Draft version January 8, 2021
Typeset using LaTeX **twocolumn** style in AASTeX63

Cluster Detection from three-dimensional Lensing Mass Map

ABSTRACT

A new method is developed to reconstruct three-dimensional mass maps from photometric weak-lensing shear measurements. The three-dimensional mass map is modeled as a summation of the NFW basis atoms, which have two-dimensional multi-scale NFW surface density profiles on the transverse plane and one-dimensional Dirac delta functions in the line of sight direction. The adaptive lasso algorithm is applied to find a sparse reconstruction of the mass maps. We study the performance of three-dimensional cluster detection from the reconstructed mass map with simulations that applies shear distortions from isolated halo to HSC-like galaxy shapes with realistic photometric redshift uncertainties . Our findings are summarized as follows: 1) The lasso reconstructed suffers from a smear of structure in the line of sight direction even in the absence of shape noise, and in contrast, the adaptive lasso algorithm efficiently removes the line of sight smear. 2) The algorithm is able to detect halo with minimal mass limits of $10^{14.0}M_{\odot}/h$, $10^{14.7}M_{\odot}/h$, $10^{15.0}M_{\odot}/h$ for the low ($z < 0.3$), median ($0.3 \leq z < 0.6$) and high ($0.6 \leq z < 0.85$) redshifts, respectively, with a false detection of $0.022/\mathrm{deg}^2$. 3) The estimated redshift of the halos detected from the reconstructed mass maps are lower than the true redshift by about 0.03 for halos at low redshifts ($z \leq 0.4$). The relative redshift bias is below 0.5% for halos at $0.4 < z \leq 0.85$.

1. INTRODUCTION

Weak lensing refers to the phenomenon that light from distant galaxies is weakly yet coherently distorted by the intervening inhomogeneous density distribution along the line of sight due to gravity's influence. As a result, the shapes of the background galaxies are distorted and the information of the foreground mass distribution is imprinted to the background galaxy images. Weak lensing offers a direct probe into the mass distribution, including both visible matter and invisible dark matter, in our universe (see Kilbinger 2015; Mandelbaum 2018, for recent reviews). Several large-scale surveys focus on studying the weak-lensing effect at a 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), NGRST (Spergel et al. 2015)).

The reconstruction of density map from shear measurements has received considerable interest since we can constrain the cosmology model using the peak counts, which are identified from the weak lensing mass map (Jain & Van Waerbeke 2000; Fan et al. 2010; Lin et al. 2016). Moreover, we can directly detect massive clusters through identifying high signal-to-noise (SNR) peaks on the mass map without any reference to the mass-to-light ratio (Schneider 1996; Hamana et al. 2004).

The two-dimensional (2-D) density map reconstruction which recovers 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). Cluster detection from 2-D mass maps has been widely studied and applied to these large-scale weak-lensing surveys (Shan et al. 2012; Miyazaki et al. 2018a; Hamana et al. 2020); however, cross matching with some (e.g., optically selected) cluster catalog is needed to assign the redshift and extract its physical parameter (such as mass) for each cluster selected from 2-D mass map.

On the other hand, we can directly reconstruct 3-D mass maps taking advantage of photometric redshift (photo-z) information (Simon et al. 2009; VanderPlas et al. 2011); however, these methods either do not have enough spatial resolution to identify clusters, or they suffer from smears along the line of sight direction, which need to be overcome for practical searches of clusters in 3D mass maps. Alternatively, Hennawi & Spergel (2005) propose to perform a maximum-likelihood detection of clusters, by convolving tomographic shear measurements with 3-D filters that match the tangential shears induced by multi-scale NFW halos, without fully reconstruct the 3-D mass maps.

In this paper, we develop a novel method for high-resolution 3-D mass map reconstruction that are free

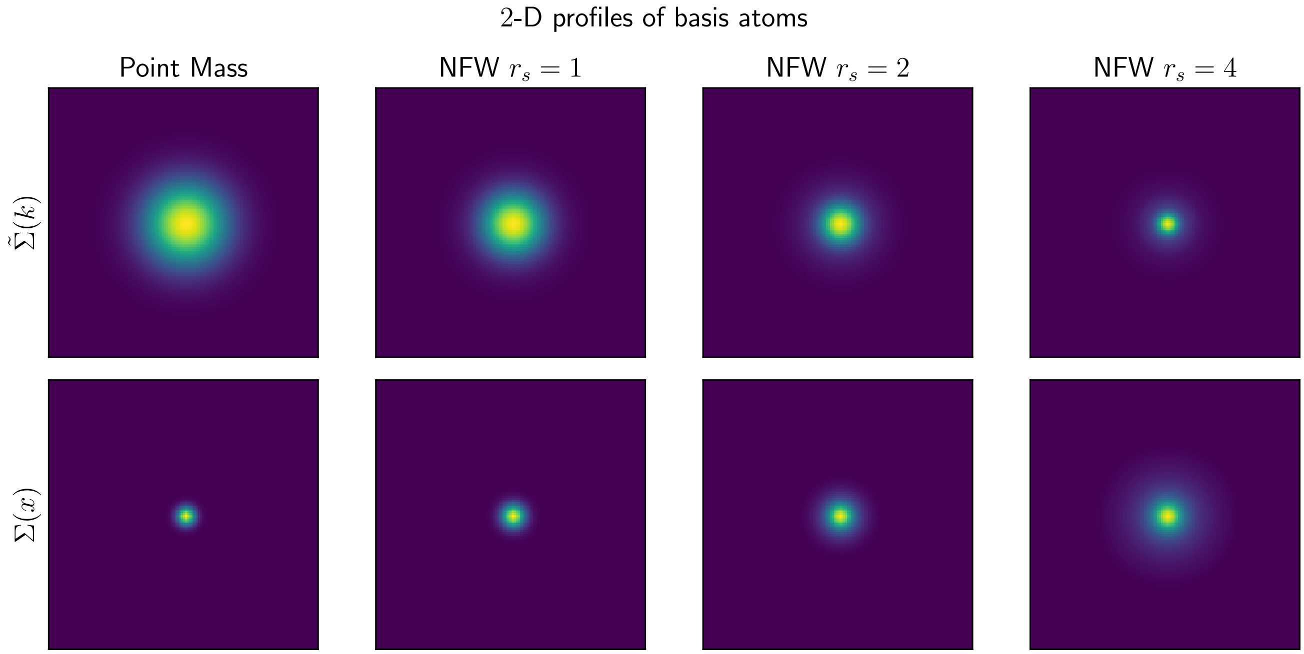


Figure 1. The smoothed pixelized basis atoms. The upper row shows the basis atoms in Fourier space, and the lower row shows the basis atoms in Real space. The leftmost column is the point mass atom, and the other columns are the multi-scale NFW atoms. The smoothing kernel is Gaussian with a standard dev iation of 1.5 pixels.

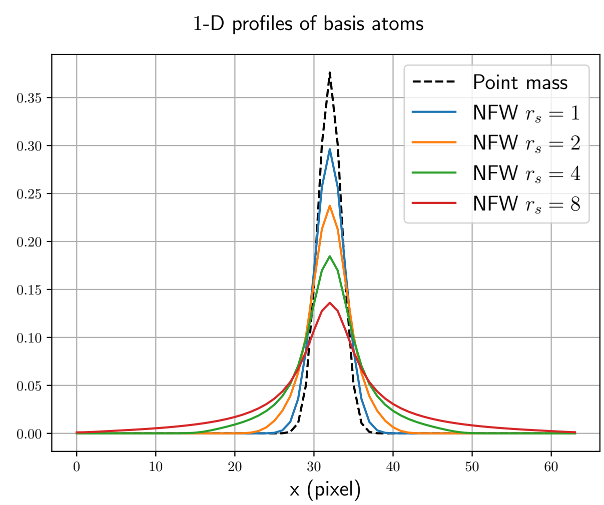


Figure 2. The 1-D slices for smoothed pixelized basis atoms at $x = 0$. The corresponding 2-D profiles are shown in Figure 1.

from line of sight smears. The 3-D mass map is modeled as a summation of the NFW (Navarro et al. 1997) basis atoms, which have 2-D multi-scale NFW surface density profiles on the transverse plane and one-dimensional Dirac delta functions in the line of sight direction. The adaptive lasso algorithm (Zou 2006) is applied to find a sparse reconstruction of the mass maps.

We study the performance of cluster detection from the reconstructed mass maps using simulations applying shear distortions from isolated halos to HSC-like galaxy shapes with realistic photo-z uncertainties.

This paper is organized as follows. In Section 2, we propose the new method for 3-D density map reconstruction. In Section 3, we study the cluster detection from the reconstructed mass map using isolated halo simulations with the HSC observational condition. In Section 4, we summarize and discuss the future development of the method.

2. 3-D MASS MAP RECONSTRUCTION

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

$$\gamma = \mathbf{T}\delta, \tag{1}$$

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

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:

$$\delta = \mathbf{\Phi}x, \tag{2}$$

where $\mathbf{\Phi}$ is the transformation operator from the projection coefficient field to the density contrast, and x

denotes the projection coefficients. Simon et al. (2009) reconstruct the density field in Fourier space, which is equivalent to modeling the mass field with sinusoidal functions. On the other hand, Leonard et al. (2014) model the mass field with Starlets (Starck et al. 2015).

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

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

where $\left\| \Sigma^{-\frac{1}{2}} (\gamma - \mathbf{T\Phi} x) \right\|_2^2$ is the l^2 chi-square term[1] measuring the difference between the prediction and the data, while $C(x)$ is the regularization term measuring the deviation of the coefficient estimation (x) from the prior assumption. The l^p norm is defined as

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

Such penalized estimation prefers the parameters that are able to describe the observations and align with the prior assumption. The regularization parameter λ adjusts the relative weight between the data and prior assumption 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 smearing along the line of sight direction with a standard deviation of $\sigma_z = 0.2 \sim 0.3$.

Leonard et al. (2014) propose the Glimpse algorithm, which uses a derivative version of l^1 lasso regulation ($C = \|x\|_1^1$) to find a sparse solution in the Starlet 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 minimum of a non-convex loss function penalized with the firm thresholding function. The Glimpse algorithm reduces the smearing along the line of sight since the coordinate descent algorithm forces the structure to grow only on the most related lens redshift plane. However, the stability of the non-convex optimization and greedy coordinate descent algorithm has not been fully justified. Moreover, the Starlet dictionary are not designed to model the profile of clumpy mass in the universe, and the Starlet dictionary does not account for the angular scale difference at different lens redshifts.

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

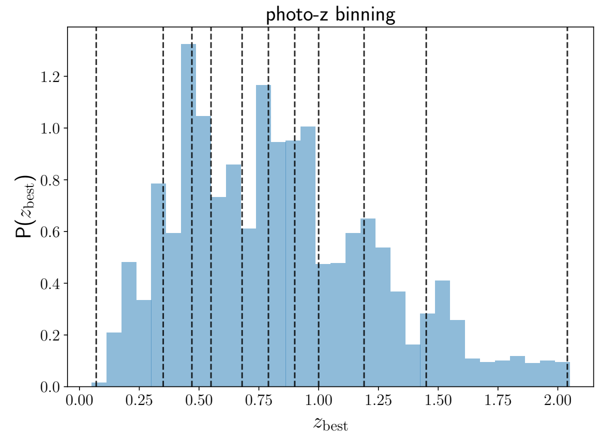


Figure 3. The source galaxies are binned into 10 redshift bins according to their Machine Learning and photo-Z (MLZ) best photo-z estimation. The blue histogram is the normalized number distribution of the best photo-z estimation. The vertical dashed lines are the boundaries of the redshift bins. The galaxies are equal-number binned.

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 model dictionary with the multi-scale NFW atoms. The atoms follow multi-scale surface density profiles of the NFW functions (Takada & Jain 2003) on the transverse plane. Following Leonard et al. (2014), we neglect the depth of halos since the resolution scale of the reconstruction in the line of sight direction is much larger than the halo scale. Therefore, we set the NFW atoms' profile in the line of sight direction as the Dirac delta function. Moreover, we assume that the halos are 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 lasso estimator that smears the structure along the line of sight, the adaptive lasso can efficiently reduce the smearing effect.

Compared with Leonard et al. (2014), our dictionary is built up to describe the clumpy mass in the universe with 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 justified.

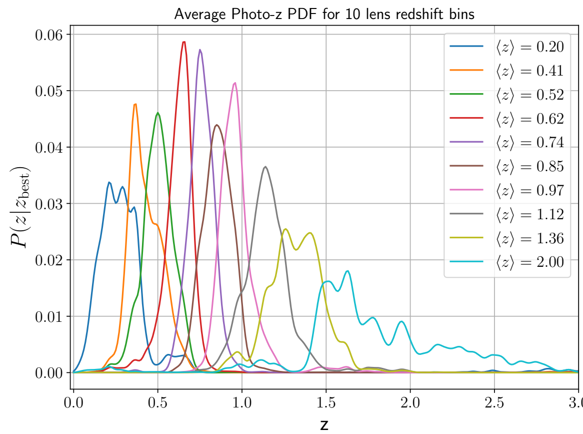


Figure 4. The average PDF of MLZ photo-z error for 10 source redshift bins.

We first review the lensing effect 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 the line of sight is

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

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.

After substituting comoving distance (χ) with redshift (z), we have

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

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

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

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 source redshift plane via

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

where

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

Here we denote the physical shear distortion as γ_L, and we note that it is not the final shear measurement since the final shear measurement is influenced by systematic errors from observations. The systematic errors will be discussed in Section 2.3.

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

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

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

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

and eq. (10) is simplified to

$$\gamma_L = \mathbf{Q}\delta. \tag{12}$$

2.2. *Dictionary*

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

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

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 models at different positions in the comoving coordinates. $x_s(\vec{r'})$ is the projection coefficient of the density contrast field onto the basis atoms at the comoving coordinate: $\vec{r'}$.

We propose to use the multi-scale NFW atoms, denoted as $\{\phi_1, ..., \phi_N\}$, as the basis atoms of our dictionary. On the transverse plane, the NFW atoms follow surface density profiles of the NFW halos (Takada & Jain 2003) with comoving scale radii r_s and concentration c. In this paper, we apply a hard truncation on the NFW profiles at the virial radius: cr_s, the influence of different truncation forms (Oguri & Hamana 2011) on the mass map reconstruction is left to our future work.

As the scales of halos are much smaller than the reachable redshift resolution, we neglect the depth of halo on

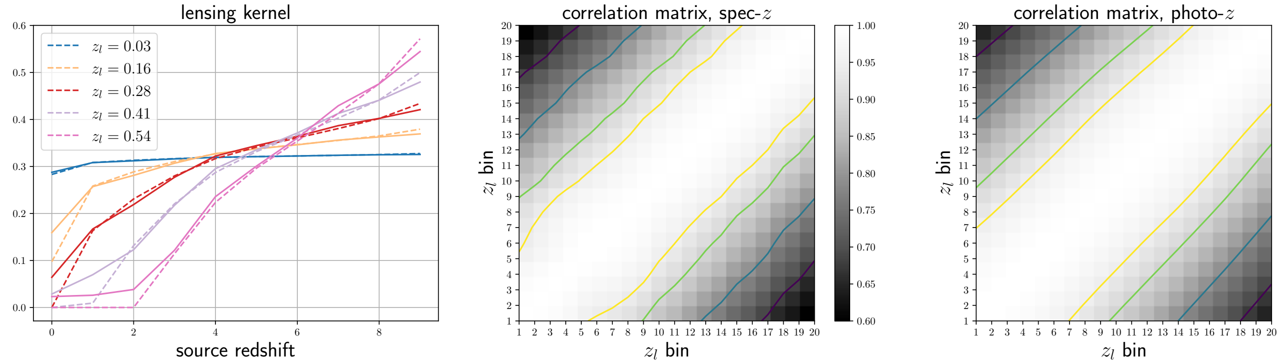


Figure 5. The left panel shows the lensing kernels for five different lens redshifts. The dashed lines are the kernels for spectroscopic redshift, which assumes that the source galaxies' redshifts are precisely estimated. The solid lines are for photometric redshifts, which accounts for the influence 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.

the line of sight direction and set the NFW atoms' profiles in the line of sight direction to one-dimensional (1-D) Dirac delta functions as suggested by (Leonard et al. 2014). The multi-scale NFW atoms are defined as

$$\begin{aligned}\phi_s(\vec{r}_\theta, z) =& \frac{f}{2\pi r_s^2} F(|\vec{r}_\theta|/r_s)\delta_D(z),\\ &(s = 1...N)\end{aligned} \tag{14}$$

where $\vec{r}_\theta$ is the comoving position in the transverse plane,

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

$f = 1/[\ln(1+c) - c/(1+c)]$. In this work, we fix $c = 4$ for the NFW atoms in different scale frames. Then, we write the basis atoms into the angular separation coordinates:

$$\begin{aligned}\phi_s(\vec{\theta}, z) =& \frac{f\chi(z)^2}{2\pi r_s^2} F(|\vec{\theta}|\chi(z)/r_s)\delta_D(z),\\ &(s = 1...N)\end{aligned} \tag{16}$$

where $\chi(z)$ is the comoving distance at the redshift z.

To simplify the notation, we compress the projection coefficients into a column vector: $x = \begin{pmatrix} x_0 \\ x_1 \\ ... \\ x_N \end{pmatrix}$, and compress the dictionary transform operator to a row vector:

$$\mathbf{\Phi} = \left(\int d^3r\phi_0(\vec{r}) \ \int d^3r\phi_1(\vec{r}) \ ... \ \int d^3r\phi_N(\vec{r}) \right). \tag{17}$$

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

$$\gamma_L = \mathbf{Q\Phi}x. \tag{18}$$

In this paper, a 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

$$\phi_{\rm PM}(\vec{\theta}, z) = \delta_D(\theta_1)\delta_D(\theta_2)\delta_D(z). \tag{19}$$

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 measurement is deviated from the physical shear prediction due to the systematic errors from the observation. The influence of systematics is carefully studied and incorporated into the forward modeling in this subsection.

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 broad 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

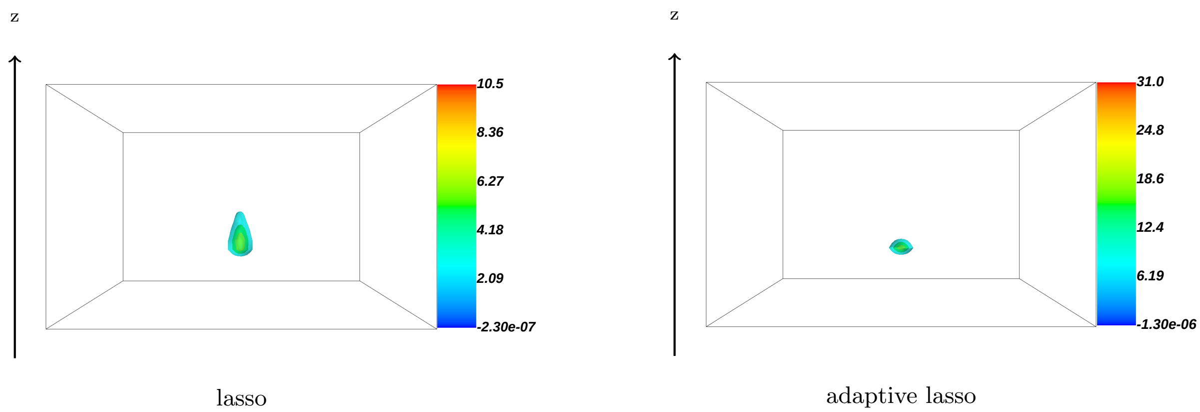


Figure 6. The density map reconstructions with the lasso (left) and the adaptive lasso (right) algorithms. The vertical direction is the line of sight direction, and the lower boundaries and upper boundaries of the boxes are $z = 0.01$ and $z = 0.85$, respectively. Noises on galaxy shape measurements are neglected in the simulation. The input halo's virial mass is $M_{200} = 10^{15}\ h^{-1} M_{\odot}$, and its redshift is $z = 0.35$. The lasso reconstruction smears the density field along the line of sight direction.

estimated redshifts of galaxies suffer from much larger uncertainty than the 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 is denoted as $P(z|z_s)$, and the expected shear distortion on the galaxy is

$$\int dz_s P(z|z_s) \gamma_L(\vec{\theta}, z_s). \tag{20}$$

With the definition of photo-z smearing operator:

$$\mathbf{P} = \int dz_s P(z|z_s), \tag{21}$$

we express the influence of photo-z uncertainty on the shear signal as

$$\gamma_L \rightarrow \mathbf{P}\gamma_L. \tag{22}$$

Figure 3 shows the histogram of the best Machine Learning and photo-Z (Carrasco Kind & Brunner 2013, MLZ) photometric redshift estimation from Tanaka et al. (2018) for galaxies in the tract 9347 of the HSC S16A data release (Aihara et al. 2018). These Galaxies are divided into ten source galaxy bins according to the photo-z best-fit estimation, and the boundaries of the bins are shown as vertical dashed lines in Figure 3. Figure 4 shows the average probability density function (PDF) 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 galaxy redshifts. The results include both the lensing kernels for spectroscopic redshifts with neglectable redshift uncertainties and 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 uncertainty on the source galaxy redshift estimation 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 correlations 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 uncertainty slightly increases the correlations between lensing kernels at different lens plane.

2.3.2. *Smoothing*

The observed galaxies have a random yet unequally-spaced spatial distribution in the transverse plane. To boost the computational speed, we smooth the shear measurements from galaxies and pixelize the smoothed shear field onto a regular grid. After the pixelization, the fast Fourier transform (FFT) can be directly conducted on the transverse plane in each source redshift bin. Another benefit of smoothing is that it reduces the bias from the aliasing effect in the pixelization process as the smoothing kernel reduces the amplitude of the shear signal at the high frequency.

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

$$\gamma_{\rm sm}(\vec{\theta}) = \frac{\sum_i W(\vec{\theta} - \vec{\theta}_i, z - z_i)\gamma_i}{\sum_i W(\vec{\theta} - \vec{\theta}_i, z - z_i)}, \tag{23}$$

where $W(\vec{\theta}, z)$ is a 3-D smoothing kernel. γ_i, z_i and θ_i are the shear, photometric reshift, 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 $W_\times(z)$ as

$$W(\vec{\theta}, z) = W_T(\vec{\theta}) W_\times(z). \tag{24}$$

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. These components of the smoothing kernel are

$$\begin{aligned} W_T(\vec{\theta}) &= \frac{1}{2\pi\beta^2} \exp\left(-\frac{|\vec{\theta}|}{2\beta^2}\right), \\ W_\times(z) &= \begin{cases} 1/\Delta z & (|z| < \Delta z/2), \\ 0 & else. \end{cases} \end{aligned} \tag{25}$$

In this paper, we set $\beta = 1.'5$.

By definition, the smoothing kernel is normalized as

$$\int d^3 r W(\vec{r}) = 1. \tag{26}$$

With the approximation that the density of galaxy number - $n(\vec{r})$ - varies slowly at the smoothing scale, the smoothed galaxy number density (in unit of number per pixel), which is defined as

$$n_{\rm sm}(\vec{r}) = \sum_i W(\vec{\theta} - \vec{\theta}_i, z - z_i), \tag{27}$$

equals the galaxy number density: $n_{\rm sm}(\vec{r}) = n(\vec{r})$. However, 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

$$\mathbf{W} = \int d^3 r' W(\vec{r} - \vec{r'}), \tag{28}$$

and the smoothing procedure influence the shear signal by

$$\gamma_L \rightarrow \mathbf{W} \gamma_L. \tag{29}$$

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 average the shear measurements in each pixel. Such a scenario is equivalent to resampling the shear field smoothed with a 3-D top-hat kernel with the same scale as the pixels.

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, 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 galaxy number density (defined in eq. (27)):

$$M(\vec{r}) = \begin{cases} 0 & n_{\rm sm} > 1, \\ 1 & \text{else.} \end{cases} \tag{30}$$

The mask changes the shear measurements by

$$\gamma_L(\vec{\theta}, z) \rightarrow M(\vec{\theta}, z) \gamma_L(\vec{\theta}, z), \tag{31}$$

We define the masking operator as

$$\mathbf{M} = \int d^3 r' M(\vec{r'}) \delta_D(\vec{r} - \vec{r'}), \tag{32}$$

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

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

$$\gamma = \mathbf{MWPQ\Phi} x. \tag{33}$$

For simplicity, we denote $\mathbf{A} = \mathbf{MWPQ\Phi}$ and eq. (33) is written as

$$\gamma = \mathbf{A} x. \tag{34}$$

2.3.4. *Pixelization*

We pixelize the smoothed shear field into an $N_\theta \times N_\theta \times N_s$ grid, where N_θ is the number of pixels for the two orthogonal axes of the transverse plane and N_s is the number of pixels for the line of sight axis. γ_α denotes the smoothed shear measurements recorded on the pixel with index α, where $\alpha = 1 ... N_\theta \times N_\theta \times N_s$. The grids on the transverse planes are equally spaced with a pixel size of $1'$. While 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 projection coefficient field x into an $N_\theta \times N_\theta \times N_l$ grid. The pixelization on the transverse plane for each scale frame is the same as that of the smoothed shear field on the transverse plane. At the same time, the projection coefficient field is pixelized into equal spaced grids in the line of sight direction ranging from redshift 0.01 to redshift 0.85. Here, we use N_l to denote the number of the lens planes and x_β to denote the projection parameter indexed as β, where $\beta = 1 ... N_\theta \times N_\theta \times N_l \times N$. 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. The l^2 norm of the i-th column vectors weighted by the inverse of the noise covariance matrix's diagonal elements is defined as

$$\mathcal{N}_i = \sum_{\alpha} A_{i\alpha} A_{i\alpha} / \Sigma_{\alpha\alpha}. \tag{35}$$

We note that the effective basis atoms have different weighted l^2 norm. Before solving the density map reconstruction problem, we normalize the column vectors of the transform matrix through a rescaling:

$$\begin{aligned} A'_{\alpha\beta} &= A_{\alpha\beta} / \mathcal{N}_{\alpha}^{\frac{1}{2}}, \\ x'_{\beta} &= x_{\beta} \mathcal{N}_{\beta}^{\frac{1}{2}}, \end{aligned} \tag{36}$$

to make the projection coefficients have the same weighted l^2 norm. Such normalization boosts the speed of the gradient descent iterations in the next subsection.

2.4. *Density map reconstruction*

2.4.1. *Adaptive lasso*

The lasso algorithm uses l^1 norm of the projection coefficient field to regularize the modeling, and the estimator is defined as

$$\hat{x'}^{\text{lasso}} = \arg\min_{x} \left\{ \frac{1}{2} \left\| \Sigma^{-\frac{1}{2}} (\gamma - \mathbf{A}' x') \right\|_2^2 + \lambda \left\| x' \right\|_1^1 \right\}, \tag{37}$$

where $\|\bullet\|_1$ and $\|\bullet\|_2$ refer to the l^1 norm and l^2 norm, respectively. λ denotes the penalization parameter for the lasso estimation.

The lasso algorithm selects the parameters relevant to the measurements and simultaneously estimates the value of the selected parameters. However, it has been shown by Zou (2006) that when the column vectors of the transforming matrix $\mathbf{A}'$ are highly correlated, the lasso cannot select the relevant parameters from the parameter space consistently. Moreover, the estimated parameters are biased due to the shrinkage in the lasso regression. We note that, for the density map reconstruction problem, the column vectors are highly correlated even in the absence of photometric redshift uncertainties because, 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 smearing in the line of sight direction even in the absence of noise on shear measurements and photo-z uncertainty .

Figure 6 shows the reconstructions of a single halo's mass map with virial mass equals $M_{200} = 10^{15}\ h^{-1} M_{\odot}$ at redshift 0.35. The shear measurement error and photo-z uncertainty are not included in the simulation. The left panel of Figure 6 is the reconstruction with the 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 projection coefficients in the l^1 penalty. The adaptive lasso algorithm is a two-steps process. In the first step, the lasso is used to estimate the parameters, and the preliminary estimation of the lasso is denoted as $\hat{x'}^{\text{lasso}}$. In the second step, the preliminary lasso estimation is used to weight the penalization. The weight on penalty is defined as

$$\hat{w} = \frac{1}{\left| \hat{x'}^{\text{lasso}} \right|^{\tau}}, \tag{38}$$

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

$$\hat{x'} = \arg\min_{x'} \left\{ \frac{1}{2} \left\| \Sigma^{-\frac{1}{2}} (\gamma - \mathbf{A}' x') \right\|_2^2 + \hat{w} \lambda_{\text{ada}} \left\| x' \right\|_1^1 \right\}. \tag{39}$$

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

We rewrite the loss function with the Einstein notation:

$$\begin{aligned} L(x') = &\frac{1}{2} (\Sigma^{-1})_{\alpha\beta} (\gamma^*_{\alpha} - A'^*_{\alpha i} x'_i)(\gamma_{\beta} - A'_{\beta j} x'_j) \\ &+ \lambda_{\text{ada}} \hat{w_{\beta}} \left| x'_{\beta} \right|. \end{aligned} \tag{40}$$

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

$$G(x') = \frac{1}{2} \Sigma^{-1}_{\alpha\beta} (\gamma^*_{\alpha} - A'^*_{\alpha i} x'_i)(\gamma_{\beta} - A'_{\beta j} x'_j). \tag{41}$$

2.4.2. *FISTA*

Beck & Teboulle (2009) propose the Fast Iterative Soft Thresholding Algorithm (FISTA) to solve the lasso problem. Since the lasso's loss function and the adaptive lasso's loss function only differ in their penalization terms, the FISTA is also applicable to solve the adaptive lasso problem. In this paper, we apply the FISTA to solve both the preliminary lasso estimation and the final adaptive lasso estimation.

Here we start from the lasso preliminary estimation. The coefficients are initialized as $x_i^{(1)} = 0$. According to the FISTA algorithm, we iteratively update the projection coefficient field (x). Taking the n'th iteration as an example, a temporary update is first calculated as

$$x_i'^{(n+1)} = \text{ST}_{\lambda} \left(x_i'^{(n)} - \mu \partial_i G(x'^{(n)}) \right), \tag{42}$$

where ST is the soft thresholding function defined as

$$\mathrm{ST}_{\lambda}\left(x'\right) = \mathrm{sign}(x') \max\left(\left|x'\right| - \lambda, 0\right). \tag{43}$$

The soft thresholding is a part of the lasso algorithm. It selects the modes with an amplitude greater than λ, and shrink the selected estimations by λ. 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 $x'^{(n)}$:

$$\partial_i G(x'^{(n)}) = \Sigma^{-1}_{\alpha\beta} \,\mathrm{Re}\left(A'^{*}_{\alpha i}(\gamma_\beta - A'_{\beta j} x'_j)\right), \tag{44}$$

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

$$\begin{aligned} t^{(n+1)} &= \frac{1 + \sqrt{1 + 4(t^{(n)})^2}}{2}, \\ x'^{(n+1)} &\leftarrow x'^{(x+1)} + \frac{t^{(n)} - 1}{t^{(n+1)}}(x'^{(n+1)} - x'^{(n)}), \end{aligned} \tag{45}$$

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

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

$$0 < \mu < \frac{1}{\left\|\mathbf{A}^\dagger \mathbf{\Sigma}^{-1} \mathbf{A}\right\|}, \tag{46}$$

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

As summarized in Algorithm 2.4.2, we first initial the projection coefficients as zero and use the FISTA algorithm to find the global minimum of the lasso loss function. Such a global minimum is the preliminary estimation. We then use the preliminary lasso estimation to weight the coefficients and construct the adaptive lasso loss function. Finally, we set the preliminary lasso estimation as the warm start of the adaptive lasso estimation and use the FISTA algorithm again to find the adaptive lasso loss function's global minimum, which is the final solution.

Algorithm Our Algorithm

Input: γ: Pixelized complex 3-D array of shear
Output: δ: 3-D array of density contrast

Normalize column vectors of $\mathbf{A}$
Estimate step size μ and $\mathbf{\Sigma}$
Initialization:
$x'^{(1)} = 0$
$\hat{w} = 1$
$t^{(1)} = 1, i = 1,\ j = 1$
while $j \leq 2$ **do**
  while $i \leq N_{\mathrm{iter}}$ **do**
    # soft thresholding
    $x_i'^{(n+1)} = \mathrm{ST}_{\hat{w}\lambda}\left(x_i'^{(n)} - \mu \partial_i G(x'^{(n)})\right)$
    # FISTA algorithm
    $t^{(n+1)} = \frac{1+\sqrt{1+4(t^{(n)})^2}}{2}$
    $x'^{(n+1)} \leftarrow x'^{(x+1)} + \frac{t^{(n)}-1}{t^{(n+1)}}(x'^{(n+1)} - x'^{(n)})$
    $i = i + 1$
  end while
  Reinitialization:
  $\hat{w} = \left|\hat{x'}^{\mathrm{lasso}}\right|^{-2}$, $\lambda \leftarrow \lambda_{\mathrm{ada}}$
  $\hat{x'}^{(1)} = x'^{(N_{\mathrm{iter}})}$
  $t^{(1)} = 1,\ i = 1$
  $j = j + 1$
end while
$\delta = \mathbf{\Phi} \mathcal{N}^{-\frac{1}{2}} x'^{(N_{\mathrm{iter}})}$

3. CLUSTER DETECTION

This Section simulates weak-lensing shear fields induced by a group of NFW halos with various halo masses and redshifts. The shear fields are used to distort the HSC mock shape catalogs with different realizations of the HSC-like shape measurement error and photo-z uncertainty (Section 3.1).

Then, we test our algorithm using the simulations with different setups of the regularization parameter. We also compare the results of our algorithm, which uses the NFW dictionary (Section 3.2), with the point mass dictionary (Section 3.3).

3.1. *Simulations*

The ΛCDM cosmology used in this paper is from the best-fit result of the final full-mission Planck observation of the cosmic microwave background (CMB) with $H_0 = 67.4$ km s^{-1}Mpc^{-1} $\Omega_M = 0.315$, $\Omega_\Lambda = 0.685$, $\sigma_8 = 0.811$, $n_s = 0.965$ (Planck Collaboration et al. 2020).

We sample halos in a two-dimensional redshift-mass plane. The redshift-mass plane is evenly divided into eight redshift bins and eight mass bins. We randomly shift the input halo redshifts and halo masses from the bins' centers by a small amount. In the simulation, we set the non-linear over-density (Δ_{vir}) to 200, and M_{200} refers to the virial mass. The concentration of the NFW

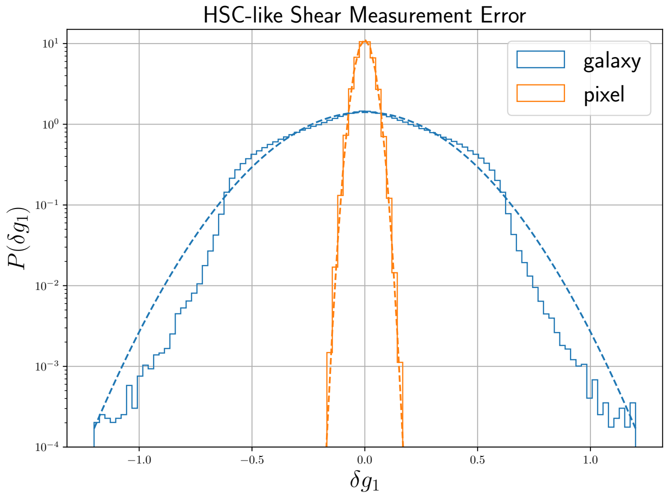


Figure 7. The solid lines show the histograms of the HSC-like shape measurement error (including both from shape noise and photon noise) on the first component of shear (g_1) for galaxies (blue lines) and smoothed pixels (orange lines). The dashed lines are the best-fit Gaussian distributions to the corresponding histograms.

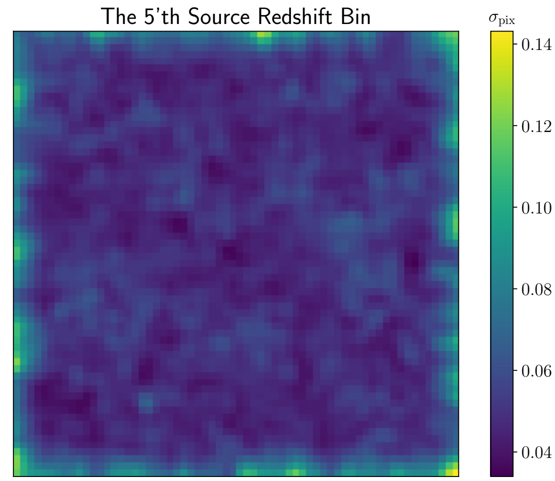


Figure 8. The standard deviation pixel map of the HSC-like shape measurement error for the fifth source galaxy bin ($0.69 \leq z < 0.80$).

halo is set to as a function of the halo's virial mass (M_{200}) and redshift (z_h) according to Ragagnin et al. (2019)

$$c_h = 6.02 \times \left(\frac{M_{200}}{10^{13} M_\odot}\right)^{-0.12} \left(\frac{1.47}{1. + z_h}\right)^{0.16}. \quad (47)$$

The halos are truncated at the virial radius. The weak-lensing shear fields of these NFW halos are simulated according to Takada & Jain (2003). The shear distortions are applied to one hundred realizations of galaxy catalogs with the HSC-like shape measurement error and photo-z uncertainty.

The mock galaxy catalogs are generated using the HSC S16A shape catalog (Mandelbaum et al. 2018). We use the galaxies in a one square degree region at the center of tract 9347 (Aihara et al. 2018). The average galaxy number density in this region is 22.94 arcmin^{-2}. The galaxies' positions are randomized to distribute homogeneously in the one-square degree stamp statistically. We randomly assign its redshift for each galaxy following the MLZ photo-z probability distribution function (Tanaka et al. 2018).

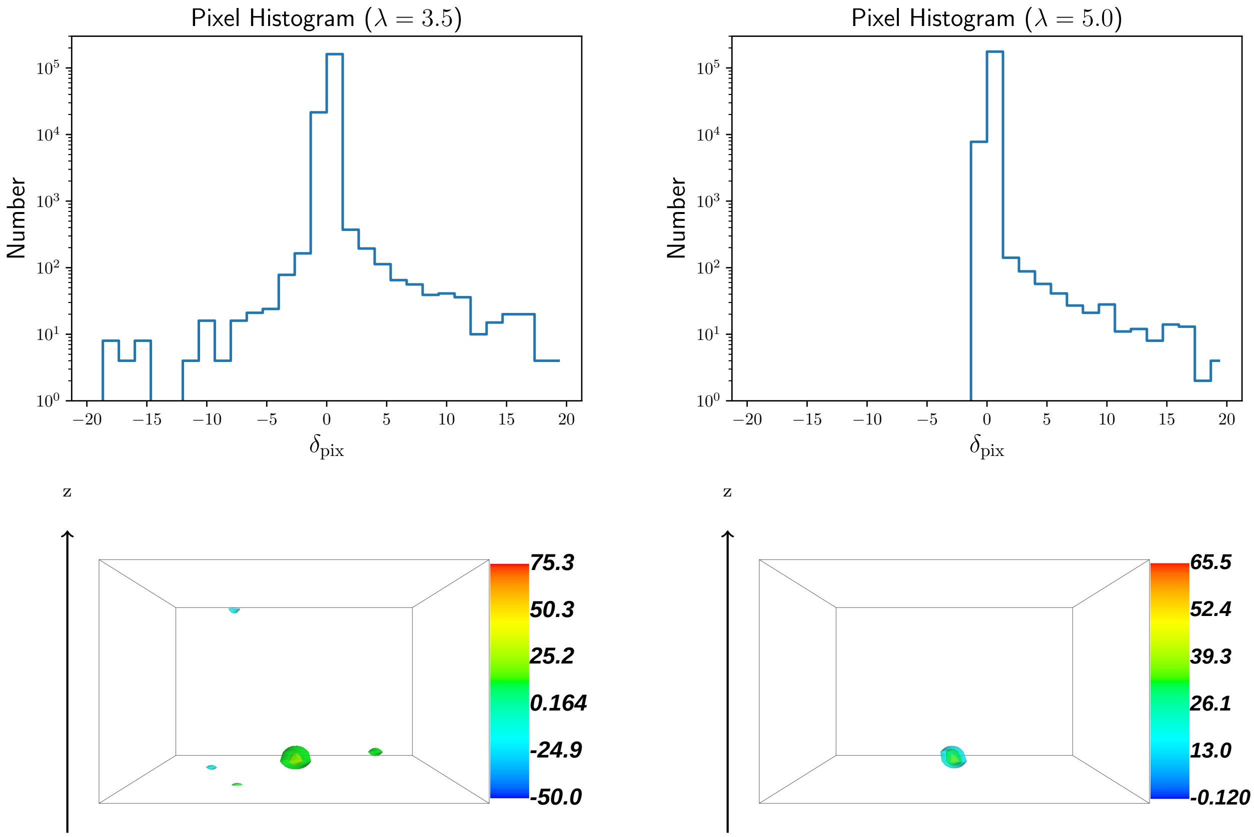
By randomly rotating the galaxies in the shape catalog, we simulate the HSC-like shape estimation errors with different realizations. The histogram of the first component of the HSC-like shape estimation error on galaxy level is shown in Figure 7. The corresponding histogram of the shape measurement error on the pixel level after the smoothing and pixelization is also shown in 7. The standard deviation map of the noise is demonstrated in Figure 8. As demonstrated in Figure 7, even though the shape measurement error on the galaxy level does not fully follow Gaussian distribution, the error is well described by Gaussian distribution after the smoothing and pixelization.

3.2. *NFW atoms*

In this subsection, we test the performance of our algorithm with the default setup that models the matter density field with multi-scale NFW atoms. The dictionary is constructed with three frames of different NFW scale radii in the comoving coordinate: 0.12 h^{-1} Mpc, 0.24 h^{-1} Mpc, and 0.36 h^{-1} Mpc. The truncation radii are set to four times the comoving scale radii for the atoms in the dictionary (concentration equals four) . We note that each frame of our dictionary fixes the scale radius in the comoving coordinates; therefore, the NFW atoms have different angular radii in different lens redshift bins.

We test the algorithm with different regularization parameters (λ) for the preliminary lasso estimation, which are 3.5, 4.0, and 5.0. The corresponding regularization parameters for the final adaptive lasso estimations are set to $\lambda_{\rm ada} = \lambda^{\tau+1}$. Here, we note that both the preliminary lasso estimation and the final adaptive lasso estimation select the pixels with the SNRs greater than λ in each gradient descent iteration and estimate the density in the selected pixels. While the final adaptive lasso estimation further enhances the growth of the pixels with preliminary estimations greater than λ.

This paper does not go beyond the resolution limit defined by the Gaussian smoothing kernel with a standard deviation of $1.'5$ and the $1'$ pixel scale as discussed



NFW: $\lambda = 3.5$ NFW: $\lambda = 5.0$

Figure 9. The lower panels show the density maps reconstructed from the mock galaxy shape catalog with the NFW dictionary. The upper panels show the pixels' number histograms. The penalization parameters are $\lambda = 3.5$ (left) and $\lambda = 5.0$ (right). The input halo mass is $M_{200} = 10^{15.02}\ h^{-1} M_{\odot}$, and its redshift is $z = 0.164$. The vertical direction is the line of sight direction. The boxes' lower boundaries and upper boundaries of correspond to $z = 0.01$ and $z = 0.85$, respectively.

in Section 2.3.2 and Section 2.3.4, respectively. Therefore, we smooth the reconstructed density with the same Gaussian kernel in each lens redshift plane.

Figure 9 shows the 3-D density maps reconstructed with different penalization parameters for a halo with $M_{200} = 10^{15.02}\ h^{-1} M_{\odot}$ at redshift 0.164. Also, the pixels' histograms are shown in Figure 9. From these plots, we conclude that the adaptive lasso algorithm sets most of the reconstructed pixels to zero and only keeps the modes strongly related to the data. Moreover, the reconstructed density maps do not suffer from the line of sight smearing. After the reconstruction for each simulation, we identify the peaks on the sparse density map.

Following Lanusse et al. (2016), we normalize the detected peaks in the l-th ($l = 1...20$) lens redshift plane to account for the peak amplitude difference due to the difference in the norm of the lensing kernels for different redshift bins:

$$\delta^{\rm n}_{\rm peak}(\vec{\theta}, z_l) = \delta_{\rm peak}(\vec{\theta}, z_l)/\mathcal{R}_l^{\frac{1}{2}}, \tag{48}$$

where the normalization matrix is defined as

$$\mathcal{R}_l = \sum_s K^2(z_l, z_s). \tag{49}$$

In Figure 10, we show the stacked histograms of the normalized peaks with different penalization parameters. Here, we stack the histograms from 100 realizations of all halos sampled in the redshift-mass plane. Also, we simulate 1000 realizations of pure noise catalogs and perform the reconstructions on these noise catalogs to study the noise properties. The histograms of normalized peaks detected from the pure noise catalogs are shown in Figure 10 and we show the best-fit Gaussian functions of the noise peaks' histograms.

Figure 10 tells that the densities of peaks (including both true and false peaks) are suppressed as the penalization parameter λ increases. Moreover, we find the standard deviation of noise peaks slightly decreases as λ increases. As a result, for a higher detection threshold ($\lambda = 5.0$), we find a clearer peak number excess for mass maps reconstructed from mock catalogs comparing with the noise peak histogram .

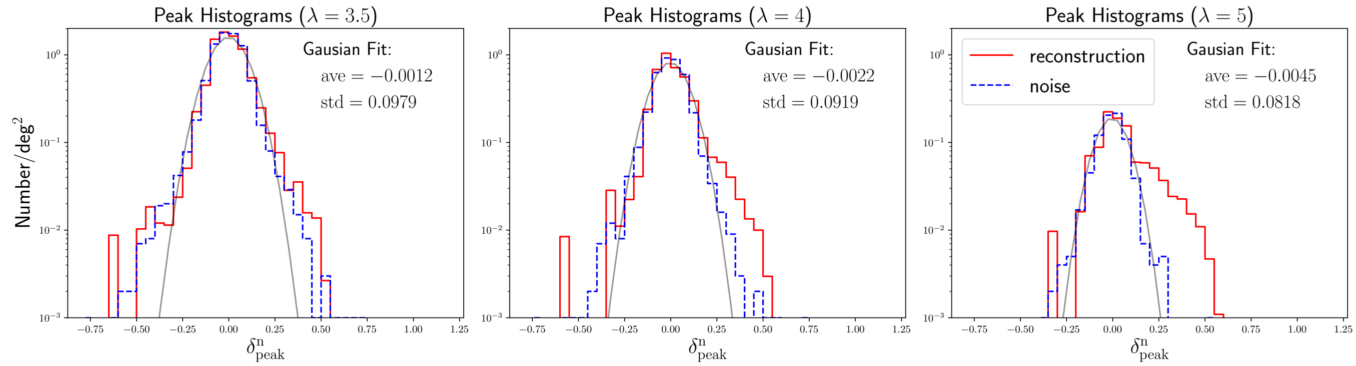


Figure 10. The number per square degree histograms of detected peak values from all of the simulations. The solid red steps result from reconstructions with the NFW dictionary penalized with different regularization parameters: $\lambda = 3.5, 4.0, 5.0$. The dashed blue steps are the corresponding results of the reconstructions from 1000 realizations of pure noise catalogs. The gray lines are the best-fit Gaussian distributions to the noises' peak histograms.

The 2-D histogram, stacked from all of the simulations, for the offsets of the detected peak positions from the input halos' positions is shown in the left panel of Figure 11. We see a clustering of peaks close to the input halo's position on the stacked position histogram. For each stamp simulation, we find the positive peak closest to the input position (in the pixel unit). If the closest peak lay inside the region denoted with the dashed box in the left panel of Figure 11, we take it as a true peak detection of the input halo. Other identified peaks, which include both positive and negative peaks, are taken as false detections.

The right panel of Figure 11 shows the average redshift of true detections for each halo. The estimated redshifts are lower than the true redshifts by about 0.03 for halos in the low-redshift range ($z \leq 0.4$). For halos at $0.4 < z \leq 0.85$, the relative redshift bias is below 0.5%.

With the intent to suppress false detections, we select peaks with values greater than an ad-hoc threshold as candidates of galaxy clusters following Miyazaki et al. (2018a). The threshold is set to a few times the standard deviation of the noise peaks. We use different detection thresholds (1.5σ and 3.0σ) to detect galaxy clusters from the mass maps reconstructed with $\lambda = 3.5, 4.0, 5.0$. The left and middle columns of Figure 12 show the detection rates for halos in the (mass, redshift) planes with detection thresholds set to 1.5σ and 3.0σ of the noise peaks' distributions, respectively. The right column of Figure 12 shows the corresponding numbers of false detections per square degree as functions of detection thresholds. The first, second, and third rows of Figure 12 correspond to the $lambda = 3.5, 4.0, 5.0$, respectively.

Figure 12 tells that the false peak density is suppressed as the detection threshold increases. Also, the detection rate of halo significantly decreases. We decide to set the detection threshold to 1.5σ and set the penalization parameter λ to 5.0 since such a setup suppresses the false detection to 0.022 while keeping a good halo detection rate. In summary, The algorithm is able to detect halo with minimal mass limits of $10^{14.0} M_{\odot}/h$, $10^{14.7} M_{\odot}/h$, $10^{15.0} M_{\odot}/h$ for the low ($z < 0.3$), median ($0.3 \leq z < 0.6$) and high ($0.6 \leq z < 0.85$) redshifts, respectively.

According to the detection rate measured from the simulation, we predict the number density of detected cluster using the halo mass function of Tinker et al. (2008). We use HMF (Murray et al. 2013), an open-source package, to calculate the halo mass function. The predicted halo detection number density for the setup ($\lambda = 5$, 1.5σ detection threshold) is shown in Figure 13. The expected cluster number density in total is 0.49 deg^{-2}, which corresponds to 78.4 clusters for the first year HSC shear catalog (Mandelbaum et al. 2018) with a survey area of $\sim 160\,\mathrm{deg}^2$. The expected number of detections is similar to the number of 2-D cluster detections on the first year HSC shape catalog (Miyazaki et al. 2018b).

3.3. *Point mass atoms*

We substitute the default NFW dictionary with the point mass dictionary and reconstruct the mass map from the mock galaxy shape catalog, to compare with the default setup.

The regularization parameter (λ) for the preliminary lasso is set to 3.5 and 5.0. Inspired by Pramanik & Zhang (2020), which propose to incorporate external group information into different adaptive lasso penalization weights by setting the penalization weights for projection coefficients in the same group to the average of the adaptive weights in this group, we smooth the preliminary lasso estimation in each lens redshift bin with a top-hat filter of comoving diameter: 0.25 h^{-1} Mpc. The smoothed preliminary lasso estimation is de-

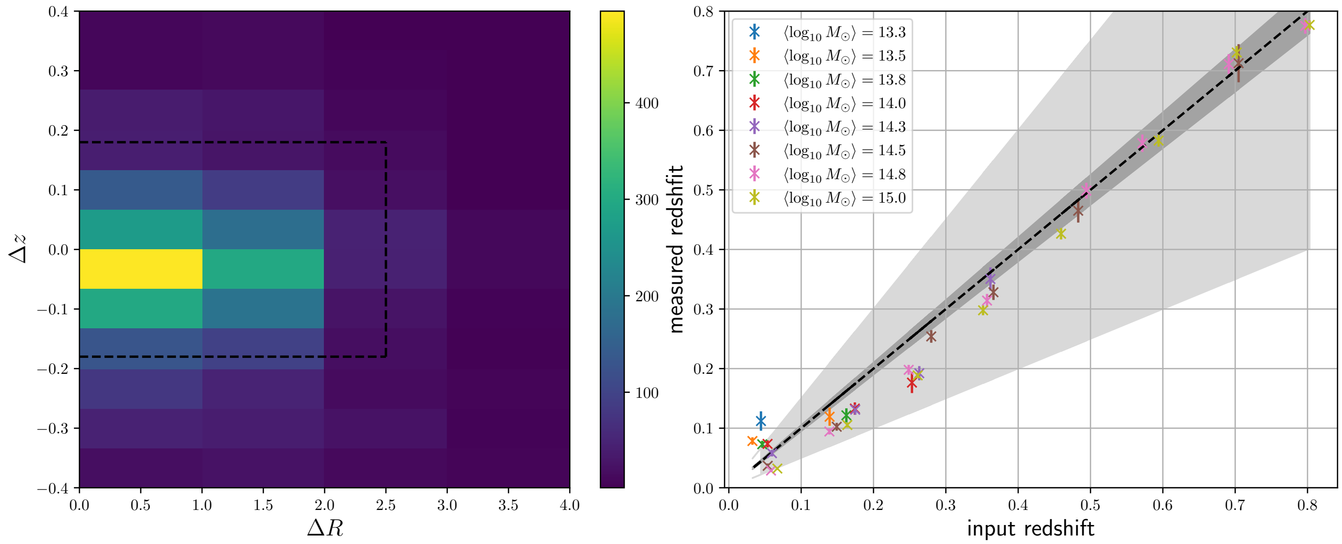


Figure 11. The left panel shows the stacked 2-D distribution of the deviations of detected peak positions from the centers of the corresponding input halos. The x-axis is for the deviated distance in the transverse plane, and the y-axis is for the deviation of the redshift. For each simulation, the positive peak inside the dashed black box with the minimal offset (in the pixel unit) from the input halo's position is taken as a true detection. The right panel focuses on the deviation of detected peaks in the line of sight direction. The x-axis is the input halo redshifts, and the y-axis is the redshift of the detected peak. The '×' denotes the average redshift of detected peaks for each halo over different noise realizations, and the error-bars are the uncertainties of the average redshifts. The deep gray area is for the relative redshift bias less than 0.05, and the light gray area is for the relative redshift bias less than 0.5. These results in this figure are based on the NFW dictionary with $\lambda = 3.5$.

noted as $\hat{x}^{\rm ls}_{\rm sm}$, and the penalization weights are set to $\hat{w} = 1/\left|\hat{x}^{\rm ls}_{\rm sm}\right|^{\tau}$.

As demonstrated in Figure 14, the mass reconstructions with the point mass dictionary tend to assign masses to several different redshift bins in the neighboring region of the halo's center. In contrast, as demonstrated in Figure 9, the NFW dictionary manages to perform consistent mass reconstructions. We think the problem of the point mass dictionary originates from the fact that the profile of the point mass atom in the transverse plane is much more compact than the profile of the input halos, especially at low redshift.

4. SUMMARY

We develop a novel method to reconstruct 3-D density contrast maps from weak-lensing shear measurements and photometric redshift estimations. Our method models 3-D density contrast maps as summations of NFW atoms with different comoving radii. With the prior assumption that the clumpy masses sparsely distribute in the 3-D space, the density map is reconstructed using the adaptive lasso algorithm (Zou 2006). The method is tested with realistic simulations using HSC-like shape estimation error and photo-z uncertainty.

The findings of this paper are summarized as follows:

(i) The lasso algorithm's solution suffers from a smear of structure in the line of sight direction even in the absence of shape noise, and the adaptive lasso algorithm efficiently removes the line of sight smear of structure.

(ii) The algorithm is able to detect halo with minimal mass limits of $10^{14.0} M_{\odot}/h$, $10^{14.7} M_{\odot}/h$, $10^{15.0} M_{\odot}/h$ for the low ($z < 0.3$), median ($0.3 \leq z < 0.6$) and high ($0.6 \leq z < 0.85$) redshifts, respectively, with a false detection of $0.022/\mathrm{deg}^2$.

(iii)The estimated redshifts of the halos detected from the reconstructed mass maps are lower than the true redshift by about 0.03 for halos at low redshifts ($z \leq 0.4$). The relative redshift bias is below 0.5% for halos at $0.4 < z \leq 0.85$.

We will apply the method to the shear measurements of the HSC survey (Mandelbaum et al. 2018; Li et al. 2020) to perform galaxy cluster detection in our future work.

ACKNOWLEDGEMENTS

XL thank Yin Li and Jiaxin Han for useful discussions.

XL was supported by Global Science Graduate Course (GSGC) program of University of Tokyo and JSPS KAKENHI (JP19J22222).

REFERENCES

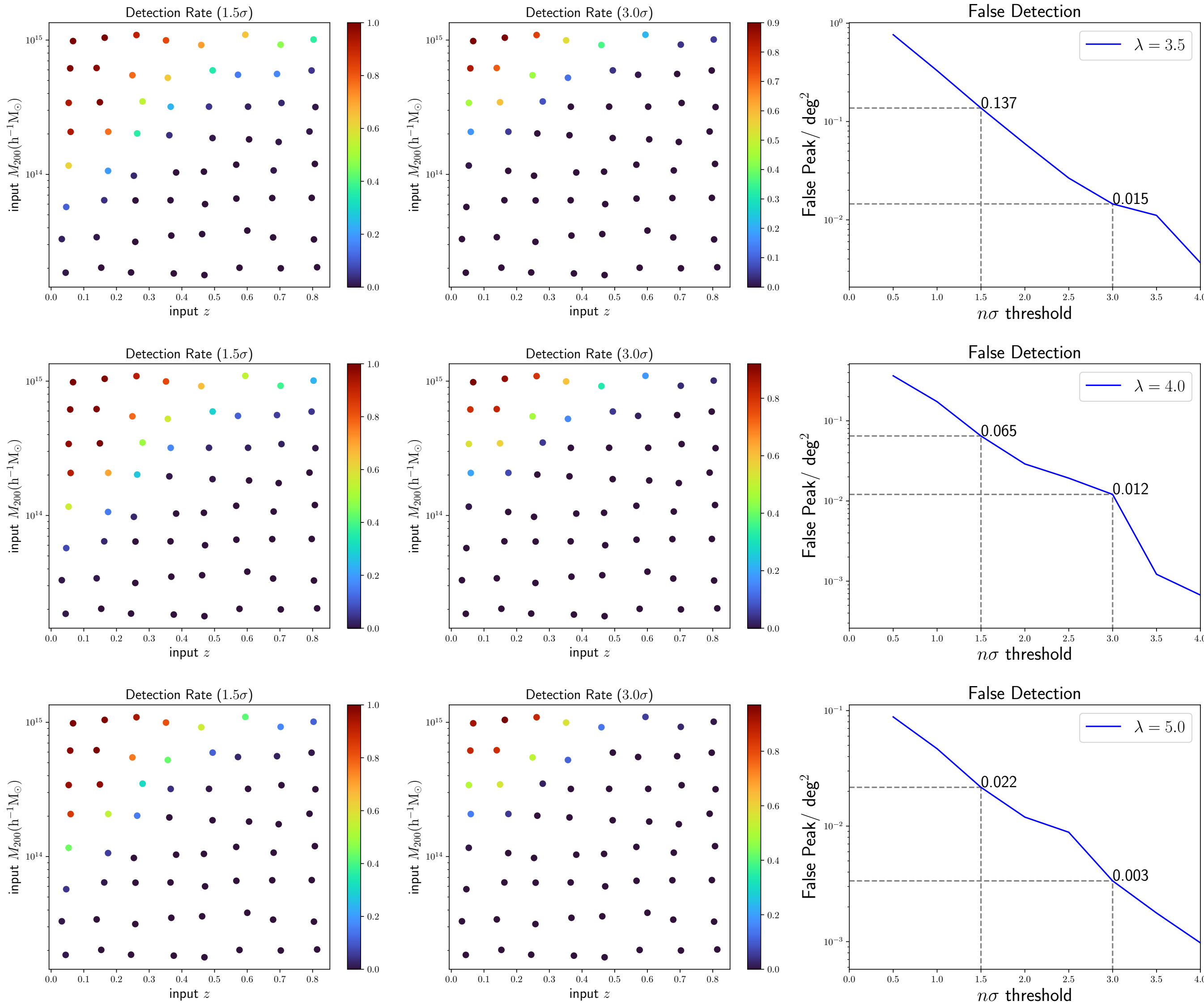


Figure 12. The detection rates and false peak densities for different penalization parameters and detection thresholds. The first, second, and third rows correspond to the results with $\lambda = 3.5, 4.0, 5.0$, respectively. The left and middle columns are the halo detection rates for detection thresholds equal 1.5σ and 3.0σ, respectively. The right column shows the density of false peaks as a function of detection threshold.

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

Beck, A., & Teboulle, M. 2009, SIAM Journal on Imaging Sciences, 2, 183

Carrasco Kind, M., & Brunner, R. J. 2013, MNRAS, 432, 1483, doi: 10.1093/mnras/stt574

Chang, C., Pujol, A., Mawdsley, B., et al. 2018, MNRAS, 475, 3165, doi: 10.1093/mnras/stx3363

de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, Experimental Astronomy, 35, 25, doi: 10.1007/s10686-012-9306-1

Fan, Z., Shan, H., & Liu, J. 2010, ApJ, 719, 1408, doi: 10.1088/0004-637X/719/2/1408

Hamana, T., Shirasaki, M., & Lin, Y.-T. 2020, PASJ, 72, 78, doi: 10.1093/pasj/psaa068

Hamana, T., Takada, M., & Yoshida, N. 2004, MNRAS, 350, 893, doi: 10.1111/j.1365-2966.2004.07691.x

Hennawi, J. F., & Spergel, D. N. 2005, ApJ, 624, 59, doi: 10.1086/428749

Hildebrandt, H., Köhlinger, F., van den Busch, J. L., et al. 2020, ap, 633, A69, doi: 10.1051/0004-6361/201834878

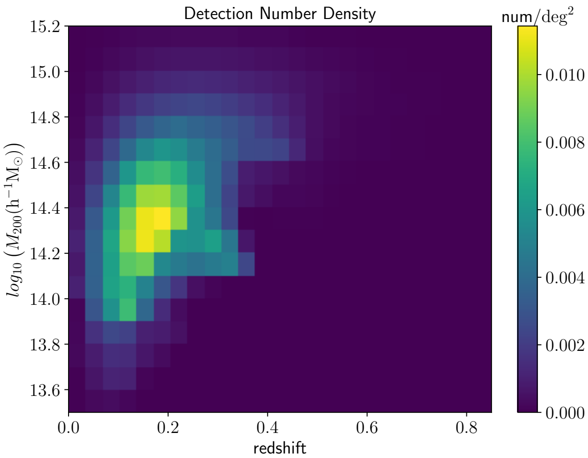


Figure 13. The expected number density of detected clusters per square degree as a function of halo's virial mass and redshift. The number density in total is 0.49 $\deg^{-2}$.

Jain, B., & Van Waerbeke, L. 2000, ApJL, 530, L1, doi: 10.1086/312480

Jeffrey, N., Abdalla, F. B., Lahav, O., et al. 2018, MNRAS, 479, 2871, doi: 10.1093/mnras/sty1252

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

Kilbinger, M. 2015, Reports on Progress in Physics, 78, 086901, doi: 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

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

Li, X., Oguri, M., Katayama, N., et al. 2020, The Astrophysical Journal Supplement Series, 251, 19, doi: 10.3847/1538-4365/abbad1

Lin, C.-A., Kilbinger, M., & Pires, S. 2016, A&A, 593, A88, doi: 10.1051/0004-6361/201628565

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

Mandelbaum, R., Miyatake, H., Hamana, T., et al. 2018, PASJ, 70, S25, doi: 10.1093/pasj/psx130

Miyazaki, S., Oguri, M., Hamana, T., et al. 2018a, PASJ, 70, S27, doi: 10.1093/pasj/psx120

—. 2018b, PASJ, 70, S27, doi: 10.1093/pasj/psx120

Murray, S. G., Power, C., & Robotham, A. S. G. 2013, Astronomy and Computing, 3, 23, doi: 10.1016/j.ascom.2013.11.001

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

Oguri, M., & Hamana, T. 2011, MNRAS, 414, 1851, doi: 10.1111/j.1365-2966.2011.18481.x

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

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, ap, 641, A6, doi: 10.1051/0004-6361/201833910

Pramanik, S., & Zhang, X. 2020, arXiv e-prints, arXiv:2006.02041. https://arxiv.org/abs/2006.02041

Price, M. A., Cai, X., McEwen, J. D., et al. 2020, MNRAS, 492, 394, doi: 10.1093/mnras/stz3453

Ragagnin, A., Dolag, K., Moscardini, L., Biviano, A., & D'Onofrio, M. 2019, MNRAS, 486, 4001, doi: 10.1093/mnras/stz1103

Schneider, P. 1996, MNRAS, 283, 837, doi: 10.1093/mnras/283.3.837

Shan, H., Kneib, J.-P., Tao, C., et al. 2012, ApJ, 748, 56, doi: 10.1088/0004-637X/748/1/56

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

Spergel, 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, 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

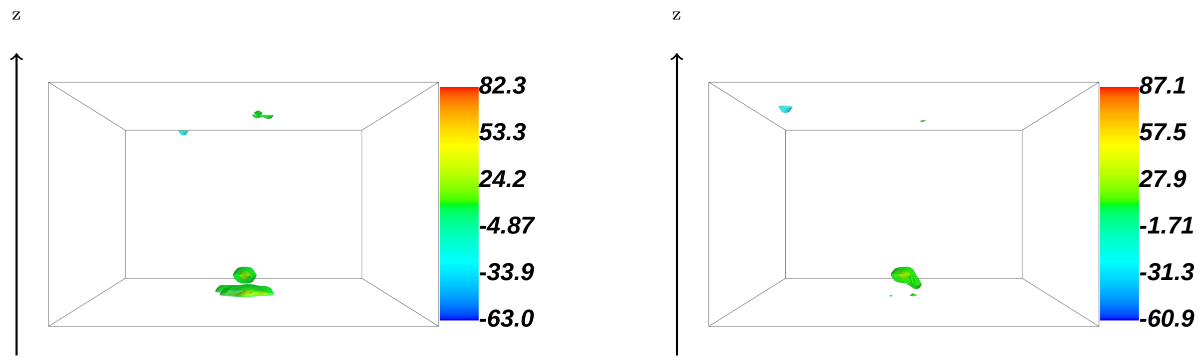
Tanaka, M., Coupon, J., Hsieh, B.-C., et al. 2018, Publications of the Astronomical Society of Japan, 70, S9, doi: 10.1093/pasj/psx077

The Dark Energy Survey Collaboration. 2005, ArXiv Astrophysics e-prints

Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, pj, 688, 709, doi: 10.1086/591439

VanderPlas, J. T., Connolly, A. J., Jain, B., & Jarvis, M. 2011, ApJ, 727, 118, doi: 10.1088/0004-637X/727/2/118

Zou, H. 2006, Journal of the American Statistical Association, 101, 1418, doi: 10.1198/016214506000000735



Point Mass: $\lambda = 3.5$ Point Mass: $\lambda = 5.0$

Figure 14. The density maps reconstructed from the mock galaxy shape catalog with the point mass dictionary. The penalization parameters are $\lambda = 3.5$ (left) and $\lambda = 5.0$ (right). The input halo mass is $M_{200} = 10^{15.02}\ h^{-1}M_{\odot}$, and its redshift is $z = 0.164$. The vertical direction is the line of sight direction. The boxes' lower boundaries and upper boundaries of correspond to $z = 0.01$ and $z = 0.85$, respectively.