Astronomy & Astrophysics

Weak-lensing mass reconstruction using sparsity and a Gaussian random field

J.-L. Starck[1], K. E. Themelis[1], N. Jeffrey[2,3], A. Peel[4], and F. Lanusse[1]

[1] AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, 91191 Gif-sur-Yvette, France
e-mail: jstarck@cea.fr
[2] Laboratoire de Physique de l'Ecole Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, Paris, France
[3] Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[4] Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland

ABSTRACT

Aims. We introduce a novel approach to reconstructing dark matter mass maps from weak gravitational lensing measurements. The cornerstone of the proposed method lies in a new modelling of the matter density field in the Universe as a mixture of two components: (1) a sparsity-based component that captures the non-Gaussian structure of the field, such as peaks or halos at different spatial scales, and (2) a Gaussian random field, which is known to represent the linear characteristics of the field well.
Methods. We propose an algorithm called MCALens that jointly estimates these two components. MCALens is based on an alternating minimisation incorporating both sparse recovery and a proximal iterative Wiener filtering.
Results. Experimental results on simulated data show that the proposed method exhibits improved estimation accuracy compared to customised mass-map reconstruction methods.

Key words. cosmology: observations – techniques: image processing – methods: data analysis – gravitational lensing: weak

1. Introduction

In recent years, interest has increased in exploring the two most dominant components of the universe: dark matter and dark energy. To this end, large-scale imaging and spectroscopic surveys are currently under development, such as the *Euclid* mission (Laureijs et al. 2011), the *Rubin* Observatory Legacy Survey of Space and Time (Abell et al. 2009), and the Roman Space Telescope (Spergel et al. 2015), which will map the sky with unprecedented accuracy. A prominent cosmological probe for these surveys is weak gravitational lensing.

Weak gravitational lensing measures correlations in the small distortions of distant galaxies caused by the gravitational potential of massive structures along the line of sight. Its effect on distant source galaxies is twofold: the galaxy shapes are magnified by a convergence field, κ, while the galaxy ellipticities are perturbed from their underlying intrinsic value by a shear field, γ. To contribute constraints on cosmological parameters and models, two-point correlation functions of the shear have been used with considerable success (Kilbinger et al. 2013; Alsing et al. 2016a). This type of correlation is sufficient to statistically describe the Gaussian structures of the lensing field, such as those expected to be present either in the early universe or at large scales that are less affected by gravitational collapse. To capture non-Gaussian structures such as those expected in smaller scales at later time, higher order moments of the shear need to be employed (Munshi & Coles 2017).

On the other hand, the convergence density field is not observed directly as a result of the mass-sheet degeneracy (Bartelmann & Schneider 2001; Kilbinger 2015). Physically,

the convergence field reveals the projected total matter density along the line of sight, weighted by a lensing kernel in the mid-distance between the observer and the galaxy sources. The density field is inhomogeneous, encompassing Gaussian-type large-scale structures, as well as non-Gaussian features, such as peaks. To shed light on the way in which the convergence density field constrains cosmology, peak statistics (e.g. Jain & Van Waerbeke 2000; Marian et al. 2011; Lin & Kilbinger 2015; Liu & Haiman 2016; Peel et al. 2017a; Fluri et al. 2018; Li et al. 2019; Ajani et al. 2020) and higher order correlation functions and moments, such as the Minkowksi functionals (e.g. Kratochvil et al. 2012; Shirasaki et al. 2012; Petri et al. 2013), have been applied directly on mass maps. It is therefore essential for mass-mapping methods to preserve both Gaussian and non-Gaussian features during the reconstruction process.

Mass-mapping methods solve an ill-posed problem because the sampling of the lensing field is irregular and the signal-to-noise ratio on small scales is low. A widely used algorithm to perform mass mapping is the Kaiser-Squires method (Kaiser & Squires 1993), which is expressed as a simple linear operator in Fourier space. However, it is inevitable that this estimator returns poor results because it does not take particular care of the noise or missing data. A different approach, motivated also by the Bayesian framework, is that of Wiener filtering (Wiener 1949). In this approach, a Gaussian random field is assumed as a prior for the convergence map, which is responsible for inserting some bias that prevents our estimate from over-fitting (Zaroubi et al. 1995). Moreover, a recently proposed customised method is the algorithm called gravitational lensing inversion and mapping using sparse estimators

(GLIMPSE2D; Lanusse et al. 2016). GLIMPSE2D is a highly sophisticated algorithm that takes advantage of the sparse regularisation framework to solve the ill-posed linear inverse problem. GLIMPSE2D is based on sparse representations (i.e. wavelets) and is therefore well designed to recover piece-wise smooth features. An analytical comparison of these three estimators is provided in Jeffrey et al. (2018a).

In this paper we propose to bridge the gap between the sparse regularisation method of GLIMPSE2D and the Wiener-filtering method by modelling the matter density field in the universe using both linear and non-linear characteristics. Specifically, we assume that the density field is modelled as a mixture of two terms: (1) a non-Gaussian term that adopts a sparse representation in a selected wavelet basis (Starck et al. 2015), and (2) a Gaussian term that is modelled using a Gaussian random field (Elsner & Wandelt 2013; Horowitz et al. 2019). The non-Gaussian signal component is able to capture the non-linear characteristics of the convergence field, such as peaks, while the Gaussian component of the signal is responsible for capturing the lower-frequency characteristics of the underlying field, such as smooth variations. To our knowledge, this is the first time that this mixture modelling is proposed for mass-map reconstruction.

This paper is structured as follows. In Sect. 2 we introduce the formalism of weak gravitational lensing and describe the mass-map reconstruction problem. To this end, we provide a brief overview of the current algorithms of Kaiser-Squires, Wiener filtering, and GLIMPSE2D. We then present our proposed mass-mapping method in Sect. 3. The method is novel in the sense that it exploits both sparsity in the wavelet domain and a Gaussian random field model. Section 4 illustrates the enhanced estimation performance of the proposed method by providing experiments conducted on both simulated and real data.

Notation: We use $(\cdot)^*$ to denote the Hermitian transpose of a matrix or the adjoint operator of a transform. With $\|\cdot\|_1$ and $\|\cdot\|_2$ we denote the ℓ_1 and ℓ_2 norm, respectively, ($\|\mathbf{x}\|_1 = \sum_{i=1}^{N} |x_i|$, $\|\mathbf{x}\|_2^2 = \mathbf{x}^T \mathbf{x}$). The determinant of a matrix or the absolute value of a scalar is denoted by $|\cdot|$, while diag($\mathbf{x}$) stands for a diagonal matrix that contains the elements of vector $\mathbf{x}$ on its diagonal. Finally, $\mathcal{R}^N$ is the N-dimensional Euclidean space, $\mathbf{0}$ denotes the zero vector, $\mathbf{1}$ the all-ones vector, and $\mathbf{I}_K$ is the $K \times K$ identity matrix.

2. Weak-lensing mass mapping

Gravitational lensing describes the phenomenon where the light emitted from distant galaxies is deflected as it passes through a foreground mass distribution. The lensing effect distorts the images of distant galaxies; the distortion is proportional to the size and shape of the projected matter distribution along the line of sight. Specifically, the mapping between the source coordinates, β, and the lensed image coordinates, θ, is given by the lens equation (e.g. Kilbinger 2015),

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \nabla\psi(\boldsymbol{\theta}), \tag{1}$$

where $\psi(\cdot)$ defines the lensing potential that conceals the deflection of light rays by the gravitational lens. Under the Born approximation, which assumes that the lensing potential is weak enough, we may linearise the coordinate transformation of Eq. (1) by using the Jacobian $\mathcal{A} = \partial\boldsymbol{\beta}/\partial\boldsymbol{\theta}$ as

$$\beta_i = \mathcal{A}_{ij}\theta_j, \tag{2}$$

where $\mathcal{A}_{ij} = \partial\beta_i/\partial\theta_j$ are the coefficients of the amplification matrix $\mathcal{A}$, and we assume the Einstein summation convention.

The symmetrical matrix $\mathcal{A}$ can also be parametrised in terms of the convergence, κ, and the shear, γ, as

$$\mathcal{A} = \left[\begin{array}{cc} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{array} \right]. \tag{3}$$

The convergence can then be defined as a dimensionless quantity that relates to the lensing potential through the Poisson equation,

$$\kappa = \frac{1}{2}(\partial_1\partial_1 + \partial_2\partial_2)\psi = \frac{1}{2}\nabla^2\psi, \tag{4}$$

while the shear is mathematically expressed as a complex field whose components also relate to ψ,

$$\gamma_1 = \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\psi \text{ and } \gamma_2 = \partial_1\partial_2\psi. \tag{5}$$

Equation (3) shows that the convergence causes an isotropic change in the size of the source image because it appears in the diagonal of $\mathcal{A}$. In comparison, the shear causes anisotropic changes of the image shapes. The convergence κ can also be interpreted via Eq. (4) as a weighted projection of the mass density field between the observation and the source. Factoring out the term $(1 - \kappa)$ in Eq. (3) leaves the amplification matrix dependent on the reduced shear,

$$\mathcal{A} = (1 - \kappa)\left[\begin{array}{cc} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{array} \right], \tag{6}$$

which is directly measured in lensing surveys, and it is defined as $g = \gamma/(1 - \kappa)$. In the weak-lensing limit, where $\gamma, \kappa \ll 1$, the reduced shear is approximately equal to the true shear, that is, $g \simeq \gamma$.

In this paper we are interested in recovering the convergence κ from the reduced shear data. This is an ill-posed inverse problem because of the finite sampling of the reduced shear over a limited area of the survey and the presence of shape noise in the measurements. In the following we review some of the customised weak-lensing mass reconstructing algorithms, namely the Kaiser-Squires, the Wiener filtering, and the GLIMPSE2D methods.

Kaiser-Squires. A theoretical framework for reconstructing convergence maps from the observable weak-lensing shear in the Fourier domain was proposed in Kaiser & Squires (1993). As we showed in Eqs. (4) and (5), the convergence and shear are both expressed as second-order derivatives of the lensing potential. Their interrelation through the lensing potential ψ is expressed by a two-dimensional convolution (Kaiser & Squires 1993),

$$\boldsymbol{\gamma}(\boldsymbol{\theta}) = \frac{1}{\pi} \int_{\mathbb{R}^2} \mathrm{d}^2\boldsymbol{\theta}' \, \mathcal{D}(\boldsymbol{\theta} - \boldsymbol{\theta}')\kappa(\boldsymbol{\theta}'), \tag{7}$$

where $\mathcal{D}(\boldsymbol{\theta}) = -1/(\theta_1 - i\theta_2)^2$. This convolution is equivalently expressed in Fourier space as the element-wise multiplication,

$$\tilde{\boldsymbol{\gamma}}(\boldsymbol{k}) = \pi^{-1}\tilde{\mathcal{D}}(\boldsymbol{k})\tilde{\kappa}(\boldsymbol{k}), \tag{8}$$

where $\boldsymbol{k}$ is the wave vector, and the Fourier transform of the kernel $\mathcal{D}(\boldsymbol{\theta})$ is given by

$$\tilde{\mathcal{D}}(\boldsymbol{k}) = \pi \frac{k_1^2 - k_2^2 + 2ik_1k_2}{k_1^2 + k_2^2}, \tag{9}$$

with k_1 and k_2 being the two frequency components of $\boldsymbol{k}$. Discretising Eq. (7) and adopting matrix notation, we may consider

that the observed shear $\boldsymbol{\gamma}$ is generated as a linear combination of the convolution matrix $\mathbf{A}$ and the unknown underlying field $\boldsymbol{\kappa}$,

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

where $\boldsymbol{n}$ is the statistical uncertainty vector associated with the data. Based on Eq. (8), $\boldsymbol{A}$ can be decomposed in Fourier space as $\boldsymbol{A} = \boldsymbol{F}\mathbf{P}\boldsymbol{F}^*$, where $\boldsymbol{F}$ denotes the discrete Fourier transform, $\boldsymbol{F}^*$ is its adjoint, and $\mathbf{P}$ is the diagonal operator that defines the convergence field-shear relation in Fourier space,

$$\tilde{\boldsymbol{\gamma}} = \mathbf{P}\tilde{\boldsymbol{\kappa}} = \left(\frac{k_1^2 - k_2^2}{k^2} + \iota \frac{2k_1 k_2}{k^2} \right) \tilde{\boldsymbol{\kappa}} \tag{11}$$

where $k^2 = k_1^2 + k_2^2$ and $\tilde{\boldsymbol{\kappa}} = \boldsymbol{F}\boldsymbol{\kappa}$.

Equation (11) corresponds to a discretised version of Eq. (8). As it stands, the Kaiser-Squires inversion of Eq. (11) has several drawbacks. First, it is not defined for $\boldsymbol{k} = 0$, which stems from the mass-sheet degeneracy (i.e. the mean value of the convergence field cannot be retrieved). Next, it is ill-posed because typically, the shear field is a discrete under-sampling of the underlying convergence field. Neither does it take masked data into account. Nonetheless, the Kaiser-Squires estimate is still used in practice because it is easy to apply.

Wiener filtering. The Wiener filter was introduced in the 1940s (Wiener 1949), and it is the optimal linear filter in the minimum mean-square sense that provides a denoised version of the desired signal. The Wiener-filtered estimate of the convergence map can be expressed using the linear equation

$$\boldsymbol{\kappa}_{\mathrm{G}} = \mathbf{W}\boldsymbol{\gamma}, \tag{12}$$

where the matrix $\mathbf{W}$ is given by

$$\mathbf{W} = (\mathbf{A}\boldsymbol{\Sigma}_\kappa \mathbf{A}^* + \boldsymbol{\Sigma}_n)^{-1} \mathbf{A}^* \boldsymbol{\Sigma}_\kappa, \tag{13}$$

and $\boldsymbol{\Sigma}_\kappa$ and $\boldsymbol{\Sigma}_n$ are the pre-defined covariance matrix of the convergence and the noise, respectively. When the noise is not stationary, the Wiener-filter solution is not straightforward, and an iterative approach is required. This is discussed in the next section.

From a Bayesian perspective, Wiener filtering is equivalent to the maximum a posteriori (MAP) estimator given that the convergence is a zero-mean Gaussian signal with covariance $\boldsymbol{\Sigma}_\kappa$. When we assume that shear measurements are distorted by uncorrelated Gaussian noise, the likelihood function shares the noise properties, that is,

$$p(\boldsymbol{\gamma}|\boldsymbol{\kappa}, \boldsymbol{\Sigma}_n) = |2\pi\boldsymbol{\Sigma}_n|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\boldsymbol{\gamma} - \mathbf{A}\boldsymbol{\kappa})^* \boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\gamma} - \mathbf{A}\boldsymbol{\kappa}) \right], \tag{14}$$

while the distribution of the Gaussian random field can be written as

$$p(\boldsymbol{\kappa}|\boldsymbol{\Sigma}_\kappa) = |2\pi\boldsymbol{\Sigma}_\kappa|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\boldsymbol{\kappa}^* \boldsymbol{\Sigma}_\kappa \boldsymbol{\kappa} \right]. \tag{15}$$

Given Eqs. (14) and (15) above, the MAP estimator is expressed as

$$\boldsymbol{\kappa}_{\mathrm{G}} = \arg\max_{\boldsymbol{\kappa}} \{p(\boldsymbol{\kappa}|\boldsymbol{\gamma})\} \propto \arg\max_{\boldsymbol{\kappa}} \{p(\boldsymbol{\gamma}|\boldsymbol{\kappa}, \boldsymbol{\Sigma}_n)p(\boldsymbol{\kappa}|\boldsymbol{\Sigma}_\kappa)\}$$

$$\propto \arg\max_{\boldsymbol{\kappa}} \exp\left[-\frac{1}{2}\boldsymbol{\kappa}^* \boldsymbol{\Sigma}_\kappa \boldsymbol{\kappa} - \frac{1}{2}(\boldsymbol{\gamma} - \mathbf{A}\boldsymbol{\kappa})^* \boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\gamma} - \mathbf{A}\boldsymbol{\kappa}) \right], \tag{16}$$

which leads to the minimisation

$$\boldsymbol{\kappa}_{\mathrm{G}} = \arg\min_{\boldsymbol{\kappa}} \left\{ \|\boldsymbol{\gamma} - \mathbf{A}\boldsymbol{\kappa}\|_{\boldsymbol{\Sigma}_n}^2 + \|\boldsymbol{\kappa}\|_{\boldsymbol{\Sigma}_\kappa}^2 \right\}. \tag{17}$$

It is easy to see that the minimum is attained at $\boldsymbol{\kappa}_{\mathrm{G}} = \mathbf{W}\boldsymbol{\gamma}$, where $\mathbf{W}$ is the Wiener filter given in Eq. (12).

The implementation of the Wiener filter requires the inversion of the matrix $\mathbf{W}$ that includes both a signal covariance matrix component and the noise covariance matrix component.

The unknown $\boldsymbol{\kappa}_{\mathrm{G}}$ is assumed to be a Gaussian random field with a covariance matrix diagonal in Fourier space, $\boldsymbol{\Sigma}_\kappa = \boldsymbol{F}^* \boldsymbol{C}_\kappa \boldsymbol{F}$, where $\boldsymbol{C}_\kappa$ is a diagonal matrix with diagonal values equal to the theoretical power spectrum P_κ. When the noise is stationary, its covariance matrix is also diagonal with diagonal elements equal to the noise power spectrum P_n. In this case, the filter solution is obtained in Fourier space by $\tilde{\boldsymbol{\kappa}}_{\mathrm{G}} = \tilde{\mathbf{W}}\tilde{\boldsymbol{\gamma}}$, where the Wiener filter is $\tilde{\mathbf{W}} = \frac{P_\kappa}{P_\kappa + P_n}$. In practice, the noise is generally not stationary and depends on the number of shear measurements in the area related to a given pixel of the shear field. Therefore the noise covariance matrix is diagonal in pixel space, not in Fourier space, and the Wiener-filter solution becomes harder to derive, requiring either making an incorrect assumption (i.e. that the noise is stationary) or inverting a very large matrix. This renders its inversion computationally complex and prone to numerical errors. To circumvent this computationally intensive operation, a forward-backward (FB) proximal iterative Wiener filtering was proposed in Bobin et al. (2012) for the denoising of cosmic microwave background spherical maps, exploiting the property that the signal and noise covariance matrices are diagonal in pixel and Fourier space, respectively. Equation (17) comprises two separable terms,

$$f_1(\boldsymbol{\kappa}) = \| \boldsymbol{\gamma} - \mathbf{A}\boldsymbol{\kappa} \|_{\boldsymbol{\Sigma}_n}^2 \quad \text{and} \quad f_2(\boldsymbol{\kappa}) = \|\boldsymbol{\kappa}\|_{\boldsymbol{\Sigma}_\kappa}^2. \tag{18}$$

Following the FB method (Starck et al. 2015), we can solve Eq. (18) by designing an iterative fixed point algorithm as

$$\boldsymbol{\kappa}^{k+1} = \mathrm{prox}_{\mu f_2}(\boldsymbol{\kappa}_{\mathrm{G}}^k + \nabla f_1(\boldsymbol{\kappa}_{\mathrm{G}}^k)), \tag{19}$$

which is known to converge when $\mu < 2/\|\mathbf{A}^* \boldsymbol{\Sigma}_n^{-1} \mathbf{A}\|_2$, Combettes & Wajs (2005).

Computing the proximal operator in Eq. (19), we have the following iterative Wiener-filtering algorithm:
– Forward step:

$$\mathbf{t} = \boldsymbol{\kappa}^n + 2\mu \mathbf{A}^* \boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\gamma} - \mathbf{A}\boldsymbol{\kappa}^n) \tag{20}$$

– Backward step:

$$\boldsymbol{\kappa}^{n+1} = \boldsymbol{F}^* \left(\mathbf{P}_\kappa \left(\mathbf{P}_\eta + \mathbf{P}_\kappa \right)^{-1} \right) \boldsymbol{F}\mathbf{t}, \tag{21}$$

where $\mathbf{t}$ is an auxiliary variable, $\mu = \min(\boldsymbol{\Sigma}_n)$, $\mathbf{P}_\eta = 2\mu\mathbf{I}$ and $\boldsymbol{\kappa}^0 = 0$. This algorithm is free from matrix inversions because both $\boldsymbol{\Sigma}_n$ used in Eq. (20) and $\mathbf{P}_\kappa$ used in Eq. (21) are diagonal matrices. A similar algorithm, exploiting transformations between pixel and harmonic space, was proposed by Elsner & Wandelt (2013) using the messenger field framework. These methods can also be adapted to efficiently sample from the posterior distribution of the unknown mass map (Alsing et al. 2017; Jeffrey et al. 2018b). The Wiener approach recovers the Gaussian component of the convergence field well, but it is far from optimal at extracting the non-Gaussian information from the data because peak-like structures are suppressed. This has motivated the development of sparse recovery methods based on wavelets.

Sparse recovery. It has been shown that sparse recovery using wavelets is a very efficient way to reconstruct convergence maps (Starck et al. 2006b; Leonard et al. 2012; Lanusse et al. 2016; Peel et al. 2017b; Price et al. 2020). The mass-mapping problem is addressed as a general ill-posed problem, which is solved through weighted ℓ_1-norm regularisation in a wavelet-based analysis sparsity model. Sparsity in the discrete cosine transform (DCT) domain was also proposed to fill the missing data area in the convergence map (Pires et al. 2009).

The GLIMPSE algorithm avoids any binning or smoothing of the input data that could potentially cause loss of information. The primary cost function is

$$\arg\min_{\kappa} \left\{ \frac{1}{2}\|\gamma - \mathcal{T}\mathbf{P}\mathcal{F}^*\kappa\|^2_{\Sigma_n} + \lambda\|\omega \odot \mathbf{\Phi}^*\kappa\|_1 + i_{\mathbb{R}}(\kappa) \right\}, \quad (22)$$

where $\mathcal{T}$ is the nonuniform discrete Fourier transform (NDFT) matrix, $\mathbf{P}$ is defined in Eq. (11), λ is a sparsity regularisation parameter, ω is a weighting vector, $\mathbf{\Phi}^*$ is the adjoint operator of the wavelet transform, and $i_{\mathbb{R}}$ is an identity function that drives the imaginary part of the convergence to zero. This cost function is also generalised in GLIMPSE2D in order to replace the shear by the reduced shear, or to incorporate the flexion information. It has been shown that GLIMPSE2D significantly outperforms Wiener filtering for peaks recovery, while Wiener filtering does better on the Gaussian map content (Jeffrey et al. 2018a).

Deep learning. Deep-learning techniques have recently been proposed and appear to be very promising (Jeffrey et al. 2020). The input of the neural network is not the shear field directly, but rather the Wiener-filter solution. We could imagine that the closer we are to the true solution with standard techniques such as Wiener filter or sparsity, the better deep learning can improve the solution. Open questions related to deep learning remain to be answered, such as its generalisation to cosmologies not present in the training data set, or the potential bias introduced by using a theoretical power spectrum in the Wiener-filter solution serving as input of the neural network.

3. Modelling with sparsity and a Gaussian random field

3.1. New convergence map model

We showed two different models in Sect. 2: the first modelling the convergence map as a Gaussian random field, which recovers the large scales of the convergence map well but suppresses peak structures, and the second assuming the convergence map is compressible in the wavelet domain (i.e. sparse modelling). The sparse recovery is clearly complementary to the Wiener-filter solution because it recovers peaks extremely well, but recovers the Gaussian content poorly.

To address these limitations, it appears to be natural to introduce a novel modelling approach, where the convergence field κ is assumed to comprise two parts, a Gaussian and a non-Gaussian:

$$\kappa = \kappa_{\mathrm{G}} + \kappa_{\mathrm{NG}}. \quad (23)$$

The non-Gaussian part of the signal κ_{NG} is subject to a sparse decomposition in a wavelet dictionary, while the component κ_{G} is assumed to be inherently non-sparse and Gaussian.

The morphological component analysis (MCA) has been proposed (Starck et al. 2004; Elad et al. 2005) to separate two components mixed in a single image when these components

have different morphological properties. This appears to be impossible because we have two unknowns and one equation, but it was shown that it is sometimes possible to extract these two components if we can exploit their morphological differences. This requires different penalisation functions C_{G} and C_{NG} on each of these two components, and we need to minimise

$$\min_{\kappa_{\mathrm{G}},\kappa_{\mathrm{NG}}} \left\{ \|\gamma - \mathbf{A}(\kappa_{\mathrm{G}} + \kappa_{\mathrm{NG}})\|^2_{\Sigma_n} + C_{\mathrm{G}}(\kappa_{\mathrm{G}}) + C_{\mathrm{NG}}(\kappa_{\mathrm{NG}}) \right\}. \quad (24)$$

The MCA performs an alternating minimisation scheme:
– Estimate κ_{G} assuming κ_{NG} is known:

$$\min_{\kappa_{\mathrm{G}}} \left\{ \|(\gamma - \mathbf{A}\kappa_{\mathrm{NG}}) - A\kappa_{\mathrm{G}}\|^2_{\Sigma_n} + C_{\mathrm{G}}(\kappa_{\mathrm{G}}) \right\}. \quad (25)$$

– Estimate κ_{NG} assuming κ_{G} is known:

$$\min_{\kappa_{\mathrm{NG}}} \left\{ \|(\gamma - \mathbf{A}\kappa_{\mathrm{G}}) - A\kappa_{\mathrm{NG}}\|^2_{\Sigma_n} + C_{\mathrm{NG}}(\kappa_{\mathrm{NG}}) \right\}. \quad (26)$$

Examples of these decompositions are shown on the MCA web page[1]. A range of MCA applications in astrophysics can be found in Starck et al. (2003), André et al. (2010), Möller et al. (2015), Bobin et al. (2016), Melchior et al. (2018), Joseph et al. (2019) and Wagner-Carena et al. (2020).

Gaussian component κ_{G}. We used the standard Wiener-filter modelling where κ_{G} is assumed to be a Gaussian random field,

$$C_{\mathrm{G}}(\kappa_{\mathrm{G}}) = \|\kappa\|^2_{\Sigma_\kappa}, \quad (27)$$

and the solution of Eq. (25) is obtained using the iterative Wiener filtering presented in the previous section.

Non-Gaussian component. There are different ways to use a sparse model in the MCA framework. The most obvious would be to use standard ℓ_1 or ℓ_0-norm regularisation in a wavelet-based sparsity model, as is done in the GLIMPSE2D algorithm. This would give

$$C_{\mathrm{NG}}(\kappa_{\mathrm{NG}}) = \lambda\|\mathbf{\Phi}^*\kappa_{\mathrm{NG}}\|_p, \quad (28)$$

where $p = 0$ or 1, Φ is the wavelet matrix, and λ is the regularisation parameter (Lagrange multiplier).

After implementing this approach, we found that large wavelet scales and Fourier low frequencies are relatively close, leading to difficulties in separating the information. We therefore investigated another approach that first involves estimating the set Ω of active coefficients, that is, the scales and positions where wavelet coefficients are above a given threshold. These are typically between three and five times the noise standard deviation relative to each wavelet coefficient. Ω can therefore be seen as a mask in the wavelet domain, where $\Omega_{j,x} = 1$ if a wavelet coefficient detected at scale j and position x, that is, when $|(\mathbf{\Phi}^*\mathbf{A}^*\gamma)_{j,x}| > \lambda\sigma_{j,x}$, and 0 otherwise. The noise $\sigma_{j,x}$ at scale j and position x can be determined using noise realisations as in the GLIMPSE algorithm. An even faster approach is to detect the significant wavelet coefficients on $\mathbf{\Phi}^*\mathbf{A}^*\Sigma_n^{-\frac{1}{2}}\gamma$ instead of $\mathbf{\Phi}^*\mathbf{A}^*\gamma$. The noise is therefore whitened because the noise factor $\mathbf{A}^*\Sigma_n^{-\frac{1}{2}}$ is Gaussian with a standard deviation equal to unity and with a uniform power spectrum. We implemented both approaches and found similar results, although the second approach is easier.

---

[1] http://jstarck.cosmostat.org/mca

When this wavelet mask Ω is estimated, we can estimate the non-Gaussian component κ_{NG} by

$$\min_{\kappa_{NG}} \left\{ \|\Omega \odot \Phi^* ((\gamma - \mathbf{A}\kappa_G) - A\kappa_{NG})) \|^2 + C_{NG}(\kappa_{NG}) \right\}, \qquad (29)$$

with $C_{NG}(\kappa_{NG}) = i_{\mathbb{R}}(\kappa_{NG})$.

This changes the original formalism because the data fidelity term is now different, but it presents a very interesting advantage. When Ω is fixed, the algorithm is almost linear and only the positivity constraint remains. Therefore we can easily derive a good approximation of the error map just by propagating noise and relaxing this positivity constraint. This is further discussed below. Similarly to the GLIMPSE method, a positivity constraint is applied on the non-Gaussian component κ_{NG}. Peaks in κ can be on top of voids, and therefore have negative pixel values. As peaks are captured by the non-Gaussian component, they are positive by construction in κ_{NG}, but the convergence map $\kappa = \kappa_G + \kappa_{NG}$ can still be negative at peak positions. The larger the non-Gaussianities, the more we can expect MCALens to improve on linear methods such as the Wiener filtering.

The prior signal auto-correlation of the Gaussian component is included within the signal covariance term of the Gaussian component. We encode no explicit prior auto-correlation for the non-Gaussian signal and no explicit prior cross-correlation between the Gaussian and non-Gaussian component. Clearly, these correlations exist, but including their contribution in the prior in this framework would be extremely difficult theoretically and in practice. However, these correlations will still appear in the final reconstruction, driven by the correlation information in the data.

3.2. MCALens algorithm

Algorithm 1 MCALens algorithm

1: **Input**: Shear map γ_1, γ_2, signal and noise covariance Σ_κ, Σ_n, and detection level λ.
2: Initialise: $\kappa_{NG}^{(0)} = \kappa_G^{(0)} = \Omega = 0$, $\mu = \min(\Sigma_n)$, $\mathbf{P}_\eta = 2\mu\mathbf{I}$.
3: Calculate wavelet coefficients: $\alpha = \Phi^* \mathbf{A}^* \Sigma_n^{-\frac{1}{2}} \gamma$.
4: $\forall j, x$, set $\Omega_{j,x} = 1$ $\quad if \quad |(\alpha)_{j,x}| > \lambda$.
5: **for** $n = 0, \ldots, N_{max} - 1$ **do**
6: $\qquad$————— Find κ_{NG} —————
7: $\qquad$ Calculate the shear residual: $\gamma_r = \gamma - \mathbf{A}(\kappa_G^{(n)} + \kappa_{NG}^{(n)})$.
8: $\qquad$ Calculate the sparse residual: $s_r = \mathbf{A}^* \Sigma_n^{-\frac{1}{2}} \gamma_r$.
9: $\qquad$ Calculate the sparse residual in the mask:
$\qquad$ $s_{mr} = \Phi(\Omega \odot (\Phi^* s_r))$.
10: $\qquad$ Get the new sparse component: $\mathbf{S} = \kappa_{NG}^{(n)} + s_{mr}$.
11: $\qquad$ Positivity constraint: $\kappa_{NG}^{(n+1)} = [\mathbf{S}]_+$.
12: $\qquad$————- Find κ_G —————
13: $\qquad$ Calculate the shear residual: $\gamma_r = \gamma - \mathbf{A}(\kappa_G^{(n)} + \kappa_{NG}^{(n+1)})$.
14: $\qquad$ Forward step: $\mathbf{t} = \kappa_G^n + 2\mu\mathbf{A}^*\Sigma_n^{-1}\gamma_r$.
15: $\qquad$ Backward step: $\kappa_G^{n+1} = F^* \left(\mathbf{P}_\kappa \left(\mathbf{P}_\eta + \mathbf{P}_\kappa \right)^{-1} \right) F\mathbf{t}$.
16: **end for**
17: **return** $\left(\kappa_G^{(N_{max})}, \kappa_{NG}^{(N_{max})} \right)$.

We solved the recovery problem of Eq. (23) using a two-step optimisation procedure. First, a gradient descent step to minimise Eq. (29) to recover the non-Gaussian component κ_{NG}. This was followed by an iteration of the iterative Wiener filtering to minimise Eq. (17). Details of the method are given in Algorithm 1.

The number of scales N_s used to compute the wavelet transform of an $N_x \times N_y$ image are automatically derived by
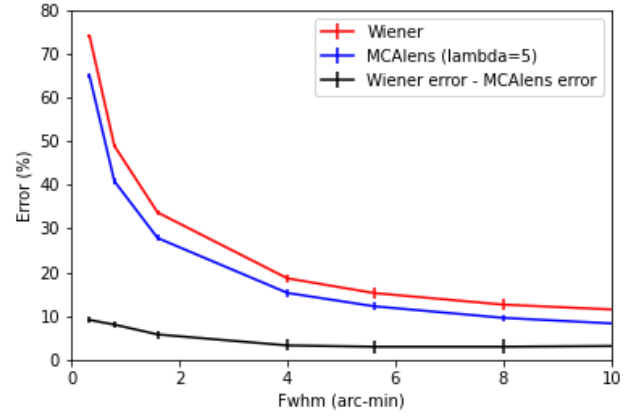

Fig. 1. RAMSES simulations: Error vs. scale for Wiener filter (red) and MCALens (blue).

$N_s = \mathbf{int}(\log(\min(N_x, N_y)))$. The λ parameter is a detection level, which was fixed for all our experiments with real and simulated data to 5, that is, five times the noise standard deviation. This is a conservative threshold and gave excellent results.

3.3. Errors

Much attention has recently been given to estimate errors or uncertainties with mass-map products. For linear methods such as Wiener filtering or Kaiser-Squires, it is easy to estimate the standard deviation (or root-mean-square, RMS) per pixel, just by propagating noise realisations using the same reconstruction filters. The uncertainty per pixel does not, however, give the probability that a clump in a reconstructed image is true or only due to some noise fluctuations. Another approach, closer to certain science cases with maps, estimates the significance of clumps. Peel et al. (2017b) used Monte Carlo simulations to address the significance of clumps. Repetti et al. (2019) proposed a hypothesis test called BUQO to perform the same task, requiring the user to define manually a mask around the clump. Similarly, Price et al. (2019) performed hypothesis tests of Abell-520 cluster structures using highest posterior density regions. With MCALens, we can include aspects of both approaches.

RMS and signal-to-noise ratio maps. Algorithm 1 includes two steps that involve a non-linear operator, first to estimate Ω in line 4 and then in line 11 to perform the positivity constraint. To propagate noise realisations, Ω has to be set to the one obtained with the data, and this non-linearity step therefore does not occur, so that only the positivity remains. A full linear algorithm could therefore be obtained just by removing this positivity constraint during the noise propagation by replacing $\kappa_{NG}^{(n+1)} = [\mathbf{S}]_+$ in Algorithm 1, line 11 by $\kappa_{NG}^{(n+1)} = \mathbf{S}$. This causes more noise to enter the solution because few pixel values with negative values in κ_{NG} are not thresholded, and the derived RMS map is therefore slightly conservative. We can therefore build noise realisations, run the MCALens algorithm on each realisation, and derive the RMS map by taking pixel per pixel the standard deviation of the obtained reconstructed maps. The signal-to-noise ratio (S/N) map is derived by dividing the absolute value of the reconstructed map by the RMS map.

Significance map. Algorithm 1 obtains in line 4 the wavelet mask Ω by comparing the wavelet coefficients, $\alpha = \Phi^* \mathbf{A}^* \Sigma_n^{-\frac{1}{2}} \gamma$, to the threshold $\lambda\sigma$ where $\sigma = 1$ as the noise is whitened with
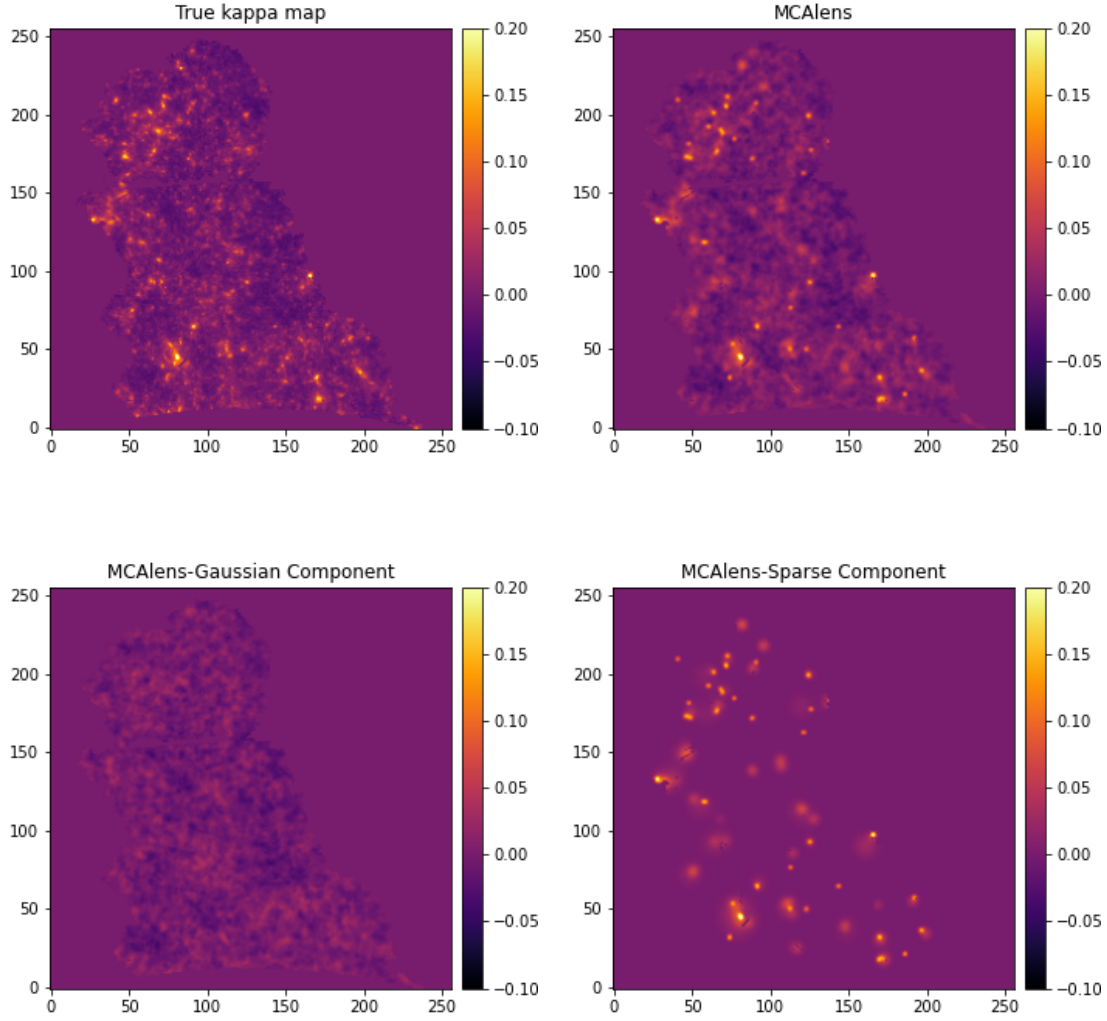
Fig. 2. RAMSES simulations. *Top panel*: true convergence map and MCAlens recovery. *Bottom panel*: Gaussian components and sparse components. The sum of these two maps is equal to the top right MCAlens map.

unit variance. This corresponds to performing a hypothesis test H_0 that the wavelet coefficient is due to noise alone, and if a given wavelet coefficient $\alpha_{j,x}$ is such that $|\alpha_{j,x}| > \lambda$, then the H_0 hypothesis is rejected and we assume that the coefficient amplitude cannot be explained by noise fluctuations and is therefore due to signal. λ is therefore directly related to the significance of the wavelet coefficients, and the mask Ω indicates which coefficients are detected with a given significance level. $\Omega_{j,x}$ is binary, and we build the significance map s by $s_x = \sum_j \Omega_{j,x}$, that is, a simple co-adding of all binary scales. Such a map could also be used as a way to automatically derive the user mask required in the BUQO method (Repetti et al. 2019). We present in Sect. 4 examples of RMS, S/N, and significance maps.

3.4. Extension to the sphere

When a wide-field map needs to be reconstructed, the flat approximation can no longer be used, and we have to build a map on the sphere. A traditional approach is to decompose the sphere into overlapping patches, assume a flat approximation on each individual patch, reconstruct each patch independently, and finally, recombine all patches on the sphere. This solution is certainly good enough to recover clumps relative to clusters, that is, the non-Gaussian component, but certainly not good enough for the Gaussian component, which contains information at low

frequencies. In the framework of the DES project, a 1500 deg^2 map has been reconstructed (Chang et al. 2018) using HEALPIX pixelisation (Górski et al. 2005) and a straightforward spherical Kaiser-Squires algorithm consisting of

1. calculating a spin transform of the HEALPIX shear map to obtein the E and B spherical harmonic coefficients,

2. smoothing by a Gaussian the E and B modes in the spherical harmonic domain,

3. applying an inverse spherical transform independently to each of these two modes to obtain the two convergence maps κ^E and κ^B.

Sparsity in a Bayesian framework (Price et al. 2021) and forward-fitting in harmonic space (Mawdsley et al. 2020) have also recently been proposed for spherical mass mapping. Using similarly a HEALPIX shear and convergence pixelisation, we can also easily derive an extension of the iterative-Wiener filtering and the MCALens algorithm by

– replacing the matrix $\mathbf{A} = \boldsymbol{F}\mathbf{P}\boldsymbol{F}^*$ in Eq. (10) by $\mathbf{A} = {}_2\mathbf{Y_0}\mathbf{Y}^*$, where ${}_s\mathbf{Y}$ and ${}_s\mathbf{Y}^*$ represent the forward and inverse spin-s spherical harmonic transforms, respectively,

– replacing the wavelet decomposition $\boldsymbol{\Phi}$ used in Eq. (28) by the spherical wavelet decomposition (Starck et al. 2006a).

Then the same MCAlens algorithm as in Algorithm 1 can be used to derive a spherical convergence map. An example is given in Sect. 4.3.
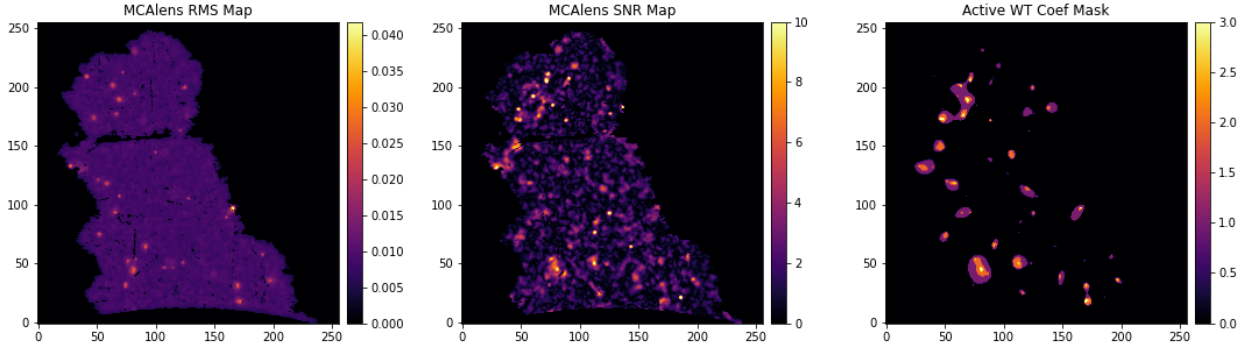
Fig. 3. RAMSES simulations. *From left to right*: RMS map, S/N map, and significance map.

3.5. B-mode

We focused earlier on the convergence map (i.e. E-mode). The MCAlens algorithm given in Algorithm 1 remains valid when E and B mode are to be estimated by adopting the following notation:

$$\boldsymbol{\gamma} = \begin{pmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{pmatrix}, \boldsymbol{\kappa} = \begin{pmatrix} \kappa_E \\ \kappa_B \end{pmatrix}, \mathbf{A} = \begin{pmatrix} \frac{k_1^2 - k_2^2}{k^2} & \frac{2k_1 k_2}{k^2} \\ \frac{2k_1 k_2}{k^2} & -\frac{k_1^2 - k_2^2}{k^2} \end{pmatrix}, \boldsymbol{\alpha} = \begin{pmatrix} \alpha_E \\ \alpha_B \end{pmatrix},$$

$$\mathbf{P}_\kappa = \begin{pmatrix} \mathbf{P}_{\kappa_E} \\ \mathbf{P}_{\kappa_B} \end{pmatrix}, \boldsymbol{\Phi} = \begin{pmatrix} \boldsymbol{\Phi}_E \\ \boldsymbol{\Phi}_B \end{pmatrix},$$

where the matrix $\boldsymbol{\Phi}$ consists of applying a sparse decomposition independently on each mode. In practice, we used the same wavelet decomposition for both modes (i.e. $\boldsymbol{\Phi}_E = \boldsymbol{\Phi}_B$). The delicate point is applying the Wiener filter to the B-mode at line 15 of the algorithm. In theory, $\mathbf{P}_{\kappa_B} = 0$, and by construction, no Gaussian component can be recovered. Because the B-mode is mainly useful for investigating systematic errors, we are not interested in recovering the B-mode least-squares estimator (which is zero), and we find it more useful to process the B-mode similarly to the E-mode. We therefore advocate use $\mathbf{P}_{\kappa_B} = \mathbf{P}_{\kappa_E}$ instead. In this way, the fluctuations in the E-mode can be properly compared to those in the B-mode.

4. Experimental results

4.1. Toy-model experiment

We used a simulation derived from RAMSES N-body cosmological simulations (Teyssier 2002), with a ΛCDM model[2], with a pixel size of $0.34' \times 0.34'$ and a galaxy redshift of 1. To obtain a realistic mask and noise behaviour, we used the MICE pixel noise covariance derived for the DES project (see Appendix A). As the pixel resolution is different, it leads to an optimistic realisation, but it has the advantage of illustrating the effect of our MCA model well.

We ran MCAlens using 100 iterations, $\lambda = 5$ (i.e. detection at 5σ), the MICE covariance matrix, and the theoretical power spectrum was the true map power spectrum. To evaluate the results, we ran 100 different noise realisations, and we applied both Wiener filtering and MCAlens. We calculated the reconstruction error at different resolutions with

$$\text{Err}_\%(\sigma) = \frac{\| \, \boldsymbol{G}_\sigma \, (M \, (\boldsymbol{\kappa} - \boldsymbol{\kappa}_t)) \, \|}{\| \, M \boldsymbol{\kappa}_t \, \|}, \tag{30}$$

where $\| \, x \, \| = \sqrt{\sum_i x_i^2}$, $\boldsymbol{\kappa}$ is the reconstructed convergence map, $\boldsymbol{\kappa}_t$ is the true convergence map, $\boldsymbol{G}_\sigma(x)$ is the convolution of x with
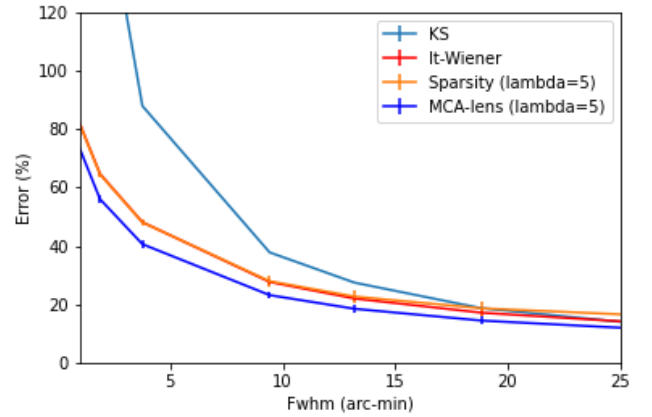
2 See http://www.projet-horizon.fr

Fig. 4. Columbia simulations. Error vs. scale for four different methods: Kaiser-Squires (light blue), Wiener (red), sparsity (orange), and MCAlens (blue).

a Gaussian with standard deviation σ, and M is the binary mask with $M_k = 1$ if the covariance matrix is not infinite at location k (i.e. we have data at this location) and 0 otherwise. Figure 1 shows the mean error $\text{Err}_\%(\sigma)$ for the Wiener filter and MCAlens solutions. The black curve shows the difference between the Wiener-filter error and the MCAlens error. This allows us to better visualise that the improvement is larger at fine scales. It is interesting to note that MCAlens is better than Wiener filter even at large scales. When the MCAlens non-Gaussian component contains a significant number of features as in the case of this experiment, these features also contribute in a a non negligible way on larger scales, which explains why MCAlens remains better than Wiener filtering at large scales. MCAlens leads to a clear improvement in terms of quadratic error. The top panels of Fig. 2 show the simulated convergence map and the MCAlens reconstructed map. The bottom panels show the Gaussian and the non-Gaussian part recovered from the noisy data. The sum of these two components is equal to the MCAlens reconstruction (top right). This shows that the non-linear and linear components can be recovered well. Figure 3 shows the RMS map, the S/N map, and the significance map.

4.2. Columbia lensing simulation

We used a convergence map, released by the Columbia lensing group[3] (Liu et al. 2018), corresponding to a cosmological model with parameters $\{M_\nu, \Omega_m, 10^9 A_s, M_\nu, h, w\} = \{0.1, 0.3, 2.1, 0, 0.7, -1\}$, and with a pixel size of $0.4' \times 0.4'$.
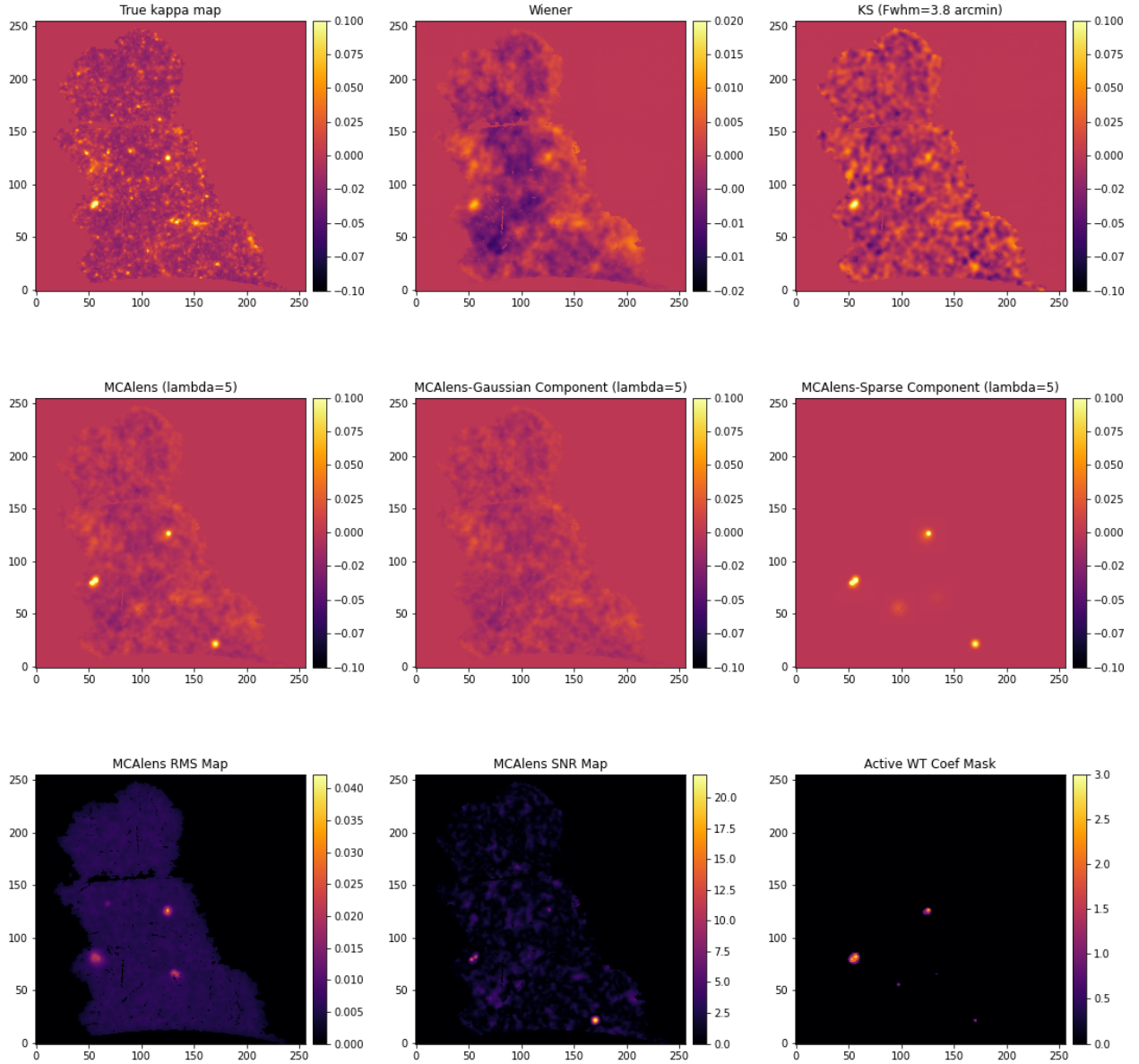
3 http://columbialensing.org

Fig. 5. Columbia convergence map recovery. *Top panel*: a true convergence map, a Wiener-filter map, and a Kaiser-Squires map smoothed with a Gaussian with an FWHM of 3.8 arcmin. *Middle panel*: MCAlens map and its Gaussian and sparse components. The sum of these two last maps is equal to the first map. *Bottom panel*: RMS, S/N, and significance maps.

We rebinned the map to $0.8' \times 0.8'$, and similarly to the previous experiment, we simulated noisy data using the same MICE covariance, applying a global rescaling in order to have realistic noise corresponding to a mean number of galaxies equal to 30 per arcmin^2, as we expect in future space lensing surveys. To evaluate the results, we ran 100 different noise realisations, and we applied Kaiser-Squires, sparse recovery, Wiener filtering, and MCAlens. For the sparse recovery and MCAlens, we used $\lambda = 5$ (i.e. detection at 5σ), and for Wiener filtering and MCAlens, we the used the theoretical power spectrum as the true convergence map power spectrum.

Figure 4 shows the error computed using Eq. (30). In addition to the well-recovered peaks, MCAlens also leads to a clear improvement in terms of quadratic error compared the other methods. In contrast to the RAMSES experiment, we see here a convergence at large scales between MCAlens and Wiener filter because the MCAlens non-Gaussian component contains only a

few peaks (see Fig. 5, middle right), which therefore contribute only negligibly to the largest scales.

The top row of Fig. 5 shows the simulated convergence map, the Wiener-filter result, and the Kaiser-Squires reconstruction with smoothing applied. The middle row shows the MCAlens map and its Gaussian and non-Gaussian parts recovered from the noisy the data. The sum of these two components is equal to the MCAlens map. Similarly to previous experiments, the non-linear and the linear components are well recovered. The bottom row of Fig. 5 shows the RMS map, the S/N map, and the significance map.

4.3. Spherical data

We created a full shear map from the full-sky MICE simulated map and its covariance matrix, which has about three galaxies per arcmin^2. We ran the spherical MCAlens method with 200
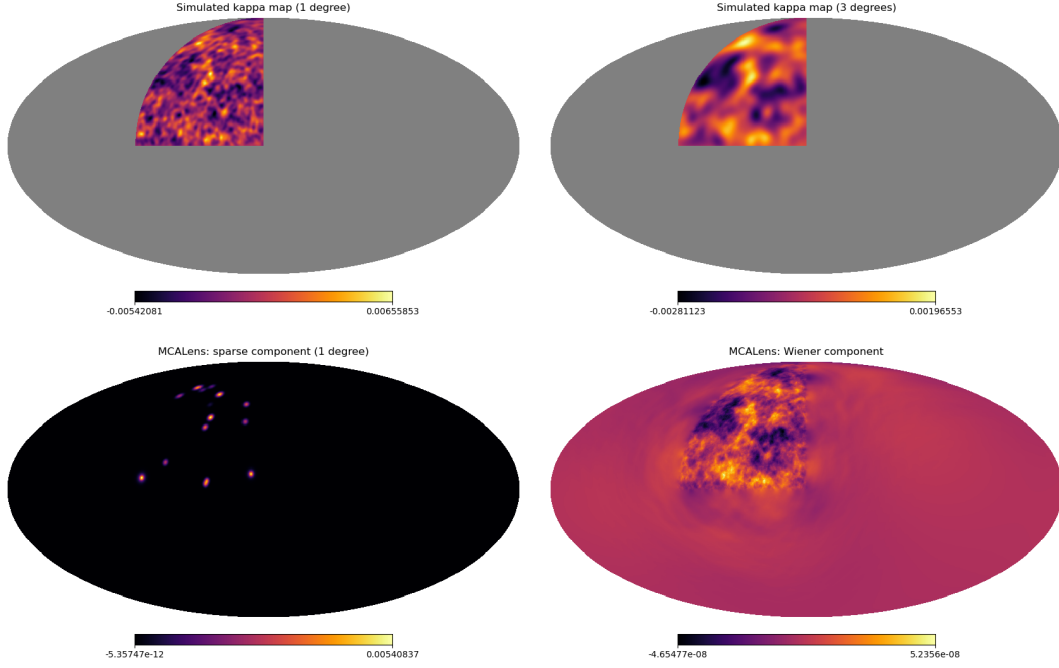
Fig. 6. MICE simulations. *Top panel*: true convergence map at 1 and 3 degrees. *Bottom*: MCAlens sparse and Gaussian components.
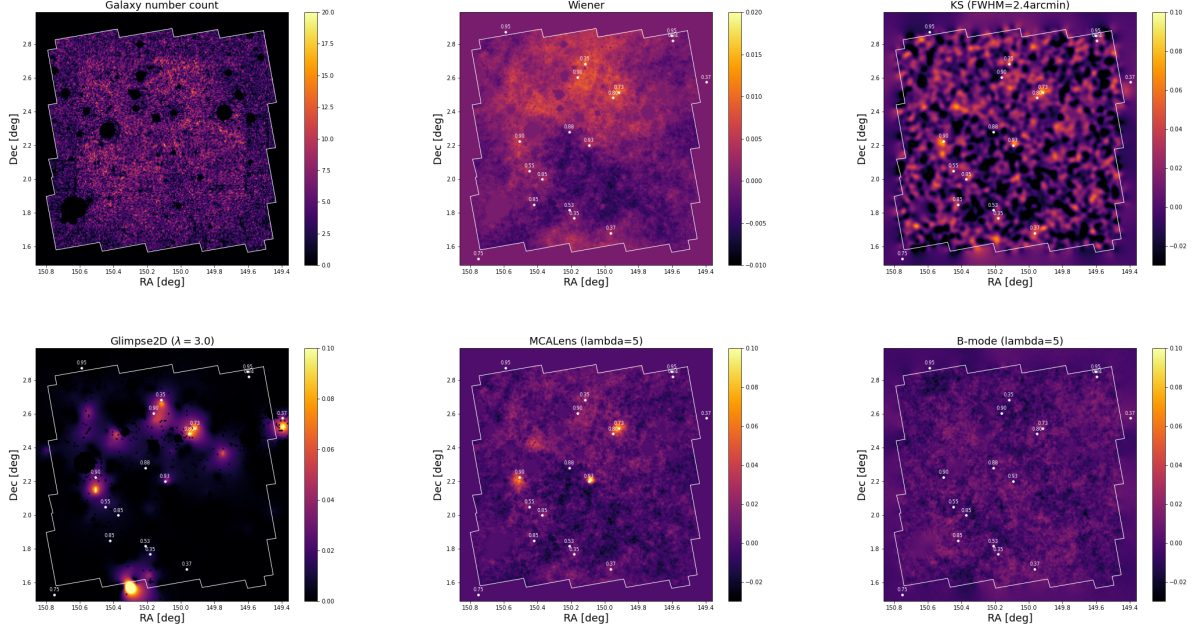

Fig. 7. COSMOS data. *Top panel*: galaxy count map, Wiener-filter map, and Kaiser-Squires map smoothed with a Gaussian with an FWHM of 2.4 arcmin. *Bottom panel*: Glimpse, MCAlens, and MCAlens *B*-mode maps.

iterations, $\lambda = 5$ (i.e. detection at 5σ), and we used the power spectrum of the simulation as the theoretical power spectrum. The top row of Fig. 6 shows the simulated noise-free map at 1 and 3 degrees, and the bottom row shows the MCAlens non-Gaussian and Gaussian components.

4.4. COSMOS field

In this last section, we apply MCAlens to reconstruct a convergence map of the 1.64 deg^2 HST/ACS COSMOS survey (Scoville et al. 2007). We used the bright galaxies shape catalogue produced in Schrabback et al. (2010).

The results after MCAlens was applied on COSMOS data are presented in Fig. 7. The top row shows the galaxy count map, the Wiener-filter map, and the Kaiser-Squires map smoothed with a Gaussian at a full width at half maximum of 2.4 arcmin. The bottom row shows the Glimpse, MCAlens *E*-mode, and MCAlens *B*-mode maps. White dots show the locations and redshifts of X-ray selected massive galaxy clusters from the *XMM-Newton* Wide Field Survey (Finoguenov et al. 2007) with $0.3 < z < 1.0$.

5. Conclusion

A novel mass-mapping algorithm has been presented that is able to recover high-resolution convergence maps from weak gravitational lensing measurements. Our proposed process involves a model with two components, a Gaussian and a non-Gaussian, for which we developed an efficient algorithm to derive the solution. We showed that we can also handle a non-diagonal covariance matrix. We extended the method so that it can deal with spherical maps, which is needed for future surveys such as the *Euclid* space mission. Our experiments clearly show a significant improvement compared to currently used algorithms.

In the spirit of reproducible research, the MCAlens algorithm is publicly available in the CosmoStat's Github package[4], including the material needed to reproduce the simulated experience (folder examples/mcalens_paper and script make_fig.py).

References

Abell, P. A., Burke, D. L., Hamuy, M., et al. 2009, Lsst science book, version 2.0, Tech. rep.
Ajani, V., Peel, A., Pettorino, V., et al. 2020, Phys. Rev. D, 102, 103531
Alsing, J., Heavens, A., Jaffe, A. H., et al. 2016a, MNRAS, 455, 4452
Alsing, J., Heavens, A., & Jaffe, A. H. 2016b, MNRAS, 466, 3272
Alsing, J., Heavens, A., & Jaffe, A. H. 2017, MNRAS, 466, 3272
André, P., Men'shchikov, A., Bontemps, S., et al. 2010, A&A, 518, L102
Bartelmann, M., & Schneider, P. 2001, Phys. Rep., 340, 291
Bobin, J., Starck, J. L., Sureau, F., & Fadili, J. 2012, Adv. Astron., 2012
Bobin, J., Sureau, F., & Starck, J. L. 2016, A&A, 591, A50
Carretero, J., Castander, F. J., Gaztañaga, E., Crocce, M., & Fosalba, P. 2015, MNRAS, 447, 646
Chang, C., Pujol, A., Mawdsley, B., et al. 2018, MNRAS, 475, 3165
Combettes, P. L., & Wajs, V. R. 2005, Multiscale Modeling & Simulation, 4, 1168
Crocce, M., Castander, F. J., Gaztanaga, E., Fosalba, P., & Carretero, J. 2015, MNRAS, 453, 1513
Elad, M., Starck, J.-L., Donoho, D., & Querre, P. 2005, Appl. Comput. Harmonic Anal., 19, 340
Elsner, F., & Wandelt, B. D. 2013, A&A, 549, A111
Finoguenov, A., Guzzo, L., Hasinger, G., et al. 2007, ApJS, 172, 182
Fluri, J., Kacprzak, T., Sgier, R., Refregier, A., & Amara, A. 2018, J. Cosmol. Astropart Phys., 2018, 051
Fosalba, P., Crocce, M., Gaztanaga, E., & Castander, F. J. 2015a, MNRAS, 448, 2987
Fosalba, P., Gaztanaga, E., Castander, F. J., & Crocce, M. 2015b, MNRAS, 447, 1319
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Hoffmann, K., Bel, J., Gaztanaga, E., et al. 2015, MNRAS, 447, 1724

Horowitz, B., Seljak, U., & Aslanyan, G. 2019, J. Cosmol. Astropart. Phys., 2019, 035
Jain, B., & Van Waerbeke, L. 2000, ApJ, 530, L1
Jasche, J., & Lavaux, G. 2014, MNRAS, 447, 1204
Jeffrey, N., Abdalla, F. B., Lahav, O., et al. 2018a, MNRAS, 479, 2871
Jeffrey, N., Heavens, A. F., & Fortio, P. D. 2018b, Astron. Comput., 25, 230
Jeffrey, N., Lanusse, F., Lahav, O., & Starck, J.-L. 2020, MNRAS, 492, 5023
Joseph, R., Courbin, F., Starck, J. L., & Birrer, S. 2019, A&A, 623, A14
Kaiser, N., & Squires, G. 1993, ApJ, 404, 441
Kilbinger, M. 2015, Rep. Prog. Phys., 78, 086901
Kilbinger, M., Fu, L., Heymans, C., et al. 2013, MNRAS, 430, 2200
Kratochvil, J. M., Lim, E. A., Wang, S., et al. 2012, Phys. Rev. D, 85
Lanusse, F., Starck, J.-L., Leonard, A., & Pires, S. 2016, A&A, 591, A2
Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints [arXiv:1110.3193]
Leonard, A., Dupé, F.-X., & Starck, J.-L. 2012, A&A, 539, A85
Li, Z., Liu, J., Matilla, J. M. Z., & Coulton, W. R. 2019, Phys. Rev. D, 99, 063527
Lin, C.-A., & Kilbinger, M. 2015, A&A, 576, A24
Liu, J., Bird, S., Matilla, J. M. Z., et al. 2018, J. Cosmol. Astropart. Phys., 2018, 049
Liu, J., & Haiman, Z. 2016, Phys. Rev. D, 94, 043533
Marian, L., Hilbert, S., Smith, R. E., Schneider, P., & Desjacques, V. 2011, ApJ, 728, L13
Mawdsley, B., Bacon, D., Chang, C., et al. 2020, MNRAS, 493, 5662
Melchior, P., Moolekamp, F., Jerdee, M., et al. 2018, Astron. Comput., 24, 129
Möller, A., Ruhlmann-Kleider, V., Lanusse, F., et al. 2015, J. Cosmol. Astropart. Phys., 2015, 041
Munshi, D., & Coles, P. 2017, J. Cosmol. Astropart. Phys., 2017, 010
Peel, A., Lin, C.-A., Lanusse, F., et al. 2017a, A&A, 599, A79
Peel, A., Lanusse, F., & Starck, J.-L. 2017b, ApJ, 847, 23
Petri, A., Haiman, Z., Hui, L., May, M., & Kratochvil, J. M. 2013, Phys. Rev. D, 88, 123002
Pires, S., Starck, J. L., Amara, A., et al. 2009, MNRAS, 395, 1265
Price, M. A., McEwen, J. D., Cai, X., et al. 2019, MNRAS, 489, 3236
Price, M. A., Cai, X., McEwen, J. D., et al. 2020, MNRAS, 492, 394
Price, M. A., McEwen, J. D., Pratley, L., & Kitching, T. D. 2021, MNRAS, 500, 5436
Repetti, A., Pereyra, M., & Wiaux, Y. 2019, SIAM J. Imaging Sci., 12, 87
Schrabback, T., Hartlap, J., Joachimi, B., et al. 2010, A&A, 516, A63
Scoville, N., Abraham, R. G., Aussel, H., et al. 2007, ApJS, 172, 38
Shirasaki, M., Yoshida, N., Hamana, T., & Nishimichi, T. 2012, ApJ, 760, 45
Spergel, D., Gehrels, N., Baltay, C., et al. 2015, ArXiv e-prints [arXiv:1503.03757]
Starck, J.-L., Candès, E., & Donoho, D. 2003, A&A, 398, 785
Starck, J.-L., Elad, M., & Donoho, D. 2004, Adv. Imaging Electron Phys., 132
Starck, J.-L., Moudden, Y., Abrial, P., & Nguyen, M. 2006a, A&A, 446, 1191
Starck, J.-L., Pires, S., & Réfrégier, A. 2006b, A&A, 451, 1139
Starck, J.-L., Murtagh, F., & Fadili, J. 2015, Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis (Cambridge University Press)
Tallada, P., Carretero, J., Casals, J., et al. 2020, Astron. Comput., 32, 100391
Teyssier, R. 2002, A&A, 385, 337
Upham, R. E., Whittaker, L., & Brown, M. L. 2020, MNRAS, 491, 3165
Wagner-Carena, S., Hopkins, M., Diaz Rivero, A., & Dvorkin, C. 2020, MNRAS, 494, 1507
Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A. 2004, Phys. Rev. D, 70, 083511
Wiener, N. 1949, Extrapolation, Interpolation, and Smoothing of Stationary Time Series: with Engineering Applications (Cambridge: MIT press), 7
Zaroubi, S., Hoffman, Y., Fisher, K. B., & Lahav, O. 1995, ApJ, 449, 446

---

4 https://github.com/CosmoStat/cosmostat

Appendix A: MICE simulations

We used the public MICE (v2) simulated galaxy catalogue, which is constructed from a light cone *N*-body dark matter simulation (Fosalba et al. 2015a,b; Crocce et al. 2015; Carretero et al. 2015; Hoffmann et al. 2015; Tallada et al. 2020). The MICE catalogue provides the calculated weak-lensing (noise-free) observables shear and convergence. In a given patch of simulated sky, we selected galaxies in the redshift[5] *z* range [0.6, 1.4]. Each galaxy corresponds to a noisy shear measurement, and we subsampled with a density of ∼8000 galaxies per deg^2.

Uncorrelated, complex shape noise values were randomly drawn from a Gaussian distribution and added to the shear value of each selected galaxy. This noise per galaxy is zero mean and has a variance $2\sigma_\epsilon^2 = 0.1636$ (as estimated from data (Jeffrey et al. 2018a)). The final pixelised noise (**n**) has a variance that depends on the number of galaxies per pixel.

In our simulated data, we mimicked these conditions by choosing to remove all galaxies in given regions. Here no shear measurements are available, and the noise variance is effectively infinite. The top row of Fig. A.1 shows a simulated convergence map with the DES SV footprint and mask, and the bottom row shows two simulated observed shear component maps, γ_1 and γ_2.
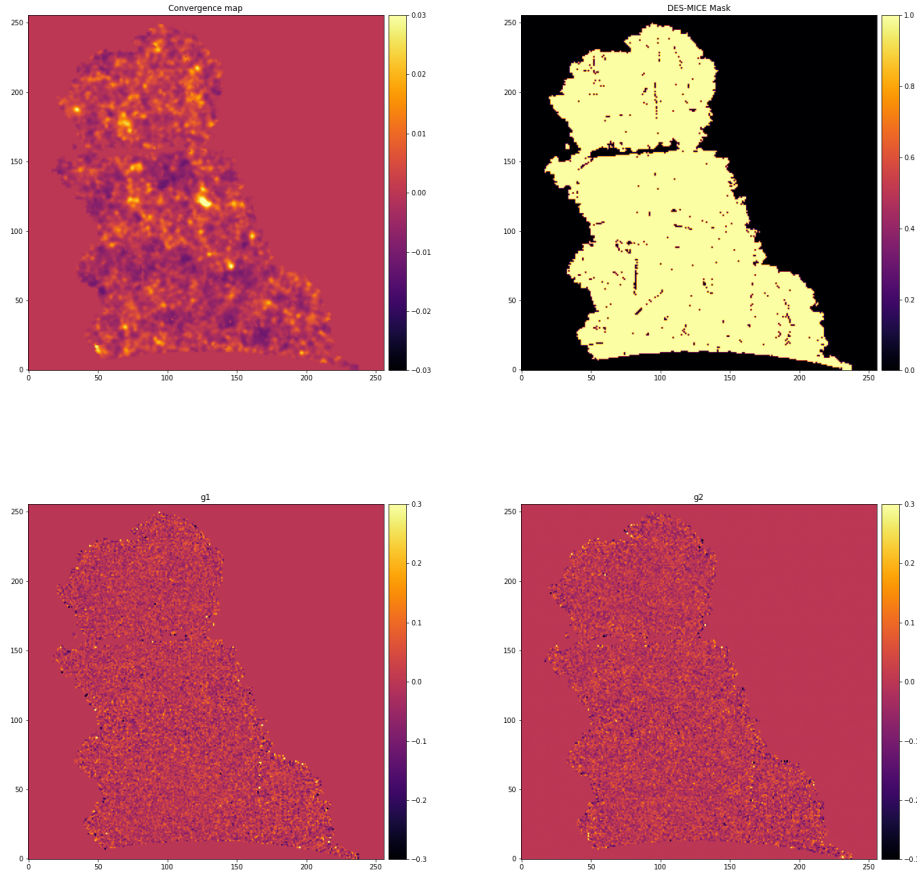

Fig. A.1. *Top*: true MICE convergence map with DES SV footprint and mask. *Bottom*: two simulated observed shear components, γ_1 and γ_2.

Appendix B: Wiener-filter tour

B.1. Wiener filter and inpainting

Missing data are a common problem for galaxy surveys because foreground objects that obscure the background galaxies have to be masked out. In addition to the non-stationary noise, observed shear fields therefore also present missing data. By noting the mask M as equal to 1 if we have shear measurements at a pixel position and zero otherwise, missing data can be handled in Wiener filtering by forcing the noise covariance matrix to be very high at locations where $M = 0$. This causes the solution to be different from zero and smoothed in the missing data area, and it remove border-effect artefacts. The Wiener filtering can therefore be seen as an inpainting technique because it fills the missing area in the image. Alternative inpainting techniques were proposed in the past through sparse recovery techniques (Pires et al. 2009; Lanusse et al. 2016) or Gaussian constraint realisations (Zaroubi et al. 1995; Jeffrey et al. 2018a). Because the Wiener model assumed the solution to be a Gaussian random field, it would make sense to have a solution where the inpainted area presents the same statistical properties as in non-inpainted area. This property is by construction verified with constraint realisations, and it was shown that this is also the case with sparse inpainting (Pires et al. 2009). A similar sparse inpainting can be very easily included in the proximal Wiener filtering, minimising the following equation:

$$\kappa_{\mathrm{G}} = \arg\min_{\kappa} \left\{ \|M(\gamma - \mathbf{A}\kappa)\|_{\Sigma_n}^2 + \beta\|\kappa\|_{\Sigma_\kappa}^2 + \lambda\|\Phi^*\kappa\|_p \right\}, \quad (B.1)$$

where $p = 0$ or 1, Φ is the discrete cosine eictionary (DCT), and λ is a Lagrangian parameter. We obtain the forward-backward algorithm, called **InpWiener**:

– Forward step:

$$\mathbf{t} = \kappa^n + 2\mu\mathbf{A}^*\Sigma_n^{-1}(M(\gamma - \mathbf{A}\kappa^n)) \quad (B.2)$$

– Backward step:

$$\kappa^{n+1} = M\mathbf{u} + (1 - M)\Delta_{\Phi,\lambda}\mathbf{u} \quad (B.3)$$

where and $\Delta_{\Phi,\lambda}$ is the proximal operator defined in Pires et al. (2009), which consists of applying a DCT transform to u, thresholding the DCT coefficients, and reconstructing an image from the thresholded coefficients, and

$$\mathbf{u} = F^* \left(\mathbf{P}_\kappa \left(\mathbf{P}_\eta + \mathbf{P}_\kappa \right)^{-1} \right) F\mathbf{t}. \quad (B.4)$$

Areas for which we possess information are processed as in the usual Wiener-filter case, while the inpainting regularisation affects area with missing data (i.e. when $M = 0$).

Concerning Eq. (B.1), it is interesting to note the following relations between **InpWiener** and other methods:

– KS: if $\beta = 0$, $\lambda = 0$ and Σ_n is diagonal with constant values along the diagonal (i.e. stationary Gaussian noise), **InpWiener** leads to the non-iterative standard Kaiser-Squires solution.

– GKS: If $\beta = 0$ and $\lambda = 0$, the least-squares estimator is derived with the iterative algorithm: $\kappa^{n+1} = \kappa^n + 2\mu\mathbf{A}^*\Sigma_n^{-1}(M(\gamma - \mathbf{A}\kappa^n))$, with $\mu = \min(\Sigma_n)$. As it generalises the Kaiser-Squires method, we call this algorithm **GKS**.

– FASTLens: If $\beta = 0$ and Σ_n is diagonal with constant values along the diagonal, **InpWiener** leads to the FASTLens inpainting algorithm (Pires et al. 2009).

– GIKS: If $\beta = 0$, **InpWiener** leads to an inpainted generalised the Kaiser-Squires solution where the **InpWiener** forward is unchanged, and the backward step becomes

$$\kappa^{n+1} = M\mathbf{t} + (1 - M)\Delta_{\Phi,\lambda}\mathbf{t}. \quad (B.5)$$

Similarly to Sects. 3.5 and 3.5, these algorithms can handle data on the sphere and jointly reconstruct E and B modes.

Inpainted Wiener-filter experiment. To test **InpWiener**, we used the public MICE (v2) simulated galaxy catalogue presented in Appendix A.

Figure B.1 shows the Wiener-filter solution (left) and the inpainted Wiener-filter solution (right) derived from the shear components shown in Fig. A.1.

B.2. Agnostic Wiener filtering

The Wiener-filter method needs to know the theoretical power spectrum P_κ, and the solution therefore varies with the assumed cosmological model used to derive P_κ. To avoid this issue, a solution could be to estimate P_κ directly from the shear measurements, for instance using a mask correction as in Upham et al. (2020). Bayesian techniques have also been used to infer the map and power spectrum (Wandelt et al. 2004; Jasche & Lavaux 2014; Alsing et al. 2016b). An alternative approach is to used the **GIKS** inpainting algorithm to first fill the missing area and then compute the power spectrum of the inpainted map. When **GIKS** is applied to the data and a set of R noise realisations, the final estimator is

$$P_\kappa = \mathrm{powspec}(\kappa_{\mathrm{Data}}) - \frac{1}{R} \sum_i \mathrm{powspec}(\kappa_{\mathrm{Rea}_i}). \quad (B.6)$$

Because the data noise P_κ will still be noisy, a final denoising step or a function fitting can be done.

As an illustration, we fitted the function $f(k, a, u, e, c) = \exp(|a * (uk)^{-e}|) + c$ to the estimated noisy P_κ. Figure B.2 shows an example of an estimated power spectrum following this approach.

Experiment: Effect of an unknown theoretical power spectrum. In this experiment, we used the same public MICE simulations, and we extracted 18 shear different shear maps with different noise realisations. For each of them, we applied the forward-backward Wiener-filter algorithm using the true theoretical power spectrum, and we applied the inpainted agnostic Wiener-filter method on the same data, re-estimating for each of the 18 shear data sets the theoretical power spectrum. Figure B.3 shows the reconstruction errors at different resolutions for Kaiser-Squires, Wiener filter, and inpainted agnostic Wiener filter. The inpainting clearly has no effect on the reconstruction error, which is expected because only the area where the mask is equal to one is used in the error calculation, and also because the agnostic approach has very little effect on the final solution. We do not claim that we should use an agnostic approach when a Wiener filtering is applied, but it is interesting to have this option available, for example when Wiener filtering is used without any assumption about the cosmology, and to give us the possibility to verify whether cosmological priors affect the results.
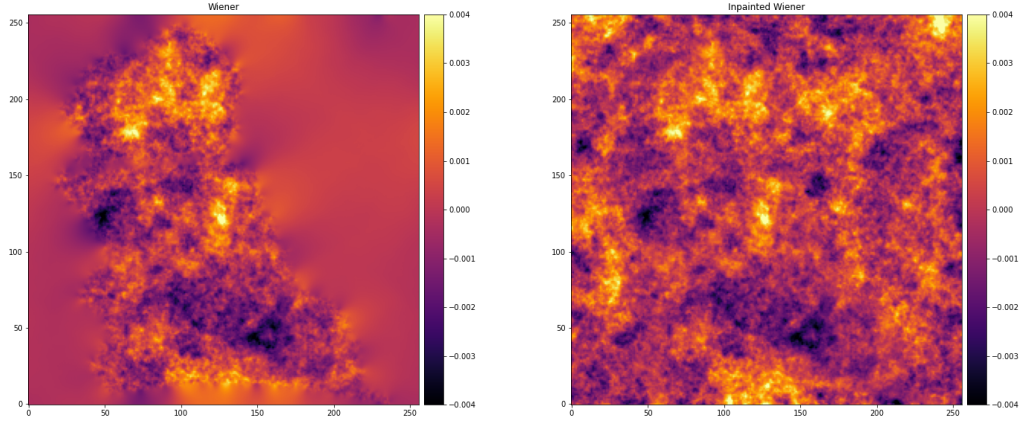
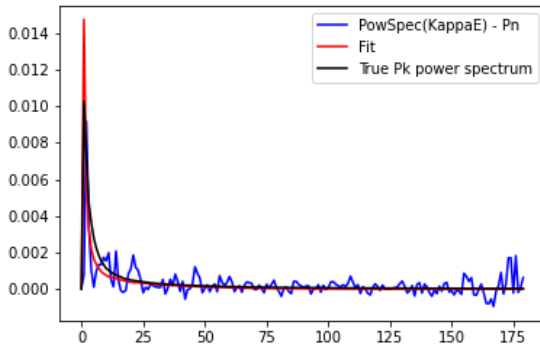**Fig. B.1.** Wiener and inpainted Wiener solutions.

Fig. B.2. Theoretical power spectrum of one convergence map. In black the true theoretical power spectrum; in red, the spectrum estimated using the **GIKS** algorithm, corrected from noise power spectrum, and in blue the fit.
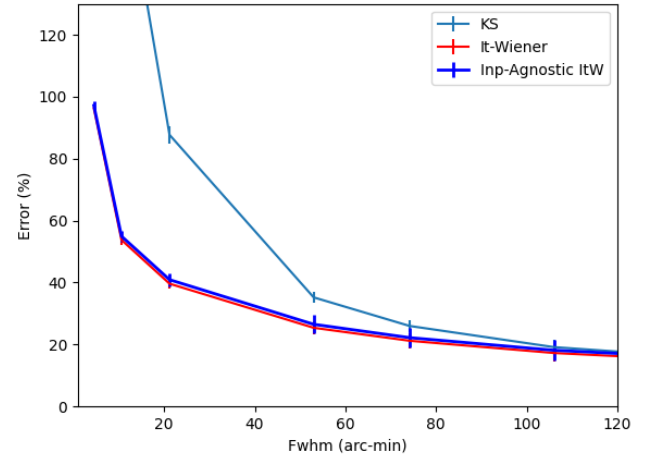

Fig. B.3. Reconstruction error at different resolution for Kaiser-Squires, Wiener filter, and inpainted agnostic Wiener filter.