

Cosmological constraints from the capture of non-Gaussianity in Weak Lensing data

Sandrine Pires^{*}, Adrienne Leonard and Jean-Luc Starck

Laboratoire AIM, CEA/DSM-CNRS-Universit  Paris Diderot, IRFU/SEDI-SAP, Service d'Astrophysique, CEA Saclay, Orme des Merisiers, 91191 Gif-sur-Yvette, France

Released 2011 Xxxxx XX

ABSTRACT

Weak gravitational lensing has become a common tool to constrain the cosmological model. The majority of the methods to derive constraints on cosmological parameters use second-order statistics of the cosmic shear. Despite their success, second-order statistics are not optimal and degeneracies between some parameters remain. Tighter constraints can be obtained if second-order statistics are combined with a statistic that is efficient to capture non-Gaussianity. In this paper, we search for such a statistical tool and we show that there is additional information to be extracted from statistical analysis of the convergence maps beyond what can be obtained from statistical analysis of the shear field. For this purpose, we have carried out a large number of cosmological simulations along the σ_8 - Ω_m degeneracy, and we have considered three different statistics commonly used for non-Gaussian features characterization: skewness, kurtosis and peak count. To be able to investigate non-Gaussianity directly in the shear field we have used the aperture mass definition of these three statistics for different scales. Then, the results have been compared with the results obtained with the same statistics estimated in the convergence maps at the same scales. First, we show that shear statistics give similar constraints to those given by convergence statistics, if the same scale is considered. In addition, we find that the peak count statistic is the best to capture non-Gaussianities in the weak lensing field and to break the σ_8 - Ω_m degeneracy. We show that this statistical analysis should be conducted in the convergence maps: first, because there exist fast algorithms to compute the convergence map for different scales, and secondly because it offers the opportunity to denoise the reconstructed convergence map, which improves non-Gaussian features extraction.

Key words: Cosmology: Weak Lensing, Methods: Data Analysis

1 INTRODUCTION

Gravitational light deflection, caused by large scale structure along the line-of-sight, produces an observable pattern of alignments in the images of distant galaxies. This distortion of the images of distant galaxies by gravitational lensing, called cosmic shear, offers an opportunity to directly probe the total matter distribution of the Universe, and not just the luminous matter. Therefore, the statistical properties of this gravitational shear field are directly linked to the statistical properties of the total matter distribution and can thus be directly compared to theoretical models of structure formation. Despite some systematics (PSF distortion, intrinsic alignments...), this approach is extremely attractive since it is unaffected by the biases characteristic of methods based only on the light distribution.

Since its first detection (Van Waerbeke et al. 2000; Kaiser et al. 2000; Wittman et al. 2000; Bacon et al. 2000), cosmic shear has rapidly become a major tool to constrain the cosmological model (for review, see e.g. Mellier 1999; Bartelmann and Schneider 2001; Refregier 2003; Hoekstra and Jain 2008; Munshi et al. 2008).

In most weak lensing studies, second-order statistics are the most commonly used statistical probe (e.g. Maoli et al. 2001; Hoekstra et al. 2006; Benjamin et al. 2007; Fu et al. 2008) because of their potential to constrain the power spectrum of density fluctuations in the late Universe. However, second-order statistics are not optimal to constrain cosmological parameters. For example, they only depend on a degenerate combination of the amplitude of matter fluctuations σ_8 and the matter density parameter Ω_m (Maoli et al. 2001; Refregier et al. 2002; Bacon et al. 2003; Massey et al. 2005; Dahle 2006).

The optimality of second-order statistics to constrain

^{*} Email: sandrine.pires@cea.fr

cosmological parameters depends heavily on the assumption of Gaussianity of the field. However the weak lensing field is composed, at small scales, of non-Gaussian features such as clusters of galaxies. These non-Gaussian signatures, which can be measured via higher-order moments, carry additional information that cannot be extracted with second-order statistics. Since the non-Gaussianity is induced by the growth of structures, it holds important cosmological information. Many studies (e.g. Bernardeau et al. 1997a; Takada and Jain 2004; Kilbinger and Schneider 2005; Pires et al. 2009a; Bergé et al. 2010) have shown that combining second-order statistics with higher-order statistics tighten the constraints on cosmological parameters.

Most non-Gaussian studies (e.g. Schneider et al. 1998; Jarvis et al. 2004; Kilbinger and Schneider 2005; Dietrich and Hartlap 2010) have been performed in the shear field because it can be directly derived from the shape of galaxies. This paper aims to produce evidence that there is additional information to be extracted from a higher-order statistical analysis of the convergence maps beyond what can be obtained from a higher-order statistical analysis of the shear field because higher-order statistics are probing the non-Gaussian features of the signal and these non-Gaussian structures can be better reconstructed in convergence maps using a denoising. In Pires et al. (2009a), the efficiency of several higher-order convergence statistics have been compared to discriminate cosmological models along the σ_8 - Ω_m degeneracy. In this paper, we are interested in showing the advantage of using these higher-order convergence statistics compared to higher-order shear statistics. This comparison cannot be performed directly because the evaluation of non-Gaussian statistics in the shear field requires to use their aperture mass definition (Schneider et al. 1998) or another different filter that is defined for a given scale θ . A fair comparison with convergence statistics requires the statistics in the convergence maps to be estimated at the same scale. A stationary wavelet transform, the "à trous" wavelet transform, has been used in this paper to compute the convergence statistics at the given scale θ .

The paper is organized as follows. In section 2, the cosmological models selected for this study are described, followed by a short description of the weak lensing simulations. Section 3 summarizes the different statistics used in this study. We give the definition of the aperture mass M_{ap} and present the three shear statistics considered in this study. Then, the "à trous" wavelet transform is defined, as well as the three related convergence statistics. Section 4 presents our results and we summarize our conclusions in section 5.

2 SIMULATIONS OF WEAK LENSING MASS MAPS

N-body simulations have been used to numerically compute the variation of the different statistics with cosmological parameters and then compare their ability to break degeneracies. We carried out N-body simulations for 5 different cosmological models along the σ_8 - Ω_m degeneracy corresponding to current constraints from a power spectrum analysis. The N-body simulations were carried out using the RAMSES code (Teyssier 2002). Fig. 1 shows the distribution of these cosmological simulations in the σ_8 - Ω_m

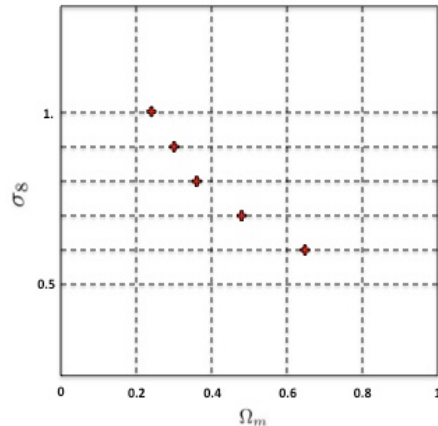


Figure 1. Location of the 5 simulated cosmological models in the σ_8 - Ω_m plane. The 5 cosmological models have been selected along the σ_8 - Ω_m degeneracy corresponding to current constraints from a power spectrum analysis.

plane. The characteristics of these cosmological models have been given in Pires et al. (2009a). For each cosmological model, we have run 100 realizations in order to quantify the observational uncertainties.

In the N-body simulations that are commonly used in cosmology, the dark matter distribution is represented by discrete massive particles. The simplest way of treating these particles is to map their positions onto a pixelized grid. In the case of multiple sheet weak lensing, we do this by taking slices through the 3D simulations. These slices are then projected into 2D mass sheets. The effective convergence can subsequently be calculated by stacking a set of these 2D mass sheets along the line of sight, using the lensing efficiency function. This is a procedure that was used before by Vale and White (2003), where the effective 2D mass distribution κ_e is calculated by integrating the density fluctuation along the line of sight. Using the Born approximation, which assumes that the light rays follow straight lines, the convergence can be numerically expressed by

$$\kappa_e \approx \frac{3H_0^2 \Omega_m L}{2c^2} \sum_i \frac{\chi_i (\chi_0 - \chi_i)}{\chi_0 a(\chi_i)} \left(\frac{n_p R^2}{N_t s^2} - \Delta r_{fi} \right), \quad (1)$$

where H_0 is the Hubble constant, Ω_m is the density of matter, c is the speed of light, L is the length of the box, and χ are comoving distances where χ_0 is the distance to the source galaxies. The summation is performed over the i^{th} box. The number of particles associated with a pixel of the simulation is n_p , the total number of particles within a simulation is N_t , and $s = L_p/L$, where L_p is the length of the plane responsible for the lensing. R is the size of the 2D maps and $\Delta r_{fi} = \frac{r_2 - r_1}{L}$, where r_1 and r_2 are comoving distances.

For each of our 5 models, we have run 21 N-body simulations, each containing 256^3 particles. We have used these 3D N-body simulations to derive 100 realizations of the convergence field for each model. The field is $3.95^\circ \times 3.95^\circ$ and it has been downsampled to 1024×1024 pixels (1 pixel = $0.23'$). Figure 2 shows one of these convergence maps. As said previously, the 5 cosmological models have been selected along the σ_8 - Ω_m degeneracy corresponding to current constraints

from a power spectrum analysis. Additionally, we have verified that the power spectrum of the 5 cosmological models are still degenerated with the survey area considered.

The convergence map κ that corresponds to the projected matter density is not directly observable but, it can be derived from the observed shear maps γ using the following relation (Kaiser and Squires 1993; Starck et al. 2006):

$$\hat{\kappa} = \hat{P}_1 \hat{\gamma}_1 + \hat{P}_2 \hat{\gamma}_2, \quad (2)$$

with :

$$\hat{P}_1(\mathbf{k}) = \frac{k_1^2 - k_2^2}{k^2},$$

$$\hat{P}_2(\mathbf{k}) = \frac{2k_1 k_2}{k^2},$$

where the hat symbol denotes Fourier transform. Inversely, the shear maps γ_i can easily be derived from the convergence map using the following relation:

$$\hat{\gamma}_i = \hat{P}_i \hat{\kappa} \quad (3)$$

In practice, observed shear maps γ_i^{obs} are obtained by averaging over a finite number of galaxies and are therefore noisy: $\gamma_i^{obs} = \gamma_i + N_i$, where N_1 and N_2 are noise contributions with zero mean and standard deviation $\sigma_n = \sigma_\epsilon / \sqrt{A \cdot n_g}$ with A being the area of the pixel in arcmin^2 and n_g the average number of galaxies per arcmin^2 . Typical values for the density of galaxies is $n_g = 30 \text{ gal/arcmin}^2$ for ground-based simulations and $n_g = 100 \text{ gal/arcmin}^2$ for space-based simulations. Although a bit optimistic, these two configurations have been considered to derive noisy shear maps and to compute the shear statistics described in §3.1. The derived noisy shear maps are downsampled to 1024×1024 pixels (1 pixel = $0.23'$) like the simulated convergence maps. An estimation of the corresponding noisy convergence maps can be derived from the equation [2]: $\hat{\kappa}_n = \hat{\kappa} + \hat{N}$, where $\hat{N} = \hat{P}_1 \hat{N}_1 + \hat{P}_2 \hat{N}_2$. As follows, the noise N in κ_n is still Gaussian and uncorrelated. The inversion does not amplify the noise, but κ_n may be dominated by the noise if N is large, which is the case in practice. Ground-based and space-based simulations of convergence maps have been derived by this way. The noisy convergence maps derived by inversion are still $3.95^\circ \times 3.95^\circ$ downsampled to 1024×1024 pixels (1 pixel = $0.23'$) and they have been used to compute the convergence statistics described in §3.3.

3 WEAK LENSING STATISTICS

The properties of the shear field γ (or associated convergence field κ) can be measured statistically and reveal precious information about cosmological parameters. Up to now, cosmic shear studies have focused mainly on second-order statistics which only probe the Gaussian part of the matter distribution. However, the matter distribution is composed of non-Gaussian features such as the clusters of galaxies that are the result of the non-linear evolution of the primordial Gaussian field. Therefore, higher-order statistics are required to probe the non-Gaussian part of the field and thus improve our constraints on cosmological parameters. This study will focus on these higher-order statistics estimated both in the shear maps γ and in the convergence maps κ .

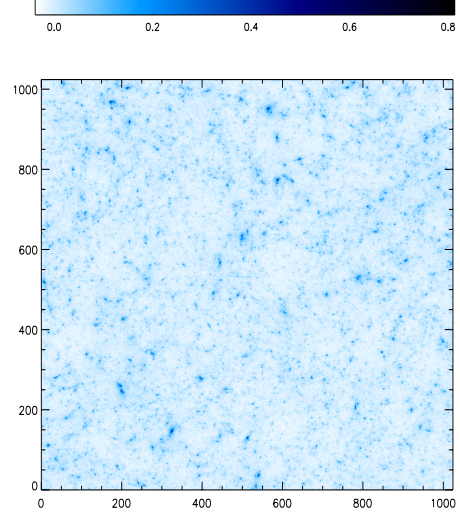


Figure 2. Simulated convergence map corresponding to a realization of a cosmological model with parameters: $\Omega_m = 0.23$, $\Omega_L = 0.77$, $h = 0.594$, $\sigma_8 = 1$. The field is $3.95^\circ \times 3.95^\circ$ downsampled to 1024×1024 (pixel scale = $0.23'$).

3.1 Shear statistics

The shear field γ can be directly derived from measurements of the shape of galaxies. For this reason, two-point statistics of the shear field have become the standard way of constraining cosmological parameters (see for example Maoli et al. 2001; Hoekstra et al. 2006; Benjamin et al. 2007; Fu et al. 2008). Most of the interest in this type of analysis comes from its potential to constrain the spectrum of density fluctuations present in the late Universe and thus the cosmological parameters. However, as said previously, cosmological parameters cannot be determined accurately using only second-order statistics because only the Gaussian features of the field are captured by this method. Therefore, higher-order statistics have been introduced to probe the non-Gaussian features of the field and thus break degeneracies. However, although the two-point correlations of the spin-2 shear field γ_i can be reduced to a scalar quantity for parity reasons, this is not the case for higher-order moments of the shear field (see Schneider and Lombardi 2003; Takada and Jain 2003; Zaldarriaga and Scoccimarro 2003). A way to get round this problem is to estimate higher-order statistics of the aperture mass M_{ap} , which has been introduced by Schneider (1996), rather than using the shear field directly.

The **aperture mass** M_{ap} is one of the most widely used techniques for probing non-Gaussianity from the shear field (e.g. Schneider et al. 1998; Jarvis et al. 2004; Kilbinger and Schneider 2005; Dietrich and Hartlap 2010).

The aperture mass M_{ap} can be expressed in terms of the tangential component of the shear γ_t :

$$M_{ap}(\theta) = \int d^2\vartheta Q_\theta(\vartheta) \gamma_t(\vartheta), \quad (4)$$

where $Q_\theta(\vartheta)$ is a radially symmetric, finite and continuous weight function and ϑ is measured from the center of the aperture. The choice of the weight function $Q_\theta(\vartheta)$ is arbitrary at this point. In this paper, we have used the form

introduced by Crittenden et al. 2002, that has been found to be more sensitive for constraining Ω_m than other forms (Zhang et al, 2003) :

$$Q_\theta(\vartheta) = \frac{\vartheta^2}{4\pi\theta^4} \exp\left(-\frac{\vartheta^2}{2\theta^2}\right). \quad (5)$$

In contrast with the shear field, which is a spin-2 field from which higher-order moments are not trivial to define, the aperture mass defined by equation [4] is a scalar (spin-0) field, the skewness and kurtosis of which are well defined.

In this section, for each noisy shear maps γ described in §2, we have estimated the aperture mass $M_{ap}(\theta)$ for several apertures θ from the relations [4] and [5]. Figure 3 shows some of these aperture mass maps for apertures $\theta = 0.46', 0.92', 1.85', 3.70', 7.40', 11.20'$. Then, the following non-Gaussian statistics have been estimated:

(i) The skewness of the aperture mass map (M_{ap}^3) is the third-order moment of the aperture mass $M_{ap}(\theta)$ and can be computed directly from shear maps filtered with different aperture mass. The skewness is a measure of the asymmetry of the probability distribution function. The probability distribution function will be more or less skewed positively depending on the abundance of dark matter haloes at the θ scale. The formalism exists to predict the skewness of the aperture mass map for a given cosmological model (see e.g. Jarvis et al. 2004; Kilbinger and Schneider 2005).

(ii) The kurtosis of the aperture mass map (M_{ap}^4) is the fourth-order moment of the aperture mass $M_{ap}(\theta)$ and can be computed directly from the different aperture mass maps. The kurtosis is a measure of the peakedness of the probability distribution function. The presence of dark matter haloes at a given θ scale will flatten the probability distribution function and widen its shoulders leading to a larger kurtosis. The formalism exists to predict the kurtosis of the aperture mass map for a given cosmological model (Jarvis et al. 2004; Kilbinger and Schneider 2005).

(iii) The peak count of the aperture mass maps $P_{M_{ap}}^\mathcal{T}$. A peak is defined as connected pixels above a detection threshold \mathcal{T} . We consider all pixels that are connected via the sides or the corners of the pixel as one structure. It means, we are not discriminating between peaks due to massive halos and peaks due to projections of large-scale structures. The formalism exists to predict the peak counts in weak-lensing surveys, including the fraction of spurious detections caused by projections effects (e.g. Maturi et al. 2010).

We have reviewed the state-of-the-art of the non-Gaussian statistics used to constrain cosmology. Interesting analytical results relative to the shear three-point correlation function or the convergence bispectrum were also reported (e.g. Ma and Fry (2000a,b); Scoccimarro and Couchman (2001); Cooray and Hu (2001)). However, when considering only the equilateral configuration of the bispectrum, it has been shown that the discrimination efficiency of the cosmological models is relatively poor (Pires et al. 2009b). An analytical comparison has been performed in Bergé et al. (2010) (for an Euclid-like survey) between the full bispec-

trum and an optimal match-filter peak count and both approaches were found to provide similar results. However, as the full bispectrum calculation has a much higher complexity than peak counting, and no public code exists to compute it (only an equilateral code is available (Pires et al. 2009a)), the bispectrum has not been considered in this study.

3.2 Convergence statistics

In this section, we have used the noisy convergence maps described in §2. The convergence has already been used in some studies (Bernardeau et al. 1997b; Hamana et al. 2004; Pires et al. 2009a; Wang et al. 2009; Bergé et al. 2010) to extract non-Gaussian information from higher-order statistics. In this paper, we want to study the ability of higher-order shear statistics compared to higher-order convergence statistics to break the σ_8 - Ω_m degeneracy. However, a fair comparison requires to compare the previous shear statistics with the convergence statistics at the same scale θ of the aperture mass. We could have used the definition of the aperture mass expressed in terms of the convergence given by:

$$M_{ap}(\theta) = \int d^2\vartheta U_\theta(\vartheta) \kappa(\vartheta), \quad (6)$$

with:

$$U_\theta(\vartheta) = \frac{1}{2\pi\theta^2} \left(1 - \frac{\vartheta^2}{2\theta^2}\right) \exp\left(-\frac{\vartheta^2}{2\theta^2}\right). \quad (7)$$

However, we have preferred to use an undecimated isotropic wavelet transform: the "à trous" wavelet transform, which computes simultaneously J aperture mass maps for dyadic scales. The "à trous" wavelet transform decomposes a convergence map κ (of size $N \times N$) as a superposition of the form:

$$\kappa(x, y) = c_J(x, y) + \sum_{j=1}^J w_j(x, y). \quad (8)$$

The algorithm outputs $J+1$ sub-band arrays of size $N \times N$ where c_J is a coarse or smooth version of the original image κ and w_j represents the details of κ at scale 2^j (see Starck et al. 1998; Starck and Murtagh 2002, for details). Leonard et al. (2011) have shown that the wavelet bands w_j are formally identical to aperture mass maps at scale $\theta = 2^j$ except the filter U_θ is replaced by the following wavelet function $\psi(x, y)$:

$$\frac{1}{4}\psi\left(\frac{x}{2}, \frac{y}{2}\right) = \varphi(x, y) - \frac{1}{4}\varphi\left(\frac{x}{2}, \frac{y}{2}\right), \quad (9)$$

where $\varphi(x, y) = \varphi(x)\varphi(y)$ and $\varphi(x)$ is a compact function (a B3-spline function) defined by:

$$\varphi(x) = \frac{1}{12}(|x-2|^3 - 4|x-1|^3 + 6|x|^3 - 4|x+1|^3 + |x+2|^3).$$

Fig. 4 displays in the Fourier domain in solid lines the aperture mass filters at scale $\theta_i = 2, 4, 8, 16, 32$ pixels (1 pixel = 0.23 arcmin) and in dashed lines, the corresponding wavelet filters at the same scale 2^j pixels with $j = 1, 2, 3, 4, 5$. To assess the response of aperture mass filters and wavelet filters, we have generated artificial shear data from a null convergence map with a single central delta function. The aperture mass algorithm has then been applied to the resulting shear maps and the wavelet transform has been computed from

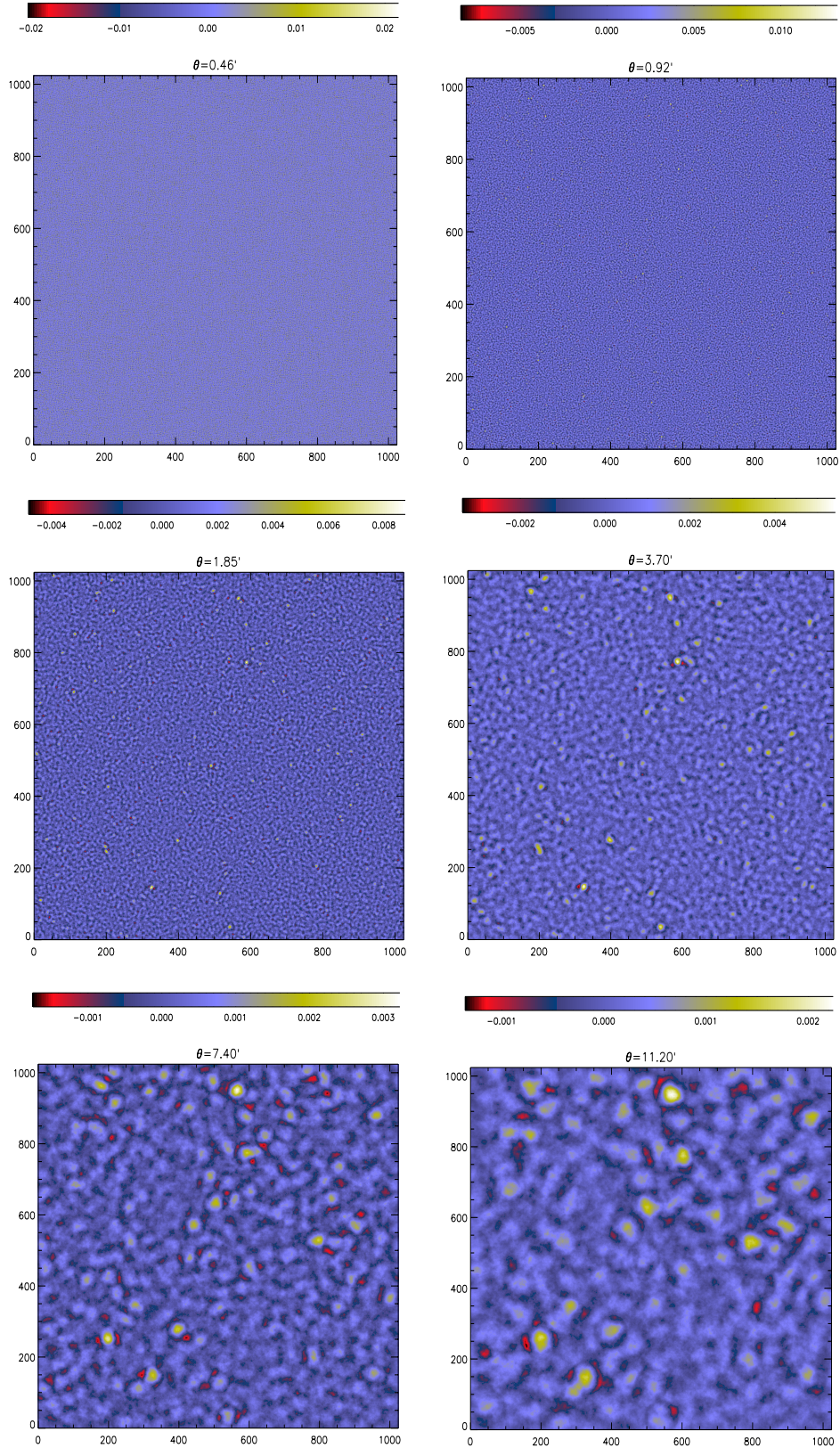


Figure 3. Aperture mass maps obtained from noisy shear maps (space-based simulations) filtered with an aperture mass with scales $\theta = 0.46', 0.92', 1.85', 3.70', 7.40', 11.20'$. The field is $3.95^\circ \times 3.95^\circ$ downsampled to 1024×1024 i.e. a pixel scale of $0.23'$.

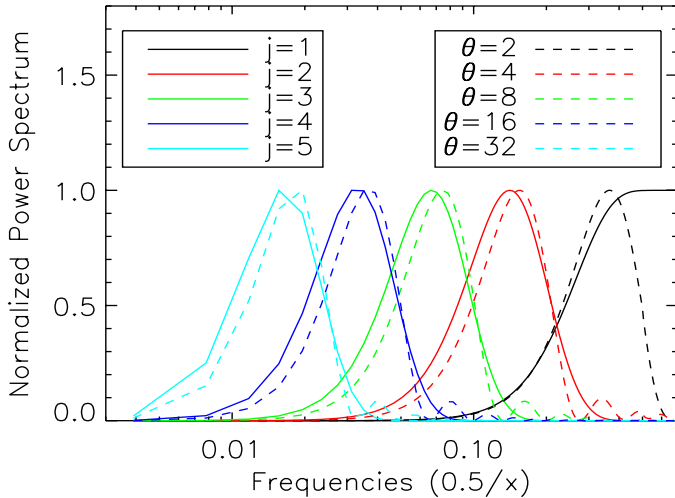


Figure 4. Frequency response of the aperture mass filters for scales $\theta_i = 0.46', 0.92', 1.85', 3.70', 7.40'$ (solid lines) and frequency response of the wavelet filters at the same scales (dashed lines).

the convergence map. Then the response of these filters in Fourier space has been obtained by considering their power spectra.

These two filter banks are really close, except the last filters (from the left). The last wavelet filter is a high-pass filter whereas the last aperture mass filter is a band-pass filter. As a consequence, we can access to the finest scales of the image using the wavelet transform method which is not the case with the aperture mass method. Nevertheless, the generalized definition of the aperture mass statistics estimated from the shear n -point correlation functions (see Schneider et al. 2005; Semboloni et al. 2011) also gives access to these small scales because it is estimated on the shear catalogue directly. However, this has not been considered in this study because it is too intensive computationally. In the same way, there exists also some wavelet transforms that can work directly on the shear catalogue (see e.g. Deriaz et al., 2012). This has not been considered in this analysis because the resolution of the maps is very good and the noise is already dominant at this scale. Considering now the other filters, we see that the aperture mass filters have some unwanted oscillations that make the wavelet filters much more localized in Fourier space. These oscillations in Fourier space are due to the fact that the aperture mass $Q_\theta(\vartheta)$ defined in [5] is truncated for $\vartheta > \theta$ (see Leonard et al. 2011, for more details). Another point in favour of the wavelet transform is its time to compute. The wavelet transform complexity is $\propto \mathcal{O}(N^2(j+1))$ compared to the aperture mass complexity, which is $\propto \mathcal{O}(N^2 \sum_i \theta_i^2)$. For one of our weak lensing simulations (of size 1024×1024), it is about 250 times faster to compute the 5 considered scales with a wavelet transform than with the aperture mass definition. The comparison between the aperture mass and the wavelet transform method has only been performed for scales : $0.92', 1.85', 3.70'$ and $7.40'$. The scales in between have not been considered in this analysis because the dyadic scales are sufficient to characterize the variation of the discrimination as a function of the scale. Nevertheless, if we want to perform a more precise

analysis, the scales in between this dyadic scaling can easily be obtained with the wavelet transform method by changing the initial pixel scale of the map. In this study, for each simulated noisy convergence map κ described in §2, we have estimated the wavelet transform from the previous definition [8] and then the following statistics have been estimated for each wavelet band w_j :

- (i) The skewness of the wavelet band $\langle w_j^3 \rangle$ that is computed directly from the different wavelet bands.
- (ii) The kurtosis of the wavelet band $\langle w_j^4 \rangle$ that is computed directly from the different wavelet bands.
- (iii) The peak count of the wavelet band $P_{w_j}^\mathcal{T}$. A peak is defined as connected pixels above a detection threshold \mathcal{T} .

3.3 Convergence statistics in denoised maps

In this section, we want to derive higher-order statistics from denoised maps because the more the data are noisy and the more the probability distribution function looks like a Gaussian, the less the higher-order statistics will be useful. We expect, for example, the skewness and the kurtosis to tend to zero with an additive Gaussian noise and the clusters to be more difficult to extract. Therefore, to extract the non-Gaussian structures and reduce the impact of the noise in the analysis, we have used the MRLens denoising proposed by Starck et al. (2006), which is a multiscale Bayesian denoising based on the sparse representation of the data. In Starck et al. (2006); Teyssier et al. (2009); Pires et al. (2009a), the authors have shown that this method outperforms several standard techniques to detect non-Gaussian structures such as Gaussian filtering, Wiener filtering and MEM filtering.

In this study, the MRLens denoising has been used to denoise each of the simulated convergence maps κ . The MRLens denoising software is available at the following address: "http://irfu.cea.fr/Ast/mrlens_software.php". We have only applied this denoising to the noisy convergence maps because of the difficulty of applying denoising to the spin-2 shear field. Then, a wavelet transform has been applied to the denoised convergence map $\tilde{\kappa}$ in order to compute the denoised wavelet bands \tilde{w}_j and estimate the following statistics:

- (i) The skewness of the denoised wavelet band $\langle \tilde{w}_j^3 \rangle$ that is computed directly from the denoised wavelet band \tilde{w}_j .
- (ii) The kurtosis of the denoised wavelet band $\langle \tilde{w}_j^4 \rangle$ that is computed directly from the denoised wavelet band \tilde{w}_j .
- (iii) The peak count of the denoised wavelet band $P_{\tilde{w}_j}^\mathcal{T}$. A peak is defined as connected pixels above a threshold ϵ . Contrary to §3.1 and §3.2 where the threshold \mathcal{T} is used to extract peaks from the noise, the threshold ϵ (set to the rms value of the denoised wavelet band \tilde{w}_j) is only used to reject spurious detections in the denoised convergence maps.

4 RESULTS

4.1 The methodology

In this study, we are interested in comparing the ability of the previous statistics to break the σ_8 - Ω_m degeneracy, using

θ_i	$\langle M_{ap}^3 \rangle$	$\langle M_{ap}^4 \rangle$	$P_{M_{ap}}^{2\sigma}$	$P_{M_{ap}}^{3\sigma}$
0.46'	04.60 %	02.30 %	39.70 %	54.95 %
0.69'	23.85 %	02.25 %	67.05 %	55.10 %
0.92'	33.40 %	03.70 %	79.30 %	76.40 %
1.40'	10.60 %	01.35 %	92.35 %	85.45 %
1.85'	03.45 %	01.45 %	91.25 %	89.20 %
2.80'	03.40 %	07.95 %	83.25 %	90.00 %
3.70'	15.15 %	23.00 %	69.40 %	86.70 %
5.60'	28.55 %	28.90 %	02.45 %	70.55 %
7.40'	26.95 %	24.30 %	4.90 %	60.50 %
11.20'	23.00 %	20.30 %	09.65 %	01.90 %

Table 1. Mean discrimination efficiency (in percent) from noisy aperture mass maps (space-based simulations) for scales $\theta_i = 0.46', 0.92', 1.40', 1.85', 2.80', 3.70', 5.60', 7.40', 11.20'$.

the set of cosmological simulations described in §2. For this purpose, we have characterized, for each statistic, its ability to discriminate between the 5 different cosmological models. As has been done in Pires et al. (2009a), we have computed a "discrimination efficiency" that expresses in percentage the ability of a statistic to discriminate between two cosmological models. For this purpose, a statistical tool called FDR (False Discovery Rate) introduced by Benjamini and Hochberg (1995) has been used to set in an adaptive way the thresholds to classify between the different cosmological models. Each threshold is estimated in such way that the rate of allowed false detections is inferior to a 0.05. The larger the discrimination efficiency is, the less the probability distributions of the statistic values for the different cosmological models overlap. The optimal statistic will be the one that maximizes the discrimination for all pairs of models. A mean discrimination efficiency can be estimated for each statistic by averaging the discrimination efficiency across all the pairs of models.

4.2 Shear statistics results

Table 1 and Table 2 show the mean discrimination efficiency for the shear statistics, described in §3.1, estimated for various aperture mass $M_{ap}(\theta_i)$ respectively for space-based and ground-based simulations.

On first glance at these two tables, we can see that the results worsen with noise whatever the statistics. As expected, the skewness and the kurtosis are very poor at discriminating between different cosmological models in noisy aperture mass maps because the aperture mass map probability distribution function tends to a Gaussian distribution as the noise increases. This makes the skewness and the kurtosis of the aperture mass map tend to zero. The result with peak counting is significantly better because a basic denoising is applied by only selecting the peaks above a given threshold in the aperture mass map. However, the choice of this threshold is relatively important because the differences between a 2σ and a 3σ threshold are significant.

θ_i	$\langle M_{ap}(\theta_i)^3 \rangle$	$\langle M_{ap}(\theta_i)^4 \rangle$	$P_{M_{ap}(\theta_i)}^{2\sigma}$	$P_{M_{ap}(\theta_i)}^{3\sigma}$
0.46'	01.95 %	02.25 %	03.45 %	06.50 %
0.69'	01.75 %	01.05 %	12.50 %	05.60 %
0.92'	02.65 %	01.30 %	31.95 %	21.90 %
1.40'	07.70 %	01.75 %	49.35 %	41.90 %
1.85'	04.05 %	02.15 %	55.85 %	48.85 %
2.80'	01.80 %	03.15 %	55.00 %	62.65 %
3.70'	04.00 %	05.20 %	54.40 %	54.45 %
5.60'	09.70 %	10.90 %	11.05 %	60.40 %
7.40'	12.00 %	10.25 %	14.15 %	51.90 %
11.20'	09.05 %	08.90 %	04.30 %	09.60 %

Table 2. Mean discrimination efficiency (in percent) from noisy aperture mass maps (ground-based simulations) for scales $\theta_i = 0.46', 0.92', 1.40', 1.85', 2.80', 3.70', 5.60', 7.40', 11.20'$.

	Space-based		Ground-based	
	Purity	Completeness	Purity	Completeness
$P_{M_{ap}}^{2\sigma}$	16.07 %	61.27 %	10.77 %	35.60 %
$P_{M_{ap}}^{3\sigma}$	47.63 %	35.11 %	33.15 %	11.56 %

Table 3. Purity and completeness for the peak count for space-based (left) and ground based (right) aperture mass maps corresponding to realizations of the cosmological model with $\Omega_m = 0.3$ and $\sigma_8 = 0.9$. $P_{M_{ap}}^{2\sigma}$ (respectively $P_{M_{ap}}^{3\sigma}$) is defined for peaks above a 2σ -threshold (respectively a 3σ -threshold) on noisy aperture mass maps for scale $\theta = 1.85'$.

Table 3 shows the purity and the completeness for the peak count estimated on noisy aperture mass maps ($\theta = 1.85'$) for space-based (left) and ground based (right). Completeness and purity are two important criteria to evaluate the performance of a peak detection method. Purity is defined as the ratio of true detections to the total number of peaks detected, and completeness is defined as the ratio of true detections to the total number of peaks in the simulation. The total number of peaks per scale θ is estimated from the simulated 2D convergence maps (without noise) with different aperture masses. It means, we are not discriminating between peaks due to massive halos and peaks due to projections of large scale-structures. With a 2σ -threshold, the completeness is maximal but the purity is poor because there is a large number of false detections due to shot noise, among the total number of detected peaks. Thus, the number of detected peaks is considerably overestimated especially at small scales for which the noise is important. Thus, the choice of the best threshold is a trade-off between purity and completeness.

Another important parameter is the scale θ of the aperture mass. The discrimination efficiency depends strongly on the scale that is considered. The best discrimination efficiency scale is displayed in bold, for each statistic, in Table

Scale	$\langle w_j^3 \rangle$	$\langle w_j^4 \rangle$	$P_{w_j}^{2\sigma}$	$P_{w_j}^{3\sigma}$
Finest scales	02.00 %	01.15 %	12.05 %	00.70 %
0.92'	37.95 %	04.75 %	86.30 %	73.05 %
1.85'	03.55 %	02.10 %	94.40 %	93.85 %
3.70'	18.25 %	25.65 %	84.05 %	87.05 %
7.40'	36.40 %	30.90 %	24.60 %	66.35 %

Table 4. Mean discrimination efficiency (in percent) from noisy convergence wavelet maps (space-based simulations) for the finest scales of the image given by the high-pass filter (first column) and for scales 0.92', 1.85', 3.70', 7.40' (other columns).

Scale	$\langle w_j^3 \rangle$	$\langle w_j^4 \rangle$	$P_{w_j}^{2\sigma}$	$P_{w_j}^{3\sigma}$
Finest scales	00.60 %	00.65 %	01.20 %	00.65 %
0.92'	07.50 %	01.60 %	38.15 %	13.45 %
1.85'	06.70 %	02.25 %	62.75 %	58.85 %
3.70'	04.15 %	06.45 %	62.25 %	52.40 %
7.40'	14.55 %	17.65 %	44.20 %	40.90 %

Table 5. Mean discrimination efficiency (in percent) from noisy convergence wavelet maps (ground-based simulations) for the finest scales of the image given by the high-pass filter (first column) and for scales 0.92', 1.85', 3.70', 7.40' (other columns).

1 and Table 2. We see that the best scale depends on the statistic that is used. This difference can be explained by the fact that the statistics are not sensitive in the same way to the different characteristics of the clusters. Skewness and kurtosis are very sensitive to the density of the clusters, e.g. very dense clusters will skew significantly the probability distribution function whereas small clusters will have a small impact. In contrast, the peak count is mainly sensitive to the number of clusters regardless of their masses or their density. A massive cluster will be accounted in the same way as a small cluster if it is detected.

The best discrimination efficiency in noisy aperture mass maps (for space-based simulations) has been obtained with the peak count with a 2σ threshold (92.35 %) for a scale of 1.40'. This is definitely better than the best result obtained with the skewness (33.40 %) as well as the best result obtained with the kurtosis (28.90 %).

4.3 Convergence statistics results

Table 4 and Table 5 show the mean discrimination efficiency for the convergence statistics, described in §3.2, estimated for different wavelet scales (2^j) respectively for space-based and ground-based simulations. As previously, the skewness and the kurtosis are very poor at discriminating between different cosmological models in noisy convergence maps because the skewness and the kurtosis tend to zero as noise is increased (see Fig. 5).

A comparison with Table 1 and Table 2 shows that the

	Space-based		Ground-based	
	Purity	Completeness	Purity	Completeness
$P_{w_3}^{2\sigma}$	22.24 %	63.20 %	14.85 %	36.31 %
$P_{w_3}^{3\sigma}$	56.66 %	38.76 %	42.02 %	13.11 %
$P_{w_3}^{mrlens}$	84.30 %	49.55 %	75.37 %	25.92 %

Table 6. Purity and completeness for the peak count for space-based (left) and ground-based (right) convergence maps corresponding to realizations of the cosmological model with $\Omega_m = 0.3$ and $\sigma_8 = 0.9$. $P_{w_3}^{2\sigma}$ (respectively $P_{w_3}^{3\sigma}$) is defined for peaks above a 2σ -threshold (respectively a 3σ -threshold) on noisy convergence maps at the third scale of a wavelet transform (1.85') and $P_{w_3}^{mrlens}$ is defined for peaks above a ϵ -threshold on MRLens denoised convergence maps at the third scale of a wavelet transform (1.85').

results obtained with the aperture mass maps at the same scale are very similar. This is not a surprising result because it has been shown by Leonard et al. (2011) that applying aperture mass filters at dyadic scales in shear maps is comparable to performing a wavelet transform of the convergence map that can be derived from the shear maps by inversion in Fourier space. This also explains the similarity of the wavelet filters compared to the aperture mass filters at the same scales in Fig. 4. However, the results are slightly improved with the wavelet transform for every statistic and every scale, which tends to show that the shape of the wavelet filters is more efficient to capture the non-Gaussian structures present in the weak lensing maps. This is possibly a consequence of the oscillations seen in the aperture mass filters (see Fig. 4), which gives rise to a small leakage of the signal into higher frequencies.

Some other studies have been conducted to find an optimal filter for detecting dark matter haloes (Maturi et al. 2005; Pace et al. 2007) and thus avoid the spurious peaks due to large-scale structure projections. However, these filters are less efficient because the projection effects that are normally a main source of uncertainty when probing the clusters, here serve as an additional source of cosmological information (see Dietrich and Hartlap 2010; Wang et al. 2009, for more details).

The best discrimination efficiency in noisy convergence maps (for space-based simulations) has been obtained with the peak count (94.40 %), for a scale of 1.85' and a 2σ -threshold.

Table 6 shows the purity and the completeness for the peak count at the third scale of a wavelet transform (1.85') for space-based (left) and ground-based (right) convergence maps. A comparison with Table 3 shows that both purity and completeness are improved with the wavelet transform. As with the aperture mass statistic, the completeness is maximal with a 2σ -threshold.

As previously, the constraints on cosmological models obtained with peak count (94.40 %) are significantly better than the ones that can be reached with the skewness (37.95 %) and the kurtosis (30.90 %).

Scale	$\langle \bar{w}_j^3 \rangle$	$\langle \bar{w}_j^4 \rangle$	$P_{\bar{w}_j}$
Finest scales	53.40 %	43.20 %	68.35 %
0.92'	47.90 %	41.15 %	92.45 %
1.85'	58.80 %	44.70 %	96.75 %
3.70'	63.30 %	48.05 %	90.40 %
7.40'	54.90 %	40.45 %	63.45 %

Table 7. Mean discrimination efficiency (in percent) from MRLens denoised convergence wavelet maps (space-based simulations) for the finest scales of the image given by the high-pass filter (first column) and for scales 0.92', 1.85', 3.70', 7.40' (other columns).

Scale	$\langle \bar{w}_j^3 \rangle$	$\langle \bar{w}_j^4 \rangle$	$P_{\bar{w}_j}$
Finest scales	42.15 %	30.05 %	38.20 %
0.92'	35.95 %	28.60 %	40.45 %
1.85'	31.65 %	20.85 %	62.35 %
3.70'	41.80 %	29.95 %	72.65 %
7.40'	44.75 %	32.25 %	54.55 %

Table 8. Mean discrimination efficiency (in percent) from MRLens denoised convergence wavelet maps (ground-based simulations) for the finest scales of the image given by the high-pass filter (first column) and for scales 0.92', 1.85', 3.70', 7.40' (other columns).

4.4 Denoised Convergence statistics results

In this section, we want to show that the convergence statistics can be improved significantly if denoising is applied to the convergence maps. As said previously, the convergence maps have been denoised using the MRLens denoising described in Starck et al. (2006). Table 7 and Table 8 show the mean discrimination efficiency for the denoised convergence statistics, described in §3.3, estimated for different wavelet scales (2^j) respectively for space-based and ground-based simulations.

As expected, the MRLens denoising improves considerably the discrimination efficiency of the skewness and kurtosis. This comes from its ability to reconstruct the non-Gaussian structures that dominate at small scales. However, the skewness and kurtosis values are significantly overestimated compared to original kurtosis, as shown in Fig. 5, because the MRLens denoising is only efficient in recovering high peaks in the signal, which affects the tails of the probability distribution function.

The MRLens denoising also improves the discrimination efficiency of the peak count at all scales, especially for ground-based simulations for which the noise is important.

Table 6 shows the purity and the completeness for peak counting on MRLens denoised maps for space-based and ground-based simulations. In the MRLens denoising, we again have the usual trade-off between purity and completeness. A different threshold is selected for each wavelet band,

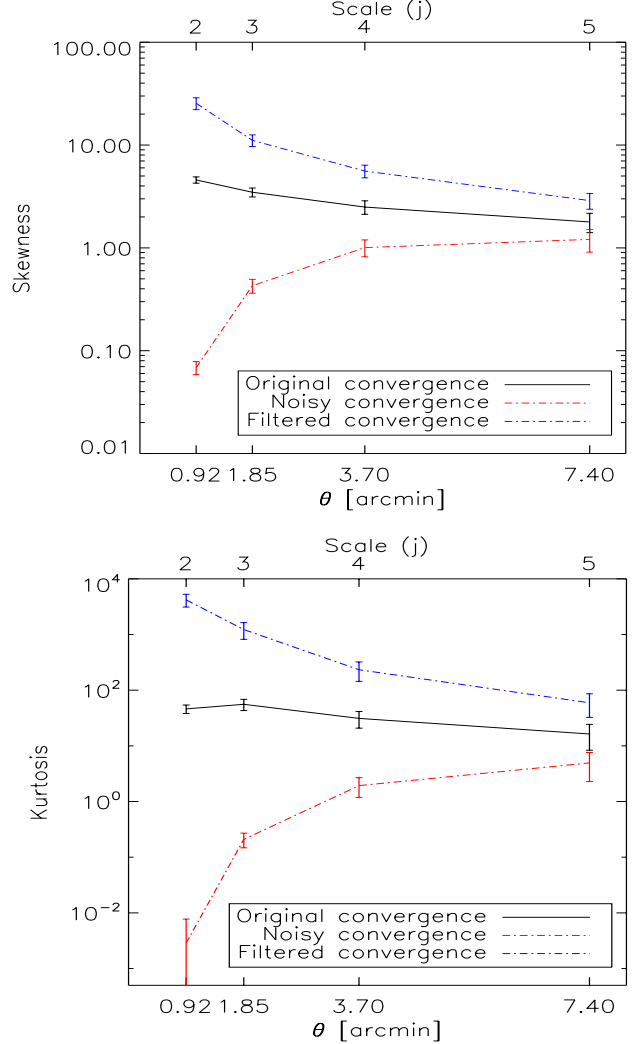


Figure 5. Top: Mean skewness per scale for original convergence maps (black), space-based noisy convergence maps (red) and MRLens denoised convergence maps (blue). Bottom: Mean kurtosis per scale for original convergence maps (black), space-based noisy convergence maps (red) and MRLens denoised convergence maps (blue). The skewness and the kurtosis of the noisy convergence maps are considerably reduced especially at small scales for which the noise is important. In contrast, the skewness and the kurtosis are significantly overestimated on MRLens denoised convergence maps. These convergence maps correspond to realizations of the cosmological model with $\Omega_m = 0.3$ and $\sigma_8 = 0.9$.

and this is done in an adaptive way, conformed to a False Discovery Rate method (Starck et al. 2006), which provides a more robust discrimination. The completeness with MRLens denoising is slightly inferior to a 2σ -threshold but its purity is maximal.

The best discrimination efficiency in denoised convergence maps (for space-based simulations) has been obtained with the peak count (96.75 %) still for a scale of 1.85', in perfect agreement with the results of Pires et al. (2009a). Table 9 shows the discrimination efficiency obtained with this statistic that enables to discriminate between the five cosmological models even for contiguous models for which the discrimination is challenging. The Table is not symmet-

	model 1	model 2	model 3	model 4	model 5
model 1	x	85 %	100 %	100 %	100 %
model 2	89 %	x	92 %	100 %	100 %
model 3	100 %	92 %	x	89 %	100 %
model 4	100 %	100 %	92 %	x	98 %
model 5	100 %	100 %	100 %	98 %	x

Table 9. Discrimination efficiency (in percent) between the 5 cosmological models obtained with the peak count on denoised convergence maps at the third scale (1.85') of a wavelet transform (space-based simulations).

ric because the probability distributions of the statistics are not symmetric and to quantify the discrimination between two cosmological models, the FDR method sets two different thresholds in an adaptive way. This result can be compared with the result obtained in Pires et al. (2009a) (Table 7) with another set of cosmological simulations. The results are very similar. The small discrepancies are only due to the limited size of the cosmological simulation sample.

5 CONCLUSION

The goal of this paper was to investigate how to best extract non-Gaussianity from weak lensing surveys to constrain the cosmological model. For this purpose, we have been interested in showing that there is an extra information that can be derived from higher-order statistical analysis of the convergence maps beyond what can be obtained from higher-order statistical analysis of the shear maps. Therefore, we have compared the efficiency of several higher-order shear and convergence statistics to break the $\sigma_8 - \Omega_m$ degeneracy, by comparing their ability to discriminate between 5 cosmological models along this degeneracy.

Most of the techniques used to estimate higher-order statistics from the spin-2 shear field are based on the aperture mass expressed in terms of the tangential component of the shear (see relation [4]). Analogous convergence statistics can be obtained by using the aperture mass defined from the convergence maps (see relation [6]). However, in accordance with Leonard et al. (2011), we have preferred to use an alternative solution that computes simultaneously multiple aperture mass maps for dyadic scales: the "à trous" wavelet transform. In this study, we have observed that this method is 250 times faster than the aperture mass method to run on our simulations and that the wavelet filters are much more localized in Fourier space compared to aperture mass filters. It follows that the results obtained in wavelet convergence maps compared to the results in aperture mass maps are very similar but slightly improved in the wavelet case for every statistic and every scale. Therefore, contrary to a generally accepted idea, the noise properties in aperture mass maps are not better than in convergence maps, if the same scale is considered.

Contrary to another accepted idea, further important cosmological information can be extracted from noisy convergence maps if a denoising such as MRLens is used. This comes from its ability to reconstruct the non-Gaussian structures that are induced by the growth of structures. In this

study, we have shown that the MRLens denoising improves considerably the discrimination efficiency of the skewness and kurtosis. It also improves the discrimination efficiency of the peak count especially for ground-based simulations ($n_g = 30$ gal/arcmin²) for which the noise is important.

For an Euclid-like survey, the density of galaxies is expected to be around $n_g = 40$ gal/arcmin² for the wide-field survey and around $n_g = 80$ gal/arcmin² for the deep-field survey. It is clear from this study, that the non-Gaussian statistical analysis should be performed in denoised convergence maps as described in §3.3.

Finally, the best non-Gaussian statistic to constrain cosmological model in combination with the power spectrum has been found to be the peak count per scale. And further cosmological information should be obtained by combining the constraints obtained with the peak count at different scale as shown by Marian et al. (2011). This will be investigated in a future work.

ACKNOWLEDGMENTS

This work has been supported by the European Research Council grant SparseAstro (ERC-228261).

REFERENCES

- Bacon, D., Refregier, A., and Ellis, R.: 2000, *MNRAS* pp 318–625
- Bacon, D. J., Massey, R. J., Refregier, A. R., and Ellis, R. S.: 2003, *MNRAS* **344**, 673
- Bartelmann, M. and Schneider, P.: 2001, *Phys. Rep.* **340**, 291
- Benjamin, J., Heymans, C., Semboloni, E., van Waerbeke, L., Hoekstra, H., Erben, T., Gladders, M. D., Hettterscheidt, M., Mellier, Y., and Yee, H. K. C.: 2007, *MNRAS* **381**, 702
- Benjamini, Y. and Hochberg, Y.: 1995, *J. R. Stat. Soc. B* **57**, 289
- Bergé, J., Amara, A., and Réfrégier, A.: 2010, *ApJ* **712**, 992
- Bernardeau, F., van Waerbeke, L., and Mellier, Y.: 1997a, *A&A* **322**, 1
- Bernardeau, F., van Waerbeke, L., and Mellier, Y.: 1997b, *A&A* **322**, 1
- Cooray, A. and Hu, W.: 2001, *ApJ* **548**, 7
- Dahle, H.: 2006, *ApJ* **653**, 954

- Dietrich, J. P. and Hartlap, J.: 2010, *MNRAS* **402**, 1049
- Fu, L., Semboloni, E., Hoekstra, H., Kilbinger, M., van Waerbeke, L., Tereno, I., Mellier, Y., Heymans, C., Coupon, J., Benabed, K., Benjamin, J., Bertin, E., Doré, O., Hudson, M. J., Ilbert, O., Maoli, R., Marmo, C., McCracken, H. J., and Ménard, B.: 2008, *A&A* **479**, 9
- Hamana, T., Takada, M., and Yoshida, N.: 2004, *MNRAS* **350**, 893
- Hoekstra, H. and Jain, B.: 2008, *Annual Review of Nuclear and Particle Science* **58**, 99
- Hoekstra, H., Mellier, Y., van Waerbeke, L., Semboloni, E., Fu, L., Hudson, M. J., Parker, L. C., Tereno, I., and Benabed, K.: 2006, *ApJ* **647**, 116
- Jarvis, M., Bernstein, G., and Jain, B.: 2004, *MNRAS* **352**, 338
- Kaiser, N. and Squires, G.: 1993, *ApJ* **404**, 441
- Kaiser, N., Wilson, G., and Luppino, G.: 2000, *AJ* pp 318–625
- Kilbinger, M. and Schneider, P.: 2005, *A&A* **442**, 69
- Leonard, A., Pires, S., and Starck, J.-L.: 2011, *MNRAS*
- Ma, C.-P. and Fry, J. N.: 2000a, *ApJ* **543**, 503
- Ma, C.-P. and Fry, J. N.: 2000b, *ApJ* **538**, L107
- Maoli, R., Van Waerbeke, L., Mellier, Y., Schneider, P., Jain, B., Bernardeau, F., Erben, T., and Fort, B.: 2001, *A&A* **368**, 766
- Marian, L., Smith, R. E., Hilbert, S., and Schneider, P.: 2011, *ArXiv e-prints*
- Massey, R., Refregier, A., Bacon, D. J., Ellis, R., and Brown, M. L.: 2005, *MNRAS* **359**, 1277
- Maturi, M., Angrick, C., Pace, F., and Bartelmann, M.: 2010, *A&A* **519**, A23
- Maturi, M., Meneghetti, M., Bartelmann, M., Dolag, K., and Moscardini, L.: 2005, *A&A* **442**, 851
- Mellier, Y.: 1999, *ARA&A* **37**, 127
- Munshi, D., Valageas, P., van Waerbeke, L., and Heavens, A.: 2008, *Phys. Rep.* **462**, 67
- Pace, F., Maturi, M., Meneghetti, M., Bartelmann, M., Moscardini, L., and Dolag, K.: 2007, *A&A* **471**, 731
- Pires, S., Starck, J.-L., Amara, A., Réfrégier, A., and Teyssier, R.: 2009a, *A&A* **505**, 969
- Pires, S., Starck, J.-L., Amara, A., Teyssier, R., Réfrégier, A., and Fadili, J.: 2009b, *MNRAS* **395**, 1265
- Refregier, A.: 2003, *ARA&A* **41**, 645
- Refregier, A., Rhodes, J., and Groth, E. J.: 2002, *ApJ* **572**, L131
- Schneider, P.: 1996, *MNRAS* **283**, 837
- Schneider, P., Kilbinger, M., and Lombardi, M.: 2005, *A&A* **431**, 9
- Schneider, P. and Lombardi, M.: 2003, *A&A* **397**, 809
- Schneider, P., van Waerbeke, L., Jain, B., and Kruse, G.: 1998, *MNRAS* **296**, 873
- Scoccimarro, R. and Couchman, H. M. P.: 2001, *MNRAS* **325**, 1312
- Semboloni, E., Schrabback, T., van Waerbeke, L., Vafaei, S., Hartlap, J., and Hilbert, S.: 2011, *MNRAS* **410**, 143
- Starck, J.-L. and Murtagh, F.: 2002, *Astronomical Image and Data Analysis*, Springer-Verlag
- Starck, J.-L., Murtagh, F., and Bijaoui, A.: 1998, *Image Processing and Data Analysis: The Multiscale Approach*, Cambridge University Press
- Starck, J.-L., Pires, S., and Réfrégier, A.: 2006, *A&A* **451**, 1139
- Takada, M. and Jain, B.: 2003, *MNRAS* **344**, 857
- Takada, M. and Jain, B.: 2004, *MNRAS* **348**, 897
- Teyssier, R.: 2002, *A&A* **385**, 337
- Teyssier, R., Pires, S., Prunet, S., Aubert, D., Pichon, C., Amara, A., Benabed, K., Colombi, S., Refregier, A., and Starck, J.-L.: 2009, *A&A* **497**, 335
- Vale, C. and White, M.: 2003, *ApJ* **592**, 699
- Van Waerbeke, L., Mellier, Y., Erben, T., Cuillandre, J., and Bernardeau, F. e. a.: 2000, *A&A* p. 318:30
- Wang, S., Haiman, Z., and May, M.: 2009, *ApJ* **691**, 547
- Wittman, D., Tyson, J., Kirkman, D. and DellAntonio, I., and Bernstein, G.: 2000, *Nature* p. 405:143
- Zaldarriaga, M. and Scoccimarro, R.: 2003, *ApJ* **584**, 559