Multiscale cluster lens mass mapping – I. Strong lensing modelling

E. Jullo* and J.-P. Kneib

Laboratoire d’Astrophysique de Marseille, CNRS-Université de Provence, 38 rue F. Joliot-Curie, 13013 Marseille, France

Accepted 2009 February 17. Received 2009 January 27; in original form 2008 July 7

ABSTRACT

We propose a novel technique to refine the modelling of galaxy cluster mass distribution using gravitational lensing. The idea is to combine the strengths of both ‘parametric’ and ‘non-parametric’ methods to improve the quality of the fit. We develop a multiscale model that allows sharper contrast in regions of higher density where the number of constraints is generally higher. Our model consists of (i) a multiscale grid of radial basis functions with physically motivated profiles and (ii) a list of galaxy-scale potentials at the location of the cluster member galaxies. This arrangement of potentials of different sizes allows us to reach a high resolution for the model with a minimum number of parameters. We apply our model to the well-studied cluster Abell 1689. We estimate the quality of our mass reconstruction with a Bayesian Monte Carlo Markov Chain sampler. For a selected subset of multiple images, we manage to halve the errors between the positions of predicted and observed images compared to previous studies. This is due to the flexibility of multiscale models at intermediate scale between cluster and galaxy scale. The software developed for this paper is part of the public lenstool package which can be found at http://www.oamp.fr/cosmology/lenstool.

Key words: gravitational lensing – methods: numerical – galaxies: clusters: individual: Abell 1689.

1 INTRODUCTION

Non-baryonic dark matter (DM) is today commonly accepted as a predominant contributor to the matter density of our Universe. DM is indeed required to explain the velocity distribution of stars and gas in galaxies (e.g. Salucci 2001; de Blok, Bosma & McGaugh 2003) or the velocity dispersion of the galaxies in galaxy clusters (e.g. Czoske et al. 2002) but also to reproduce the large-scale galaxy distribution (e.g. Seljak et al. 2005; Guzzo et al. 2008) and the cosmological microwave background (CMB) fluctuations (e.g. the review by Hu & Dodelson 2002). A particularly striking demonstration of the need of DM is the direct detection of DM in the ‘Bullet Cluster’ (Bradač et al. 2006; Clowe et al. 2006) a cluster merger of two massive clusters where the X-ray gas and DM are spatially segregated due to the difference in nature of baryons and DM. However, beyond this general agreement about the need of DM, very little is known about its nature. The increasing amount of observational evidences steadily rules out possibilities. For instance, the CMB seems to rule out the warm DM (Spergel et al. 2003) and the measured extension of the truncation radius of galaxies in clusters seems to rule out fluid-like, strongly interacting DM at 5σ (Natarajan, De Lucia & Springel 2007). On another aspect, numerical simulations show that the collapsed non-collisional DM forms NFW density profiles (Navarro, Frenk & White 1997), with a central density peak. In contrast, if self-interacting particles are considered, this peak can be replaced by a profile with flat core (Spergel & Steinhardt 2000). On the observational side, measuring the slope of the DM density profile is still a hot topic either at the galaxy scale, with discrepancies in the rotation curves of stars (Gentile, Tonini & Salucci 2007; Valenzuela et al. 2007), or at the cluster scale, with discrepancies related to the slope of density profiles determined by gravitational lensing (Meneghetti et al. 2007; Limousin et al. 2008; Sand et al. 2008). Finally, numerical simulations with self-interacting particles predict less small-scale haloes than simulations with non-collisional particles. For a long time, the missing satellite problem in the Local Group was considered as evidence in favour of the self-interacting particles hypothesis. However, Simon & Geha (2007) have shown that this problem could merely be an observational issue. On the cluster scale, Natarajan et al. (2007) have shown that the mass function of substructures was in agreement with simulations with non-collisional particles, at least for a few strong lensing clusters. Estimating the mass distribution of cosmological objects with great accuracy is therefore a unique way to unveil the nature of DM.

Since the early 1990s, gravitational lensing has appeared as a robust tool to model the mass distribution of cosmological objects like galaxies, galaxy clusters and large-scale structures (Gavazzi et al. 2007; Limousin et al. 2007; Massey et al. 2007; Fu et al. 2008). With deep HST/ACS observations of massive clusters of galaxies, a large number of multiple images have been uncovered. In particular in the case of Abell 1689, Broadhurst et al. (2005) were the first to identify more than 100 multiple images, part of more than 30 multiple image systems. However, in several

*E-mail: eric.jullo@oamp.fr
recent works, strong lensing modelling has been unable to re- 
produce the numerous systems of multiple images observed in 
massive galaxy clusters with less than typically 1 arcsec resid- 
ual rms. In Abell 1689, Limousin et al. (2007) obtain an rms of 
2.87 arcsec for 34 systems of multiple images, Halkola, Seitz & 
Pannella (2006) report an rms of 2.7 arcsec and Broadhurst et al. 
(2005) an rms of 3.2 arcsec. In Abell 1703, Limousin et al. (2008) 
obtain an rms of 1.45 arcsec with 13 systems of multiple images and 
in Abell 2218 Eliasdottir et al. (2009) obtain an rms of 1.49 arcsec 
with eight systems. The physical origin of this systematic error is not 
yet fully understood. Do we miss invisible small-scale subhaloes in 
our models or are we badly reconstructing the large-scale mass dis- 
tribution? These large residual errors are likely highlighting a lack of 
resolution and/or flexibility in the lensing mass models. Indeed, 
mass models traditionally consist of an analytical density profile 
centred with respect to the light distribution and fitted to the posi- 
tions of the multiple images. In addition, Kneib et al. (1996) have 
shown that the complementary modelling of galaxy-scale haloes 
hosting bright cluster member galaxies significantly improves the fit 
(see Jullo et al. 2007 for a thorough description of the analytical 
modelling of clusters). In contrast to traditional ‘parametric’ mod- 
elling of galaxy clusters, partisans of ‘non-parametric’ models claim 
that their methods may allow a perfect fit, however, at the expense of 
sometimes unphysical solutions. In the case of the ‘non-parametric’ 
method, the mass distribution is generally tessellated into a regu- 
lar grid of small elements of mass called pixels (Saha & Williams 
1997; Diego et al. 2005). Alternatively, Bradač et al. (2005) pre- 
ter tessellating the gravitational potential because its derivatives 
directly yield the surface density and other important lensing quan-
tities. Point-like pixels can also be replaced by radial basis functions 
(RBF). RBF are real-valued functions with radial symmetry. Several 
density profiles for the RBF have been tested so far. Liesenborgs 
et al. (2007) use Plummer profiles and Diego et al. (2007) use RBF 
with Gaussian profiles. They also study the properties of power-law 
and isothermal profiles as well as Legendre and Hermite poly nomi-
als. They advise to use compact-like profiles such as the Gaussian 
or the power-law profiles, since too extended profiles produce a 
constant shear excess in the resulting surface mass density. Finally, 
instead of using a regular grid, Coe et al. (2008) and Deb, Gold- 
berg & Ramdass (2008) use the actual distribution of images as an 
irregular grid. Then, they either place RBF on this grid or directly 
estimate the derivatives of the potential at the images’ location. 
Whatever their implementation, the multiple images reproduction 
is generally greatly improved with respect to traditional ‘parametric’ 
modelling. However, the qualification of these models is still a mat-
ter of debate. Indeed, because of their large number of free param- 
ters with respect to the number of constraints, many different models 
can perfectly fit the data. In order to identify the best physically moti- 
vated model and eventually learn something on DM distribution in 
galaxy clusters, external criteria (e.g. mass positivity) or regular-
ization terms (e.g. to avoid unwanted high spatial frequencies) 
are required. In addition, galaxy mass scales are never taken into 
account, although traditional modelling has demonstrated that they 
effectively affect the positions of multiple images. This final step 
makes such ‘non-parametric’ models a little uncertain. None the 
less, ‘non-parametric’ models are useful because their large flex- 
ibility allows the exploration of the mass distribution regardless of 
any a priori assumptions. For example, these ‘non-parametric’ 
methods are efficient to reveal complex mass distribution such as 
found in the ‘Bullet Cluster’.

In this article, we study the properties of a model made of a 
multiscale grid of RBF and a sample of analytically defined galaxy-
scale DM haloes. We analyse how this model compares with a 
traditional ‘parametric’ model. We apply our analysis to the galaxy 
cluster Abell 1689 for its large amount of systems of multiple im-
ages. In Section 1, we present the analytic definition of our RBF. 
In Section 2, we evaluate the ability of our multiscale grid model 
reproducing a simple NFW profile. In Section 3, we use the 
mass model of Limousin et al. (2007) as an input to build a mul-
tiscle grid model of the galaxy cluster Abell 1689. In Section 4, 
we fit this model to a subset of multiple images and compare the 
produced mass map, deviation angle map and shear map to the 
one obtained with a traditional model. In addition, we perform an 
overfitting check. To do so, we assume that if a model opti-
imized with a subset of an image catalogue can accurately predict 
the rest of it, it does not overfit the data. Therefore, we compare 
the rms between predicted and observed images for the part of 
the image catalogue not used as constraints. Finally in Section 5, we 
study how different values of parameters related to the grid build-
ning affect the quality of the fit, the density profile of the cluster 
and the estimated properties of the galaxy-scale haloes. When nec-
ecessary, we use a flat Λ cold dark matter concordance cosmology 
with Ωm = 0.3 and H0 = 74 km s^{-1} Mpc^{-1}. At the redshift of 
Abell 1689 z = 0.184, an angular scale of 1 arcsec corresponds to 
2.992 kpc.

2 LENSING EQUATIONS

The lens equation,
\[ \beta(\theta) = \theta - \alpha(\theta). \]  
(1)

defines the transformation between the image position \( \theta \) and 
the source position \( \beta, \alpha(\theta) \) is the deflection angle due to the 
lens (e.g. Schneider, Ehlers & Falco 1992).

The amplification \( \mu \) of an image located in \( \theta \) is inversely propor-
tional to the determinant of the amplification matrix \( A \):
\[ \mu(\theta) = \frac{1}{|\det(A)|}. \]  
(2)

where the amplification matrix \( A \) is the derivative of the lens equa-
tion at the image location
\[ A = \frac{\partial \beta}{\partial \theta} = \begin{bmatrix} 1 - \kappa + \gamma & 0 \\ 0 & 1 - \kappa - \gamma \end{bmatrix}. \]  
(3)

defined here in the amplification basis. \( \kappa \) is the convergence and 
\( \gamma \) is the shear.

Through the Fermat Principle, it is possible to demonstrate that 
the deflection angle \( \alpha(\theta) \) is proportional to the gradient of the two-
dimensional Newtonian potential (Blandford & Narayan 1986)
\[ \alpha(\theta) = \frac{2}{c^2} D_{\Omega S} \nabla \phi(\theta), \]  
(4)

which in turn is related to the surface density \( \Sigma \) and the conver-
gence \( \kappa \) through the Poisson relation in two dimensions:
\[ 2 \frac{D_{\Omega S}^2}{c^2} \nabla^2 \phi(\theta) = 2 \frac{D_{\Omega L}}{D_{\Omega S}} 4\pi G \Sigma = 2 \frac{\Sigma}{\Sigma_{crit}} = 2\kappa. \]  
(5)

\( \Sigma_{crit} \) is the critical density above which strong lensing is possible. 
\( D_{\Omega L}, D_{\Omega S} \) and \( D_{\Omega S} \) are cosmological angular distances between 
the observer O, the lens L and the source S.

Deflection angles are additive quantities. For instance, if in a 
cluster we consider \( N \) clumps of mass located in \( \theta_i \), then each of 
them independently deflects a light beam crossing the cluster by an
angle $\alpha_i$. The total deflection angle computed at an observed image position $\theta$ is then

$$\alpha(\theta) = \sum_{i=1}^{N} \alpha_i(|\theta - \theta_i|).$$  \hspace{1cm} (6)

Let us now define the RBF used to build the multiscale grid. Its density profile is given by the analytical expression of the truncated isothermal mass distribution (TIMD), which is a circular version of the truncated Pseudo Isothermal Elliptical Mass Distribution (PIEMD, Kassiola & Kovner 1993; Kneib et al. 1996; Limousin, Kneib & Natarajan 2005; Elisadottir et al. 2009). The analytical expression of its surface density is

$$\Sigma(R) = \sigma^2 \, f(R, r_{\text{core}}, r_{\text{cut}})$$  \hspace{1cm} (7)

with

$$f(R, r_{\text{core}}, r_{\text{cut}}) = \frac{1}{2G} \frac{r^2_{\text{cut}}}{r^2_{\text{cut}} - r^2_{\text{core}}} \times \left( \frac{1}{\sqrt{r^2_{\text{core}} + R^2}} - \frac{1}{\sqrt{r^2_{\text{cut}} + R^2}} \right).$$  \hspace{1cm} (8)

$f$ defines the profile and $\sigma^2$ defines the weight of the RBF. Note that this profile is characterized by two changes in its slope marked by the $r_{\text{core}}$ and $r_{\text{cut}}$ radius (see fig. 1 in Jullo et al. 2007). Within $r_{\text{core}}$, the density is roughly constant, between $r_{\text{core}}$ and $r_{\text{cut}}$ it is isothermal $\Sigma \propto r^{-1}$ and beyond $r_{\text{cut}}$ it falls as $\Sigma \propto r^{-3}$. Compared to the Gaussian RBF profile used by Diego et al. (2007) or the Plummer RBF profile used by Liesenborgs et al. (2007), the TIMD profile has a shallower slope in the centre but falls in a steeper manner after $r_{\text{cut}}$, thus preventing the mass sheet excess noted by Diego et al. (2007) with the pure isothermal profile. In addition, the TIMD profile is physically motivated. Its total mass is finite as well as its central density. In this respect, this profile is more physical than the notorious NFW potential (Navarro et al. 1997) which fits non-collisional DM numerical simulations but has an infinite central density and an infinite total mass (see Limousin et al. 2005, in which TIMD and NFW potentials are compared). Finally, thanks to its flat core, the TIMD potential can produce extended flat regions, in particular in the centre of clusters if necessary.

3 THE MULTISCALE GRID

3.1 Definition and motivation

In this section, we detail how we build the multiscale model and demonstrate its capabilities in reproducing a singular NFW mass profile. For the moment, we do not include lensing constraints.

As proposed by Diego et al. (2005), we create a coarse multiscale grid from a pixelated input mass map and recursively refine it in the densest regions. Doing so, the huge range in mass observed in galaxy clusters is efficiently sampled with a minimum number of grid pixels. In contrast to many previous works, we do not use a squared grid but an hexagonal grid on the ground so that it better fits the generally rounded shape of galaxy clusters. With such a geometry, it is straightforward to generate a triangular mesh, which is the best way to pack a set of RBF.

In practice, we start by bounding the field of interest with a hexagon centred on the cluster centre. We split it into six equilateral triangles as shown in Fig. 1. Then, we choose a simple splitting criterion. We have tested several criteria: total mass; standard deviation or amplitude of surface density variations in a triangle and density of constraints, but none of them worked as well as the surface density threshold. Considering for instance an input mass map in an image with pixels of 1 arcsec$^2$, a triangle on this image is split into four subtriangles if it contains a single pixel (i.e. a region of 1 arcsec$^2$) that exceeds a user-defined surface density threshold. For instance, in order to trigger strong lensing regions, the threshold can be set equal to the critical density in $M_{\odot}$ arcsec$^{-2}$ at the cluster redshift. Overcritical regions will be split whereas subcritical regions will not. In the extreme case where the mass map is everywhere greater than the threshold, it results in a regular grid in which the number of triangles increases as $3 \times 2^n + 1$ and the number of triangle summits, or grid nodes, as $N = 1 + 3 \times (2^n + 2^2)$, where $n$ is the level of recursive splitting. Table 1 summarizes for some levels of splitting the maximum number of triangles and nodes a grid can contain. The level of splitting or equivalently the finest grid resolution ($\sim 2^{-n}$) is set by the user. As stated above, we have tested models where the grid is refined at the location of the multiple images. Similarly, as in Coe et al. (2008), such models lead to perfect fits to the data. However, we note that they also easily get biased by the chosen set of data.

Finally, RBF described by TIMD potentials are placed at the grid node location $\theta_i$. Their core radius is set equal to the size of the smallest nearby triangle and their cut radius is set equal to three times the core radius (this is discussed in Section 6). Their weights $\sigma^2_i$ are obtained by inverting the following system of $N$ equations:

$$\begin{bmatrix} M_{11} & \ldots & M_{1N} \\ \vdots & \ddots & \vdots \\ M_{N1} & \ldots & M_{NN} \end{bmatrix} \begin{bmatrix} \sigma^2_1 \\ \vdots \\ \sigma^2_N \end{bmatrix} = \begin{bmatrix} S_1 \\ \vdots \\ S_N \end{bmatrix}$$  \hspace{1cm} (9)

with

$$M_{ij} = f_j(|\theta_i - \theta_j|, r_{\text{core}}, r_{\text{cut}}).$$  \hspace{1cm} (10)

$S_i$ is the surface density read from the input mass map at the grid node location $\theta_i$. $M_{ij}$ is the value of a RBF with $\sigma^2 = 1$, centred on the grid node location $\theta_i$ and computed at a radius $R = |\theta_i - \theta_j|$. The product $M_{ij}\sigma^2_i$ gives the contribution of this RBF to the surface density $S_i$ (see equation 7).
3.2 Reproducing a NFW profile with a multiscale model

With enough resolution, multiscale models can reproduce all sorts of input mass distributions, provided of course that no slope steeper than $\Sigma \propto r^{-\alpha}$ (the steepest slope of the TIMD potential). As an exercise, let us consider the input mass map produced by a cluster at a redshift $z = 0.184$ whose density profile follows a circular NFW profile with a concentration of 5 and a scale radius of 150 arcsec. We compute a $200 \times 200$ arcsec$^2$ pixelated mass map of this object [i.e. about the field of view of the Advanced Camera for Surveys (ACS) on board the Hubble Space Telescope (HST)] to be used as an input mass map. To build the multiscale grid, we start with an hexagon whose centre matches the centre of the NFW mass distribution and whose radius is set to 100 arcsec. We set the splitting threshold to $4.7 \times 10^{10} M_{\odot}$ arcsec$^{-2}$ and limit to $n = 6$ the number of recursive splittings. Fig. 2 shows the produced multiscale grid. The smallest triangle, i.e. the smallest element of resolution, is $R_0 = 100$ arcsec/2$^6 = 1.54$ arcsec wide. The grid contains 229 nodes, to which we associate RBF with TIMD profile. Their weight $\sigma_i^2$ are obtained by inverting equation (9). Note, however, that we prefer an iterative method rather than a direct matrix inversion. Indeed, if we directly invert matrix $M$, we obtain a perfect fit but the solution vector contains negative $\sigma_i^2$, i.e. RBF with negative density profiles. Although Liesenborgs et al. (2008) allow some RBF to be negative, we are more conservative and prevent any element from the $\sigma_i^2$ vector from becoming negative. To do so, we minimize the following quantity:

$$Z = \sum_{j=1}^{N} \left[ S_{\text{input}}(\theta_j) - S_{\text{pred}}(\theta_j) \right]^2,$$

where $S_{\text{input}}$ and $S_{\text{pred}}$ are the input and predicted $S_j$ quantities of equation (9). With this iterative inversion technique, we can force the RBF to have positive weights $\sigma_i^2$, hence make sure that the overall surface density is positive. This way, we also avoid an additional regularization term (and possible related effects on the final results) to control the sign of the surface density in the lensing $\chi^2$ defined below in equation (12).

![Figure 2. Multiscale grid made of 229 RBF, mapping an input mass distribution with a NFW profile. The radius of the circles corresponds to the core radius of the RBF, i.e. locally the grid resolution. The central grey disc of radius $2 \times R_0$ with $R_0 = 1.54$ arcsec represents the central grid resolution beyond which mass map comparison is meaningless in Fig. 3.](image)

Fig. 3 shows the reconstructed radial density profile. Note how well the arrangement of RBF fits the input NFW density profile. The residual is lower than 5 per cent on the meaningful domain from twice $R_0$ to the hexagon inner limit. Note that a similar residual is also achieved with three levels of splitting, though on the corresponding meaningful domain. Therefore, a large number of splittings is not expressly justified unless the very centre of the mass distribution is of particular interest regarding some strong lensing constraints (e.g. radial arcs).

![Figure 3. Reproduction of the input NFW profile (in black) by a grid with six levels of splitting (solid red line) and three levels of splitting (dashed red line). The recovered profile with the six levels grid is the sum of the 229 RBF (dotted lines) shown in Fig. 2. The two vertical lines at $2R_0$ and $2R_3$ mark the lower bounds of validity for the six levels grid and the three levels grid, respectively. The dashed vertical line marks the hexagon limit at $100\sqrt{2}$ arcsec. Residuals show that both models present similar levels of errors (<5 per cent) within their respective domain of validity.](image)
lensing regions, within which $\kappa > 1$. Since the particular distribution and the shape of the observed multiple images depend on the mass distribution in these regions, it is important to give them as much flexibility as possible. In a first attempt, we stop the levels of recursive splitting at 3. We obtain a grid of 120 RBF shown in Fig. 4. The size of the smallest triangle hence our smallest element of resolution is $R_1 = 12.3$ arcsec. In comparison, Limousin et al. (2007) estimate the core radius of the cluster to be 33 arcsec which is more than twice $R_s$. According to the Shannon rule, we have therefore enough resolution to model it and reproduce central systems of multiple images. The weights $\sigma_i^2$ of the RBF are computed by iterative inversion of equation (9) as detailed previously.

Fig. 5 shows a disagreement between the convergence map produced by this model and the input one, above all at the location of galaxy-scale clumps. Consequently, the multiscale model underestimates by 9 per cent the mass inside the hexagon. Beyond the hexagon limit, note the increasing disagreement between the two models. This is a modelling artefact of the multiscale model due to the difference in slope between the RBF, which falls as $\Sigma \propto r^{-3}$, and the input profile slope, which falls as $\Sigma \propto r^{-1}$. Of course, such inaccuracies in the convergence map induce serious discrepancies in terms of shear and deflection angle. Fig. 5(c) shows large errors in the shear map at the location of galaxy-scale clumps, but smaller errors far from these regions. In contrast, Fig. 5(d) shows that in the deflection angle map discrepancies are weaker, with only 10 per cent of disagreement throughout. Note that the discrepancy observed at the very centre is due to the lack of resolution of the multiscale model. Therefore, we conclude that the lack of resolution of the multiscale model at the location of galaxy-scale clumps severely affects the lensing properties of the model.

Of course, more levels of splitting could improve the recovery, as demonstrated in the previous section with the NFW profile. However, here, with four levels of splitting, the number of RBF raises to 318 and with six levels it amounts to 4287. Optimizing grids with so many clumps is currently beyond our computational resources.

4.1 Adding galaxy haloes

Instead of blindly increasing the level of splitting, we better build a multiscale model including both a multiscale grid and galaxy-scale clumps. The grid thus becomes a flexible mass component to fit the cluster-scale distribution of mass, whereas galaxy-scale clumps fit small-scale irregularities. This appears as a cheap solution to our problem of resolution.

5 STRONG LENSING CONSTRAINTS ON MULTISCALE MODEL

5.1 Methodology

In this section, we investigate the ability to constrain such multiscale models with strong lensing and compare the results in terms of mass distribution and image prediction rms to a reference analytical model. This comparison is performed on the observational data of the galaxy cluster Abell 1689. First, we describe the Bayesian Monte Carlo Markov Chain (MCMC) sampler we use to optimize the parameters. Secondly, we present the multiscale and the reference models, their free parameters, assumed priors and the strong lensing constraints. Finally, we analyse the results.

5.2 Bayesian MCMC sampler

Given their large number of free parameters and the small number of constraints provided by strong lensing data, ‘non-parametric’ models are usually underconstrained, hence the regularization terms to obtain a smooth mass distribution (Marshall et al. 2002; Bradač et al. 2005; Suyu et al. 2006). In contrast, we do not use an explicit regularization term because our arrangement of RBF into a regular grid, and the smooth and everywhere positive TIMD profile describing
them, constitutes an intrinsic regularization scheme. In addition, equating the TIMD core radius of the RBF to the grid resolution impedes strong discontinuities in the produced mass maps. Finally, the overlap of nearby RBF correlates the model’s parameters and makes them dependent, thus reducing the effective number of free parameters. The number of constraints may then become larger than the number of effective parameters, and multiscale models may not be underconstrained (see Section 6).

In order to check our Bayesian MCMC sampler [based on the BAYESYS library (Skilling 2004) implemented in LENSTOOL.1 (Kneib et al. 1993; Jullo et al. 2007)], we generate a mock catalogue of 12 sources and 35 multiple images with the multiscale model described above. Then, we check that the mass distribution estimated from this catalogue matches the input one. The $\chi^2$ is computed in the source plane for simplicity and computation time. For each system $i$ containing $n_i$ multiple images, we define $\chi^2_i$ as

$$\chi^2_i = \sum_{j=1}^{n_i} \left( \frac{\beta_j - \langle \beta \rangle}{\mu_j \sigma_j} \right)^2,$$

where we use the scaling equation (equation 1) to compute the source position $\beta_j$ of the observed image $j$, $\langle \beta \rangle$ is the barycentre of the $\beta_j$, $\mu_j$ is the magnification for image $j$ and $\sigma_j$ is the observational error at the position of image $j$.

Fortunately, we detect no systematics in the recovery. The error between the recovered and the input density maps is about 1 per cent on average in the hexagon. The most outstanding result is that the algorithm converged towards the best-fitting region with only 3 $\times$ $10^8$ $\chi^2$ per second on average and computed only 1.17 $\chi^2$ points per dimension. Note that a standard gradient method would have required at least 3 $\chi^2$ points per dimension to find a minimum i.e. $3^{122} = 10^{58} \chi^2$ points.

5.3 Multiscale model and priors

We consider the multiscale model of Section 4 to which we add galaxy-scale clumps. The model then contains a multiscale grid of 120 RBF; its smallest element of resolution is $R_1 = 12.3$ arcsec and it has a ratio $r_{\text{cut}} / r_{\text{core}} = 3$. We fix the profile and positions of the RBF. Their weights $\sigma_j^2$ are the only free parameters. We allow them to vary along a wide flat prior defined in the range 10–$3.0$.

$\sum_{j=1}^n r_{\text{cut}} = 1.0$ and $\sum_{j=1}^n r_{\text{core}} = 0.1$.

$r_{\text{cut}} = r_{\text{core}}^* \left( \frac{L}{L^*} \right)^{1/2}$, $r_{\text{core}} = r_{\text{cut}}^* \left( \frac{L}{L^*} \right)^{1/2}$, $\sigma_0 = \sigma^*_0 \left( \frac{L}{L^*} \right)^{1/4}\text{ (13)}$

where $L^*$, $r_{\text{core}}^*$, $r_{\text{cut}}^*$ and $\sigma^*_0$ are, respectively, the luminosity and the three PIEMD parameters of a typical early-type galaxy at the cluster redshift. We consider a cluster galaxy with luminosity $L^*$ corresponding to an F775W magnitude of 17.54. We know from previous studies in this cluster that $r_{\text{cut}}^* \approx 13$ arcsec (Halkola, Seitz & Pannella 2007; Limousin et al. 2007). Still, we assume a flat prior between 1 and 30 arcsec in order to investigate the possible interactions between the grid and galaxy-scale clumps. Accordingly, we assume $\sigma_0^*$ vary along a wide flat prior defined in the range 10–400 km s$^{-1}$. Finally, recent studies have shown that the density profile of early-type galaxies in the field is singular up to observational limits (Koopmans et al. 2006; Czoske et al. 2008). Therefore, we fix $r_{\text{core}}$ to 0.03 arcsec (i.e. a small value but different from zero for numerical reasons). In total, this multiscale model sums 122 free parameters, whose priors are reported in Table 2.

5.4 Reference model

As a reference model, we use a modified version of the Limousin et al. model with two cluster-scale clumps (for the main and the north-east haloes), three galaxy-scale clumps described by individual PIEMD potentials to model the brightest cluster galaxy (BCG) and Galaxy 1 and Galaxy 2 that strongly affect Systems 6, 24 and Systems 1, 2, respectively, and the 60 galaxy-scale clumps of the multiscale model described by scaling relations. In total, the reference model sums 33 free parameters, whose priors are reported in Table 3.

5.5 Strong lensing constraints in Abell 1689

From Limousin et al. (2007), we select a catalogue of 28 images in 12 systems of multiple images (see Table 4). We will use the rest of the images later to check for model predictability and overfitting.

5.6 Computational considerations

On a 2.4 GHz processor, the Bayesian estimation of the 122 free parameters took 15 days to produce about 20 000 MCMC samples. Although it could be considered as quite computationally intensive, considering the number of free parameters and the fact that we not only find the best-fitting region but also explore the parameter space in its neighbourhood, it is in fact very efficient. In the last section, we discuss some issues related to the number of samples.
Table 2. Priors for the multiscale model.

| Number of clumps | $r_{\text{cut}}/r_{\text{core}}$ | Level of splitting | $\sigma_i$ | $r_{\text{core}}^*$ | $r_{\text{cut}}^*$ | $\sigma_0^*$ | $m_K^*$ |
|------------------|-------------------------------|------------------|-------------|----------------|-----------------|--------------|----------|
| 120 + 60         | 3                             | 3                | [0,1922] km s$^{-1}$ | 0.03 arcsec | [1,40] arcsec | [10,400] km s$^{-1}$ | 16 |

Table 3. Priors for the reference model.

| ID    | RA       | Dec.    | $\epsilon$ | $\theta$ (deg) | $r_{\text{core}}$ (kpc) | $r_{\text{cut}}$ (kpc) | $\sigma_0$ (km s$^{-1}$) |
|-------|----------|---------|-------------|-----------------|--------------------------|------------------------|-------------------------|
| Clump 1 | [−15,15] | [−15,15] | [0,1.0,0.55] | [0,180] | [30,150] | 1500 | [1000,1700] |
| Clump 2 | [−90,−35] | [5,79] | [0,0.4,0.9] | [0,180] | [25,90] | 500 | [300,650] |
| BCG    | [−10,10] | [−10,10] | [0,0.6] | [0,180] | [0,1.1,0] | [9,550] | [300,680] |
| Galaxy 1 | 49.0 | 31.5 | [0,0.9] | [0,180] | [0,1.3,0] | [9,180] | [150,280] |
| Galaxy 2 | [−49,−45] | [27,35] | [0,0.9] | [0,180] | [0,1.2,0] | [4,190] | [200,580] |
| $L^*$ elliptical galaxy | ... | ... | ... | ... | 0.15 | [20,60] | [150,280] |

Note. Coordinates are given in arcseconds with respect to the BCG (RA = 13:11:29, Dec. = −01:20:27). The ellipticity $\epsilon$ is that of the mass distribution, expressed as $a^2 - b^2/a^2 + b^2$.

Table 4. Multiply imaged systems used as constraints. Bold values correspond to averaged values for the whole system.

| ID    | RA       | Dec.    | $\langle \chi^2 \rangle$ | $z_{\text{spec}}$ | Image plane rms (arcsec) Multiscale model | Image plane rms (arcsec) Reference model |
|-------|----------|---------|--------------------------|-----------------|-----------------------------------------|---------------------------------------|
| 1.     | 13:11:26.44 | −1:19:56.37 | 1.91 | 3.0 | 0.71 | 0.67 |
| 1.1    | 13:11:29.76 | −1:21:07.31 | 1.28 | 2.5 | 0.28 | 0.59 |
| 2.     | 13:11:26.52 | −1:19:55.07 | 5.07 | 1.1 | 0.22 | 0.67 |
| 4.1    | 13:11:32.16 | −1:20:57.33 | 4.35 | 2.6 | 0.36 | 1.01 |
| 7.1    | 13:11:25.44 | −1:20:51.52 | 1.93 | 4.8 | 0.16 | 0.23 |
| 10.1   | 13:11:33.96 | −1:20:50.99 | 0.84 | 1.8 | 0.10 | 0.27 |
| 10.11  | 13:11:28.05 | −1:20:12.28 | 0.68 | 1.8 | 0.08 | 0.18 |
| 10.2   | 13:11:33.34 | −1:20:13.79 | 0.77 | 1.8 | 0.08 | 0.19 |
| 24.1   | 13:11:29.18 | −1:20:56.04 | 3.34 | 2.6 | 0.17 | 0.52 |
| 24.11  | 13:11:30.29 | −1:19:33.86 | 3.34 | 2.6 | 0.17 | 0.52 |
| 24.13  | 13:11:33.71 | −1:20:19.82 | 2.92 | 2.6 | 0.18 | 0.40 |
| 24.2   | 13:11:29.22 | −1:20:55.28 | 2.92 | 2.6 | 0.18 | 0.40 |
| 24.23  | 13:11:30.25 | −1:19:33.26 | 2.92 | 2.6 | 0.18 | 0.40 |
| 24.24  | 13:11:33.69 | −1:20:18.80 | 2.92 | 2.6 | 0.18 | 0.40 |
5.7 Results

5.7.1 Image prediction

Results confirm the ability of multiscale models at being used as lens models. Indeed, the rms per system and per image reported in Table 4 and shown in Fig. 7(a) highlight the good precision we obtain. The rms averaged over all the systems is 0.28 arcsec. In contrast, the reference model optimized with the same catalogue of images produces a mean rms of 0.54 arcsec. For all the systems except System 1, the rms with the multiscale model are lower than the rms with the reference model. System 1 has a slightly larger rms because the galaxy-scale clump located 3.2 arcsec to the west may not fit the same scaling relation as the other galaxy-scale clumps. Indeed, if we assume that two galaxies do not follow the same scaling relations, there is no perfect solution for the scaling relation parameters. The images producing the largest $\chi^2$ bias the scaling relations in their favour. The fit of images with a lower $\chi^2$ but close to galaxies following other scaling relations worsens. In this respect, note that for System 1 $\chi^2 = 1.91$ and rms = 0.71 arcsec whereas for System 24.1 $\chi^2 = 3.34$ and rms = 0.17 arcsec. Both systems are at the same distance of a galaxy-scale clump. System 24.1 is therefore likely biasing the scaling relations in its favour.

Given the large number of free parameters compared to the number of constraints, we could have expected better rms and smaller $\chi^2$. Indeed, it is usually accepted that models with more free parameters than constraints allow an infinity of perfect solutions. Since our solutions are not perfect, it means that the number of useful effective parameters in our model is actually lower than the number of constraints. To estimate this number, we analyse the distribution of the MCMC samples in the parameter space by means of the principal component analysis (PCA) technique. Fig. 6 shows that 90 per cent of the variance of the distribution is reproduced with only 12 effective parameters out of the 122 ‘real’ parameters. With the reference model, 90 per cent of the variance is reproduced with only 10 effective parameters out of the 33 ‘real’ ones. In both cases, the number of effective parameters is lower than the 32 constraints, hence the non-perfect fit. Though very interesting, investigating the physical meaning of these effective parameters is beyond the scope of this paper.

As a definite confirmation that our multiscale model is not underconstrained and does not overfit the data, we apply the cross-checking technique: with both the multiscale and the reference models, we predict the image positions for the part of the image catalogue not used as constraints and compare the rms between predicted and observed positions. This way, if a model is biased towards the subset of images used as constraints, it should be unable to give accurate predictions. Fig. 7(b) shows the rms given for the two models for the non-optimized part of the image catalogue. On average, we find an rms of 3.32 arcsec for the multiscale model and an rms of 3.49 arcsec for the reference model. Since the two models give similar predictions, we conclude that the multiscale model does not overfit the data.

In addition, Fig. 7(b) also shows that when images get closer to galaxies their rms increases. This increasing rms again suggests that galaxy-scale clumps do not perfectly fit the imposed scaling relations. There must exist a scatter in the scaling relations that images seem to be sensitive to, and that should be included in our future models.

5.7.2 Convergence, deviation and shear maps

Fig. 8 compares mean convergence, deviation and shear maps produced with the reference and the multiscale models. To produce these maps, we compute the convergence, deviation and shear maps for each mass model of the MCMC chain and compute the mean maps by averaging the values of each pixel of each map. First of all, note that the two convergence maps are very similar. The better rms obtained with the multiscale model does not originate from any particular missing clump in the reference model. The mass enclosed at the Einstein radius $M(< 45$ arcsec) is also very similar with less than 1 per cent difference. In other words, the ‘mass follows light’ assumption in traditional modelling holds. None the less, the better rms obtained with the multiscale model demonstrates a significantly higher degree of flexibility. In particular, we note that in the
convergence map the north-east clump, at intermediate scale between cluster and galaxy scale, looks smoother and more detached from the main clump than in the convergence map produced by the reference model. The deviation maps produced by the two models are also very similar with less than 5 per cent of difference at the image position. Finally, the errors in the shear maps are mostly lower than 10 per cent at the image position. This is currently below what we can observationally constrain by measuring the ellipticity of multiple images. Consequently, the better rms with the multiscale model is merely due to its large number of free parameters, which allow a refined modelling of the mass distribution irregularities at intermediate scale between cluster and galaxy scale.
5.7.3 Error mass map

In addition to the mean convergence map, the Bayesian approach also allows us to compute the standard deviation and the signal-to-noise ratio (S/N) convergence maps. Fig. 9 shows that the mass distribution is estimated with high confidence, since the S/N is everywhere larger than 10 inside the hexagon, i.e. less than 10 per cent error. This means that the correlations between the RBF parameters, highlighted by the PCA technique, must strongly restrict the range allowed by the priors, hence the small error. Furthermore, the smooth isomass contours confirm the ability of multiscale models at producing smooth mass maps.

6 DISCUSSION

In the previous section, we have worked at demonstrating the strength of our model. We are now going to highlight some critical aspects which will give the reader a more complete view of the multiscale models, but also of ‘non-parametric’ models in general, for which the effects of grid parameters on the final results are rarely investigated in detail. In this section, we particularly investigate two parameters: (i) the $r_{\text{cut}}/r_{\text{core}}$ ratio defining the RBF concentration and (ii) the level of splitting. We do not investigate the threshold parameter because we only aim at oversampling strong lensing regions. Therefore hereafter, in all the models, the surface density threshold is set equal to $\Sigma_{\text{cut}} = 2 \times 10^{10} \, M_\odot \, \text{arcsec}^{-2}$.

To begin with, we build five multiscale models similar to the one used so far, but with different $r_{\text{cut}}/r_{\text{core}}$ ratios and levels of splitting. Some models built with three levels of splitting have 127 clumps because the grid has been shifted by 6 arcsec with respect to the input mass map. We found that this shift does not affect the final results. We optimize each model with our catalogue of 28 images and report the mean rms and mean likelihood ($\log(L)$) = $-0.5 (\chi^2)$ in Table 5. Note also that fewer MCMC samples have been gathered with models A and C in order to evaluate how the number of MCMC samples affects the error estimation.

### 6.1 The $r_{\text{cut}}/r_{\text{core}}$ ratio

The $r_{\text{cut}}/r_{\text{core}}$ ratio characterizes the concentration of the RBF. Models with small ratios are more flexible because their RBF are more independent. Table 5 shows that such models better fit multiple images but can also overfit them. In contrast, models with larger ratios produce more extended and overlapping RBF, which worsens the fit to multiple images. For instance, model A with $r_{\text{cut}}/r_{\text{core}} = 2$ (i.e. with very concentrated RBF) gives a better fit than model D with $r_{\text{cut}}/r_{\text{core}} = 10$. If we calculate their likelihood ratio, we find $\log(L_{\text{D}}/L_{\text{A}}) = 113$, which means that model D is clearly ruled out.

In addition to the fit quality, the $r_{\text{cut}}/r_{\text{core}}$ ratio affects the extension of the RBF. The $r_{\text{cut}}$ extension parameter becomes large, in order for the galaxy-scale clumps to accommodate the mass distribution must then be very smooth with only a few very extended and effective RBF. The $r_{\text{cut}}$ extension parameter becomes large, in order for the galaxy-scale clumps to accommodate the mass. Besides, whatever the threshold is set equal to $10$ inside the hexagon, i.e. less than 10 per cent error. This means that the correlations between the RBF parameters, highlighted by the PCA technique, must strongly restrict the range allowed by the priors, hence the small error. Furthermore, the smooth isomass contours confirm the ability of multiscale models at producing smooth mass maps.

### 6.2 The level of splitting

Another prior involved in the building of the grid is the level of splitting. Fig. 13 shows that in contrast to the $r_{\text{cut}}/r_{\text{core}}$ ratio the

---

**Table 5. List of multiscale models used in this section.**

| ID | $r_{\text{cut}}/r_{\text{core}}$ ratio | Number of splitting levels | Number of clumps | Number of samples | $M_{\text{gal}}/M_{\text{tot}}$ ratio (per cent) | Mean image plane rms (arcsec) | $\langle \log(L) \rangle$ |
|----|-------------------------------------|---------------------------|------------------|------------------|-----------------------------------------------|-------------------------------|-----------------|
| A  | 2                                   | 3                         | 127              | 439              | 14                                            | 0.23                          | −14.8           |
| B  | 3                                   | 3                         | 120              | 20 920           | 13                                            | 0.27                          | −13.4           |
| C  | 4                                   | 3                         | 127              | 369              | 12                                            | 0.33                          | −28.2           |
| D  | 10                                  | 3                         | 127              | 25 029           | 11                                            | 0.86                          | −228.2          |
| E  | 3                                   | 4                         | 318              | 38 399           | 5                                             | 0.22                          | −13.1           |
| Reference | —                                | —                         | —                | 1000             | 22                                            | 0.57                          | −59.4           |
Multiscale cluster lens mass mapping

Figure 10. Confidence intervals at 68 and 99 per cent of the scaling relation parameters $r_{\text{cut}}^*$ and $\sigma^*$ obtained with models A, C and D. Dashed lines show the curves of constant $M/L$ ratio within a 60 arcsec aperture (i.e. the mean distance from multiple images to galaxy-scale clumps). In black is shown the same curve but for the Limousin et al. model. The vertical dashed limit at $2R_3 = 24.6$ arcsec marks the grid resolution of model A. The grid resolution of models C and D at $4R_3$ and $10R_3$, respectively, is beyond the abscissa scale limit of this plot. This figure shows that galaxy-scale clumps efficiently increase the grid resolution.

Figure 11. Histograms of the RBF velocity dispersions in models A, B, C and D, as distributed before strong lensing optimization (these distributions slightly shrink after optimization, i.e. small values get larger and large values get smaller). Note that multiscale models are mainly made of RBF with small velocity dispersion. As RBF get more extended, more RBF are assigned small velocity dispersions.

Figure 12. Convergence profiles obtained with models A, C and D. Larger $r_{\text{cut}}/r_{\text{core}}$ ratios make the profile shallower. The vertical arrows radially locate the strong lensing constraints.

Figure 13. Density profile comparison between multiscale models obtained with three and four levels of splitting, hence made of 120 and 318 RBF, respectively, and 60 galaxy-scale clumps. The vertical arrows radially locate the strong lensing constraints. In this region, the density profiles are similar. Outside, they are mostly driven by the priors.

Fig. 14 shows that the level of splitting affects the scaling relation parameter $r_{\text{cut}}^*$. With model E, $r_{\text{cut}}^* \simeq 15$ arcsec whereas with model B $r_{\text{cut}}^* \simeq 25$ arcsec. In addition, it shows that with model E $r_{\text{cut}}^*$ is close to the grid resolution $3R_4 = 18.4$ arcsec, whereas with model B it is significantly smaller, since $3R_3 = 36.9$ arcsec. It seems therefore that by increasing the level of splitting, we just replace the galaxy-scale clumps by the RBF of the grid. The large errors in model E indeed indicate that the galaxy-scale clumps are not constrained anymore by the data. Besides, Fig. 14 and Table 5 show that the $M/L$ ratio decreases as the level of splitting increases. The contribution of the galaxy-scale clumps to the total mass decreases from 13 per cent in model B to 5 per cent in model D. In other words, with
four levels of splitting, galaxy-scale clumps do not help anymore in increasing the model resolution. The scaling relation inflexibility might impede them from decreasing in size, to make the model gain in resolution and improve the fit. Again, one solution could be to introduce a scatter in the scaling relations. We will explore this solution in a forthcoming paper.

6.3 Invariant quantity

So far, we have found that both the $r_{\text{cut}}/r_{\text{core}}$ ratio and the level of splitting affect the estimation of the scaling relation parameters $r_{\text{cut}}^*$, $\sigma^*_i$ as well as the $M/L$ ratio of the galaxy-scale clumps. In this context, studying the physical properties of the galaxy-scale clumps seems a little risky. Nevertheless, it is possible to get more reliable values by making measurements directly on the surface density maps. In Fig. 15, we compute the aperture mass of galaxy-scale clumps with SExtractor (Bertin & Arnouts 1996) and pixelated mass maps obtained for models A, B, C, D and E. We choose SExtractor because it provides a multithreshold algorithm to assign the mass in a pixel to the most credible of two nearby galaxies. We find that the masses enclosed in a 4 arcsec aperture (i.e. the smallest distance between two nearby galaxy-scale clumps) are almost unaffected by the grid parameters. The scatter is of the order of 10 per cent, in agreement with the errors found in the density profiles. Estimating the galaxy-scale clump properties directly from the surface density maps seems therefore a more reliable and better constrained solution (given the priors) than simply considering the $r_{\text{cut}}^*$ and $\sigma^*_i$ modelling parameters.

6.4 Error analysis

As already stressed, given the size of the parameter spaces, it is encouraging to see our Bayesian MCMC sampler converging to the best-fitting region in less than 5000 samples. However, Fig. 14 shows that the error bars increase between model B and D. Since model D sums twice as many samples as model B, it seems that the estimation of the errors depends on the number of computed}

![Figure 14](https://example.com/figure14.png)

**Figure 14.** Confidence levels at 68 and 99 per cent of the scaling parameters $r_{\text{cut}}^*$ and $\sigma^*_i$ recovered with multiscale models B (green) and D (blue). Dashed lines show the curves of constant $M/L$ ratio within a 60 arcsec aperture. Dashed limits at $3R_e$ and $3R_e$ mark the extension of the smallest RBF for models B and D, respectively. Note that with four levels, the galaxy-scale clumps are completely unconstrained because they directly compete with the RBF.

![Figure 15](https://example.com/figure15.png)

**Figure 15.** Galaxy-scale clump mass measurements within a 4 arcsec aperture, performed on pixelated mass maps resulting from the fit of models A, B, C, D and E to the same subset of multiple images. Measurements are very consistent with a scatter of about 10 per cent, in agreement with the errors observed on the radial density profiles.

![Figure 16](https://example.com/figure16.png)

**Figure 16.** Evolution of the cumulative relative error of the parameters of model B as a function of the number of MCMC samples. The error does not stabilize even after more than 20 000 samples. To study the evolution of the error estimation with the number of MCMC samples, we define the cumulative relative error quantity computed over all the free parameters of the model as

$$\text{cumulative parameters error} = \sqrt{\sum_i \left( \frac{\sigma[X_i]}{E[X_i]} \right)^2},$$

where $\sigma[X_i]$ and $E[X_i]$ are the standard deviation and the mean value of the $X_i$ MCMC random variable for parameter $i$. Fig. 16 shows that this error has not converged even after more than 20 000 samples. Similarly, we define the error on the density profile as

$$\text{density profile error} = \sqrt{\int_0^{100} \left( \frac{\sigma[S_i]}{E[S_i]} \right)^2 \, dR},$$

where $S_i$ is a random variable for the radial density profile integrated in the range 0–100 arcsec. In contrast to the error on the
or 3 for instance, are difficult to estimate accurately and might actually be underestimated.

7 CONCLUSION

In this paper, we present a multiscale model sufficiently flexible to reproduce the observed systems of multiple images with high accuracy, but also robust against overfitting. The model combines a grid of RBF, and galaxy-scale clumps hosting cluster member galaxies, described by PIEMD potentials and scaling relations.

We apply this model to the galaxy cluster Abell 1689. We constrain the model with a subset of multiple images extracted from the Limousin et al. (2007) catalogue. We obtain an rms between observed and predicted positions of 0.28 arcsec in the image plane, i.e. about half the rms obtained with a slightly modified version of the Limousin et al. model (the reference model). We confirm the predictability of our multiscale model by cross-checking the optimized model with the part of the image catalogue not used as constraints. We find similar rms with the multiscale and the reference models. This confirms that we do not overfit the data, despite the large number of free parameters. We propose a PCA technique to estimate the effective number of free parameters. We find that our multiscale model with 122 parameters contains only 12 effective parameters. Our multiscale model produces smooth mass maps and radial density profiles.

Then, we compare the convergence, deviation and shear maps obtained with both the multiscale and the reference models. We find no major difference between the two sets of maps. This means that the better rms is due to minor changes allowed by the flexibility of the multiscale model at intermediate scale between cluster and galaxy scale. We also note that the galaxy-scale clumps are better integrated to the north-east cluster-scale clump with the multiscale grid than with the traditional reference model. The better modelling of this cluster-scale component, embedded in projection in the main cluster core, thus illustrates how multiscale models can reproduce irregular mass distributions.

Finally, we study how changes on the grid parameters affect the density profile, the scaling relation parameters $\sigma_{\text{g}}$, $r_{\text{cut}}^*$ and the M/L ratio of the galaxy-scale clumps. We find that galaxy-scale clumps efficiently increase the resolution of the grid of RBF without modifying the density profile. Excessively raising the level of splitting in order to increase the grid resolution is thus not expressly justified. We also find that the scaling relation parameter $r_{\text{cut}}^*$, related to the galaxy-scale clump extension, decreases as the RBF get more concentrated. This degeneracy leads to several models producing the same radial density profile. Fortunately, the likelihood is also significantly affected, hence allowing an effective model selection. In spite of this degeneracy between the model’s parameters, we propose a reliable solution to measure the galaxy-scale clump mass in an aperture directly on the mass maps produced by the optimized models. Indeed, we note that the model’s parameters degenerate, but the resulting mass distributions are little affected by the priors on the grid.

This work raises the issue of accurately measuring substructure properties in galaxy clusters (Halkola et al. 2007; Natarajan et al. 2007; Smith & Taylor 2008). The degeneracy between galaxy-scale and cluster-scale mass components can potentially lead to spurious conclusions. In traditional modelling of observationally quiet clusters, whatever the priors on the adopted profile for the cluster-scale component (e.g. PIEMD or NFW profile), the substructure properties are little affected, given the constraints. However, in unrelaxed perturbed clusters, pursuing this study with the same technique appears to be more risky. Our multiscale model as described above offers an interesting solution.

In terms of computation time, this method is still slow. We could make it faster by restricting the number of gradients and Laplacians calculated per image to solely the most significant. However, this has to be treated with care since the sum of negligible gradients can become significant. Missing such gradation could lead to spurious results. Refer for instance to Deb et al. (2008) to understand how complex it is to accurately select the relevant clumps for gradient calculation.

Multiscale models open interesting avenues for the modelling of galaxy clusters. Indeed, today standard parametric methods are facing their own limitations. New lensing results reveal irregular or merging clusters (e.g. Bradač et al. 2006; Jee et al. 2007; Mahdavi et al. 2007). Multiscale models with RBF provide the robustness of parametric methods as well as the flexibility of non-parametric methods. Since we use analytical potentials, the extension of the method to weak lensing is straightforward. We are currently working at combining strong and weak lensing signals to extend the accuracy of our modelling to larger radius. In inner regions, the strong lensing constraints will need a dense network of RBF to be reproduced accurately whereas in the outskirts a coarser sampling will perfectly fit with the lower density of weak lensing constraints (see Jullo et al., in preparation). This method can also be extended to multiplane lensing, allowing thus tomographic analysis.

ACKNOWLEDGMENTS

The authors thank Phil Marshall, Harald Ebeling, Marceau Limousin and Johan Richard, as well as the referee for enlightening and fruitful comments. J-PK and EJ acknowledge support from CNRS and SL2S grant 06-BLAN-0067 from the Agence National de la Recherche.

REFERENCES

Bertin E., Arnouts S., 1996, A&AS, 117, 393
Blandford R., Narayan R., 1986, ApJ, 310, 568
Bradač M., Schneider P., Lombardi M., Erben T., 2005, A&A, 437, 39

Figure 17. Evolution of the relative error on the density profile of model B as a function of the number of MCMC samples. The error stabilizes soon after 1500 samples.
This paper has been typeset from a \TeX/LaTeX file prepared by the author.