Improving Weak Lensing Mass Map Reconstructions using Gaussian and Sparsity Priors: Application to DES SV

N. Jeffrey,[1]⋆ F. B. Abdalla,[1,2] O. Lahav,[1] F. Lanusse,[3] J.-L. Starck,[4] A. Leonard,[1] D. Kirk,[1] C. Chang,[5] E. Baxter,[6] T. Kacprzak,[7] S. Seitz,[8,9] V. Vikram,[10] L. Whiteway,[1] T. M. C. Abbott,[11] S. Allam,[12] S. Avila,[13,14] E. Bertin,[15,16] D. Brooks,[1] A. Carnero Rosell,[17,18] M. Carrasco Kind,[19,20] J. Carretero,[21] F. J. Castander,[22,23] M. Crocce,[22,23] C. E. Cunha,[24] C. B. D'Andrea,[6] L. N. da Costa,[17,18] C. Davis,[24] J. De Vicente,[25] S. Desai,[26] P. Doel,[1] T. F. Eifler,[27,28] A. E. Evrard,[29,30] B. Flaugher,[12] P. Fosalba,[22,23] J. Frieman,[12,5] J. García-Bellido,[14] D. W. Gerdes,[29,30] D. Gruen,[24,31] R. A. Gruendl,[19,20] J. Gschwend,[17,18] G. Gutierrez,[12] W. G. Hartley,[1,7] K. Honscheid,[32,33] B. Hoyle,[8,9] D. J. James,[34] M. Jarvis,[6] K. Kuehn,[35] M. Lima,[36,17] H. Lin,[12] M. March,[6] P. Melchior,[37] F. Menanteau,[19,20] R. Miquel,[38,21] A. A. Plazas,[28] K. Reil,[31] A. Roodman,[24,31] E. Sanchez,[25] V. Scarpine,[12] M. Schubnell,[30] I. Sevilla-Noarbe,[25] M. Smith,[39] M. Soares-Santos,[12,40] F. Sobreira,[41,17] E. Suchyta,[42] M. E. C. Swanson,[20] G. Tarle,[30] D. Thomas,[13] A. R. Walker[11]

(DES Collaboration)

The authors' affiliations are shown in Appendix A.

ABSTRACT

Mapping the underlying density field, including non-visible dark matter, using weak gravitational lensing measurements is now a standard tool in cosmology. Due to its importance to the science results of current and upcoming surveys, the quality of the convergence reconstruction methods should be well understood. We compare three different mass map reconstruction methods: Kaiser-Squires, Wiener filter, and Glimpse. Kaiser-Squires is a direct inversion method, taking no account of survey masks or noise. The Wiener filter is well motivated for Gaussian density fields in a Bayesian framework. The Glimpse method uses sparsity, with the aim of reconstructing non-linearities in the density field. We compare these methods with a series of tests on the public Dark Energy Survey (DES) Science Verification (SV) data and on realistic DES simulations. The Wiener filter and Glimpse methods offer substantial improvement on the standard smoothed Kaiser-Squires with a range of metrics. For both the Wiener filter and Glimpse convergence reconstructions we present a 12% improvement in Pearson correlation with the underlying truth from simulations. To compare the mapping methods' abilities to find mass peaks, we measure the difference between peak counts from simulated ΛCDM shear catalogues and counts from catalogues with no mass fluctuations. This is a standard data vector when inferring cosmology from peak statistics. The maximum signal-to-noise value of these peak statistic data vectors was increased by a factor of 3.5 for the Wiener filter and by a factor of 9 using Glimpse. Glimpse shows significant signal-to-noise in the peak statistic data vectors at high values of convergence, whereas the Wiener filter and KS suppress these high peaks. With simulations we measure the reconstruction of the harmonic phases in the maps, showing that the concentration of the phase residuals is improved 17% by Glimpse and 18% by the Wiener filter. Finally, we show that the correlation between the reconstructions from data and the foreground redMaPPer clusters is increased 18% by the Wiener filter and 32% by Glimpse.

Key words: gravitational lensing: weak – large-scale structure of Universe– methods: statistical

⋆ E-mail: niall.jeffrey.15@ucl.ac.uk

1 INTRODUCTION

Mass map reconstruction from weak gravitational lensing recovers the underlying matter distribution in the universe from measurements of galaxy shapes. Images of distant galaxies are deformed by the inhomogeneous matter distribution along the line of sight. Any matter can contribute to the lensing effect, making it a direct probe of non-visible dark matter.

Weak lensing, which takes advantage of the statistical power from many small distortions (that is, gravitational lensing induced "shears"), is now a well established tool in constraining cosmology. The Dark Energy Survey (DES) has used the 2-point correlation function of shear to contribute to excellent constraints on cosmological parameters and models, including the nature of dark energy (DES Collaboration et al. 2017). Shear 2-point correlation functions have been used to constrain cosmology from many other survey datasets (van Uitert et al. 2017, Kilbinger et al. 2013). These methods use the shear measurements directly, as the shear can be related to the underlying matter distribution without needing to explicitly reconstruct mass maps.

A zero-mean Gaussian random field can be characterised entirely by its 2-point correlations. The matter density field in the early universe is expected to be highly Gaussian, a property which persists into the late universe for the large scales that were less affected by gravitational collapse. For the smaller scales at late times, non-linear gravitational collapse has led to a highly non-Gaussian density field. Much valuable information can be extracted from this non-Gaussianity, although this requires additional methods beyond 2-point statistics.

Popular proposed methods to extract this information include N-point statistics and higher order moments (Cooray & Hu 2001), peak statistics (Dietrich & Hartlap 2010, Kacprzak et al. 2016, Peel et al. 2017, Shan et al. 2017, Martinet et al. 2017), and Minkowksi functionals (Kerscher et al. 1996, Petri et al. 2013). It is often either essential or convenient to apply these methods to the density field directly (rather than in the space of the shear measurements), thereby necessitating a reliable mass map reconstruction.

Peak statistics are particularly promising, as peaks in the density field probe the non-Gaussian structure directly. Peaks can be identified from aperture mass maps, which are derived by convolving the shear data with a kernel, or from the reconstructed density field. The first approach has the advantage of having local noise, while the second is "closer" to the underlying density field and often has faster algorithms. Both methods often require simulations to provide a link between the theory and data, with the exception of proposed semi-analytic models (Lin & Kilbinger 2015, Shan et al. 2017).

In addition to using mass maps for higher order statistics to constrain cosmological parameters and models, the mass maps can themselves be intrinsically useful. Clerkin et al. (2017), using the original DES Science Verification (SV) mass map, show evidence that the 1-point distribution of the density field is more consistent with Log-Normal than Gaussian. Combining mass maps with the spatial distributions of stellar mass, galaxies, or galaxy clusters allows the relationship between the visible baryonic matter and invisible dark matter to be studied. Using mass maps to constrain

galaxy bias (Chang et al. 2016), the relation between the distribution of galaxies and matter, can in turn aid cosmological probes other than weak lensing. Maps also enable simple tests for systematic error in the galaxy shape catalogues.

Since the first application of mass mapping methods to wide-field surveys with the Canada-France Hawaii Telescope Lensing Survey (CFHTLenS) data (Van Waerbeke et al. 2013), mass maps have been a standard product of large weak lensing surveys. In addition to DES, current surveys reconstructing the density field from weak lensing data include the Kilo-Degree Survey (Giblin et al. in prep.) and the Hyper Supreme-Cam Subaru Strategic Program (HSC-SSP) (Oguri et al. 2017). Mapping dark matter is key to the science goals of the future Euclid Mission (Amendola et al. 2016) and the Large Synoptic Survey Telescope (LSST Science Collaboration et al. 2009).

DES is a ground based photometric galaxy survey, observing in the southern sky from the 4m Blanco telescope at the Cerro Tololo Inter-American Observatory (CTIO) in Chile with five photometric filters covering the optical and near-infrared spectrum using the Dark Energy Camera (Flaugher et al. 2015, Dark Energy Survey Collaboration et al. 2016). The SV data come from an initial run over a fraction of the final sky coverage, but to almost the full exposure time of the final survey. The sky coverage is still large, 139 $\deg^2$, and the nearly full exposure (Chang et al. 2015) gives a galaxy density almost equal to what is expected after the complete 5 years of DES observations.

This paper uses the public DES SV data to compare the quality of mass mapping reconstruction methods. The maps are of the two-dimensional convergence, κ, a weighted projection of the density field in the foreground of the observed background galaxies. Recovering the convergence from the shear data is an ill-posed inverse problem, troubled by survey masks and galaxy "shape noise".

This work follows on from that of Chang et al. (2015) and Vikram et al. (2015) in which the original DES SV mass map was created using the Kaiser & Squires (1993) method. In this paper we compare three quite different methods: Kaiser-Squires (KS); Wiener filtering (Wiener 1949); and GLIMPSE (Leonard et al. 2014, Lanusse et al. 2016), a sparsity-based reconstruction method. The Kaiser-Squires method is a direct inversion from shear to convergence, taking no account of missing data or the effect of noise. The Wiener filter and GLIMPSE assume different prior knowledge about the underlying convergence to account for the effects of noise and missing data.

In Sec. 2 we describe the theoretical foundation for weak lensing mass mapping and the three different methods used for this work. In Sec. 3 we describe the DES SV shear data, the accompanying simulations, and the redMaPPer galaxy cluster catalogue. Foreground galaxy clusters are expected to trace the true density field, and therefore should be correlated with the convergence reconstruction. The different methods are also applied to realistic data simulations where the true convergence is known. In Sec. 4 we present our results on data and simulation, using various quality metrics for the reconstruction. On simulations these metrics are the Pearson correlation coefficient, the pixel root-mean-square error (RMSE), the variance of the 1-point distribution of pixel values, the phase residuals, and peak statistics. On

data we compare the convergence reconstructions to the foreground galaxy clusters. We conclude in Sec. 5.

2 METHODOLOGY

2.1 Weak gravitational lensing

We can use measurements of the distortion of background galaxy shapes by weak gravitational lensing to learn about the mass distribution in the foreground without making many physical assumptions or relying on phenomenological models. For convenience, here we summarise some of the existing literature relevant for mass mapping from weak lensing (Bartelmann & Schneider 2001, Kilbinger 2015).

The weak lensing formalism follows photon paths along geodesics in a perturbed Friedmann-Robertson-Walker (FRW) metric. The perturbations are sourced by the density field of large scale structure. Throughout we assume that the perturbations are small, and that the measurements are made over a small enough patch of the sky that the sky geometry is Euclidian. Consistent with the Planck CMB results (Planck Collaboration et al. 2016) and motivated by inflationary theory, we assume that the global geometry of the universe is flat.

The density contrast, $\delta = (\rho - \bar{\rho})/\bar{\rho}$, of a pressureless fluid is related to the scalar gravitational potential perturbation, Φ, through the Poisson equation,

$$\nabla^2 \Phi = \frac{3 H_0 \Omega_m}{2a} \delta \ , \tag{1}$$

where H_0 is the present value of the Hubble parameter, a is the cosmological scale factor, and ρ and $\bar{\rho}$ are the local and mean density respectively.

For a flat universe, the lensing potential is given by

$$\psi(\vec{\theta}, \omega) = 2 \int_0^\omega d\omega' \left[\frac{\omega - \omega'}{\omega \omega'} \right] \Phi(\vec{\theta}, \omega') \ , \tag{2}$$

where ω is the comoving distance.

The Born approximation assumes that the observed angle to a point, $\vec{\theta}$, deviates only a small amount from the true angle $\vec{\beta}$, so the change in distance of the photon's path is negligible. We can characterise the effect of lensing on the galaxies using the Jacobian of the transformation, $\mathcal{A}_{ij} = \partial \beta_i / \partial \theta_j$, which is decomposed into the functions $\kappa(\vec{\theta})$ and $\gamma(\vec{\theta}) = \gamma_1 + i\gamma_2$, and which is given by

$$\begin{aligned} \mathcal{A} &= \left(\delta_{ij} - \frac{\partial^2 \psi(\vec{\theta})}{\partial \theta_i \partial \theta_j} \right) \\ &= \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \ . \end{aligned} \tag{3}$$

Using the definition of the lensing potential and the Poisson equation, the convergence can be expressed as an integral over the density along the line of sight,

$$\kappa(\vec{\theta}, \omega) = \frac{3 H_0^2 \Omega_m}{2} \int_0^\omega d\omega' \frac{\omega'(\omega - \omega')}{\omega} \frac{\delta(\vec{\theta}, \omega')}{a(\omega')} \ . \tag{4}$$

For a distribution $n(\omega)$ of lensed galaxies, the lensing efficiency kernel is defined to be

$$p(\omega') = \int_{\omega'}^{\omega_\infty} \left(\frac{\omega - \omega'}{\omega} \right) n(\omega) d\omega \ ; \tag{5}$$

this weights the contribution of the foreground density fluctuations to give the convergence weighted over the redshift distribution of source galaxies,

$$\begin{aligned} \kappa(\vec{\theta}) &= \int_0^\infty n(\omega) \kappa(\vec{\theta}, \omega) d\omega \\ &= \frac{3 H_0^2 \Omega_m}{2} \int_0^\infty d\omega' p(\omega') \omega' \frac{\delta(\vec{\theta}, \omega')}{a(\omega')} \ . \end{aligned} \tag{6}$$

The shear, $\gamma(\vec{\theta})$, which is assumed to be an observable in the weak lensing limit, is given by

$$\begin{aligned} \gamma(\vec{\theta}) &= \frac{1}{\pi} \int_{\mathbb{R}^2} d^2 \theta' \mathcal{D}(\vec{\theta} - \vec{\theta}') \kappa(\vec{\theta}') \\ &\text{where } \mathcal{D}(\vec{\theta}) = -(\theta_1 - i\theta_2)^{-2} \ . \end{aligned} \tag{7}$$

For surveys where the integral is over large angles on the sky, this formulation breaks down, and requires a full treatment in spherical bases. Wallis et al. (2017) show that errors can be introduced at an $\mathcal{O}(1\%)$ level for correlations between points at DES SV angular separation depending on the projection. All of the methods used here use the small angle approximation, and should suffer equally.

The real and imaginary parts of the shear γ represent a chosen two dimensional coordinate system. In weak lensing, the observed ellipticity[1] of a galaxy ϵ_{obs} is related to the reduced shear g plus the intrinsic ellipticity of the source galaxy ϵ_s through

$$\begin{aligned} \epsilon_{obs} &\approx g + \epsilon_s \\ &\text{where } g = \frac{\gamma}{1 - \kappa} \ . \end{aligned} \tag{8}$$

The reduced shear is approximately the true shear, $g \approx \gamma$, in the weak lensing limit. This allows a standard definition of observed shear, $\gamma_{obs} = \epsilon_{obs}$, where the measurements are degraded by "shape noise", caused by the ϵ_s values of the observed galaxies:

$$\gamma_{obs} \approx \gamma + \epsilon_s \ . \tag{9}$$

The shape noise for a given galaxy is modelled as a randomly-drawn Gaussian variate, $\epsilon_s \sim G(0, \sigma_\epsilon)$, where σ_ϵ is estimated from data. The distribution of the ellipticity from the SV data in figure 2 is not an exact Gaussian, as the true distribution is the result of galaxy astrophysics, though a Gaussian still has properties that make it a good approximation. The Gaussian would be the maximum entropic, least informative, distribution for known mean and variance, and, by the central limit theorem, would be the correct distribution in the limit of large numbers of galaxies averaged in pixels.

---

[1] Using the Bartelmann & Schneider (2001) equation 4.10 ellipticity definition for ϵ.

It is possible to extend the simple Kaiser-Squires method (Sec. 2.2) to use the reduced shear, g, for the mildly non-linear lensing regime when it is no longer appropriate to assume $g \approx \gamma$ (Schneider & Seitz 1995, Seitz & Schneider 1995, Seitz & Schneider 2001). This is also done by GLIMPSE (Sec. 2.4).

In matrix notation, the problem as given by equations 7 and 9 can be expressed as a linear model, with a data vector of observed shear measurements

$$\gamma = \mathbf{A}\kappa + \mathbf{n} \,, \tag{10}$$

where $\mathbf{A}$ is a discretised version of equation 7 and $\mathbf{n}$ is a noise vector due to shape noise (equation 9). The elements of the data vector can either correspond to the individual shear measurements or to measurements binned into angular pixels (in which case the noise vector would be the average noise in the pixel).

The convergence need not be reconstructed with the same pixelisation as the shear measurements, giving κ and γ vectors of different length. Missing data due to survey masks would correspond to a shorter γ vector; here one may wish to fill in the convergence in the masked region — this is known as *inpainting*. Different sized κ and γ vectors result in a non-square $\mathbf{A}$ matrix, potentially causing inversion problems.

2.2   Kaiser-Squires reconstruction

2.2.1   Theory

The convergence-to-shear relationship, equation 7, is a convolution in the two dimensional angular plane. The two-dimensional Fourier transforms of the shear and convergence, defined for κ as

$$\tilde{\kappa}(\vec{k}) = \int_{\mathbb{R}^2} \mathrm{d}^2\theta \kappa(\vec{\theta}) \exp(i\vec{\theta} \cdot \vec{k}) \,, \tag{11}$$

are related through an elementwise product via the convolution theorem

$$\tilde{\gamma}(\vec{k}) = \pi^{-1} \tilde{\mathcal{D}}(\vec{k}) \tilde{\kappa}(\vec{k}) \,, \tag{12}$$

where the Fourier transform of the kernel is given by

$$\tilde{\mathcal{D}}(\vec{k}) = \pi \frac{(k_1^2 - k_2^2 + 2ik_1k_2)}{|\vec{k}|^2} \,; \tag{13}$$

here k_1 and k_2 are the components of $\vec{k}$. Using $\tilde{\mathcal{D}}\tilde{\mathcal{D}}^* = \pi^2$, equation 12 can be rewritten:

$$\tilde{\kappa}(\vec{k}) = \pi^{-1} \tilde{\gamma}(\vec{k}) \tilde{\mathcal{D}}^*(\vec{k}) \quad \text{for} \quad \vec{k} \neq \vec{0} \,. \tag{14}$$

The inverse Fourier transform then returns the convergence reconstruction in configuration space (Kaiser & Squires 1993).

The real and imaginary parts of the reconstruction are the E- and B-modes respectively, where $\kappa_{\mathrm{recon}} = \kappa_E + i\kappa_B$. In standard cosmology (equation 7), the convergence sourced by a real density field should be a pure E-mode. Errors, noise or other systematic effects can lead to B-mode contributions to the reconstruction.

2.2.2   Implementation

In the matrix formulation of equation 10, this deconvolution corresponds to multiplying the Fourier space shear field with the inverse of $\mathbf{A}$ in Fourier space. For a case with no shape noise, that is

$$\tilde{\gamma} = \tilde{\mathbf{A}}\tilde{\kappa} \,, \tag{15}$$

the Kaiser-Squires method is identical to using the inverse matrix

$$\begin{aligned} \left[\tilde{\mathbf{A}}^{-1}\right]_{ij} &= \frac{k_{1,i}^2 - k_{2,i}^2 - 2ik_{1,i}k_{2,i}}{k_{1,i}^2 + k_{2,i}^2} \delta_{ij} \\ &= \left[\tilde{\mathbf{A}}^\dagger\right]_{ij} \,, \end{aligned} \tag{16}$$

where the Kronecker delta function, δ_{ij}, relates the element-wise multiplication in Fourier space to a diagonal matrix operator, and † is the conjugate transpose.

For the Kaiser-Squires inversion in configuration space, the $\mathbf{A}$ and $\mathbf{A}^\dagger$ matrices are not diagonal, and therefore are slower to compute. The discretisation of the underlying smooth shear field into finite configuration space makes the property $\mathbf{A}\mathbf{A}^\dagger = \mathbf{I}$ inexact. As a result of these factors, we choose to implement the Kaiser-Squires reconstruction in Fourier space.

The shear due to lensing is much smaller than the shape noise, and not all places on the sky contain usable galaxies. Both the shape noise and the random sampling of background galaxies propagate error through this noisy reconstruction. Binning the shear measurements into larger pixels can reduce the shape noise per pixel and ensure that there are no empty pixels, but this comes at a loss of the small scale information and cannot deal with masks or the edges of the survey.

A smoothing filter is applied to the Kaiser-Squires reconstruction to reduce the noise. This will similarly lose any small scale structure, and especially suppress peaks in the convergence. In this work, matching Chang et al. (2015), we smooth the Kaiser-Squires maps with a Gaussian kernel. The standard deviation scale, σ_{smooth}, of this Gaussian kernel is free to be chosen, where $\sigma_{smooth} = 0$ corresponds to standard, unsmoothed Kaiser-Squires.

2.3   Wiener Filter

2.3.1   Theory

The Wiener filter is the linear minimum-variance solution to linear problems of the type in equation 10, where the noise is uncorrelated. The Wiener filter reconstruction (Lahav et al. 1994, Zaroubi et al. 1995) is given by

$$\begin{aligned} \kappa_W &= \mathbf{W}\gamma \\ \mathbf{W} &= \mathbf{S}_\kappa \mathbf{A}^\dagger \left[\mathbf{A}\mathbf{S}_\kappa \mathbf{A}^\dagger + \mathbf{N}\right]^{-1} \,. \end{aligned} \tag{17}$$

Here $\mathbf{S}_\kappa$ and $\mathbf{N}$ are the signal and noise covariance matrices respectively, which are $\langle \kappa\kappa^\dagger \rangle$ and $\langle \mathbf{n}\mathbf{n}^\dagger \rangle$ for this problem.

This filter is the linear minimum-variance solution, as $\mathbf{W}$ is a linear operator that minimises the variance

$$\langle (\mathbf{W}\gamma - \kappa)^\dagger (\mathbf{W}\gamma - \kappa) \rangle \,. \tag{18}$$

If the chosen prior on κ does not constrain the reconstruction, so that $\mathbf{S}_\kappa^{-1} \rightarrow 0$ (Simon et al. 2009), or if the data are noise-free, $\mathbf{N} = 0$, then the linear minimum variance filter becomes the Kaiser-Squires reconstruction. Setting $\mathbf{S}_\kappa^{-1} \rightarrow 0$ is equivalent to removing the signal prior in the following Bayesian framework.

From a different starting point, for the Wiener posterior we begin by assuming a Gaussian likelihood (Jasche & Lavaux 2015)

$$Pr(\gamma|\kappa) = \frac{1}{\sqrt{(\det 2\pi\mathbf{N})}}\exp\left[-\frac{1}{2}(\gamma - \mathbf{A}\kappa)^\dagger \mathbf{N}^{-1}(\gamma - \mathbf{A}\kappa)\right], \quad (19)$$

where it is assumed that $\mathbf{N}$ is known and the noise is both uncorrelated and Gaussian, as assumed in equation 10. Intrinsic alignments of clustered galaxies will violate this uncorrelation condition.

The prior on the convergence is that of a Gaussian random field, which is applicable for the density field on large scales at late times,

$$Pr(\kappa|\mathbf{S}_\kappa) = \frac{1}{\sqrt{(\det 2\pi\mathbf{S}_\kappa)}}\exp\left[-\frac{1}{2}\kappa^\dagger \mathbf{S}_\kappa^{-1}\kappa\right]. \quad (20)$$

Using Bayes' theorem and the fact that $Pr(\gamma|\mathbf{S}_\kappa, \kappa) = Pr(\gamma|\kappa)$, the full posterior is given by

$$\begin{aligned} Pr(\kappa|\mathbf{S}_\kappa, \gamma) &= \frac{Pr(\gamma|\kappa)Pr(\kappa|\mathbf{S}_\kappa)}{Pr(\gamma)} \\ &\propto \frac{1}{\sqrt{(\det 2\pi\mathbf{S}_\kappa)}}\frac{1}{\sqrt{(\det 2\pi\mathbf{N})}}\times \\ &\quad \exp\left[-\frac{1}{2}\kappa^\dagger\mathbf{S}_\kappa^{-1}\kappa - \frac{1}{2}(\gamma - \mathbf{A}\kappa)^\dagger\mathbf{N}^{-1}(\gamma - \mathbf{A}\kappa)\right] \\ &\propto \exp\left[-\frac{1}{2}(\kappa - \mathbf{W}\gamma)^\dagger(\mathbf{S}_\kappa^{-1} + \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^\dagger)(\kappa - \mathbf{W}\gamma)\right], \end{aligned}$$
$$(21)$$

where $\mathbf{W}$ is the Wiener filter, so the *maximum a posteriori* (MAP) solution is that of the Wiener reconstruction.

If the aim of the reconstruction is to infer cosmology from the non-Gaussian component of the density field, the Wiener filter may not be the ideal method for mass map recovery. The small scale modes with less power are often suppressed, losing the peak structure. Qualitatively it can be thought of as either the Gaussian prior being inappropriate or as the linear filter being insufficient.

2.3.2 Implementation

Using the exact Fourier space property $\tilde{\mathbf{A}}^{-1} = \tilde{\mathbf{A}}^\dagger$ we rewrite equation 17 as

$$\begin{aligned} \tilde{\kappa}_W &= \tilde{\mathbf{A}}^{-1}\tilde{\mathbf{S}}_\gamma\left[\tilde{\mathbf{S}}_\gamma + \tilde{\mathbf{N}}\right]^{-1}\tilde{\gamma} \\ &= \tilde{\mathbf{A}}^{-1}\tilde{\gamma}_W \\ &= \tilde{\mathbf{A}}^\dagger\tilde{\gamma}_W, \end{aligned}$$
$$(22)$$

where we have used $\tilde{\mathbf{A}}\tilde{\mathbf{S}}_\kappa\tilde{\mathbf{A}}^\dagger = \langle\tilde{\mathbf{A}}\tilde{\kappa}\tilde{\kappa}^\dagger\tilde{\mathbf{A}}^\dagger\rangle = \langle\tilde{\gamma}\tilde{\gamma}^\dagger\rangle = \tilde{\mathbf{S}}_\gamma$. This shows that applying the Wiener filter to the shear to recover γ_W and then applying the Kaiser-Squires inversion in Fourier

space is equivalent to directly calculating the Wiener filter of the convergence.

In configuration space, the noise covariance matrix is given by

$$\left[\mathbf{N}\right]_{ij} = \frac{2\sigma_\epsilon^2}{p_i}\delta_{ij}, \quad (23)$$

where p_i is the galaxy count per pixel. Empty pixels in the masked region have infinite variance, absorbing the mask into a special case of the Wiener filter denoising.

The signal properties for a Gaussian random field are constrained entirely by the mean and the signal covariance matrix, which in harmonic space is identical to the power spectrum. The cosmological principle implies that the angular distribution of a field on the sky is statistically isotropic, so the angular power spectrum, C_ℓ, can contain all the 2-point statistical information. The angular power spectrum of the physical shear E-mode shear signal is defined as

$$\begin{aligned} C_{\ell,E} &= \frac{1}{2\ell+1}\sum_{m=-\ell}^{+\ell}\langle|a_{\ell m,E}|^2\rangle \\ &= C_{\ell,\kappa}, \end{aligned}$$
$$(24)$$

where $a_{\ell m}$ are the spherical harmonic coefficients and the brackets $\langle\ \rangle$ average over realisations of the signal. The second equality assumes the flat sky approximation for high[2] ℓ.

We generate a theoretical power spectrum using the Limber approximation with the Cosmosis package (Zuntz et al. 2015) with our prior fiducial cosmological parameters: $\Omega_m = 0.286$, $\Omega_\Lambda = 0.714$, $\Omega_b = 0.047$, $h = 0.7$, $\sigma_8 = 0.82$, $n_s = 0.962$ and w $= -1$. We use a background galaxy distribution defined from equation 31, and shown in figure 1.

It is commonly asked whether it is reasonable to assume cosmological parameters in the map reconstruction, if the maps are then used to infer cosmological parameters. Though we assume a specific set of cosmological parameters, it would still be possible to use the maps for cosmological parameter estimation, from peak statistics for example, if the same prior is used on the simulations and the data identically. If simulations are not used, the power spectrum can be jointly inferred from the data (Jasche & Lavaux 2015) using Gibbs sampling.

In order to generate the power spectrum in flat Fourier space, rather than on the curved sky, we again use a flat sky approximation

$$k_\theta^2 P(k_\theta) = \left(\frac{\mathcal{N}}{2\pi}\right)^2\ell(\ell+1)C_\ell, \quad (25)$$

adapted from Loverde & Afshordi (2008), where $\mathcal{N}$ is the total number of pixels in the map, k_θ is the magnitude of the projected Fourier mode, and where we have defined our projected angular power spectrum as

$$P(k_\theta)\delta(\mathbf{k}_\theta - \mathbf{k}_\theta') = \langle\ \tilde{\gamma}(\mathbf{k}_\theta)\ \tilde{\gamma}^\dagger(\mathbf{k}_\theta)\ \rangle. \quad (26)$$

The largest scale mode is $\ell = 20.51$, which corresponds to an angular separation of 17.55 deg.

---

[2] We omit a prefactor which goes as $1 - O(\ell^{-2})$ for high ℓ.

Though the signal covariance matrix is diagonal in harmonic space (equation 26), and the independent noise has covariance which is diagonal in configuration space (equation 23), there is no natural basis in which both are sparse. Inversion of dense matrices to evaluate the Wiener filter is bypassed using the algorithm presented in Elsner & Wandelt (2013), where an additional messenger field is used to pass information between harmonic and configuration space, iteratively converging to the Wiener filter solution.

These messenger field methods were extended by Jasche & Lavaux (2015) to draw Markov chain Monte Carlo (MCMC) samples from the whole Wiener posterior (equation 21). The first application of messenger field methods to weak lensing data was by Alsing et al. (2016, 2017), who drew samples from the Wiener posterior and generated Wiener filtered shear (not convergence) maps from CFHTLenS data. By comparison, in this work we do not sample from the Wiener posterior; instead, we use the original messenger field algorithm of Elsner & Wandelt (2013) to calculate the Wiener filter reconstruction of the convergence map from DES SV shear data and simulations.

2.4 Sparsity reconstruction

2.4.1 Theory

Consider the coefficients α of the decomposition of a signal $\mathbf{x}$ in a representation space (or "dictionary") $\mathbf{\Phi}$, so that $\mathbf{x} = \mathbf{\Phi}\alpha$. Example dictionaries include the Fourier transform or wavelet transforms. Assuming a sparse prior on the signal $\mathbf{x}$ in the dictionary $\mathbf{\Phi}$ means that its representation α is expected to be sparse, that is, with most of the coefficients equal to 0 (Starck et al. 2015). A simple example is a cosine function signal and a Fourier transformation; in this sparse basis only two coefficients have a non-zero value (corresponding to the frequency of the cosine function).

Formally most signals cannot strictly be made sparse, and are merely compressible with a choice of an appropriate transformation, such as a wavelet transform (Starck et al. 2015, Leonard et al. 2014). For a compressible signal the magnitude-ordered sparse coefficients, α_i, are expected to have exponential decay and therefore to have a Laplace distribution (Tibshirani 1994).

Consider a generic linear inverse problem of the form $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$. A robust estimate of the signal $\mathbf{x}$ can be recovered by solving the ("LASSO") optimisation problem

$$\underset{\alpha}{\arg\min} \quad ||\mathbf{y} - \mathbf{A}\mathbf{\Phi}\alpha||_2^2 + \lambda ||\alpha||_1 \ , \tag{27}$$

where λ is a Lagrangian multiplier (Tibshirani 1994). Here the first term corresponds to a χ^2 minimisation, ensuring fidelity of the signal reconstruction, while the second is the sparsity-promoting regularisation term.

We can include non-constant noise variance by weighting the first χ^2 according to the variance. If the noise variance is included in the χ^2 term, the λ value can be interpreted as a signal-to-noise level in the transformed (e.g. wavelet) space.

The second term does not use the Euclidan l_2 norm, but instead uses the sparsity-promoting l_1 norm, defined as

$$||\alpha||_1 = \sum_i |\alpha_i| \ . \tag{28}$$

These methods are non-linear, so it can be difficult to derive properties analytically. With realistic simulations of the data and true signal, the value for λ can be chosen to maximise some success metric. This is analogous to selecting a theoretical power spectrum for the Wiener filter, or a smoothing scale for Kaiser-Squires.

Sparse recovery methods are non-linear and are not necessarily formulated in the Bayesian framework of the Wiener filter. The Wiener filter reconstruction is that which maximises the Wiener posterior, which is known analytically provided the noise and signal are Gaussian with known covariance. However, one may make a frequentist estimate of the error of the sparse reconstruction by propagating the noise properties of the data using bootstrapping or Monte Carlo techniques.

2.4.2 Implementation/GLIMPSE

The choice of dictionary depends on the structures contained in the signal. Theory of structure formation in the universe predicts the formation of quasi-spherical halos of bound matter. It is standard practice to represent the spatial distribution of matter in halos with spherically symmetric Navarro-Frenk-White (Navarro et al. 1996) or Singular Isothermal Sphere profiles. Coefficients of Isotropic Undecimated Wavelets (Starck et al. 2015) in two dimensions are well suited to the observed convergence of a dark matter halo. The wavelet transform used in the GLIMPSE algorithm is the Starlet (Starck et al. 2007), which can represent positive, isotropic objects.

The sparsity prior in the starlet basis enforces a physical model that the matter field is a superposition of spherically symmetric dark matter halos. This is not wholly correct, but is an approximation which is true for the non-linear regime in the standard model of structure formation, similarly to how the assumption of Gaussianity holds in the linear regime. On large scales, where the density field is expected to be Gaussian, the GLIMPSE sparsity prior is less appropriate.

The GLIMPSE algorithm aims to solve the optimisation problem

$$\hat{\kappa} = \underset{\kappa}{\arg\min} \quad ||\mathbf{N}^{-\frac{1}{2}} \left[\gamma - \mathbf{T}^\dagger \hat{\mathbf{A}} \mathbf{F}\kappa\right]||_2^2$$
$$+ \lambda ||\omega \mathbf{\Phi}^\dagger \kappa||_1 + i_{\mathrm{Im}(\kappa)=0} \ , \tag{29}$$

where $\mathbf{F}$ is the Fourier transform matrix, $\mathbf{T}$ is the Non-equispaced Discrete Fourier Transform (NDFT) matrix, $\hat{\mathbf{A}}$ is defined in equation 16, ω is a diagonal matrix of weights, and $\mathbf{\Phi}^\dagger$ is the inverse wavelet transform. The indicator function $i_{\mathrm{Im}(\cdot)=0}$ (defined in Appendix B) in the final term imposes realness on the reconstruction (no B-modes). The use of NDFT allows the first term to perform a forward fitted Kaiser-Squires-like step without binning the shear data, allowing the smaller-scales to be retained in the reconstruction. The full algorithm, including the calculation of the weights, is described in Sec. 3.2 in Lanusse et al. (2016).

Though the problem presented in equation 29 is an optimisation using the shear data γ, in fact it is the reduced shear (equation 8) that GLIMPSE uses to recover κ (Lanusse et al. 2016). As an extension, the GLIMPSE algorithm can also perform the joint reconstruction with reduced shear and flexion, a third-order weak gravitational lensing effect (Bacon et al. 2006) (although no flexion data are available for our galaxy shear catalogue).

As the prior knowledge in this reconstruction relates to the quasi-spherical clustering of bound matter, enforced through a sparsity prior in Starlet space, this method should better reconstruct the smaller scale non-Gaussian structure than the Wiener filter.

3 DATA AND SIMULATIONS

3.1 Dark Energy Survey Science Verification Data

The shear data are from the 139 deg^2 SPT-E field of the public DES SV data. This initial test data set was taken during an observing run before the official start of the full science survey. The galaxy catalogue comes from the SVA1 (Science Verification) Data Release[3]. Due to changes to the catalogues before final release (more galaxy shear measurements are now available to us), the catalogue used in this work is not identical to that used by Chang et al. (2015), even when the same data selections are made. All maps are therefore new, and slightly different to the previously published SV map.

The photometric redshifts from five optical filters ($grizY$) were estimated using the Bayesian Photometric Redshifts (BPZ) code (Benítez 2000, Coe et al. 2006, & Bonnett et al. 2016). The final median depth estimates are g ~24.0, r ~23.0, i ~23.0 and z ~22.4 (10-σ galaxy limiting magnitude). The "background galaxies", the ones from which the shear is measured, are taken in the range $0.6 < z_{mean} < 1.2$. The z_{mean} value for each galaxy is the mean of the posterior probability distribution function (PDF) estimated using the BPZ code. The PDF for each galaxy is very broad, giving a total stacked PDF of background galaxies that extends beyond the [0.6, 1.2] redshift range, as can be seen in figure 1.

Using the `ngmix` shape catalogue, we apply a selection of `sva1_flag = 0` & `ngmix_flag = 0` to obtain galaxies with a well-measured shear. The `ngmix` catalogue contains corrections to measurement bias, in the form of "sensitivities", which can be applied to a weighted ensemble of hundreds or thousands of galaxies, but which cannot be applied per galaxy (which is not ideal for mass mapping). The structure of equation 7 implies that a multiplicative shear bias would lead to a convergence amplitude bias. Under the assumption that multiplicative shear bias will not vary across the survey area, we correct all measured ellipticities by the same debiasing factor

$$\epsilon_{obs,i} = \epsilon_{measured,i} \times \bar{s}^{-1} \, , \tag{30}$$

where i is a galaxy index and $\bar{s}$ (≈ 0.82) is the mean sensitivity correction from all galaxies in our `ngmix`-selected cat-

---

[3] http://des.ncsa.illinois.edu
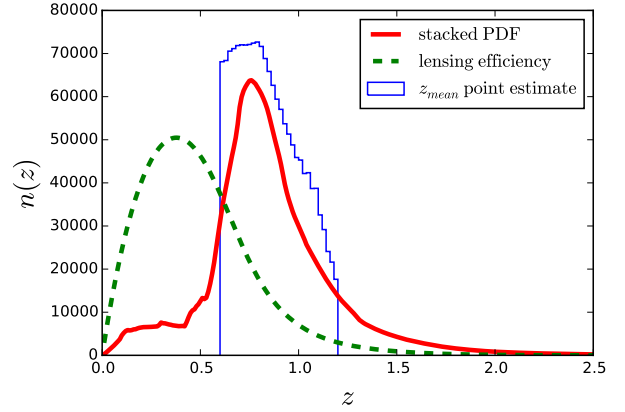
Figure 1. The redshift distribution from BPZ of the selected background galaxies with $0.6 < z_{mean} < 1.2$. The blue solid histogram is of the galaxies' point estimate mean redshifts in bins of $\Delta z = 0.02$. The red line is the stacked redshift probability density function (PDF) of all selected galaxies. The green dashed line is the lensing efficiency (equation 5) of the background galaxies.

alogue. The total number of galaxies after the redshift and shape measurement selection is 1, 628, 663.

For the Kaiser-Squires reconstruction, the shear measurements are binned into angular pixels in a 256×256 map, with average pixel size of 4.11 arcmin, using a sinusoidal projection with a centre at RA=71.0 deg. This is similar to the 5 arcmin pixel scale of the original Chang et al. (2015) map. The choice of central RA for Kaiser-Squires is to minimise the mask in the square projection, which is a large source of systematic error. For the Wiener filter, where the mask is taken into account, the shear measurements are also binned into angular pixels in a 256×256 map, but sinusoidally projected with a central RA=81.3 deg, to make the square maximally isotropic. The GLIMPSE algorithm does not bin the input shear measurements, but requires a pixel scale for the reconstruction, which we set as 3 arcmin using its gnomonic projection centred on RA=76.95 deg and DEC=−52.23 deg.

3.2 redMaPPer Clusters

Groups and clusters of galaxies are expected to trace the highest density regions in the foreground. They are luminous objects that correspond to regions of highly non-linear growth, where the density field has deviated from Gaussianity.

The public redMaPPer cluster catalogue (Rykoff et al. 2016) used the redMaPPer algorithm to optically identify clusters and to estimate each cluster's richness, λ_{RM}. The richness is defined as the sum of the membership probabilities over all galaxies within a scale radius (chosen to minimise the scatter in the mass-richness relation); it gives an estimate for the number of galaxies in a cluster. Cluster mass is expected to scale approximately linearly with richness. The redshift uncertainty is excellent, around $\sigma_z/(1 + z) \sim 0.01$, due to the clusters containing large numbers of well modelled, red galaxies. The public redMaPPer catalogue used in this work contains only clusters with $\lambda_{RM} \geqslant 20$, so that the clusters with less certainty of detection and characterisation are not used.
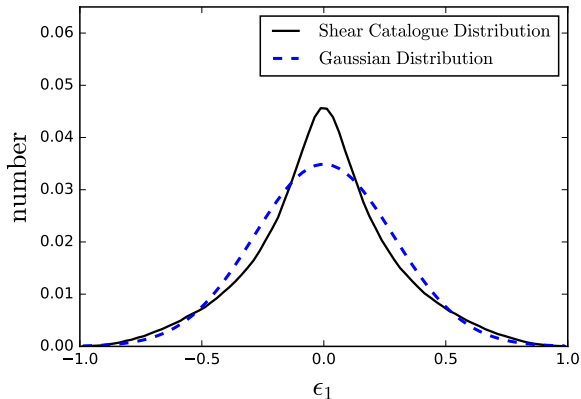
Figure 2. Distribution of the first component of ellipticity, ϵ_1, from the selected SV catalogue. A Gaussian distribution with the same mean and standard deviation shows that the ellipticity distribution is not a true Gaussian, though the noise per pixel will be more closely Gaussian due to the central limit theorem.

3.3 Simulations

To compare the reconstructions between different methods, we use a simulated catalogue with a known true convergence. We use a set of N-body simulations developed for the DES collaboration and designed to be representative of the DES data (Busha et al. 2013). The simulations used are N-body light cones composed from three boxes (1400^3, 2048^3, and 2048^3 particles in boxes of comoving length 1050 Mpc/h, 2600 Mpc/h, and 4000 Mpc/h respectively). The cosmological parameters for the simulations are: $\Omega_m = 0.286$, $\Omega_\Lambda = 0.714$, $\Omega_b = 0.047$, $\sigma_8 = 0.82$, $h_0 = 0.7$, $n_s = 0.96$, $w = -1$. We apply a mask to match the SV data.

Source galaxies have randomly-assigned positions in the simulations, as correlation between the background galaxy positions and the weak lensing shear signal is expected to be negligible. The simulated catalogues contain the lensing matrix components, $\mathcal{A}_{ij}$, for each galaxy, calculated with the ray-tracing code `CALCLENS` (Becker 2013). This provides the true κ and γ per galaxy, from which we derive the reduced shear. The shape noise due to the intrinsic ellipticities of the source galaxies, ϵ_s, is simulated by adding an ellipticity component to the reduced shear. Each noise realisation is generated from the data by randomly exchanging the ellipticity values between galaxies in the catalogue to remove the weak lensing signal and leave the shape noise.

We attempt to match the redshift distribution of the simulated galaxies to the observed redshift distribution, $n(z)$. We use the stacked posterior probability density functions of individual galaxy redshifts from the selected data catalogue (figure 1), giving an estimate of the true underlying distribution. This assumes that

$$n(z) = \sum_i p_i(z), \tag{31}$$

where $p_i(z)$ are the individual probability distributions for the galaxies from BPZ. This is not necessarily exact, due to errors in $p_i(z)$ per galaxy (Leistedt et al. 2016), but is a reasonable choice for a simulated catalogue. Using rejec-

tion sampling in bins of $\Delta z = 0.02$ we select galaxies with a probability equal to the ratio between the desired $n(z)$ from the data and the distribution in the simulation. One typical simulated catalogue contained $1,629,024$ galaxies, slightly different to the data catalogue due to the sampling scheme, but with the desired $n(z)$.

4 RESULTS

To ensure that the mass map tests are consistent with different output formats, all maps were converted onto a spherical pixelisation using `HEALPix` (Górski et al. 2005). A `HEALPix` map comprises twelve subdivisions on the sphere, which are then each partitioned into `NSIDE × NSIDE` grids. Each pixel of a `HEALPix` supersampled `NSIDE = 4096` map was filled according to the value at the corresponding RA and DEC in the reconstructed maps. The supersampled high `NSIDE` maps were then degraded to `NSIDE = 1024`. The true convergence maps from the simulations were directly binned from the convergence values at galaxy positions to `NSIDE = 1024`. For all maps the same mask is applied, where pixels with no galaxies are masked.

Figure 3 shows the mass map reconstructions from the SV shear data using the three different methods. An example simulation with truth and the three reconstructed maps is shown in figure 4. The "tuning parameters", $\sigma_{smooth} = 10.0$ arcmin for Kaiser-Squires and $\lambda = 3.0$ for GLIMPSE, are tuned to maximise the Pearson correlation coefficient r with the underlying truth when tested on simulations.

Using a suite of 10 simulations, in Sec. 4.1 we calculate the Pearson correlation coefficient between the truth and the reconstruction with different methods as a test of the reconstruction's quality. In Sec. 4.2, we calculate the root-mean-square error of the residuals between the truth and the reconstruction. In Sec. 4.3 we calculate the variance of the 1-point distribution of the pixel values in the reconstruction and compare with the truth. In Sec. 4.4 and Sec. 4.5 we quantify the quality of the reconstruction of the phase and peak statistics respectively, by comparing to the simulated truth. The final result presented in Sec. 4.6 compares the reconstruction from the DES SV shear data with foreground galaxy clusters from the redMaPPer catalogue (which are expected to trace non-linearities in the underlying density field).

In this work we do not use correlation functions as a test of the map reconstruction. None of the mass mapping methods here are expected to reproduce the correct correlation functions or power spectra. It is simple to show this analytically with the Wiener filter, where despite the filter giving the MAP pixel values, the pixel variance, and therefore the power spectrum, is suppressed.

4.1 Pixel Cross Correlation

We quantify the correlation between the true convergence from simulation and the reconstructed convergence of the simulated catalogue using the Pearson correlation coefficient. As with other metrics of success for mass map reconstruction, this can be used to tune the sparsity λ parameter and the smoothing scale for Kaiser-Squires.
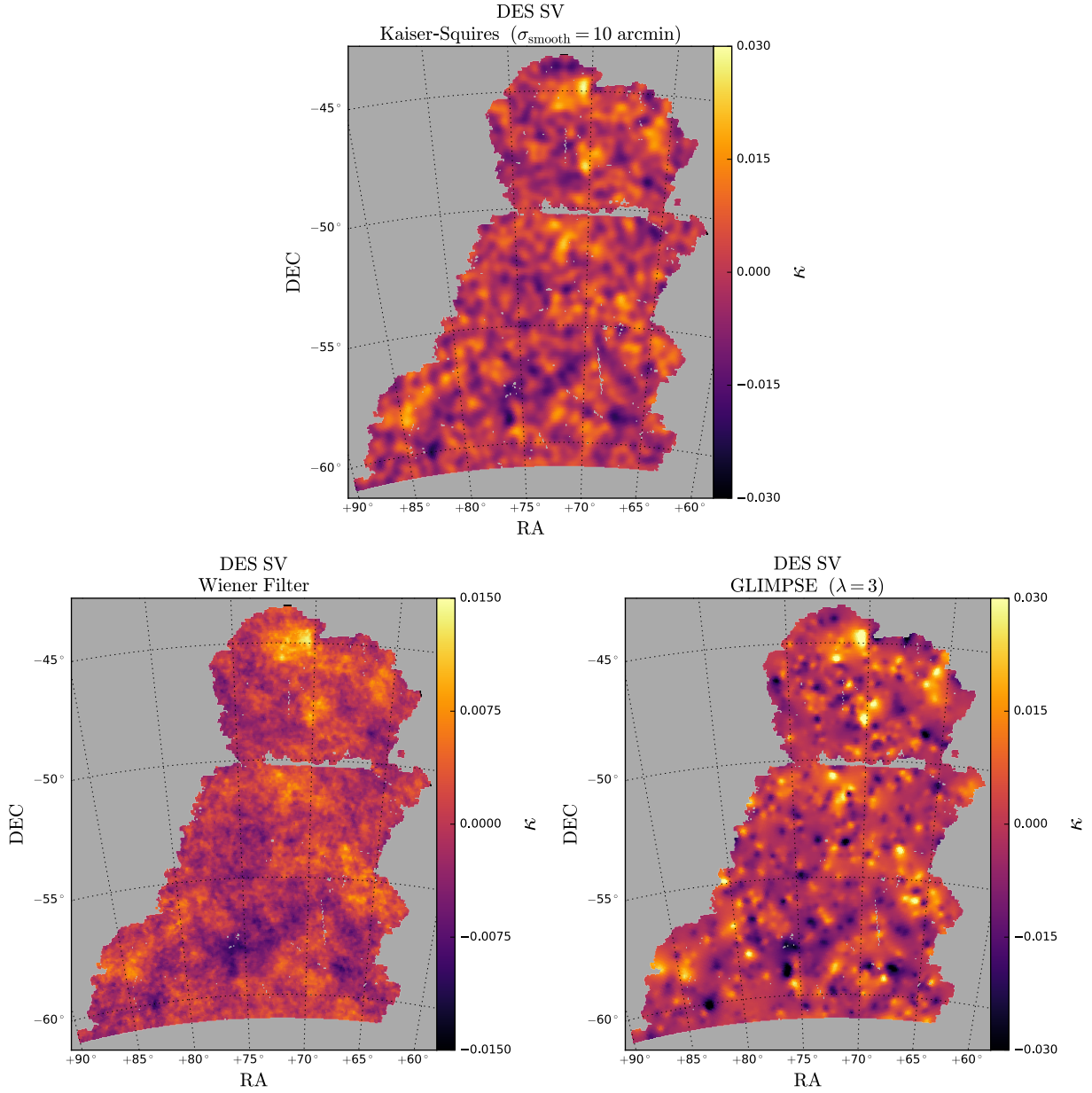
Figure 3. The convergence (κ) map reconstructions using the DES SV shear data with the three different methods. *Top panel*: Kaiser-Squires reconstruction with a smoothing scale σ_{smooth} = 10 arcmin. *Right panel*: The GLIMPSE reconstruction with a regularisation parameter λ = 3.0. Both tuning parameters were chosen to maximise the Pearson correlation coefficient r when tested on simulations (See Sec. 4.1). *Left panel*: The Wiener filter reconstruction. **Note** that the colour scale for the Wiener filter is less than that for the other reconstructions, as the pixel values are closer to zero.

The Pearson correlation coefficient, r, between the pixels' true convergence, κ^{truth}, and the reconstruction, κ^{recon}, is given by

$$r = \frac{\sum_{i=1}^{n}(\kappa_i^{truth} - \bar{\kappa}^{truth})(\kappa_i^{recon} - \bar{\kappa}^{recon})}{\sqrt{\sum_{i=1}^{n}(\kappa_i^{truth} - \bar{\kappa}^{truth})^2}\sqrt{\sum_{i=1}^{n}(\kappa_i^{recon} - \bar{\kappa}^{recon})^2}} \ , \quad (32)$$

where the summations are over all pixels i in the map and $\bar{\kappa}$ is the mean convergence in the map.

In the left panels of figure 5, the Pearson r value from 10 simulations is plotted for varying tuning parameters. Almost all of the simulations and also their mean have a maximal Pearson r value at σ_{smooth} = 10.0 arcmin for Kaiser-Squires and at λ = 3.0 for GLIMPSE.

Table 1 presents the mean value from the 10 simulations, where the tuning parameter is chosen to maximise r when relevant. All methods show good correlation with the underlying true convergence. Both the Wiener filter and GLIMPSE have the same highest value of r = 0.37, 12% higher than Kaiser-Squires.

Note that the Pearson correlation coefficient as pre-
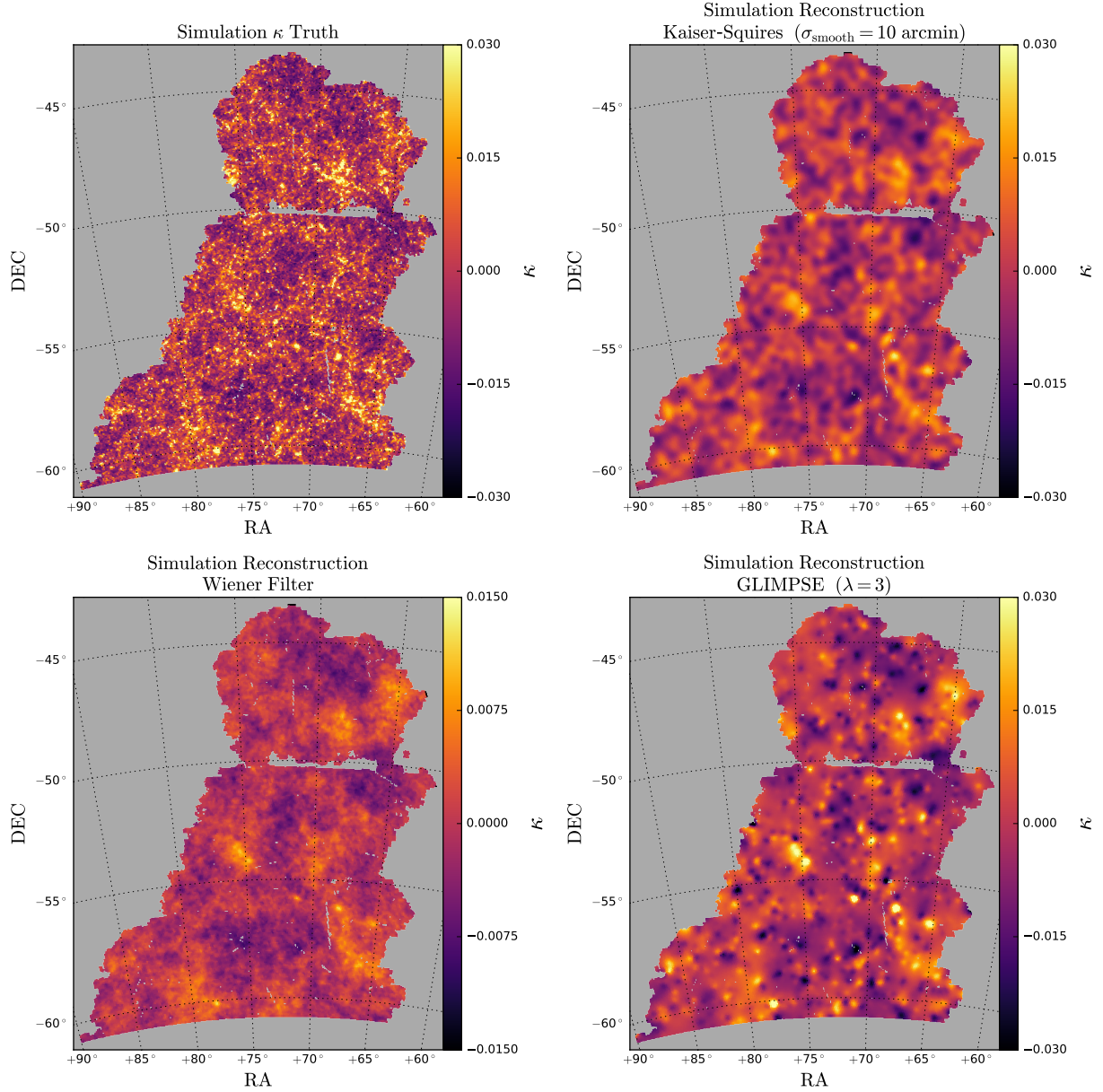
Figure 4. The *top left* panel is an example of a true convergence map, κ^{recon}, from simulation. The *top right* panel is the Kaiser-Squires reconstruction with a smoothing scale $\sigma_{smooth} = 10$ arcmin. The *bottom right* panel is the GLIMPSE reconstruction with regularisation parameter $\lambda = 3.0$. Both tuning parameters were chosen to maximise the Pearson correlation coefficient r when tested on simulations (see Sec. 4.1). The *bottom left* panel is the Wiener filter reconstruction. **Note** that the colour scale for the Wiener filter is less than that for the other reconstructions, as the pixel values are closer to zero.

sented in equation 32 is invariant under a rescaling of the reconstruction. Despite the Wiener filter reconstruction having values closer to zero, the Wiener filter maps still have good correlation to the truth. This second aspect is addressed in Sec. 4.3 and in the second column of table 1.

4.2 Pixel Residuals

The difference between the true convergence from simulation and the reconstruction in pixel i is defined as

$$\Delta \kappa_i = \kappa_i^{truth} - \kappa_i^{\mathrm{recon}} \ . \tag{33}$$

We define the root-mean-square error (RMSE) as

$$\mathrm{RMSE}(\kappa^{\mathrm{truth}}, \kappa^{\mathrm{recon}}) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \Delta \kappa_i^2} \tag{34}$$

Kaiser-Squires Reconstruction

GLIMPSE Reconstruction
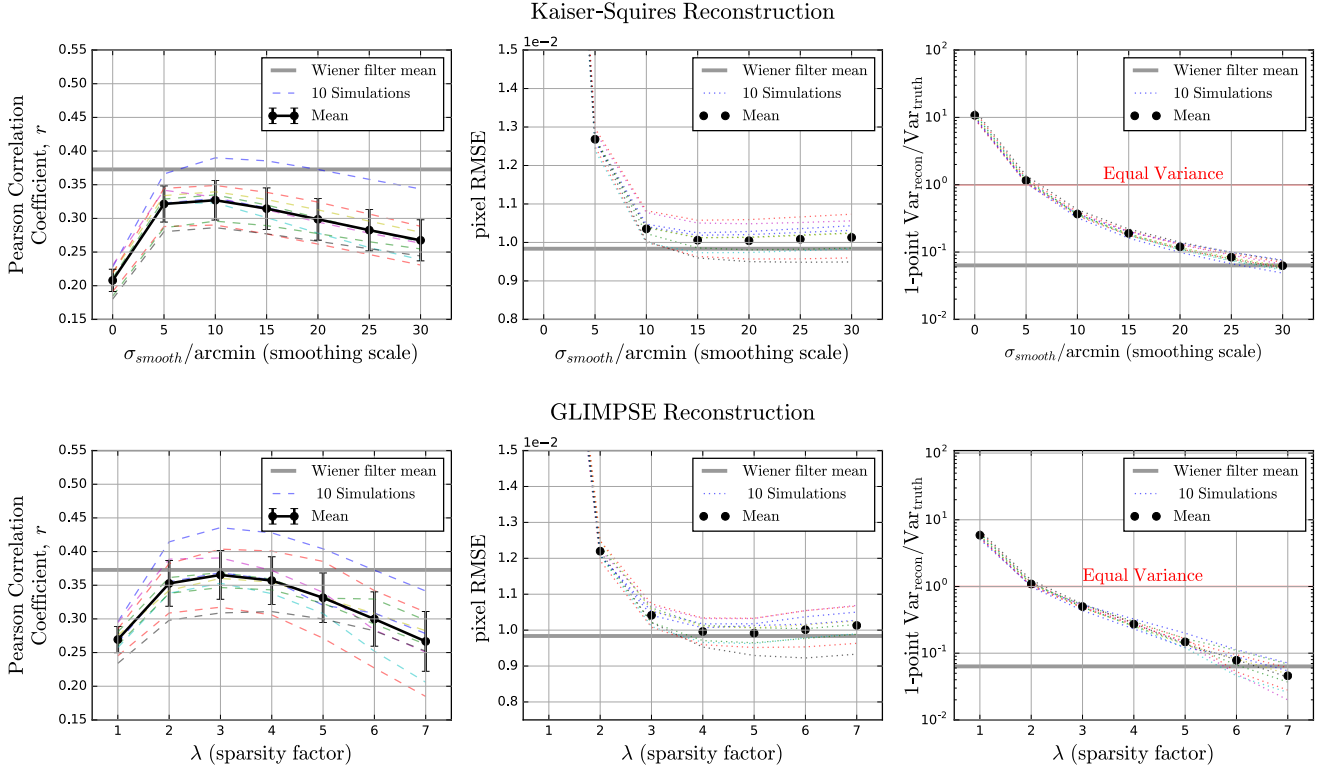
Figure 5. Kaiser-Squires (top) and GLIMPSE (bottom). Three different statistics comparing the true κ map and the reconstruction with 10 simulations. *Left panel*: The Pearson correlation coefficient, r (equation 32). The errorbar on the mean is the standard deviation of the sample. The better the reconstruction, the higher the value of r. *Middle panel*: The lower the pixel RMSE (equation 34), the better the reconstruction. *Right panel*: Ratio of variances between the 1-point distribution of the pixels in the reconstruction and pixels in the true map (equation 35).

where n is the number of pixels.

A smaller value of RMSE for a given method implies a better reconstruction according to this metric. It is this RMSE that the Wiener filter attempts to minimise using a linear filter, as defined in equation 18, by using an assumed signal covariance $\langle \kappa \kappa^{\dagger} \rangle$ (see Sec. 2.3.1).

The centre panel of figure 5 shows that increasing the smoothing scale, σ_{smooth}, for Kaiser-Squires or the regularisation parameter, λ, for GLIMPSE initially reduces the pixel RMSE, but increased filtering contributes little beyond $\sigma_{smooth} = 10.0$ arcmin for Kaiser-Squires or $\lambda = 3.0$ for GLIMPSE.

The smallest mean pixel RMSE is 10.0×10^{-3} for Kaiser-Squires and 9.9×10^{-3} for GLIMPSE. The Wiener filter, whose smoothing is constrained by the prior on C_ℓ and which therefore cannot be tuned, has a pixel RMSE of 9.4×10^{-3}.

4.3 Pixel 1-Point Variance

The 1-point distribution can be thought as a histogram of the pixel values. Figure 6 shows an example of such a histogram (derived from the simulated truth map and reconstructions of figure 4).

The mean of this distribution is unconstrained by weak lensing, due to an integration constant in equation 7. The variance of the 1-point distribution is increased compared to the underlying truth due to shape noise in the un-
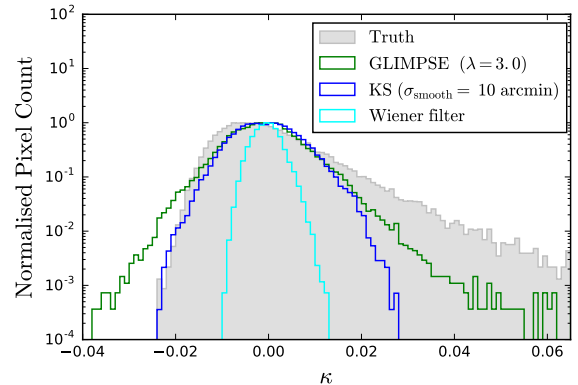
Figure 6. Pixel histograms (1-point distributions) for various map reconstructions from the simulated data shown in figure 4. The histograms are normalised such that the largest value of each is equal to one. The ratio of the variance between the reconstructions and the truth is presented in table 1.

smoothed Kaiser-Squires reconstruction. A reconstruction method would aim to reduce the variance of the 1-point pixel distribution to match that of the underlying truth.

Table 1. The centre column gives the average Pearson correlation coefficient r (equation 32) between κ^{truth} and κ^{recon} from 10 simulations. The choices of σ_{smooth} = 10 arcmin and λ = 3.0 maximise the Pearson r value. The right column gives the ratio of the pixel variance between κ^{recon} and κ^{truth} (equation 36).

Method	Pearson r	Variance Ratio
KS (σ_{smooth} = 10 arcmin)	0.33	3.7×10^{-1}
Wiener filter	0.37	6.3×10^{-2}
GLIMPSE (λ = 3.0)	0.37	5.0×10^{-1}

We define the estimate of the variance of the 1-point distributions of the truth or reconstructed κ as

$$\text{Var}_{\text{truth}} = \frac{1}{n-1} \sum_{i=1}^{n} (\kappa_i^{\text{truth}} - \bar{\kappa}^{\text{truth}})^2$$
$$\text{Var}_{\text{recon}} = \frac{1}{n-1} \sum_{i=1}^{n} (\kappa_i^{\text{recon}} - \bar{\kappa}^{\text{recon}})^2 , \quad (35)$$

where the notation matches equation 34. The ratio of these variances is given by

$$\frac{\text{Var}_{\text{recon}}}{\text{Var}_{\text{truth}}} = \frac{\sum_{i=1}^{n} (\kappa_i^{\text{recon}} - \bar{\kappa}^{\text{recon}})^2}{\sum_{i=1}^{n} (\kappa_i^{truth} - \bar{\kappa}^{truth})^2} , \quad (36)$$

The closer this value is to 1, the better the variance of the pixel distribution matches the truth. Using 10 simulations we can calculate this quantity for different reconstruction methods (and at different smoothing scales or λ regularisation values where relevant).

In figure 5 the right panel shows the result of this test for GLIMPSE and Kaiser-Squires. Both methods show a pixel distribution that has too high variance for insufficient filtering, and too low variance for over-filtering. For Kaiser-Squires, the ratio is closest to 1 at a smoothing scale of σ_{smooth} = 5 arcmin. For GLIMPSE, the ratio is closest to 1 at a sparsity regularisation value of λ = 2.

Both of these reconstruction methods have a matching variance at a smoothing parameter value less than that which maximises the Pearson correlation coefficient r. If one chose this parameter to maximise the Pearson r value, such that λ = 3 and σ_{smooth} = 10 arcmin, a good reconstruction should also have the ratio of the variances as close to 1 as possible.

The right column of table 1 gives the mean variance ratio from 10 simulations with the different methods. The choice of λ = 3.0 and σ_{smooth} = 10 arcmin are the tuning parameters that maximise the Pearson r value for GLIMPSE and Kaiser-Squires respectively. Though GLIMPSE and the Wiener filter reconstructions both have the same Pearson r value, the variance of the pixel values of the Wiener filter is much lower with respect to the underlying truth than is the case for GLIMPSE. This can also be seen in the reconstructions of figure 4, where the Wiener filter pixel values are closer to zero than the simulated true convergence.

The histogram of figure 6 shows, for one single example, the distributions matching what the results of the second column of table 1 describe. GLIMPSE outperforms the other methods at matching the variance of the underlying truth, however it still falls short. Also, all methods, including GLIMPSE, have distributions which are symmetric, unlike the asymmetric, heavy-tailed distribution of the true κ values.

Though GLIMPSE reconstructs maps with the 1-point distribution variance closest to the truth, it is also the only method to have convergence values dropping below the truth. These unphysical "negative peaks" can also be seen in the map reconstructions from data (figure 3) and from simulated catalogues (figure 4), and are likely to come from enforcement of sparsity for positive and negative wavelets equally. The physical motivation for GLIMPSE comes from a density field of superimposed halos. Though there should be no negative halos, negative wavelets are included to map the underdense regions, clearly at the expense of producing these very negative regions.

4.4 Phase reconstruction

The summation over all m modes at each ℓ multipole in the angular power spectrum (equation 24) loses all phase information; only the magnitudes are retained. This phase information corresponds to the spatial distribution of anisotropies. As the phases are dependent on the physical underlying structure, they contain information beyond what can be gained by 2-point statistics. Their retention is a well-motivated, desired property of a mass mapping reconstruction.

Inspired by Chapman et al. (2013), who use phases to test the reconstruction after foreground removal from simulated Epoch of Reonization 21-cm maps, we use the phase residual as a metric of success between our three methods.

The phase difference between the true map and the reconstruction is defined as

$$\Delta\theta_{\ell m} = \theta_{\ell m}^{truth} - \theta_{\ell m}^{\text{recon}}$$
$$= \arg(a_{\ell m}^{truth}) - \arg(a_{\ell m}^{\text{recon}}) , \quad (37)$$
$$\text{where} \quad \arg(z) = \arctan\left(\frac{\text{Im}(z)}{\text{Re}(z)}\right) .$$

A small phase difference $\Delta\theta_{\ell m}$ between the truth and the reconstruction implies that the phase has been well reconstructed. For random variables drawn from a Gaussian distribution, this would correspond to a small standard deviation. Here, however, a Gaussian distribution would be an inappropriate choice as it assumes the data are defined on an unbounded Euclidean space.

The two dimensional data space of phase pairs, $\{\theta_{\ell m}^{truth}, \theta_{\ell m}^{\text{recon}}\}$, is a torus, T^2, and the projected data space of the phase difference, $\Delta\theta_{\ell m}$, is a circle, S^1. On a circle, the maximum entropy, least informative, distribution for specified mean and variance is the von Mises (Jammalamadaka & Sengupta 2001), which in one dimension is given by

$$Pr(\Delta\theta_{\ell m}|C, \mu) = \frac{1}{2\pi I_0(C)} \exp\left[C \cos(\Delta\theta_{\ell m} - \mu)\right] , \quad (38)$$

where I_0 is the modified Bessel function of order 0, and C is a concentration parameter. For $\mu = 0$, a large concentration parameter (analogous to $1/\sigma^2$) would correspond to a small dispersion in the phase reconstruction error. The aim

Kaiser-Squires

GLIMPSE
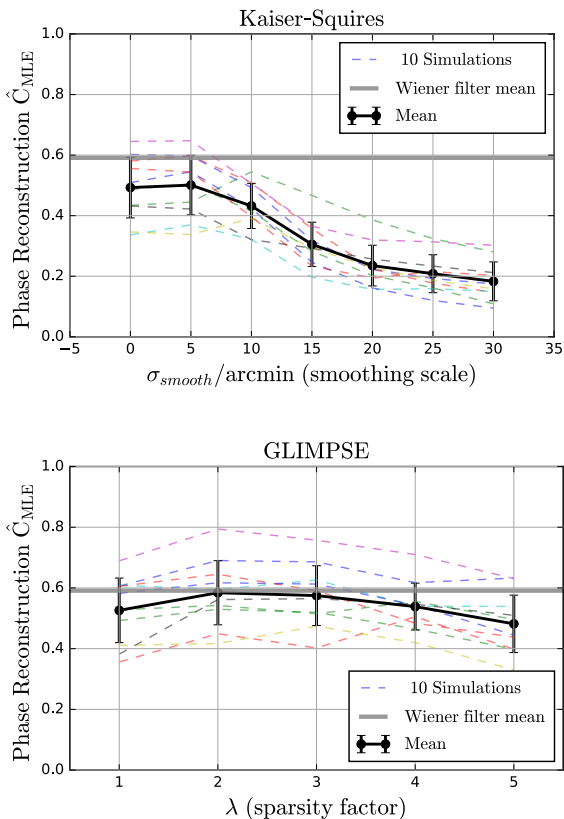
Figure 7. The maximum likelihood value of the concentration of the phase residual distribution, $\hat{C}_{\rm MLE}$, as described by equation 39. The $\hat{C}_{\rm MLE}$ values are shown for 10 different simulations and with Kaiser-Squires (*top panel*) at varying smoothing scale, σ_{smooth}, and GLIMPSE (*bottom panel*) at varying regularisation parameter λ. The phase reconstruction is best for $\sigma_{smooth} = 5$ arcmin and $\lambda = 3.0$ respectively.

is therefore to compare the inferred value of the concentration, C, between different mass mapping methods, with a larger value of C implying a better phase reconstruction.

By assuming that the error on the phase reconstruction is independent between phases, we can say that the phase differences, $\Delta\vec{\theta}$, are independent and identically distributed random variables, with a likelihood distribution given by

$$
\begin{aligned}
Pr(\Delta\vec{\theta}|C,\mu) &= \prod_{\ell m} \frac{1}{2\pi I_0(C)} \exp\left[C\cos(\Delta\theta_{\ell m} - \mu)\right] \\
&= \frac{1}{[2\pi I_0(C)]^n} \exp\left[C\sum_{\ell m}\cos(\Delta\theta_{\ell m} - \mu)\right] .
\end{aligned}
\tag{39}
$$

As only the relative values of C are needed to compare different mass mapping methods, the full posterior distribution is not required. Additionally, any reasonable prior distribution, $Pr(C)$, will be either flat or monotonically decreasing above zero, so the ranking of maps by the largest maximum likelihood value or maximum posterior value of C will be identical. For the purposes of this comparison the simpler maximum likelihood estimate, $\hat{C}_{\rm MLE}$, will therefore do.

We calculate the maximum likelihood values of μ and C by taking the spherical harmonic transform of our `HEALPix` map to recover the $a_{\ell m}$ coefficients up to $\ell_{max} = 1024$, calcu-

Table 2. The mean over 10 simulations of the von Mises concentration maximum likelihood estimate, $\hat{C}_{\rm MLE}$, from phase residuals (equation 39).

Method	Phase reconstruction Concentration $\hat{C}_{\rm MLE}$
KS ($\sigma_{smooth} = 5$ arcmin)	0.501
Wiener filter	0.591
GLIMPSE ($\lambda = 3.0$)	0.584

lating the phase residual as defined by equation 37 between the truth and the reconstruction for each coefficient, and then maximising the likelihood (equation 39). The maximisation is performed using the `scipy` package BFGS algorithm (Byrd et al. 1995, Zhu et al. 1997, Morales & Nocedal 2011), using 3 random initialisation values to test for robustness.

Figure 7 show the results for the phase reconstruction from 10 simulations using Kaiser-Squires and GLIMPSE with varying tuning parameters. For Kaiser-Squires the mean phase reconstruction value, $\hat{C}_{\rm MLE}$, is maximised at $\sigma_{smooth} = 5.0$ arcmin. For larger smoothing scales the phase reconstruction quality drops, as phase information is lost. For GLIMPSE the mean phase reconstruction value, $\hat{C}_{\rm MLE}$, is maximised at $\lambda = 3.0$. The maximum value of $\hat{C}_{\rm MLE}$ is not particularly pronounced, and the $\hat{C}_{\rm MLE}$ values are quite stable over a range of λ.

Table 2 presents the mean values of $\hat{C}_{\rm MLE}$ with the best tuning parameters for the three map reconstruction methods. Both GLIMPSE and the Wiener filter do much better than Kaiser-Squires for reconstructing the phases. Though the variance from these 10 different simulations is large, the Wiener filter does slightly better than GLIMPSE, as can be seen in figure 7.

4.5 Peak Statistics

Peak statistics are a promising method for inferring cosmological parameters from data, as they access information beyond what can be inferred from 2-point correlation functions. Unlike higher order correlation functions, such as the bispectrum, peak statistics are inherently high signal-to-noise. They also probe the highly non-linear regions, where non-Gaussianity is greatest. The effect of masking is trivially taken into account by applying the identical mask to the suite of simulations used to construct a likelihood.

We cannot truly test which mass mapping method best constrains cosmology with the statistics of density peaks without fully deriving the posterior probability distributions of cosmological parameters. It is possible to test which method returns peaks which are distinguishable from noise and at which convergence values. Distinguishing a large number of peaks from noise at high values of κ would mean the map is reconstructing the non-linear regions well.

For a given convergence map, we can define a function, $n(\kappa)$, that gives the number of peaks as a function of convergence. For a given mass reconstruction method we can compare the peaks in reconstructions from simulated data with the peaks in reconstructions from catalogues of "randoms", with shape noise but no weak lensing shear signal (equiva-
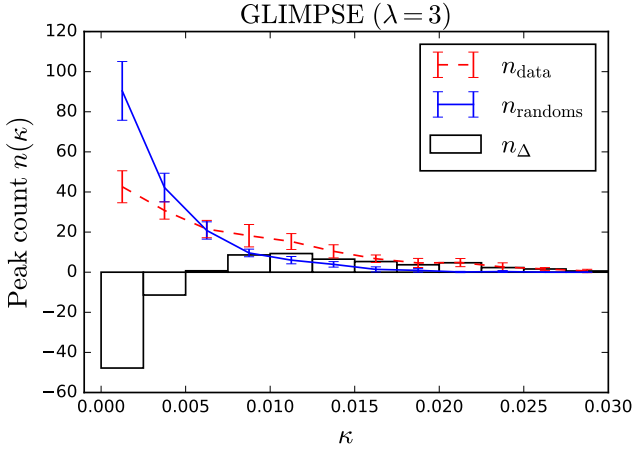
GLIMPSE ($\lambda = 3$)

Figure 8. The mean n_{data}, n_{randoms} and n_Δ functions with 10 simulated data catalogues and 10 random catalogues from GLIMPSE ($\lambda = 3$) reconstructions. n_Δ is defined in equation 40. Errorbars are standard deviation sample estimates from the 10 simulations, and are consistent with Poissonian noise.

lent to $\gamma = 0$ in equation 9). If a given map from data or from a simulated catalogue has the same $n(\kappa)$ as the random catalogues, then the mass mapping method used has been useless for peak statistics. On the other hand, if the map from data or simulation has a very different $n(\kappa)$ function to that from the reconstruction from the random catalogue, then the map reconstruction method has recovered "true", physical κ peaks.

In the DES SV cosmology constraints from peak statistics, Kacprzak et al. (2016) use this difference as the data vector used to constrain cosmology,

$$n_\Delta(\kappa_i) = n_{\text{data}}(\kappa_i) - n_{\text{randoms}}(\kappa_i) \ . \tag{40}$$

This function is far from zero at a given κ if there is a large difference between the number of peaks counted in maps reconstructed (a) from data and (b) from random catalogues.

It is reasonable to believe that the number of peaks, $n(\kappa_i)$, in the ith bin, κ_i, is drawn from a Poisson distribution. The difference between two Poissonian random variables follows the Skellam distribution. Using this distribution, we expect the difference in the number peaks in maps from data and from random catalogues to have a mean given by

$$\mu_\Delta(\kappa_i) = \mu_{\text{data}}(\kappa_i) - \mu_{\text{randoms}}(\kappa_i) \ , \tag{41}$$

and a variance given by

$$\sigma_\Delta^2(\kappa_i) = \mu_{\text{data}}(\kappa_i) + \mu_{\text{randoms}}(\kappa_i). \tag{42}$$

We can therefore define a peak signal-to-noise estimate

$$\text{SNR}(\kappa_i) = \frac{\mu_\Delta(\kappa_i)}{\sqrt{\sigma_\Delta^2(\kappa_i)}} \ . \tag{43}$$

Figure 8 shows n_{data}, n_{randoms}, and n_Δ from GLIMPSE ($\lambda = 3$) from 10 simulations and 10 random catalogues.

Here we define a peak as a local maxima in the HEALPIX map. Across different methods and smoothing parameters, the predicted variance from equation 42 matches well with the estimated sample variance, verifying that the peak distribution is indeed Poissonian for a given κ.

Figure 9 shows the peak signal-to-noise (SNR) estimates from 10 simulations and from 10 random catalogues as a function of κ and smoothing scale, for Kaiser-Squires, or λ, for GLIMPSE. As the peaks in the maps from data have higher convergence values than those from random catalogues, the SNR(κ) function is negative for low values of κ.

In the figures, the GLIMPSE reconstruction gives better signal-to-noise estimates on the peaks than does the Kaiser-Squires reconstruction. For Kaiser-Squires, the largest positive and negative signal-to-noise values are 1.52 and -1.28. For GLIMPSE, the largest positive and negative are signal-to-noise values of 2.32 and -13.72. For the Wiener filter these values are 4.20 and -5.41.

The Wiener filter therefore has the highest signal-to-noise of the peak function $n_\Delta(\kappa)$, though the κ values of these peaks are very low. As can be seen in the reconstruction from the SV data in figure 3, the pixel values of the Wiener filter are much closer to zero. This is reflected in the peak statistic signal-to-noise values. In the left panel of figure 9, the Wiener filter detects negligibly few peaks with $\kappa > 0.0125$, whereas GLIMPSE detects peaks with positive signal-to-noise up to higher values of κ. It is at these high values where the non-Gaussian information due to non-linear structure formation can be probed.

4.6 Foreground Clusters

Comparisons with foreground clusters of galaxies is an independent test of the mass map reconstructions, as it uses data (unlike our tests on simulations).

In figure 10 the redMaPPer clusters described in Sec. 3.2 are overlaid on the DES SV κ map reconstructions shown in figure 3. The maps show good spatial correlation between the locations of the clusters and the κ peaks in the map.

The size of a cluster marker is the effective lensed cluster richness λ_{RM}^{eff}, rather than the redMaPPer cluster richness. This concept is adapted from the definition of κ_g presented in Chang et al. (2015). For a given cluster, this is defined as

$$\lambda_{RM}^{eff} = \frac{p(z)\omega(z)}{a(z)} \times \lambda_{RM} \times \frac{\langle \lambda_{RM} \rangle}{\langle \lambda_{RM}^{eff} \rangle} \ , \tag{44}$$

where z is the redshift of the cluster, $p(z)$ is the lensing efficiency at the location of the cluster (see figure 1), and $\omega(z)$ is the comoving distance to the cluster (so that the first term matches the integrand of equation 6). The final term normalises the mean, where $\langle \lambda_{RM} \rangle$ is the average richness over all galaxy clusters. The effective lensed cluster richness gives the richness as "seen" by the lensing effect, where clusters at the peak of the lensing efficiency should contribute more to the κ map. We therefore calculate the correlation between λ_{RM}^{eff} for each cluster and the reconstructed κ value at the cluster centre.

This method does not take into account multiple clusters overlapping in a given line of sight. In figure 10, many small clusters overlap on large peaks in the reconstructed κ
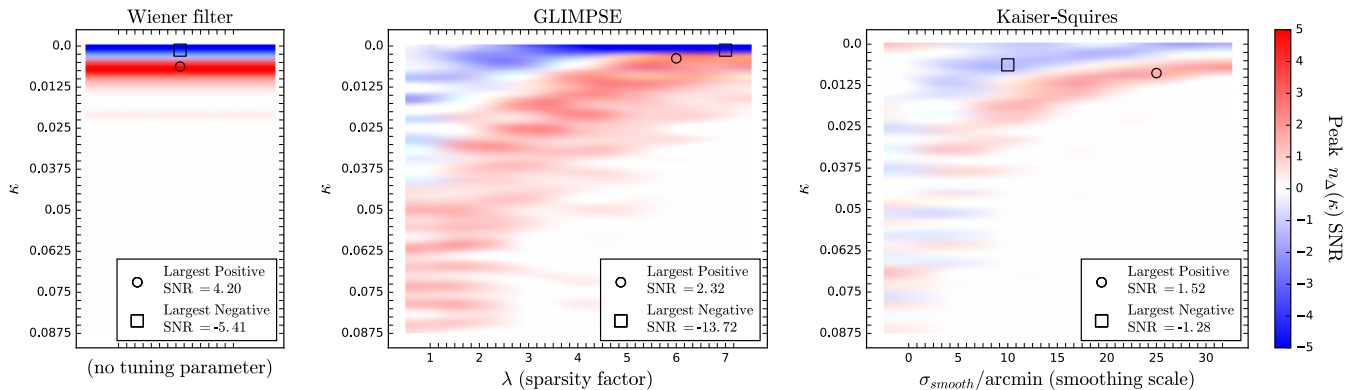
Figure 9. The estimated signal-to-noise (SNR) of $n_\Delta(\kappa)$ (equation 40) using 10 simulated data catalogues and 10 random catalogues with the three different mass mapping methods. Kaiser-Squires and GLIMPSE maps can be tuned by their respective parameters. The width of the left (Wiener filter) panel is purely nominal; it does not actually have a "flat structure", just no parameter to tune. (SNR interpolated with a sinc function.)

Table 3. The Pearson correlation coefficient value, r, between effective richness, λ_{RM}^{eff}, of the foreground redMaPPer clusters and the reconstructed convergence map at the location of each galaxy cluster.

Method	redMaPPer Cluster λ_{RM}^{eff} Pearson r
KS (σ_{smooth} = 10 arcmin)	0.116
Wiener filter	0.129
GLIMPSE (λ = 3.0)	0.152

map. The naive one-to-one correspondence between cluster and κ would mistake this for an excess of κ in the reconstruction. However, all methods will suffer equally from this assumption. A more thorough treatment of this overlapping effect is left for future work.

Table 3 presents the Pearson correlation coefficient r between the λ_{RM}^{eff} value of each cluster and the κ^{recon} value at the corresponding pixel. The tuning parameters for Kaiser-Squires and GLIMPSE are chosen to maximise the Pearson correlation coefficient r between the reconstruction and the truth from simulations (see Sec. 4.1).

Though both GLIMPSE and the Wiener filter take into account the noise and the mask in the data, and therefore do better than Kaiser-Squires, the GLIMPSE reconstructions show higher correlation with the effective richness of the foreground clusters than do the Wiener filter reconstructions. This is no surprise, as GLIMPSE is expected to do better at reconstructing non-Gaussian κ, which would correspond to the non-linear matter structures in which clusters of galaxies form.

5 CONCLUSIONS

In this work we have presented convergence map reconstructions using the public DES SV shear data with three different methods: Kaiser-Squires, Wiener filter, and GLIMPSE.

Kaiser-Squires is a simple inversion from shear to convergence, whereas the Wiener filter and GLIMPSE use prior knowledge about the true convergence to help regularise the reconstruction and to reduce the effects of noise and missing data. The Wiener filter is a Bayesian MAP estimate if the signal and noise are Gaussian and the respective covariance matrices are known. The GLIMPSE method enforces a sparsity-promoting l_1 norm in a wavelet space where the wavelets represent positive, isotropic, quasi-spherical objects well. GLIMPSE is therefore expected to do well at reconstructing non-linear structures. The Wiener filter and GLIMPSE therefore aim to reconstruct different regimes: the linear and non-linear density field.

The three methods were applied to realistic simulations of the DES SV shear data, for which an underlying true convergence is known. Using these simulations we are also able to tune the Kaiser-Squires smoothing scale, σ_{smooth}, and the GLIMPSE sparsity regularisation parameter, λ.

With these simulations we measure the Pearson correlation coefficient, r, between the truth and the reconstruction with different methods. Compared to the Kaiser-Squires reconstructions we find a 12% improvement in Pearson correlation with both the Wiener filter and GLIMPSE. The tuning parameters of σ_{smooth} = 10 arcmin for Kaiser-Squires and λ = 3 for GLIMPSE maximise the Pearson correlation. We also measure the variance of the 1-point distribution of the reconstructed convergence. The Wiener filter suppresses the variance to 6.3 % of the truth, Kaiser-Squires to 37 % and GLIMPSE to 50 % of the truth. The tunable parameters here were those which maximised the Pearson correlation with the truth.

A large motivation for creating these maps is to reconstruct the convergence while still retaining the non-Gaussian information (which cannot be accessed with 2-point statistics such as the power spectrum). As such, we test the reconstruction of the harmonic phases, which is averaged out in the power spectrum, and the signal-to-noise of a peak statistic data vector, which is a popular probe of non-Gaussian information. The phase residuals between the truth and the reconstruction have the highest von Mises concentration
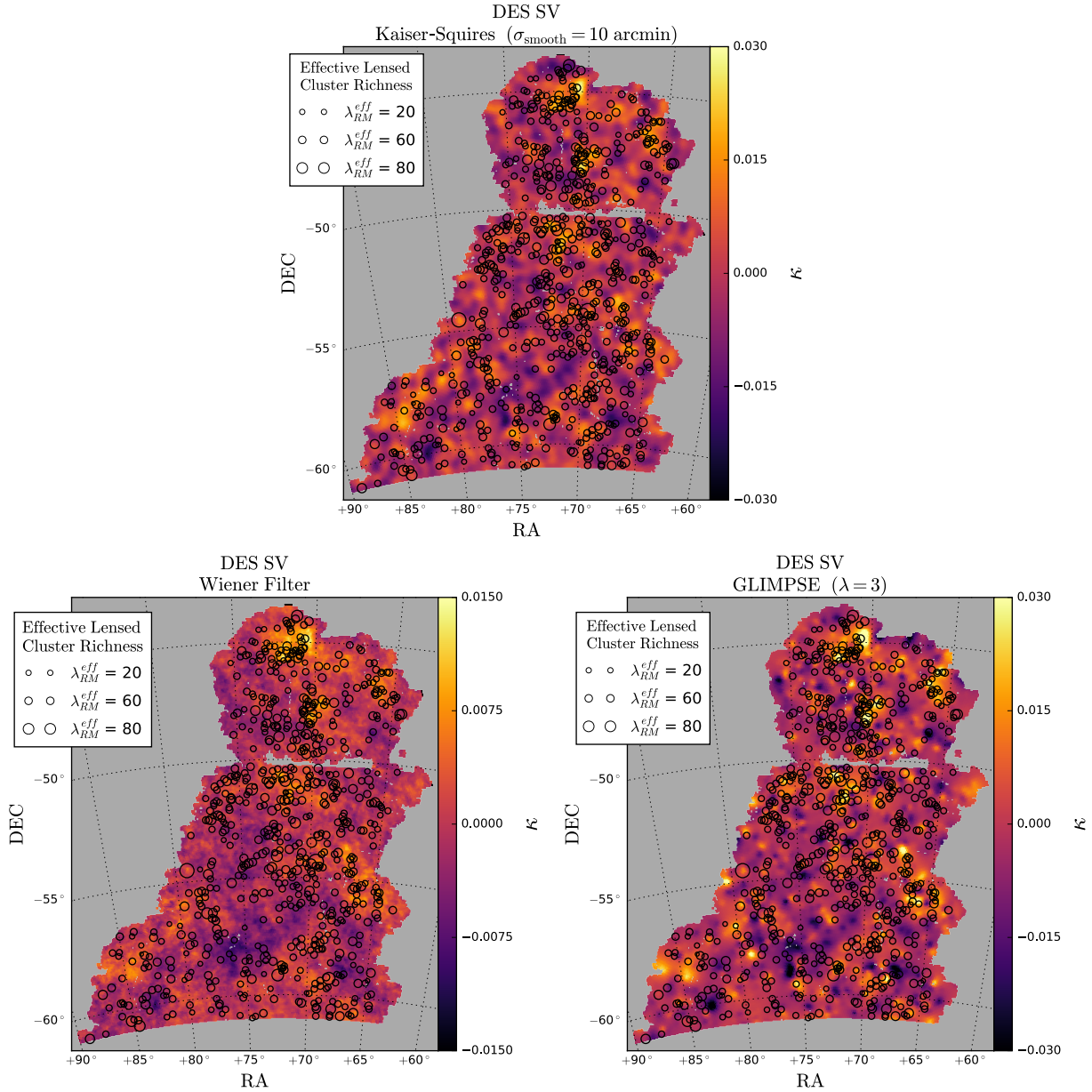
Figure 10. The mass map reconstruction from DES SV shear data with the three different methods, as presented in figure 3), with the locations of redMaPPer clusters overlaid. The size of the cluster marker is the effective richness of the cluster, as defined in equation 44. **Note** that the colour scale for the Wiener filter is less than that for the other reconstructions, as the pixel values are closer to zero.

with the Wiener filter ($\hat{C}_{\text{MLE}} = 0.591$), with the GLIMPSE reconstruction performing comparably ($\hat{C}_{\text{MLE}} = 0.584$). Both methods outperformed the Kaiser-Squires reconstruction ($\hat{C}_{\text{MLE}} = 0.501$).

With realistic data vectors for peak statistics generated from simulations, the maximum signal-to-noise value was increased by a factor of 3.5 for the Wiener filter and by a factor of 9 for GLIMPSE, compared to Kaiser-Squires. The signal-to-noise of the peak statistic data vector ($n_\Delta(\kappa)$) is shown in figure 9, where GLIMPSE has significant signal-to-noise with high convergence peaks, where non-linearities in the underlying density field are highest. We predict these

high value peaks are most useful for constraining cosmology beyond Gaussianity. In order to constrain cosmology with these different reconstruction methods, realistic simulations with different cosmological parameters or models must be used and the same reconstruction method should be applied to the simulations and data. As seen from our results, different reconstruction methods can produce convergence maps with different properties.

Finally, we switched from using simulations to instead using real observations (DES SV data). Here we measured the correlation between the reconstructed maps and the effective richness of the foreground redMaPPer clusters (this

is the cluster richness as "seen" by the lensing effect). Table 3 shows the results. Compared with Kaiser-Squires, the Wiener filter shows a 18% increase and GLIMPSE shows a 32% increase in correlation. This demonstrates with independent, cosmological data the ability of the methods to reconstruct non-linear structures.

The metrics we have used for comparing the three reconstruction methods are generic, and they have been inspired by recent applications of weak lensing mass maps to cosmological studies (e.g. Chang et al. 2016, Kacprzak et al. 2016). These metrics may not be optimal for evaluating every application of mass maps. Future studies can compare the efficiency of the three and other methods in end-to-end analyses; for example, with the estimation of cosmological parameters or identification of galaxy clusters.

Applying the Wiener filter and GLIMPSE methods to the DES Year 1 (Y1) shear catalogue would require extensions of the methods to account for the curved sky at large angular scales. The Y1 data covers $\approx 1500 \deg^2$ and contains $\approx 34,800,000$ galaxies, so is a large increase in data volume from DES SV. This modification has already been done with an extension of Kaiser-Squires to the sphere by Chang et al. (2017) for the Y1 DES data. These extensions would also be useful for the upcoming $\approx 5000 \deg^2$ DES Y3 shear catalogue.

Of future interest would be to use the Wiener filter or GLIMPSE convergence maps for scientific results, as we have shown that they reconstruct the convergence better than Kaiser-Squires according to many different metrics.

We have made our map reconstructions (as shown in figure 3) available at **https://github.com/NiallJeffrey/DES_SV_mass_maps**.

ACKNOWLEDGEMENTS

manuscript, or allow others to do so, for United States Government purposes.

We use the visualization software package `SkyMapper`[4] for the map figures.

REFERENCES

Alsing J., Heavens A., Jaffe A. H., Kiessling A., Wandelt B., Hoffmann T., 2016, MNRAS, 455, 4452

Alsing J., Heavens A., Jaffe A. H., 2017, MNRAS, 466, 3272

Amendola L., et al., 2016, preprint, (arXiv:1606.00180)

Bacon D. J., Goldberg D. M., Rowe B. T. P., Taylor A. N., 2006, MNRAS, 365, 414

Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291

Becker M. R., 2013, MNRAS, 435, 115

Benítez N., 2000, ApJ, 536, 571

Bonnett C., et al., 2016, Phys. Rev. D, 94, 042005

Busha M. T., Wechsler R. H., Becker M. R., Erickson B., Evrard A. E., 2013, in American Astronomical Society Meeting Abstracts #221. p. 341.07

Byrd R. H., Lu P., Nocedal J., Zhu C., 1995, SIAM Journal on Scientific Computing, 16, 1190

Chang C., et al., 2015, Physical Review Letters, 115, 051301

Chang C., et al., 2016, MNRAS, 459, 3203

Chang C., et al., 2017, preprint, (arXiv:1708.01535)

Chapman E., et al., 2013, MNRAS, 429, 165

Clerkin L., et al., 2017, MNRAS, 466, 1444

Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, AJ, 132, 926

Cooray A., Hu W., 2001, ApJ, 548, 7

DES Collaboration et al., 2017, preprint, (arXiv:1708.01530)

Dark Energy Survey Collaboration et al., 2016, MNRAS, 460, 1270

Dietrich J. P., Hartlap J., 2010, MNRAS, 402, 1049

Elsner F., Wandelt B. D., 2013, A&A, 549, A111

Flaugher B., et al., 2015, AJ, 150, 150

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759

Jammalamadaka S. R., Sengupta A., 2001, Topics in circular statistics. Series on multivariate analysis ; v. 5, World Scientific, River Edge, N.J.

Jasche J., Lavaux G., 2015, MNRAS, 447, 1204

Kacprzak T., et al., 2016, MNRAS, 463, 3653

Kaiser N., Squires G., 1993, ApJ, 404, 441

Kerscher M., Schmalzing J., Buchert T., 1996, in Coles P., Martinez V., Pons-Borderia M.-J., eds, Astronomical Society of the Pacific Conference Series Vol. 94, Mapping, Measuring, and Modelling the Universe. p. 247 (arXiv:astro-ph/9512105)

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

Kilbinger M., et al., 2013, MNRAS, 430, 2200

LSST Science Collaboration et al., 2009, preprint, (arXiv:0912.0201)

Lahav O., Fisher K. B., Hoffman Y., Scharf C. A., Zaroubi S., 1994, Astrophysical Journal Letters, 423, L93

Lanusse F., Starck J.-L., Leonard A., Pires S., 2016, A&A, 591, A2

Leistedt B., Mortlock D. J., Peiris H. V., 2016, MNRAS, 460, 4258

Leonard A., Lanusse F., Starck J.-L., 2014, Monthly Notices of the Royal Astronomical Society, 440, 1281

Lin C.-A., Kilbinger M., 2015, A&A, 576, A24

Loverde M., Afshordi N., 2008, Phys. Rev. D, 78, 123506

Martinet N., et al., 2017, preprint, (arXiv:1709.07678)

Morales J. L., Nocedal J., 2011, ACM Trans. Math. Softw., 38, 7:1

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Oguri M., et al., 2017, preprint, (arXiv:1705.06792)

Peel A., Lin C.-A., Lanusse F., Leonard A., Starck J.-L., Kilbinger M., 2017, A&A, 599, A79

Petri A., Haiman Z., Hui L., May M., Kratochvil J. M., 2013, Phys. Rev. D, 88, 123002

Planck Collaboration et al., 2016, A&A, 594, A13

Rykoff E. S., et al., 2016, ApJS, 224, 1

Schneider P., Seitz C., 1995, A&A, 294, 411

Seitz C., Schneider P., 1995, A&A, 297, 287

Seitz S., Schneider P., 2001, A&A, 374, 740

Shan H., et al., 2017, preprint, (arXiv:1709.07651)

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

Starck J.-L., Fadili J., Murtagh F., 2007, IEEE Transactions on Image Processing, 16, 297

Starck J.-L., Murtagh F., Fadili J., 2015, Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis, 2nd edn. Cambridge University Press, New York, NY, USA

Tibshirani R., 1994, Journal of the Royal Statistical Society, Series B, 58, 267

Van Waerbeke L., et al., 2013, MNRAS, 433, 3373

Vikram V., et al., 2015, Phys. Rev. D, 92, 022006

Wallis C. G. R., McEwen J. D., Kitching T. D., Leistedt B., Plouviez A., 2017, preprint, (arXiv:1703.09233)

Wiener N., 1949, Extrapolation, interpolation, and smoothing of stationary time series.   Vol. 7, MIT press Cambridge, MA

Zaroubi S., Hoffman Y., Fisher K. B., Lahav O., 1995, ApJ, 449, 446

Zhu C., Byrd R. H., Lu P., Nocedal J., 1997, ACM Trans. Math. Softw., 23, 550

Zuntz J., et al., 2015, Astronomy and Computing, 12, 45

van Uitert E., et al., 2017, preprint, (arXiv:1706.05004)

---

[4] https://github.com/pmelchior/skymapper

APPENDIX A: AUTHOR AFFILIATIONS

[1] Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

[2] Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa

[3] McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[4] Laboratoire AIM, UMR CEA-CNRS-Paris 7, Irfu, SAp/SEDI, Service d'Astrophysique, CEA Saclay, 91191 Gif-sur-Yvette Cedex, France

[5] Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

[6] Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA

[7] Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland

[8] Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany

[9] Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany

[10] Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA

[11] Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, Casilla 603, La Serena, Chile

[12] Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA

[13] Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK

[14] Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, 28049 Madrid, Spain

[15] CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France

[16] Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France

[17] Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

[18] Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil

[19] Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA

[20] National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA

[21] Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain

[22] Institut d'Estudis Espacials de Catalunya (IEEC), 08193 Barcelona, Spain

[23] Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain

[24] Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

[25] Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

[26] Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India

[27] Department of Astronomy/Steward Observatory, 933 North Cherry Avenue, Tucson, AZ 85721-0065, USA

[28] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA

[29] Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA

[30] Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

[31] SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

[32] Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA

[33] Department of Physics, The Ohio State University, Columbus, OH 43210, USA

[34] Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA

[35] Australian Astronomical Observatory, North Ryde, NSW 2113, Australia

[36] Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, São Paulo, SP, 05314-970, Brazil

[37] Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA

[38] Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain

[39] School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK

[40] Department of Physics, Brandeis University, Waltham, MA 02453, USA

[41] Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-859, Campinas, SP, Brazil

[42] Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

APPENDIX B: INDICATOR FUNCTION

We define the indicator function i_C (as used in equation 29) of a set C as

$$i_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise .} \end{cases} \tag{B1}$$

This paper has been typeset from a TEX/LATEX file prepared by the author.