors*

3.2.1. *Shear Measurement Error*

HSC-like Shear Measurement Error
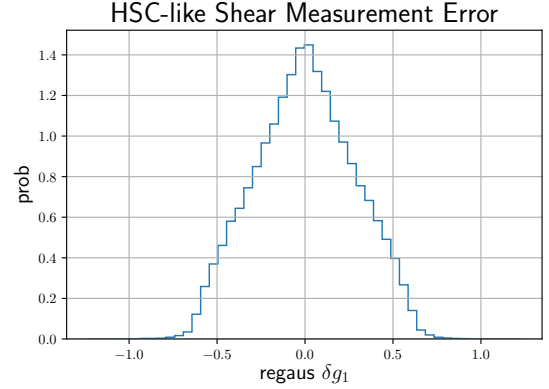
Figure 7. HSC-like measurement error on the first component of shear (g_1).

3.2.2. *Photo-z Error*

4. RESULTS

4.1. *Detection Rate*

4.2. *Mass and Redshift Estimation*

5. SUMMARY

REFERENCES

Aihara, H., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, S8, doi: 10.1093/pasj/psx081

Friedman, J., Hastie, T., Höfling, H., & Tibshirani, R. 2007, The Annals of Applied Statistics, 1, 302, doi: 10.1214/07-aoas131

Kaiser, N., & Squires, G. 1993, ApJ, 404, 441, doi: 10.1086/172297

Kilbinger, M. 2015, Reports on Progress in Physics, 78, 086901, doi: 10.1088/0034-4885/78/8/086901

Leonard, A., Lanusse, F., & Starck, J.-L. 2014, MNRAS, 440, 1281, doi: 10.1093/mnras/stu273

Mandelbaum, R. 2018, ARA&A, 56, 393, doi: 10.1146/annurev-astro-081817-051928

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493, doi: 10.1086/304888

Oguri, M., Miyazaki, S., Hikage, C., et al. 2018, PASJ, 70, S26, doi: 10.1093/pasj/psx070

Simon, P., Taylor, A. N., & Hartlap, J. 2009, MNRAS, 399, 48, doi: 10.1111/j.1365-2966.2009.15246.x

Takada, M., & Jain, B. 2003, MNRAS, 340, 580, doi: 10.1046/j.1365-8711.2003.06321.x

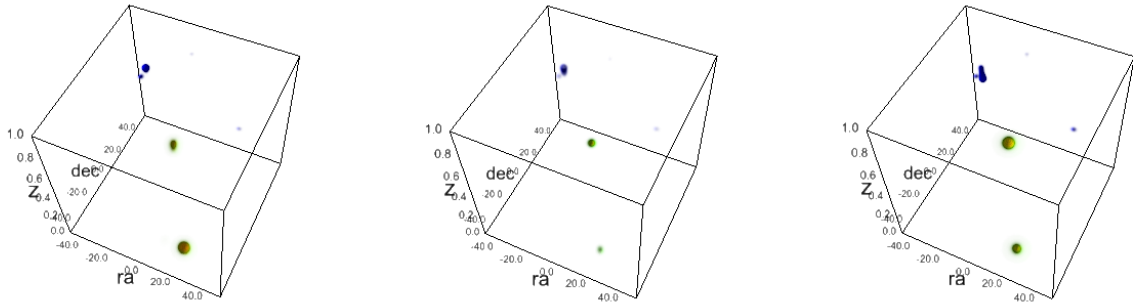Wright, S. J. 2015, Mathematical Programming, 151, 3, doi: 10.1007/s10107-015-0892-3

Figure 8. The density map reconstructed for a halo with mass=$3.16 \times 10^{14} M_\odot/h$, $z = 0.51$. The left panel shows the results with $\tau = 0.10$, the middle panel is the result for $\tau = 0.15$, the right panel is for $\tau = 0.20$.
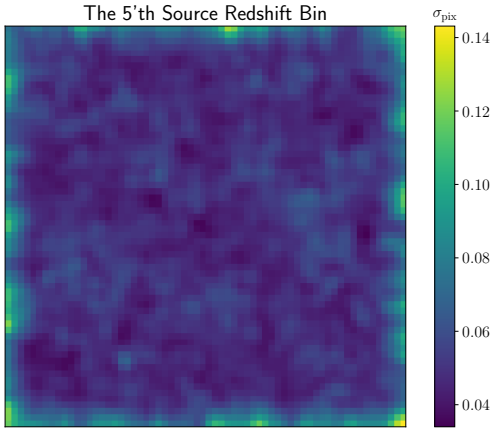

Figure 9. The standard deviation map of shear measurement error for the fifth source bin ($0.69 \leq z < 0.80$).

APPENDIX