Astron. Astrophys. 338, 375–382 (1998)

ASTRONOMY AND ASTROPHYSICS

The effects of source clustering on weak lensing statistics

F. Bernardeau

Service de Physique Théorique, C.E. de Saclay, F-91191 Gif-sur-Yvette cedex, France

Received 30 October 1997 / Accepted 15 May 1998

Abstract. I investigate the effects of source clustering on the weak lensing statistics, more particularly on the statistical properties of the local convergence, κ, at large angular scales. The Perturbation Theory approach shows that the variance is not affected by source clustering at leading order but higher order moments such as the third and fourth moments can be.

I compute the magnitude of these effects in case of an Einstein-de Sitter Universe for the angular top-hat filtered convergence. In these calculations the so-called Broadhurst and multiple lens coupling effects are neglected. The source clustering effect is found to be particularly important when the redshift distribution is broad enough so that remote background sources can be significantly lensed by closer concentrations of galaxy sources. The source clustering effects are shown to remain negligible, for both the skewness and the kurtosis, when the dispersion of the redshift of the sources is less than about 0.15.

Key words: cosmology: dark matter – cosmology: gravitational lensing – cosmology: large-scale structure of Universe

1. Introduction

The construction of gravitational distortion maps, for tracing the large scale structure of the Universe, is a promising tool for cosmology. For the first time, it would indeed provides us with an unbiased representation of the mass distribution in the universe (Blandford et al. 1991, Miralda-Escudé 1991, Kaiser 1992). And recent works tend to prove that indeed it should be possible to have reliable distortion maps at the level expected for the distortion induced by the large-scale structures (Schneider et al. 1998a).

The cosmological interpretation of such maps is however challenging. The main difficulty is that the background sources used to make such measurements are very faint galaxies, whose distances and distribution is not necessarily well known. It has been pointed out in recent papers (Villumsen 1997, Bernardeau et al. 1997, Jain & Seljak 1997) that it was crucial to know with a good accuracy the redshift distribution of those sources. Their distances determine indeed the magnitude of the distortion effect: the more distant they are the larger the effect is. The r.m.s. of the distortion is however not the only information of cosmological interest to be extracted from distortion maps. In particular the departure of the gravitational convergence from a Gaussian statistics in case of Gaussian initial conditions is an indicator of the amount of nonlinearity reached by the cosmic density field (Bernardeau et al. 1997, Jain & Seljak 1997, Schneider et al. 1998b). To get a reliable description of such detailed analysis, one has however to take into account the possible effects produced by the intrinsic statistical properties of the sources.

In this paper I investigate the effect of source clustering for high-order moment of the convergence. In Sect. 2, the observational schemes for the distortion and convergence fields are recalled, and the couplings between the lens density fluctuations and the background galaxy fluctuations are explicitly written. The calculations are made for an Einstein-de Sitter Universe only. In Sect. 3, the implications of such couplings for the skewness, third moment of the convergence are investigated. In Sect. 4, the expression of the kurtosis, fourth moment, is computed, taking into account the source clustering properties. The numerical applications, in Sect. 5, are made for different models of source distributions.

Send offprint requests to: F. Bernardeau, fbernardeau@cea.fr

2. Observational schemes for the distortion field

2.1. The filtering schemes

In the weak lensing regime, and for small-size background objects, the local gravitational distortion effects can be entirely described by the local deformation matrix $\mathcal{A}$. This matrix expresses the local linear transform between the source and the image plane induced by all the lenses present along a line-of-sight. Its inverse can be related to the second order derivative of the gravitational potential through the equation,

$$\mathcal{A}^{-1}(\gamma) = \mathrm{Id} - \left\{ \frac{3}{2} \int_0^{\mathcal{D}_s} \mathrm{d}\mathcal{D}\, \frac{(\mathcal{D}_s - \mathcal{D})\,\mathcal{D}}{\mathcal{D}_s}\, \frac{\phi_{,ij}(\mathcal{D},\gamma)}{a(\mathcal{D})} \right\} \quad (1)$$

where Id is the identity matrix, $\mathcal{D}$ is the comoving angular distance along the line-of-sight, $a(\mathcal{D})$ is the expansion factor and $\phi_{,ij}(\mathcal{D},\gamma)$ are second order derivatives of the local gravitational

potential at the position $(\mathcal{D},\gamma)$ along the directions i and j orthogonal to the line-of-sight. Expression (1) is valid for an Einstein-de Sitter only, but can be easily extended to any background geometry. It is written for a given distance of the source, $\mathcal{D}_s$, that may vary for different line-of-sights.

In practice, the deformation matrix is not directly measurable. The quantities that are directly accessible are the shape parameters of the observed galaxies, $\mathcal{S}^I$, in the image plane. For objects with a large enough extension (compared to the width of the point spread function), this matrix can be related to the shape matrix in the source plane, $\mathcal{S}^S$, through,

$$\mathcal{S}^I = \mathcal{A}^{-1}\cdot\mathcal{S}^S\cdot\mathcal{A}^{-1}\,\frac{1}{\det(\mathcal{A}^{-1})}. \tag{2}$$

Then the determination of the direction and amplitude of the ellipticities of the galaxies gives a local estimation of the deformation matrix[1]. Because of the intrinsic ellipticities of the sources, the cosmic distortion can be detected only when a large number of galaxies are taken into account. Averaging over few hundred of galaxies, distortion signals down to a few percents are in principle detectable (Blandford et al. 1991).

The particular quantity that can thus be reconstructed is the local filtered convergence, κ. It is directly related to the trace of $\mathcal{A}^{-1}$ with,

$$\kappa = 1 - \mathrm{tr}\left(\mathcal{A}^{-1}\right)/2. \tag{3}$$

This quantity can be obtained directly from the observed galaxy shapes, when it is filtered with a compensated filter (i.e. convolved with a function of zero integral, see Kaiser 1995, Schneider et al. 1998b). In general, however, it is always possible to obtain a convergence map from a distortion map by solving a differential equation (Kaiser 1995). In the following I will therefore focus my analysis on the statistical properties of the local convergence, filtered by a top-hat window funtion.

In the expression (1) the relation between the deformation matrix and the gravitational potential is given for a unique distance of the source plane. But actually the measured convergences result from averages made over many background galaxies that can be at different distances. More specifically the measured local convergence at scale θ_0 is obtained from background objects taken for instance in a solid angle of radius θ_0, so that it reads,

$$\kappa_{\theta_0} = \frac{1}{N_s}\sum_{i=0}^{N_s}\kappa(\gamma_i), \tag{4}$$

where γ_i is the direction of the $i^{\rm th.}$ source galaxy. The number density of source galaxies that can be used is about 40 per arcmin^2 (with the usual deep exposures in the I band). For a filtering radius of $20'$, we expect then to have about $50\,000$ galaxies. This number is large enough to assume that the discretization of the background field is not important. We can then work in the coutinuous limit in the source plane. The number density of sources at distance D_s and in the direction γ can be written,

$$n_s(\mathcal{D}_s,\gamma) = n_s(\mathcal{D}_s)\,(1+\delta_s(\mathcal{D}_s,\gamma)), \tag{5}$$

where $n_s(\mathcal{D}_s)$ is the average number density of sources[2](that fulfill the selection criteria at a given distance), and $\delta_s(\mathcal{D}_s,\gamma)$ is their local density contrast. The density n_s is normalized to unity, $\int_0^2 \mathrm{d}\mathcal{D}_s\, n_s(\mathcal{D}_s)=1$. Writting the Eq. (4) in the continuous limit we have,

$$\begin{aligned}\kappa_{\theta_0} \;=\; & -\frac{3}{2}\int \mathrm{d}^2\gamma\, W_{\theta_0}(\gamma)\int_0^2 \mathrm{d}\mathcal{D}_s\int_0^{\mathcal{D}_s}\mathrm{d}\mathcal{D}\,\frac{\mathcal{D}\,(\mathcal{D}_s-\mathcal{D})}{a\,\mathcal{D}_s}\\ & \times\delta_{\rm mass.}(\mathcal{D},\gamma)\,n_s(\mathcal{D}_s)\,(1+\delta_s(\mathcal{D}_s,\gamma))\\ & \Big/\int\mathrm{d}^2\gamma\, W_{\theta_0}(\gamma)\int_0^2\mathrm{d}\mathcal{D}_s\, n_s(\mathcal{D}_s)\,[1+\delta_s(\mathcal{D}_s,\gamma)].\end{aligned} \tag{6}$$

In this expression, the two filters are a priori both top-hat filters,

$$\begin{aligned} &W_{\theta_0}(\gamma) = 1 \quad \text{for}\ \ |\gamma|\le\theta_0;\\ &W_{\theta_0}(\gamma) = 0 \quad \text{for}\ \ |\gamma|>\theta_0.\end{aligned} \tag{7}$$

Note however that in general the filters are not necessary the same. For instance one can use a compensated filter for the convergence, and still use a top-hat window for the selection of the sources. More generally, instead of giving an equal weight to all sources in a given area, it is always possible to weight them proportionally to the inverse of the local density. We would then have,

$$\kappa^{\rm surf.}_{\theta_0} = \frac{\sum_{i=0}^{N_s} w_i\,\kappa(\gamma_i)}{\sum_{i=0}^{N_s} w_i}, \tag{8}$$

where w_i is a weight associated with each background object. We encounter here a situation similar to the cosmic velocity field statistics (see discussion in Bernardeau & van de Weygaert 1996). To get a weight inversely proportional to the local density, there are a priori different possible schemes. One can for instance use a two step filtering, with a usual filtering scheme on a small grid, and then a final filtering on the larger scale, corresponding to the actual smoothing scale of the statistical analysis. One could also think of using similar approaches as the Voronoi and Delaunay methods developed by Bernardeau & van de Weygaert. In such a case the weigth w_i would be for instance the inverse of the surface of the Voronoi cell occupied by a given galaxy. It would lead to a proper 'surface-average' filtering scheme,

$$\begin{aligned}\kappa^{\rm surf.}_{\theta_0} \;=\; & -\frac{3}{2}\int\mathrm{d}^2\gamma\, W_{\theta_0}(\gamma)\\ & \times\frac{\int_0^2\mathrm{d}\mathcal{D}_s\int_0^{\mathcal{D}_s}\mathrm{d}\mathcal{D}\,\frac{\mathcal{D}\,(\mathcal{D}_s-\mathcal{D})}{a\,\mathcal{D}_s}\,\delta_{\rm mass.}\,n_s(\mathcal{D}_s)\,(1+\delta_s)}{\int_0^2\mathrm{d}\mathcal{D}_s\,n_s(\mathcal{D}_s)\,[1+\delta_s]},\end{aligned} \tag{9}$$

[1] see however an alternative method proposed by van Waerbeke et al. 1997, called the Pixel Autocorrelation Function, that does not use the shape matrices.

[2] Note that for an Einstein-de Sitter Universe the angular distance at $z\to\infty$ equals 2.

where $\delta_{\rm mass.}$ is taken at position $(\mathcal{D},\gamma)$ and δ_s at position $(\mathcal{D}_s,\gamma)$. Unlike in the velocity case, the dependence on the local tracer density contrast does not vanish, because of the finite width of the redshift distribution of the sources. In all these formulae both the density contrast of the cosmic density field and of the sources are present. They are both random fields and moreover their cross-correlations are a priori comparable to their autocorrelation properties. The aim of this study is to investigate the consequence of such couplings.

2.2. The physical effects induced by source clustering

First of all let me stress that the effects discussed here are different in nature from other couplings that have been described previously in the literature. A possible source a coupling between the lens and the source population is, for instance the amplification effect that is expected to affect the local number density of detected tracers (Broadhurst et al. 1995). This effect is due to the fact that the apparent size of background objects depends on the local gravitational amplification making them more, or less, easily detectable. This effect is always present, even if the sources are not clustered. It is also expected to vanish when the sources are selected on surface brightness criteria only. Here I assume that the sources are, in this sense, perfectly selected, so that the amplification effect is entirely negligible.

What are then the effects of source clustering? It comes from the fact that the source 'plane' is actually thick and inhomogeneous. One mechanism originates from the fact that there might be a significant overlapping between the distribution of sources and the distribution of lenses. Let me imagine now for instance there is a large potential well at a very large distance, in the overlapping area. Then because the source galaxies trace somehow the matter field, you expect to have *at the same time* a relative *excess* of close sources. The presence of these sources tends to reduce the gravitational signal of the remote lenses. One expects then that the gravitational distortions of the furthest lenses will be systematically underestimated. In Fig. 1 a sketch of the actual situation is proposed.

The implications of such a coupling depends on the level of description one wishes for the local convergence. At linear order the expression of the local convergence is not affected. Therefore the variance is not expected to be much changed at large scale. At small scales, the extra couplings with the source density fluctuations can compete with the intrinsic non-linear evolution of the projected density. It should then be addressed with a complete numerical study.

Another possible effect is due to the fact that the source plane is expected to be 'bumpy', i.e. the average distance of the sources may vary from one direction to another. For instance if one observes the gravitational distortion induced by a perfectly round potential with a 'bumpy' source plane, the efficiency of the gravitational effect is expected to vary from one direction to another creating apparent substructures in the potential well. The importance of this effect once again depends on the level of description ones whishes. It is expected in general to create more power at small scale, but it can be significant (if it is present at all) only at small angular scale. Perturbatively, this effect is expected to play a role only for the kurtosis and moments of higher order. Note that contrary to the previous mechanism, this effect is not due to the cross-correlation between the lenses and the sources, but to the intrinsic correlation properties of the sources.

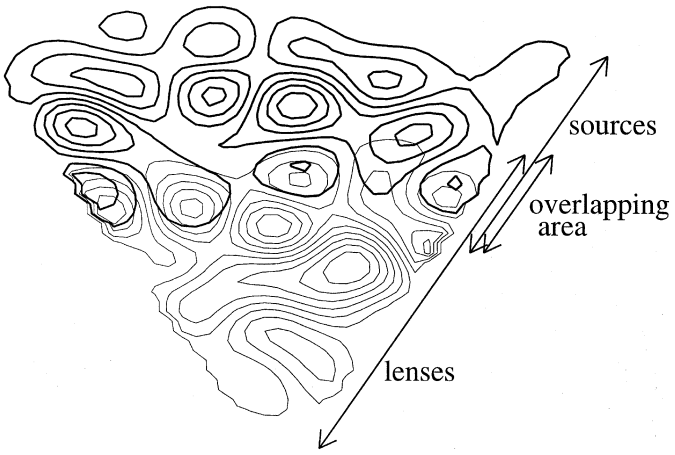

Fig. 1. Sketch of the lens (thin lines) and source (thick lines) density fluctuations. The source distribution is not expected to be smooth. Moreover the source density fluctuations are correlated to the lens density fluctuations. This is particularly important when the overlapping area is large.

2.3. Models for lens and source correlations

The aim of the coming sections is to investigate the implications of source clustering on high-order moments by means of Perturbation Theory. Calculations can be pursued only with a model for the mass-galaxy and galaxy-galaxy correlation functions. In the following I will assume that a local bias holds and that the local galaxy density contrast can be expanded in terms of the *linear* mass density contrast,

$$\delta_{\rm s}(\mathcal{D},\gamma) \;=\; b_1(\mathcal{D})\,\delta^{(1)}_{\rm mass}(\mathcal{D},\gamma) + b_2(\mathcal{D})\,\left[\delta^{(1)}_{\rm mass}(\mathcal{D},\gamma)\right]^2 + \dots \qquad (10)$$

Although there is no complete justification for such an assumption, it is a quite natural from the linear analysis of Bardeen et al. (1986) or even from the non-linear description[3] proposed by Bernardeau & Schaeffer (1992). Note that a skewness of about 3 (Bouchet et al. 1993, Kim & Strauss 1998) for the galaxies implies, $b_2 \approx 0.5\,b_1^2$. Furthermore I have to assume a given function for the evolution of the bias factors b_1 and b_2 with redshift. In the following I will assume that

$$b_1(\mathcal{D}) \propto 1/a(\mathcal{D}) \quad \text{and} \quad b_2(\mathcal{D}) \propto 1/a^2(\mathcal{D}). \qquad (11)$$

According to the results of Bardeen et al. (1986) it corresponds to objects defined with a fixed threshold in units of the variance.

[3] It implies in particular that the 3-point correlation function of the galaxy field can be exactly factorized in products of 2-point functions

Finally the mass (and galaxy) density field is fruitfully described by its Fourier transform,

$$\delta(\mathcal{D},\gamma) = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{3/2}}\,\delta(\mathcal{D},\mathbf{k})\,\exp[\mathrm{i}(\mathcal{D}\mathbf{k}_\perp\cdot\gamma + \mathcal{D}\,k_r)] \quad (12)$$

where k_r is the radial part of the wave vector $\mathbf{k}$ and $\mathbf{k}_\perp$ is its perpendicular part. In the linear regime, the Fourier components $\delta(\mathcal{D},\mathbf{k})$ grow like $a(\mathcal{D})$ for an Einstein-de Sitter Universe, and they are assumed to obey a Gaussian statistics characterized by the power spectrum $P(k)$,

$$\begin{aligned}\delta^{(1)}(\mathcal{D},\mathbf{k}) &= a(\mathcal{D})\,\delta_{\rm init.}(\mathbf{k});\\ \langle\delta_{\rm init.}(\mathbf{k})\,\delta_{\rm init.}(\mathbf{k}')\rangle &= \delta_{\rm Dirac}(\mathbf{k}+\mathbf{k}')\,P(k).\end{aligned} \quad (13)$$

As the numerical calculations will be done at a fixed smoothing scale, it is reasonable to assume that $P(k)$ follows a power law behavior,

$$P(k)\propto k^n, \quad (14)$$

and at the scales of interest we expect (see Bernardeau et al. 1997),

$$n\approx -1.5. \quad (15)$$

As mentionned before, from a Perturbation Theory point of view, the variance is not affected at leading order by the source clustering. In the following, the calculation will concentrate on the skewness and kurtosis of the local convergence.

3. Implications for the skewness

At large scale, for Gaussian initial conditions, the third moment is given by a combination of the first order and second order terms of the local convergence with respect to the initial density field (e.g. Peebles 1980, Fry 1984, Goroff et al. 1986),

$$\left\langle\kappa^3\right\rangle \approx 3\left\langle\left[\kappa^{(1)}\right]^2\kappa^{(2)}\right\rangle. \quad (16)$$

In the following, it is assumed that the effect of multiple lenses and spurious observational couplings are negligible (see Bernardeau et al. 1997).

3.1. Expressions of the first and second order convergence

The presence of source clustering does not change the expression of the first order term,

$$\begin{aligned}\kappa^{(1)}_{\theta_0} &= -\frac{3}{2}\int \mathrm{d}^2\gamma\,W_{\theta_0}(\gamma)\int_0^2\mathrm{d}\mathcal{D}_s\int_0^{\mathcal{D}_s}\mathrm{d}\mathcal{D}\\ &\quad\times\frac{\mathcal{D}\,(\mathcal{D}_s-\mathcal{D})}{\mathcal{D}_s}\,\frac{\delta^{(1)}_{\rm mass}(\mathcal{D},\gamma)}{a(\mathcal{D})}\,n_s(\mathcal{D}_s).\end{aligned} \quad (17)$$

Written in terms of $\delta(\mathbf{k})$ it reads,

$$\begin{aligned}\kappa^{(1)}_{\theta_0} &= -\frac{3}{2}\int_0^2\mathrm{d}\mathcal{D}_s\int_0^{\mathcal{D}_s}\mathrm{d}\mathcal{D}\,\frac{\mathcal{D}\,(\mathcal{D}_s-\mathcal{D})}{\mathcal{D}_s}\,n_s(\mathcal{D}_s)\\ &\quad\times\int\frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{3/2}}\delta_{\rm init.}(\mathbf{k})\,\exp[\mathrm{i}\,k_r\,\mathcal{D}]\,W(\mathbf{k}_\perp\,\mathcal{D}\theta_0),\end{aligned} \quad (18)$$

where the filter W is expressed here in Fourier space. One can further simplify this expression by introducing the efficiency function, $\omega(\mathcal{D})$, defined by

$$\omega(\mathcal{D}) = \frac{3}{2}\int_{\mathcal{D}}^2\mathrm{d}\mathcal{D}_s\,\frac{(\mathcal{D}_s-\mathcal{D})\mathcal{D}}{\mathcal{D}_s}\,n_s(\mathcal{D}_s), \quad (19)$$

so that

$$\begin{aligned}\kappa^{(1)}_{\theta_0} &= -\int_0^2\mathrm{d}\mathcal{D}\,\omega(\mathcal{D})\int\frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{3/2}}\delta_{\rm init.}(\mathbf{k})\\ &\quad\times\exp[\mathrm{i}\,k_r\,\mathcal{D}]\,W(\mathbf{k}_\perp\,\mathcal{D}\theta_0).\end{aligned} \quad (20)$$

When the source clustering is neglected, the second order term of the local convergence is given by

$$\begin{aligned}\kappa^{(2)}_{\theta_0} &= -\int_0^2\mathrm{d}\mathcal{D}\,\omega(\mathcal{D})\int\frac{\mathrm{d}^3\mathbf{k}\,\mathrm{d}^3\mathbf{k}'}{(2\pi)^3}\,\delta_{\rm init.}(\mathbf{k})\delta_{\rm init.}(\mathbf{k}')\,a(\mathcal{D})\\ &\quad\times F_2(\mathbf{k},\mathbf{k}')\exp[\mathrm{i}\,(k_r+k'_r)\,\mathcal{D}]\\ &\quad\times W\left[|\mathbf{k}_\perp+\mathbf{k}'_\perp\,|\,\mathcal{D}\theta_0\right],\end{aligned} \quad (21)$$

where F_2 is an homogeneous function of the wave vectors (e.g. Goroff et al. 1986).

When source clustering is taken into account, two extra terms for the second order should be added,

$$\begin{aligned}\kappa^{\rm s.c.(2)}_{\theta_0} &= \kappa^{(2)}_{\theta_0} - \frac{3\,b_1}{2}\int_0^2\mathrm{d}\mathcal{D}_s\int_0^{\mathcal{D}_s}\mathrm{d}\mathcal{D}\,n_s(\mathcal{D}_s)\,\frac{(\mathcal{D}_s-\mathcal{D})\mathcal{D}}{\mathcal{D}_s}\\ &\quad\times\int\frac{\mathrm{d}^3\mathbf{k}\,\mathrm{d}^3\mathbf{k}'}{(2\pi)^3}\,\delta(\mathbf{k})\,\delta(\mathbf{k}')\,\exp[\mathrm{i}\,(k_r\mathcal{D}+k'_r\mathcal{D}_s)]\\ &\quad\times W[|\mathcal{D}\mathbf{k}_\perp+\mathcal{D}_s\mathbf{k}'_\perp|\,\theta_0]\\ &\quad+b_1\int_0^2\mathrm{d}\mathcal{D}\,\omega(\mathcal{D})\int_0^2\mathrm{d}\mathcal{D}_s n_s(\mathcal{D}_s)\\ &\quad\times\int\frac{\mathrm{d}^3\mathbf{k}\,\mathrm{d}^3\mathbf{k}'}{(2\pi)^3}\,\delta(\mathbf{k})\,\delta(\mathbf{k}')\,\exp[\mathrm{i}\,(k_r\mathcal{D}+k'_r\mathcal{D}_s)]\\ &\quad\times W[|\mathcal{D}\mathbf{k}_\perp+\mathcal{D}_s\mathbf{k}'_\perp|\,\theta_0],\end{aligned} \quad (22)$$

In the following, the moments will be calculated with an angular top-hat filter, that in k-space reads,

$$W(k) = 2\frac{J_1(k)}{k}, \quad (23)$$

where J_1 is the spherical Bessel function.

3.2. Expression of the variance

The variance can be calculated straightforwardly from the expression of the linear convergence. Using the small angle approximation we have (Bernardeau et al. 1997),

$$\begin{aligned}\left\langle\kappa^2\right\rangle &= \frac{\Gamma[(1-n)/2]\,\Gamma[1+n/2]}{\Gamma[1-n/2]\,\Gamma[2-n/2]\,\pi^{3/2}}\,\theta_0^{-(n+2)}\\ &\quad\times\int_0^2\mathrm{d}\mathcal{D}\,\omega^2(\mathcal{D})\,\mathcal{D}^{-(n+2)}.\end{aligned} \quad (24)$$

And in order to simplify the following expressions I define the integral I_2 with,

$$I_2\equiv\int_0^2\mathrm{d}\mathcal{D}\,\omega^2(\mathcal{D})\,\mathcal{D}^{-(n+2)}. \quad (25)$$

3.3. Expression of the skewness

For a top-hat filter, using the small angle approximation and when the source clustering is neglected, the skewness is given by (Bernardeau et al. 1997),

$$s_3 \equiv \frac{\langle\kappa^3\rangle}{\langle\kappa^2\rangle^2} \tag{26}$$

$$s_3 = -\left[\frac{36}{7}-\frac{3\,(n+2)}{2}\right] \times\int_0^2 \mathrm{d}\mathcal{D}\,\omega^3(\mathcal{D})\,\mathcal{D}^{-2(n+2)}\,a(\mathcal{D})/I_2^2. \tag{27}$$

This result has been obtained from specific properties of the angular top-hat filter (see Bernardeau 1995). For this filter no further approximation than the small angle approximation is required. To compute the skewness taking into account the source clustering, it is of interest to assume that,

$$\frac{1}{2\pi}\int_0^{2\pi}\sin(\theta)\mathrm{d}\theta\,W\left(|\mathbf{k}+\mathbf{k}'|\right)=W(k)\,W(k'), \tag{28}$$

where θ is the angle between the wave vectors $\mathbf{k}$ and $\mathbf{k}'$. This property is not exact, but the error it induces is extremely weak, less than 1% for $n\approx -1.5$. This property implies in particular that the two filtering schemes Eqs. (6) and (9) give the same results for top-hat filtering.

Taking advantage of this expression, we have,

$$s_3^{\mathrm{s.c.}} = s_3-\frac{9}{I_2^2}\,b_1\int_0^2\mathrm{d}\mathcal{D}\int_0^2\mathrm{d}\mathcal{D}'\,\omega(\mathcal{D})\,\frac{\mathcal{D}(\mathcal{D}'-\mathcal{D})}{\mathcal{D}'} \times n_s(\mathcal{D}')\,\omega(\mathcal{D}')\,\mathcal{D}^{-(n+2)}\,\mathcal{D}'^{-(n+2)} +\frac{6}{I_2}\,b_1\int_0^2\mathrm{d}\mathcal{D}\,\omega(\mathcal{D})\,n_s(\mathcal{D})\,\mathcal{D}^{-(n+2)}/I_2, \tag{29}$$

when source clustring is taken into account. It of course depends only on b_1. In Table 1, I give the results for the source models described in Sect. 5.

4. The kurtosis

In the frame of Perturbation Theory, the expression of the kurtosis is given by

$$\left\langle\kappa^4\right\rangle_c \equiv \left\langle\kappa^4\right\rangle-3\left\langle\kappa^2\right\rangle^2 = 6\left\langle\left[\kappa^{(1)}\right]^2\left[\kappa^{(2)}\right]^2\right\rangle_c+4\left\langle\left[\kappa^{(1)}\right]^3\kappa^{(3)}\right\rangle_c. \tag{30}$$

When the clustering effects is neglected the third order κ is given by the integral over the line-of-sight of the local third order density. Then, using the general formulae of Bernardeau (1995) one can compute the kurtosis,

$$s_4 \equiv \frac{\left\langle\kappa^4\right\rangle_c}{\left\langle\kappa^2\right\rangle^3}; \tag{31}$$

$$s_4 = \left[\frac{2540}{49}-33\,(n+2)+\frac{21\,(n+2)^2}{4}\right] \times\int_0^2\mathrm{d}\mathcal{D}\,\omega^4(\mathcal{D})\,\mathcal{D}^{-3(n+2)}\,a^2(\mathcal{D})/I_2^3. \tag{32}$$

The general expression of the third order moment contains extra terms,

$$\begin{aligned}\kappa^{\mathrm{s.c.}(3)} = {}& \kappa^{(3)}-\frac{3}{2}\int_0^2\mathrm{d}\mathcal{D}_s\int_0^{\mathcal{D}_s}\mathrm{d}\mathcal{D}\,\frac{(\mathcal{D}_s-\mathcal{D})\mathcal{D}}{\mathcal{D}_s} \times\delta^{(2)}_{\mathrm{mass.}}(\mathcal{D})\,n_s(\mathcal{D}_s)\,\delta^{(1)}_{\mathrm{s}}(\mathcal{D}_s)\\ &-\frac{3}{2}\int_0^2\mathrm{d}\mathcal{D}_s\int_0^{\mathcal{D}_s}\mathrm{d}\mathcal{D}\,\frac{(\mathcal{D}_s-\mathcal{D})\mathcal{D}}{\mathcal{D}_s} \times\delta^{(1)}_{\mathrm{mass.}}(\mathcal{D})\,n_s(\mathcal{D}_s)\,\delta^{(2)}_{\mathrm{s}}(\mathcal{D}_s)\\ &+\int_0^2\mathrm{d}\mathcal{D}\,\omega(\mathcal{D})\delta^{(2)}_{\mathrm{mass.}}(\mathcal{D})\int_0^2\mathrm{d}\mathcal{D}_s\,n_s(\mathcal{D}_s)\,\delta^{(1)}_{\mathrm{s}}(\mathcal{D}_s)\\ &+\frac{3}{2}\int_0^2\mathrm{d}\mathcal{D}_s\,n_s(\mathcal{D}_s)\int_0^{\mathcal{D}_s}\mathrm{d}\mathcal{D}\,\frac{(\mathcal{D}_s-\mathcal{D})\mathcal{D}}{\mathcal{D}_s} \times\delta^{(1)}_{\mathrm{mass.}}(\mathcal{D})\,\delta^{(1)}_{\mathrm{s.}}(\mathcal{D}_s)\int_0^2\mathrm{d}\mathcal{D}_s\,n_s(\mathcal{D}_s)\,\delta^{(1)}_{\mathrm{s}}(\mathcal{D}_s)\\ &+\int_0^2\mathrm{d}\mathcal{D}\,\omega(\mathcal{D})\,\delta^{(1)}_{\mathrm{mass.}}(\mathcal{D}) \times\left(\int_0^2\mathrm{d}\mathcal{D}_s\,n_s\,\delta^{(2)}_{\mathrm{s}}-\left[\int_0^2\mathrm{d}\mathcal{D}_s\,n_s\,\delta^{(1)}_{\mathrm{s}}\right]^2\right).\end{aligned} \tag{33}$$

As a result, when one wants to take fully into account the source-clustering effects, many new terms should be included.

In order to simplify the expression of the resulting kurtosis, I introduce the function,

$$\mathcal{V}(\mathcal{D},\mathcal{D}') = -\frac{3}{2}\mathcal{H}(\mathcal{D}'-\mathcal{D})\,\frac{(\mathcal{D}'-\mathcal{D})\mathcal{D}}{\mathcal{D}'}\,n_s(\mathcal{D}') +\omega(\mathcal{D})\,n_s(\mathcal{D}'), \tag{34}$$

where $\mathcal{H}$ is the Heaviside function, $\mathcal{H}(\mathcal{D}'-\mathcal{D})=0$ if $\mathcal{D}'<\mathcal{D}$ and $\mathcal{H}(\mathcal{D}'-\mathcal{D})=1$ if $\mathcal{D}'\geq\mathcal{D}$.

Then, $s_4^{\mathrm{s.c.}}$ is given by

$$\begin{aligned}s_4^{\mathrm{s.c.}} = {}& s_4-\frac{4}{I_2^3}\,b_1\left[\frac{36}{7}-\frac{3}{2}(n+2)\right] \times\int_0^2\mathrm{d}\mathcal{D}_1\,\omega(\mathcal{D}_1)\,\mathcal{D}_1^{-(n+2)}\int_0^2\mathrm{d}\mathcal{D}_2\,\omega^2(\mathcal{D}_2)\\ &\times\left[\mathcal{V}(\mathcal{D}_1,\mathcal{D}_2)+\mathcal{V}(\mathcal{D}_2,\mathcal{D}_1)\right]\mathcal{D}_2^{-2(n+2)}\\ &+\frac{12}{I_2^3}\,b_1^2\int_0^2\mathrm{d}\mathcal{D}_1\,\omega(\mathcal{D}_1)\,\mathcal{D}_1^{-(n+2)}\\ &\times\int_0^2\mathrm{d}\mathcal{D}_2\,\left[\mathcal{V}(\mathcal{D}_1,\mathcal{D}_2)+\mathcal{V}(\mathcal{D}_2,\mathcal{D}_1)\right]\,\mathcal{D}_2^{-(n+2)}\\ &\times\int_0^2\mathrm{d}\mathcal{D}_3\,\left[\mathcal{V}(\mathcal{D}_2,\mathcal{D}_3)+\mathcal{V}(\mathcal{D}_2,\mathcal{D}_3)\right]\mathcal{D}_3^{-(n+2)}\,\omega(\mathcal{D}_3)\\ &-\frac{24}{I_2^3}\,b_2\int_0^2\mathrm{d}\mathcal{D}_1\,\omega(\mathcal{D}_1)\,\mathcal{D}_1^{-(n+2)}\\ &\times\int_0^2\mathrm{d}\mathcal{D}_2\,\omega^2(\mathcal{D}_2)\,\mathcal{V}(\mathcal{D}_1,\mathcal{D}_2)\,\mathcal{D}_2^{-2(n+2)}\\ &-\frac{24}{I_2^3}\,b_1^2\int_0^2\mathrm{d}\mathcal{D}_1\,\omega(\mathcal{D}_1)\,\mathcal{D}_1^{-(n+2)}\\ &\times\int_0^2\mathrm{d}\mathcal{D}_2\,\omega(\mathcal{D}_2)\,n_s(\mathcal{D}_2)\,\mathcal{D}_2^{-(n+2)}\end{aligned}$$

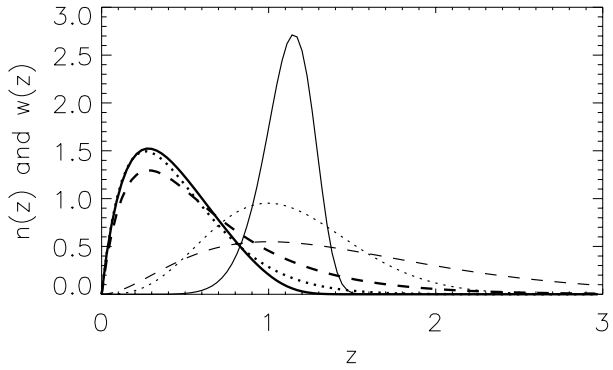

Fig. 2. Shapes of the distribution functions of the sources (thin lines) and lens efficiency functions (thick lines) for the three models (models 1, 2, 3 are respectively plotted with solid, dotted and dashed lines). All functions are arbitrarily normalized to unity

$$\times \int_{\mathcal{D}_1}^{2} \mathrm{d}\mathcal{D}_3\, \omega(\mathcal{D}_3)\, n_s(\mathcal{D}_3)\, \mathcal{D}_3^{-(n+2)} \times \left[\frac{3}{2}\,\frac{(\mathcal{D}_3-\mathcal{D}_1)\mathcal{D}_1}{\mathcal{D}_3} - \omega(\mathcal{D}_1)\right] \quad (35)$$

Note that, contrary to the skewness case, there is a contributing term

$$s_4^{\rm bumps} \equiv \frac{12}{I_2^3}\, b_1^2 \int_0^2 \mathrm{d}\mathcal{D}_1\, \omega(\mathcal{D}_1)\, \mathcal{D}_1^{-(n+2)} \times \int_0^2 \mathrm{d}\mathcal{D}_2\, \mathcal{V}(\mathcal{D}_1,\mathcal{D}_2)\, \mathcal{D}_2^{-(n+2)} \times \int_0^2 \mathrm{d}\mathcal{D}_3\, \mathcal{V}(\mathcal{D}_2,\mathcal{D}_3)\, \mathcal{D}_3^{-(n+2)}\, \omega(\mathcal{D}_3), \quad (36)$$

which is due to the source auto-correlation function only. It corresponds to the second mechanism described in Sect. 2.2. This term does not disappear a priori when the overlapping between the source distribution and the lens efficiency function is arbitrarily small. The results shown in Table 1 prove however that in practice this contribution is always negligible compared to the cross-correlation effects.

5. Discussions

5.1. Numerical results

In order to get numerical results one should choose a specific model for the redshift source distribution. I will assume that n_s takes the form,

$$n_s(z) \propto z^\alpha \,\exp\left[-(z/z_0)^\beta\right]. \quad (37)$$

I make the calculations for three hypothesis,

$$\text{model } 1 : z_0 = 1.15, \quad \alpha = 8., \quad \beta = 8.; \quad (38)$$

$$\text{model } 2 : z_0 = 0.75, \quad \alpha = 3., \quad \beta = 1.8; \quad (39)$$

$$\text{model } 3 : z_0 = 0.5, \quad \alpha = 2., \quad \beta = 1. \quad (40)$$

For the first two models the mean redshift of the sources is about unity, with a larger distribution in the second case. The first model has somewhat arbitrary parameters, whereas the second is motivated by results of galaxy evolution models (Charlot & Fall, in preparation). It would correspond to galaxies with a I magnitude between 22 and 24. It has a redshift dispersion higher than model 1. In model 3, I have assumed a very broad distribution of the redshift distribution. The resulting values for the skewness and the kurtosis in such models are given in Table 1. As expected the first two models give roughly the same answers, because the mean source redshifts are very close. In the third model there is a significant number of sources at very high redshift. It lowers the values of s_3 and s_4.

Skewness and kurtosis show a similar sensitivity with source clustering effects. For the skewness, the remaining corrective term can be as large as 25% of the signal when the redshift distribution of the sources is large. The correction is about 30% for the kurtosis assuming that $b_1 \approx 1$ and $b_2 \approx 0.5$. However for a narrow redshift distribution, the corrective terms remain small (about 3% for model 1 for the skewness and 2% for the kurtosis). It shows that the high order moments are slightly dependent on source clustering but that this dependence can be controlled if the redshift dispersion of the sources is low enough. If one does not want to include pre-knowledge on the galaxy-mass cross-correlation to analyze the data, it will be important to reduce as much as possible the width of the redshift distribution in the adopted selection criteria. Note that it will be anyway possible to get constraints on the mass-source and source-source correlations from counts in cells statistics applied to the selected sources (Schneider 1997, van Waerbeke 1998).

It is also worth to have in mind that these results depend on the shape of the filter. In particular if one does not use a top-hat filter the two filtering schemes are not equivalent. This would be the case in particular for compensated filters. The "proper" surface-weighting scheme is expected to be, in general, less sensitive to the source clustering.

5.2. Skewness and kurtosis to measure the non-Gaussian effects

In the preliminary investigations of the non-Gaussian properties expected to be observable in convergence maps, the focus has been mainly put on the skewness. It is indeed the first non-trivial cumulant expected to emerge due of mode couplings and its calculation is possible in the frame of Perturbation Theory. However, it would be stupid to limit the search of non-Gaussian effects to the skewness only. Even at the level of the shape of the local convergence PDF, the skewness cannot entirely characterize the departure from a gaussian distribution. This departure manifests itself by the apparition of a whole set of cumulants. In the context we are interested in, the PDF of κ is essentially a 2 parameter familly (when a given population of sources is considered): it depends on the amplitude of the density fluctuations and on Ω (neglecting at first view the Λ dependence). It implies that all cumulants are somehow connected together. In particular the kurtosis s_4 appears simply to be an other way of constraining Ω (it is in perturbation theory independent of σ_8) which means that s_3 and s_4 are naturally related. This relation

Table 1. Skewness and kurtosis with source clustering effects

	model 1	model 2	model 3
$\langle z_s \rangle$	1.11	1.11	1.50
$\Delta(z_s)$	0.15	0.42	0.87
$s_3^{\rm s.c.}$	$-37.6 + 1.05\,b_1$	$-39.9 + 6.4\,b_1$	$-29.1 + 7.8\,b_1$
$s_4^{\rm s.c.}$	$2890 - 49\,b_1 + 12\,b_1^2 - 24\,b_2$	$3330 - 410\,b_1 - 217\,b_1^2 - 280\,b_2$	$1790 - 390\,b_1 - 4\,b_1^2 - 290\,b_2$
$s_4^{\rm bumps}$	$1.3\,b_1^2$	$19\,b_1^2$	$21\,b_1^2$

can be easily identified from the results of Table 1. Indeed one can check that, in the absence of source clustering,

$$\frac{s_4}{s_3^2} \approx 2.05 \text{ to } 2.1, \tag{41}$$

for the three models. The dependence of the ratio s_4/s_3^2 on n is expected to be weak, as well as its dependence on the cosmological parameters. We encounter here the same situation that has been noticed by Bernardeau (1994) for the cumulants of the cosmic density or the cosmic velocity divergence filtered with a 3D top-hat window function. This property is a priori valid in the quasilinear regime only. In the strongly non-linear regime the validity of this relationship remains an open question although results obtained in a phenomenological description of the cumulant behavior of the 3D cumulants (Colombi et al. 1997) strongly suggests that such a relation may survive in the non-linear regime. The consequences are two fold,

- s_4 might be used as an alternative method of measuring Ω;
- the fact that s_4/s_3^2 should be approximatly 2 in all cases might also reveal a precious property to use for testing the correctness of observational results. For instance, when the bias effects are too strong this ratio tends indeed to increase up to about 2.7.

5.3. Systematics and spurious couplings

To summarize, the possible spurious couplings that may affect the statistical properties of the convergence maps that have been identified so far are:

- Multiple lens couplings. This is due to the fact that the convergence effects of lenses that are aligned do not add linearly. Actually, to compute the effect of combined lenses one should multiply the amplification matrices. Departure from the Born approximation appears at the same level of approximation. Both are shown to induce at most a few percent effect correction on the skewness (see Bernardeau et al. 1997).
- Magnification effects. This is due to the fact that the population of selected sources may depend on the local magnification (see again Bernardeau et al. 1997). This effect is difficult to quantify a priori. It should be tested with the selection algorithms which are actually used.
- Source density fluctuations. This is one of the effects which is investigated in this paper. It comes from the fact that the source plane appears "bumpy". At large scale, it induces corrective terms at the level of the kurtosis only, but it might play a more important role at small scale.
- Source-lens correlations. This effect, the main effect investigated in this paper, is due to the fact that the sources might also somehow trace the foreground lenses when the distribution function of the source redshift is large enough. The latter two effects are shown to be negligible when the width of the redshift distribution is small enough (say less than 0.15).

All these effects have been shown to intervene for the high order moments only, in the frame of perturbation theory that is at large enough scale (above about 10 arcmin scale). In this approach it is possible to estimate the magnitude of these effects and to show that for a reasonable redshift distribution they are negligible.

In practice however it is likely that the scales of interest will be much smaller than scale at which the perturbation theory regime is valid. It would therefore be interesting to extend these results to the small scales. In particular the effect of source clustering on the variance might not be totally negligible.

Acknowledgements. The author would like to thank Yannick Mellier and Ludovic van Waerbeke for many discussions and the referee, Bhuvnesh Jain, for suggestions that have contributed to the improvement of the manuscript. The author is also grateful to IAP where most of this work has been conducted.

References

Bardeen J.M., Bond J.R., Kaiser N., Szalay A.S. 1986, ApJ 304, 15
Bernardeau F. 1994, ApJ 433, 1
Bernardeau F. 1995, A&A 301, 309
Bernardeau F., Schaeffer R. 1992, A&A 255, 1
Bernardeau F., van de Weygaert R. 1996, MNRAS 279, 693
Bernardeau F., van Waerbeke L., Mellier Y. 1997, A&A 322, 1
Blandford R.D., Saust A.B., Brainerd T.G., Villumsen J.V. 1991, MNRAS 251, 600.
Bouchet F.R., Strauss M.A., Davis M., et al. 1993, ApJ 417, 36
Broadhurst T.J., Taylor A.N., Peacock J.A. 1995, ApJ 438, 49
Colombi S., Bernardeau F., Bouchet F.R., Hernsquist L. 1997, MNRAS 287, 241
Fry J.N. 1984, ApJ 279, 499
Goroff M.H. Grinstein B., Rey S.-J., Wise M.B. 1986, ApJ 311, 6
Jain B., Seljak U. 1997, ApJ 484, 560
Kaiser N. 1992, ApJ 388, L72
Kaiser N. 1995, ApJ 439, L1

Kim R.S., Strauss M.A. 1998, ApJ 493, 39
Miralda-Escudé J. 1991, ApJ 380, 1
Peebles P.J.E. 1980,*The large scale Structure of the Universe*, Princeton Univ. Press
Schneider P. 1997, astro-ph/9708269
Schneider P., van Waerbeke L., Mellier Y., Jain B., Seitz S., Fort B. 1998a, MNRAS 333, 767
Schneider P., van Waerbeke L., Jain B., Kruse G. 1998b, MNRAS 296, 873
Van Waerbeke L., 1998, A&A 334, 1
Van Waerbeke L., Mellier Y., Schneider P., Fort B., Mathez G. 1997, A&A 317, 303
Villumsen J.V. 1996, MNRAS 281, 369