

Steps towards nonlinear cluster inversion through gravitational distortions

III. Including a redshift distribution of the sources

Carolín Seitz and Peter Schneider

Max-Planck-Institut für Astrophysik, Postfach 1523, D-85740 Garching, Germany

Received 16 January 1996 / Accepted 15 June 1996

Abstract. In a series of previous papers we have considered the reconstruction of the surface mass density of a cluster of galaxies from images of lensed faint background galaxies. We showed that the reconstructed surface mass density is not uniquely determined, but that there exists a global invariance transformation that leaves the shape of the images of the lensed galaxies unchanged. Because of this, only lower limits on the total mass of a cluster can be derived if no further information besides image ellipticities are used. Throughout these papers we used the simplifying assumption that all sources are at the same redshift.

In this paper we account for a redshift distribution of the faint galaxies, and in particular, some of these galaxies can lie in front of the cluster or can be cluster members. We show how the mass distribution of a cluster of galaxies can be obtained from images of these faint galaxies, if the redshift distribution of these galaxies is known. We demonstrate that for the reconstruction of non-critical clusters we need less information on the redshift distribution of the galaxies, i.e., we only need to know two or three moments of the distribution.

We show that the mean mass density across the data field is still a free variable, i.e., there remains a global invariance transformation of the resulting mass density field. For non-critical clusters we can derive the transformation explicitly and it is similar to that derived previously for the case of a single redshift of the sources. We discuss several theoretical ideas to break the mass degeneracy; of those considered, we find that only the magnification effect on the number density of galaxy images can be used successfully in practice.

Key words: gravitational lensing – dark matter – cosmology: observations – galaxies: clusters of

1. Introduction

This is the third and final paper in a series in which we have considered the reconstruction of the surface mass density of galaxy

Send offprint requests to: P. Schneider

clusters from the (weak) image distortion it imposes onto faint background galaxies. Inspired by the pioneering work of Kaiser & Squires (1993, hereafter KS), who derived a parameter-free inversion equation for the surface mass density of the cluster in terms of the tidal deflection field [which has been discovered earlier by Tyson, Valdes & Wenk (1990); earlier work on cluster mass determination by weak lensing effects include Kochanek (1990) and Miralda-Escudé (1991)] we have started to generalize the KS inversion method to include also strong clusters, i.e., cluster which are – at least nearly – capable to produce giant arcs. In Paper I (Schneider & Seitz 1995) we have analyzed the basic observable from image distortions and pointed out a global invariance transformation of the surface mass density which leaves the observable distortion invariant. In Paper II (Seitz & Schneider 1995) we have then developed an iterative procedure to reconstruct the density field, which was then applied to synthetic data and shown to work well.

In the present paper, we want to generalize the treatment of Paper II in two different ways. First, the inversion procedure constructed in Paper II is not unbiased and contains boundary artefacts in the same way as the original KS method, due to the fact that the observations are always limited to a finite data field. There are now several different inversion methods which are unbiased (Schneider 1995; Kaiser et al. 1995; Bartelmann 1995, Seitz & Schneider 1996) – all of them are based on a relation found by Kaiser (1995). As was demonstrated in Seitz & Schneider (1996), these finite-field methods work well and can be applied efficiently. Second, in the earlier papers cited above it was assumed that all sources have the same effective redshift; by that we mean that all sources have about the same ratio $D(z_d, z)/D(z)$ of the angular diameter distance as measured from the lens and from the observer. This assumption is fairly well justified if the cluster redshift is relatively small, $z_d \lesssim 0.2$, say. For weak cluster lenses, this assumption can be easily dropped since then the (linear) reconstruction proceeds equivalently to the case that all sources are at about the ‘mean redshift’ of the population (this will be made more precise in Sect. 4.3 below). However, if the cluster is not assumed to be weak, the consider-

ation of the source redshift distribution becomes more difficult. We shall discuss this problem in some detail below.

The rest of the paper is organized as follows: in Sect. 2 we present basic equations and our notation. In Sect. 3 we derive the dependence of the local observables on the local values of the surface mass density and the shear. In Sect. 4, we describe the nonlinear reconstruction method; after briefly reviewing the unbiased finite-field inversion method developed by Seitz & Schneider (1996) for the case of a single source redshift, we generalize this technique to the case of a redshift distribution. As in the case of a single source redshift, we also have a global invariance transformation, which, however, cannot be written in closed form in general. However, restricting the consideration to moderately strong clusters, we construct the invariance transformation explicitly. As we show in Sect. 5 and in Appendix 3, the presence of a redshift distribution of the sources provides several methods to break the invariance transformation.

However, from simulations we found that among the methods considered, only that is successful in practice which makes use of the magnification effect on the number density of galaxy images (see also Broadhurst, Taylor & Peacock 1995). In contrast to Broadhurst, Taylor & Peacock (1995) and Broadhurst (1995) we do not use the local magnification to derive the local mass density, but we use the magnification averaged over the field to constrain the mean mass density in the field (in order to combine information from the shear with local magnification information, a maximum likelihood approach seems to be the best strategy – see Bartelmann et al. 1996). Essential for the success of this method is that the source counts deviate sufficiently strongly from the $n(S) \propto S^{-2}$ behaviour, which seems to be the case for galaxies with red colour. Furthermore, the method depends strongly on the assumption that the faint background galaxies are distributed rather smoothly, i.e., that no strong correlations in their angular position is present, as seems to be justified by recent investigations (e.g., Infante & Pritchett 1995).

We apply our methods to synthetic data in Sect. 6 to demonstrate its feasibility. The application to an HST exposure of the cluster Cl0939+4713 will be published elsewhere (Seitz et al. 1995c). We discuss our results in Sect. 7.

2. Basic equations

The distortion of images of background galaxies depends on the dimensionless surface mass density of the lens, which is the physical surface mass density $\Sigma(\boldsymbol{\theta})$ divided by the critical surface mass density Σ_{crit} . If we now consider the background sources to be distributed in redshift, then the critical surface mass density depends on the redshift z of the source:

$$\Sigma_{\text{crit}}(z) = \begin{cases} \infty & \text{for } z \leq z_d, \\ \frac{c^2 D(z)}{4\pi G D(z_d) D(z, z_d)} & \text{for } z > z_d. \end{cases} \quad (2.1)$$

Here $D(z_d)$ and $D(z)$ are the angular diameter-distances from the observer to the lens at redshift z_d and to the source at redshift

z , and $D(z_d, z)$ is the angular diameter-distance from the lens to the source. Defining

$$w(z; z_d) = \frac{\lim_{z \rightarrow \infty} \Sigma_{\text{crit}}(z)}{\Sigma_{\text{crit}}(z)} = \frac{\Sigma_{\text{crit}\infty}}{\Sigma_{\text{crit}}(z)} \quad (2.2)$$

we obtain for the dimensionless surface mass density $\kappa(\boldsymbol{\theta}, z)$ at angular position $\boldsymbol{\theta}$ for a source at redshift z

$$\kappa(\boldsymbol{\theta}, z) = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_{\text{crit}}(z)} = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_{\text{crit}\infty}} \frac{\Sigma_{\text{crit}\infty}}{\Sigma_{\text{crit}}(z)} \equiv \kappa(\boldsymbol{\theta}) w(z; z_d). \quad (2.3)$$

The function $w(z; z_d)$ relates the ‘lensing strength’ for a source with redshift z to that of a hypothetical source at ‘infinite redshift’, and its form depends on the geometry of the universe. For an Einstein–de Sitter universe we have

$$w(z; z_d) = \begin{cases} 0 & \text{for } z \leq z_d \\ \frac{\sqrt{1+z} - \sqrt{1+z_d}}{\sqrt{1+z} - 1} & \text{for } z > z_d. \end{cases} \quad (2.4)$$

In particular, for sources with redshift smaller than that of the lens, the ‘lensing strength’ vanishes. For the rest of this paper, we consider a single cluster lens at redshift z_d , and drop the second argument of w , i.e., $w(z; z_d) \equiv w(z)$.

Since the shear γ is related linearly to the surface mass density, its dependence on source redshift is the same as for κ , so the $\gamma(\boldsymbol{\theta}, z) \equiv w(z) \gamma(\boldsymbol{\theta})$. The magnification $\mu(\boldsymbol{\theta}, z)$ of an image with position $\boldsymbol{\theta}$ and a source redshift z then becomes

$$\begin{aligned} \mu(\boldsymbol{\theta}, z) &:= \frac{1}{|\det A(\boldsymbol{\theta}, z)|} \\ &= \frac{1}{\left| [1 - w(z)\kappa(\boldsymbol{\theta})]^2 - w^2(z) |\gamma(\boldsymbol{\theta})|^2 \right|}. \end{aligned} \quad (2.5)$$

The cluster is non-critical for sources at redshift z if $\det A(\boldsymbol{\theta}, z) > 0$ everywhere; it is non-critical for all source redshifts if $[1 - \kappa(\boldsymbol{\theta})]^2 - |\gamma(\boldsymbol{\theta})|^2 > 0$ for all $\boldsymbol{\theta}$.

3. Local observables and their dependence on local lens parameters κ and γ

Throughout this section we assume that these lensing parameters can be considered as constant over a small solid angle around $\boldsymbol{\theta}$, and we will therefore suppress the argument $\boldsymbol{\theta}$ in all equations.

In contrast to Papers I & II, we use here the ellipticity parameter

$$\epsilon = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2\sqrt{Q_{11}Q_{22} - Q_{12}^2}} \quad (3.1)$$

where Q_{ij} are the components of the tensor of second brightness moments, as defined in Paper I. Defining the intrinsic source ellipticity ϵ_s in complete analogy to (3.1), and using the locally

linearized lens equation, yielding $Q_s = AQA$, one finds for the transformation of intrinsic to observed ellipticity:

$$\epsilon(\epsilon_s, z) = \begin{cases} \frac{\epsilon_s - g(z)}{1 - g^*(z)} \epsilon_s & \text{for } |g(z)| \leq 1, \\ \frac{1 - g(z)}{\epsilon_s^* - g^*(z)} \epsilon_s^* & \text{for } |g(z)| > 1, \end{cases} \quad (3.2)$$

where

$$g(z) = \frac{\gamma(z)}{1 - \kappa(z)} = \frac{w(z)\gamma}{1 - w(z)\kappa}. \quad (3.3)$$

The condition $|g(z)| \leq 1$ ($|g(z)| > 1$) in Eq. (3.2) is equivalent to the condition $\det A(z) \geq 0$ ($\det A(z) < 0$).

Now, let $p_{\epsilon_s}(y) y \, dy \, d\varphi$ be the probability that the source ellipticity $\epsilon_s = y \exp(2i\varphi)$ is within $y \, dy \, d\varphi$ around ϵ_s . Then, for fixed redshift z , the expectation value of the n -th moment ϵ^n is given through

$$\begin{aligned} \langle \epsilon^n \rangle_{\epsilon_s}(z) &\equiv \int_0^1 dy y p_{\epsilon_s}(y) \int_0^{2\pi} d\varphi \epsilon^n(\epsilon_s, z) \\ &= \begin{cases} [-g(z)]^n & \text{for } |g(z)| \leq 1, \\ \left(\frac{-1}{g^*(z)}\right)^n & \text{for } |g(z)| > 1. \end{cases} \end{aligned} \quad (3.4)$$

The remarkable fact that the expectation values $\langle \epsilon^n \rangle_{\epsilon_s}$ do not depend on the source ellipticity distribution, whereas $\langle \chi \rangle_{\chi_s}$ does (Papers I & II), is the reason for choosing ϵ as the ellipticity parameter in this paper. We derive (3.4) in the Appendix 1.

If the galaxies are distributed in redshift according to the probability density $p_z(z)$, the expectation values of the moments ϵ^n become:

$$\begin{aligned} \langle \epsilon^n \rangle_{\epsilon_s, z} &\equiv \int_0^\infty dz p_z(z) \langle \epsilon^n \rangle_{\epsilon_s}(z) \\ &= \int_{\det A(z) \geq 0} dz p_z(z) [-g(z)]^n \\ &+ \int_{\det A(z) < 0} dz p_z(z) \left(\frac{-1}{g^*(z)}\right)^n \\ &= \gamma^n \int_{\det A(w) \geq 0} dw p_w(w) \left(\frac{-w}{1 - \kappa w}\right)^n \\ &+ \left(\frac{1}{\gamma^*}\right)^n \int_{\det A(w) < 0} dw p_w(w) \left(\frac{1 - \kappa w}{-w}\right)^n \\ &\equiv \gamma^n X_n(\kappa, \gamma) + \frac{1}{\gamma^{*n}} Y_n(\kappa, \gamma) \\ &= \gamma^n \left(X_n(\kappa, \gamma) + \frac{1}{|\gamma|^{2n}} Y_n(\kappa, \gamma) \right). \end{aligned} \quad (3.5)$$

In the third line we used the transformation of the source redshift distribution into their w -distribution, which consists of a delta ‘function’ at $w = 0$ for galaxies with $z \leq z_d$ and is given by $p_w(w) dw = p_z(z) dz$ for galaxies with $z > z_d$. For a single redshift z_0 of the sources, the expectation values reduce to $\langle \epsilon^n \rangle_{\epsilon_s, z} = \langle \epsilon^n \rangle_{\epsilon_s}(z_0)$ given in (3.4).

Generally, the boundaries of the integrals in (3.5), and therefore X_n and Y_n , depend on κ and γ ,

$$\begin{aligned} X_n(\kappa, \gamma) &= \left(\int_0^{\min(1, 1/|\kappa+|\gamma|)} + \int_{\frac{1}{\max(1, \kappa-|\gamma|)}}^1 \right) \\ &\times dw p_w(w) \left(\frac{-w}{1 - \kappa w}\right)^n \end{aligned} \quad (3.6)$$

$$Y_n(\kappa, \gamma) = \int_{\min(1, 1/|\kappa+|\gamma|)}^{\frac{1}{\max(1, \kappa-|\gamma|)}} dw p_w(w) \left(\frac{1 - \kappa w}{-w}\right)^n.$$

One sees that for some values of the parameters κ and γ there exist two intervals of w for which $\det A(w) \geq 0$, separated by that interval for which $\det A(w) < 0$.

In the case of weak lensing we obtain from (3.5) that $\langle \epsilon^n \rangle_{\epsilon_s, z} \approx (-1)^n \langle w^n \rangle \gamma^n$, and the shear is – modulo $\langle w^n \rangle$ – an observable.

An estimate of these expectation values for the moments of the image ellipticities can be obtained by considering (locally) an ensemble of sources with ellipticities ϵ_k and defining the means

$$\bar{\epsilon}^n \equiv \frac{\sum_{k=1}^{N_{\text{gal}}} \epsilon_k^n u_k}{\sum_{k=1}^{N_{\text{gal}}} u_k}, \quad (3.7)$$

where u_k is an appropriately chosen weight factor (see Paper II or Seitz & Schneider 1996). These mean values $\bar{\epsilon}^n$ are statistically distributed around the expectation values $\langle \epsilon^n \rangle_{\epsilon_s, z}$, and we use them as an estimate for $\langle \epsilon^n \rangle_{\epsilon_s, z}$.

To summarize, the mean image ellipticity and their higher moments $\langle \epsilon^n \rangle_{\epsilon_s, z}$ do no longer directly provide us with a local estimate for $\gamma/(1 - \kappa)$ or its inverse as in the case where all sources are at the same redshift, which was assumed in previous papers. In fact, the dependence of $\langle \epsilon^n \rangle_{\epsilon_s, z}$ on the local lens parameters κ and γ can be quite complicated, depending on the redshift distribution of the sources and on the local lens parameters itself.

To obtain information on the local lens parameters κ, γ from $\bar{\epsilon} \approx \langle \epsilon \rangle_{\epsilon_s, z}$ we have to know the redshift distribution $p_z(z)$ of the galaxy population which enters Eq. (3.5). We point out that the average in (3.5) extends over all galaxies, i.e., also over those situated in front of the lens, which, however, do not provide any information on the local lens parameters, since $\langle \epsilon^n \rangle_{\epsilon_s}(z) = 0$ for $z \leq z_d$, but contribute to the noise which is inherent in the estimate of the local lens parameters derived from $\bar{\epsilon}$. We want to stress that galaxies situated in the foreground of the cluster do not affect the shear estimate systematically, as long as their fraction is known and properly included in (3.5) via $p_w(w)$. In the rest of this paper we assume that this redshift distribution of galaxies is known or at least can be estimated, e.g., by the lensing effect itself (see Bartelmann & Narayan 1995, Kneib et al. 1995).

For illustration we assume that the redshift distribution is given by a function of the form

$$p_z(z) = \frac{\beta z^2}{\Gamma(3/\beta) z_0^3} \exp\left(- (z/z_0)^\beta\right), \quad (3.8)$$

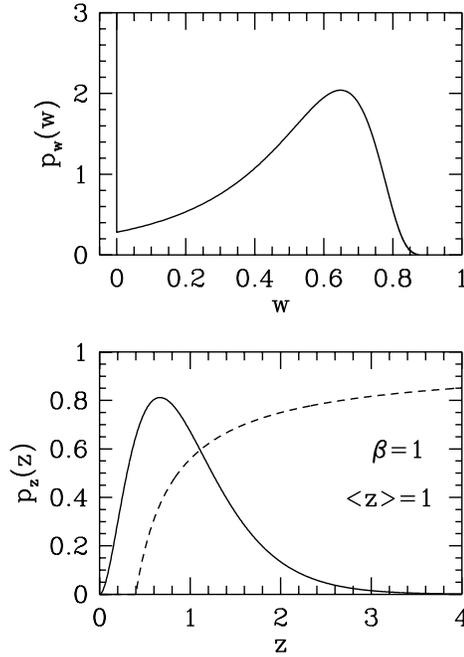


Fig. 1. The lower panel (solid line) shows the redshift distribution given in Eq. (3.8) for $\langle z \rangle = 3z_0 = 1$ and $\beta = 1$. The dashed line shows the function $w(z)$ defined in Eq. (2.4) for a cluster redshift of $z_d = 0.4$. The upper panel shows the corresponding distribution $p_w(w)$. Note that $p_w(w)$ has a delta-‘function’ peak at $w = 0$, with an amplitude given by the probability that a source has a redshift smaller than that of the lens

taken from Brainerd et al. (1995), with moments $\langle z \rangle = 3z_0$ for $\beta = 1$, $\langle z \rangle = 1.5z_0$ for $\beta = 1.5$ and $\langle z \rangle = \Gamma(4/3)z_0$ for $\beta = 3$.

In Fig. 1 we show this distribution $p(z)$ for $\beta = 1 = \langle z \rangle$ and the corresponding distribution $p_w(w)$ for a lens with redshift $z_d = 0.4$.

4. The reconstruction of the cluster mass distribution

4.1. Inversion relations for a single source redshift

The reconstruction of the surface mass density of a cluster from lensed images of background sources has been described in several papers (see introduction) for sources at the same redshift z . Let $\kappa(\boldsymbol{\theta})$ be the dimensionless surface mass density of the cluster, scaled by the appropriate critical surface mass density, and let $\gamma(\boldsymbol{\theta})$ be the corresponding complex shear. Both quantities are given as second partial derivatives of the deflection potential, and it was shown by Kaiser (1995) that the following relation between third partial derivatives of the deflection potential, or first partial derivatives of κ and γ , is valid:

$$\nabla \kappa(\boldsymbol{\theta}) = - \begin{pmatrix} \gamma_{1,1} + \gamma_{2,2} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix} \equiv \mathbf{U}(\boldsymbol{\theta}) \quad (4.1)$$

Hence, the surface mass density can be obtained in terms of the shear by integrating this first-order differential equation. However, the shear is not an observable in general (Kochanek 1990;

Paper I), but the quantity $g(\boldsymbol{\theta}) = \gamma(\boldsymbol{\theta})/[1 - \kappa(\boldsymbol{\theta})]$, if we confine our considerations to non-critical clusters. (For critical clusters only $2g/(1 + |g|^2)$ is observable.) Inserting $\gamma = g(1 - \kappa)$ into (4.1) yields (Kaiser 1995)

$$\nabla K(\boldsymbol{\theta}) = \frac{1}{1 - g_1^2 - g_2^2} \begin{pmatrix} 1 + g_1 & g_2 \\ g_2 & 1 - g_1 \end{pmatrix} \begin{pmatrix} g_{1,1} + g_{2,2} \\ g_{2,1} - g_{1,2} \end{pmatrix} \quad (4.2)$$

$$\equiv \mathbf{u}(\boldsymbol{\theta}) \quad ,$$

where

$$K(\boldsymbol{\theta}) := \ln [1 - \kappa(\boldsymbol{\theta})] \quad . \quad (4.3)$$

Hence, it is possible to derive the gradient of the quantity K in terms of the observable quantity g . Obviously, the integration of (4.2) allows an arbitrary integration constant, i.e., $K(\boldsymbol{\theta})$ can only be determined up to an additive constant. Several methods exist to perform an integration of (4.2); see Schneider (1995), Kaiser et al. (1995), Bartelmann (1995), and Seitz & Schneider (1996). All of them would be equivalent if $\mathbf{u}(\boldsymbol{\theta})$ were a gradient vector field. However, since \mathbf{u} is determined observationally, it is not exact and thus in general not rotation-free. The integration derived in Seitz & Schneider (1996),

$$K(\boldsymbol{\theta}) = \int_{\mathcal{U}} d^2\theta' \mathbf{H}(\boldsymbol{\theta}'; \boldsymbol{\theta}) \cdot \mathbf{u}(\boldsymbol{\theta}') + \bar{K} \quad , \quad (4.4)$$

was explicitly constructed to account for this ‘rotational noise component’ and has been demonstrated in Seitz & Schneider (1996) to work better than the other proposed methods. Here, $\mathbf{H}(\boldsymbol{\theta}'; \boldsymbol{\theta})$ is a vector field explicitly constructed in Seitz & Schneider (1996), and \bar{K} is the average of $K(\boldsymbol{\theta})$ over the data field \mathcal{U} , i.e., the region where image ellipticities have been measured. Of course, \bar{K} is an undetermined constant, so that the surface mass density $\kappa(\boldsymbol{\theta})$ is determined up to the transformation

$$\kappa(\boldsymbol{\theta}) \rightarrow (1 - \lambda) + \lambda\kappa(\boldsymbol{\theta}) \quad . \quad (4.5)$$

4.2. General inversion method

In the general case of a redshift distribution of the sources, we again make use of (4.1). The formal integration of (4.1) proceeds in the same way as that of (4.2), i.e.,

$$\kappa(\boldsymbol{\theta}) = \int_{\mathcal{U}} d^2\theta' \mathbf{H}(\boldsymbol{\theta}'; \boldsymbol{\theta}) \cdot \mathbf{U}(\boldsymbol{\theta}') + \bar{\kappa} \quad , \quad (4.6)$$

where $\bar{\kappa}$ is the average of κ over the data field \mathcal{U} . The vector field $\mathbf{U}(\boldsymbol{\theta})$ is defined in (4.1) and given by first partial derivatives of the shear. The shear in turn is related to the mean image ellipticity via (3.5),

$$\langle \epsilon \rangle_{\epsilon_{s,z}} = \gamma \left(X_1(\kappa, \gamma) + \frac{1}{|\gamma|^2} Y_1(\kappa, \gamma) \right) \approx \bar{\epsilon} \quad . \quad (4.7)$$

Note that Eq. (4.7) is a local relation, valid at every point $\boldsymbol{\theta}$. The complicated dependence of $\langle \epsilon \rangle_{\epsilon_{s,z}}$ on κ and γ suggests an iterative approach for the solution of the inversion problem:

let $\bar{\epsilon}$ be an ‘measured’ estimate for $\langle \epsilon \rangle_{\epsilon_{s,z}}$ and let $\gamma^{(n)}(\boldsymbol{\theta})$ and $\kappa^{(n)}(\boldsymbol{\theta})$ be an estimate for the shear field and the surface mass density. From that, an updated estimate for the shear field can be obtained, using (4.7):

$$\gamma^{(n+1)}(\boldsymbol{\theta}) = \bar{\epsilon} \left[X_1(\kappa^{(n)}(\boldsymbol{\theta}), \gamma^{(n)}(\boldsymbol{\theta})) + \frac{Y_1(\kappa^{(n)}(\boldsymbol{\theta}), \gamma^{(n)}(\boldsymbol{\theta}))}{|\gamma^{(n)}(\boldsymbol{\theta})|^2} \right]^{-1} \quad (4.8)$$

Then, by differentiation, the vector field $\mathbf{U}^{(n+1)}(\boldsymbol{\theta})$ can be calculated from (4.1), by using the shear field $\gamma^{(n+1)}(\boldsymbol{\theta})$. And finally, an updated estimate for the surface mass density field is obtained from (4.6),

$$\kappa^{(n+1)}(\boldsymbol{\theta}) = \int_{\mathcal{D}} d^2\theta' \mathbf{H}(\boldsymbol{\theta}'; \boldsymbol{\theta}) \cdot \mathbf{U}^{(n+1)}(\boldsymbol{\theta}') + \bar{\kappa} \quad (4.9)$$

This iteration process is started by choosing $\kappa^{(0)}(\boldsymbol{\theta}) = 0$, $\gamma^{(0)}(\boldsymbol{\theta}) = 0$. It is clear that the integration constant $\bar{\kappa}$ is still a free variable, i.e., with the method described here there remains a global invariance transformation of the resulting surface mass density field; in contrast to the case considered in Sect. 4.1, this transformation cannot be explicitly determined, due to the highly nonlinear relations occurring here. For critical clusters we need at most 10 steps to achieve a convergence of the iteration algorithm. For less massive clusters about 5 iteration steps are sufficient. We find that the iteration algorithm is more stable and converges faster than in the case of a single source redshift (see Paper II), mainly because there are no well-defined critical curves as function of $\boldsymbol{\theta}$, since their location depends on the source redshift.

4.3. The case for non-critical clusters

In the case of weak lensing ($\kappa \ll 1$, $|\gamma| \ll 1$), the mass reconstruction depends only on the mean value of w , so that $\kappa \propto 1/\langle w \rangle$ (Kaiser & Squires 1993). This ceases to be true if the cluster is not weak. Here we show how the generalization of (4.5) reads in the case of non-critical clusters and a redshift distribution of sources.

If the cluster is non-critical for all redshifts of the sources, the inversion problem can be simplified because $Y_n(\kappa, \gamma) = 0$ for all n – see (3.6) – and the X_n depend only on κ ,

$$X_n(\kappa) = \int_0^1 dw p_w(w) \left(\frac{-w}{1-\kappa w} \right)^n \quad (4.10)$$

Then, (3.5) simplifies to

$$\langle \epsilon^n \rangle_{\epsilon_{s,z}} = \gamma^n X_n(\kappa) \quad (4.11)$$

As a result, the iteration procedure described in the last subsection can be applied in a somewhat simpler way, by using (4.11) instead of (4.7); in addition, one can use the approximation (A2.4) derived in the Appendix 2 for $X_1(\kappa)$ which yields

$$X_1(\kappa) \approx \frac{-\langle w \rangle}{1 - \kappa \frac{\langle w^2 \rangle}{\langle w \rangle}} \quad (4.12)$$

This approximation is found to be sufficiently accurate to describe $X_1(\kappa)$ for (generic) non-critical clusters (see Fig. 8 in Appendix 2).

Combining (4.11) for $n = 1$, (4.12), and replacing the expectation value $\langle \epsilon \rangle_{\epsilon_{s,z}}$ with the observed local average $\bar{\epsilon}$, we obtain

$$\gamma = -\frac{\bar{\epsilon}}{\langle w \rangle} (1 - f \langle w \rangle \kappa) \quad , \quad (4.13)$$

with the definition

$$f := \frac{\langle w^2 \rangle}{\langle w \rangle^2} \quad . \quad (4.14)$$

Inserting this expression for γ into (4.1), one obtains after some manipulations

$$V \nabla \kappa = \frac{1}{\langle w \rangle} (1 - f \langle w \rangle \kappa) \begin{pmatrix} \bar{\epsilon}_{1,1} + \bar{\epsilon}_{2,2} \\ \bar{\epsilon}_{2,1} - \bar{\epsilon}_{1,2} \end{pmatrix} \quad , \quad (4.15)$$

with the matrix

$$V = \begin{pmatrix} 1 + f\bar{\epsilon}_1 & f\bar{\epsilon}_2 \\ f\bar{\epsilon}_2 & 1 - f\bar{\epsilon}_1 \end{pmatrix} \quad .$$

Thus,

$$\nabla \kappa = \frac{1}{\langle w \rangle} (1 - f \langle w \rangle \kappa) V^{-1} \begin{pmatrix} \bar{\epsilon}_{1,1} + \bar{\epsilon}_{2,2} \\ \bar{\epsilon}_{2,1} - \bar{\epsilon}_{1,2} \end{pmatrix} \quad , \quad (4.16)$$

with the inverse

$$V^{-1} = \frac{1}{1 - f^2(\bar{\epsilon}_1^2 + \bar{\epsilon}_2^2)} \begin{pmatrix} 1 - f\bar{\epsilon}_1 & -f\bar{\epsilon}_2 \\ -f\bar{\epsilon}_2 & 1 + f\bar{\epsilon}_1 \end{pmatrix} \quad . \quad (4.17)$$

If we now define

$$K(\boldsymbol{\theta}) := \ln(1 - f \langle w \rangle \kappa(\boldsymbol{\theta})) \quad , \quad (4.18)$$

(4.16) can be written as

$$\nabla K(\boldsymbol{\theta}) = f V^{-1} \begin{pmatrix} \bar{\epsilon}_{1,1} + \bar{\epsilon}_{2,2} \\ \bar{\epsilon}_{2,1} - \bar{\epsilon}_{1,2} \end{pmatrix} \equiv \mathbf{u}(\boldsymbol{\theta}) \quad . \quad (4.19)$$

Note the similarity between (4.19) and (4.2). Thus, the vector field $\mathbf{u}(\boldsymbol{\theta})$ can be constructed directly from the observable $\bar{\epsilon}$, as in the case of all sources being at the same redshift, and the same inversion equation (4.4) should be used, but with the current definitions of K and \mathbf{u} . Thus, by using the approximation (4.12), the inversion of a non-critical cluster is no more complicated than in the case of a single source redshift. In particular, (4.18) immediately shows that K can be determined only up to an additive constant, which implies the invariance transformation

$$\kappa(\boldsymbol{\theta}) \rightarrow \frac{(1 - \lambda)}{f \langle w \rangle} + \lambda \kappa(\boldsymbol{\theta}) \quad ; \quad (4.20)$$

in other words, $(1 - \langle w^2 \rangle \kappa(\boldsymbol{\theta}) / \langle w \rangle)$ can be determined only up to a multiplicative constant.

5. Breaking the mass invariance

In Sect. 4, we have shown that there exists an invariance transformation for the surface mass density map which leaves the observed local mean image ellipticities constant (4.5, 4.9 & 4.20). The occurrence of the invariance can be traced back to the fact that the inversion equation was derived from a first-order differential equation – see (4.1) – which has a free integration constant. We now discuss several ideas to determine this constant.

5.1. Using only shape information

In the case of a single source redshift the invariance transformation for the surface mass density can be formulated most generally by saying that the probability distribution $p(\epsilon)$ of the observed image ellipticities is unchanged under the transformation (4.5); this was explicitly shown in Paper I. The invariance transformation as expressed in (4.9) and (4.20) is of a different kind, since it does *not* leave $p(\epsilon)$ invariant, but only its first order moment $\langle \epsilon \rangle$. Therefore, one might be tempted to try to break the invariance transformation by using information of the image ellipticity distribution only.

5.1.1. A maximum-likelihood approach

We consider a fixed set of ‘observed’ galaxies with ellipticities ϵ_i and positions θ_i . Then we reconstruct the surface mass density of a cluster as described in Sect. 4.2 assuming a value $\bar{\kappa}$ for the mean mass density in the field [see Eq. (4.9)]. From that we calculate the mass density $\kappa(\theta_i)$ and shear $\gamma(\theta_i)$ at the positions θ_i ($i = 1, \dots, N$). Next we calculate the probability $p(\epsilon_i, \kappa(\theta_i), \gamma(\theta_i))$ to observe the image ellipticity ϵ_i at the galaxy position θ_i . Finally we maximize the likelihood $\mathcal{L}_1 = \prod_i p(\epsilon_i, \kappa(\theta_i), \gamma(\theta_i))$ using different values for $\bar{\kappa}$ in the mass reconstruction. We find that the mass degeneracy is broken, but unfortunately only weakly, and only in those regions of the cluster where the weak lensing regime is violated. From our simulation we find that the number density of galaxy images is orders of magnitudes too low to allow for the breaking of the mass sheet degeneracy in practice. Furthermore, the likelihood depends much stronger on the assumed redshift distribution – which in practice is weakly constrained – than on the value of the mean mass density $\bar{\kappa}$ and is therefore a much better tool to investigate the redshift distribution of the sources than to determine the mean surface mass density of a particular cluster.

5.1.2. Second moments of image ellipticities

For non-critical clusters, the ratio

$$R = \frac{\left| \langle \epsilon^2 \rangle_{\epsilon_s, z} \right|}{\left| \langle \epsilon \rangle_{\epsilon_s, z} \right|^2} = \frac{X_2(\kappa)}{X_1(\kappa)^2} \approx f \frac{\left(1 - \kappa \frac{\langle w^2 \rangle}{\langle w \rangle} \right)^2}{\left(1 - \kappa \frac{\langle w^3 \rangle}{\langle w^2 \rangle} \right)^2} \quad (5.1)$$

depends only on the local surface mass density κ . We derive (5.1) in Appendix 3. Measuring R could therefore provide a

direct estimate for κ . However, we can not measure $\langle \epsilon^2 \rangle_{\epsilon_s, z}$ and $\langle \epsilon \rangle_{\epsilon_s, z}$ accurately enough to determine the local mass density κ or the average mass density $\bar{\kappa}$ from a reasonable number density of image ellipticities (see Appendix 3).

5.2. Magnification information through number density of images

Let us assume that the number density $n_0(S, z)$ of the unlensed faint galaxies with flux S and redshift z is given through

$$n_0(S, z) = p_z(z) F(S), \quad (5.2)$$

where $p_z(z)$ is the normalized redshift distribution and $F(S)$ the distribution in flux. Note that the factorization of $n_0(S, z)$ as given in (5.2) is a fairly special assumption for the galaxy distribution, which will not be valid in general. However, for simplicity we shall make use of (5.2) in this paper. This factorization may in fact be an approximate distribution over a limited range of flux S . Since the magnifications are not very large except perhaps in the very central parts of the cluster, the ‘dynamic range’ over which (5.2) is applied is not large and might be a valid description. It should be noted that this factorization has already been used implicitly in (3.5). If (5.2) is not assumed, then the redshift distribution of sources locally – where the surface mass density is κ and the shear is γ – will differ from $p_z(z)$ due to the magnification, so that the functions X_n and Y_n will attain an additional dependence on κ and $|\gamma|$. This will not cause any additional conceptual problem.

Then, consider a position θ in the cluster with surface mass density $\kappa(\theta)$ and shear $\gamma(\theta)$. The observed number density of galaxies with redshift z and flux larger than S is (see Broadhurst et al. 1995)

$$n(> S, z, \theta) = p_z(z) \frac{1}{\mu(\theta, z)} F\left(> \frac{S}{\mu(\theta, z)}\right), \quad (5.3)$$

with the magnification defined in (2.5). The total number density of galaxies observed with flux larger than S is obtained through integration of (5.3) and yields

$$n(> S, \theta) = \int_0^\infty dz p_z(z) \frac{1}{\mu(\theta, z)} F\left(> \frac{S}{\mu(\theta, z)}\right). \quad (5.4)$$

If $F(> S) \propto S^{-\alpha}$, then we obtain from (5.4)

$$\begin{aligned} n(> S, \theta) &= n_0(> S) \int_0^\infty dz p_z(z) \mu^{\alpha-1}(\theta, z) \\ &\equiv n_0(> S) \langle \mu^{\alpha-1}(\theta, z) \rangle_z. \end{aligned} \quad (5.5)$$

From Eq. (5.5) we conclude that the number density is not changed if $\alpha = 1$ and it is increased (decreased) for $\alpha > 1$ ($\alpha < 1$). Next, averaging (5.5) over the data field \mathcal{U} with area U , we obtain

$$\begin{aligned} \langle n(> S, \theta) \rangle_{\mathcal{U}} &\equiv \frac{1}{U} \int_{\mathcal{U}} d^2\theta n(> S, \theta) \\ &= n_0(> S) \langle \mu^{\alpha-1}(\theta, z) \rangle_{z, \mathcal{U}}. \end{aligned} \quad (5.6)$$

Hence, the ratio of the number of observed galaxies in the data field \mathcal{U} to the number $U n_0(> S)$ which would be observed in the absence of a foreground lens gives $\langle \mu^{\alpha-1}(\boldsymbol{\theta}, z) \rangle_{z, \mathcal{U}}$. The number density $n_0(> S)$ of the unlensed galaxies is regarded as an universal function and has been measured in several colours down to very faint magnitudes (see for example Smail et al. 1995). Because of this, the local observables in the case of cluster lensing are the moments of the image ellipticities $\langle \epsilon^n \rangle(\boldsymbol{\theta})$ and $\langle \mu(\boldsymbol{\theta}, z) \rangle_z$. We note that the result that $\langle \mu(\boldsymbol{\theta}, z) \rangle_z$ is a local observable is only true for the ansatz (5.2) with $F(> S) \propto S^{-\alpha}$. If either the ansatz (5.2) does not hold, or if $F(> S)$ is not a power law, then the observable quantity is a different one and can in particular have a much more complicated dependence on the magnification. However, $n(> S, \boldsymbol{\theta})$ can still be compared – may be not analytically – with $n_0(> S)$ and provides information on the local magnification.

In the following we consider a fixed value of the flux threshold S and use for notational simplicity $n(\boldsymbol{\theta})$ and n_0 instead of $n(> S, \boldsymbol{\theta})$ and $n_0(> S)$.

5.2.1. Do galaxy positions break the mass invariance?

The probability to observe a galaxy within the area $d^2\theta$ around $\boldsymbol{\theta}$ is

$$p(\boldsymbol{\theta}) d^2\theta = \frac{n(\boldsymbol{\theta})}{\int_{\mathcal{U}} d^2\theta n(\boldsymbol{\theta})} d^2\theta. \quad (5.7)$$

Using Eqs. (5.5) and (5.6) we obtain this probability in terms of the local (redshift-averaged) magnification

$$p(\boldsymbol{\theta}) d^2\theta = \frac{1}{U} \frac{\langle \mu^{\alpha-1}(\boldsymbol{\theta}, z) \rangle_z}{\langle \mu^{\alpha-1}(\boldsymbol{\theta}, z) \rangle_{z, \mathcal{U}}} d^2\theta. \quad (5.8)$$

The likelihood \mathcal{L} for observing N galaxies with positions $\boldsymbol{\theta}_i$ ($i = 1, \dots, N$) is

$$\mathcal{L} = \prod_i p(\boldsymbol{\theta}_i). \quad (5.9)$$

On the other hand, we can perform the mass reconstruction according to Eq. (4.9) and obtain $\langle \mu^{\alpha-1}(\boldsymbol{\theta}, z, \bar{\kappa}) \rangle_z$ as a function of the assumed mean mass density $\bar{\kappa}$ in the field. From that we can calculate with Eq. (5.8) the probability density $p(\boldsymbol{\theta}_i)$ and finally according to (5.9) the likelihood $\mathcal{L}(\bar{\kappa})$ for the assumed mean mass density. However this does not work. To understand this, assume for a moment that all galaxies are at the same redshift; for simplicity we take this redshift to be ‘infinity’ here. Then the reconstructed mass density κ is related to the true one κ_{true} through the invariance transformation (4.5) which gives $(1 - \kappa_{\text{true}})\lambda = 1 - \kappa$ and $\gamma_{\text{true}}\lambda = \gamma$. For the magnification we obtain

$$\mu_{\text{true}}^{\alpha-1}(\boldsymbol{\theta}) = \mu^{\alpha-1}(\boldsymbol{\theta})\lambda^{2-2\alpha}, \quad (5.10)$$

and from that we conclude that $p(\boldsymbol{\theta})$ is independent of λ , or, equivalently, independent of $\bar{\kappa}$. This means that for all sources at the same redshift, the *positions* of the observed galaxies do not give any information on the *mean mass density* in the field.

If the sources are distributed in redshift then the mean mass $\bar{\kappa}$ assumed for the reconstruction changes the ratio

$$p(\boldsymbol{\theta}, \bar{\kappa}) = \frac{1}{U} \frac{\langle \mu^{\alpha-1}(\boldsymbol{\theta}, z, \bar{\kappa}) \rangle_z}{\langle \mu^{\alpha-1}(\boldsymbol{\theta}, z, \bar{\kappa}) \rangle_{z, \mathcal{U}}}. \quad (5.11)$$

However, the dependence on $\bar{\kappa}$ is very weak as can be verified for non-critical clusters through calculating, for fixed $\kappa(\boldsymbol{\theta})$ and $\gamma(\boldsymbol{\theta})$, the function $\langle \mu^{\alpha-1}(\boldsymbol{\theta}, z) \rangle_z(\lambda)$ using the invariance transformation (4.20). As a result, the *positions* of the observed galaxies provide no or not enough information to break the mass degeneracy.

Up to here all our attempts to break the mass degeneracy in practice have failed. This is because for all cases considered, the mass degeneracy is – in theory – broken because of the presence of a redshift distribution of the sources and is no longer broken if all sources are at the same redshift. We conclude that in order to derive $\bar{\kappa}$ one has to use a method which breaks the mass degeneracy not because of the presence of a redshift distribution of the sources, but independent of that.

5.2.2. Breaking the mass degeneracy through the total number of observed galaxies

We shall now consider the magnification effect on the distribution of galaxy images. We do not use the magnification effect locally as proposed by Broadhurst et al. (1995) and Broadhurst (1995) in order to derive the local mass density from it, but we use the ‘global’ magnification effect, i.e., we compare the total number of observed galaxies with the expected (unlensed one), and relate that to the mean mass density inside the observed field.

In the above subsection we have not used the number density n_0 of galaxies observable in the absence of lensing. If we use this, then we can calculate from Eq.(5.6) the expected number $\langle N \rangle(\bar{\kappa})$ of galaxies as a function of the mean mass density $\bar{\kappa}$ assumed in the mass reconstruction,

$$\langle N \rangle(\bar{\kappa}) = U n_0 \langle \mu^{\alpha-1}(\boldsymbol{\theta}, z, \bar{\kappa}) \rangle_{z, \mathcal{U}} \quad (5.12)$$

and compare that with the number N of galaxies actually observed. We compare the deviation in terms of the standard deviation $\sigma = \sqrt{\langle N \rangle}$ and obtain

$$\chi^2(\bar{\kappa}) = \frac{[N - \langle N \rangle(\bar{\kappa})]^2}{\langle N \rangle(\bar{\kappa})}. \quad (5.13)$$

The value $\bar{\kappa}$ for which $\chi^2(\bar{\kappa})$ shows a minimum (in fact, at which $\chi^2 = 0$) gives the most probable value for the mean surface mass density of the cluster.

We note that in the case of a single redshift of the sources (see Sect. 4.1) the determination of the mean mass density is even simpler: the mass reconstruction is done according to (4.4) with an assumed value for \bar{K} , say $\bar{K} = 0$. From the resulting mass and shear maps $\kappa(\boldsymbol{\theta})$ and $\gamma(\boldsymbol{\theta})$, one can calculate $\langle \mu^{\alpha-1}(\boldsymbol{\theta}) \rangle_{\mathcal{U}}$.

Because of Eq. (5.6) we can relate that to the true magnification and from Eq. (5.10) we conclude that

$$\frac{N}{Un_0} = \frac{\langle n(\theta) \rangle_{\mathcal{U}}}{n_0} = \langle \mu_{\text{true}}^{\alpha-1}(\theta) \rangle_{\mathcal{U}} = \lambda^{2-2\alpha} \langle \mu^{\alpha-1}(\theta) \rangle_{\mathcal{U}}. \quad (5.14)$$

Thus, we can immediately calculate λ from (5.14), and with the invariance transformation (4.5) the true mass density $[1 - \kappa_{\text{true}}(\theta)] \lambda = 1 - \kappa(\theta)$.

In Sect. 6.2, we apply this method to break the mass invariance to a numerically simulated cluster ($z_d = 0.4$), for the case that the sources are distributed in redshift. We demonstrate that the method yields good estimates for the mean surface mass density, i.e., it yields good estimates on the total mass inside the observed field. However, we point out that the method is not applicable in the case that $\alpha \approx 1$, because then no magnification bias or antibias occurs, and the observed number density of galaxy images provides no information on the mean mass density $\bar{\kappa}$ in the field. We also remark that we do not use the magnification effect to determine the *shape* of the mass distribution, but only for the overall normalization of the mass.

We note that the method suggested by Bartelmann & Narayan (1995) still works in this case: they made use of the magnification of individual galaxy images, at fixed surface brightness which is known to be unchanged by gravitational light deflection. Their method will become extremely useful once sufficient HST data become available to obtain the mean size-surface brightness relation for these faint galaxies.

6. Application to simulated data

For testing the reconstruction technique we use the same numerically modelled cluster as in Paper II. Using its surface mass density shown in Fig. 2, we generate a distribution of synthetic galaxy images using the source ellipticity distribution

$$p_s(|\epsilon_s|) = \frac{1}{\pi R^2 [1 - \exp(-1/R^2)]} \exp(-|\epsilon_s|^2/R^2) \quad (6.1)$$

with $R = 0.15$,

and the redshift-distribution (3.8) with $\beta_{\text{true}} = 1$ and $\langle z \rangle_{\text{true}} = 1$. The galaxies are randomly distributed in the source plane, and the number density of galaxy images is $n_{\text{gal}} = 60/\text{arcmin}^2$. The cluster is placed at a redshift $z_d = 0.4$ so that $\langle w \rangle = 0.44$, $\langle w^2 \rangle = 0.25$ and $f = \langle w^2 \rangle / \langle w \rangle^2 = 1.29$.

6.1. The derived mass distribution for different assumed redshift distributions of the sources

We show in Fig. 3 the resulting surface mass density map for $\bar{\kappa} = \bar{\kappa}_{\text{true}} = 0.12$, $\beta = \beta_{\text{true}} = 1$ and $\langle z \rangle = \langle z \rangle_{\text{true}} = 1$, i.e., the assumed redshift distribution is the true one. We find a good agreement with the original mass density displayed in Fig. 2. Compared to the mass reconstruction shown Paper II – where all sources were at the same redshift – the noise is increased due

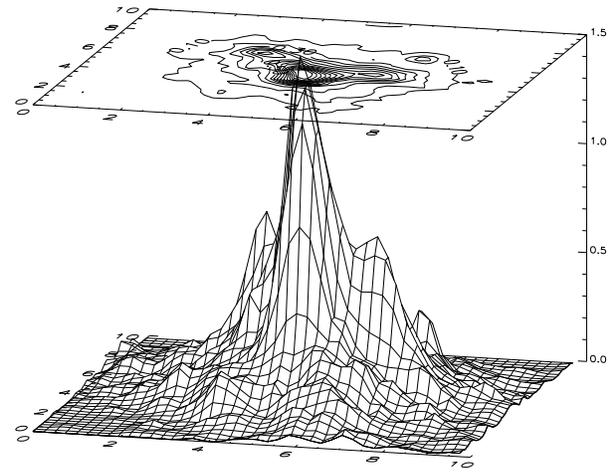


Fig. 2. The surface mass density distribution of the numerically modelled cluster. The sidelength is about $10'$ corresponding to 3.8 Mpc for $H_0 = 50 \text{ km/s/Mpc}$ and an EdS universe. The levels of the contour lines are 0.1, 0.2, \dots , 1.3.

to the dependence of the lensing strength of the cluster on the redshift of the sources.

However, in practice we do not know $\bar{\kappa}_{\text{true}}$ and we therefore choose $\bar{\kappa}$ such that the minimum of the – locally smoothed – reconstructed surface mass density is zero. The reason for not requiring the minimum of the reconstructed mass map to be zero but that of the locally smoothed mass density (gaussian smoothing function with smoothing length of $1'$) is that the former quantity is much more affected by noise. Choosing $\bar{\kappa}$ in that way leads to a good reconstruction only if the minimum of the cluster mass distribution is indeed close to zero, which is valid in practice if the observed field is several Mpc wide.

With that procedure we obtain almost the same mass density as that shown in Fig. 3 and the mean mass density detected inside the field is $\bar{\kappa} = 0.133$.

In Fig. 4 we show the mass distribution obtained using $\langle z \rangle = 1.5$ and $\beta = 2$. The mean mass density $\bar{\kappa} = 0.098$ is chosen such that the minimum of the locally-smoothed mass density is zero. Overestimating the mean redshift of the galaxies leads to an underestimation of the overall mass distribution and vice versa. This is shown in Fig. 5, where we use $\langle z \rangle = 0.6$ and $\beta = 3$ for the reconstruction and obtain a mean mass density of $\bar{\kappa} = 0.228$. Common to the Figs. 3 to 5 is that the substructure is nicely recovered, regardless of the assumed redshift distribution.

In Fig. 6 we show the mean mass density obtained for different assumed redshift distributions. We find that this is approximately proportional to $1/\langle w \rangle$, i.e., it is inversely proportional to the assumed mean effective lensing strength.

6.2. Breaking the mass invariance in practice

We now use Eq. (5.12) to break the mass invariance. First we choose the parameter α describing the number counts versus flux of the faint galaxies. Again, the source ellipticity- and redshift distribution is given through (3.8) and (6.1) with param-

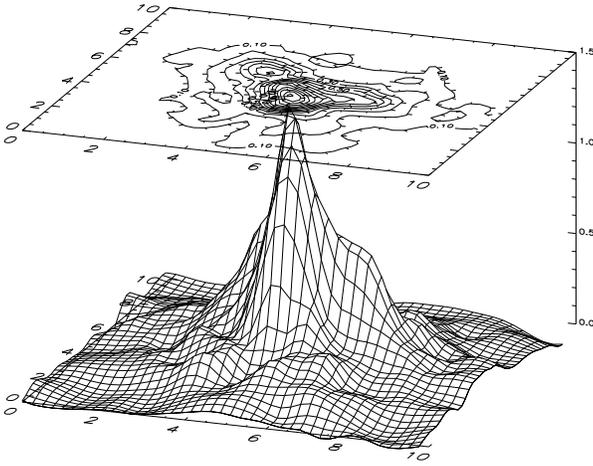


Fig. 3. The reconstructed mass density obtained solving Eq. (4.9) iteratively, assuming that the redshift distribution is known and assuming $\bar{\kappa} = \bar{\kappa}_{\text{true}} = 0.12$. We obtain almost the same result if we choose $\bar{\kappa}$ such that the minimum of the locally-smoothed mass density is zero. This assumption leads to $\bar{\kappa} = 0.133$. The levels of the contour lines are the same as in Fig. 2

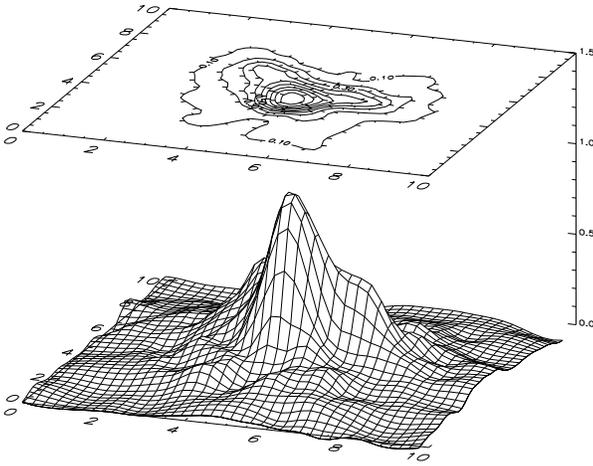


Fig. 4. The reconstructed mass density obtained solving Eq. (4.9) iteratively assuming that the redshift distribution given through (3.8) with $\langle z \rangle = 1.5$ and $\beta = 2$. The true redshift distribution is described by $\langle z \rangle_{\text{true}} = 1$ and $\beta_{\text{true}} = 1$. For the mean mass density we obtain $\bar{\kappa} = 0.098$. The levels of the contour lines are the same as in Fig. 2

eters as given in Table 1. The spatial distribution of the galaxy images is given through Eq. (5.5), where $\langle \mu^{\alpha-1}(\boldsymbol{\theta}, z) \rangle_z$ is calculated from integrating the ‘true’ local magnification weighted with the redshift distribution of the sources. We keep the number of galaxy images (lens plane) fixed and use $N = 5260$ galaxy images for our simulations.

Next we perform the mass reconstruction according to Eq. (4.9) for different values of $\bar{\kappa}$. From the resulting mass and shear map we calculate the map $\langle \mu^{\alpha-1}(\boldsymbol{\theta}, z, \bar{\kappa}) \rangle_z$, where we assume to know the true redshift distribution of the sources and the true slope α . Finally, we derive from Eq. (5.12) the number of

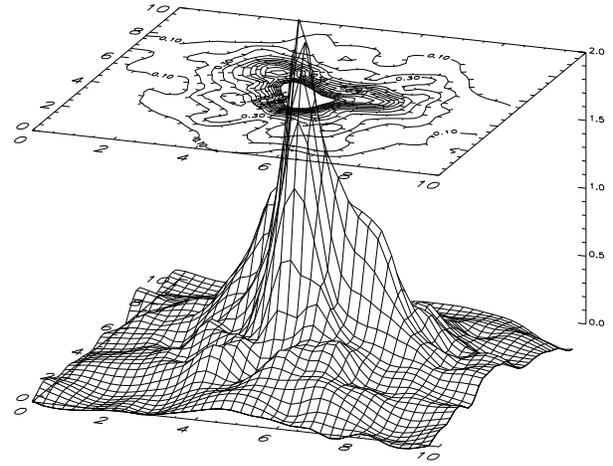


Fig. 5. Same as Fig. 4 but assuming that the redshift distribution is given through (3.8) with $\langle z \rangle = 0.6$ and $\beta = 3$. For the mean mass density we obtain $\bar{\kappa} = 0.228$. The levels of the contour lines are the same as in Fig. 2. Note that the scale of the z -axis is increased in this figure

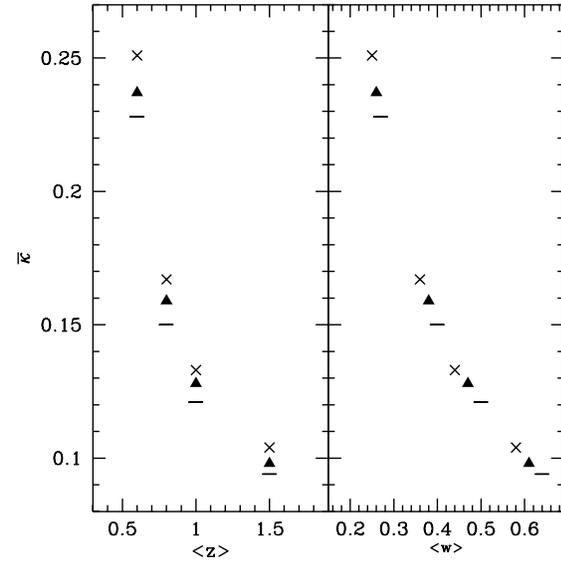


Fig. 6. The mean mass density $\bar{\kappa}$ averaged over the data field \mathcal{U} as a function of the assumed redshift distribution of the sources. The mean mass density is obtained such that the minimum of the locally averaged mass density is zero. For redshift distributions with $\beta = 1$ we use crosses, for $\beta = 1.5$ triangles and for $\beta = 3$ dashes. The left panel shows the mean mass density obtained from the reconstruction as a function of the mean redshift $\langle z \rangle$, the right panel as a function of $\langle w \rangle$ defined in (A2.1). For producing the synthetic galaxy images we use for their redshift distribution $\langle z \rangle_{\text{true}} = 1$ and $\beta_{\text{true}} = 1$, i.e., $\langle w \rangle = 0.44$. The true mean mass density of the cluster is $\bar{\kappa}_{\text{true}} = 0.12$

expected galaxies for the assumed value of $\bar{\kappa}$ and calculate from that with Eq. (5.13) the function $\chi^2(\bar{\kappa})$. The resulting function $\chi^2(\bar{\kappa})$ is shown in Fig. 7 for different values of the parameter α .

We find that all curves $\chi^2(\bar{\kappa})$ show a minimum close to $\bar{\kappa}_{\text{true}} = 0.12$, but the minima are narrower the more the slope

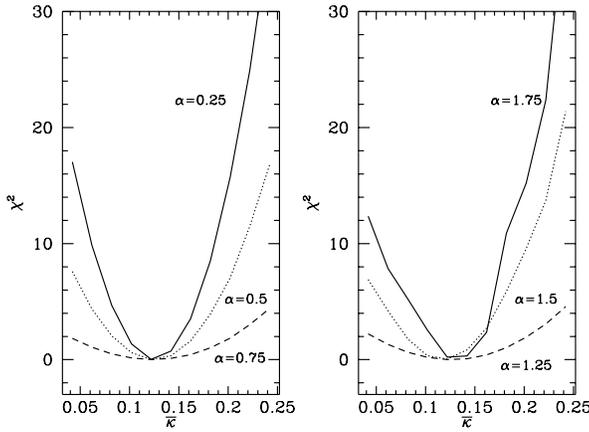


Fig. 7. The $\chi^2(\bar{\kappa})$ as a function of the assumed value of the mean mass density $\bar{\kappa}$ used for the mass reconstruction. We use different parameters α to describe the dependence of the number counts on the flux ($F(>S) \propto S^{-\alpha}$) of the faint galaxies. We assume to know the redshift distribution of the sources and the parameter α . The true values of the mean mass density is $\bar{\kappa}_{\text{true}} = 0.12$

α deviates from one. This is because the larger $|1 - \alpha|$, the stronger the magnification bias (or anti-bias) and the stronger the dependence of $\langle N \rangle(\bar{\kappa})$ is on $\bar{\kappa}$. Hence, for the same number density of galaxy images, the significance of rejecting values of $\bar{\kappa}$ increases with increasing $|1 - \alpha|$. From our simulations we find that the mean surface mass density is $0.09 \leq \bar{\kappa} \leq 0.16$ (2σ level) for $\alpha = 0.25$, $0.06 \leq \bar{\kappa} \leq 0.18$ for $\alpha = 0.5$ and $0 \leq \bar{\kappa} \leq 0.23$ for $\alpha = 0.75$. The true value of the mean surface mass density is $\bar{\kappa}_{\text{true}} = 0.12$. Increasing the number density of the observed galaxy images improves the significance levels, because the χ^2 function is roughly proportional to the number of observed galaxies. We note that the galaxies used for determining $\bar{\kappa}$ can be different from that used for the mass reconstruction, since the shape of the galaxies has not to be measured for the former purpose, increasing the number density of available galaxy images considerably. A number density of $n_{\text{gal}} = 100/\text{arcmin}^2$ would reduce the 2σ confidence levels to (0.08, 0.17) for $\alpha = 0.5$ and to (0.10, 0.15) for $\alpha = 0.25$.

In practice, the redshift distribution of the sources is poorly known and the mean redshift $\langle z \rangle$ can be over- or underestimated in a cluster mass reconstruction. If the true mean redshift is $\langle z \rangle_{\text{true}} = 1$ and if we assume for the reconstruction $\langle z \rangle = 0.8$ [1.5], then we obtain the 2σ confidence limits $\bar{\kappa} \in (0.11, 0.23)$ [$\bar{\kappa} \in (0.08, 0.16)$]. To derive these limits we used a slope $\alpha = 0.5$ for the number counts of the faint galaxies.

7. Discussion

The present paper should be considered as a further step in building up the ground for nonlinear cluster reconstruction techniques. In contrast to Papers I & II we have taken into account a redshift distribution of the sources. Our main results can be summarized as follows:

(1) We have related the local expectation values of the moments ϵ^n of the image ellipticities to the local lens parameters γ and κ in Eq. (3.5). These moments depend in a complicated way on the redshift distribution of the sources as shown in Eqs. (3.5) and (3.6). In Sect. 4.3 we have shown that in the case of non-critical clusters the local expectation values $\langle \epsilon^n \rangle$ depend only on the local lens parameters γ and κ and on few moments $\langle w^k \rangle$ of the effective distance w of the sources from the lens, weighted with the redshift distribution of the sources [see Eqs. (4.12), (4.15), and (4.16)].

(2) In Sect. 4.2, we have generalized the inversion method developed in Paper II such that the redshift distribution of the sources is accounted for. This is important if the cluster is at a high redshift ($z_d \gtrsim 0.2$), because then the effective lensing strength for sources with $z_s \approx 0.5$ is significantly different from that for sources with $z_s \approx 1$. For the reconstruction the redshift distribution of the sources has to be assumed and the inversion equation is solved iteratively. In Sect. 6 we have applied the reconstruction technique to synthetic data and shown that the reconstructed mass density is in good agreement with the original one. Compared to the mass reconstruction in the case of a single source redshift, the noise is slightly increased if the same number density of galaxy images is used.

(3) For non-critical clusters we have simplified the inversion method in Sect. 4.3 such that only one integration is necessary. Then the inversion equation is very similar to that developed previously for a single redshift of the sources (see Sect. 4.1). For the reconstruction of a non-critical cluster, one only has to assume the first two moments $\langle w \rangle$ and $\langle w^2 \rangle$ of the effective distance w weighted with the redshift distribution. The weaker the cluster, the less information on the redshift distribution has to be available for the reconstruction of the mass distribution. For a weak cluster it is sufficient to know the first moment $\langle w \rangle$.

(4) We have shown that the mean surface mass density $\bar{\kappa}$ is still a free variable in the inversion equation (4.9), i.e., there still remains a global invariance transformation for the surface mass density field which leaves the field of $\langle \epsilon \rangle$ unchanged. For non-critical clusters, this invariance transformation is given explicitly in Eq. (4.20) and it is similar to that derived previously for a single redshift of the sources [see Eq. (4.5)].

(5) In Sect. 5 we have discussed some ideas to break the mass invariance: maximizing the likelihood for the observed image ellipticities, using the second moments $\langle \epsilon^2 \rangle$ of image ellipticities or the magnification effect on the position of the galaxy images. For all these cases, the mass invariance is broken only in the presence of a redshift distribution of the sources. We find that the invariance is broken too weakly to make use of it in practice.

(6) Using the magnification effect on the number density of the galaxy images (see Broadhurst, Taylor & Peacock 1995) is the most promising way to break the mass degeneracy. In contrast to Broadhurst et al. (1995) and Broadhurst (1995), we do not use the observed local number density to derive the local surface mass density (because this local estimate has a considerably larger noise) but we use the total number of observed images to determine the mean mass density: we assume a value

for $\bar{\kappa}$, perform the mass reconstruction, then calculate the local magnification and from that the number of expected galaxy images in the field. Comparing this with the number of ‘observed’ galaxy images via a χ^2 -analysis gives a confidence interval for $\bar{\kappa}$. This method works only if the distribution $F(> S) \propto S^{-\alpha}$ of the unlensed sources with flux larger than S has a slope $\alpha \neq 1$, because for $\alpha = 1$ the number density of lensed objects is not changed compared to the unlensed one. The confidence interval obtained for $\bar{\kappa}$ is the narrower, the larger $|\alpha - 1|$ is, because the magnification bias (anti-bias) increases with increasing $\alpha > 1$ (decreasing $\alpha < 1$).

(7) In Sect. 6 we have used the same numerically modelled cluster as in Paper II for testing the reconstruction technique. We reconstruct the mass density for the case that (i) the assumed redshift distribution is the true one, (ii) the mean redshift is overestimated and (iii) underestimated. The mean mass density $\bar{\kappa}$ is adjusted such that the minimum of the reconstructed mass density is zero. We find that the substructure of the cluster is nicely recovered regardless of the assumed redshift distribution. The total mass detected is within 10% of the true mass for (i), underestimated for (ii) and overestimated for (iii). Using the method as describe above in (6), we find that reliable limits on the mean mass density can be derived with a significance increasing with increasing number density of galaxy images and increasing $|\alpha - 1|$. We note that for a successful application of this method on real data it is essential (a) to know the number density of the unlensed sources which would have been detected in the absence of lensing but under the *same observing conditions*, (b) to choose a colour which gives an appropriate slope of the faint galaxies in flux ($|\alpha - 1|$ large !) and (c) to know the mean redshift of the sources considered.

Finally, we should mention that although the cluster reconstruction method developed in this and our earlier papers is considerably more complicated than straight application of the original Kaiser & Squires (1993) method, these modifications are essential if applied to WFPC2 observations of a cluster center. For the one case we considered, i.e., the cluster Cl 0939+4713 (Seitz et al. 1996), the small field-of-view of the WFPC2, the lensing strength of the cluster center, and the fairly high redshift $z_d = 0.4$ of the cluster make it absolutely necessary to apply an unbiased finite-field inversion technique, to use a non-linear reconstruction method, and to account for a redshift distribution of the galaxies.

Acknowledgements. We thank Achim Weiss for carefully reading, and very useful comments on our manuscript and an anonymous referee for his very constructive comments which helped improving the presentation. This work was supported by the ‘‘Sonderforschungsbereich 375-95 für Astro-Teilchenphysik’’ der Deutschen Forschungsgemeinschaft.

Appendix A1

Let us derive Eq.(3.4), first considering the case $|g(z)| \leq 1$. We define $\epsilon_s = y \exp(2i\varphi)$ and $u = \exp(2i\varphi)$, so that $d\varphi = -i du/u$, use Eq. (3.2) and begin with the φ -integration:

$$\begin{aligned} \int_0^{2\pi} d\varphi \epsilon^n(\epsilon_s, z) &= \int_0^{2\pi} d\varphi \left(\frac{\epsilon_s - g(z)}{1 - g^*(z)\epsilon_s} \right)^n \\ &= -i \oint du \left(\frac{yu - g(z)}{1 - g^*(z)yu} \right)^n \frac{1}{u} \\ &= 2\pi [-g(z)]^n . \end{aligned} \quad (\text{A1.1})$$

In the last step, we used that the integrand has one pole at $u = 0$ where the corresponding residuum is $[-g(z)]^n$; the other pole at $u = 1/(g^*y)$ lies outside the unit circle and thus does not contribute. As a result we obtain for $|g(z)| \leq 1$,

$$\langle \epsilon^n \rangle_{\epsilon_s}(z) = \int_0^1 dy y p_{\epsilon_s}(y) [-g(z)]^n = [-g(z)]^n . \quad (\text{A1.2})$$

Similarly, we derive for $|g(z)| > 1$:

$$\begin{aligned} \int_0^{2\pi} d\varphi \epsilon^n(\epsilon_s, z) &= \left[\int_0^{2\pi} d\varphi [\epsilon^*(\epsilon_s, z)]^n \right]^* \\ &= \left[\int_0^{2\pi} d\varphi \left(\frac{1 - g^*(z)\epsilon_s}{\epsilon_s - g(z)} \right)^n \right]^* \\ &= \left[-i \oint du \left(\frac{1 - g^*(z)uy}{yu - g(z)} \right)^n \frac{1}{u} \right]^* = 2\pi \left(\frac{-1}{g^*(z)} \right)^n \end{aligned} \quad (\text{A1.3})$$

since the pole at $u = g/y$ lies again outside the unit circle; we thus find that for $|g(z)| > 1$,

$$\langle \epsilon^n \rangle_{\epsilon_s}(z) = \int_0^1 dy y p_{\epsilon_s}(y) \left(\frac{-1}{g^*(z)} \right)^n = \left(\frac{-1}{g^*(z)} \right)^n . \quad (\text{A1.4})$$

Appendix A2

Here we derive approximations for $X_1(\kappa)$ and $X_2(\kappa)$ for the case of non-critical clusters. This non-criticality implies that $1 - w\kappa > 0$. We can expand the integral in (4.10), using

$$\langle w^n \rangle := \int_0^1 dw p_w(w) w^n = \int_{z_d}^{\infty} dz p_z(z) w^n(z) , \quad (\text{A2.1})$$

to obtain

$$\begin{aligned} X_1(\kappa) &= \int_0^1 dw p_w(w) \frac{-w}{1 - \kappa w} \\ &= - \int_0^1 dw p_w(w) w \sum_{k=0}^{\infty} (\kappa w)^k \\ &= - \sum_{k=0}^{\infty} \kappa^k \langle w^{k+1} \rangle . \end{aligned} \quad (\text{A2.2})$$

It is easily seen from (4.10) that for $n \geq 1$ the following relation holds:

$$X_{n+1}(\kappa) = - \frac{1}{n} \frac{dX_n(\kappa)}{d\kappa} . \quad (\text{A2.3})$$

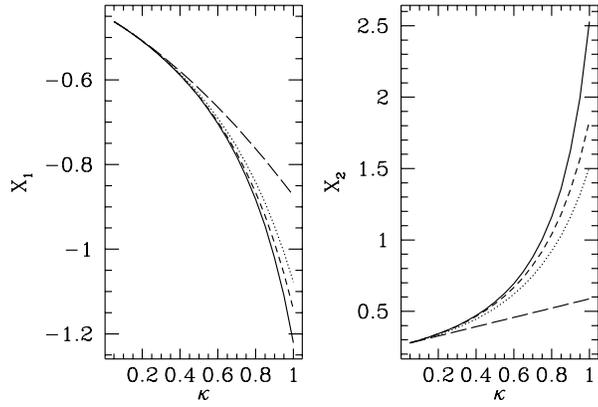


Fig. 8. Comparing different approximations for $X_1(\kappa)$ (left panel) and $X_2(\kappa)$ (right panel) as defined in Eq.(3.6) as a function of κ for the redshift distribution $p_z(z)$ shown in Fig. 1. The dotted lines show the approximations given in Eq. (A2.4), the dashed one those of Eq. (A2.5) and the long dashed line the expansion given in Eq. (A2.2) up to $k = 2$ as an approximation for $X_1(\kappa)$, and its first derivative as an approximation for $X_2(\kappa)$ ($X_2 = -\partial X_1/\partial \kappa$), respectively. The solid lines show the exact result for X_1 and X_2 obtained from performing the integration (4.10)

As it turns out, whereas the power series in (A2.2) converges, a Padé approximation behaves much better. From (A2.2) and (A2.3), we obtain for the lowest-order approximation (see e.g. Press et al. 1992, page 194 ff.)

$$\begin{aligned} X_1(\kappa) &\approx \frac{-\langle w \rangle}{1 - \kappa \frac{\langle w^2 \rangle}{\langle w \rangle}}, \\ X_2(\kappa) &\approx \frac{\langle w^2 \rangle}{\left(1 - \kappa \frac{\langle w^2 \rangle}{\langle w \rangle}\right)^2}, \end{aligned} \quad (\text{A2.4})$$

and for the next higher order,

$$\begin{aligned} X_1(\kappa) &\approx -\frac{\langle w \rangle + \kappa \left[\langle w^2 \rangle - \frac{\langle w^3 \rangle \langle w \rangle}{\langle w^2 \rangle} \right]}{1 - \kappa \frac{\langle w^3 \rangle}{\langle w^2 \rangle}}, \\ X_2(\kappa) &\approx \frac{\langle w^2 \rangle}{\left(1 - \kappa \frac{\langle w^3 \rangle}{\langle w^2 \rangle}\right)^2}. \end{aligned} \quad (\text{A2.5})$$

In Fig. 8, left panel, we compare $X_1(\kappa)$ with the approximations (A2.4) and (A2.5), and the expansion (A2.2) up to the term $k = 2$ for a cluster with $z_d = 0.4$ and the redshift distribution shown in Fig. 1 (i.e., for which $\langle z \rangle = 1$). We find that the approximations (A2.4) and (A2.5) are excellent for $\kappa \lesssim 0.6$, good for $0.6 \lesssim \kappa \lesssim 0.8$, and clearly deteriorate for $\kappa \gtrsim 0.8$. As expected, approximation (A2.5) is always better than (A2.4). In Fig. 8, right panel, we compare $X_2(\kappa)$ with the approximations given above. We find qualitatively the same result, but the deviation of $X_2(\kappa)$ from the approximations (A2.4) and (A2.5) becomes important for smaller values of κ , $\kappa \gtrsim 0.6$, say. We

note that these approximations become even better if the difference between mean redshift of the sources and the cluster redshift increases.

As a result, we obtain that the approximation (A2.4) for $X_1(\kappa)$ is sufficiently accurate over a wide range of κ (if κ approaches values close to 0.8, the cluster most likely is critical anyway). Therefore, one can use approximation (A2.4) for (most) non-critical clusters.

Appendix A3

In the case of a non-critical cluster (or at positions of a critical cluster where $\kappa + |\gamma| < 1$) the ratio

$$R := \left| \frac{\langle \epsilon^2 \rangle}{\langle \epsilon \rangle^2} \right| \quad (\text{A3.1})$$

becomes according (4.10)

$$R = X_2(\kappa)/X_1^2(\kappa) \quad (\text{A3.2})$$

Therefore, for a fixed redshift distribution of the galaxies, the ratio R depends on κ only. If we identify the expectation values again with the local averages,

$$R \approx \left| \frac{\bar{\epsilon}^2}{\langle \bar{\epsilon} \rangle^2} \right|, \quad (\text{A3.3})$$

a comparison of (A3.2) with (A3.3) can in principle determine the local value of κ . In Fig. 9 we show R for a non-critical cluster as a function of κ (solid line) for the redshift distribution shown in Fig. 1 ($\beta = 1 = \langle z \rangle$). The following approximations for $R(\kappa)$ can be derived from (4.11) and the relations in Appendix 2:

$$R(\kappa) \approx \frac{\langle w^2 \rangle}{\langle w \rangle^2} = f = R(0), \quad (\text{A3.4a})$$

for $\kappa \rightarrow 0$,

$$R(\kappa) \approx f \left[1 + \kappa \left(\frac{\langle w^2 \rangle}{\langle w \rangle} - \frac{\langle w^3 \rangle}{\langle w^2 \rangle} \right) \right]^{-2} \quad (\text{A3.4b})$$

good for $\kappa \lesssim 0.2$, and

$$R(\kappa) \approx f \frac{\left(1 - \kappa \frac{\langle w^2 \rangle}{\langle w \rangle}\right)^2}{\left(1 - \kappa \frac{\langle w^3 \rangle}{\langle w^2 \rangle}\right)^2} \quad (\text{A3.4c})$$

which is good for $\kappa \lesssim 0.8$. We obtain Eq. (A3.4a) using the approximations (A2.4) for $X_1(\kappa)$ and $X_2(\kappa)$. For Eq. (A3.4b) we use the approximations (A2.5) for X_1 and X_2 , and in the case of Eq. (A3.4c) we use the approximation (A2.5) for X_2 and the approximation (A2.4) for X_1 . We compare these approximations with $R = X_2(\kappa)/X_1^2(\kappa)$ in Fig. 9: for the chosen redshift distribution, (A3.4a) gives $R \approx 1.29$, approximation (A3.4b) is shown as the dotted line and is obviously a good approximation for small $\kappa \lesssim 0.2$, whereas approximation (A3.4c) is shown as the

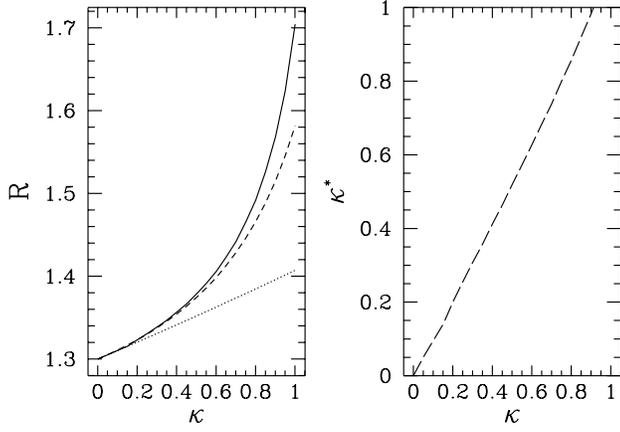


Fig. 9. The solid line (left panel) shows the ratio $R = |\langle \epsilon^2 \rangle| / |\langle \epsilon \rangle|^2$ as a function of the surface mass density κ in the case of a non-critical cluster, for which $R = X_2(\kappa)/X_1^2(\kappa)$. The dotted line shows the approximation (A3.4b) and the dashed line (A3.4c). We can invert (A3.4c) to derive the approximation (A3.5) for κ . The resulting values κ^* are compared to the true one κ in the right panel

dashed line and fits well up to $\kappa \lesssim 0.8$. For higher surface mass densities the cluster is most probably critical and Eq. (A3.2) does not hold anyway, but has to be replaced according to Eq. (3.5); then, R becomes a function of γ and κ . Eq. (A3.4c) can be inverted to obtain κ as a function of the observable quantity R ,

$$\kappa \approx \frac{\sqrt{\frac{R}{f}} - 1}{\frac{\langle w^3 \rangle}{\langle w^2 \rangle} \sqrt{\frac{R}{f}} - \frac{\langle w^2 \rangle}{\langle w \rangle}}. \quad (\text{A3.5})$$

In Fig. 9 (right panel) we compare the surface mass density κ derived from Eq. (A3.5) with the true one and find indeed a good agreement for $\kappa \lesssim 0.8$.

From Eq. (A3.5) one might conclude that $\kappa(\theta)$ can be directly obtained from observing $R(\theta)$ without using any of the inversion techniques derived in the past. Unfortunately this is not the case because $R(\theta)$ can not be determined precisely enough from local observations. To obtain a reliable estimate for a local surface mass density $\kappa \gtrsim 0.4$, one would need about 1000 galaxy images.

Thus, we try to derive the mean surface mass density $\bar{\kappa}$ by measuring $R(\theta)$: assuming a value for $\bar{\kappa}$ we perform the reconstruction according to (4.9) and derive from the mass and shear map the corresponding map of $\tilde{R}(\theta)$, which then depends on the assumed value for the mean mass density $\bar{\kappa}$. Finally, we compare this with the measured $R(\theta)$, i.e., we search for $\bar{\kappa}$ which minimizes $\chi^2 = \langle (\tilde{R}(\theta) - R(\theta))^2 \rangle_{\mathcal{U}}$, averaged over the data-field \mathcal{U} . Again, we find that the mean surface mass density can not be determined in that way, because the local mean image ellipticities $\langle \epsilon \rangle$ and $\langle \epsilon^2 \rangle$ can not be derived with sufficient accuracy given the observed number density of galaxy images even in very deep exposures.

We note that we tried to determine $\bar{\kappa}$ in various other ways from moments of image ellipticities: assuming a value for $\bar{\kappa}$ we

performed the reconstruction and calculated from the mass and shear map the local expectation values of $\langle \epsilon^n \rangle(\bar{\kappa})$ according to Eq. (3.5). Then, we compared that with the ‘measured’ local means $\bar{\epsilon}^n$ and minimized $\chi^2 = |\langle \epsilon^n \rangle(\bar{\kappa}) - \bar{\epsilon}^n|_{\mathcal{U}}^2$ varying the value $\bar{\kappa}$ used for the reconstruction. All these attempts failed for reasonable assumptions for the number density of galaxy images, showing that the image ellipticities provide not enough information to break the mass degeneracy in practice.

References

- Bartelmann, M. 1995, A&A, 303, 643.
 Bartelmann, M. & Narayan, R. 1995, ApJ, 451, 60.
 Bartelmann, M., Narayan, R., Seitz, S. & Schneider, P. 1996, ApJ Letters, submitted.
 Broadhurst, T.J., Taylor, A.N. & Peacock, J.A. 1995, ApJ 438, 49.
 Broadhurst, T.G. 1995, preprint astro-ph/9505013.
 Brainerd, T., Blandford, R. & Smail, I. 1995, preprint astro-ph/9503073.
 Infante, L. & Pritchet, C.J. 1995, preprint.
 Kaiser, N. & Squires, G. 1993, ApJ 404, 441 (KS).
 Kaiser, N. 1995, ApJ 493, L1.
 Kaiser, N., Squires, G., Fahlman, G.G., Woods, D. & Broadhurst, T. 1994, preprint astro-ph/9411029.
 Kneib, J.P., Ellis, R.S., Smail, I.R., Couch, W.J., Sharples, R., 1995, ApJ submitted.
 Kochanek, C.S. 1990, MNRAS 247, 135.
 Press, W.H., Teukolsky, S.A., Vetterlin, W.T. & Flannery, B.P. 1992, *Numerical Recipes*, Cambridge University Press.
 Miralda-Escude, J. 1991, ApJ 370, 1.
 Schneider, P. 1995, A&A 302, 639.
 Schneider, P. & Seitz, C. 1995, A&A 294, 411 (Paper I).
 Seitz, C. & Schneider, P. 1995, A&A 297, 287 (Paper II).
 Seitz, C., Kneib, J.P., Schneider, P. & Seitz, S. 1995, A&A, submitted.
 Seitz, S. & Schneider, P. 1996, A&A 305, 383.
 Smail, I., Hogg, D.W., Yan, L. & Cohen, J.G. 1995b, ApJ, 449, L105.
 Tyson, J.A., Valdes, F. & Wenk, R.A. 1990, ApJ 349, L1.