

Density mapping with weak lensing and phase information

Rafał M. Szepietowski,^{1,2★} David J. Bacon,¹ Jörg P. Dietrich,^{3,4} Michael Busha,⁵
Risa Wechsler^{6,7,8} and Peter Melchior^{9,10}

¹*Institute of Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building, Burnaby Road, Portsmouth PO1 3FX, UK*

²*Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, UK*

³*Universitäts-Sternwarte, Ludwig-Maximilians-Universität München, Scheinerstr. 1, D-81679 München, Germany*

⁴*Excellence Cluster Universe, Boltzmannstr. 2, D-85748 Garching b. München, Germany*

⁵*Institute for Theoretical Physics, Universität Zürich, CH-8057 Zürich, Switzerland*

⁶*Kavli Institute for Particle Astrophysics and Cosmology, 452 Lomita Mall, Stanford University, Stanford, CA 94305, USA*

⁷*Physics Department, Stanford University, Stanford, CA 94305, USA*

⁸*SLAC National Accelerator Laboratory, 2575 Sand Hill Rd, MS 29, Menlo Park, CA 94025, USA*

⁹*Center for Cosmology and Astro-Particle Physics, The Ohio State University, 191 W. Woodruff Ave, Columbus, OH 43210, USA*

¹⁰*Department of Physics, The Ohio State University, 191 W. Woodruff Ave, Columbus, OH 43210, USA*

Accepted 2014 February 25. Received 2014 February 4; in original form 2013 June 21

ABSTRACT

The available probes of the large-scale structure in the Universe have distinct properties: galaxies are a high resolution but biased tracer of mass, while weak lensing avoids such biases but, due to low signal-to-noise ratio, has poor resolution. We investigate reconstructing the projected density field using the complementarity of weak lensing and galaxy positions. We propose a maximum-probability reconstruction of the 2D lensing convergence with a likelihood term for shear data and a prior on the Fourier phases constructed from the galaxy positions. By considering only the phases of the galaxy field, we evade the unknown value of the bias and allow it to be calibrated by lensing on a mode-by-mode basis. By applying this method to a realistic simulated galaxy shear catalogue, we find that a weak prior on phases provides a good quality reconstruction down to scales beyond $l = 1000$, far into the noise domain of the lensing signal alone.

Key words: gravitational lensing: weak – methods: data analysis – large-scale structure of Universe.

1 INTRODUCTION

Weak lensing is a promising cosmological probe, allowing the mass distribution in the Universe to be investigated without assumptions about the dynamics of the baryonic component.

In the pioneering work of Kaiser & Squires (1993), it has been shown that weak lensing can be used to map the distribution of dark matter in galaxy clusters. Following this, several methods for making so-called mass maps have been developed, with much attention given to reconstruction methods such as maximum-likelihood (ML) approaches (Bartelmann et al. 1996). However, there is a substantial level of noise in the resulting maps, due to the effect of galaxies having intrinsic ellipticities in addition to the sought-after gravitational shear. Therefore, it was immediately realized that the reconstruction methods require smoothing or regularization (Squires & Kaiser 1996). A significant proposal in this regard is the maximum-entropy (MaxEnt) method known from image reconstruction studies (Bridle

et al. 1998; Seitz, Schneider & Bartelmann 1998; Marshall et al. 2002).

These methods work well when applied to clusters, but the lensing ellipticity measurements of galaxies are still sufficiently noisy that reconstruction of the low-contrast large-scale structure is not possible with significant signal-to-noise (S/N). In this study, we develop a methodology attempting to make maps of the projected density with higher S/N, by utilizing a maximum-probability (MP) reconstruction with a physically motivated prior probability term: we will examine the usefulness of using Fourier phase information from the distribution of galaxies in the lensing map area. This is related to other recent methods that use galaxy positions to improve density reconstruction (Simon 2012) or combine weak lensing and galaxy positions to measure bias (Amara et al. 2012); in our case, we do not need to assume an amplitude for the bias.

The paper is organized as follows. In Section 2, we review the relevant theoretical background, including weak gravitational lensing quantities and the Fourier description of fields. We also emphasize the importance of Fourier phases in mapping cosmological fields. In Section 3, we introduce the MP method. We define the likelihood and the prior term for our reconstruction

★E-mail: rafal.szepietowski@manchester.ac.uk

method, describe the phase prior in detail, and outline the practical implementation of our method. Section 4 describes the simulated data set used in the analysis and the results of applying the reconstruction method. Finally, we discuss the implications of our work in Section 5.

2 THEORY

2.1 Lensing quantities

Here we briefly discuss the necessary lensing theory; full details can be found in e.g. Bartelmann & Schneider (2001) and Munshi et al. (2008).

The flat perturbed Friedman–Robertson–Walker metric of the standard cosmological model is

$$ds^2 = (1 + 2\Phi/c^2)dt^2 - a^2(t)(1 - 2\Phi/c^2) [dr^2 + r^2 d\Omega^2], \quad (1)$$

where Φ is the usual Newtonian gravitational potential and a is the scale factor. The potential is related to the matter density field by Poisson's equation

$$\nabla_{\text{com}}^2 \Phi = 4\pi G \bar{\rho} \delta a^2 = \frac{3}{2} H_0^2 \Omega_m \frac{\delta}{a}, \quad (2)$$

where $\delta = \rho/\bar{\rho} - 1$ describes the perturbation around the mean density of matter in the Universe.

In this space–time a lensing potential can be defined as

$$\phi(\boldsymbol{\theta}, r) \equiv \frac{2}{c^2} \int_0^r dr' \frac{r - r'}{r r'} \Phi(\boldsymbol{\theta}, r'), \quad (3)$$

where r is the comoving distance of the source and the integration is along the line of sight, and $\boldsymbol{\theta}$ is the position on the sky. This can be understood as a two-dimensional projection of the gravitational potential. The way in which an image of a source is distorted when passing through a gravitational field depends on a combination of the second-order derivatives of the lensing potential

$$\kappa = \frac{1}{2} (\partial_1^2 + \partial_2^2) \phi, \quad (4)$$

$$\gamma_1 = \frac{1}{2} (\partial_1^2 - \partial_2^2) \phi, \quad (5)$$

$$\gamma_2 = \partial_1 \partial_2 \phi, \quad (6)$$

where κ is called the convergence, γ_1 and γ_2 are the two components of the shear γ , and ∂_1, ∂_2 denote angular derivatives in the x and y directions, respectively. These quantities are found in the Jacobian matrix

$$\mathbf{A} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}, \quad (7)$$

which maps the source plane coordinates β_i to the image plane coordinates θ_j

$$\mathbf{A}_{ij} = \frac{\partial \beta_i}{\partial \theta_j}. \quad (8)$$

The convergence κ describes the projection of the overdensity field on the sky

$$\kappa(\boldsymbol{\theta}, r) = \frac{3H_0^2 \Omega_m}{2c^2} \int_0^r dr' \frac{r'(r - r')}{r} \frac{\delta(\boldsymbol{\theta}, r')}{a(r')}, \quad (9)$$

and this projected density is the quantity which we seek to reconstruct as a map.

2.2 Fourier description of fields

In our reconstruction method, we will use a prior term which involves the phase of lensing fields, so here we define the required quantities for this term. A real space field such as κ can be expanded in a Fourier superposition of plane waves:

$$\kappa(\boldsymbol{\theta}) = \sum \tilde{\kappa}(\mathbf{l}) \exp(i\mathbf{l} \cdot \boldsymbol{\theta}). \quad (10)$$

The Fourier transform $\tilde{\kappa}$ of such a field is complex and is described by an amplitude $|\tilde{\kappa}(\mathbf{l})|$ and phase α_l , where

$$\tilde{\kappa}(\mathbf{l}) = |\tilde{\kappa}(\mathbf{l})| \exp(i\alpha_l). \quad (11)$$

A Gaussian random field will have phases distributed independently¹ and uniformly on the interval $[-\pi, \pi)$. The statistical properties of the field are then fully specified by its power spectrum $P(l) = \langle |\tilde{\kappa}(\mathbf{l})|^2 \rangle_l$, where $\langle \rangle_l$ denotes an average over all modes at a wavenumber l .

However, the phase information contained in the κ field is interesting for two reasons:

(i) *Morphology*: in cases where one is interested in a specific realization of a density field, the phases describe features of its spatial pattern (Chiang 2001). For instance, one might be examining a region of the Universe where one wants to know the spatial distribution of matter, to understand the relationship between density and astrophysical properties (e.g. star formation).

(ii) *Non-Gaussianity*: due to primordial physics (e.g. Komatsu et al. 2009) and non-linear evolution on scales probed by weak lensing, the κ field will have non-zero higher order statistics beyond the power spectrum. This higher order information is encoded in a combination of phase and amplitude of the Fourier transformed field. If we can obtain a full estimate of phase and amplitude, we will be able to extract information about the growth of structure and the early Universe (Watts & Coles 2003; Chiang, Naselsky & Coles 2004).

3 METHOD

3.1 Maximum-probability reconstruction

Our reconstruction method seeks to find a hypothesis field which has the maximum probability of accounting for the observed data. We suppose that we have a data vector \mathbf{d} , which contains estimates of shear from observed galaxy ellipticities. We parametrize the hypothesis field by the values \mathbf{p} of projected density in a grid of pixels. The best-fitting set of parameters is then found by maximizing the posterior probability $P(\mathbf{p}|\mathbf{d}, M)$ according to Bayes' theorem

$$P(\mathbf{p}|\mathbf{d}, M) = \frac{L(\mathbf{d}|\mathbf{p}, M)P(\mathbf{p}|M)}{P(\mathbf{d}|M)} \propto L(\mathbf{d}|\mathbf{p}, M)P(\mathbf{p}|M), \quad (12)$$

where $L(\mathbf{d}|\mathbf{p}, M)$ is the likelihood and $P(\mathbf{p}|M)$ is the prior probability. The evidence $P(\mathbf{d}|M)$ is useful to compare various models M , whereas for a particular model M we can simply deal with the proportional term on the right-hand side. If we have no knowledge of how the parameters of the model should be distributed, we may assume that all values are equally likely a priori, i.e. the prior distribution is flat. Then, $P(\mathbf{p}|\mathbf{d}, M) \propto L(\mathbf{d}|\mathbf{p}, M)$ and the posterior distribution is found by maximizing the likelihood. This is the basis of ML methods.

¹ There is a caveat to this statement. For a real valued field κ , its Fourier modes have to satisfy the Hermitian relation $\tilde{\kappa}^*(l) = \tilde{\kappa}(-l)$.

However, the ML method (Bartelmann et al. 1996) will typically overfit the data by fitting the noise. Due to finite sampling of the shear field at galaxy positions, and further contamination of the signal by galaxy ellipticity noise, the reconstruction methods require smoothing or regularization (Squires & Kaiser 1996). We can consider two classes of prior which try to achieve this: *informative* and *uninformative* priors, differing in the assumptions which they make about the signal. If the purpose of introducing extra information is to regularize rather than inform an inference we can speak of a *weakly informative* prior.

Over the past two decades, different forms of regularization have been considered. An important example is the MaxEnt regularization known from image reconstruction (Bridle et al. 1998; Seitz et al. 1998; Marshall et al. 2002) which, while being an uninformative prior, benefitted from inferring information about the correlations in the data (Marshall et al. 2002). In addition, methods have been studied with informative priors; these make some assumptions about the nature of the signal, e.g. Wiener filtering (Hu & Keeton 2002; Simon, Taylor & Hartlap 2009; Simon et al. 2012). Here we will consider a MP approach with a weakly informative prior.

3.2 Likelihood

We would like to find a best-fitting hypothesized model for the convergence, κ , given a set of shear observations γ^d . In the flat sky approximation, we can relate the convergence and shear fields most easily in Fourier space (Kaiser & Squires 1993):

$$\tilde{\gamma}_1(\mathbf{l}) = \frac{l_1^2 - l_2^2}{l_1^2 + l_2^2} \tilde{\kappa}(\mathbf{l}), \quad (13)$$

$$\tilde{\gamma}_2(\mathbf{l}) = \frac{2l_1 l_2}{l_1^2 + l_2^2} \tilde{\kappa}(\mathbf{l}). \quad (14)$$

As the field of observations will be limited, a simple application of these transformations introduces edge effects, which we will mitigate by making reconstructions over larger patches than the data (see Section 3.4).

The data vector γ^d consists of estimates of the shear components γ_1 and γ_2 in each pixel of a 2D grid. These are obtained by averaging over galaxy ellipticities in each pixel, so that the error on the mean shear in a pixel is

$$\sigma_\gamma \approx \sigma_\varepsilon / \sqrt{n}, \quad (15)$$

where σ_ε is the intrinsic scatter of shear estimators for galaxies, and n the mean number of galaxies in a pixel. This error is approximately Gaussian by the central limit theorem.

If our hypothesized convergence field has corresponding shear pixel values γ_i^k , and the data shear pixel values are γ_i^d , then the likelihood for our hypothesized reconstruction is

$$L(\gamma^d | \kappa) \propto \prod_{i,j} \exp \left(- \frac{(\gamma_i^k - \gamma_i^d)^T \mathbf{C}_{ij}^{-1} (\gamma_j^k - \gamma_j^d)}{2} \right), \quad (16)$$

where \mathbf{C}^{-1} is the noise covariance matrix. Assuming the noise in each pixel is uncorrelated makes the covariance matrix diagonal and simplifies the likelihood to

$$L(\gamma^d | \kappa) \propto \prod_i \exp \left(- \frac{(\gamma_i^k - \gamma_i^d)^2}{2\sigma_\gamma^2} \right) = \exp \left(- \frac{\chi_\gamma^2}{2} \right). \quad (17)$$

This assumption is trivially true for shape noise, which dominates on all scales considered. However, intrinsic correlations between

galaxy shapes will introduce non-zero off-diagonal terms in the covariance matrix (Catelan, Kamionkowski & Blandford 2001; Hirata & Seljak 2004).

We turn now to consider the prior term for our MP reconstruction.

3.3 Phase prior

A prior term that accounts for the claim that galaxies trace mass, even if very poorly, can be achieved by constructing a prediction of the lensing convergence based on galaxy count overdensities

$$\delta_g(\boldsymbol{\theta}, z) = \frac{n_z(\boldsymbol{\theta})}{\bar{n}_z} - 1, \quad (18)$$

where $n_z(\boldsymbol{\theta})$ is the number density of galaxies at position $\boldsymbol{\theta}$ and \bar{n}_z is the mean number density of galaxies at redshift z . We could suppose that the overall matter overdensity $\delta \simeq b^{-1} \delta_g$, where b is the galaxy bias. Then, we can project δ according to equation (9) to find the count-estimated convergence κ_g . For a sample divided into N_z redshift bins the projection becomes

$$\kappa_g(\boldsymbol{\theta}, z) = \frac{3H_0\Omega_m}{2c^2} \sum_{i=1}^{N_z} \Delta r_i \frac{r(z_i)[r(z) - r(z_i)]}{r(z)} \frac{\delta(\boldsymbol{\theta}, z_i)}{a(z_i)}, \quad (19)$$

where $\Delta r_i = r(z_i) - r(z_{i-1})$. It would then be possible to require that the hypothesized final convergence field is close to this κ_g , within some tolerance.

However, there is a problem with this approach: the bias b is unknown, and the claim of linear bias introduces another assumption into the reconstruction.

An easy way of avoiding this problem is to consider only the information about the phases of the Fourier modes of κ_g , neglecting their amplitudes. Fig. 1 shows the relation between the phases of the true convergence κ and count convergence κ_g found in DES mock catalogue v4.02 (see Section 4.1).

As expected for a close-to-Gaussian field, the histograms of phases for both κ and κ_g fields are close to uniform in the range $[-\pi, \pi)$. However, the overlaid histogram of the phase difference $\Delta\alpha = \alpha^\kappa - \alpha^{\kappa_g}$ between the true κ and κ_g is visibly spiked around $\Delta\alpha = 0$, indicating a strong correlation between the phases of the two fields. We now discuss how this phase difference is calculated in detail.

3.3.1 Phase difference distribution

As the phases are distributed on the interval $[-\pi, \pi)$ their differences will have values on the interval $(-\pi, \pi)$. However, since the phases are a cyclic quantity, absolute phase difference $|\Delta\alpha| > \pi$ will correspond to a phase difference smaller than π . This is easily accounted for: if $\Delta\alpha$ is less than $-\pi$, we add 2π to $\Delta\alpha$; if $\Delta\alpha$ is greater than or equal to π , then we subtract 2π from $\Delta\alpha$.

We can construct the correlation matrix for the phase difference between true convergence phase and galaxy-count derived convergence phase. In our simulations (Section 4.1), this is constructed from 36 different $2^\circ \times 2^\circ$ areas including κ and κ_g information, as for each area only one galaxy distribution realization is available. By the ergodic principle, this should give an estimate of how much the phases usually differ between the density and galaxy fields in an area. We find that the correlation matrix constructed for 2 arcmin \times 2 arcmin pixels is strongly diagonal with the median absolute value of the correlation coefficient $\simeq 0.06$.

The histograms of $\Delta\alpha$ for the whole field (Fig. 1) as well as for individual wavenumbers (Fig. 2) are well fitted by a wrapped

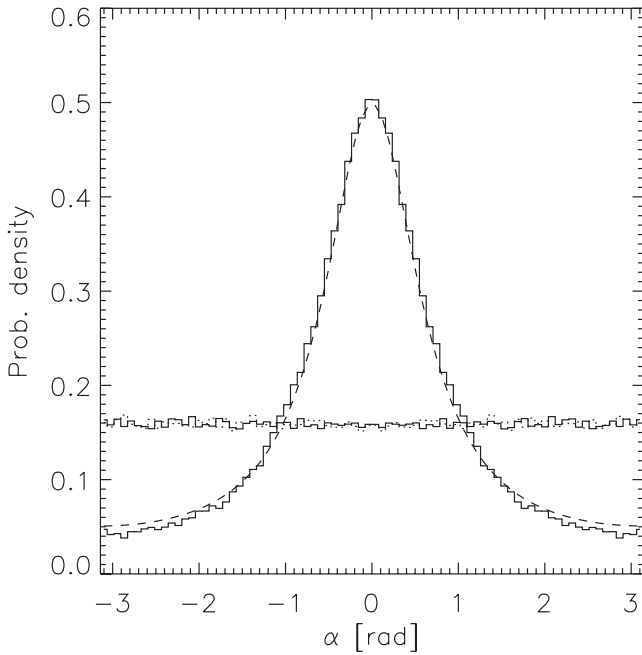


Figure 1. Distribution of convergence Fourier phases and their difference. Histogram of the phases α of the true convergence field κ (solid line) and those obtained from the galaxy distribution κ_g (dotted line) for all wavenumbers. The distributions are close to uniform, as expected for fields which have a distribution close to that of a Gaussian random field. Overlaid (peaked curves), the histogram of the phase difference $\Delta\alpha$ between the true convergence κ and the approximation κ_g (solid line), for all wavenumbers. The distribution is well approximated by a wrapped Cauchy distribution (dashed line). We see a strong correlation between the phases of the two fields.

Cauchy probability distribution function:

$$P_{\text{prior}}(\kappa|\alpha_{\text{gal}}) \propto \prod_i \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\Delta\alpha_i)}. \quad (20)$$

We note that the distribution is symmetric around zero. The parameter describing the width of the distribution is $\rho = e^{-\sigma_\alpha}$, where σ_α is the half-width of an unwrapped Cauchy distribution. For small values, σ_α can be estimated using the median absolute deviation (MAD)

$$\sigma_\alpha \approx 1.1 \cdot \text{MAD}_{\Delta\alpha}. \quad (21)$$

We provide further details on this distribution in Appendix A. However, we want to use the phase information as a weakly informative prior, so we are free to relax this width; we will allow more tolerance in phase difference between our reconstructed κ and the κ_g field by choosing $\sigma_\alpha = 2.2 \cdot \text{MAD}_{\Delta\alpha}$. Using $\sigma_\alpha = 1.1 \cdot \text{MAD}_{\Delta\alpha}$ would take us in the direction of a *joint* reconstruction of the density field from shear and galaxy position data, which is also of interest; some of our runs in Section 4.2 explore this possibility.

It is to be expected that σ_α will be a function of l , with the phase differences between galaxies and dark matter for large-scale modes being more constrained than for small-scale ones. We indeed find this to be the case in our simulations, as shown in Fig. 3. The phase difference distribution for each l also follows a wrapped Cauchy distribution. This distribution is naturally generated when the difference between κ and κ_g comes from a white noise contribution, such as shot-noise, and possibly a further contribution from the stochasticity of the bias relation (Dekel & Lahav 1999; Manera & Gaztañaga 2011). Hence, the low- l modes have smaller phase dif-

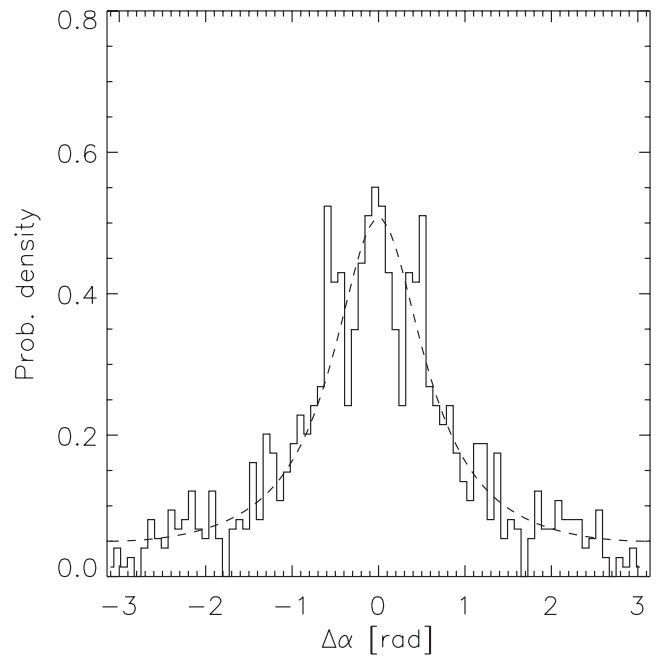
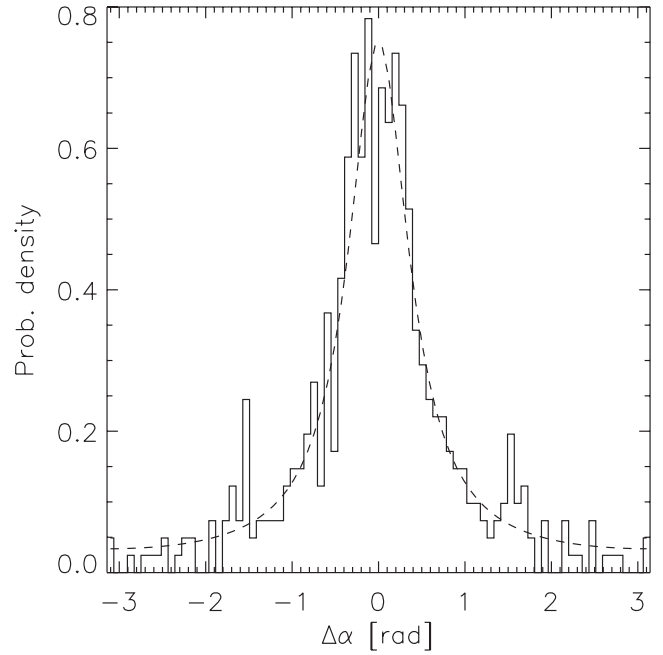


Figure 2. The histogram of the phase difference $\Delta\alpha$ between the true convergence κ and the approximation κ_g (solid lines) at $l = 1200$ (top) and $l = 2250$ (bottom). The distributions are well approximated by a wrapped Cauchy distribution (dashed lines).

ferences, as this white noise offset is smaller as a proportion of the signal on these scales.

In the mock catalogue, the galaxy biasing is roughly linear and deterministic. It could be that the wrapped Cauchy pdf of the phase differences is typical only for this type of bias, but might be quite different for more complex scenarios. Hence, further studies of how the phase difference distribution arises are important. However, as we permit very large errors on the phase difference, moderate

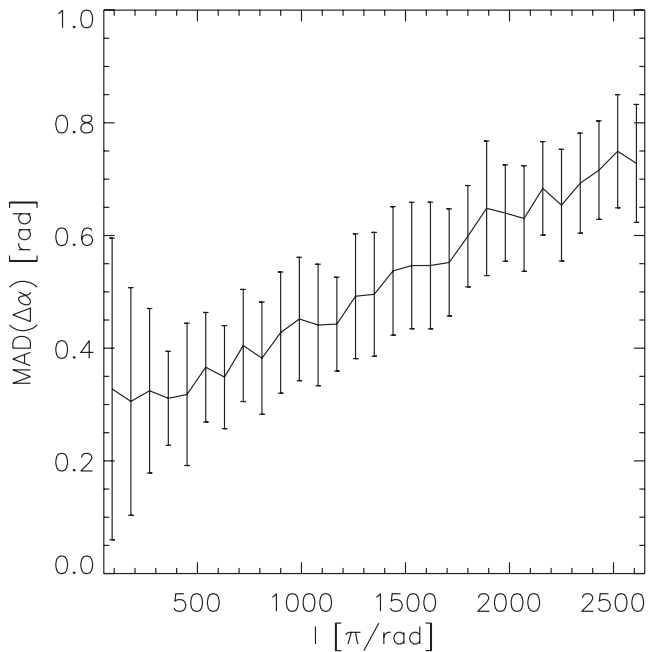


Figure 3. MAD of the phase difference $\Delta\alpha$ between the true convergence κ and the approximation obtained from the galaxy distribution κ_g as a function of l . The solid line shows the mean $\text{MAD}(\Delta\alpha)$ of the phase difference obtained in shells of radius l from the origin with error bars showing the standard deviation, across the 36 simulated fields.

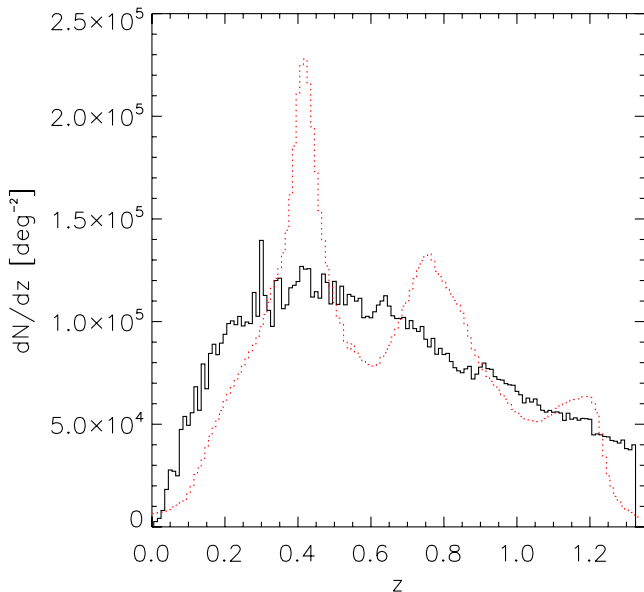


Figure 4. Distribution of galaxies with redshift. True redshift (black solid) and photometric redshifts obtained using the ANNz code (red dotted).

deviations from our simulations’ bias model should not change the conclusions of the paper.

In reality, the estimation of δ_g will suffer from systematics originating, for example, from an inhomogeneous galaxy survey. These could be mitigated by methods used for the matter power spectrum estimation, where pixels are reweighted to account for the mask (Feldman, Kaiser & Peacock 1994; Percival, Verde & Peacock 2004). A further systematic will arise from using photometric redshifts to estimate distances (Fig. 4). However, this will be

mitigated by the fact the convergence is projected; nevertheless, careful tests of this systematic will be necessary.

3.4 Practical implementation

We are now ready to discuss our approach to finding a reconstructed convergence field. Rather than estimating the posterior distribution of our convergence hypotheses, we will seek a *maximum a posteriori* (MAP) solution. The reconstruction is performed by seeking a $\tilde{\kappa}^{\text{trial}}$ that maximizes the posterior probability. The posterior pdf will be generally strongly peaked so it is convenient to work with its logarithm

$$-\ln P(\tilde{\kappa}|\gamma^d, \alpha_{\text{gal}}) \propto -\ln L - \ln P_{\alpha}^{\text{prior}}, \quad (22)$$

which varies more slowly with the change in $\tilde{\kappa}$.

As the shape of the posterior pdf is generally unknown, we use a simple heuristic optimizer. We use the idea of simulated annealing (Kirkpatrick, Gelatt & Vecchi 1983), but replace the usual Metropolis–Hastings sampler (Metropolis et al. 1953; Hastings 1970) with a Multi Try Metropolis (Liu, Liang & Wong 2000) one. In each step t a set of trial convergence fields $\{\tilde{\kappa}_i^{\text{trial}}\}$ is generated from the current field

$$\tilde{\kappa}_i^{\text{trial}} = \tilde{\kappa}^{\text{current}} + \delta\tilde{\kappa}_i, \quad (23)$$

where components of each $\delta\tilde{\kappa}_i$ are drawn from normal distribution $\mathcal{N}(0, \sigma_t\sqrt{P(l)})$, where the scaling $P(l)$ is proportional to the expected signal (see below). A proposal field $\tilde{\kappa}^{\text{propos}}$ is then chosen. To limit the random walk behaviour, the field with the highest probability different from the current one is chosen. Then, a reference set $\{\kappa_j^{\text{ref}}\}$ that includes $\tilde{\kappa}^{\text{current}}$ is formed from that field. The proposal field is then accepted with the probability

$$P(\tilde{\kappa}^{\text{propos}} | \{\kappa_j^{\text{ref}}\}) = 1 \quad \text{for} \quad \frac{\sum_j P(\tilde{\kappa}_j^{\text{ref}})}{\sum_i P(\tilde{\kappa}_i^{\text{trial}})} \geq 1, \quad (24)$$

$$P(\tilde{\kappa}^{\text{propos}} | \{\kappa_j^{\text{ref}}\}) = T_t \frac{\sum_j P(\tilde{\kappa}_j^{\text{ref}})}{\sum_i P(\tilde{\kappa}_i^{\text{trial}})} \quad \text{otherwise.} \quad (25)$$

In addition to a cooling schedule for the acceptance rate

$$T_t = \frac{T_0}{\log_{10}(t + 10)}, \quad (26)$$

we have added a similar schedule to decrease the step size in the sampling algorithm

$$\sigma_t = \frac{\sigma_0}{\log_{10}(t^2 + 10)}, \quad (27)$$

to allow for more refined changes as the optimizer gets closer to the solution we seek (Elson et al. 2007; Kotze 2009). The solution with the highest probability κ_{best} is stored and used as the output of the optimizer.

Operations on the fields, such as calculating the shears from the convergence, are performed in Fourier space; hence, edge effects such as periodic boundaries of the reconstruction will be present. This would mean that the largest scales would not be recovered accurately. This is partially solved by introducing a larger reconstruction grid as suggested in Bridle et al. (1998) and here we use a grid four times bigger than the reconstruction area.

To aid the optimization process, we choose a starting position for our hypothesis which is expected to be close to the MAP solution. The initial guess for the reconstruction, $\tilde{\kappa}^{\text{initial}}$, is a field fully consistent with the prior; that is, we choose phases from the galaxy

convergence map. We also apply a power spectrum filter to the $\tilde{\kappa}_g$ field

$$\tilde{\kappa}^{\text{initial}}(l) = \tilde{\kappa}_g(l) \sqrt{\frac{P(l)}{P_g(l)}}, \quad (28)$$

which gives the κ_g field the required amplitude of power spectrum and suppresses the high- l noise. As this is only a starting guess, any $P(l)$ with a very approximately correct shape and amplitude should suffice. Here, we choose the true average κ power spectrum from simulations. By choosing this starting point, the optimizer evolves the reconstruction from the prior to the posterior under the influence of lensing.

However, to check for possible local maxima in the posterior, we also try running the code from a noisy position such as κ_g without applying any filters.

4 APPLICATION TO SIMULATED DATA

4.1 Simulated galaxy catalogue

For this study, we have used the mock galaxy catalogues created for the Dark Energy Survey based on the algorithm Adding Density Determined GALaxies to Lightcone Simulations (ADDGALS; Wechsler et al., in preparation; Busha et al., in preparation). This algorithm attaches synthetic galaxies, including multiband photometry, to dark matter particles in a lightcone output from a dark matter N -body simulation and is designed to match the luminosities, colours and clustering properties of galaxies. The catalogue used here was based on a single ‘Carmen’ simulation run as part of the LasDamas of simulations (McBride et al., in preparation).² This simulation modelled a flat Λ cold dark matter universe with $\Omega_m = 0.25$ and $\sigma_8 = 0.8$ in a 1 Gpc h^{-1} box with 1120³ particles. A 220 deg² light cone extending out to $z = 1.33$ was created by pasting together 40 snapshot outputs.

The galaxy distribution for this mock catalogue was created by first using an input luminosity function to generate a list of galaxies, and then adding the galaxies to the dark matter simulation using an empirically measured relationship between a galaxy’s magnitude, redshift and local dark matter density, $P(M_r, z|\delta_{\text{dm}})$ – the probability that a galaxy with magnitude M_r and redshift z resides in a region with local density δ_{dm} . This relation was tuned using a high-resolution simulation combined with the SubHalo Abundance Matching technique that has been shown to reproduce the observed galaxy two-point function to high accuracy (Kravtsov et al. 2004; Conroy, Wechsler & Kravtsov 2006; Reddick et al. 2013).

For the galaxy assignment algorithm, we choose a luminosity function that is similar to the SDSS luminosity function as measured in Blanton et al. (2003), but evolves in such a way as to reproduce the higher redshift observations (e.g. SDSS-Stripe 82, AGES, GAMA, NDWFS and DEEP2). In particular, ϕ_* and M_r are varied as a function of redshift in accordance with the recent results from GAMA (Loveday et al. 2012).

Once the galaxy positions have been assigned, photometric properties are added. Here, we use a training set of spectroscopic galaxies taken from SDSS DR5. For each galaxy in both the training set and simulation we measure Δ_5 , the distance to the fifth nearest galaxy on the sky in a redshift bin. Each simulated galaxy is then assigned an SED based on drawing a random training-set galaxy with the

appropriate magnitude and local density, k -correcting to the appropriate redshift, and projecting on to the desired filters. When doing the colour assignment, the likelihood of assigning a red or a blue galaxy is smoothly varied as a function of redshift in order to simultaneously reproduce the observed red fraction at low and high redshifts as observed in SDSS and DEEP2.

For the simulation of gravitational lensing, weak lensing shear at each galaxy position was computed using the multiple plane ray tracing code CALCLENS (Becker 2013). Then, an intrinsic ellipticity is assigned to each galaxy. The intrinsic shape distribution and dispersion σ_ϵ in these simulations are magnitude dependent and are modelled after those found in deep SuprimeCam i' -band data with excellent seeing (0.6 arcsec), with fainter galaxies having a higher intrinsic ellipticity dispersion. Averaged over all galaxies $\sigma_\epsilon = 0.4$.

4.2 Results

From the simulated catalogue described in Section 4.1, we select a large square patch of 12° × 12°. To study the behaviour of the reconstructions, 100 areas (with replacement) of 2° × 2° were randomly selected from this patch. These were divided into pixels of 2 arcmin × 2 arcmin containing $\simeq 116$ galaxies. Hence, the number density of sources is 29 gal arcmin⁻². We use the same galaxies as sources and tracers of the density field.

The reconstruction code was run for 30 000 trial steps for each sub-field, with 300 trial fields generated in each optimization step. The reconstructed maps span 4° × 4°, containing 14 400 pixels of 2 arcmin × 2 arcmin, i.e. we reconstruct a larger patch than the 2° × 2° data patch in each case.

The reconstructions were performed for each of the 100 fields using different phase distribution parameters and initial guesses that are summarized in Table 1. Using 100 different fields allowed us to examine the noise properties of the reconstruction method.

Reconstructions were performed using an ML method (i.e. no prior terms) and the MP approach with the phase prior. In this set of runs, the phase prior included a phase tolerance $\sigma_\alpha = 2.2 \cdot \text{MAD}_{\Delta\alpha}$ in order to provide a weakly informative prior. To obtain a reasonable starting point, $\tilde{\kappa}^{\text{initial}}$ was filtered according to equation (28).

Fig. 5 shows examples of maps obtained using both methods of reconstruction (b and c) with the true simulated convergence map (a) and the convergence estimated from galaxy positions (d) shown for comparison (using $\delta = \delta_g$, i.e. $b = 1$, see Section 3.3). The ML method reconstructs only the most prominent peaks, with a high level of contamination by spurious peaks. The inclusion of the phases prior appears to improve the map considerably, but it also maps features from κ_g that are not necessarily present in the true convergence, e.g. RA = 40 arcmin, $\delta = 115$ arcmin. However, these are consistent with the lensing only reconstruction.

Table 1. List of the reconstructions carried out, with different combinations of priors, phase distribution parameters and initial reconstruction hypothesis.

Posterior	Phases tolerance	κ^{initial}
$L(\gamma^{\text{d}} \kappa)$	–	Filt.
$L(\gamma^{\text{d}} \kappa)P(\kappa \alpha_{\text{gal}})$	$\sigma_\alpha = 1.1 \cdot \text{MAD}_{\Delta\alpha}$	Filt.
$L(\gamma^{\text{d}} \kappa)P(\kappa \alpha_{\text{gal}})$	$\sigma_\alpha = 2.2 \cdot \text{MAD}_{\Delta\alpha}$	Filt.
$L(\gamma^{\text{d}} \kappa)P(\kappa \alpha_{\text{gal}})$	$\sigma_\alpha = 2.2 \cdot \text{MAD}_{\Delta\alpha}$	Noisy
$L(\gamma^{\text{d}} \kappa)P(\kappa \alpha_{\text{gal}})$	$\sigma_\alpha = 3.3 \cdot \text{MAD}_{\Delta\alpha}$	Filt.

² Further details regarding the simulations can be found at <http://lss.phy.vanderbilt.edu/lasdamas/simulations.html>

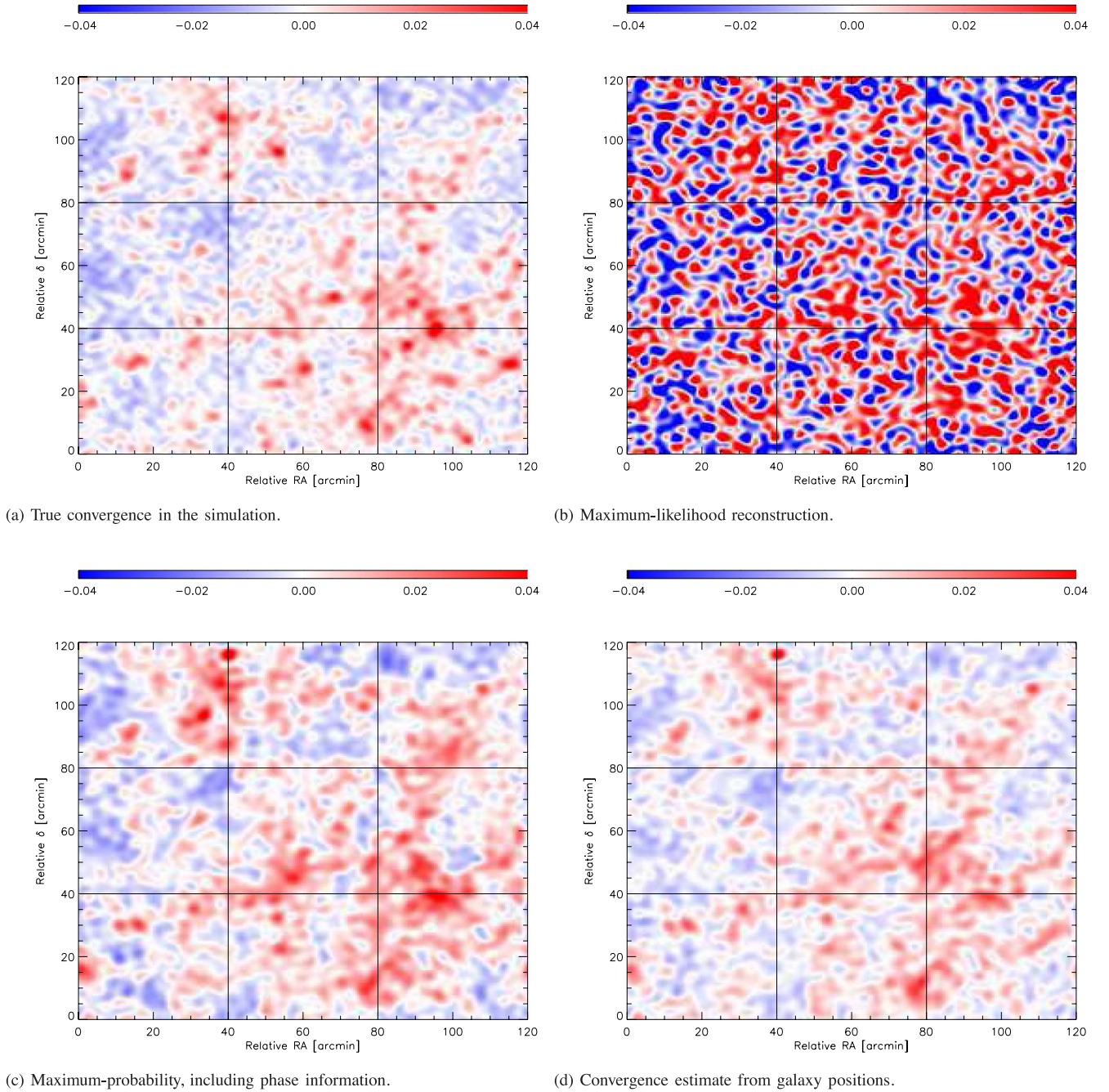


Figure 5. Resulting reconstructed maps of the convergence field. The maps are showing an example of a $2^\circ \times 2^\circ$ reconstruction field with $2 \text{ arcmin} \times 2 \text{ arcmin}$ pixels. The maps were zero-padded in Fourier space to have a smoother appearance. The true convergence is shown along with reconstructions obtained using the ML method and the MP method with the phase prior. The galaxy convergence κ_g from which the prior was computed is also shown for comparison. (a) True convergence in the simulation. (b) ML reconstruction. (c) MP, including phase information. (d) Convergence estimate from galaxy positions.

To quantify the quality of the reconstruction, we construct a power spectrum of the error per mode in the reconstruction,

$$P_{\text{err}}(l) = \langle |\tilde{\kappa}_l^{\text{rec}} - \tilde{\kappa}_l^{\text{true}}|^2 \rangle_l. \quad (29)$$

A faithful reconstruction will have small $P_{\text{err}}(l)$, preferably smaller than the true power in order to achieve good S/N (i.e. the errors in the reconstruction are preferably smaller than the signal of the reconstructed structures for a given scale). $P_{\text{err}}(l)$ shows the scale dependence of the reconstruction faithfulness. However, it is not

intended as a metric of how well we can reconstruct the power spectrum from the maps.

Fig. 6 shows the power spectra (dashed) and error power spectra (dotted) of the reconstruction averaged over 100 fields. The ML reconstruction (red) is dominated by noise on most scales. Including the phase prior (blue) leads to a reconstruction that has higher S/N than the ML reconstruction on all scales, and has S/N > 1 even beyond $l = 1000$, far into the domain where the initial shear data is noise-dominated. On a pixel by pixel basis, the phase prior improves

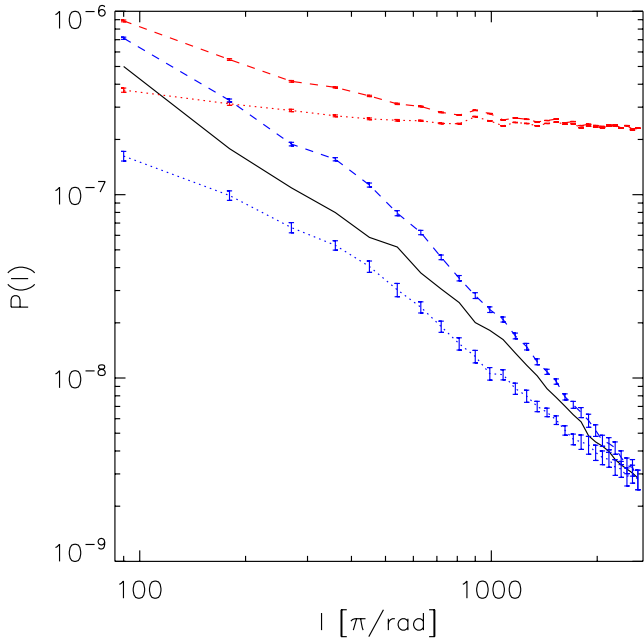


Figure 6. Power spectra (dashed) and error power spectra (dotted) for the reconstructions. The true convergence power spectrum (black solid line) is plotted for comparison. Red: ML approach. Blue: MP reconstruction including the phase prior. The reconstructions including the phase prior have $S/N > 1$ even beyond $l = 1000$, far into the domain where the shear data is noise-dominated.

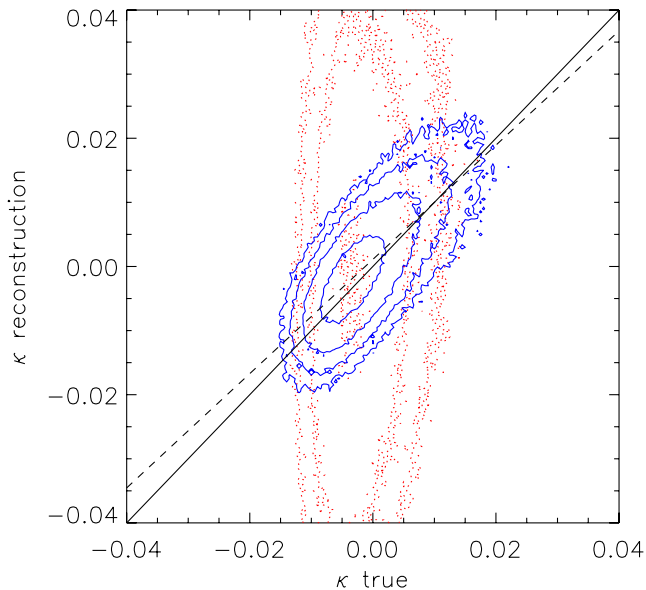


Figure 7. The contours for a 2D histogram of pixels in the reconstruction versus pixels in the true convergence. Contours are for 10, $10^{1.5}$, 10^2 , $10^{2.5}$ values, and the histogram shows a concatenation of reconstructions for 100 different fields. Results are shown with phase prior (blue solid) and ML approaches (red dotted). The best-fitting line to the phase reconstruction contours (black dashed) has a gradient of 0.89 and offset of 0.001.

the correlation between the true convergence and the reconstruction as shown in Fig. 7. The Pearson correlation coefficient changes from 0.22 for the ML reconstruction to 0.72 in the case of the MP reconstruction.

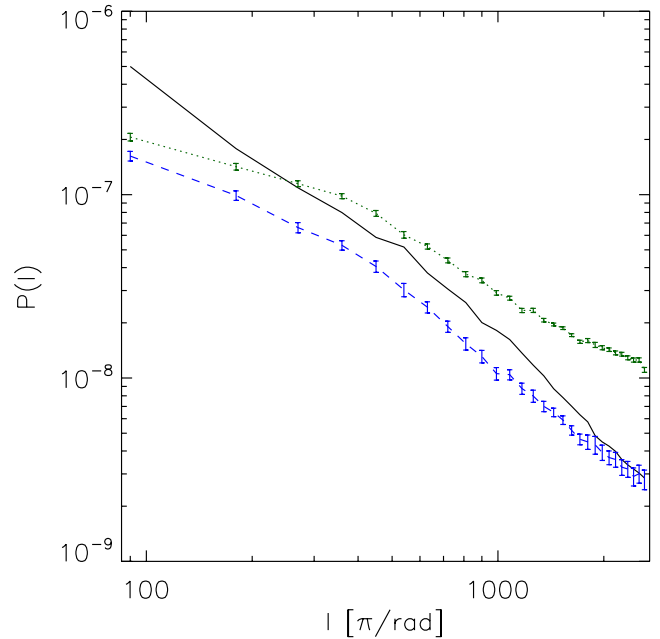


Figure 8. Dependence on the starting position. The true convergence power spectrum (black solid line) is plotted for comparison. We show the error power for a reconstruction starting from a κ^{initial} filtered according to equation (28) (blue dashed) and an unfiltered one (green dotted).

The reduction of the noise visible in Fig. 6 is due to the interplay between the galaxy phases and both the phase and amplitude of the lensing. Given noisy shear data, and if the phases of the two fields disagree strongly, the only permitted hypothesis that satisfies both the phase prior and the likelihood with modest probability, has low amplitude for the signal. On the other hand, where the phases agree, a higher amplitude is permitted.

To assess the errors on curves in Fig. 6, an additional 100 runs different starting points were performed on a single $2^\circ \times 2^\circ$ field, to see the variation in reconstructions permitted by the optimizer. The different $\kappa_i^{\text{initial}}$ fields were generated by multiplying each mode in $\tilde{\kappa}_g$ by a complex random number with each component drawn from a standard normal distribution $\mathcal{N}(0, 1)$. The error bars on different power spectra in Figs 6, 8 and 9 show the standard deviation in error powers of this set of runs. We see that these errors are substantially smaller than the variation between the ML and MP runs (Fig. 6) and also between MP runs with different values of the σ_α parameter (Fig. 9).

To check the dependence of the reconstruction on the initial guess κ^{initial} , further reconstructions with the phase prior were performed. The phase tolerance was again set to $\sigma_\alpha = 2.2 \cdot \text{MAD}_{\Delta\alpha}$ but κ^{initial} was left unfiltered. Fig. 8 shows the errors on these reconstruction compared to the analogous filtered one. The reconstruction with an unfiltered starting guess (green dotted) deviates more from the reconstruction with a filtered one (blue dashed) on small scales, $l \gtrsim 1000$ suggesting that the posterior probability surface is very flat in some directions (or multimodal). Although, the difference is visible on all scales, the reconstruction remains a substantial improvement over the ML reconstruction in Fig. 6.

The tolerance we permit on the phases has a moderate impact on the reconstruction, as shown in Fig. 9. The lines show error power spectra for reconstruction with phase tolerance of $\sigma_\alpha = 1.1 \cdot \text{MAD}_{\Delta\alpha}$ (red dotted), $\sigma_\alpha = 2.2 \cdot \text{MAD}_{\Delta\alpha}$ (green dashed) and $\sigma_\alpha = 3.3 \cdot \text{MAD}_{\Delta\alpha}$ (blue dot-dashed), and the error power grows

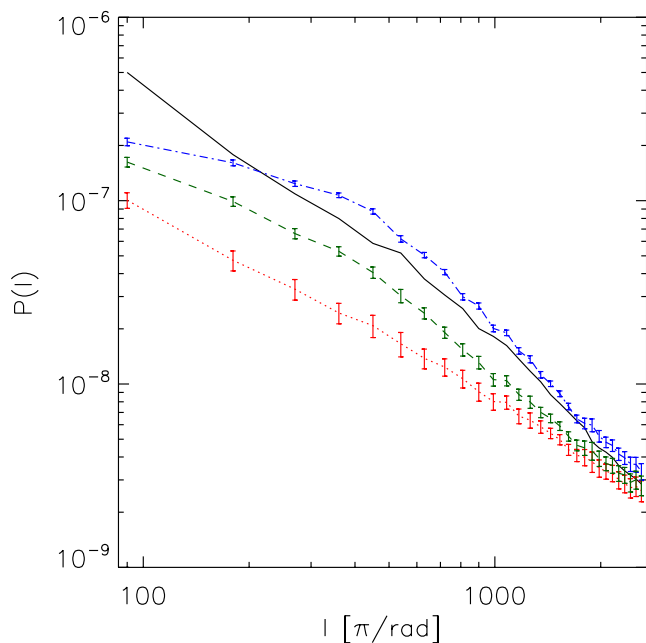


Figure 9. Dependence on the phase tolerance. The true convergence power spectrum (black solid line) is plotted for comparison. The lines show errors for reconstructions with phase tolerance of $\sigma_\alpha = 1.1 \cdot \text{MAD}_{\Delta\alpha}$ (red dotted), $\sigma_\alpha = 2.2 \cdot \text{MAD}_{\Delta\alpha}$ (green dashed) and $\sigma_\alpha = 3.3 \cdot \text{MAD}_{\Delta\alpha}$ (blue dot-dashed).

by a factor of 2 on intermediate scales between the tightest and weakest of these tolerances. However, independent of the phase tolerance the reconstructions are similar on small scales where the reconstruction is noise dominated, and on the largest scales where the likelihood term is large.

5 CONCLUSIONS

In this paper, we have proposed an MP reconstruction method for the lensing convergence, and have studied the impact of a physically motivated prior term.

To put a weakly informative prior on the Fourier phases of the modes, we made a prediction of the convergence from the galaxy number overdensity, and used this to inform the preferred phases of the reconstructed convergence field. In this way, by using only the phases of this field, we avoid the use of the unknown amplitude of the linear galaxy bias. We also do not require a deterministic bias, as we allow a phase deviation between the galaxy distribution and the underlying matter density.

By implementing and testing this method with a realistic simulated galaxy shear catalogue, we have found that a weak prior on phases provides a good quality 2D density reconstruction with $S/N \geq 1$ on scales up to and beyond $l = 1000$ (Fig. 6).

The sensitivity of the phase prior reconstruction to initial conditions (Fig. 8) shows that the probability surface is flat in directions associated with noise-dominated modes, as expected. However, an approximate knowledge of the power spectrum can help to select a solution with modest S/N even on the smallest scales. The phase difference tolerance can be made more or less strict, depending on whether one wishes to make a joint reconstruction using weak lensing and phases, or instead to make a reconstruction from weak lensing weakly informed by phases. In either case, the reconstruc-

tion is found to be an improvement over ML reconstruction (contrast Figs 9 and 6).

Although, most of the phase information is coming from the galaxy field, the amplitude of the modes is determined by the interplay between these and the lensing, which includes both phase and amplitude information. It is important to emphasize that in Fig. 5(d) the amplitude is an assumption, whereas in Fig. 5(c) it is derived purely from data.

In summary, using the phase information from the galaxy distribution to inform weak lensing density reconstruction, appears to be a very powerful addition to the tools we can use for mass mapping. As these maps combine information from the weak lensing and galaxy fields, they can potentially be used to improve our understanding of the relation between dark matter and galaxies, i.e. the bias.

ACKNOWLEDGEMENTS

We thank Bruce Bassett, Mathew Becker, Rob Crittenden, Alan Heavens and Phil Marshall for useful discussions. RS also thanks B. Bassett for organizing the Cape Town Cosmology School 2012 which inspired some of the ideas presented in this paper.

This work was partially supported by STFC grant ST/K00090X/1 and a Royal Society-NRF International Exchange Grant. RS acknowledges support from STFC in the form of a Research Studentship. PM is supported by the US Department of Energy under Contract no. DE-FG02-91ER40690.

Please contact the authors to request access to research materials discussed in this paper.

REFERENCES

- Amara A. et al., 2012, MNRAS, 424, 553
- Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
- Bartelmann M., Narayan R., Seitz S., Schneider P., 1996, ApJ, 464, L115
- Becker M. R., 2013, MNRAS, 435, 115
- Blanton M. R. et al., 2003, ApJ, 592, 819
- Bridle S. L., Hobson M. P., Lasenby A. N., Saunders R., 1998, MNRAS, 299, 895
- Catelan P., Kamionkowski M., Blandford R. D., 2001, MNRAS, 320, L7
- Chiang L.-Y., 2001, MNRAS, 325, 405
- Chiang L.-Y., Naselsky P. D., Coles P., 2004, ApJ, 602, L1
- Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
- Dekel A., Lahav O., 1999, ApJ, 520, 24
- Elson E. C., Bassett B. A., van der Heyden K., Vilakazi Z. Z., 2007, A&A, 464, 1167
- Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23
- Hastings W. K., 1970, Biometrika, 57, 97
- Hirata C. M., Seljak U., 2004, Phys. Rev. D, 70, 063526
- Hu W., Keeton C. R., 2002, Phys. Rev. D, 66, 063506
- Kaiser N., Squires G., 1993, ApJ, 404, 441
- Kirkpatrick S., Gelatt C. D., Vecchi M. P., 1983, Science, 220, 671
- Komatsu E. et al., 2009, preprint (arXiv:0902.4759)
- Kotze J., 2009, PhD thesis, Univ. Cape Town
- Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35
- Liu J. S., Liang F., Wong W. H., 2000, J. Am. Stat. Assoc., 95, 121
- Loveday J. et al., 2012, MNRAS, 420, 1239
- Manera M., Gaztañaga E., 2011, MNRAS, 415, 383
- Marshall P. J., Hobson M. P., Gull S. F., Bridle S. L., 2002, MNRAS, 335, 1037

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, *J. Chem. Phys.*, 21, 1087
 Munshi D., Valageas P., van Waerbeke L., Heavens A., 2008, *Phys. Rep.*, 462, 67
 Percival W. J., Verde L., Peacock J. A., 2004, *MNRAS*, 347, 645
 Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, *ApJ*, 771, 30
 Seitz S., Schneider P., Bartelmann M., 1998, *A&A*, 337, 325
 Simon P., 2012, *A&A*, 560, 15
 Simon P., Taylor A. N., Hartlap J., 2009, *MNRAS*, 399, 48
 Simon P. et al., 2012, *MNRAS*, 419, 998
 Squires G., Kaiser N., 1996, *ApJ*, 473, 65
 Watts P., Coles P., 2003, *MNRAS*, 338, 806

APPENDIX A: WRAPPED CAUCHY DISTRIBUTION

The Cauchy pdf is given by

$$f_C(x; x_0, \sigma) = \frac{1}{\pi} \cdot \frac{\sigma}{\sigma^2 + (x - x_0)^2}, \quad x \in (-\infty, \infty). \quad (\text{A1})$$

The Wrapped Cauchy pdf is defined as

$$f_{WC}(\beta; \beta_0, \gamma) = \sum_{n=-\infty}^{\infty} \frac{\sigma}{\pi(\sigma^2 + (x - x_0 + 2\pi n)^2)}, \quad (\text{A2})$$

which gives

$$f_{WC}(\beta; \beta_0, \gamma) = \frac{1}{2\pi} \cdot \frac{1 - \gamma^2}{1 + \gamma^2 + 2\gamma \cos(\beta - \beta_0)}, \quad (\text{A3})$$

where $\gamma = e^{-\sigma}$ and $\beta \in [-\pi, \pi)$.

A Cauchy distributed random number x can be generated from two independent normally distributed random numbers $y_1, y_2 \sim \mathcal{N}(0, 1)$ as

$$x = x_0 + \sigma \frac{y_1}{y_2}. \quad (\text{A4})$$

Then, a wrapped Cauchy distributed random number is obtained by taking

$$\beta = x \bmod 2\pi, \quad (\text{A5})$$

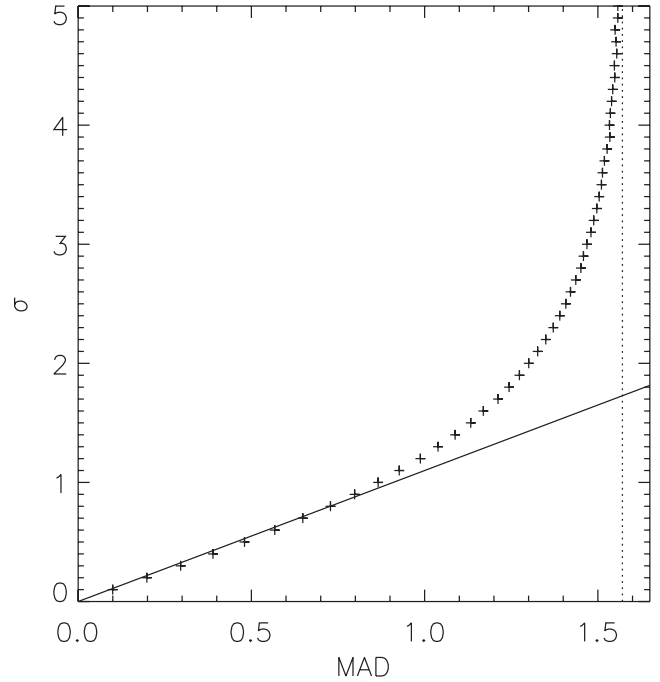


Figure A1. MAD as an estimator of σ_α . Assuming σ_α can be estimated as $1.1 \cdot \text{MAD}_{\Delta\alpha}$ (solid line) is justified for values of $\sigma_\alpha \lesssim 1$ (see Fig. 3). For larger values, $\text{MAD}_{\Delta\alpha}$ will tend to a constant (here $\pi/2$).

and applying a procedure similar to the one in Section 3.3.1, i.e. if β is less than $-\pi$, we add 2π to β ; if β is greater than or equal to π , then we subtract 2π from β .

For a distribution with $\beta_0 = 0$ the parameter σ_α can be approximated (for small values) as $1.1 \cdot \text{MAD}_{\Delta\alpha}$ (Fig. A1). For high values of σ_α this approximation breaks down; as $\sigma_\alpha \rightarrow \infty$ the wrapped Cauchy tends to a uniform distribution, and $\text{MAD}_{\Delta\alpha}$ goes to a constant equal to the standard deviation of the uniform distribution, here $\pi/2$ (see Fig. A1).

This paper has been typeset from a \TeX/L\AA\TeX file prepared by the author.