A Compressed Sensing Approach to 3D Weak Lensing

Adrienne Leonard
Service d'Astrophysique
CEA - Saclay, Orme des Merisiers
91191 Gif sur Yvette CEDEX, France
Email: adrienne.leonard@cea.fr

François-Xavier Dupé
Laboratoire d'Informatique Fondamentale
UMR CNRS 7272
Aix-Marseille Université
Email: francois-xavier.dupe@lif.univ-mrs.fr

Jean-Luc Starck
Service d'Astrophysique
CEA - Saclay, Orme des Merisiers
91191 Gif sur Yvette CEDEX, France
Email: jstarck@cea.fr

***Abstract*—Weak gravitational lensing is an ideal probe of the dark universe. In recent years, several linear methods have been developed to reconstruct the density distribution in the Universe in three dimensions, making use of photometric redshift information to determine the radial distribution of lensed sources. We aim to address three key problems seen in these methods; namely, the bias in the redshifts of detected objects, the line-of-sight smearing seen in reconstructions, and the damping of the amplitude of the reconstruction relative to the underlying density. We also aim to detect structures at higher redshifts than have previously been achieved, and to improve the line-of-sight resolution of our reconstructions. We consider the problem under the framework of compressed sensing (CS). Under the assumption that the data are sparse or compressible in an appropriate dictionary, we construct a robust estimator and employ state-of-the-art convex optimisation methods to reconstruct the density contrast. We demonstrate that our method is able to accurately reproduce cluster haloes up to a redshift of $z_{\rm cl} = 1.0$, deeper than state-of-the-art linear methods. We directly compare our method with these linear methods, and demonstrate minimal radial smearing and redshift bias in our reconstructions, as well as a reduced damping of the reconstruction amplitude as compared to the linear methods. In addition, the CS framework allows us to consider an underdetermined inverse problem, thereby allowing us to reconstruct the density contrast at finer resolution than the input data.**

I. Introduction

Weak gravitational lensing has become a powerful tool for studying the dark universe, allowing us to place constraints on key cosmological parameters, and offering the possibility to place independent constraints on the dark energy equation of state parameter, w (Levy and Brustein, 2009; Hoekstra and Jain, 2008; Munshi et al., 2008; Albrecht et al., 2006; Peacock et al., 2006; Schneider, 2006; Van Waerbeke and Mellier, 2003).

Until recently, weak lensing studies considered the shear signal, and recovered the mass distribution, in two-dimensional projection (see Schneider, 2006, for a review of weak lensing). However, with improved data quality and wide-band photometry, it is now possible to recover the mass distribution in three dimensions by using photometric redshift information to deproject the lensing signal along the line of sight (Simon et al., 2009; Massey et al., 2007a,b; Taylor et al., 2004).

Under the assumption of Gaussian noise, various linear methods have been developed to recover the 3D matter distribution, which rely on the construction of a pseudo-inverse operator to act on the data, and which include a penalty function encoding the prior that is to be placed on the signal (VanderPlas et al., 2011; Simon et al., 2009; Castro et al., 2005; Taylor et al., 2004; Bacon and Taylor, 2003; Hu and Keeton, 2002).

These methods produce promising results; however, they show a number of problematic artefacts. Notably, structures detected using these methods are strongly smeared along the line of sight, the detected amplitude of the density contrast is damped (in some cases, very strongly), and the detected objects are shifted along the line of sight relative to their true positions. These effects result from the choice of method used; Simon et al. (2009) note that their choice of filter naturally gives rise to a biased solution, and VanderPlas et al. (2011) suggest that linear methods might be fundamentally limited in the resolution attainable along the line of sight as a result of the smearing effect seen in these methods.

Furthermore, these methods are restricted to deal solely with the overdetermined inverse problem. In other words, the resolution obtainable on the reconstruction of the density is limited to be, at best, equal to that of the input data. Thus, the resolution of the reconstruction is entirely limited by the quality of the data and its associated noise levels, with no scope for improvement by judicious choice of inversion or denoising method.

In this paper, we consider the 3D lensing problem as an instance of compressed sensing, where the sensing operator models the line-of-sight integration of the matter density giving rise to the lensing signal. For simplicity in implementation, and as a proof of concept of our method, we first reduce the 3D weak lensing problem to a one-dimensional inversion, with each line of sight in an image treated as independent.

Despite the loss of information this implies, we show that using such an approach produces reconstructions with minimal bias and smearing in redshift space, and with reconstruction amplitudes $\sim 75\%$ of the true amplitude (or better, in some cases). This is a significant improvement over current linear methods. In addition, our method exhibits an apparent increased sensitivity to high-redshift structures, as compared with linear methods. Our reconstructions do exhibit some noise, with false detections appearing along a number of lines

of sight. However, these tend to be localised to individual pixels, rather than forming coherent structures, and clearly arise from overfitting of the data.

We further present a full 3D treatment of the reconstruction problem, adopting a simple greedy algorithm to solve the inverse problem, and demonstrate much improved results, with a significant reduction in the number of false detections seen in the reconstruction.

Throughout the text, we assume a ΛCDM cosmology, with $\Omega_\Lambda = 0.736,\ \Omega_\mathrm{M} = 0.264,\ h = 0.71,\ \sigma_8 = 0.801$, consistent with the WMAP-7 results (Larson et al., 2011).

II. 3D Weak Lensing

The distortion of galaxy images caused by the weak lensing effect is described, on a given source plane, by the Jacobian matrix of the coordinate mapping between source and image planes:

$$A = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} , \tag{1}$$

where κ is the projected dimensionless surface density, and $\gamma = \gamma_1 + \mathrm{i}\gamma_2$ is the complex shear. The shear is related to the convergence via a convolution in two dimensions:

$$\gamma(\boldsymbol{\theta}) = \frac{1}{\pi}\int d^2\boldsymbol{\theta}'\mathcal{D}(\boldsymbol{\theta}-\boldsymbol{\theta}')\kappa(\boldsymbol{\theta}') , \tag{2}$$

where

$$\mathcal{D}(\boldsymbol{\theta}) = \frac{1}{(\boldsymbol{\theta}^*)^2} , \tag{3}$$

$\boldsymbol{\theta} = \theta_1 + \mathrm{i}\theta_2$, and an asterisk * represents complex conjugation.

The convergence, in turn, can be related to the 3D density contrast $\delta(\boldsymbol{r}) \equiv \rho(\boldsymbol{r})/\overline{\rho} - 1$ by

$$\kappa(\boldsymbol{\theta}, w) = \frac{3H_0^2\Omega_M}{2c^2}\int_0^w dw' \frac{f_K(w')f_K(w-w')}{f_K(w)}\frac{\delta[f_K(w')\boldsymbol{\theta}, w']}{a(w')} , \tag{4}$$

where H_0 is the hubble parameter, Ω_M is the matter density parameter, c is the speed of light, $a(w)$ is the scale parameter evaluated at comoving distance w, and

$$f_K(w) = \begin{cases} K^{-1/2}\sin(K^{1/2}w), & K>0 \\ w, & K=0 \\ (-K)^{-1/2}\sinh([-K]^{1/2}w) & K<0 \end{cases} , \tag{5}$$

gives the comoving angular diameter distance as a function of the comoving distance and the curvature, K, of the Universe.

If the shear (or convergence) data is divided into N_sp redshift bins, and the density contrast reconstruction is divided into N_lp redshift bins (where N_sp is not necessarily equal to N_lp), we can write the convergence $\kappa^{(i)}$ on each source plane as

$$\kappa^{(i)}(\boldsymbol{\theta}) \simeq \sum_{\ell=1}^{N_\mathrm{lp}} Q_{i\ell}\delta^{(\ell)}(\boldsymbol{\theta}) , \tag{6}$$

where

$$Q_{i\ell} = \frac{3H_0^2\Omega_M}{2c^2}\int_{w_\ell}^{w_{\ell+1}} dw \frac{\overline{W}^{(i)}(w)f_K(w)}{a(w)} , \tag{7}$$

and

$$\overline{W}^{(i)}(w) = \int_0^{w^{(i)}} dw' \frac{f_K(w-w')}{f_K(w')}\left(p(z)\frac{dz}{dw}\right)_{z=z(w')} . \tag{8}$$

Thus, for each line of sight, equation (6) describes a matrix multiplication, encoding a convolution along the line of sight. It is the inversion of this transformation:

$$\kappa(z) = \mathbf{Q}\delta(z) , \tag{9}$$

that is the focus of this paper.

A. Linear inversion methods

We focus here on the method presented in Simon et al. (2009). For a review of other linear methods, the reader is referred to Hu and Keeton (2002), and for an alternative linear inversion method, see VanderPlas et al. (2011).

The 3D lensing problem is effectively one of observing the density contrast convolved with the linear operator $\mathbf{R}$, and contaminated by noise, which is assumed to be Gaussian. Formally, we can write the 3D lensing problem as

$$\boldsymbol{d} = \mathbf{R}\boldsymbol{s} + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2) , \tag{10}$$

where $\boldsymbol{d}$ is the observation, $\boldsymbol{s}$ the real density and ε the Gaussian noise.

The general idea behind linear inversion methods is to find a linear operator $\mathbf{H}$ that acts on the data vector to yield a solution that minimises some functional, such as the variance of the residual between the estimated signal and the true signal, subject to some regularisation or prior-based constraints.

Simon et al. (2009) opt to use a Saskatoon filter (Tegmark, 1997; Tegmark et al., 1997), which combines a Wiener filter and an inverse variance filter, with a tuning parameter α introduced that allows switching between the two. This gives rise to a minimum variance filter, expressed as

$$\hat{\boldsymbol{s}}_{MV} = [\alpha\mathbf{1} + \mathbf{S}\mathbf{R}^\dagger\boldsymbol{\Sigma}^{-1}\mathbf{R}]^{-1}\mathbf{S}\mathbf{R}^\dagger\boldsymbol{\Sigma}^{-1}\boldsymbol{d} , \tag{11}$$

where $\mathbf{S} \equiv \langle \boldsymbol{s}\boldsymbol{s}^\dagger \rangle$ encodes prior information about the signal covariance, $\Sigma \equiv \langle \boldsymbol{n}\boldsymbol{n}^\dagger \rangle$ gives the covariance matrix of the noise, and $\mathbf{1}$ is the identity matrix.

This switching is designed to allow a balance between the increased constraining power offered by the Wiener filter over the inverse variance filter – which yields an improved signal-to-noise in the reconstruction – and the biasing that the Wiener filter imposes on the solution.

As discussed extensively in Simon et al. (2009) and VanderPlas et al. (2011), linear methods give rise to a significant bias in the location of detected peaks, damping of the peak signal, and a substantial smearing of the density along the line of sight. The compressed sensing (CS) theory, described below, allows us to address the lensing inversion problem under a new perspective, and we will show that these three aspects are significantly improved using a non-linear CS approach.

III. Compressed sensing approach

A. Compressed sensing theory

We consider some data Y_i ($i \in [1, .., m]$) acquired through the linear system

$$Y = \mathbf{\Theta} X \ , \tag{12}$$

where $\mathbf{\Theta}$ is an $m \times n$ matrix. Compressed sensing (Candès and Tao, 2006; Donoho, 2006) is a sampling/compression theory based on the sparsity of the observed signal, which shows that, under certain conditions, one can exactly recover a k-sparse signal (a signal for which only k pixels have values different from zero, out of n total pixels, where $k < n$) from $m < n$ measurements.

This recovery is possible from undersampled data only if the sensing matrix $\mathbf{\Theta}$ verifies the *restricted isometry property* (RIP) (see Candès and Tao, 2006, for more details). This property has the effect that each measurement Y_i contains some information about the entire pixels of X; in other words, the sensing operator $\mathbf{\Theta}$ acts to spread the information contained in X across many measurements Y_i.

Under these two constraints – sparsity and a transformation meeting the RIP criterion – a signal can be recovered exactly even if the number of measurements m is much smaller than the number of unknown n. This means that, using CS methods, we will be able to outperform the well-known Shannon sampling criterion by far.

The solution X of (12) is obtained by minimizing

$$\min_X \|X\|_1 \quad s.t. \ \ Y = \mathbf{\Theta} X \ , \tag{13}$$

where the ℓ_1 norm is defined by $\|X\|_1 = \sum_i \mid X_i \mid$. The ℓ_1 norm is well-known to be a sparsity-promoting function; i.e. minimisation of the ℓ_1 norm yields the most sparse solution to the inverse problem. Many optimisation methods have been proposed in recent years to minimise this equation. More details about CS and ℓ_1 minimisation algorithms can be found in Starck et al. (2010).

In real life, signals are generally not "strictly" sparse, but are *compressible*; i.e. we can represent the signal in a basis or frame (Fourier, Wavelets, Curvelets, etc.) in which the curve obtained by plotting the obtained coefficients, sorted by their decreasing absolute values, exhibits a polynomial decay. Note that most natural signals and images are compressible in an appropriate basis.

We can therefore reformulate the CS equation above (Equation (13)) to include the data transformation matrix $\mathbf{\Phi}$:

$$\min_\alpha \|\alpha\|_1 \quad s.t. \ \ Y = \mathbf{\Theta}\mathbf{\Phi}\alpha \ , \tag{14}$$

where $X = \mathbf{\Phi}^* \alpha$, and α are the coefficients of the transformed solution X in $\mathbf{\Phi}$, which is generally referred to as the *dictionary*. Each column represents a vector (also called an *atom*), which ideally should be chosen to match the features contained in X. If $\mathbf{\Phi}$ admits a fast implicit transform (e.g. Fourier transform, Wavelet transform), fast algorithms exist to minimise Equation (14).

One problem we face when considering CS in a given application is that very few matrices meet the RIP criterion. However, it has been shown that accurate recovery can be obtained as long the mutual coherence between $\mathbf{\Theta}$ and $\mathbf{\Phi}$, $\mu_{\mathbf{\Theta},\mathbf{\Phi}} = \max_{i,k} |\left\langle \mathbf{\Theta}_i, \mathbf{\Phi}_k, \right\rangle|$, is low (Candes and Plan, 2010). The mutual coherence measures the degree of similarity between the sparsifying basis and the sensing operator. Hence, in its relaxed definition, we consider a linear inverse problem $Y = \mathbf{\Theta}\mathbf{\Phi} X$ as being an instance of CS when

1) the problem is underdetermined,
2) the signal is compressible in a given dictionary $\mathbf{\Phi}$,
3) the mutual coherence $\mu_{\mathbf{\Theta},\mathbf{\Phi}}$ is low. This will happen every time the matrix $\mathbf{A} = \mathbf{\Theta}\mathbf{\Phi}$ has the effect of spreading out the coefficients α_j of the sparse signal on all measurements Y_i.

Most CS applications described in the literature are based on such a soft CS definition. Compressed sensing was introduced for the first time in astronomy for data compression (Bobin et al., 2008; Barbey et al., 2011), and a direct link between CS and radio-interferometric image reconstruction was recently established in Wiaux et al. (2009), leading to dramatic improvement thanks to the sparse ℓ_1 recovery (Li et al., 2011).

The 3D weak lensing reconstruction problem can be seen to completely meet the soft-CS criteria above. The problem is underdetermined, as we seek a higher resolution than can be attained in the noise-limited observations, the matter density in the Universe is sparsely distributed, and the lensing operator spreads out the underlying density in a compressed sensing way.

Sparsity prior: Sparse priors have been shown to be very useful in regularising ill-posed inverse problems (see Fadili and Starck, 2009, and references therein). In addition, a sparse prior using a wavelet basis has been used in many areas of signal processing in astronomy, such as denoising, deconvolution, and inpainting to recover missing data in images (Starck et al., 2010) . The idea underlying these priors is that there exists a dictionary in which a given dataset is sparsely represented. The dictionary used should therefore match the shapes of the structures that we aim to detect as closely as possible.

Many experiments, such as N-body simulations, have shown that the matter density in the universe is mainly distributed in localised clusters, which are connected by thin filaments. Because structures in the Universe appear to be physically sparse, we may therefore assume that the matter density is sparse in a domain adapted for cluster- and curve-like structures. These domains exist and can be constructed by gathering several well-chosen transforms inside a dictionary.

For the reconstruction of clusters of galaxies, we are in a perfect situation for sparse recovery because clusters are localised in the angular domain, and are not resolved along the line of sight owing to the bin size. They can therefore be modelled as Dirac $\delta-$functions along the line of sight, while an isotropic wavelet basis can be used in the angular domain.

B. Problem statement

Under the CS framework, the reconstruction of the matter density amounts to finding the most sparse solution that is consistent with the data. There are many different ways to formulate such an optimisation problem, and we opt for the following:

$$\min_{\boldsymbol{s}\in\mathbb{R}^n} \|\boldsymbol{\Phi}^*\boldsymbol{s}\|_1 \;\; s.t. \;\; \tfrac{1}{2}\|\boldsymbol{d}-\mathbf{R}\boldsymbol{s}\|^2_{\boldsymbol{\Sigma}^{-1}} \leq \epsilon \quad \boldsymbol{s}\in\mathcal{C}\ , \qquad (15)$$

where the term to minimise is a sparsity-penalty function over the dictionary coefficients. The second term in Equation (15) above is a data fidelity constraint, with $\boldsymbol{\Sigma}$ being the covariance matrix of the noise and ϵ the allowed distance between the estimation and the observation, while the final term forces the solution to have values inside a given interval, usually $\mathcal{C} = [-1, +\infty[^n$ for matter overdensity.

Note that this latter constraint, encoding a hard minimum on the signal to be recovered, is not possible with linear methods, and is therefore an additional strength of our method. Enforcing these physical constraints on our solution helps to ensure the recovery of the most physically compelling solution given the data.

In order to solve (15), we use the primal-dual splitting method of Chambolle and Pock (2011), which is described in full in Leonard et al. (2012a).

IV. Simulations and Noise Modelling

A. Cluster simulations

In order to test our method, we needed to simulate a realistic data set. To this end, we consider here a fiducial survey with a background galaxy number density $n_g = 100\,\mathrm{arcmin}^{-2}$ distributed in redshift according to

$$p(z) \propto z^2 \mathrm{e}^{-(1.4z/z_0)^{1.5}}\ , \qquad (16)$$

with $z_0 = 1.0$ (Taylor et al., 2007; Kitching et al., 2011), and the distribution truncated at a maximum redshift of $z_{max} = 2$. Figure 1 shows this probability distribution, normalised arbitrarily.

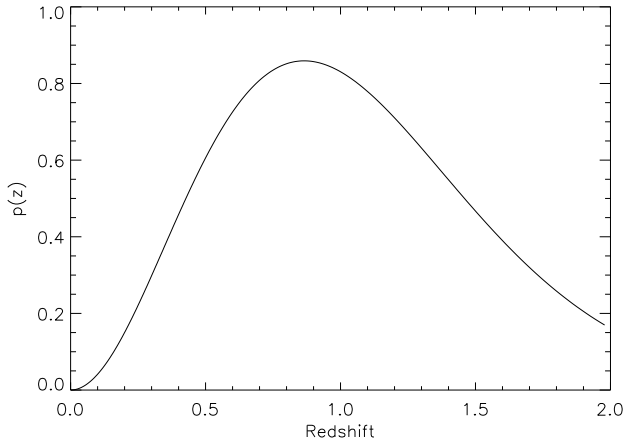

Fig. 1. Redshift probability distribution $p(z)$ of the sources as a function of source redshift z for the simulations described in the text.

We took the intrinsic dispersion in shear measurements to be $\sigma_\gamma = 0.2$, and considered a field of $1^\circ \times 1^\circ$ divided into a grid of 60×60 pixels. These parameters were chosen to mimic the data quality expected from next-generation surveys such as Euclid (Refregier et al., 2010) and LSST (Ivezic et al., 2008; LSST Science Collaboration et al., 2009).

In each simulated image, clusters were generated following an NFW density profile with $M_{200} = 10^{15}\,M_\odot$, $c = 3$ binned into $N_{\rm sp}$ redshift bins. The effective convergence and shear were computed by integrating the lensing signal within each source redshift bin, and Gaussian noise was added, scaled appropriately by the number density of galaxies within that bin.

B. Noise modelling and control

We assumed that the redshifts of the sources are known exactly, so there is no correlation between the noise in each source bin. Therefore, the covariance matrix of the noise along the line of sight is diagonal, with

$$\boldsymbol{\Sigma}_{ii} = \sigma_g^2(z_i) = \frac{\sigma_\gamma^2}{n_g(z_i)A_{pix}}\ , \qquad (17)$$

where A_{pix} is the pixel area, $n_g(z_i)$ is the number density of sources in the bin at redshift z_i, and σ_γ is the intrinsic dispersion in galaxy ellipticity, taken throughout to be 0.2.

V. 1D implementation of the CS approach

A. Reconstructions in 1D

As noted in § II, the 3D weak lensing problem can be reduced to a one-dimensional problem, by taking as our data vector the (noisy) lensing convergence along each line of sight, which is related to the density contrast through Equation (9). Therefore, we took $\boldsymbol{d} = \boldsymbol{\kappa}_{ij}(z)$ and $\mathbf{R} = \mathbf{Q}$, and considered each line of sight in our images independently. Furthermore, we took $\boldsymbol{\Phi}$ to be a δ-function dictionary, owing to the fact that clusters are poorly resolved along the line of sight at the resolution of our reconstructions.

Clearly, a one-dimensional implementation throws away information, because we do not account at all for the correlation between neighbouring lines of sight that will arise in the presence of a large structure in the image; however, reducing the problem to a single dimension is fast and easy to implement, and allows us to test the efficacy of the algorithm using a particularly simple basis function through which we impose sparsity.

However, the algorithm used is entirely general; therefore, with appropriate choice of a 3D basis set and taking $\boldsymbol{d} = \boldsymbol{\gamma}(\boldsymbol{\theta}, z)$ and $\mathbf{R} = \mathbf{P}_{\gamma\kappa}\mathbf{Q}$, one can implement this algorithm as a fully 3D treatment of the data with no modification to the algorithm itself.

B. Comparison with linear methods

Firstly, to demonstrate the effectiveness of our method, we compared our method directly with the radial Wiener filter method of Simon et al. (2009), which is shown in Leonard et al. (2012a) to be the most effective linear method.

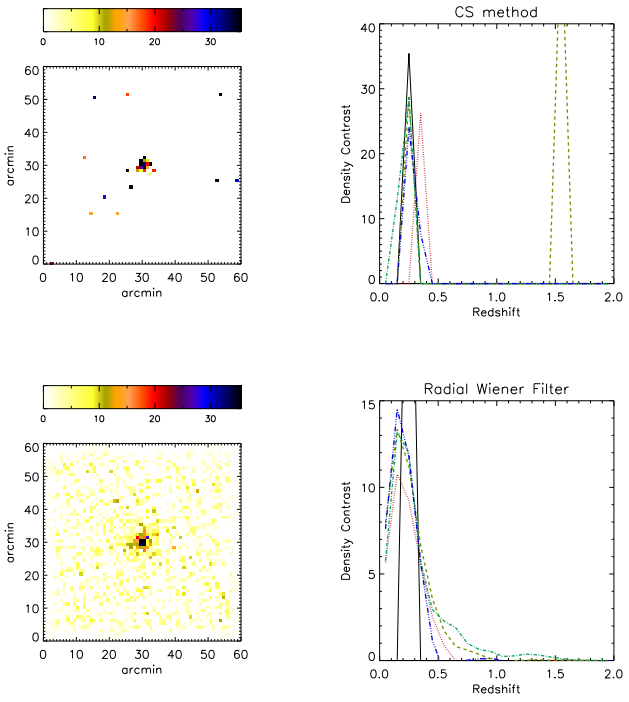

Fig. 2. Comparison of our method with the radial Wiener filter. The left column shows the 2D projection of the reconstruction, while the right column shows the 1D reconstructions along the four central lines of sight (dashed lines). Note that, owing to the amplitude damping effect in the Wiener reconstruction, the y-axis scaling is different in each of the line-of-sight plots.

Because linear methods are only defined for $N_{\rm lp} \leq N_{\rm sp}$, we considered the case of $N_{\rm lp} = N_{\rm sp} = 20$, and generated a single cluster halo at a redshift of $z_{\rm cl} = 0.25$ following an NFW halo profile with $M_{200} = 10^{15} M_\odot$ and $c = 3$. We chose the Wiener filter tuning parameter to be $\alpha = 0.05$. For further details of the algorithm parameters set in this comparison, the reader is referred to Leonard et al. (2012a)

Note that while linear methods take $\boldsymbol{d} = \boldsymbol{\gamma}(\boldsymbol{\theta}, z)$, $\mathbf{R} = \mathbf{P}_{\gamma\kappa}\mathbf{Q}$, our method takes $\boldsymbol{d} = \boldsymbol{\kappa}(\boldsymbol{\theta}, z)$, $\mathbf{R} = \mathbf{Q}$ as before. The noise levels in each case are identical.

The results are presented in figures 2 and 3. Figure 2 presents the 2D projections of the reconstructions, computed by integrating the reconstruction along each line of sight, and the 1D reconstructions along the four central lines of sight. In the 3D renderings of Figure 3, the reconstructions are thresholded at $\delta = 3$ (i.e. the plot only shows $\delta_{\rm rec} \geq 3$), and each is smoothed with a Gaussian of width $\sigma = 0.7\,$pix in all three directions.

The radial and Wiener filter method shows very little noise in the 2D projections, and little smearing of the reconstruction along the line of sight. However, this method fails to correctly recover correctly the redshift of the cluster. Futhermore, the amplitude of the cluster reconstruction is a factor of ~ 2.5 smaller than the input density.

Our results are seen to suffer from several prominent, pixel-scale false detections along noisy lines of sight not associated with the cluster. However, along the four central lines of sight an excellent correlation between the input density contrast and our reconstruction is seen. One line of sight exhibits a prominent high-redshift false detection; but this does not appear in the remaining three lines of sight, and the overall amplitude of the reconstruction is around 75% of the true value. The 3D rendering demonstrates that the noise in our reconstruction shows very little coherent structure (i.e. tends to be restricted to isolated pixels), and is largely low-level. Moreover, the cluster is incredibly well-localised in redshift space, with the smearing seen in the figure primarily arising from the applied smoothing.

C. *Improving the line-of-sight resolution*

Given the success of our method at reconstructing lines of sight at the same resolution as the input data, it is interesting to consider whether we are able to improve on the output resolution of our reconstructions. It is also worthwhile to test our ability to detect clusters at higher redshifts than those considered above, given that compressed sensing is specifically designed to tackle underdetermined inverse problems. Indeed, noise-free simulations suggest that a resolution improvement of up to a factor of 4 in the redshift direction may be possible with this method.

Therefore, we generated clusters as before, with our data binned into $N_{\rm sp} = 20$ redshift bins, but aimed to reconstruct onto $N_{\rm lp} = 25$ redshift bins. We furthermore consider clusters at redshifts of $z_{\rm cl} = 0.2,\ 0.6,$ and 1.0. Given the changed reconstruction parameters, we modified our noise control parameters slightly and, for all results that follow, used $\lambda = 7.5,\ \epsilon = 2.8$.

Figure 4 shows the two-dimensional projection of the reconstruction and the 1D reconstruction of the four central lines of sight as before. Figure 5 shows the reconstructions of these haloes using our method as a 3D rendering. The 3D rendering is, as before, thresholded at $\delta_{\rm rec} = 3$ and smoothed with a Gaussian of width $\sigma = 0.7\,$pix.

Several features are immediately apparent. Firstly, all three clusters are clearly identified by our method. This is particularly impressive for the $z_{\rm cl} = 1.0$ cluster, because linear methods have, thus far, been unable to reconstruct clusters at such a high redshift (see, e.g. Simon et al., 2009; VanderPlas et al., 2011). We note that this detection is dependent on the redshift distributions of sources, however, and the lack of detection in Simon et al. (2009) and VanderPlas et al. (2011) may be due, in part, to their choice of probability distribution. However, we also note that the background source density in our sample is highly diminished behind the $z_{\rm cl} = 1.0$ cluster ($32.5\,\mathrm{arcmin}^{-2}$).

Again, the 3D renderings indicate that there is very little smearing of the reconstruction along the line of sight, in contrast with linear methods. This is also shown by the line-of-sight plots, in which the unsmoothed reconstructions show very localised structure. Furthermore, the reconstructions exhibit minimal redshift bias, and some lines of sight are seen to recover the amplitude of the density contrast without any notable damping.

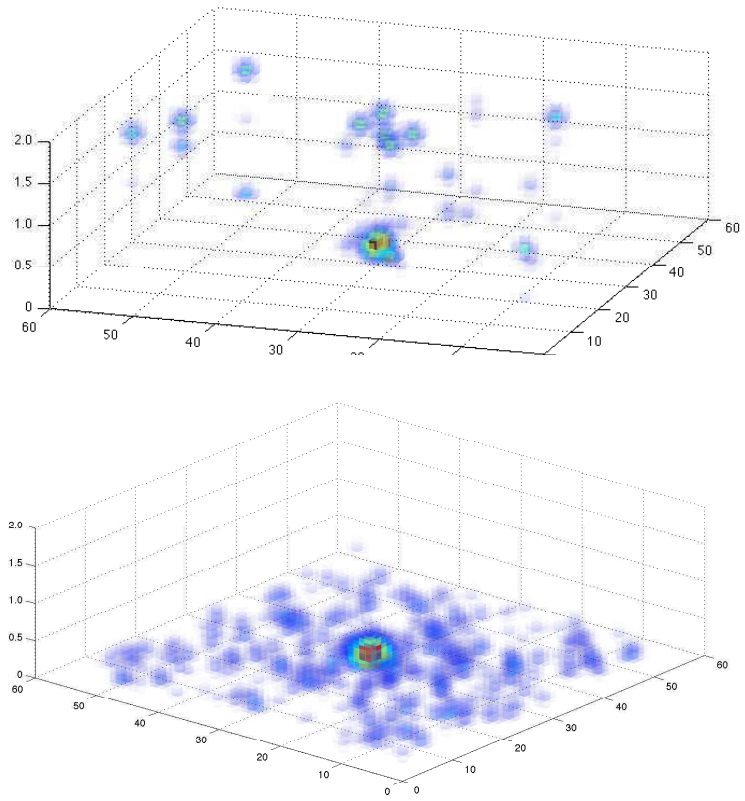

Fig. 3. 3D rendering of the reconstruction of a $z_{\rm cl} = 0.25$ cluster using our method (top), and the radial Wiener filter method (bottom $\alpha = 0.05$) as described in the text.

However, again we see several prominent "hot pixels" or false detections along noisy lines of sight. These detections are more evident as the cluster moves to high redshift, and may be significantly larger than the expected density contrast of the cluster. This is expected, because the number density of sources behind the lens diminishes.

We note that these false detections often manifest themselves at high redshift, arising out of the overfitting of the data at high redshift. We also note that the false detections are very well localised in both angular and redshift space, and do not form coherent large structures, which makes them easily identifiable as false detections; they tend to be localised to isolated pixels.

VI. 3D IMPLEMENTATION OF THE CS APPROACH

Considering each line of sight as independent throws away a large amount of information about the correlation between neighbouring lines of sight in the vicinity of matter concentrations. Moreover, such an algorithm imposes the sparsity constraint only locally – along each line of sight – rather than globally over the entire 3D data cube. Thus, a weaker constraint is imposed on the data, and the method is not optimal for finding the sparsest solution consistent with the data.

In order to treat the problem effectively in three dimensions, we still consider solving equation 9, however we must now impose our sparsity constraint on the full three-dimensional field.This is done by finding an appropriate three dimensional basis in which our data can be sparsely represented. As we are primarily seeking clusters of galaxies, the Dirac dictionary is still valid along the line of sight, while in the angular domain clusters are extended, largely isotropic structures for which a the starlet wavelet basis (Starck et al., 1998; Starck and Murtagh, 2006) is an appropriate choice.

While the 1D algorithm described above and in Leonard et al. (2012a) is entirely general, and can be applied to this 3D problem, we choose to approach the problem with a simpler algorithm based on a greedy method proposed by Nam et al. (2011). This method involves iterative hard-thresholding of the estimate of the solution, with the threshold chosen such that a certain fraction of the coefficients – chosen to be the smallest coefficients in the sparsifying dictionary – are discarded at each iteration.

We test this method on the $z_{\rm cl} = 0.2$ cluster simulation used previously, and the 3D rendering of the result is shown in Figure 6. For this reconstruction, we retain only 95% of the coefficients at each iteration. The cluster is again well localised in the redshift direction, with no smearing or bias seen along the line of sight, and the reconstruction exhibits very little damping of the cluster amplitude. More importantly, with this 3D treatment many of the pixel-scale false detections do not appear when a global sparsity constraint is applied.

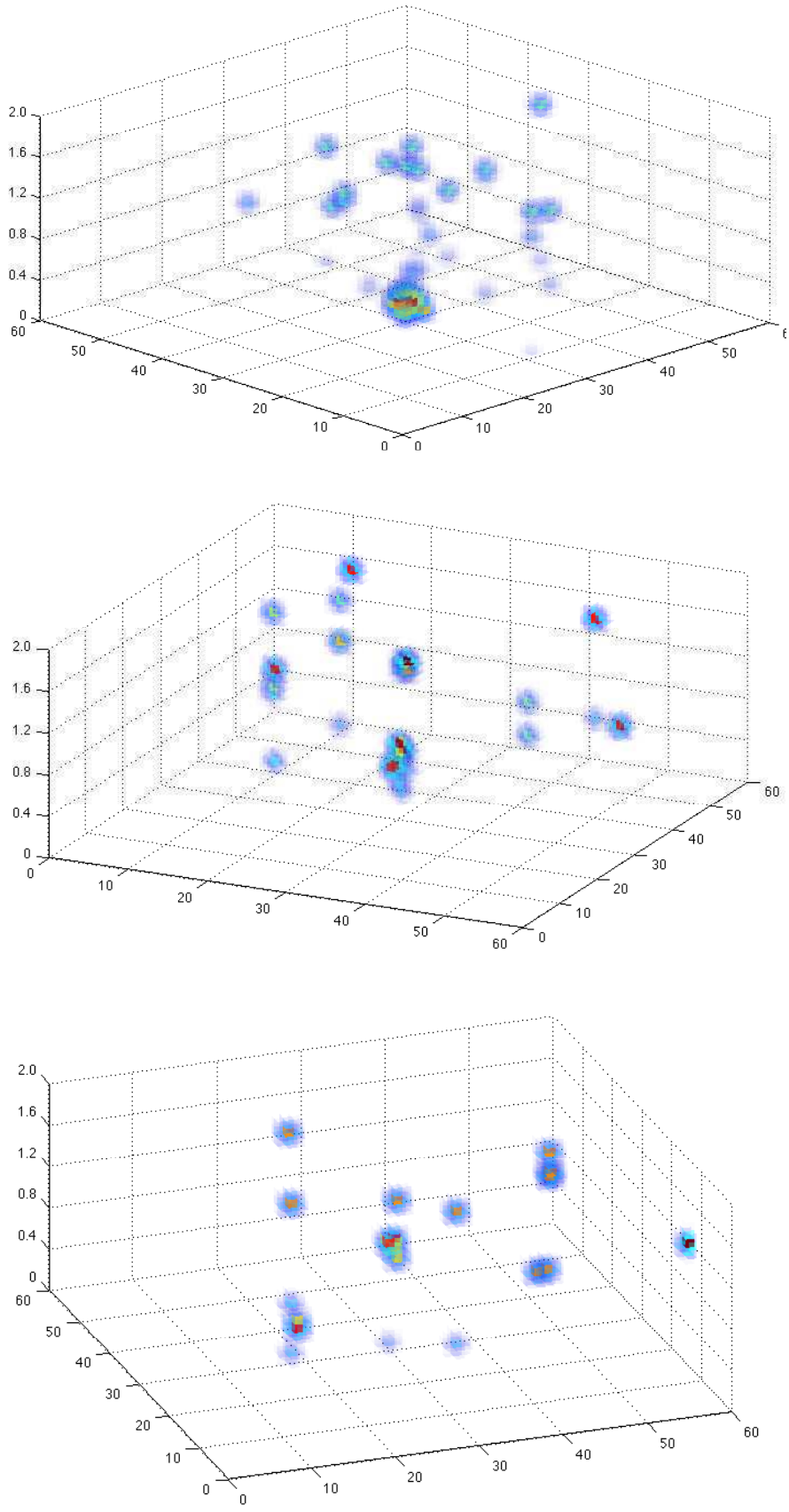

Fig. 5. Reconstructions of single clusters located at a redshift of $z_{\rm cl} = 0.2$ (top), $z_{\rm cl} = 0.6$ (middle) and $z_{\rm cl} = 1.0$ (bottom). The reconstructions are thresholded at $\delta = 3$ and smoothed with a Gaussian of width $\sigma = 0.7\,$pix.

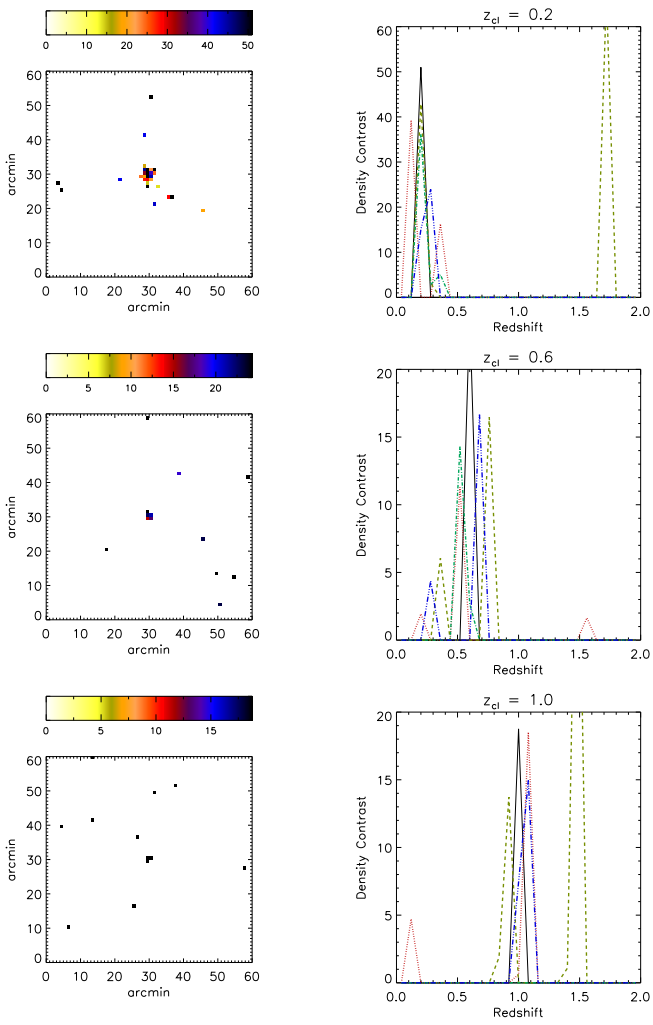

Fig. 4. Reconstructions of single clusters located at a redshift of $z_{\rm cl} = 0.2$ (top row), $z_{\rm cl} = 0.6$ (middle row) and $z_{\rm cl} = 1.0$ (bottom row). As before, the left column shows the two-dimensional integrated projection of the reconstruction, while the right panel shows the input density contrast along the line of sight (solid line) and the 1D reconstruction along each of the four central lines of sight (dashed lines).

This result is impressive, and demonstrates the power of the compressive sensing approach. One drawback of this particular choice of solver is that the choice of the fraction of coefficients to discard at each iteration is set rather arbitrarily. While the appropriate value for this threshold may be approximated considering simulations, it would be preferable to be able to select the important coefficients by means of a more physically-motivated criterion, and one which is more closely linked to the data.

Such a threshold, based on computing the noise level over the data cube at each wavelet scale and thresholding at an appropriate level for denoising (such as at $3-5\sigma$), is the subject of ongoing work (Leonard et al., 2012b).

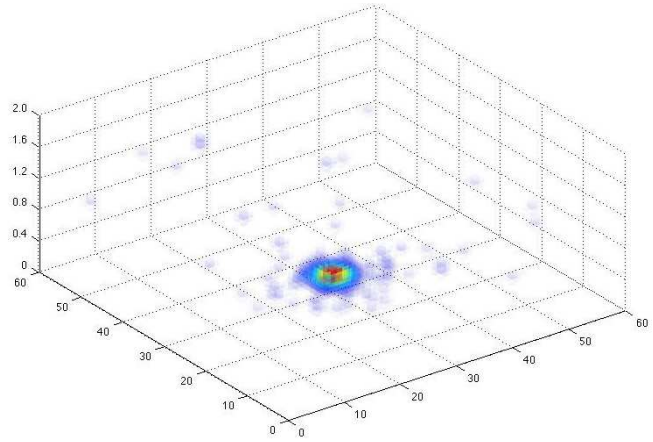

Fig. 6. 3D rendering of the reconstruction of a cluster at $z_{\rm cl} = 0.2$ using a 3D wavelet+Dirac basis and a greedy algorithm to solve the full 3D inversion.

VII. Summary and conclusions

Current approaches to 3D weak lensing involve linear inversion, where a pseudo-inverse operator is constructed that incorporates prior constraints on the statistical distribution of the measurement noise and the underlying density. These methods are straightforward to construct and implement, make use of common tools, and are usually fairly fast. This makes them a convenient choice when approaching the 3D weak lensing problem.

However, reconstructions obtained in this way suffer from line-of-sight smearing, bias in the detected redshift of structures, and a damping of the reconstruction amplitude relative to the input. It has furthermore been noted (VanderPlas et al., 2011) that the reconstructions obtained using these techniques may be fundamentally limited regarding the resolution attainable along the line of sight, because of smearing effects resulting from these linear methods. In addition, these methods are unable to treat an underdetermined inversion, and therefore are limited in their output resolution by the resolution of the input data which, in turn, is limited by the measurement noise.

We have presented a new approach to 3D weak lensing reconstructions by considering the weak lensing problem to be an instance of compressed sensing, where the underlying structure we aim to reconstruct is sparsely represented in an appropriate dictionary. Under such a framework, we were able to consider underdetermined transformations, thereby relaxing the constraints on the resolution of the reconstruction obtained using our method.

In a first implementation, we have reduced the problem to that of one-dimensional reconstructions along the line of sight. Whilst this is clearly not optimal because it throws away a lot of information, it allowed us to simplify the problem and to employ a particularly simple basis through which we imposed sparsity. We demonstrated that our method closely reproduces the position, radial extent and amplitude of simulated structures, with very little bias or smearing. This is a significant improvement over current linear methods.

Furthermore, we demonstrated an ability to reconstruct

clusters at higher redshifts than has been attainable using linear methods. Although our reconstructions exhibit false detections resulting from the noise, these noisy peaks do not form coherent structures, and are therefore well-localised and easily identifiable as noise peaks.

We have also presented a first implementation of a fully three-dimensional treatment of this problem, and demonstrated that – even while using a very simple algorithm – this treatment dramatically improves the quality of the reconstructions.

The quality of data available for weak lensing measurements continues to improve, and the methods by which we measure the weak lensing shear are becoming ever more sophisticated. So, too, must the methods we use to analyse the data to reconstruct the underlying density. While linear methods appear to be limited in resolution, and offer biased estimators, it is clear that a nonlinear approach such as ours does not, and – in a fully 3D implementation – may therefore allow us to map the cosmic web in far greater detail than has previously been achieved.

VIII. Acknowledgments

We would like to thank Jalal Fadili for useful discussions, and Jake VanderPlas for supplying the code for the linear reconstructions and for helpful comments. This work is supported by the European Research Council grant SparseAstro (ERC-228261).

References

Albrecht, A., Bernstein, G., Cahn, R., Freedman, W. L., Hewitt, J., Hu, W., Huth, J., Kamionkowski, M., Kolb, E. W., Knox, L., Mather, J. C., Staggs, S., and Suntzeff, N. B.: 2006, *ArXiv astro-ph/0609591*

Bacon, D. J. and Taylor, A. N.: 2003, *MNRAS* **344**, 1307

Barbey, N., Sauvage, M., Starck, J.-L., Ottensamer, R., and Chanial, P.: 2011, *A&A* **527**, A102+

Bobin, J., Starck, J.-L., and Ottensamer, R.: 2008, *IEEE Journal of Selected Topics in Signal Processing* **2**, 718

Candès, E. and Tao, T.: 2006, *IEEE Transactions on Information Theory* 52(12), 5406–5425

Candes, E. J. and Plan, Y.: 2010, *A Probabilistic and RIPless Theory of Compressed Sensing*, submitted

Castro, P. G., Heavens, A. F., and Kitching, T. D.: 2005, *Phys. Rev. D* **72(2)**, 023516

Chambolle, A. and Pock, T.: 2011, *Journal of Mathematical Imaging and Vision* **40**, 120

Donoho, D.: 2006, *IEEE Transactions on Information Theory* 52(4), 1289–1306

Fadili, M. and Starck, J.-L.: 2009, in *Proceedings of the International Conference on Image Processing, ICIP 2009, 7-10 November 2009, Cairo, Egypt*, pp 1461–1464, IEEE

Hoekstra, H. and Jain, B.: 2008, *Annual Review of Nuclear and Particle Science* **58**, 99

Hu, W. and Keeton, C. R.: 2002, *Phys. Rev. D* **66(6)**, 063506

Ivezic, Z., Tyson, J. A., Allsman, R., Andrew, J., Angel, R., and for the LSST Collaboration: 2008, *ArXiv: 0805.2366*

Kitching, T. D., Heavens, A. F., and Miller, L.: 2011, *MNRAS* pp 426–+

Larson, D., Dunkley, J., Hinshaw, G., Komatsu, E., Nolta, M. R., Bennett, C. L., Gold, B., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Page, L., Smith, K. M., Spergel, D. N., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L.: 2011, *ApJS* **192**, 16

Leonard, A., Dupé, F.-X., and Starck, J.-L.: 2012a, *A&A* **539**, A85

Leonard, A., Lanusse, F., Dupé, F.-X., and Starck, J.-L.: 2012b, *in prep.*

Levy, D. and Brustein, R.: 2009, *JCAP* **6**, 26

Li, F., Cornwell, T. J., and de Hoog, F.: 2011, *A&A* **528**, A31+

LSST Science Collaboration, Abell, P. A., Allison, J., Anderson, S. F., Andrew, J. R., Angel, J. R. P., Armus, L., Arnett, D., Asztalos, S. J., Axelrod, T. S., and et al.: 2009, *ArXiv: 0912.0201*

Massey, R., Rhodes, J., Ellis, R., Scoville, N., Leauthaud, A., Finoguenov, A., Capak, P., Bacon, D., Aussel, H., Kneib, J., Koekemoer, A., McCracken, H., Mobasher, B., Pires, S., Refregier, A., Sasaki, S., Starck, J., Taniguchi, Y., Taylor, A., and Taylor, J.: 2007a, *Nature* **445**, 286

Massey, R., Rhodes, J., Leauthaud, A., Capak, P., Ellis, R., Koekemoer, A., Réfrégier, A., Scoville, N., Taylor, J. E., Albert, J., Bergé, J., Heymans, C., Johnston, D., Kneib, J., Mellier, Y., Mobasher, B., Semboloni, E., Shopbell, P., Tasca, L., and Van Waerbeke, L.: 2007b, *ApJS* **172**, 239

Munshi, D., Valageas, P., van Waerbeke, L., and Heavens, A.: 2008, *Phys. Rep.* **462**, 67

Nam, S., Davies, M., Elad, M., and Gribonval, R.: 2011, in *CAMSAP - 4th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing - 2011*, San Juan, Puerto Rico

Peacock, J. A., Schneider, P., Efstathiou, G., Ellis, J. R., Leibundgut, B., Lilly, S. J., and Mellier, Y.: 2006, *ESA-ESO Working Group on "Fundamental Cosmology"*, Technical report, Ed. Universirty

Refregier, A., Amara, A., Kitching, T. D., Rassat, A., Scaramella, R., Weller, J., and Euclid Imaging Consortium, f. t.: 2010, *ArXiv: 1001.0061*

Schneider, P.: 2006, *Weak Gravitational Lensing*, pp 269–+, Springer

Simon, P., Taylor, A. N., and Hartlap, J.: 2009, *MNRAS* **399**, 48

Starck, J. and Murtagh, F.: 2006, *Astronomical Image and Data Analysis*

Starck, J., Murtagh, F. D., and Bijaoui, A.: 1998, *Image Processing and Data Analysis*, Cambridge University Press

Starck, J.-L., Murtagh, F., and Fadili, J. M.: 2010, *Sparse Image and Signal Processing*, Cambridge University Press

Taylor, A. N., Bacon, D. J., Gray, M. E., Wolf, C., Meisenheimer, K., Dye, S., Borch, A., Kleinheinrich, M., Kovacs, Z., and Wisotzki, L.: 2004, *MNRAS* **353**, 1176

Taylor, A. N., Kitching, T. D., Bacon, D. J., and Heavens, A. F.: 2007, *MNRAS* **374**, 1377

Tegmark, M.: 1997, *ApJL* **480**, L87+

Tegmark, M., de Oliveira-Costa, A., Devlin, M. J., Netterfields, C. B., Page, L., and Wollack, E. J.: 1997, *ApJL* **474**, L77+

Van Waerbeke, L. and Mellier, Y.: 2003, *ArXiv astro-ph/0305089*

VanderPlas, J. T., Connolly, A. J., Jain, B., and Jarvis, M.: 2011, *ApJ* **727**, 118

Wiaux, Y., Jacques, L., Puy, G., Scaife, A. M. M., and Vandergheynst, P.: 2009, *MNRAS* **395**, 1733