Astron. Astrophys. 335, 1–11 (1998)

ASTRONOMY AND ASTROPHYSICS

Improving the accuracy of mass reconstructions from weak lensing: from the shear map to the mass distribution

Marco Lombardi and Giuseppe Bertin

Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

Received 27 January 1998 / Accepted 16 March 1998

Abstract. In this paper we provide a statistical analysis of the parameter-free method often used in weak lensing mass reconstructions. It is found that a proper assessment of the errors involved in such a non-local analysis requires the study of the relevant two-point correlation functions. After calculating the two-point correlation function for the reduced shear, we determine the expected error on the inferred mass distribution and on other related quantities, such as the total mass, and derive the error power spectrum. This allows us to optimize the reconstruction method, with respect to the kernel used in the inversion procedure. The optimization process can also be carried out on the basis of a variational principle. In particular, we find that curl-free kernels are bound to lead to more accurate mass reconstructions. Our analytical results clarify the arguments and the numerical simulations by Seitz & Schneider (1996).

Key words: gravitational lensing – dark matter – galaxies: clusters

1. Introduction

One of the most interesting applications of gravitational lenses is the determination of the projected mass distribution from weak lensing observations. As noted, among others, by Webster (1985), the mean orientation of a large number of distant galaxies gives a measure of the *shear* associated with the lens. The observed shear can then be used to derive the two-dimensional mass distribution of the lens responsible for the deformation induced on the background. This last step can be carried out in two different ways. The easier route is to use a specific model for the lens with a number of free parameters that will be determined by a comparison between the observed and the predicted shear (see, e.g., Kneib *et al*. 1996). A more general procedure is the so called "parameter-free reconstruction" (Kaiser & Squires 1993; see also Bartelmann *et al*. 1996). In this latter method the mass distribution can be directly determined from the shear map, provided that the shear is known with sufficient accuracy and detail, which requires the existence of a large number of source galaxies.

Such reconstruction techniques are, of course, a powerful tool to study the matter distribution in clusters (see e.g. Tyson, Valdes, Wenk 1990, Fahlman *et al.* 1994, Smail *et al.* 1994) and for large scale structures. It is then important to optimize the reconstruction process in order to make the best use of the observations. For this purpose, we have to assess the expected error of a specific reconstruction method, which is the main goal of the present paper.

In this article we focus our attention on the parameter-free method, mainly because this is more general and does not depend on the particular lens under consideration. In a previous paper (Lombardi & Bertin 1998, hereafter Paper I) we have provided expressions for the error involved in the *local* measurements of the shear (or the reduced shear) of the lens as a function of the parameters characterizing the distribution of source galaxies. Here we extend the statistical analysis to the inferred global mass distribution.

The text is organized as follows. In Sect. 2 we introduce the spatial weight function and we briefly describe various reconstruction methods used to infer the lens mass distribution. In Sect. 3 we calculate the expected error on the measured shear in the regime of weak lensing and in the more general case as a function of position in a given field of the sky; here the formulae of Paper I are generalized to the two-point correlation function for the shear map (see Eq. (25)). This important result is then used in Sect. 4 to calculate the expected errors on the mass distribution associated with the various reconstruction methods. The results are then compared, in Sect. 5, to the simulations by Seitz & Schneider (1996).

The main result of the paper is contained in Eq. (35) (together with Eq. (26)) that describes the two-point correlation function for the mass density κ obtained from weak lensing analysis. This proves that, in order to optimize the reconstruction process for observations in a finite area of the sky, a curl-free kernel should be used (see Eq. (36)). This behavior is confirmed by numerical simulations.

2. From the shear map to the mass distribution

We consider a field of the sky with N source galaxies located at $\boldsymbol{\theta}^{(n)}$ and characterized by observed quadrupole $Q^{(n)}$ and ellipticity $\chi^{(n)}$ (see Appendix A for a summary of the adopted

Send offprint requests to: M. Lombardi notation). Here we suppose that the galaxies are observed inside a field Ω of area A, with mean spatial density equal to $\rho = N/A$. In the rest of the paper we will reserve the term "weak lensing" to the limit of small lens density $\kappa \ll 1$, i.e. $\gamma \simeq g$.

2.1. Spatial weight function

Source galaxies located close to a given position $\boldsymbol{\theta}$ will better constrain the value of the reduced shear $g(\boldsymbol{\theta})$ at such location. In order to describe this effect, we may thus introduce a suitable weight function $W(\boldsymbol{\theta}, \boldsymbol{\theta}')$. The first argument of the weight function, $\boldsymbol{\theta}$, represents the point of the sky under consideration and for which we want to measure the shear $g(\boldsymbol{\theta})$, while the second argument $\boldsymbol{\theta}'$ represents the location of one observed galaxy. The weight function should penalize galaxies far from $\boldsymbol{\theta}$, i.e. $W(\boldsymbol{\theta}, \boldsymbol{\theta} + \boldsymbol{\vartheta})$ should decrease for increasing $\|\boldsymbol{\vartheta}\|$. Some additional "natural" conditions can be given to further characterize a specific choice of weight function and these are most convenient when applied (beginning with Eq. (22)) to a spatially continuous distribution of source galaxies with density ρ. Here we list a few possible assumptions, where the first is obviously the least restrictive:

- 1. The weight function is *even* with respect to $\boldsymbol{\vartheta}$, i.e.

$$W(\boldsymbol{\theta}, \boldsymbol{\theta} + \boldsymbol{\vartheta}) = W(\boldsymbol{\theta}, \boldsymbol{\theta} - \boldsymbol{\vartheta}) \; . \tag{1}$$

- 2. The weight function is said *invariant upon translations*, if it is *even* (see above) and if it depends only on the difference $\boldsymbol{\theta} - \boldsymbol{\theta}'$:

$$W(\boldsymbol{\theta}, \boldsymbol{\theta}') = W(\boldsymbol{\theta} - \boldsymbol{\theta}') \; . \tag{2}$$

- 3. One natural choice is that of a *Gaussian* dependent only on the distance $\|\boldsymbol{\theta} - \boldsymbol{\theta}'\|$,

$$W(\boldsymbol{\theta}, \boldsymbol{\theta}') = \frac{1}{2\pi\rho\sigma_W^2} \exp\left(-\frac{\|\boldsymbol{\theta} - \boldsymbol{\theta}'\|^2}{2\sigma_W^2}\right) \; , \tag{3}$$

  where the angular scale σ_W should be sufficiently large to ensure the presence of an adequate number of galaxies in a disk of radius σ_W centered on the generic point $\boldsymbol{\theta}$.

The value of the weight function at a given point $\boldsymbol{\theta}'$, of course, has no particular meaning: only *relative* values are significant. Indeed all the following results can be shown to be unaffected if we merely multiply the weight function by a constant. Thus, we may always choose a *normalized* weight function, so that

$$\sum_{n=1}^{N} W\bigl(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)}\bigr) = 1 \tag{4}$$

for every $\boldsymbol{\theta}$. Of course, such a normalization will remove the translation invariance property, if initially present in W. Still, the invariance may be retained when we will move to a continuous description (see comment after Eq. (22)).

The spatial weight function operates much like the "shape" weight functions considered in Paper I (see Eqs. (23) and (21) there). In particular, using the isotropy condition, we can obtain the shear map $g(\boldsymbol{\theta})$ either from

$$\sum_{n=1}^{N} W\bigl(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)}\bigr) \chi^{\mathrm{s}}\bigl(\chi^{(n)}, g(\boldsymbol{\theta})\bigr) = 0 \; , \tag{5}$$

or from

$$\chi^{\mathrm{s}}\left(\sum_{n=1}^{N} W\bigl(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)}\bigr) Q^{(n)}, g(\boldsymbol{\theta}) \right) = 0 \; . \tag{6}$$

In Paper I we discussed the different merits of the two options. In the limit of "sharp" distributions for the source galaxies ($c \ll 1$), they both lead to the same determination of the true reduced shear g_0 (apart from the ambiguity associated with the $g \mapsto 1/g^*$ invariance).

As we will see the angular scale of the weight function W, i.e. the diameter of the set where $W(\boldsymbol{\theta}, \boldsymbol{\theta}')$ is significantly different from zero, determines a lower bound for the smallest details shown in the reconstructed map $\kappa(\boldsymbol{\theta})$. For example, in the weak lensing limit and for a weight function invariant upon translations, the mean value of the measured density κ is related to the true density κ_0 through the expression (see Appendix C)

$$\langle \kappa(\boldsymbol{\theta}) \rangle = \rho \int W(\boldsymbol{\theta} - \boldsymbol{\theta}') \kappa_0(\boldsymbol{\theta}') \, \mathrm{d}^2\theta' \; . \tag{7}$$

This relation, similar to that found when an image is degraded by a PSF, suggests the possible application of deconvolution techniques (see Lucy 1994) to the present context.

2.2. Weak lensing regime

We first consider the case where Ω is very large, so that the field is identified with the whole plane (the effects of the boundaries will be discussed soon, in Sect. 2.4).

There are basically two ways to reconstruct the mass distribution $\kappa(\boldsymbol{\theta})$ from the shear map $g(\boldsymbol{\theta})$ (Seitz & Schneider 1996). The first, more natural method is based on the integral relation

$$\kappa(\boldsymbol{\theta}) = \int D_i(\boldsymbol{\theta} - \boldsymbol{\theta}') \gamma_i(\boldsymbol{\theta}') \, \mathrm{d}^2\theta' \; , \tag{8}$$

where the kernel D_i is given by (Kaiser & Squires 1993)

$$\begin{pmatrix} D_1(\boldsymbol{\theta}) \\ D_2(\boldsymbol{\theta}) \end{pmatrix} = \frac{1}{\pi \|\boldsymbol{\theta}\|^4} \begin{pmatrix} \theta_1^2 - \theta_2^2 \\ 2\theta_1\theta_2 \end{pmatrix} \; . \tag{9}$$

In the weak lensing limit $\gamma \simeq g$, and thus the reduced shear map can be used directly in Eq. (8) to derive $\kappa(\boldsymbol{\theta})$. Note that the inverse relation holds with the *same* kernel

$$\gamma_i(\boldsymbol{\theta}) = \int D_i(\boldsymbol{\theta} - \boldsymbol{\theta}') \kappa(\boldsymbol{\theta}') \, \mathrm{d}^2\theta' \; . \tag{10}$$

This relation will turn out to be useful in Appendix C.

A second possibility, which can be proved to be mathematically equivalent to the first, is based on the exact relation (see Kaiser 1995)

$$\nabla\kappa = \boldsymbol{u}(\theta) = -\begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix} \; , \tag{11}$$

which is a direct consequence of the thin lens equations. Here $\gamma_{i,j} = \partial \gamma_i / \partial \theta_j$. By analogy with the condition used to derive Eq. (8), if we assume that $\kappa(\boldsymbol{\theta})$ vanishes for large values of $\|\boldsymbol{\theta}\|$, Eq. (11) can be inverted to give

$$\kappa(\boldsymbol{\theta}) = \int H_i^{\mathrm{KS}}(\boldsymbol{\theta} - \boldsymbol{\theta}') u_i(\boldsymbol{\theta}') \,\mathrm{d}^2\theta' \,, \tag{12}$$

with the kernel

$$H_i^{\mathrm{KS}}(\boldsymbol{\theta}) = \frac{\theta_i}{2\pi \|\boldsymbol{\theta}\|^2} \,. \tag{13}$$

In Eq. (12) the shear map enters through the vector $\boldsymbol{u}$, which, in the weak lensing limit, involves the derivatives of $g(\boldsymbol{\theta})$. This second method thus introduces undesired differentiations, but it has the advantage that it is more easily generalized to include the effects of the boundaries (see Sect. 2.4 below).

2.3. *The general case*

When the lens is not weak, Eq. (8)

$$\kappa(\boldsymbol{\theta}) = \int D_i(\boldsymbol{\theta} - \boldsymbol{\theta}') \bigl(1 - \kappa(\boldsymbol{\theta}')\bigr) g_i(\boldsymbol{\theta}') \,\mathrm{d}^2\theta' \tag{14}$$

can be solved by iteration (Seitz & Schneider 1995).

The second method, related to Eq. (12), has been generalized by Kaiser (1995) for strong lenses. If we introduce $\tilde\kappa(\boldsymbol{\theta}) = \ln\bigl(1 - \kappa(\boldsymbol{\theta})\bigr)$ and the new vector

$$\tilde u_i = \frac{1}{1 - |g|^2} \begin{pmatrix} 1 + g_1 & g_2 \\ g_2 & 1 - g_1 \end{pmatrix} \begin{pmatrix} g_{1,1} + g_{2,2} \\ g_{2,1} - g_{1,2} \end{pmatrix} \,, \tag{15}$$

then it is possible to show that the relation $\nabla \tilde\kappa(\boldsymbol{\theta}) = \tilde{\boldsymbol{u}}(\boldsymbol{\theta})$ holds. As a result $\tilde\kappa$ can be obtained from $\tilde{\boldsymbol{u}}$ via the same integral equation (12) used earlier. The fact that $\tilde\kappa$ is determined only up to a constant here translates into a non-trivial invariance for the density distribution $\kappa(\boldsymbol{\theta})$, under the transformation (see Schneider & Seitz 1995)

$$\kappa(\boldsymbol{\theta}) \longmapsto (1 - C)\kappa(\boldsymbol{\theta}) + C \,, \tag{16}$$

consistent with Eq. (14).

2.4. *Effect of the boundaries*

The methods described so far assume an infinite domain of integration. In practice, one can measure the shear only in a finite area (e.g. the CCD area), which is often small compared to the angular size of the lensing cluster. Therefore, the relations given earlier should be properly modified.

We briefly noted that the second method is better suited for the purpose. In the following, for simplicity, we consider only the weak lensing limit, but the equations that we will provide can be easily extended to the general case by replacing $(\boldsymbol{u}, \kappa) \mapsto (\tilde{\boldsymbol{u}}, \tilde\kappa)$. The relations suggested by Seitz & Schneider (1996) for mass reconstruction in a field Ω of finite area A are of the form

$$\kappa(\boldsymbol{\theta}) - \bar\kappa = \int_\Omega G_i(\boldsymbol{\theta}, \boldsymbol{\theta}') u_i(\boldsymbol{\theta}') \,\mathrm{d}^2\theta' \,. \tag{17}$$

Here $\bar\kappa$ is a constant representing the average of κ, while $\boldsymbol{G}$ is a suitable kernel. The kernel is chosen so as to give the correct mass distribution if $\boldsymbol{u}$ could be measured with no errors (see Eq. (C14)). There is however some freedom left in the choice of the kernel, mainly because it returns a scalar field (κ) from a vector field ($\boldsymbol{u}$). This freedom will be further discussed later on. One interesting kernel, called *noise filtering*, has been introduced by Seitz & Schneider (1996)

$$\boldsymbol{H}^{\mathrm{SS}}(\boldsymbol{\theta}, \boldsymbol{\theta}') = -\nabla_{\boldsymbol{\theta}'} \mathcal{H}^{\mathrm{SS}}(\boldsymbol{\theta}, \boldsymbol{\theta}') \,, \tag{18}$$

where $\mathcal{H}^{\mathrm{SS}}$ is the solution of Neumann's boundary problem ($\boldsymbol{n}$ is the unit vector orthogonal to $\partial\Omega$)

$$\nabla^2_{\boldsymbol{\theta}'} \mathcal{H}^{\mathrm{SS}}(\boldsymbol{\theta}, \boldsymbol{\theta}') = \delta(\boldsymbol{\theta} - \boldsymbol{\theta}') - \frac{1}{A} \,, \tag{19}$$

$$\frac{\partial \mathcal{H}^{\mathrm{SS}}(\boldsymbol{\theta}, \boldsymbol{\theta}')}{\partial \theta'_i} n_i(\boldsymbol{\theta}') = 0 \qquad \forall \boldsymbol{\theta}' \in \partial\Omega \,. \tag{20}$$

The term related to the area A ensures the proper applicability of the Gauss theorem. Note that the kernel $\boldsymbol{H}^{\mathrm{SS}}$ has vanishing curl, i.e.

$$\nabla_{\boldsymbol{\theta}'} \wedge \boldsymbol{H}^{\mathrm{SS}}(\boldsymbol{\theta}, \boldsymbol{\theta}') = 0 \,. \tag{21}$$

3. Measurements of the reduced shear map and of the two-point correlation function

In this section we will give an expression for the reduced shear map measured using Eqs. (5) or (6). In Paper I we have calculated the statistics associated with a *local* shear measurement under the hypothesis that the probability distribution for the source ellipticity χ^{s} is *sharp*, i.e. most source galaxies are nearly round. Now we consider situations where the reduced shear is a function of the position $\boldsymbol{\theta}$, but we assume that $g(\boldsymbol{\theta})$, a smooth function of $\boldsymbol{\theta}$, does not change significantly on the angular scale σ_W of the weight function $W(\boldsymbol{\theta}, \boldsymbol{\theta}')$. An important new aspect of the analysis that has to be addressed here, in view of the goal of determining the error on the reconstructed mass, is the calculation of the two-point correlation function for the shear map.

So far we have considered source galaxies with random orientation but with fixed position on the sky (corresponding to $\boldsymbol{\theta}^{(n)}$ on the observer's sky). It is interesting to average all results by assuming that galaxies have random positions. The result of the average can be approximated by considering a continuous distribution of galaxies with density ρ (number of galaxies per steradian). This leads us to ignore, for the moment, the effects of Poisson noise associated with the finite number of source galaxies (further comments are given at the end of Sect. 4.1). Following Seitz & Schneider (1996), we consider a homogeneous distribution of galaxies in $\boldsymbol{\theta}$, i.e. in the *observer's plane*. If ρ is independent of $\boldsymbol{\theta}$, we may change summations with integrals using the rule $\sum_n \mapsto \rho \int \mathrm{d}^2\theta'$. Here, to simplify the derivations and the discussion, we adopt, for every $\boldsymbol{\theta}$, the normalization

$$\rho \int W(\boldsymbol{\theta}, \boldsymbol{\theta}') \,\mathrm{d}^2\theta' = 1 \tag{22}$$

for the weight function. For an infinite field Ω or for the case described in Fig. 1, this normalization does not break the translation invariance of W. Then as shown in Appendix B, the relation between expected and true value of g, corresponding to Eq. (7), is

$$\big\langle g(\boldsymbol{\theta}) \big\rangle = \rho \int W(\boldsymbol{\theta}, \boldsymbol{\theta}') g_0(\boldsymbol{\theta}') \, \mathrm{d}^2\theta' \; . \qquad (23)$$

As is intuitive, "near" galaxies give the most important contribution to the measured value of g.

The correct generalization of the covariance matrix $\mathrm{Cov}_{ij}(g)$ when g is a function of the position $\boldsymbol{\theta}$ is a two-point correlation function:

$$\mathrm{Cov}_{ij}(g; \boldsymbol{\theta}, \boldsymbol{\theta}') = \Big\langle \big(g_i(\boldsymbol{\theta}) - \overline{g_i(\boldsymbol{\theta})} \big) \big(g_j(\boldsymbol{\theta}') - \overline{g_j(\boldsymbol{\theta}')} \big) \Big\rangle . \qquad (24)$$

Note that the knowledge of the "diagonal" values $\mathrm{Cov}_{ij}(g; \boldsymbol{\theta}, \boldsymbol{\theta})$ is not sufficient to calculate the error on other variables, such as the density distribution κ, determined from g (cf. Eq. (32) and Eq. (35)).

If we assume that the weight function is *even* (property 1 of Sect. 2.1), then the two-point correlation function of g can be written in the simple form (see Appendix B)

$$\begin{aligned} \mathrm{Cov}_{ij}(g; \boldsymbol{\theta}, \boldsymbol{\theta}') \; &= \; \frac{c}{4} \big(1 - \big| \langle g(\boldsymbol{\theta}) \rangle \big|^2 \big) \big(1 - \big| \langle g(\boldsymbol{\theta}') \rangle \big|^2 \big) \delta_{ij} \\ & \times \; \rho \int W(\boldsymbol{\theta}, \boldsymbol{\theta}'') W(\boldsymbol{\theta}', \boldsymbol{\theta}'') \, \mathrm{d}^2\theta'' \; . \end{aligned} \qquad (25)$$

Here c is the covariance of the ellipticity distribution of the source galaxies. In this equation, as noted for Eq. (23), we suppose the weight function $W(\boldsymbol{\theta}, \boldsymbol{\theta}')$ to be normalized.

In the weak lensing limit Eq. (25) then reduces to

$$\begin{aligned} \mathrm{Cov}_{ij}(\gamma; \boldsymbol{\theta}, \boldsymbol{\theta}') \; &= \; \frac{c\rho\delta_{ij}}{4} \int W(\boldsymbol{\theta}, \boldsymbol{\theta}'') W(\boldsymbol{\theta}', \boldsymbol{\theta}'') \, \mathrm{d}^2\theta'' \\ &= \; \frac{c\delta_{ij}}{16\pi\rho\sigma_W^2} \exp\left(- \frac{\|\boldsymbol{\theta} - \boldsymbol{\theta}'\|^2}{4\sigma_W^2} \right) \; . \end{aligned} \qquad (26)$$

The last relation holds for a Gaussian weight function of the form given in Eq. (3); here $\mathrm{Cov}_{ij}(\gamma; \boldsymbol{\theta}, \boldsymbol{\theta}')$ is a simple Gaussian with variance $2\sigma_W^2$ and depends only on $\|\boldsymbol{\theta} - \boldsymbol{\theta}'\|$. The variance $\big\langle \big(\gamma_i(\boldsymbol{\theta}) - \overline{\gamma_i(\boldsymbol{\theta})} \big)^2 \big\rangle$ of $\gamma_i(\boldsymbol{\theta})$ is simply $\mathrm{Cov}_{ii}(\gamma; \boldsymbol{\theta}, \boldsymbol{\theta}) = c/(16\pi\rho\sigma_W^2)$ (without summation on i) and thus increases if σ_W decreases. This behavior can be explained by considering that the number of galaxies used for a single point is of the order of $\rho\sigma_W^2$. Notice also that σ_W sets the scale length of the covariance of γ: measurements of $\gamma(\boldsymbol{\theta})$ and $\gamma(\boldsymbol{\theta}')$ are uncorrelated if $\|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \gg \sigma_W$.

4. Measurements of the mass distribution

It is not difficult, at least in principle, to calculate the error on $\kappa(\boldsymbol{\theta})$ from the two-point correlation function of g. The error on κ, of course, depends on the reconstruction method used. For this reason, following Sect. 2, we consider different methods separately. For simplicity, we suppose that the weight function W is invariant upon translations. Moreover, we suppose that the angular scale of the weight function W is much smaller than the angular scale of κ (i.e. the scale where κ varies significantly). In general, if we ignore edge effects, the relation between the error on g and that on κ is given by $\mathrm{Cov}_{ij}(\gamma; \boldsymbol{\theta}, \boldsymbol{\theta}') = \delta_{ij} \mathrm{Cov}(\kappa; \boldsymbol{\theta}, \boldsymbol{\theta}')$. Here we show only the results obtained, referring to Appendix C for a derivation. As in Sect. 2, we first refer to the case where Ω is identified with the whole plane (finite field effects will be addressed in Sect. 4.3 below).

4.1. Weak lensing regime

In this case we can use either Eq. (8) or Eq. (12). A rather surprising result is that both methods lead to the same mean values and errors for κ. The result for the mean value has already been stated in Eq. (7), i.e. the measured mass distribution $\big\langle \kappa(\boldsymbol{\theta}) \big\rangle$ is the convolution of the weight function W with the true mass distribution $\kappa_0(\boldsymbol{\theta})$.

For an "isolated lens" (a case where κ is taken to vanish outside a certain domain Ω_{in}) the ambiguity associated with Eq. (16) is resolved and the concept of total mass of the lens becomes meaningful. In the weak lensing limit, from any reconstructed $\kappa(\boldsymbol{\theta})$ one can in principle accept also $\kappa_C(\boldsymbol{\theta}) = \kappa(\boldsymbol{\theta}) + C$. Now if we know that the density vanishes outside Ω_{in}, the constant C can be determined by requiring

$$\int_{\Omega_{\mathrm{out}}} \mathrm{d}^2\theta \, \kappa_C(\boldsymbol{\theta}) = 0 \; , \qquad (27)$$

where $\Omega_{\mathrm{out}} = \Omega \setminus \Omega_{\mathrm{in}}$ is the part of the field not contained in Ω_{in}. Therefore, the appropriate density to be used is

$$\kappa_C(\boldsymbol{\theta}) = \kappa(\boldsymbol{\theta}) - \frac{1}{A_{\mathrm{out}}} \int_{\Omega_{\mathrm{out}}} \kappa(\boldsymbol{\theta}') \, \mathrm{d}^2\theta' \; . \qquad (28)$$

The associated total mass is

$$\begin{aligned} M \; &= \; \int_\Omega \kappa_C(\boldsymbol{\theta}) \, \mathrm{d}^2\theta = \int_{\Omega_{\mathrm{in}}} \kappa_C(\boldsymbol{\theta}) \, \mathrm{d}^2\theta \\ &= \; \int_{\Omega_{\mathrm{in}}} \kappa(\boldsymbol{\theta}) \, \mathrm{d}^2\theta - \frac{A_{\mathrm{in}}}{A_{\mathrm{out}}} \int_{\Omega_{\mathrm{out}}} \kappa(\boldsymbol{\theta}) \, \mathrm{d}^2\theta \; . \end{aligned} \qquad (29)$$

Therefore,

$$\begin{aligned} \langle M \rangle \; &= \; \int_{\Omega_{\mathrm{in}}} \big\langle \kappa(\boldsymbol{\theta}) \big\rangle \, \mathrm{d}^2\theta - \frac{A_{\mathrm{in}}}{A_{\mathrm{out}}} \int_{\Omega_{\mathrm{out}}} \big\langle \kappa(\boldsymbol{\theta}) \big\rangle \, \mathrm{d}^2\theta \\ &= \; \int_{\Omega_{\mathrm{in}}} \kappa_0(\boldsymbol{\theta}) \, \mathrm{d}^2\theta = M_0 \; , \end{aligned} \qquad (30)$$

where M_0 is the true mass of the lens. In other words, the *smoothing effect* associated with W does not change the measured total mass M of the lens.

The covariance of the lens distribution κ can be shown to be equal to (for both Eqs. (5) and (6)):

$$\mathrm{Cov}(\kappa; \boldsymbol{\theta}, \boldsymbol{\theta}') = \frac{c\rho}{4} \int W(\boldsymbol{\theta}, \boldsymbol{\theta}'') W(\boldsymbol{\theta}', \boldsymbol{\theta}'') \, \mathrm{d}^2\theta'' \; . \qquad (31)$$

In comparing Eq. (31) to Eq. (26), one should note that the similarity of results refers statistically to average errors, but not to the individual errors of one reconstruction. The variance in the measure of the total mass is the integral of the covariance of κ:

$$\begin{aligned}\mathrm{Var}(M) \;=\; & \int_{\Omega_\mathrm{in}} \mathrm{d}^2\theta \int_{\Omega_\mathrm{in}} \mathrm{d}^2\theta' \, \mathrm{Cov}(\kappa; \boldsymbol\theta, \boldsymbol\theta') \\ & + \left(\frac{A_\mathrm{in}}{A_\mathrm{out}} \right)^2 \int_{\Omega_\mathrm{out}} \mathrm{d}^2\theta \int_{\Omega_\mathrm{out}} \mathrm{d}^2\theta' \, \mathrm{Cov}(\kappa; \boldsymbol\theta, \boldsymbol\theta') \simeq \frac{cA}{4\rho} \, ,\end{aligned} \tag{32}$$

where, we recall, A is the area used.[1] Obviously, the latter approximate expression holds when $A_\mathrm{in}/A_\mathrm{out} \ll 1$. Curiously, this result does not depend explicitly on the weight function W. The derivation given in Appendix C assumes that the weight function is of the form of Eq. (3), but a similar expression for the variance of M is expected to hold in the more general case.

The results of this subsection can be clarified by a simple example. Instead of introducing the weight function W, we consider the unweighted Eqs. (5) or (6) on small patches of the sky. For simplicity, we refer to a square set Ω of length L divided into s^2 equal square patches: thus we expect $\mathcal{N} = \rho L^2 / s^2$ galaxies per patch. In this case the expected variance of γ is (see Paper I) $c/4\mathcal{N}$, and the variance of κ is of the same order of magnitude. The expected variance of M is then $(c/4\mathcal{N})\bigl(L^2/s^2\bigr)^2 s^2$, where the first factor is the variance of κ in every patch, the second factor is the area of every patch (the square is necessary because we are dealing with variances), and the third factor arises because we must add s^2 independent variables. The final result for the variance of M is $cL^2/4\rho$, exactly as stated by Eq. (32).

Here we may come back to the issue of the Poisson noise, only mentioned at the beginning of Sect. 3. Strictly speaking, the relation $\mathcal{N} = \rho L^2 / s^2$ in the previous paragraph should be replaced by $\langle \mathcal{N} \rangle = \rho L^2 / s^2$, with $\mathcal{N}$ following a Poisson distribution. We now consider the variance of γ as a function of $\mathcal{N}$. An estimate of the effect of the Poisson noise can be obtained in the limit $\langle \mathcal{N} \rangle \gg 1$ by expanding

$$\mathrm{Var}(\gamma, \mathcal{N}) \simeq \frac{c}{4 \langle \mathcal{N} \rangle} \left[1 - \frac{\mathcal{N} - \langle \mathcal{N} \rangle}{\langle \mathcal{N} \rangle} + \left(\frac{\mathcal{N} - \langle \mathcal{N} \rangle}{\langle \mathcal{N} \rangle} \right)^2 \right] . \tag{33}$$

Averaging over the ensemble thus yields

$$\bigl\langle \mathrm{Var}(\gamma, \mathcal{N}) \bigr\rangle \simeq \frac{c}{4 \langle \mathcal{N} \rangle} \left[1 + \frac{1}{\langle \mathcal{N} \rangle} \right] . \tag{34}$$

The effect of the Poisson noise is here contained in the second term in brackets, which is negligibly small. Additional discussion is postponed to the end of Appendix B.

4.2. The general case

The situation is, in principle, quite similar to the weak lensing limit, but, in practice, the calculations are much more difficult. If the angular scale of κ is much greater than the angular scale of W, then we can prove that the mean value of the measured mass distribution given by Eq. (7) holds unchanged.

Difficulties in the calculation of the covariance of κ mainly arise from the form of the covariance of g given by Eq. (25), because of the dependence on $\boldsymbol\theta$ and $\boldsymbol\theta'$ of the first factor. However, if the lens has $\bigl| g_0(\boldsymbol\theta) \bigr| < 1$, the covariance given by Eq. (25) is smaller than that of the weak lensing limit of Eq. (26) and thus we can consider all the results given in the weak lensing regime as upper limits for the errors.

[1] In principle, all integrations should be performed in the whole plane. However, here we suppose to perform integrations over a domain Ω of area A large with respect to σ_W^2.

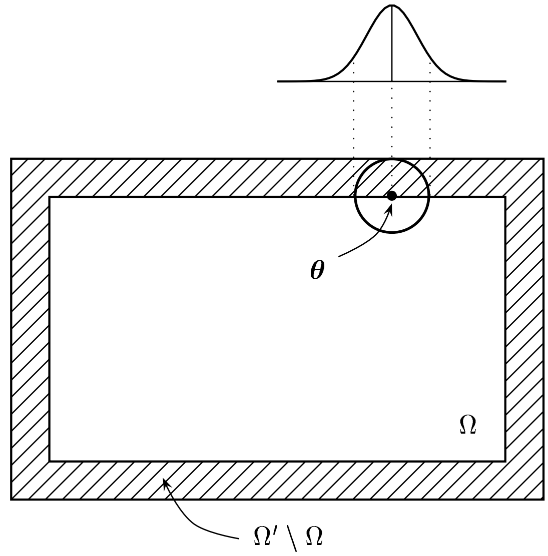

Fig. 1. Sketch of the observation area used in the mass reconstruction.

4.3. Edge effects

Finite boundaries introduce interesting effects, and make the errors depend on the kernel $\boldsymbol G$ used in Eq. (17). For simplicity we take two different sets for $\boldsymbol\theta$ (see Fig. 1). The first set is Ω', i.e. the observation area that includes all the lensed galaxies used in the reconstruction. The second set is $\Omega \subset \Omega'$, i.e. the set where we measure $g(\boldsymbol\theta)$. We suppose that every point in Ω has a neighborhood with radius of the order of the angular scale of W completely enclosed in Ω'. This assumption greatly simplifies calculations and does not have major practical consequences, except that it leads to discarding a small strip $\Omega' \setminus \Omega$ around the boundary $\partial\Omega'$ of Ω'. With this hypothesis the expected measured mass distribution is again given by Eq. (7) as long as $\boldsymbol\theta \in \Omega$. However this is strictly true only if we choose correctly the mean mass distribution $\bar\kappa$ (see Eq. (17)).

In general, the covariance depends on the kernel $\boldsymbol G$ used, and in particular on its divergence-free component (see Appendix C). [We recall that a vector field $\boldsymbol G(\boldsymbol\theta, \boldsymbol\theta')$ can be decomposed as $\boldsymbol G = \boldsymbol G' + \boldsymbol G''$, where $\boldsymbol G'$ has vanishing curl and $\boldsymbol G''$ has vanishing divergence (as usual, in the above notation and decomposition, the emphasis is on the variable $\boldsymbol\theta'$, since $\boldsymbol\theta$ is taken to be fixed).] The result is

$$\mathrm{Cov}(\kappa; \boldsymbol\theta, \boldsymbol\phi) = \mathrm{Cov}(\gamma; \boldsymbol\theta, \boldsymbol\phi)$$

$$+ \int_\Omega \mathrm{d}^2\theta' \int_\Omega \mathrm{d}^2\,\phi' G''_j(\boldsymbol{\theta},\boldsymbol{\theta}')\mathrm{Cov}(u;\boldsymbol{\theta}',\boldsymbol{\phi}')G''_j(\boldsymbol{\phi},\boldsymbol{\phi}')\ , \quad (35)$$

where $\boldsymbol{u}$ is the quantity defined in Eq. (11). The first term is clearly independent of the kernel $\boldsymbol{G}$ used, while the second term can be shown to be positive definite, i.e.

$$\mathrm{Cov}(\kappa;\boldsymbol{\theta},\boldsymbol{\theta}) \geq \mathrm{Cov}(\gamma;\boldsymbol{\theta},\boldsymbol{\theta})\ . \quad (36)$$

Thus the error on κ is minimized if a curl-free kernel $\boldsymbol{G}$ is used. This suggests that only curl-free kernels should be used in weak lensing reconstructions. In fact, the kernels so far judged to be "good" by means of simulations, all have vanishing curl (see Sect. 5). For a curl-free kernel, such as the noise filtering kernel given in Eqs. (18–20), the result is independent of the kernel used and of the set Ω.

We now investigate the class of kernels $\boldsymbol{G}(\boldsymbol{\theta},\boldsymbol{\theta}')$ that satisfy the following properties:

- *i.* $\boldsymbol{G}$ inverts Eq. (11) when $\boldsymbol{u}(\boldsymbol{\theta})$ is measured with no error;
- *ii.* $\boldsymbol{G}$ is curl-free. From the second property we can write

$$\boldsymbol{G}(\boldsymbol{\theta},\boldsymbol{\theta}') = -\nabla_{\boldsymbol{\theta}'}\mathcal{G}(\boldsymbol{\theta},\boldsymbol{\theta}')\ . \quad (37)$$

Thus, if $\boldsymbol{u}(\boldsymbol{\theta}') = \nabla\kappa(\boldsymbol{\theta}')$ we find

$$\begin{aligned}\kappa(\boldsymbol{\theta}) - \bar\kappa &= -\int_\Omega \nabla\mathcal{G}(\boldsymbol{\theta},\boldsymbol{\theta}')\cdot\nabla\kappa(\boldsymbol{\theta}')\,\mathrm{d}^2\theta' \\ &= \int_\Omega \kappa(\boldsymbol{\theta}')\nabla^2\mathcal{G}(\boldsymbol{\theta},\boldsymbol{\theta}')\,\mathrm{d}^2\theta' \\ &\quad - \int_{\partial\Omega} \kappa(\boldsymbol{\theta}')\nabla\mathcal{G}(\boldsymbol{\theta},\boldsymbol{\theta}')\cdot\boldsymbol{n}\,\mathrm{d}\theta'\ . \quad (38)\end{aligned}$$

As in Sect. 2, $\boldsymbol{n}$ is the unit vector orthogonal to $\partial\Omega$. The last relation shows that, in order to satisfy point *i*., we must have

$$\nabla^2\mathcal{G}(\boldsymbol{\theta},\boldsymbol{\theta}') = \delta(\boldsymbol{\theta}-\boldsymbol{\theta}') - \frac{1}{A}\ . \quad (39)$$

If $\kappa(\boldsymbol{\theta})$ is not known on the boundary of Ω, we should also consider:

$$\frac{\partial\mathcal{G}(\boldsymbol{\theta},\boldsymbol{\theta}')}{\partial\theta'_i}n_i = 0 \qquad \forall\boldsymbol{\theta}'\in\partial\Omega\ . \quad (40)$$

In this case the kernel to be used is simply the one obtained from $\mathcal{G} = \mathcal{H}^{\mathrm{SS}}$, i.e. $\boldsymbol{G} = \boldsymbol{H}^{\mathrm{SS}}$ (cfr. Eqs. (18–21)). Otherwise, Eq. (40) is to be dropped and the kernel $\mathcal{G}$ is determined up to a term $\mathcal{L}$, where $\mathcal{L}$ is a harmonic function ($\nabla^2\mathcal{L} = 0$). This free function can be used to dispose of the contribution that would arise from the boundary term in Eq. (38).

As, in general, the measured $\boldsymbol{u}$ field is not curl-free, the inversion can only be approximate. The best inversion can thus be found by searching for the function $\kappa(\boldsymbol{\theta})$ that minimizes the functional

$$\mathcal{S} = \int_\Omega \big\|\nabla\kappa(\boldsymbol{\theta}) - \boldsymbol{u}(\boldsymbol{\theta})\big\|^2\,\mathrm{d}^2\theta\ . \quad (41)$$

If we vary the distribution $\kappa \mapsto \kappa + \delta\kappa$, the functional would in general change to $\mathcal{S} + \delta\mathcal{S}$, with

$$\begin{aligned}\delta\mathcal{S} &= 2\int_\Omega \nabla(\delta\kappa)\cdot(\nabla\kappa - \boldsymbol{u})\,\mathrm{d}^2\theta \\ &= -2\int_\Omega \delta\kappa(\nabla^2\kappa - \nabla\cdot\boldsymbol{u})\,\mathrm{d}^2\theta \\ &\quad + 2\int_{\partial\Omega} \delta\kappa(\nabla\kappa - \boldsymbol{u})\cdot\boldsymbol{n}\,\mathrm{d}\theta\ . \quad (42)\end{aligned}$$

By setting $\delta\mathcal{S} = 0$ we readily find the associated Euler-Lagrange equation

$$\nabla^2\kappa(\boldsymbol{\theta}) = \nabla\cdot\boldsymbol{u}(\boldsymbol{\theta})\ . \quad (43)$$

This equation should be supplemented by

$$\nabla\kappa(\boldsymbol{\theta})\cdot\boldsymbol{n} = \boldsymbol{u}(\boldsymbol{\theta})\cdot\boldsymbol{n} \qquad \forall\boldsymbol{\theta}\in\partial\Omega\ , \quad (44)$$

unless we fix the value of κ on $\partial\Omega$, so that the boundary term in Eq. (42) vanishes because $\delta\kappa = 0$ on $\partial\Omega$. Eqs. (43) and (44) define a Neumann boundary problem equivalent to Eqs. (39) and (40) above, in the sense that $\mathcal{G}$ is precisely the Green function associated with it. This clarifies the interesting properties of $\boldsymbol{H}^{\mathrm{SS}}$.

4.4. Power spectrum

In order to express in a simple manner the errors involved in the reconstruction process, Seitz & Schneider (1996) introduce a "power spectrum" $P(k)$. In the weak lensing limit, their definitions are

$$\Delta(\boldsymbol{k}) = \frac{1}{A}\int_\Omega \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{\theta}}\big(\kappa(\boldsymbol{\theta}) - \kappa_0(\boldsymbol{\theta})\big)\,\mathrm{d}^2\theta\ , \quad (45)$$

$$P(k) = \big\langle\Delta(\boldsymbol{k})\Delta^*(\boldsymbol{k})\big\rangle\ , \quad (46)$$

where the mean in Eq. (46) is also over the various directions of $\boldsymbol{k}$. As a result, the power spectrum $P(k)$ is simply the variance of the complex map $\Delta(\boldsymbol{k})$, i.e. the Fourier transform of the reconstruction error. Thus, for example, the value of $P(0)$ is proportional to $\mathrm{Var}(M)$, the error on the total mass, while its behavior for larger values of k is related to the angular scale of the weight function used.

Within our framework it is not difficult to evaluate the relevant power spectrum. A simple calculation (see Appendix D) for a Gaussian weight function gives

$$P(k) \simeq \frac{c}{4\rho A}\exp\big[-\sigma_W^2 k^2\big]\ , \quad (47)$$

i.e. a simple Gaussian with variance $1/2\sigma_W^2$. One might anticipate a significant contribution to the power spectrum coming from the error associated with the difference between $\langle\kappa\rangle$ and κ_0, but we argue in Appendix D that such contribution is negligible in the weak lensing limit.

5. Comparison with numerical simulations

In this section we compare our predictions with the results obtained by Seitz & Schneider (1996) from numerical simulations. Simulations start by defining a lens mass distribution and a random sample of source galaxies. Each galaxy is traced to the lens plane and the reduced shear g is then calculated from the observed ellipticities using Eq. (5). Finally the shear map is inverted into the lens mass distribution κ using various methods. For Ω Seitz & Schneider take a square $7.5' \times 7.5'$. Source galaxies have random orientations and their ellipticities follow truncated Gaussian distributions. Simulations have been performed with three different variances for $\chi^{\rm s}$: $c_1 = 0.069109$, $c_2 = 0.13323$ and $c_3 = 0.19689$.

The reconstruction method used by Seitz & Schneider is similar to the one described in Sect. 2, with the following differences:

- *i.* Their weight function is not invariant upon translations because it is a Gaussian of argument $\|\boldsymbol{\theta} - \boldsymbol{\theta}'\|$ with the variance depending on $\boldsymbol{\theta}$.
- *ii.* An *outer smoothing* is added to the final lens mass distribution. The first point is a device introduced in order to have better resolution in the stronger parts of the lens. The second point is used in order to have a smooth lens distribution from a discrete map of κ.

Our result (47) for the power spectrum can be easily generalized in order to take into account the outer smoothing:

$$P(k) \simeq \frac{c}{4\rho A} \exp\bigl[-\bigl(\sigma_W^2 + \sigma_s^2\bigr)k^2\bigr] \; . \tag{48}$$

Here σ_s^2 is the variance associated with the outer smoothing. Note however that the expression given above does *not* take into account the variable-scale smoothing used by Seitz & Schneider. Even if this result has been derived with some approximations (weak lensing limit, large area A of Ω, fixed inner smoothing σ_W), a comparison with the simulations shows that Eq. (48) can reproduce the main features of the simulated power spectrum.

Fig. 2 shows the results of simulations together with the power spectrum predicted by Eq. (48). Thus Eq. (48) underestimates the error. This difference can be attributed to the following factors:

- 1. The weight function W considered is not precisely of the form of Eq. (3), because of the change of normalization near $\partial\Omega$. This last factor should increase the variance of κ near the boundary of Ω (the variance of γ, with direct influence on κ, should double near a side of Ω and quadruple near a corner; cf. top-right frame of Fig. 10 in Seitz & Schneider 1996).
- 2. The weight function W is not of the form of Eq. (3) also because of differential smoothing.
- 3. The constant term $\bar{\kappa}$ has not been estimated properly (see first paragraph of Sect. 4.3). In principle, this should be traced to a counterpart in $P(0)$, but for finite sets Ω there is an additional term in $P(k)$.

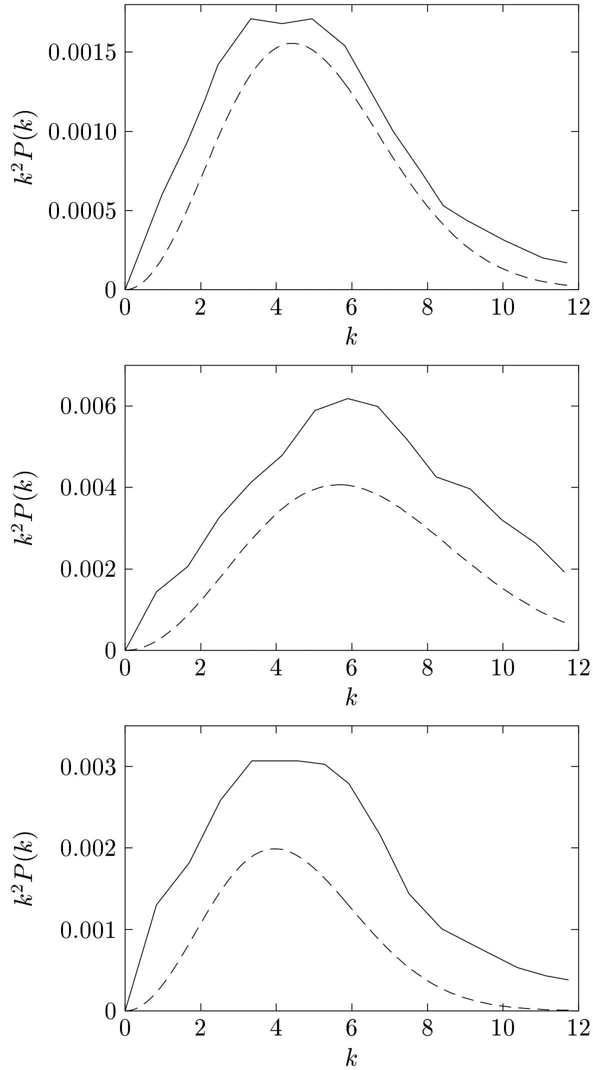

Fig. 2. Comparison of predicted power spectrum (dashed lines) with measured power spectrum (solid lines) for the simulations made by Seitz & Schneider (1996). All frames refer to a source population characterized by $c = 0.069109$. Top frame: $\rho A = 80$, $\sigma_W = 0.212'$, $\sigma_s = 0.0778'$. Middle frame: $\rho A = 50$, $\sigma_W = 0.177'$, $\sigma_s = 0'$. Bottom frame: $\rho A = 50$, $\sigma_W = 0.240'$, $\sigma_s = 0.0778'$.

- 4. The set Ω is not the whole plane.
- 5. The lens is not weak (see Eq. (25) and the extra contribution in Eq. (D4)).
- 6. Poisson noise is associated with the finite number of source galaxies.
- 7. The population of source galaxies is characterized by sizable c.

In spite of these limitations, the general behavior of $P(k)$ is reasonably well reproduced by Eq. (48). In particular, the maximum of the simulated points corresponds exactly to the maximum of our theoretical curve.

5.1. Curl-free kernels

In order to check the result of Eq. (35), we have considered different kernels used by various authors and we have compared our predictions with other aspects of the numerical simulations performed by Seitz & Schneider (1996).

The first kernel considered is the noise-filtered "SS-inversion" (Seitz & Schneider 1996) described above in Eqs. (18–20). This method has been especially designed to reduce the statistical errors and performs very well in simulations. In fact, as stated in Eq. (21), this kernel is curl-free.

Another kernel considered is the "S-inversion" (see Schneider 1995). Simulations show that errors of the S-inversion are nearly the same as for the SS-inversion. The S-inversion operates by averaging over radial paths made of two segments. Inside the segments the kernel is easily shown to be curl-free.

The last kernel is the so called "B-inversion" (Bartelmann 1995; see also Squires & Kaiser 1996). From the results of the simulations it is clear that the B-inversion leads to larger errors on the map distribution. This behavior is once again explained by Eq. (35), since the B-inversion kernel is *not* curl-free. Notice that in this case it is difficult to estimate analytically the exact error on κ.

Acknowledgements. We would like to thank Peter Schneider for several comments and suggestions that have helped us improve the paper. This work has been partially supported by MURST and by ASI of Italy.

Appendix A: notation

We collect here the main symbols used in this paper.

- X^{s} Superscript s identifies *source* (unlensed) quantities.
- X_i Subscripts refer to the complex representation, e.g. $\chi = \chi_1 + \mathrm{i}\chi_2$ or, when indicated, to the vector representation, e.g. $\boldsymbol{\theta} = (\theta_1, \theta_2)$.
- $\boldsymbol{\theta}^{\mathrm{s}}$, $\boldsymbol{\theta}$ Unlensed and observed position of a point source.
- Σ_{c} Critical density.
- $\Sigma(\boldsymbol{\theta})$ Projected mass distribution of the lens.
- $\kappa(\boldsymbol{\theta})$ Dimensionless mass distribution: $\kappa(\boldsymbol{\theta}) = \Sigma(\boldsymbol{\theta}/\Sigma_{\mathrm{c}}$.
- Q^{s}, Q Unlensed and observed quadrupole moment of an extended image.
- χ^{s}, χ Unlensed and observed (complex) ellipticity of an extended image: $\chi^{\mathrm{s}} = \bigl(Q^{\mathrm{s}}_{11} - Q^{\mathrm{s}}_{22} + 2\mathrm{i}Q^{\mathrm{s}}_{12}\bigr) / \bigl(Q^{\mathrm{s}}_{11} + Q^{\mathrm{s}}_{22}\bigr)$, and similarly for χ.
- γ Complex shear.
- g Complex reduced shear: $g = \gamma/(1-\kappa)$.
- c Covariance of the source ellipticity for an isotropic distribution: $\bigl\langle \chi^{\mathrm{s}}_i \chi^{\mathrm{s}}_j \bigr\rangle = c\delta_{ij}$.
- Id, δ_{ij} Identity matrix.

Appendix B: the shear

In this Appendix we derive Eq. (23) and Eq. (25) of Sect. 3. We assume that the reduced shear g has been measured through Eq. (5). Calculations based on Eq. (6) are very similar. As explained in Appendix A of Paper I, the mean value of g obeys the relation

$$\sum_{n=1}^{N} W\bigl(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)}\bigr) \chi^{\mathrm{s}} \left(\bigl\langle \chi^{(n)} \bigr\rangle, \bigl\langle g(\boldsymbol{\theta}) \bigr\rangle \right) = 0 \; . \tag{B1}$$

Here $\bigl\langle \chi^{(n)} \bigr\rangle$ is the mean value of χ in $\boldsymbol{\theta}^{(n)}$ and thus depends on $g_0\bigl(\boldsymbol{\theta}^{(n)}\bigr)$. We now assume that $g_0(\boldsymbol{\theta})$ does not change significantly on the angular scale of $W(\boldsymbol{\theta}, \boldsymbol{\theta}')$. This implies that we can expand the previous equation to first order in g_0. We choose, as starting point, the value $g_0 = g_*$ given by

$$g_*(\boldsymbol{\theta}) = \sum_{n=1}^{N} W\bigl(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)}\bigr) g_0\bigl(\boldsymbol{\theta}^{(n)}\bigr) \; . \tag{B2}$$

Then we find easily

$$\sum_{n=1}^{N} W \chi^{\mathrm{s}} \bigl(\chi_*, \bigl\langle g(\boldsymbol{\theta}) \bigr\rangle \bigr) + \sum_{n=1}^{N} W \left. \frac{\partial \chi^{\mathrm{s}}}{\partial \chi} \right|_{\chi = \chi_*} \frac{\partial \chi_*}{\partial g_*} \Bigl(g_0\bigl(\boldsymbol{\theta}^{(n)}\bigr) - g_* \Bigr) = 0 \; , \tag{B3}$$

where χ_* is the expected value of the ellipticity when the reduced shear is equal to g_*. Eq. (B3) has the obvious solution $\bigl\langle g(\boldsymbol{\theta}) \bigr\rangle = g_*(\boldsymbol{\theta})$, because the first term, $\chi^{\mathrm{s}}(\chi_*, g_*)$, vanishes by definition and the second vanishes due to the choice of g_* (notice that the partial derivatives in the latter do not depend on $\boldsymbol{\theta}^{(n)}$). When moving to a continuous description, we have to calculate the average expected value of g over all possible positions $\bigl\{ \boldsymbol{\theta}^{(n)} \bigr\}$ of the source galaxies

$$\bigl\langle g(\boldsymbol{\theta}) \bigr\rangle = \frac{1}{A^N} \int_\Omega \mathrm{d}^2\theta^{(1)} \int_\Omega \mathrm{d}^2\theta^{(2)} \cdots \int_\Omega \mathrm{d}^2\theta^{(N)} \times \sum_{n=1}^{N} W\bigl(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)}\bigr) g_0\bigl(\boldsymbol{\theta}^{(n)}\bigr) \; . \tag{B4}$$

The weight function $W\bigl(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)}\bigr)$ depends on all $\bigl\{ \boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(N)} \bigr\}$ because of the adopted normalization. However, in the limit $N \gg 1$, the weight function $W\bigl(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)}\bigr)$ can be considered to depend only on $\boldsymbol{\theta}^{(n)}$ alone, so that the above integral can be approximated by

$$\bigl\langle g(\boldsymbol{\theta}) \bigr\rangle \simeq \frac{1}{A} \sum_{n=1}^{N} \int_\Omega \mathrm{d}^2\theta^{(n)} \, W\bigl(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)}\bigr) g_0\bigl(\boldsymbol{\theta}^{(n)}\bigr) = \rho \int_\Omega \mathrm{d}^2\theta' \, W(\boldsymbol{\theta}, \boldsymbol{\theta}') g_0(\boldsymbol{\theta}') \; . \tag{B5}$$

This proves Eq. (23). Our result simply states that the use of the first order expansion in g reduces every mean to a weighted arithmetic mean.

Calculations for the covariance of g are much more difficult but basically repeat those given for the unweighted situation in App. A.1 of Paper I. In particular, if we call $F\big(\{\boldsymbol{\theta}^{(n)},\chi^{(n)}\},g(\boldsymbol{\theta})\big)$ the function defined in the l.h.s. of Eq. (5), we have

$$\mathrm{Cov}(g;\boldsymbol{\theta},\boldsymbol{\theta}') = B^{-1}(\boldsymbol{\theta})$$

$$\left[\sum_{n=1}^{N} A^{(n)}(\boldsymbol{\theta})\mathrm{Cov}\big(\chi^{(n)}\big)A^{(n)T}(\boldsymbol{\theta}')\right]\big(B^{-1}(\boldsymbol{\theta}')\big)^T \; , \tag{B6}$$

where

$$B(\boldsymbol{\theta}) \;=\; \frac{\partial F}{\partial g(\boldsymbol{\theta})} \tag{B7}$$

$$A^{(n)}(\boldsymbol{\theta}) \;=\; \frac{\partial F}{\partial \chi^{(n)}} \; . \tag{B8}$$

All functions have to be calculated in the mean value of their arguments. Some calculations then lead to

$$B(\boldsymbol{\theta}) \simeq \frac{2}{1-\big|\langle g(\boldsymbol{\theta})\rangle\big|^2}\,\mathrm{Id} \; . \tag{B9}$$

The term in brackets in Eq. (B6) can be written in the form

$$\sum_{n=1}^{N} A^{(n)}(\boldsymbol{\theta})\mathrm{Cov}\big(\chi^{(n)}\big)A^{(n)T}(\boldsymbol{\theta}')$$

$$= c\sum_{n=1}^{N} W\big(\boldsymbol{\theta},\boldsymbol{\theta}^{(n)}\big)W\big(\boldsymbol{\theta}',\boldsymbol{\theta}^{(n)}\big) + \text{linear terms} \; . \tag{B10}$$

Here "linear terms" means additional terms linear with respect to the quantity $\big[g_0\big(\boldsymbol{\theta}^{(n)}\big) - g_*\big]$, based on the same expansion defined by (B2).

By averaging over the source positions and by moving to a continuous description (basically following the same steps indicated in (B4) and (B5)), we thus obtain Eq. (25). Notice that the "linear terms" in Eq. (B10) do not give any contribution when averaged over the source positions. We stress that the results stated here are valid only if the weight function is even (property 1 of Sect. 2.1).

Finally, we point out that the approximation that takes us from (B4) to (B5) is precisely associated with neglecting the Poisson noise.

Appendix C: the lens mass distribution

In this Appendix we will derive Eq. (7) and the results stated in Sect. 4, assuming a weight function invariant upon translations (case 2 of Sect. 2.1).

C.1. Weak lensing limit

Calculations in the weak lensing limit are not difficult. As explained in Sect. 2.2, we can use either Eq. (8) or Eqs. (11) and (12) to convert the reduced shear into the mass distribution.

In the case of Eq. (8) we can write

$$\langle\kappa\rangle = D_i \star \langle\gamma_i\rangle = D_i \star (\rho W \star \gamma_{0i}) \; . \tag{C1}$$

Here the star denotes convolution, while γ_{0i} is the component i of the true shear map γ_0. The second step is justified by Eq. (23) applied in the weak lensing limit. By using the associative and commutative properties of the convolution and by noting that $\kappa_0 = D_i \star \gamma_{0i}$, we can write

$$\begin{aligned}\langle\kappa\rangle \;&=\; D_i \star \rho W \star \gamma_{0i} \\ &=\; \rho W \star D_i \star \gamma_{0i} = \rho W \star \kappa_0 \; .\end{aligned} \tag{C2}$$

This proves Eq. (7).

About the covariance of κ we can write

$$\begin{aligned}\mathrm{Cov}(\kappa;\boldsymbol{\theta},\boldsymbol{\phi}) = \int \mathrm{d}^2\theta' \int \mathrm{d}^2\phi'\, D_i(\boldsymbol{\theta}-\boldsymbol{\theta}') \times \\ \times\, \mathrm{Cov}(\gamma;\boldsymbol{\theta}'-\boldsymbol{\phi}')D_i(\boldsymbol{\phi}-\boldsymbol{\phi}') \; .\end{aligned} \tag{C3}$$

As D_i and $\mathrm{Cov}(\gamma)$ are even functions, this is simply a double convolution, and thus the result depends only on the difference between $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$. Therefore we can write for $\mathrm{Cov}(\kappa;\boldsymbol{\theta},\boldsymbol{\phi}) = \mathrm{Cov}(\kappa;\boldsymbol{\theta}-\boldsymbol{\phi})$ the expression

$$\mathrm{Cov}(\kappa) = D_i \star \mathrm{Cov}(\gamma) \star D_i = \mathrm{Cov}(\gamma) \; , \tag{C4}$$

i.e. Eq. (31). Here we have used again the commutative property of convolutions and the relation $D_i \star D_i = \delta$ given by Eqs. (8) and (10). [Hereafter δ means the Dirac delta distribution.]

The results are the same if we use Eqs. (11) and (12). In fact we can write Eq. (11) as a convolution between γ_i and the operator

$$T_{ij}(\boldsymbol{\theta}) = \begin{pmatrix} \delta_{,1}(\boldsymbol{\theta}) & \delta_{,2}(\boldsymbol{\theta}) \\ -\delta_{,2}(\boldsymbol{\theta}) & \delta_{,1}(\boldsymbol{\theta}) \end{pmatrix} \; , \tag{C5}$$

where $\delta_{,i}(\boldsymbol{\theta}) = \partial\delta(\boldsymbol{\theta})/\partial\theta_i$. Thus we are allowed to use the properties of convolutions. It is obvious then that the convolution with the weight function W in $\langle\gamma\rangle$ can be moved to the true lens distribution κ_0, and we find again the result of Eq. (C2). The covariance matrix of $\boldsymbol{u}$ can be calculated using the operator (C5). As a result, we find

$$\mathrm{Cov}(u) = -\nabla^2\mathrm{Cov}(\gamma) \; . \tag{C6}$$

We then have

$$\begin{aligned}\mathrm{Cov}(\kappa) \;&=\; -H_i^{\mathrm{KS}} \star \nabla^2\mathrm{Cov}(\gamma) \star H_i^{\mathrm{KS}} \\ &=\; -\nabla^2(H_i^{\mathrm{KS}} \star H_i^{\mathrm{KS}}) \star \mathrm{Cov}(\gamma) \; .\end{aligned} \tag{C7}$$

A simple calculation shows that $-\nabla^2(H_i^{\mathrm{KS}} \star H_i^{\mathrm{KS}})(\boldsymbol{\theta}) = \nabla^2\big(\ln\|\boldsymbol{\theta}\|/2\pi\big) = \delta(\boldsymbol{\theta})$. This leads again to Eq. (31).

Now let us prove Eq. (32). From Eq. (C4) and Eq. (26) we find

$$\mathrm{Cov}(\kappa) = \frac{c\rho}{4} W \star W \; . \tag{C8}$$

In the limit $A_{\mathrm{in}}/A_{\mathrm{out}} \ll 1$ explained in Sect. 4.1, the main contribution to the variance of M derives from a double integration of $\mathrm{Cov}(\kappa)$. A simple change of variables gives

$$\begin{aligned}\mathrm{Var}(M) \;&=\; \frac{c\rho}{4}\int \mathrm{d}^2\theta \int \mathrm{d}^2\theta'\,(W \star W)(\boldsymbol{\theta}') \\ &=\; \frac{c\rho A}{4}\hat{W}(\mathbf{0})\hat{W}(\mathbf{0}) = \frac{cA}{4\rho} \; .\end{aligned} \tag{C9}$$

Here $\hat{W}$ is the Fourier transform of W and the last equality holds because of the normalization (22) of the weight function.

C.2. The general case

In the general case we restrict ourselves to estimating the mean value of the lens distribution because calculations for the covariance are too difficult. Under the hypothesis that the angular scale of W is much smaller than the angular scale of κ (or g), the situation is much like that of the weak lensing limit. As shown in Appendix B, this basically implies that all averages are weighted arithmetic averages. Simple calculations show that we have $\langle \tilde{\boldsymbol{u}} \rangle = \tilde{\boldsymbol{u}}_0 \star \rho W$, and hence $\langle \tilde{\kappa} \rangle = \tilde{\kappa}_0 \star \rho W$. As usual the assumed ordering of scale lengths leads again to Eq. (C2).

C.3. Edge effects

For simplicity we refer to $\bar{\kappa} = 0$. We rewrite Eqs (17), (11) and (10) with a different notation

$$\pi_\Omega \kappa \;=\; G_i u_i \;, \tag{C10}$$

$$u_i \;=\; T_{ij} \star \gamma_j \;, \tag{C11}$$

$$\gamma_i \;=\; D_i \star \kappa \;. \tag{C12}$$

Here π_Ω is the characteristic operator for the set Ω:

$$(\pi_\Omega f)(\boldsymbol{\theta}) = \begin{cases} f(\boldsymbol{\theta}) & \text{if } \boldsymbol{\theta} \in \Omega \;, \\ 0 & \text{otherwise} \;. \end{cases} \tag{C13}$$

Eq. (C10) is equivalent to Eq. (17) with $\bar{\kappa} = 0$ if we redefine the kernel $\boldsymbol{G}(\boldsymbol{\theta}, \boldsymbol{\theta}')$ for every $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ so that $\boldsymbol{G}(\boldsymbol{\theta}, \boldsymbol{\theta}') = 0$ if either $\boldsymbol{\theta} \notin \Omega$ or $\boldsymbol{\theta}' \notin \Omega$. With this simple definition we can extend the integration domain (usually Ω) to the whole plane. Notice that while T_{ij} and D_i are used in convolutions, G_i is a generic linear operator. From these equations we have

$$\pi_\Omega \kappa = G_i u_i = G_i (T_{ij} \star D_j \star \kappa) \;, \tag{C14}$$

and thus we find the identity

$$G_i T_{ij} \star D_j = \pi_\Omega \;. \tag{C15}$$

Eq. (12) with the new notation is

$$\kappa = H^{\rm KS}_i \star u_i \;. \tag{C16}$$

This, together with Eqs. (C11) and (C12), gives us another identity:

$$u_i = T_{ij} \star D_j \star H^{\rm KS}_k \star u_k \;. \tag{C17}$$

Using Eqs. (C10), (C12), and the relation $\langle u_i \rangle = \rho W \star u_{0i}$, we can easily obtain the mean value for measures of κ:

$$\begin{aligned} \langle \kappa \rangle \;&=\; G_i \langle u_i \rangle = G_i \rho W \star u_{0i} \\ &=\; G_i T_{ij} \star D_j \star H^{\rm KS}_k \star u_{0k} \star \rho W \\ &=\; \pi_\Omega H^{\rm KS}_k \star u_{0k} \star \rho W = \pi_\Omega \rho W \star \kappa_0 \;. \end{aligned} \tag{C18}$$

As usual, subscript 0 indicates the true value of a quantity. This equation, rewritten in the more standard notation, is Eq. (7) for $\boldsymbol{\theta} \in \Omega$.

Let us calculate the covariance of κ. First of all note that, while Eq. (C14) implies Eq. (C15), from Eq. (C17) we cannot deduce that $T_{ij} \star D_j \star H^{\rm KS}_k$ is the identity. This happens because the two components of u_i are not functionally independent, as one can see from the relation $\nabla \wedge \boldsymbol{u} = 0$. In fact, using Fourier transforms it is easy to prove that the operator $R_{ik} = T_{ij} \star D_j \star H^{\rm KS}_k$ selects the curl-free component of a vector field. Its Fourier transform is

$$\hat{R}_{ij}(\boldsymbol{k}) = \frac{k_i k_j}{\|\boldsymbol{k}\|^2} \;. \tag{C19}$$

From (C10) we have

$$\mathrm{Cov}(\kappa) = G_j \mathrm{Cov}(u) G_j \;. \tag{C20}$$

Every vector field can be written as the sum of two vector fields, of which one is curl-free and the other is divergence-free. Hence, if we consider $\boldsymbol{G}(\boldsymbol{\theta}, \boldsymbol{\theta}')$ a vector field with respect to $\boldsymbol{\theta}'$, we can write

$$\boldsymbol{G}(\boldsymbol{\theta}, \boldsymbol{\theta}') = \boldsymbol{G}'(\boldsymbol{\theta}, \boldsymbol{\theta}') + \boldsymbol{G}''(\boldsymbol{\theta}, \boldsymbol{\theta}') \;, \tag{C21}$$

where

$$\begin{aligned} \nabla_{\boldsymbol{\theta}'} \wedge \boldsymbol{G}'(\boldsymbol{\theta}, \boldsymbol{\theta}') \;&=\; 0 \;, \\ \nabla_{\boldsymbol{\theta}'} \cdot \boldsymbol{G}''(\boldsymbol{\theta}, \boldsymbol{\theta}') \;&=\; 0 \;. \end{aligned} \tag{C22}$$

Thus $\boldsymbol{G}'$ and $\boldsymbol{G}''$ can be written as the gradient and the "curl" of two scalar fields:

$$\boldsymbol{G}'(\boldsymbol{\theta}, \boldsymbol{\theta}') \;=\; \nabla_{\boldsymbol{\theta}'} s'(\boldsymbol{\theta}, \boldsymbol{\theta}') \;, \tag{C23}$$

$$\boldsymbol{G}''(\boldsymbol{\theta}, \boldsymbol{\theta}') \;=\; \nabla_{\boldsymbol{\theta}'} \wedge s''(\boldsymbol{\theta}, \boldsymbol{\theta}') = \begin{pmatrix} s''_{,2}(\boldsymbol{\theta}, \boldsymbol{\theta}') \\ -s''_{,1}(\boldsymbol{\theta}, \boldsymbol{\theta}') \end{pmatrix} \;. \tag{C24}$$

There is some freedom in the choice of $\boldsymbol{G}'$ and $\boldsymbol{G}''$ (or equivalently of s' and s''). However, it is always possible to choose $\boldsymbol{G}'$ and $\boldsymbol{G}''$ so that they vanish for $\|\boldsymbol{\theta}'\| \to \infty$. With the decomposition (C21) we have $G_i R_{ij} = G'_j$ and thus

$$\mathrm{Cov}(\kappa) = (G_i R_{ij} + G''_j) \mathrm{Cov}(u) (G_k R_{kj} + G''_j) \tag{C25}$$

Recalling now the definition of R_{ij} and using Eq. (C15) we find

$$\mathrm{Cov}(\kappa) = (\pi_\Omega H^{\rm KS}_j + G''_j) \mathrm{Cov}(u) (\pi_\Omega H^{\rm KS}_j + G''_j) \tag{C26}$$

If the kernel $\boldsymbol{G}$ has vanishing curl, then $\boldsymbol{G}'' = 0$ and we find the final result

$$\begin{aligned} \mathrm{Cov}(\kappa; \boldsymbol{\theta}, \boldsymbol{\phi}) = \int \mathrm{d}^2\theta' \int \mathrm{d}^2\phi' \, H^{\rm KS}_j(\boldsymbol{\theta} - \boldsymbol{\theta}') \times \\ \times \mathrm{Cov}(u; \boldsymbol{\theta}' - \boldsymbol{\phi}') H^{\rm KS}_j(\boldsymbol{\phi} - \boldsymbol{\phi}') \qquad \text{for } \boldsymbol{\theta}, \boldsymbol{\phi} \in \Omega \;. \end{aligned} \tag{C27}$$

In general however we must evaluate three additional terms. Two of them are of the form

$$\begin{aligned} G''_j \mathrm{Cov}(u) G'_j \;&=\; \int \mathrm{d}^2\theta' \int \mathrm{d}^2\phi' \, \boldsymbol{G}''(\boldsymbol{\theta}, \boldsymbol{\theta}') \times \\ &\times\; \mathrm{Cov}(u; \boldsymbol{\theta}' - \boldsymbol{\phi}') \cdot \nabla_{\boldsymbol{\phi}'} s'(\boldsymbol{\phi}, \boldsymbol{\phi}') \;, \end{aligned} \tag{C28}$$

where we have used Eq. (C24). By the change of variable $\boldsymbol{\phi}' \mapsto \boldsymbol{\theta}' - \boldsymbol{\phi}'$ and after integrating by parts we find

$$\begin{aligned} G''_j \mathrm{Cov}(u) G'_j \;&=\; \int \mathrm{d}^2\theta' \int \mathrm{d}^2\phi' \, \nabla_{\boldsymbol{\theta}'} \cdot \boldsymbol{G}''(\boldsymbol{\theta}, \boldsymbol{\theta}') \\ &\times\; \mathrm{Cov}(u; \boldsymbol{\theta}' - \boldsymbol{\phi}') s''(\boldsymbol{\phi} - \boldsymbol{\phi}') = 0 \;, \end{aligned} \tag{C29}$$

where the last relation holds in virtue of Eq. (C23). We finally rewrite Eq. (C27) in a simplified form:

$$\mathrm{Cov}(\kappa) = \pi_\Omega H_j^{\mathrm{KS}} \mathrm{Cov}(u) \pi_\Omega H_j^{\mathrm{KS}} + G_j'' \mathrm{Cov}(u) G_j'' \qquad \text{(C30)}$$

Here the first term is independent of the specific kernel $\boldsymbol{G}$ used, while the second term depends only on $\boldsymbol{G}''$. As, by definition, $\mathrm{Cov}(u)$ is positive definite, the last term in Eq. (C31) is also positive definite. In other words if $\boldsymbol{G}'' \neq 0$ the error on κ will increase.

Appendix D: power spectrum

The power spectrum reported in Eq. (47) can be deduced from the expression of the mean and covariance of the measured mass distribution. In particular we have

$$\bigl\langle \Delta_i(\boldsymbol{k}) \bigr\rangle = \frac{1}{A} \int_\Omega \Bigl(\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\theta}} \Bigr)_i \bigl(\bigl\langle \kappa(\boldsymbol{\theta}) \bigr\rangle - \kappa_0(\boldsymbol{\theta}) \bigr) \, \mathrm{d}^2\theta \; , \qquad \text{(D1)}$$

$$\begin{aligned} \mathrm{Cov}_{ij}(\Delta; \boldsymbol{k}, \boldsymbol{k}') \; = \; & \frac{1}{A^2} \int_\Omega \mathrm{d}^2\theta \int_\Omega \mathrm{d}^2\theta' \left(\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\theta}} \right)_i \\ & \times \left(\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{\theta}'} \right)_j \mathrm{Cov}(\kappa; \boldsymbol{\theta}, \boldsymbol{\theta}') \; . \qquad \text{(D2)} \end{aligned}$$

The subscripts in the exponentials denote real ($i, j = 1$) and imaginary ($i, j = 2$) parts. The power spectrum $P(k)$ is directly related to the covariance of Δ. In fact, we have

$$P(k) = \mathrm{Cov}_{ii}(\Delta; \boldsymbol{k}, \boldsymbol{k}) + \bigl\langle \Delta_i(\boldsymbol{k}) \bigr\rangle \bigl\langle \Delta_i(\boldsymbol{k}) \bigr\rangle \; , \qquad \text{(D3)}$$

with summation implied on i and mean over all directions of $\boldsymbol{k}$. For a large set Ω we can perform integrations over the whole plane. Thus we find

$$P(k) = \frac{1}{A} \widehat{\mathrm{Cov}}(\kappa; \boldsymbol{k}) + \frac{1}{A^2} \bigl| \hat\kappa_0(\boldsymbol{k}) \bigr|^2 \bigl| \rho \hat{W}(\boldsymbol{k}) - 1 \bigr|^2 \; . \qquad \text{(D4)}$$

Here, as usual, hats indicate Fourier transform. In the weak lensing limit the second term of this expression can be dropped. Hence, if W is a Gaussian of the form of Eq. (3) we find Eq. (47). [Note that no averaging over the direction of $\boldsymbol{k}$ is needed in the weak lensing limit if the weight function has the form (3).]

References

Bartelmann M., 1995, A&A 303, 643
Bartelmann M., Narayan R., Seitz S., Schneider P., 1996, ApJ 464, L115
Fahlman G.G., Kaiser N., Squires G., Woods D., 1994, ApJ 437, 56
Falco E.E., Gorenstein M.V., Shapiro I.I., 1985, ApJ 289, L1
Kaiser N., 1995, ApJ 493, L1
Kaiser N. & Squires J., 1993, ApJ 404, 441
Kneib J.P, Ellis R.S., Smail I., Couch W.J., Sharples R.M., 1996, ApJ 417, 643
Lombardi M., Bertin G., 1998, A&A 330, 791 (Paper I)
Lucy L.B, 1994, A&A 289, 983
Schneider P., 1996, A&A 302, 639
Schneider P., Ehlers J., Falco E.E., 1992, Gravitational Lenses, Springer, Heidelberg
Schneider P., Seitz C., 1995, A&A 294, 411
Seitz C., Schneider P., 1995, A&A 297, 287
Seitz S., Schneider P., 1996, A&A 305, 383
Smail I., Ellis R.S., Fitchett M.J., Edge A.C., 1994, MNRAS 237, 257
Squires G., Kaiser N., 1996, ApJ 473, 65
Tyson J.A., Valdes F., Wenk R.A., 1990, ApJ 349, L1
Webster R.L., 1985, MNRAS 213, 871