Modelling Data with both Sparsity and a Gaussian Random Field: Application to Dark Matter Mass Mapping in Cosmology

Konstantinos E. Themelis*, Francois Lanusse†, Niall Jeffrey‡
, Austin Peel*, Jean-Luc Starck*, and Filipe B. Abdalla‡
* Laboratoire AIM, UMR CEA-CNRS-Paris, Irfu, DAp, CEA Saclay, 91191 Gif-sur-Yvette Cedex, France
† McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA
‡ Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

***Abstract*—In this paper we present a novel method for dark matter mass mapping reconstruction from weak gravitational lensing measurements. The crux of the proposed method lies in a new modelling of the matter density field in the Universe as a mixture of two components: a) a sparsity-based component that captures the non-Gaussian structure of the field, such as peaks or halos at different spatial scales; and b) a Gaussian random field, which is known to well represent the linear component of the field. This new model represents the distribution of matter in the universe much better than previously proposed models. We have developed a new algorithm that also takes into account the experimental problems we meet in practice, such as a non-diagonal covariance matrix of the noise or missing data. Experimental results on simulated data show that the proposed method exhibits improved estimation accuracy compared to state-of-the-art methods.**

I. Introduction

Images of distant galaxies are distorted by the gravitational potential of intervening massive structures along the line of sight. This effect is known as gravitational lensing, in analogy with the bending of light by optical lenses. Provided the right geometrical conditions, changes in the apparent shapes of galaxies due to lensing can sometimes be dramatic. For example, elongated arcs and multiple images are possible and belong to the strong lensing regime. The vast majority of galaxies we observe, however, experience only tiny shape distortions such that the signal is not detectable on an individual galaxy. This regime of small shape distortions is known accordingly as *weak gravitational lensing*. It is a relatively recent development in astrophysics and cosmology to exploit the statistical properties of galaxy shape measurements in order to reconstruct mass maps – maps of density fluctuations in the Universe. This is possible due to the imprint that cosmic structures leave in the coherent alignment of background galaxy shapes. Weak lensing is thus a powerful tool to probe the mass distribution of various structural components of the universe, including non-visible dark matter.

In signal processing terms, mass map reconstruction from weak lensing measurements can be viewed as an ill-posed inverse problem that usually involves two terms, the *shear* and the *convergence*. Shear quantifies the distortion of the shapes of background galaxies and is acquired through measurements available in galaxy surveys. Convergence is a dimensionless surface mass density (the projection of the 3D mass distribution on the 2D sky) that is related to the shear via some lensing potential and that we wish to estimate.

A widely used algorithm to perform mass mapping is the Kaiser-Squires method [1]. It is actually a simple least square estimator that can be carried out by an inversion formula in Fourier space, but it exhibits poor results since it takes no account of noise and missing data. A different approach, motivated also by the Bayesian framework, is that of Wiener filtering, [2]. In this approach a Gaussian random field structure is assumed as a prior for the convergence map, which is responsible for inserting some bias that prevents our solution from over-fitting, [3]. Moreover, a recently proposed state-of-the-art method is the Gravitational Lensing Inversion and MaPping using Sparse Estimators (GLIMPSE) algorithm, [4]. GLIMPSE is a highly sophisticated algorithm that takes advantage of the sparse regularization framework to solve the ill-posed linear inverse problem. GLIMPSE is based on sparse representations (i.e. wavelets), and is therefore well designed to recover piece-wise smooth features. It outperforms significantly other methods for peaks recovery. An analytical comparison between these three estimators is provided in [5].

In this paper we propose to bridge the gap between the sparse regularization method of GLIMPSE 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 now modelled as a mixture of two terms, a) a non-Gaussian term that adopts a sparse representation in a selected wavelet basis, [6], and b) a Gaussian term that is modelled using a Gaussian random field. The non-Gaussian signal component is able to capture the non-linear characteristics of the field, such as *peaks*, while the Gaussian component of the signal is responsible of capturing the lower-frequency characteristics of the underlying field, such as smooth variations. Under this assumption, we formulate a two-step optimization process. First, we utilize the GLIMPSE algorithm to estimate the non-Gaussian component of the density, whereas, in the second, we employ a Wiener filtering approach on the shear residuals in order to estimate the Gaussian signal component. It should be noted that for

ISBN 978-90-827970-1-5

implementing Wiener filtering we employ a proximal calculus method recently presented in [7]. Experimental results on simulated data provide evidence that the proposed method improves the estimation performance of the GLIMPSE algorithm in mass mapping recovery.

II. State of the art

Let us consider a set of noisy weak lensing shear measurements, denoted by the vector $\mathbf{y} \in \mathbb{C}^M$ in the following, where M is the number of galaxies in our survey. We assume that $\mathbf{y}$ is generated as a linear combination of the known $M \times N$ matrix $\mathbf{H}$ and the unknown underlying field $\mathbf{x} \in \mathbb{R}^N$, i.e.,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where the additive noise $\mathbf{n}$ is assumed to be zero-mean Gaussian, $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_n)$. The matrix $\mathbf{H}$ represents a set of known linear operations that are responsible for the transformation from convergence to shear in the weak lensing limit. Typically $\mathbf{H}$ can be decomposed as $\mathbf{H} = \mathcal{F}\mathbf{P}\mathcal{F}^H$, where $\mathcal{F}$ denotes the Fourier transform, $\mathcal{F}^H$ is its adjoint, and $\mathbf{P}$ is the diagonal operator that defines the convergence field-shear connection in Fourier space, i.e.,

$$\tilde{\mathbf{y}} = \mathbf{P}\tilde{\mathbf{x}} = \left(\frac{k_1^2 - k_2^2}{k^2} + \imath\frac{2k_1k_2}{k^2}\right)\tilde{\mathbf{x}} \tag{2}$$

where $k^2 = k_1^2 + k_2^2$ and k_1 and k_2 are frequency components of $\tilde{\mathbf{x}} = \mathcal{F}\mathbf{x}$. The objective of this paper is to solve the inverse problem of estimating the convergence field $\mathbf{x}$ after observing the noisy shear measurements $\mathbf{y}$.

Recently published state-of-the-art methods address this problem by considering either a Wiener filtering approach that utilizes a Gaussian-type prior over $\mathbf{x}$, e.g., [3], [5], or by exploiting a sparse representation of the convergence field in the wavelet domain, e.g., [4]. Unfortunately, both these approaches may not be sufficient to actually model the real data. Specifically, the Gaussian prior considered in the Wiener filtering approach can only capture the latent low-frequency part of the signal and has a typical smoothing effect on the recovered field. On the other hand, the sparsity-enforcing ℓ_1 norm regularization approach of [4] may as well capture the high-frequency components of the data, such as peaks, but it is unable to model the smoothly varying signal components. In the following section, we show how our two mixtures modelling leads to get the best of these two approaches.

III. Modelling with sparsity and a Gaussian random field

To address the limitations above, a novel modelling approach is proposed here, where the convergence $\mathbf{x}$ is assumed to comprise two parts, a Gaussian and a non-Gaussian, i.e.,

$$\mathbf{x} = \mathbf{x}_{\mathrm{NG}} + \mathbf{x}_{\mathrm{G}}. \tag{3}$$

The non-Gaussian part of the signal $\mathbf{x}_{\mathrm{NG}}$ is subject to a sparse decomposition in a wavelet dictionary, while the component $\mathbf{x}_{\mathrm{G}}$ is assumed to be inherently non-sparse and Gaussian. Under this assumption, we may recast the data model in (1) as

$$\mathbf{y} = \mathbf{H}\mathbf{x}_{\mathrm{NG}} + \mathbf{n}', \tag{4}$$

where $\mathbf{n}' = \mathbf{H}\mathbf{x}_{\mathrm{G}} + \mathbf{n}$. Note that (4) expresses a sparse recovery problem under the presence of correlated noise. To tackle this problem, we develop a sequential optimization scheme, where both GLIMPSE and Wiener filtering are utilized in order to estimate the convergence components $\mathbf{x}_{\mathrm{NG}}$ and $\mathbf{x}_{\mathrm{G}}$, respectively.

A. Proposed optimization

We solve the sparse recovery problem of (4) using a two-step optimization procedure. In the first step, we utilize the GLIMPSE algorithm to recover a non-Gaussian component of the convergence. In the second, a Wiener filtering approach is used to recover any residual Gaussian component. Hence, adopting the GLIMPSE and Wiener cost functions, we can express the proposed minimization tasks as,

$$\min_{\mathbf{x}_{NG}} \left\{ \frac{1}{2}\| \underbrace{\mathbf{y} - \mathbf{H}\mathbf{x}_{NG}}_{\mathbf{y}_r} \|_{\mathbf{\Sigma}}^2 + \lambda\|\mathbf{w} \odot \mathbf{\Phi}^*\mathbf{x}_{NG}\|_1 + \mathcal{I}_{\mathbb{R}^+}(\mathbf{x}_{NG}) \right\} \tag{5}$$

and

$$\min_{\mathbf{\Omega}, \mathbf{x}_G} \left\{ \|\mathbf{\Omega}\mathbf{y}_r - \mathbf{H}\mathbf{x}_G\|_{\mathbf{\Sigma}_n^{1/2}}^2 \right\}, \tag{6}$$

where $\mathbf{\Sigma}$ is the shear covariance matrix, λ is the sparsity regularization parameter, $\mathbf{w}$ is a vector containing the weighting coefficients of the ℓ_1-norm, $\odot$ denotes the Hadamard product, $\mathbf{\Phi}$ is the wavelet basis, $\mathcal{I}_{\mathbb{R}^+}(\cdot)$ is the indicator function defined as $\mathcal{I}_{\mathbb{R}^+}(\mathbf{x}) = 0(\infty)$ if $\mathbf{x} \in \mathbb{R}^+ (\mathbf{x} \notin \mathbb{R}^+)$, $\mathbf{\Sigma}_n$ is the covariance matrix of the additive noise $\mathbf{n}$. and $\mathbf{\Omega}$ is the matrix containing the Wiener filter coefficients. Using GLIMPSE to solve (5), we get an estimate $\hat{\mathbf{x}}_{NG}$ of the non-Gaussian component, that allows us to compute the corresponding shear residuals vector $\mathbf{y}_r = \mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_{NG}$. In turn, this vector is binned in pixel space in order to serve as input for the second optimization task in (6). The Wiener filtering solution of (6) can then be expressed as,

$$\hat{\mathbf{x}}_G = \left[\mathbf{H}^T\mathbf{\Sigma}_n^{-1/2}\mathbf{H} + \mathbf{\Sigma}_{\mathbf{x}_G}^{-1}\right]^{-1}\mathbf{H}^T\mathbf{\Sigma}_n^{-1/2}\mathbf{y}_r, \tag{7}$$

where $\mathbf{\Sigma}_{\mathbf{x}_G}$ denotes the covariance matrix of $\mathbf{x}_G$.

At this point, it is important to elaborate further on the implementation of the Wiener filtering solution in (7). Note first that the matrix $\mathbf{\Sigma}_n$ is diagonal in pixel space, while $\mathbf{\Sigma}_{\mathbf{x}_G}$ is diagonal in Fourier space. However, the matrix to be inverted in expression (7) is clearly dense and of high dimension, i.e. of the size of the number of pixels in the mass map. This renders its inversion computationally complex and prone to numerical errors. To circumvent this, we employ a proximal calculus-based forward-backward splitting (FBS) technique for the Wiener filtering step, similar to the one proposed in [7].

ISBN 978-90-827970-1-5 © EURASIP 2018

To this end, we express the Wiener optimization function (6) in an equivalent form as,

$$\min_{\mathbf{x}_G} \left\{ \|\mathbf{y}_r - \mathbf{H}\mathbf{x}_G\|^2_{\mathbf{\Sigma}_n^{1/2}} + \|\mathbf{x}_G\|^2_{\mathbf{\Sigma}_{\mathbf{x}_G}} \right\}. \tag{8}$$

The advantage of the formulation in (8) is that it comprises separable terms, e.g. $f_1(\mathbf{x}_G) = \|\mathbf{y}_r - \mathbf{H}\mathbf{x}_G\|^2_{\mathbf{\Sigma}_n^{1/2}}$ and $f_2(\mathbf{x}_G) = \|\mathbf{x}_G\|^2_{\mathbf{\Sigma}_{\mathbf{x}_G}}$. Following the FBS methodology, we can solve (8) by designing an iterative fixed point algorithm as

$$\mathbf{x}_G^{k+1} = \mathrm{prox}_{\gamma f_2}(\mathbf{x}_G^k - \gamma \nabla f_1(\mathbf{x}_G^k)), \tag{9}$$

which is known to converge when $\gamma < 2/\|\mathbf{H}^T\mathbf{\Sigma}_n^{-1/2}\mathbf{H}\|_2$, [8]. Computing the proximal operator in (9), we end up with the following iterative Wiener filtering algorithm,

$$\mathbf{t} = \mathbf{x}_G^k + \mathbf{T}\mathbf{H}^T\mathbf{\Sigma}_n^{-1/2}\left(\mathbf{y}_r - \mathbf{H}\mathbf{x}_G^k\right) \tag{10}$$

$$\mathbf{x}_G^{k+1} = \mathcal{F}^H \left(\frac{\mathbf{P}_{\mathbf{x}_G}}{\mathbf{P}_\mathbf{t} + \mathbf{P}_{\mathbf{x}_G}} \right) \mathcal{F}\mathbf{t} \tag{11}$$

where $\mathbf{t}$ is an auxiliary variable with covariance matrix $\mathbf{T} = \mathrm{diag}(2 \times \min(\Sigma_\mathrm{n}^{1/2}))$ and power spectrum $\mathbf{P}_\mathbf{t}$. Notice that the proposed algorithm is free from matrix inversions, since both $\mathbf{\Sigma}_n$ used in (10) and $\mathbf{P}_{\mathbf{x}_G} = \mathcal{F}\mathbf{\Sigma}_{\mathbf{x}_G}\mathcal{F}^H$ used in (11) are diagonal.

Furthermore, in the following we propose two methods to estimate the signal power spectrum $\mathbf{P}_{\mathbf{x}_G}$, which is required in (11). Let us first define the signal

$$\mathbf{r}^k = \mathbf{T}\mathbf{H}^T\mathbf{\Sigma}_n^{-1/2}\left(\mathbf{y}_r - \mathbf{H}\mathbf{x}_G^k\right) \tag{12}$$

so that we can express (10) as $\mathbf{t} = \mathbf{x}_G^k + \mathbf{r}^k$. We observe from (12) that $\mathbf{r}^k$ is whitened, i.e. its power spectrum is equal to one. Hence, to get a noisy estimate of $\mathbf{P}_{\mathbf{x}_G}$ at each iteration it suffices to calculate the power spectrum of $\mathbf{t}$ and subtract one for the contribution of $\mathbf{r}^k$, i.e.,

$$\hat{\mathbf{P}}_{\mathbf{x}_G} = [\mathbf{P}_\mathbf{t} - 1]_+. \tag{13}$$

A second approach is to assume that the signal $\mathbf{x}_G$ and the noise $\mathbf{n}$ are uncorrelated and, hence, their power spectra are associated as

$$\mathbf{P}_{\mathbf{x}_G} = [\mathbf{P}_\mathbf{x} - \mathbf{P}_{\mathbf{x}_{NG}}]_+ \tag{14}$$

where $\mathbf{P}_{\mathbf{x}_{NG}}$ can be easily computed from the data, while $\mathbf{P}_\mathbf{x}$ can be known from cosmological models. In this paper we utilized the approach in (14) to obtain a good approximation of $\mathbf{P}_{\mathbf{x}_G}$.

IV. Experimental results

To test the quality of the proposed reconstruction method we use realistic simulated data for which an underlying truth is known. We use the public MICE (v1) simulated galaxy catalogue, which is constructed from a lightcone N-body dark matter simulation, [9]–[14]. The MICE catalogue provides the calculated weak lensing (noise-free) observables: shear and convergence. In a given patch of simulated sky we select galaxies in the redshift (Redshift due to the expansion of the Universe is used as a proxy for distance to a galaxy, as it is an observable that monotonically increases with distance from an observer). z range $[0.6, 1.4]$. Each galaxy corresponds to a noisy shear measurement, and we subsample with a density of ~ 8000 galaxies per deg^2.

Uncorrelated, complex shape noise values are 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 variance $2\sigma_\epsilon^2 = 0.1636$ (as estimated from data [5]). The final pixelised noise ($\mathbf{n}$) has variance that depends on the number of galaxies per pixel.

Missing data is a common problem for galaxy surveys as foreground objects obscuring the background galaxies have to be "masked out". In our simulated data, we mimic these conditions by choosing to remove all galaxies in given regions. Here there are no shear measurements available and the noise variance is effectively infinite. Maps of the noisy input shear data along with the noise variance are shown in Fig. 1 with the mask applied. Panel (c) shows the true convergence map of this field. Given the high level of noise, which is common in real observations, correlations between the true convergence and the shear field are not detectable by eye.

Figure 2c shows the results with our new method. It shows that both the non-linear and the linear components can be recovered, which has not been achieved by any method in the past. For comparison, panels (a) and (b) show the results from GLIMPSE, which recovers only the non-Gaussian features, and the result from Wiener filtering, which recovers only the Gaussian part of the signal.

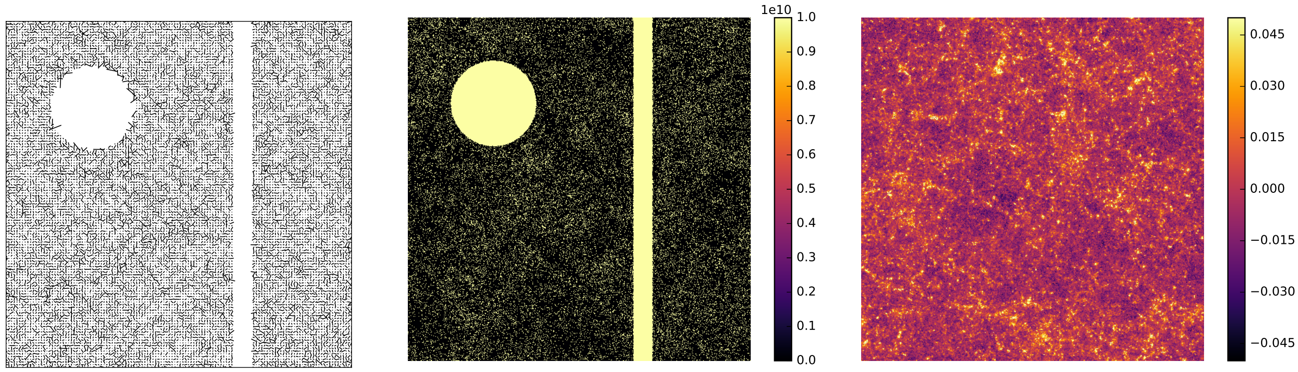
V. Conclusion

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 modelling with two components, a Gaussian one and a non-Gaussian one, and we have developed an efficient algorithm to derive the solution. We have shown that we can also handle a non-diagonal covariance matrix and missing data, so the method can be used in future experiments such as the Euclid space mission. The experiments clearly show a significant improvement compared to the state of art.

Acknowledgment

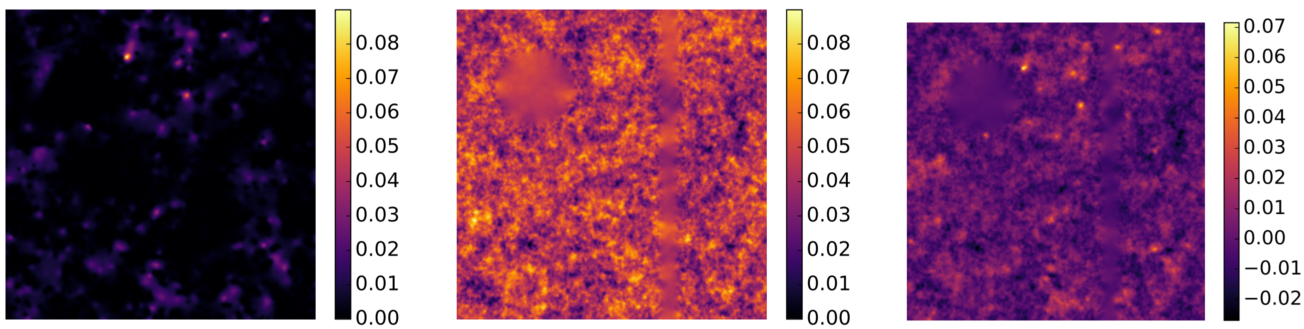
This work is funded by the DEDALE project, contract no. 665044, within the H2020 Framework Program of the European Commission.

We acknowledge the use of data from the MICE simulations, publicly available at "http://www.ice.cat/mice". This work has been made possible thanks to CosmoHub. CosmoHub has been developed by the Port d'Informaci Cientfica (PIC), maintained through a collaboration of the Institut de Fsica d'Altes Energies (IFAE) and the Centro de Investigaciones Energticas, Medioambientales y Tecnolgicas (CIEMAT). The work was partially funded by the "Plan Estatal de Investigacin Cientfica y Tcnica y de Innovacin" program of the Spanish government.



(a) Noisy input shear map with mask (b) Noise variance map with mask (c) True convergence map

Fig. 1. Simulated weak lensing maps extracted from MICE data. Shown from left to right are (a) the masked shear with noise serving as input to the reconstruction algorithms, (b) the masked noise variance map, and (c) the convergence map corresponding to the true shear for this field.



(a) GLIMPSE convergence map (b) Wiener filtering convergence map (c) Proposed map

Fig. 2. Convergence maps reconstructed from the shear using three algorithms. The maps show (a) the GLIMPSE result, which recovers the non-Gaussian part of the signal (cf. Fig. 1c), (b) the result from Wiener filtering, which recovers the Gaussian information, and (c) the result from our proposed method, which recovers both parts simultaneously.

REFERENCES

[1] N. Kaiser and G. Squires., "Mapping the dark matter with weak gravitational lensing," *Astrophysical Journal*, vol. 404, pp. 441–450, feb 1993.

[2] N. Wiener, "Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications," vol. 7, 1949.

[3] S. Zaroubi, Y. Hoffman, K. B. Fisher, and O. Lahav, "Wiener reconstruction of the large-scale structure," *Astrophysical Journal*, vol. 449, p. 446, Aug. 1995.

[4] F. Lanusse, J.-L. Starck, A. Leonard, and S. Pires, "High resolution weak lensing mass mapping combining shear and flexion," *A&A*, vol. 591, p. A2, 2016.

[5] N. Jeffrey, F. B. Abdalla, O. Lahav, F. Lanusse, J. L. Starck *et al.*, "Improving weak lensing mass map reconstructions using gaussian and sparsity priors: Application to DES SV," 2018.

[6] J.-L.Starck, F. Murtagh, and J. Fadili, *Sparse image and signal processing: wavelets, curvelets, morphological diversity*. Cambridge university press, 2010.

[7] J. Bobin, J.-L. Starck, F. Sureau, and J. Fadili, "CMB map restoration," *Advances in Astronomy*, vol. 2012, 2012.

[8] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[9] P. Fosalba, M. Crocce, E. Gaztañaga, and F. J. Castander, "The MICE grand challenge lightcone simulation - I. Dark matter clustering," *Monthly Notices of the Royal Astronomical Society*, vol. 448, pp. 2987–3000, Apr. 2015.

[10] M. Crocce, F. J. Castander, E. Gaztañaga, P. Fosalba, and J. Carretero, "The MICE Grand Challenge lightcone simulation - II. Halo and galaxy catalogues," *Monthly Notices of the Royal Astronomical Society*, vol. 453, pp. 1513–1530, Oct. 2015.

[11] P. Fosalba, E. Gaztañaga, F. J. Castander, and M. Crocce, "The MICE Grand Challenge light-cone simulation - III. Galaxy lensing mocks from all-sky lensing maps," *Monthly Notices of the Royal Astronomical Societyy*, vol. 447, pp. 1319–1332, Feb. 2015.

[12] J. Carretero, F. J. Castander, E. Gaztañaga, M. Crocce, and P. Fosalba, "An algorithm to build mock galaxy catalogues using MICE simulations," *Monthly Notices of the Royal Astronomical Society*, vol. 447, pp. 646–670, Feb. 2015.

[13] K. Hoffmann, J. Bel, E. Gaztañaga, M. Crocce, P. Fosalba *et al.*, "Measuring the growth of matter fluctuations with third-order galaxy correlations," *Monthly Notices of the Royal Astronomical Society*, vol. 447, pp. 1724–1745, Feb. 2015.

[14] J. Carretero, P.Tallada, J.Casals, M.Caubet, C.Neissner *et al.*, "Cosmohub on hadoop: a web portal to analyze and distribute cosmology data." *in prep.*, 2016.

ISBN 978-90-827970-1-5 © EURASIP 2018