relensing: Reconstructing the mass profile of galaxy clusters from gravitational lensing

Daniel A. Torres-Ballesteros,1,* and Leonardo Castañeda1†
1Observatorio Astronómico Nacional, Universidad Nacional de Colombia, Carrera 30 Calle 45-03, P.A. 111321 Bogotá, Colombia

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
In this work we present relensing, a package written in python whose goal is to model galaxy clusters from gravitational lensing. With relensing we extend the amount of software available, which provides the scientific community with a wide range of models that help to compare and therefore validate the physical results that rely on them. We implement a free-form approach which computes the gravitational deflection potential on an adaptive irregular grid, from which one can characterize the cluster and its properties as a gravitational lens. Here, we use two alternative penalty functions to constrain strong lensing. We apply relensing to two toy models, in order to explore under which conditions one can get a better performance in the reconstruction. We find that by applying a smoothing to the deflection potential, we are able to increase the capability of this approach to recover the shape and size of the mass profile of galaxy clusters, as well as its magnification map. This translates into a better estimation of the critical and caustic curves. The power that the smoothing provides is also tested on the simulated clusters Ares and Hera, for which we get an rms on the lens plane of ~ 0.17 arcsec and ~ 0.16 arcsec, respectively. Our results represent an improvement with respect to reconstructions that were carried out with methods of the same nature as relensing. At the same time, the smoothing also increases the stability of our implementation, and decreases the computation time. In its current state, relensing is available upon request.

Key words: gravitational lensing – galaxy cluster – cosmology

1 INTRODUCTION
Gravitational lensing has proven to be a powerful tool for understanding the universe at different scales (e.g. Blandford & Narayan (1992); Schneider et al. (2006)), where its power relies upon its simplicity, since it only depends on the mass distribution that acts as lens, irrespective of its nature (baryonic or not) or its dynamical state. It is of particular interest its application to the study of galaxy clusters and their mass distribution. As the most massive objects in the universe, galaxy clusters provide a framework for cosmological probes, such as testing the underlying gravitational theory (e.g. Lam et al. (2012); Pizzuti et al. (2016); Cataneo & Rapetti (2018)) and cosmological model (e.g. Gilmore & Natarajan (2009); D’Aloisio & Natarajan (2011); Julio et al. (2010)), and also provide strong constraints to the paradigm of structure formation and evolution (e.g. see Allen et al. (2011); Kravtsov & Borgani (2012) for reviews). In addition, galaxy clusters also act as effective cosmic telescopes, since their angular size and mass turn them into the biggest and strongest manifestation of gravitational lensing in the universe. Therefore, galaxy clusters are capable of magnifying faint background sources that would otherwise remain unseen, or for which it would be difficult to infer any significant information. This is for example the case for galaxies with high redshift (the first galaxies ever formed in the universe) (e.g. Kneib et al. (2004); Bradač et al. (2009); Coe et al. (2013); Richard et al. (2014); Oesch et al. (2015); McLeod et al. (2016); Furtak et al. (2021)). For a deeper description of the role that gravitation lensing plays in the context of galaxy clusters refer e.g. Kneib & Natarajan (2011); Hoekstra et al. (2013); Umetsu (2020) and the references therein.

The problem of making an estimation of the mass profile of a galaxy cluster from gravitational lensing has been approached from different perspectives over the years. As a result, several methods have been developed. In general terms, they are usually classified as parametric or free-form (sometimes also labeled as non-parametric) methods. Parametric methods are characterized by employing an input model that depends on a set of parameters which needs to reproduce the input data as well as possible. This can be done for example, by means of Bayesian methods such as MCMC or Nested Sampling. The input model can be for example, the mass distribution composed of a superposition of baryonic and dark matter components for the different cluster members, and which are chosen taking into account certain physical and geometrical properties that are expected to satisfy the properties shown by the galaxy cluster. On the other hand, free-form methods generally work on a grid on which some quantity of interest is defined, for example the convergence or the deflection potential, and they need to be found. Therefore, free-form methods are parametric on the inside, since there is a set of parameters that need to be found in order to achieve the reconstruction, but they are different to pure parametric methods in the sense that does not exist an imposed input model that constrains the reconstruction.

Among parametric approaches we can find e.g. LENSTOOL Kneib

* E-mail: daatorresba@unal.edu.co
† E-mail: lcastanedac@unal.edu.co

© 2021 The Authors
et al. (1996); Jullo et al. (2007); Newman et al. (2013); Schäfer et al. (2020), Light-Traces-Mass (LTM) Broadhurst et al. (2005), and glafic Oguri (2010). Meanwhile, for free-form approaches we have e.g. Bartelmann et al. (1996), Abdelsalam et al. (1998), Saha & Williams (1997), Seitz et al. (1998), Cacciato et al. (2006), Jee et al. (2007), Deb et al. (2008), GRALE Lieneborgs et al. (2006, 2007, 2009, 2020); Ghosh et al. (2020), WSLAP Diigo et al. (2005, 2007), SWLensed Bradač et al. (2005a, b, 2009), LensPerfect Coe et al. (2008), and SaWLens (2) Merten et al. (2009, 2011); Merten (2016); Huber et al. (2019). There are also available hybrid methods that combine elements of both approaches e.g. Julio & Kneib (2009); Niemiec et al. (2020) (which are part of LINESTOOL), and WSLAP+ Sendra et al. (2014); Diigo et al. (2016) (an extension of WSLAP).

Nowadays, the relevant observations in this regard are open access, as it is the case for Hubble Frontier Fields (HFF) Lotz et al. (2017), Reionization Lensing Cluster Survey (RELICS) Coe et al. (2019), Cluster Lensing and Supernova Survey with Hubble (CLASH) Postman et al. (2012). Additionally, there is an expected increase in both quality and quantity of observations from upcoming ground/space based observatories such as James Webb Space Telescope (JWST) Gardner et al. (2006), Euclid Laureijs (2009); Laureijs et al. (2011), and Vera C. Rubin Observatory (LSST) Ivezić et al. (2019). Therefore, it is pertinent to have access to a variety of open access software that aims to model galaxy clusters from gravitational lensing, in order to be able to compare, and validate those models, as well as the physical results that rely on them (e.g. Priewe et al. (2017); Meneghetti et al. (2017); Straut et al. (2018); Remolina González et al. (2018)).

Therefore, in this work we describe what is behind relensing, an easy-to-use package currently written in python, in which a free-form reconstruction method is implemented. The method that we discuss here is based upon the work previously presented by Bradač et al. (2005a, b, 2009), which itself is an extension of Bartelmann et al. (1996).

We explore the effectiveness of an alternative finite difference approach. Also, we extend the grid refinement to an adaptive irregular one, that is intended to follow the observational data from strong lensing as well as the shape of the main deflectors. Additionally, we discuss two different approaches to the strong lensing penalty function, which along with the introduction of a smoothing process enhance the performance of this approach, reducing the computation time and providing a more accurate reconstruction.

This paper is organized as follows. In Sec. 2 we discuss the relevant aspects of gravitational lensing theory, and also introduce the notation used throughout the paper. Then, in Sec. 3 we describe the reconstruction method in detail. Once the reconstruction framework is settled, in Sec. 4 we apply the reconstruction to two simple distributions, in order to explore how relensing behaves and how to get the best out of it. In Sec. 5 we apply the reconstruction to Ares and Hera (see Meneghetti et al. (2017)), which provide a way of testing the performance of relensing in a more realistic scenario. Finally, in Sec. 6 we present the conclusions of this work.

For the mock catalogues discussed in Sec. 4.1, and the reconstructions presented in this work, we consider a flat $\Lambda$CDM cosmology with density parameter $\Omega_{m,0} = 0.3$ and Hubble constant $H_0 = 70 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1}$.

2 GRAVITATIONAL LENSING NOTATION

In this work we focus on the gravitational lensing effect under the thin lens approximation. A presentation of this topic deeper than the one intended here, can be found, for example in Schneider et al. (1992, 2006).

In this context, the galaxy cluster with redshift $z_l$ takes the place of the lens, and it is characterized by the projected mass density $\Sigma = \Sigma(\theta)$, with $\theta$ being an angular position from the observer perspective. Then, the aim is to recover $\Sigma$ somehow. For this task, we need to compare observations to theory. Therefore, as observables we will make use of multiply imaged systems (strong regime) (e.g. Hattori et al. (1999); Schneider et al. (2006)), since galaxy clusters are rich in those systems, and we will also consider the ellipticity of weakly distorted and singly imaged background sources (weak regime) (e.g. Bartelmann & Schneider (2001); Schneider et al. (2006); Hoekstra (2013); Mandelbaum (2018); Umetsu (2020)).

The aforementioned observables can be characterized by means of the convergence $\kappa = \kappa(\mathbf{z}, \theta)$, the shear $\gamma = \gamma(\mathbf{z}, \theta)$, and also the reduced (or scaled) deflection angle $\alpha = \alpha(\mathbf{z}, \theta)$ (hereafter we will refer to it as deflection angle for simplicity), which are defined in terms of potential $\psi = \psi(\mathbf{z}, \theta)$. Here $\kappa$ is a direct estimator of the mass profile, since it is defined as

$$\kappa := \frac{\Sigma}{\Sigma_{cr}} = \frac{1}{2} \nabla^2 \psi$$

with $\Sigma_{cr} := \frac{c^2 D(z)}{4\pi G D(z)D(z_l, z)}$, where $\Sigma_{cr} = \Sigma_{cr}(z, z_l)$ is known as critic mass density, and it is defined in terms of the angular diameter distances from the observer to the lens $D(z_l)$, from the observer to source $D(z)$ of interest, and finally from the lens to such source $D(z_l, z)$. Thus, we refer to $\kappa$ as the mass profile. Similarly, $\gamma$ is a dimensionless complex quantity whose components are

$$\gamma_1 := \frac{\partial^2 \psi}{\partial \theta_1^2} - \frac{\partial^2 \psi}{\partial \theta_2^2}$$

and

$$\gamma_2 := \frac{\partial^2 \psi}{\partial \theta_1 \partial \theta_2},$$

with $\gamma := \gamma_1 + i\gamma_2$. In addition, $\alpha$ reads

$$\alpha := \nabla \psi.$$  

From (1) it is clear that once $\psi$ is known, the mass profile is recovered. Now, since $\psi$ depends on $z$ it is convenient to define a reference redshift $z_r > z_l$ so that given $\psi_r$, one can write $\psi = \psi_r + \kappa \gamma + \alpha$ in terms of $\psi_r$, as well as in terms of $\kappa_r$, $\gamma_r$ and $\alpha_r$, respectively. This transition is done by using the cosmological weight $Z = Z(z, z_r, z_l)$ defined as

$$Z := \frac{D(z, z_l)D(z_r)}{D(z)D(z_l, z_r)} U(z),$$

with

$$U(z) = \begin{cases} 1 & \text{if } z > z_l \\ 0 & \text{if } z \leq z_l \end{cases},$$

from which one gets $\psi = Z\psi_r$, $\kappa = Z\kappa_r$, $\gamma = Z\gamma_r$ and $\alpha = Z\alpha_r$. In this way, we only need to deal with $\psi_r$, which simplifies the problem significantly.

With respect to the observables, in the strong regime the multiply imaged systems can in principle be characterized by solving the lens equation

$$\beta = \theta - Z\alpha_r,$$

which relates the true angular position $\beta$ to the observed angular position $\theta$. Alternatively, it is simpler to work on the source plane, taking into account that the multiple images of a given system have to converge to a unique source.

On the other hand, with respect to the weak regime, distortions (shape, size and orientation) at this level are characterized by means of source ellipticity. Among the different definitions of ellipticity
be achieved by minimizing (10) with respect to every tentials, since they are the parameters to be found. This task can needs to be explicitly written in terms of the regime, and both \( \psi \) in Sec. 2. To do so, we implement the penalty function intrinsic or true ellipticity (e.g. Schneider et al. (2006)), the more suited for the task in hand is the complex ellipticity \( \epsilon = \epsilon(z, \theta) \), which is related to the source intrinsic or true ellipticity \( \epsilon_s \) by means of the expression
\[
e = \begin{cases} 
\frac{\epsilon_s + g}{1 + g^2} & \text{if } |g| \leq 1 \\
\frac{1 + g^2}{\epsilon_s + g^2} & \text{if } |g| > 1 
\end{cases},
\]
where the asterisk stands for complex conjugate, and the reduced shear \( g = g(z) \) is defined as
\[
g = \frac{Zy_\gamma}{1 - Zk_r}.
\]
Here \( \epsilon_s \) is not an observable, therefore (7) cannot be used directly. This problem is avoided by considering the average intrinsic ellipticity for several (as many as possible) sources in a local neighborhood to be \( \langle \epsilon_s \rangle = 0 \). As a result, for a given redshift the average observed ellipticity reads
\[
\langle \epsilon \rangle = \begin{cases} 
g & \text{if } |g| \leq 1 \\
1 & \text{if } |g| > 1
\end{cases},
\]
yielding a direct estimation of the local reduced shear, which allows us to directly contrast observations to theory at weak regime. Further details can be found in Schneider & Seitz (1995); Seitz & Schneider (1997).

3 RECONSTRUCTION METHOD

The aim of those methods derived from Bartelmann et al. (1996) and its extensions is to constrain the deflection potential \( \psi \) in a given grid by making use of gravitational lensing observations. No assumption about mass traces light is made.

Consider (in general) an irregular grid defined on the lens plane, which is composed of \( N \) nodes whose angular position is \( \theta_j \) (with \( j = 1, 2, \ldots, N \)). On each node a deflection potential \( \psi_j \) is assigned, and such potentials are constrained from the observables described in Sec. 2. To do so, we implement the penalty function
\[
\chi^2(\psi_j) = \chi^2_\alpha(\psi_j) + \chi^2_\gamma(\psi_j) + \chi^2_\kappa(R)(\psi_j) + \chi^2_\gamma(R)(\psi_j),
\]
where \( \chi^2_\alpha \) constrains the strong regime, \( \chi^2_\gamma \) constrains the weak regime, and both \( \chi^2_\kappa(R) \) and \( \chi^2_\gamma(R) \) are regularization terms. Here, (10) needs to be explicitly written in terms of the \( N \) deflection potentials, since they are the parameters to be found. This task can be achieved by minimizing (10) with respect to every \( \psi_k \) (with \( k = 1, 2, \ldots, N \)), which leads to \( N \) equations of the form
\[
\frac{\partial \chi^2_{W_k}}{\partial \psi_k} + \frac{\partial \chi^2_{\kappa(R)}}{\partial \psi_k} + \frac{\partial \chi^2_{\gamma(R)}}{\partial \psi_k} = 0,
\]
each of them with \( N \) unknowns. The system of equations is in general nonlinear, making it difficult to solve. Linearizing (if possible) such system simplifies the implementation and reduces the computation time, which is an advantage in practical terms, considering that reliable results are obtained. In Sec. 3.3 we explain how to perform such linearization.

Now, from Sec. 2 it is clear that the observables we are interested in can be characterized by \( \kappa, \gamma, \alpha \), which are defined in terms of either first or second order derivatives of \( \psi \). Since we are working on a grid, the most direct way to explicitly write these quantities in terms of \( \psi \) is by using finite differences. In this context, the advantage of finite differences over other methods lies in its capability to express the derivatives (of any order) of \( \psi \) as a linear combination of \( \psi_j \) (defined on the grid), and it also makes the linearization of (11) simple. In this direction, \( \kappa, \gamma, \alpha \), and \( \psi \) itself, evaluated at an arbitrary angular position \( \theta_i \) (it does not need to correspond to a node) can be written in terms of every \( \psi_j \) as
\[
\psi_i = \sum_{j=1}^N P_{ij} \psi_j,
\]
with \( n = 1, 2 \) representing the components of \( \gamma \) and \( \alpha \), respectively.

\[\text{The coefficients (or nodes weight) } P_{ij}, K_{ij}, G_{n,ij} \text{ and } D_{n,ij} \text{ depend on what form the grid has, i.e. if it is regular (e.g. Bradač et al. (2005a); Deb et al. (2008); Merten et al. (2009)) or irregular (e.g. Bradač et al. (2009); Merten (2016)), and also on how finite differences and the interpolation are carried out. In Sec. 3.1 we describe how we obtain such weights. Since } \theta_i \text{ is arbitrary (within the region of interest), equations (12) - (15) also work to interpolate.}

Once the different penalty functions in (10) are explicitly written in terms of \( \psi_j \), we have that their contribution to (11) after the corresponding linearization yields an expression of the form
\[
\frac{\partial \chi^2_{W_k}}{\partial \psi_j} = 2 \sum_{j=1}^N W_{k,j} \psi_j - 2 W_{k,j}^{(p)},
\]
where \( p = s \) stands for strong regime, \( p = w \) for weak regime, and \( p = \kappa = \gamma \) for both regularization terms. In Sec. 3.3 we describe how \( W_{k,j}^{(p)} \) and \( W_{k,j}^{(p)} \) are defined in each case. From (16) it is clear that (11) reduces to
\[
\sum_{j=1}^N W_{k,j} \psi_j = \mathcal{V}_k,
\]
where \( \mathcal{W}_{k,j} = \mathcal{W}_{k,j}^{(w)} + \mathcal{W}_{k,j}^{(s)} + \mathcal{W}_{k,j}^{(\kappa)} + \mathcal{W}_{k,j}^{(\gamma)} \) and \( \mathcal{V}_k = \mathcal{V}_k^{(w)} + \mathcal{V}_k^{(s)} + \mathcal{V}_k^{(\kappa)} + \mathcal{V}_k^{(\gamma)} \). Here the coefficient \( \mathcal{W}_{k,j} \) represent the \( j \)-th node weight for the \( k \)-th linear equation.

The deflection potential admits two transformations that introduce degrees of freedom in the reconstruction, which may affect different aspects of the results. On one hand, transformation
\[
\psi \rightarrow \psi' = \psi + \psi_0 + c \cdot \theta
\]
introduces three degrees of freedom, one through the constant term \( \psi_0 \), and two from the bidimensional constant vector \( c \). On the other hand, transformation
\[
\psi \rightarrow \psi' = (1 - \lambda) |\theta|^2 / 2 + \lambda \psi,
\]
responsible for the well known Mass Sheet Degeneracy (MSD for short), introduces an extra degree of freedom through the constant term \( \lambda \). Following Bradač et al. (2005a), we have that such degrees
of freedom can be fixed by keeping the deflection potential fixed on one node for each degree of freedom. Since we have four degrees of freedom, we will hold \( \psi_j = 0 \) fixed on each of the four corners of the grid if both transformations are relevant, or only on three of the four corners of the grid if the MSD in not considered to be a concern. This leaves us respectively with a \((N - 4) \times (N - 4)\) system of linear equations to be solved.

Observations of galaxy clusters have shown the existence of several systems of multiple images, and for many of them the redshift is known; at least up to some extent. When such redshifts are sufficiently diverse, in principle the MSD breaks. Under these conditions the MSD is not a concern, so that we are left only with three degrees of freedom from (18). Thus, in this case we will have a \((N - 3) \times (N - 3)\) system of linear equations. In this work this condition is met. Hence, as well as in Bradac et al. (2005a), we will not consider the MSD to be a problem, which is supported by the results that we get.

Now, before we describe which finite difference approach that we use, let us continue describing, in general, how the reconstruction method works. At this point, one might think that solving (17) is just enough for getting the reconstruction done. Unfortunately, this leads to a low resolution and under reconstructed profile. In this regard, Bradac et al. (2005a) found that two levels of iterations solve the problem. Therefore, keeping their notation, we will refer to them as the inner and outer levels. The inner level is in charge of computing the deflection potential iteratively, while the outer level is responsible for producing the grid refinement that produce a finer and adapted grid on which a new estimation of \( \psi \) takes place.

In general terms, a description of how this method works is presented in Algorithm 1.

Algorithm 1 Reconstruction

1: Input: Initial conditions and the initial guess \( \psi(0) \).
2: Compute \( \kappa(0), \gamma_1(0) \) and \( \gamma_2(0) \) from \( \psi(0) \).
3: while outer_end = False do
4: Compute the nodes weight.
5: while inner_end = False do
6: Compute the system of linear equations needed and solve it \(\rightarrow \psi(n)\).
7: Smoothing of \( \psi(n) \rightarrow \) New \( \psi(n) \) (optional).
8: Compute \( \kappa(n) \) from \( \psi(n) \).
9: if \( |k_j^{(n)} - k_j^{(n-1)}| \leq \) tolerance then
10: inner_end = True.
11: Compute \( \chi_2^2 \).
12: if \( \chi_2^2 \leq q(2N_{img}) \) then
13: outer_end = True. \( \textit{if} \) outer_end = False then
15: Apply grid refinement.
16: Compute \( \psi(m), \kappa(m), \gamma_1(m) \) and \( \gamma_2(m) \) on the new grid.
17: return \( \psi \).

This algorithm has as input (line 1) a set of parameters that are used along the reconstruction. Among them, we have the grid initial and refinement conditions (more details in Sec. 3.2), regularization weights \( \eta_x \) and \( \eta_y \) (more details in Sec. 3.3.3), whether constrains from both strong and weak regimes are used or only from the strong regime, and the type of \( \chi_2^2 \) that is going to be used (more details in Sec. 3.3.1). Last but not least, as input we have the initial guess \( \psi(0) \) for the deflection potential, from which we compute \( \kappa(0), \gamma_1(0) \) and \( \gamma_2(0) \) (line 2). These quantities are needed to initialize the reconstruction (more details in Sec. 3.3).

Once we have set the initial conditions, the outer level starts (line 3). At the beginning of each outer iteration we compute the node weights for every node and data point (line 4). Such weights come from the finite difference scheme and the interpolation (when required), and have to be computed once for each outer iteration. With these weights in hand, the inner level starts (line 5).

The inner level is in charge of computing the system of linear equations (17) and its solution, so that we get \( \psi(n) \); a new version of the deflection potential. Here, \( n \) represents the current inner iteration (line 6). The reconstruction might be quite noisy, (as we show in Sec. 4.3), nonetheless, the results in the overall reconstruction can be improved by introducing a smoothing on \( \psi(n) \). Such smoothing is the result of recomputing \( \psi(n) \) with equation (12), where for this case \( \psi_i \) represents the smoothed potential at the \( i \)-th node, while \( \psi_j \) comes from the solution of (17). This step is optional but advised, as we justify in Sec. 4.3 (line 7).

Then, from \( \psi(n) \) (smoothed or not) we compute the convergence \( k(i) \) on the grid (line 8) by using finite differences, such that, when the condition

\[
|k_j^{(n)} - k_j^{(n-1)}| \leq \text{tolerance}
\]

is satisfied for every node \( (j = 1, \ldots, N) \), (line 9), the inner level finishes its task (line 10). In (20) \( k^{(n-1)} \) represents the convergence from the previous inner iteration.

When the inner level finishes, we compute either the type 1 or type 2 penalty function \( \gamma_2 \) depending on the initial conditions (line 11). We have seen that optimal results are obtained when the reconstruction reaches \( \chi_2^2 \leq 2N_{img} \) with \( N_{img} \) being the total number of multiple images being used. This condition is intended to avoid an overfit, without compromising the quality of the reconstruction. With that in mind, we consider that once the condition \( \chi_2^2 \leq q(2N_{img}) \) (for \( q > 1 \)) is satisfied (line 12), the outer level as well as the reconstruction finish (line 13). Here we have taken \( q = 1.4 \), for which the reconstruction tends to satisfy \( (q - 1)(2N_{img}) \leq \chi_2^2 \leq q(2N_{img}) \).

If condition \( \chi_2^2 \leq q(2N_{img}) \) is not met, a new outer iteration is required (line 14). Thus, the grid refinement is performed (line 15) (more details in Sec. 3.2). Then, from the deflection potential result of the last inner iteration for the current outer iteration, we compute \( \psi(m), \kappa(m), \gamma_1(m) \) and \( \gamma_2(m) \) on this new grid (line 16) (more details in Sec. 3.1). They are required to initialize the next outer iteration (more details in Sec. 3.3), taking the place of \( \psi^{(m-1)} \), \( \kappa^{(m-1)} \), \( \gamma_1^{(m-1)} \) and \( \gamma_2^{(m-1)} \), or \( \psi(0) \), \( \kappa(0) \), \( \gamma_1(0) \) and \( \gamma_2(0) \) if we are running the first outer iteration. Here, \( m \) stands for the current outer iteration.

Finally, if no further outer iterations are required, as return we have the final deflection potential, which corresponds to the last \( \psi(n) \) that we know (line 17).

Due to the nature of the method, negative densities may appear. This happens mainly for blind reconstruction (i.e. \( \psi_j = 0 \) at every node), as well as in situations with little observational constraints. The negative values tend to appear towards the outskirts of the distribution. The presence of negative densities is also heavily related to the initial conditions. For instance, small values of \( \eta_x \) and \( \eta_y \) produce an overfit of the intrinsic noise in the reconstruction. Likewise, a too dense initial grid restrains the capability of the method to adapt to the data. One way to solve this problem (when it occurs), is to use as an initial guess a distribution other than the blind one, which, along with the smoothing reduce the noise in the reconstruction, particularly towards the outskirts.
3.1 Finite difference approach

In this regard, we follow the generalized finite difference approach (GFD) (e.g. Benito et al. (2001); Gavete et al. (2003); Benito et al. (2003)), with the difference that we use it not only for finding the derivatives on a grid from the deflection potential defined on such a grid, but also for interpolating the deflection potential and its derivatives at a given angular position \( \theta_0 \), in case it is not part of the grid.

Let us consider the \( Q \) nearest nodes (NN) (nearest neighbors) to \( \theta_0 \), as it is depicted in Fig. 1. On each of such NN we apply the Taylor series expansion to is corresponding deflection potential around \( \theta_0 \) up to second order derivatives. Therefore, for the \( i \)-th NN we get

\[
\psi_i = \psi_0 + h_1 \frac{\partial \psi_0}{\partial \theta_1} + k_1 \frac{\partial \psi_0}{\partial \theta_2} + \frac{h_1^2}{2} \frac{\partial^2 \psi_0}{\partial \theta_1^2} + h_1 k_1 \frac{\partial^2 \psi_0}{\partial \theta_1 \partial \theta_2} + \frac{k_1^2}{2} \frac{\partial^2 \psi_0}{\partial \theta_2^2},
\]

(21)

with \( h_i = \theta_i - \theta_{0,1} \) and \( k_i = \theta_{i,2} - \theta_{0,2} \). In case for example flexion is being considered in the reconstruction, higher order derivatives have to be taken in the expansion (21). Here \( X_1 = \psi_0, X_2 = \partial \psi_0/\partial \theta_1, X_3 = \partial \psi_0/\partial \theta_2, X_4 = \partial^2 \psi_0/\partial \theta_1^2, X_5 = \partial^2 \psi_0/\partial \theta_1 \partial \theta_2, \) and \( X_6 = \partial^2 \psi_0/\partial \theta_2^2 \) are the unknowns, which, following GFD, can be obtained by minimizing the penalty function

\[
X_{\text{GFD}}^2 = \sum_{i=1}^{Q} \left( \Psi_i \psi_i - W_i \right)^2,
\]

(22)

with \( W_i = W(h_i, k_i) \) being a weighting function, and

\[
\Psi_i = -\psi_i + X_1 + h_1 X_2 + k_1 X_3 + \frac{h_1^2}{2} X_4 + h_1 k_1 X_5 + \frac{k_1^2}{2} X_6.
\]

(23)

Therefore, when minimizing (22) with respect to the \( j \)-th unknown we get

\[
\frac{\partial X_{\text{GFD}}^2}{\partial X_j} = 2 \sum_{i=1}^{Q} W_i \frac{\partial \Psi_i}{\partial X_j} = 0,
\]

(24)

which leads to a system of linear equations of the form \( AX = b \), with

\[
A = \sum_{i=1}^{Q} W_i^2 \begin{pmatrix}
1 & h_i & k_i & \frac{h_1^2}{2} & h_1 k_i & \frac{k_1^2}{2} \\
h_i^2 & h_1 k_i & \frac{h_1^3}{2} & h_1 k_i^2 & \frac{k_1^3}{2} \\
k_i^2 & \frac{h_1^2 k_i}{2} & \frac{h_1 k_i^2}{2} & \frac{h_1^3 k_i}{4} & \frac{h_1 k_i^3}{4} \\
\frac{h_1^3}{4} & \frac{h_1^2 k_i}{2} & \frac{h_1 k_i^2}{2} & \frac{h_1^3 k_i}{4} & \frac{h_1 k_i^3}{4} \\
\frac{h_1 k_i^3}{4} & \frac{h_1 k_i^3}{4} & \frac{h_1 k_i^3}{4} & \frac{h_1 k_i^3}{4} & \frac{h_1 k_i^3}{4}
\end{pmatrix},
\]

(25)

and

\[
b = \sum_{i=1}^{Q} W_i^2 \psi_i \begin{pmatrix} 1, h_i, k_i, h_1^2/2, h_1 k_i, k_1^2/2 \end{pmatrix},
\]

(26)

while the elements of \( X \) were defined above. We find \( A^{-1} \) numerically, such that

\[
X_j = \sum_{i=1}^{Q} X_{ji} \psi_i,
\]

(27)

where the weight of the \( i \)-th NN for the \( j \)-th unknown is given by

\[
X_{ji} = W_i^2 \left( A^{-1}_{j1} + A^{-1}_{j2} h_1 + A^{-1}_{j3} k_1 + A^{-1}_{j4} \frac{h_1^2}{2} + A^{-1}_{j5} h_1 k_1 + A^{-1}_{j6} \frac{k_1^2}{2} \right).
\]

(28)

From (27) it is clear that only the weights of the \( Q \) NN are (in principle) different from zero. Here, from (28) it becomes straightforward to compute the weights required in equations (12) - (15).

Now, we have found from mock reconstructions that for \( Q \sim 9 \) the reconstruction still works, but for less NN the reconstruction tends to fail, since (25) becomes singular quickly. Under this condition, if the reconstruction converges, most likely the output will not be reliable. On the other hand, if \( Q \) is beyond \( Q \sim 36 \) the reconstruction flattens. We have seen that \( 16 \leq Q \leq 36 \) is good enough for the reconstruction to work properly, with optimal results for \( Q \sim 25 \), since the output is neither irregular in excess nor flattened in excess.

In this paper we use a quartic spline as \( W_i \), given by

\[
W_i = 1 - 6 \left( \frac{d_1}{dm} \right)^2 + 8 \left( \frac{d_1}{dm} \right)^3 - 3 \left( \frac{d_1}{dm} \right)^4,
\]

(29)

with \( d_1 = \sqrt{(\theta_{i,1} - \theta_{0,1})^2 + (\theta_{i,2} - \theta_{0,2})^2} \), and \( dm = \delta_{\text{max}} \), where \( \delta_{\text{max}} \) is the distance of the furthest NN from \( \theta_0 \) (as it is shown in Fig. 1). In this work we have taken \( n = 2 \).

3.2 Grid refinement

The reconstruction begins with a rectangular grid with \( N_1 \times N_2 \) nodes with respect to the rectangular coordinates \( \theta_1 \) and \( \theta_2 \), defined over the region where the reconstruction takes place.

For the refinement process, by means of a Gaussian distribution with standard deviation \( \sigma_{\text{d}} \), we draw \( N_d \) new nodes around each of the main deflectors when they are known. If that is not the case, by default, such new nodes are drawn with respect to the center of the region of reconstruction. We keep only those new nodes that lie within a radius \( R_d > \sigma_{\text{d}} \) from the given main deflector. The aim of \( R_d \) is to regulate the density of nodes, whereas weak lensing is dominant. With this distribution of nodes, the reconstruction is capable to provide a better description of those regions with high mass concentrations.

Additionally, around each of the multiple images, we set a circular neighborhood of radius \( r_d \) and determine the number of nodes...
3.3 Penalty functions

In this section we discuss how the different contributions to the penalty function (10) are defined, and how they contribute to (17).

As we described before, the reconstruction method consists of two levels of iterations, where the system of linear equations is computed at the inner level. So that, when we compute such system of equations, those terms that mess up with the linearity with respect to the deflection potential are computed using the deflection potential terms available, as discussed in Sec. 2. In particular, we focus on their identification of the multiply imaged systems. We apply this process only to the strong regime, since, as needed.

3.3.1 Strong regime

The strong regime is constrained by using the multiply imaged systems available, as discussed in Sec. 2. In particular, we focus on their angular position, which makes natural to define the penalty function for this regime as

\[ A_s^2 := \sum_{i=1}^{N_s} \left( \sum_{n=1}^{N_i} p_{in} T_{in}^{-1} p_{in} \right), \]

(30)

where \( N_s \) and \( N_i \) are, respectively, the number of multiply imaged systems that are being used, and the number of images the \( i \)-th system has. On the other hand, \( p_{in} := \theta_{in} - \theta'_{in} \) compares the \( n \)-th observed image position \( \theta'_{in} \) from the \( i \)-th system to the corresponding position \( \theta_{in} \) computed from the reconstruction. The uncertainties in \( \theta'_{in} \) at each component are \( \sigma^2_{1,in} \) and \( \sigma^2_{2,in} \), which are introduced through

\[ S_{i,in} = \begin{pmatrix} \sigma^2_{1,in} & 0 \\ 0 & \sigma^2_{2,in} \end{pmatrix}. \]

(31)

The problem with (30) relies on the need to explicitly solve the lens equation, which makes difficult to explicitly write it in terms of \( \phi \) on every node. For that reason, working on the source plane instead, turns out to be an advantage. In this direction, let us define two different penalty functions.

**Type 1:** Since any multiply imaged system is linked to a unique source, when we apply the lens equation (6) to such images, they have to take us to the same source position \( \beta \) predicted by the reconstruction for the \( n \)-th image to the source true position \( \beta' \), however, since such true position is unknown, the average source position predicted by the reconstruction

\[ \langle \beta_i \rangle = \frac{1}{N_i} \sum_{n=1}^{N_i} \hat{\beta}_{in} \]

(32)

is commonly used instead. Here \( \hat{\beta}_{in} = \theta'_{in} - Z_i \alpha_{in} \). This approach also constrains the multiple images to converge to the same source as needed.

One cannot simply work on the source plane and make use of the observed uncertainties, since the effect of the lens is important. As it is discussed in (Schneider et al. 2006, Part 2, Section 4.6), the magnification induced on the images by the mapping \( \beta \rightarrow \theta \) also affects their position uncertainties. Therefore, as a correction to this matter, by applying the Taylor expansion to the lens equation (6) up to first order derivatives around \( \langle \beta_i \rangle \), we get

\[ \hat{\beta}_{in} = M_{in} (\beta_{in} - \langle \beta_i \rangle), \]

(33)

where

\[ M_{in} = \mu_{in} \begin{pmatrix} 1 - Z_i \kappa_{in} + Z_i \gamma_{1,in} & Z_i \gamma_{2,in} \\ Z_i \gamma_{2,in} & 1 - Z_i \kappa_{in} - Z_i \gamma_{1,in} \end{pmatrix} = \mu_{in} \begin{pmatrix} m_{1,in} & m_{2,in} \\ m_{2,in} & m_{3,in} \end{pmatrix}. \]

(34)

1 We use \( \langle \beta_i \rangle \) since the true source position \( \beta_i \) is unknown.
is the magnification matrix at \( \theta_{in}^i \), and
\[
\mu_{in} = \frac{1}{(1 - Z_i \kappa_{in})^2 - Z_i^2 (\gamma_{1,1,in}^2 + \gamma_{2,2,in}^2)}
\]
is magnification of the corresponding image as well. It is important to take into account that the approximation (33) holds as long as the predicted images are as close as possible to the observed ones.

It has been shown that only the magnification \( \mu_{in} \) is enough to account for this correction (e.g. Bradač et al. (2005b)). We stick to this for the reconstructions presented in this work. Nevertheless, due to the nature of this reconstruction method, considering only the magnification (at least for our implementation) can lead to a divergent solution; the method may become numerically unstable. For instance, considering \( M_{in} \), as a hole helps to control the impact of \( \mu_{in} \) making the reconstruction more stable, with the cost of losing some accuracy around the mass peaks. Now, the implementation of either \( \mu_{in} \) or \( M_{in} \) introduce non-linearities on (17), so that they are not explicitly written in terms of \( \psi_j \). Instead, they are computed from the last \( \psi \) that is known, and manipulated as constant terms.

With respect to \( \langle \beta_i \rangle \), one can compute it in the same way as \( \mu_{in} \) and \( M_{in} \) are computed (as it is done in e.g. Bradač et al. (2005a, 2009)). However, at least for our implementation this approach is not effective, since when either \( \mu_{in} \) or \( M_{in} \) is included the reconstruction shows the tendency to produce an overestimation of \( \kappa \), close to some of the grid corners, and it cannot be effectively solved simply by considering a wider region for the reconstruction or changing the grid size. This problem is solved by explicitly writing the components of \( \langle \beta_i \rangle \) in terms of \( \psi_j \) as
\[
\langle \beta_{q,i} \rangle = \frac{1}{N_i} \sum_{p=1}^{N_i} \theta_{q,ip} - \frac{1}{N_i} \sum_{j=1}^{N} \sum_{p=1}^{N_i} D_{q,j,p} \psi_j,
\]
from which we have that
\[
\langle \theta_{q,i} \rangle := \frac{1}{N_i} \sum_{p=1}^{N_i} \theta_{q,ip},
\]
and
\[
\frac{\partial \langle \beta_{q,i} \rangle}{\partial \psi_k} := -\frac{Z_i}{N_i} \sum_{p=1}^{N_i} D_{q,ip,k},
\]
with \( q = 1, 2 \) representing each component. Under this approach, we also get an improvement in the performance. For a tolerance of \( 10^{-3} \) we observed in the different reconstructions presented in this work a reduction in the number of inner iterations of at least ten times. Also, structure was clearly shown after the first or second inner iteration, and the condition (20) was fulfilled within 4–6 inner iterations.

At this point, we have that our first penalty function is easily obtained by replacing (33) into (30). Having \( \chi_2^2 \), its contribution to (17) is given through
\[
W_{k,j}^{(s)} := \sum_{i=1}^{N_i} \sum_{n=1}^{N_n} \sum_{q=1}^{2} Z_i B_{q,ink} \left( D_{q,inj} - \frac{1}{N_i} \sum_{p=1}^{N_i} D_{q,ip} \right),
\]
and
\[
V_k^{(s)} = \sum_{i=1}^{N_i} \sum_{q=1}^{2} B_{q,ink} \left( Z_i D_{q,ink} + \frac{\partial \langle \beta_{q,i} \rangle}{\partial \psi_k} \right) - B_{q,ink} \langle \beta_{1,i} \rangle,
\]
where we have defined
\[
B_{1,ink} := \mu_{in}^2 \left( Z_i \Delta m_{1,ink} D_{1,ink} + Z_i \Delta m_{2,ink} D_{2,ink} + \Delta m_1 \frac{\partial \langle \beta_{1,i} \rangle}{\partial \psi_k} + \Delta m_2 \frac{\partial \langle \beta_{2,i} \rangle}{\partial \psi_k} \right),
\]
\[
B_{2,ink} := \mu_{in}^2 \left( Z_i \Delta m_{2,ink} D_{1,ink} + Z_i \Delta m_{3,ink} D_{2,ink} + \Delta m_2 \frac{\partial \langle \beta_{1,i} \rangle}{\partial \psi_k} + \Delta m_3 \frac{\partial \langle \beta_{2,i} \rangle}{\partial \psi_k} \right),
\]
\[
b_{1,ink} := \mu_{in}^2 \left( \Delta m_{1,ink} \theta_{1,ink} + \Delta m_{2,ink} \theta_{2,ink} \right),
\]
and
\[
b_{2,ink} := \mu_{in}^2 \left( \Delta m_{2,ink} \theta_{1,ink} + \Delta m_{3,ink} \theta_{2,ink} \right),
\]
with
\[
\Delta m_{1,ink} := \frac{m_{2,ink}^2}{\sigma_{1,ink}^2} + \frac{m_{3,ink}^2}{\sigma_{2,ink}^2},
\]
\[
\Delta m_{2,ink} := \frac{m_{1,ink} m_{2,ink}}{\sigma_{1,ink}^2} + \frac{m_{3,ink} m_{4,ink}}{\sigma_{2,ink}^2},
\]
\[
\Delta m_{3,ink} := \frac{m_{2,ink}^2}{\sigma_{1,ink}^2} + \frac{m_{4,ink}^2}{\sigma_{2,ink}^2}.
\]

Type 2: In this case, for each of the multiply imaged systems, we compare the possible pairs formed from the position of the source predicted for each of the corresponding multiple images. Therefore, the penalty function reads
\[
\chi_s^2 := \sum_{i=1}^{N_i} \sum_{m=1}^{N_n} b_{m,n}^T S_{m,n}^{-1} b_{m,n},
\]
where \( b_{m,n} = b_{im} - b_{in} \), with \( b_{in} = \theta_{in}^i - Z_i \alpha_{in} \) and \( b_{im} = \theta_{im}^i - Z_i \alpha_{im} \). In order to account for the correction discussed above, that is needed in order to be able to work at source plane, we have defined
\[
S_{m,n} := \frac{1}{2} \begin{pmatrix} \sigma_{1,im}^2 + \sigma_{1,im}^2 & 0 \\ \sigma_{2,im}^2 + \mu_{im}^2 & \sigma_{2,im}^2 + \mu_{im}^2 \end{pmatrix},
\]
Here, the contribution of \( \chi_s^2 \) to (17) yields
\[
W_{k,j}^{(s)} := \sum_{i=1}^{N_i} \sum_{n=1}^{N_n} \sum_{m=1}^{N_m} \sum_{q=1}^{2} Z_{i,m,n} \Delta D_{q,inm} \Delta D_{q,inmj},
\]
and
\[
V_k^{(s)} := \sum_{i=1}^{N_i} \sum_{n=1}^{N_n} \sum_{m=1}^{N_m} \sum_{q=1}^{2} \left( Z_i \Delta D_{q,inm} \Delta \theta_{q,in} \right),
\]
where we have defined
\[
\Delta D_{q,in} := D_{q,in} - D_{q,imp},
\]
with \( p = k, j \), and
\[
\Delta \theta_{q,in} := \theta_{q,in}^i - \theta_{q,in}^j.
\]
3.3.2 Weak regime

Following Bradač et al. (2005a), we constrain the weak regime by means of sources apparent ellipticity, as discussed in Sec. 2. Therefore, for this regime the penalty function is defined as

\[ \chi^2_w := \sum_{i=1}^{N_w} \frac{|\epsilon_i - (\epsilon_i)\rangle^2}{\sigma_i^2}, \]  

(54)

with \(N_w\) being the number of background galaxies for which their ellipticity and redshift is known. Additionally Schneider et al. (2006)

\[ \sigma_i^2 = \frac{1}{N_{p,i}} \left( 1 - \min\left( |g_i|^2, |g_i|^2 \right) \right)^2 \sigma_{\epsilon_i}^2 + \sigma_{err}^2, \]  

(55)

where \(N_{p,i}\) corresponds to the number of averaged apparent ellipticities in (9). Such ellipticities are from sources with the same redshift, and located in a neighborhood within which the properties of the lens do not change significantly Bartelmann & Schneider (2001); Schneider et al. (2006). Now, since in practice it is not easy to fulfill these conditions, the best choice is to take \(N_{p,i} = 1\), and so consider \(\epsilon_i\) to be the best representative of its sample at its neighborhood Bradač et al. (2005a). That is why in (54) \(\epsilon_i\) is compared to \((\epsilon_i)\). The quantities \(\sigma_{\epsilon_i}\) and \(\sigma_{err}\) in (55) correspond to the intrinsic and observational ellipticity standard deviations, respectively. In this paper we take \(\sigma_{\epsilon_i} \sim 0.2 - 0.3\) and \(\sigma_{err} = 0.1\) Bradač et al. (2005a); Cain et al. (2016).

An alternative approach to \(\chi^2_w\) can be found in e.g. Merten et al. (2009, 2011); Merten (2016).

According to (9), the penalty function (54) has two possible forms. To begin with, for \(|g_i| \leq 1\) (54) turns into

\[ \chi^2_w = \sum_{i=1}^{N_w} \frac{1}{\sigma_i^2} \left( \epsilon_i - Z_i \gamma_i \right)^2 = \sum_{i=1}^{N_w} \frac{\epsilon_i - Z_i \epsilon_i \gamma_i - Z_i \gamma_i}{1 - Z_i} \sigma_i^2 \]  

(56)

whereas for \(|g_i| > 1\) (54) becomes

\[ \chi^2_w = \sum_{i=1}^{N_w} \frac{1}{\sigma_i^2} \left( \epsilon_i - Z_i \gamma_i \right)^2 = \sum_{i=1}^{N_w} \frac{Z_i \epsilon_i \gamma_i^2 + Z_i \gamma_i - Z_i} {Z_i^2} \sigma_i^2 \gamma_i^2 \]  

(57)

In practice, the observations mainly come from sources that satisfy \(|g_i| < 1\). However, we have implemented both possibilities in relensing, such that, for every inner iteration we evaluate \(|g_i|\) from the last deflection potential that we have. That, relensing decides whether to use (56) or (57) as the corresponding penalty function for the weak regime.

For both versions of \(\chi^2_w\), the coefficients \(W_{k,j}^{(w)}\) and \(V_k^{(w)}\) can be written as

\[ W_{k,j}^{(w)} := \sum_{i=1}^{N_w} \left[ A_{i,j} G_{1,ik} K_{ijk,j} + A_{j,k} G_{1,jk} G_{1,ik} \right] + A_{2,i} G_{2,ik} K_{ijk,j} + A_{3,i} G_{1,ik} G_{1,jk} + A_{4,i} G_{2,ik} G_{2,jk} + A_{5,i} K_{ik} K_{ij} , \]  

(58)

and

\[ V_k^{(w)} := \sum_{i=1}^{N_w} \left[ A_{1,i} G_{1,ik} + A_{2,i} G_{2,ik} + A_{3,i} K_{ij} \right] , \]  

(59)

such that for \(|g_i| \leq 1\) we get

\[ A_{1,i} = \frac{Z_i^2}{\sigma_i^2} \epsilon_{1,i} , \quad A_{2,i} = \frac{Z_i^2}{\sigma_i^2} \epsilon_{2,i} , \quad A_{3,i} = A_{4,i} = \frac{Z_i^2}{\sigma_i^2} \]  

\[ A_{5,i} = \frac{Z_i^2}{\sigma_i^2} \epsilon_{1,i} , \quad A_{6,i} = \frac{Z_i^2}{\sigma_i^2} \epsilon_{2,i} , \quad A_{7,i} = \frac{Z_i}{\sigma_i^2} \epsilon_{2,i} , \quad A_{8,i} = \frac{Z_i}{\sigma_i^2} \epsilon_{2,i} \]  

(60)

\[ a_{3,i} = \frac{Z_i}{\sigma_i^2} \epsilon_{1,i} , \quad a_{4,i} = \frac{Z_i}{\sigma_i^2} \epsilon_{2,i} , \quad a_{5,i} = \frac{Z_i}{\sigma_i^2} \epsilon_{2,i} \]  

(61)

\[ \text{where} \quad \sigma_{\epsilon_{1,i}} := (1 - Z_i) \sigma_{\epsilon_{1,i}}^2 \quad \text{and} \quad \sigma_{\epsilon_{2,i}} := Z_i^2 \sigma_{\epsilon_{2,i}}^2 . \]  

3.3.3 Regularization terms

This sort of grid based reconstruction methods require a term which provides stability along the reconstruction. Such term helps to deal with numerical noise, which otherwise will most likely lead to divergent solutions for (17).

Different solutions have been proposed to overcome these difficulties, either as a regularization term (e.g Seitz et al. (1998); Bradač et al. (2005a,b); Merten (2016)), or by introducing the signal to noise ratio (e.g Deb et al. (2008)). Here, as we have stated before, we have chosen to implement the former approach, particularly by using the regularization terms

\[ \chi^2_{\kappa(R)} = \eta_k \sum_{j=1}^{N} \left( \kappa_j - \kappa_j^{(m-1)} \right)^2 , \]  

(62)

and

\[ \chi^2_{\gamma(R)} = \eta_{\gamma} \sum_{j=1}^{N} \sum_{q=1}^{N} \left( \gamma_{q,j} - \gamma_{q,j}^{(m-1)} \right)^2 , \]  

(63)

with \(\eta_k\) and \(\eta_{\gamma}\) being positive defined constants. Such constants or regularization weights, control how smooth the reconstruction turns. The term (62) acts mainly within the cluster inner region dominated by the strong regime, while the term (63) helps to control the reconstruction outside the inner region.

On the other hand, if we are in the \(m\)-th outer iteration, the terms \(\kappa_j^{(m-1)}, \gamma_{1,j}^{(m-1)}, \text{ and } \gamma_{2,j}^{(m-1)}\) in (62) and (63), respectively, are computed after the grid refinement at the end of the previous outer iteration, or they come from the initial guess \((m = 1)\). Therefore, from (62) we get the coefficients

\[ W_{k,j}^{(s)} := \eta_k \sum_{i=1}^{N} K_{ik} K_{ij} \]  

(64)

and

\[ V_k^{(s)} := \eta_k \sum_{i=1}^{N} \kappa_{i}^{(m-1)} K_{ik} , \]  

(65)
meanwhile for (63) we have
\[ W'_{k} := \eta_{y} \sum_{i=1}^{N} \sum_{q=1}^{2} G_{q,ik} G_{q,iq}, \quad (66) \]
and
\[ V'_{k} := \eta_{y} \sum_{i=1}^{N} \sum_{q=1}^{2} \gamma_{q,i} (m^{-1}) G_{q,ik}. \quad (67) \]

4.1 Catalogues

For each model, the catalogues involved are drawn under the same conditions. For the strong regime, we consider 40 sources randomly selected with redshift \( z_{s} > z_{f} \), and for which more than one image is obtained after solving the lens equation (6). Whereas for the weak regime we consider 1000 background sources with redshift \( z_{b} > z_{f} \), whose ellipticities catalogue is constructed in the same direction as in Bradač et al. (2005a); Cain et al. (2016). Thus, the intrinsic ellipticities catalogue is constructed in the same direction as in Bradač et al. (2005a); Cain et al. (2016). Thus, the intrinsic ellipticities are drawn from a Gaussian distribution with \( \sigma_{e} = 0.2 \), then, to each component of the observed ellipticities (which are computed from (7)) an error is added. Such an error is drawn from a Gaussian distribution with shape and rate parameters \( \alpha = 3 \) and \( \beta = 3/2 \), respectively. For each redshift, the corresponding added uncertainty is drawn from a Gaussian distribution with \( \sigma_{e,i} = 0.05(1 + z_{i}) \) (with \( i = 1, \ldots, N_{w} \)). We keep those sources with only one image after solving the lens equation.

4.2 NIS and 2NIS

The deflection potential for a NIS is defined as
\[ \psi(\theta) = R_{0} \sqrt{\Delta \theta^{2} + \theta_{0}^{2}} \quad \text{with} \quad R_{0} := \frac{4 \pi \sigma^{2} D_{\text{rel}}(z_{f}, z)}{c^{2} D(z)}, \quad (68) \]
where \( \Delta \theta := \theta - \theta_{c} \). Such deflection potential depends on the angular diameter distance between the lens and the given source \( D(z_{f}, z) \), and also between the observer and the source \( D(z) \). As well, it depends on the velocity dispersion \( \sigma \), the speed of light \( c \), the distribution core \( \theta_{0} \), and distribution center \( \theta_{c} \).

The first model we consider is a NIS with redshift \( z_{f} = 0.4 \), with center at \( \theta_{c} = (0, 0) \, \text{arcmin} \). It is characterized by the parameters \( \theta_{0} = 0.4 \, \text{arcmin} \) and \( \sigma = 1500 \, \text{km/s} \). The corresponding convergence map for a source with redshift \( z_{r} = 9 \) is depicted in Fig. 3 (True).

The second model consists of two identical NIS with redshift \( z_{f} = 0.4 \), whose center is at \( \theta_{c1} = (-0.2, -0.4) \, \text{arcmin} \), and \( \theta_{c2} = (0.2, 0.4) \, \text{arcmin} \), respectively. The parameters describing each of those NIS are \( \theta_{0} = 0.2 \, \text{arcmin} \) and \( \sigma = 1000 \, \text{km/s} \), such that the resulting convergence map for a source with redshift \( z_{r} = 9 \) is shown in Fig. 5 (True). We refer to this model as 2NIS.

4.3 Reconstruction

The reconstructions are carried out considering \( \eta_{k} = \eta_{y} = 200 \) irrespectively of the initial guess \( \psi^{(0)} \). Here, for the NIS we only consider blind reconstructions, and then we explore the effect of the smoothing on such reconstructions. Whereas, for the 2NIS we compare blind reconstructions against those for which \( \psi^{(0)} \) is given by a Non-singular Isothermal Ellipsoid (NIE for short), as we discuss later.

Now, for the initial grid, when only the strong regime is being considered \( s \), we use \( 20 \times 20 \) nodes over a region of \( 3 \times 3 \, \text{arcmin}^{2} \). Meanwhile, when both regimes are being considered \( s+w \), we take \( 22 \times 22 \) nodes over a region of \( 3.2 \times 3.2 \, \text{arcmin}^{2} \). For the last scenario, the extra nodes and grid size help to avoid overestimations of the different quantities of interest on the boundary (particularly \( \kappa \)), which arise when some observations are close to such boundary.

With respect to the refinement process, the NIS shows only one deflector, so that we add \( N_{d} = 600 \) nodes for each refinement. On the other hand, the 2NIS consists of two deflectors, so that we add \( N_{d} = 300 \) nodes for each deflector. In both cases, those new nodes are distributed considering \( \sigma_{d} = 1 \, \text{arcmin} \), and \( R_{d} = 2 \, \text{arcmin} \). Also, we take \( r_{d} = 0.2 \, \text{arcmin} \) for the first refinement process, which is reduced by a factor \( u = 0.8 \) at each of the subsequent refinements, except for the last one for which the refinement is not applied. As we discussed in Sec. 3.1, we set \( Q = 25 \) nodes to perform the finite differences.

4.3.1 Convergence

In Fig. 3 we can see eight different reconstructed convergence maps \( \kappa \), along with the true mass for the NIS; which is the smallest of the distributions that we are considering in this work. This distribution posess spherical symmetry, for which we have only considered blind reconstructions here, since a parametric fitting to the data from the strong regime either to a NIS itself or to a NIE leads to the true lens (or at least a distribution close to it), and that is not what we want. One of the aims of this work is to show the power of the method to reveal important features of the lens, based upon little knowledge about it, which is closer to what one has to face with real data.

Let us start with the effect of the smoothing. The Fig. 3 shows in the upper row the reconstructions for both penalty functions \( \chi_{r}^{2} \) without considering the smoothing (reconstructions (1) – (4)), meanwhile, in the middle row the reconstructions were carried out under the very same conditions, but with considering the smoothing (reconstructions (5) – (8)); so that reconstruction (1) is directly compared to reconstruction (5), and so on. It is clear that the reconstructions reveal the central peak regardless of the smoothing, the penalty function used, or whether weak lensing is being considered. However, in Fig. 4 we can see that the inclusion of the smoothing provides a better reconstruction once we move towards the outskirts.

Comparing Type 1 and Type 2 reconstructions (carried out under the same conditions), we can see that there are no significant differences when it comes to the results given by both penalty functions. Instead, a significant difference appears when the smoothing is included. Without smoothing, the contours in \( \kappa \) are noisier and less accurate, particularly outside the inner region. This is clearer when weak lensing constraints are being considered. Here, for the inner region we refer to the region that encloses all the multiple
images known, as it is depicted in Fig. 4 (purple dashed contour). That behaviour stands irrespective of the lens distribution or the input deflection potential $\psi(0)$. This behavior is accentuated when the complexity of the lens increases.

When only the strong regime is included, the reconstruction effectiveness is strongly restricted to the inner region, as we can see in Fig. 4, where most of such region presents a relative difference within $[\Delta_r \kappa] < 0.1$ (10%). In particular, we can see an improvement at the inner region boundary when the smoothing is included. Outside the inner region, the reconstruction loses its effectiveness and there is an underestimation of $\kappa$, which is less accentuated when the smoothing is added, but still present. Once strong and weak regimes are combined, there is an overall enhancement in the reconstruction outside the inner region, but without smoothing the increase in $\kappa$ is less uniform, and we can still see marked peaks of both underestimations and overestimations of $\kappa$.

Now, we put our attention on $M = M(\leq \theta)$, which represents the mass enclosed within an angular radius $\theta$ measured from $\theta = (0,0)$ arcmin. This applies to all the reconstructions considered in this work. Here, Fig. 7 shows $M$ for the eight reconstructions, given for $\theta_{in} \leq \theta \leq 1.5$ arcmin. The curves have been separated into Type 1 (upper panel) and Type 2 (lower panel), so that we can directly compare the effects of using the strong/weak regime as well as the smoothing. We can see that $M$ behaves similarly for Type 1 and Type 2 reconstructions. Again, we see that the best performance appears within the inner region, where we can see a relative difference within $[\Delta_r M] \leq 0.05$ (5%).

The vertical lines in Fig 7, given for $\theta_{in} \approx 0.86$ arcmin and $\theta_{out} \approx 1.04$ arcmin, correspond, respectively, to the inner and outer radius of the inner region. It is within such limits, particularly close to $\theta_{out}$, that $M$ starts to deviate from the true curve. Beyond the inner region, the effectiveness decreases, and for the range within which the reconstruction is being considered, the relative difference gets up to $[\Delta_r M] \approx 0.16$ (16%) for those reconstructions where only strong lensing is being used; irrespectively of the smoothing. With the inclusion of the weak regime, we see an improvement in $M$ of ~7%.

We see that the estimation of $M$ is not severely affected by the smoothing, since the main features of the lens are recovered in both cases. However, the smoothing provides a clear improvement in the mass profile, as well as in estimation of the magnifications maps, and therefore the critical curves; as we discuss later. Hence, from here on we consider the smoothing in our reconstructions.

For the 2NIS we focus on the comparison between blind reconstructions and those for which the input potential $\psi(0)$ or initial guess contains some previous information about the lens. We want to keep this initial guess as simple as possible, so that we cannot interfere in excess, imposing many assumptions about the lens. For that reason, as it has been shown in Bradač et al. (2005a), fitting a NIE to the multiply image systems provides a good enough estimation of $\psi(0)$ to start working with. Depending on the complexity of the cluster, or its morphological properties, it may be convenient to use a combination of NIE as the initial guess. The deflection potential of such NIE is defined by making the transformation

$$|\Delta \theta|^2 \rightarrow (1 - \epsilon)(\Delta \theta_1 \cos \varphi + \Delta \theta_2 \sin \varphi)^2 + (1 + \epsilon)(-\Delta \theta_1 \sin \varphi + \Delta \theta_2 \cos \varphi)^2$$

(69)

into (68), which accounts for the introduction of the ellipticity $\epsilon$ and the rotation $\varphi$ of the semi major axis with respect to the horizontal, and measured counterclockwise. For the fitting we have constrained $0 \leq \epsilon \leq 0.25$, so the profile does not get distorted into a peanut shaped one, as it is discussed in e.g. Golse & Kneib (2002). Here we have let $\theta_0 = 0.2$ arcmin fixed. Therefore, for the 2NIS, besides the blind reconstruction we consider a NIE as the initial guess.

In Fig. 5 we can see the eight reconstructions taken into account. The upper row shows the blind reconstructions (reconstructions (1) – (4)), whereas the middle row shows the reconstruction carried out under the same conditions except that the initial guess is given by the NIE (reconstructions (5) – (8)). The lower panel shows the true profile.

The 2NIS is characterized by two prominent peaks or main deflectors, which are effectively recovered in all reconstructions. Therefore, the initial guess is not relevant in this matter. In terms of the blind reconstructions, we can see that they struggle to recover the distribution beyond the inner region. It is expected for reconstructions (1) and (3), which only use strong lensing, as we saw for the NIS. In contrast, reconstructions (2) and (4), which also use weak lensing show an improvement in $\Delta_r \kappa$ of about 10% – 20% towards the outskirts, as it is depicted in Fig. 6 (upper row). Within most of the inner region $\kappa$ shows a relative difference $[\Delta_r \kappa] < 0.1$, with few peaks that fall into 0.1 (10%) $\leq [\Delta_r \kappa] < 0.2$ (20%).

At this point, it is clear that weak lensing strengthens the reconstructions outside the inner region, so that the smoothing does not have a negative impact on such regime. We focus now on how the reconstruction behaves once one goes beyond blind reconstructions. For reconstructions (5) – (8) we can see a considerable improvement with respect to their blind counterpart. Comparing reconstructions (1) and (5), as well as reconstructions (3) and (7) (which only use strong lensing), we find a closer approach to the true profiles, as it can be seen in the contour curves in $\kappa$ and also in the $\Delta_r \kappa$ maps in Fig. 6. In addition, comparing reconstructions (2) and (6), as well as reconstructions (4) and (8) (which use strong and weak lensing), we can see that outside the inner region they go from an underestimation to an overestimation of $\kappa$, notwithstanding, the region where $[\Delta_r \kappa] < 0.1$ is wider, and the overestimation lies beyond $[\Delta_r \kappa] > 0.2$, mostly close to the boundaries, where the constraints are weaker. In either case, within the inner region, there are no significant changes in the reconstruction. Likewise, Type 1 and Type 2 reconstructions behave equally well.

With respect to $M$, which is computed within $0 \text{arcmin} \leq \theta \leq 1.5$ arcmin, we have a good agreement with the true mass with a relative difference within $[\Delta_r M] < 0.05$ (5%) for $\theta \leq \theta_{in} \approx 0.74$ arcmin, and close to $\theta_{out} \approx 0.90$ arcmin the curve starts to deviate from the true one; except for the orange curves (reconstructions (6) and (8)) as it is depicted in Fig. 8. For $\theta > \theta_{out}$ the worst estimation of $M$ is given by reconstructions (1) and (3) (red curves) for which $\Delta_r M \approx -0.25$. They are followed by reconstructions (2) and (4) (blue curves), which show an improvement of at most ~7%. As expected from the results in $\kappa$, a better scenario is provided by reconstructions (5) and (7) (green curves), as well as (6) and (8) (orange curves), for which the relative difference is at most $[\Delta_r M] \approx 0.08$ (8%) and $\Delta_r M \approx 0.01$ (1%), respectively. It is clear that there are not significant differences between Type 1 and Type 2 reconstructions.

Blind reconstructions work incredibly well, particularly if one is interested in the inner region. However, in order to achieve a better understanding of the outskirts, it is worth using an initial guess other than the blind one. The drawback relies on what initial guess we use. For the inner region there are no significant effects if there are enough multiply imaged systems close to the mass peaks, but the outskirts are more sensitive to such changes, where the method presents particular difficulties in scaling down/up to values close to the true $\kappa$, coming from overestimations/underestimations given by the initial guess.
relensing: Reconstructing the mass profile of galaxy clusters from gravitational lensing

Figure 3. Convergence maps ($\kappa$) for the NIS, given for a source with redshift $z_s = 9$. The upper row shows the blind reconstructions without considering the smoothing, while the middle row shows the blind reconstructions considering the smoothing. The true map is shown in the lower panel.

Figure 4. Relative difference $\Delta_r \kappa$ between the reconstructed and true convergence maps for the NIS shown in Fig. 3. Here, the red and green solid contours correspond to $\Delta_r \kappa = 0.1$ and $\Delta_r \kappa = -0.1$, respectively. Meanwhile, the purple dashed contour delimits the region within which we have multiple images available.
Figure 5. Convergence maps ($\kappa$) for the 2NIS, given for a source with redshift $z_r = 9$. The upper row shows the blind reconstructions, while the middle row shows the reconstructions with an input $\psi^{(0)}$ given by a NIE. For all reconstructions, the smoothing has been considered. The true map is shown in the lower panel.

Figure 6. Same as in Fig. 4 but for the 2NIS.
4.3.2 Magnification

Magnification maps are key elements in the study of galaxy clusters, since they are needed to characterize the galaxy cluster, for example, as a cosmic telescope, since the biggest magnifications occur close to the critical curves, allowing us to study distant objects. Also, they help to identify new multiply imaged systems, as well as to extend those already known. So then, we do not only need the critical curves, but also the magnification itself, close to such curves. The quality of the reconstructed magnification maps has been studied in e.g. Meneghetti et al. (2017), where they compare the reconstruction of the simulated galaxy clusters, Ares and Hera, given by different methods and approaches, in order to have a better understanding of the results obtained for the HFF clusters. We present our own reconstructions for Ares and Hera in Sec. 5.1.

In Fig. 9 and Fig. 11 we have the magnification maps $|\mu|$ for the NIS and 2NIS, respectively. Those maps are depicted in concordance with the convergence maps shown in Fig. 3 and Fig. 5.

At critical curves, by definition $\mu$ diverges, so that they are depicted as regions of high $|\mu|$. Due to this behavior, even true deviations from true critical curves are expected to make it difficult to accurately recover $\mu$.

For the NIS, we have that reconstructions (1) – (4) (upper row in Fig. 9) produce critical curves that actually follow the form and size of the true ones, but are noisy and quite irregular, particularly the outer or tangential curve. Such irregularities in fact provide a $|\mu|$ map with high error close to the critical curves, as it is depicted in Fig. 10 (upper row). Now, by including the smoothing in reconstructions (5) – (8) (middle row in Fig. 9) we can see a better recovery of critical curves. They are not as smooth as the true ones, but are less irregular than curves (1) – (4). The improvement is not restricted to the recovery of such critical curves in shape and size, but it extends to the overall estimation of $|\mu|$, particularly when weak lensing is used, as it can be seen for reconstructions (6) and (8) in Fig. 10 (lower row). Still, the highest deviations in $|\mu|$ appear close to the critical curves.

Now, with respect to the 2NIS, in Fig. 11 we can see that there are no significant differences between the critical curves provided by the eight reconstructions. All of them provide accurate reproductions of the critical curves both in shape and size. However, for reconstructions (5) – (8) we get a better reproduction of $|\mu|$ as a whole, as it is depicted in Fig. 12. Again, the highest deviations appear close to the critical curves.

Additionally, in Fig. 13 and Fig. 14 we can see the critical and
Figure 9. Magnification maps (|\(\mu\)|) for the NIS, given for a source with redshift \(z_r = 9\). They correspond to the reconstructions shown in Fig. 3. The true map is shown in the lower panel.

Figure 10. Relative difference \(\Delta |\mu|\) between the reconstructed and true magnification maps for the NIS shown in Fig. 9. Here, the red and green solid contours correspond to \(\Delta \kappa = 0.2\) and \(\Delta \kappa = -0.2\), respectively. The black solid lines depict the true critical curves.
Figure 11. Same as in Fig. 9 but for the 2NIS.

Figure 12. Same as in Fig. 10 but for the 2NIS.
curve is supposed to be mapped into a point, but instead the reconstructed caustic is wider and more irregular. This effect has to be taken into account when one is interested in predicting the existence of new images or multiply imaged systems, as well as when it comes to verifying the reproduction of the input data.

It is worth noting that transformation (18) does not affect the recovery of neither the critical curves nor the convergence map and thus the mass profile, since $\kappa$ and $\gamma$ are invariant under this transformation. Nevertheless, $\alpha$ is not invariant, since $\alpha \rightarrow \alpha' = \alpha + c$ translates the source plane. It is not an observable, but it results in a translation of the caustic curves. We see that fixing the corresponding three degrees of freedom allows the reconstruction to properly account for the caustic curves, as we have in Fig. 13 and Fig. 14 (lower row).

5 REALISTIC DISTRIBUTIONS: ARE S AND HERA

At this point, we have explored under which conditions relensing provides the best performance in its current state, by exploring the reconstructions of the NIS and 2NIS. We move forward in order to explore the validity of the reconstructions provided by relensing in more realistic scenarios. We focus on the simulated clusters Ares and Hera, which are intended to reproduce the complex structures and lensing properties observed in the Hubble Frontier Fields (HFF) clusters.

Ares is shown in Fig. 15 (left panel). It corresponds to a cluster with redshift $z_1 = 0.5$, for which a semi-analytical approach was used for its simulation, under a flat $\Lambda$CDM cosmology with density parameter $\Omega_{m,0} = 0.272$ and Hubble constant $H_0 = 70.4\ km\ s^{-1}\ Mpc^{-1}$. Ares is characterized by two main distributions or main deflectors (MD) located at $\approx (-0.335, -0.525)\ arcmin$ (MD 1) and $\approx (0.666, 0.666)\ arcmin$ (MD 2). Also, it is rich in well defined substructures as it is shown in Fig. 16 (lower panel).

On the other hand, Hera, with redshift $z_2 = 0.507$ is shown in Fig. 15 (right panel). Unlike Ares, the simulation of Hera was carried out using a N-body approach, under a flat $\Lambda$CDM cosmology with density parameters $\Omega_{m,0} = 0.24, \Omega_{b,0} = 0.04$ and Hubble constant $H_0 = 72\ km\ s^{-1}\ Mpc^{-1}$. Like Ares, Hera shows two prominent peaks corresponding to the MD, located at $\approx (-0.001, 0.015)\ arcmin$ (MD 1) and $\approx (-0.361, 0.045)\ arcmin$ (MD 2). It also exhibits a clear substructure at $\approx (0.113, -0.883)\ arcmin$, whose influence appears in the convergence map Fig. 17 (lower panel), as well as in the magnification map Fig. 22 (lower panel).

See Meneghetti et al. (2017) and the references therein for further details on how the simulations of Ares and Hera were carried out, as well as their lensing properties.

5.1 Reconstruction

For Ares, in the strong regime, there is available a catalogue that consists of 85/242 systems/images. This amount of constraints exceeds those found e.g. in the HFF clusters, for which the number of constrains used/identified for their reconstructions have been e.g. 60/188 systems/images in Mahler et al. (2018) for Abell 2744, 45/138 systems/images in Lagattuta et al. (2019) for Abell 370, 19/52 systems/images in Karman et al. (2017) for Abell S1063, 48/138 systems/images in Bergamin et al. (2021); Vanzella et al. (2021) for MACS J0416.1-2403 (MACS 0416), 60/165 systems/images in Limousin et al. (2016) for MACS J0717.5+3745 (MACS 0717), and 45/143 systems/images in Jauzac et al. (2016) for MACS J1149.5+2223 (MACS 1149). The number of systems/images actually used for
the reconstructions is usually smaller, depending on the uncertainty in their identification. For the weak regime the ellipticities redshifts are not available. Hence, we only use the strong regime for Ares reconstructions.

In the case of Hera, for the strong regime we have fewer constraints, being 19/65 systems/images. The impact on the number of constraints is noticeable in the reconstruction, as we discuss below. Here, for the weak regime the redshifts are available, therefore, for Hera we use both strong and weak regimes to perform the reconstruction. The reconstruction takes place in a region of $3.4 \times 3.4 \text{arcmin}^2$ for which we have 123 sources with their ellipticity and redshift.

**relensing** approaches the refinement of the grid with respect to the main deflectors. So then, in order to identify what cluster members can be considered as main deflectors, one can simply apply a quick blind reconstruction, which will show the peaks of mass corresponding to those main deflectors. Following this process, we identified the two main deflectors present in Ares, as well as in Hera. Then, we estimated their location, which were aforementioned.

Now, with respect to the reconstructions, for each cluster we take as initial guess a NIE as well as a 2NIE; since both cluster have two main deflectors. Additionally, from our previous experience with the NIS and 2NIS we have that Type 1 and Type 2 reconstructions produce similar results. Hence, it is best to let them work together.

Considering the input potential and the type of reconstruction, we are left with four combinations to carry out the reconstructions, namely NIE + Type 1, NIE + Type 2, 2NIE + Type 1, and 2NIE + Type 2. For each combination we perform 10 realizations of the reconstructions, which give us 40 reconstructions in total. The results show in this section correspond to the average of those 40 reconstructions. We do this for Ares and Hera.

The conditions to perform the reconstructions are the same for Areas and Hera. For all realizations we consider the smoothing. We consider an initial grid of size $n_x \times n_y$ nodes over a region of $3.4 \times 3.4 \text{arcmin}^2$; which covers the region shown in Fig. 15. Since we are working on a square region, we consider $n_x = n_y$, which are randomly selected between 20 – 25 nodes. Likewise, $\eta_x = \eta_y$ are randomly selected between 100 – 200.

With respect to the refinement process, we add $N_d$ new nodes for each refinement for each main deflector, which is randomly between 300 – 500. Such nodes are distributed with $\sigma_d = 1 \text{arcmin}$, and $R_d = 2 \text{arcmin}$. The adaptive refinement was done with $r_d = 0.2 \text{arcmin}$ and $u = 0.8$. For the finite differences we use $Q = 25$ nodes.

We have used almost the same parameters that we used for the NIS and 2NIS, except for $n_x, n_y, \eta_x, \eta_y$, and $N_d$, which were drawn from a uniform distribution defined in the given ranges. The selection took place for each realization of the reconstruction.

### 5.1.1 Convergence

In Fig. 16 we find the reconstructed (left panel) and true (right panel) $\kappa$ maps for Ares. It is clear that our reconstruction retrieves successfully the two MD, as well as the general morphological characteristics found in Ares. However, the reconstructions cannot account for the rich substructure present in Ares.

Within the inner region we have that most of such region satisfies $|\Delta_\kappa| < 0.1 (10\%)$ as it is depicted in Fig. 18 (left panel), where we also can find few regions for which $0.1 (10\%) < |\Delta_\kappa| < 0.2 (20\%)$ (red contour) and $-0.2 (20\%) < |\Delta_\kappa| < 0.1 (10\%)$ (green contour). Beyond that, we have $|\Delta_\kappa| > 0.2 (20\%)$ at those locations with substructure. Outside the inner region the quality of the reconstruction decreases as expected, since we are using strong lensing only. The discrepancy is higher towards the upper left and lower right corners, due to the lack of observations there, as we can see in Fig. 15 (left panel). For Hera, we can see in Fig. 17 that our reconstruction (left panel) recovers both MD, as well as the general shape and size shown by Hera. However, they do not exhibit the substructure (enclosed by the white dashed curve) found in the true profile (right panel). In Fig. 18 (right panel) we can see that such substructure (enclosed by the black dashed curve) lies outside the inner region, so that...
there are not multiple images there that allow relensing to account for it. Therefore, the absence of this substructure is due to the lack of data there, and not a consequence of the reconstruction method used. This behaviour is consistent with the reconstructions discussed in Meneghetti et al. (2017). Now, similarly to Ares, we can see that $|\Delta_\text{r} \kappa| < 0.1$ (10%), for most of the inner region, with few regions where $0.1$ (10%) $\leq \Delta_\text{r} \kappa \leq 0.2$ (20%) (red contour) and $-0.2$ (20%) $\leq \Delta_\text{r} \kappa \leq -0.1$ (10%) (green contour). For the substructure we have $\Delta_\text{r} \kappa > -0.3$ (30%).

In addition, for both Ares and Hera we can see in Fig. 18 that our reconstructions present a deviation $|\Delta_\text{r} \kappa| \gtrsim 0.1$ (10%) around the MD (violet solid contours), where the underestimations are present at the exact location of the MD (green contours).

With respect to $M$, it is computed in the range $0 \text{arcmin} \leq \theta \leq 1.7 \text{arcmin}$. For Ares, the maximum relative difference that our reconstruction produces is $|\Delta_\text{r} M| \approx 0.024$ (2.4%), which appears close to the center of the distribution. The deviation from the true $M$ is stable even between $\theta_{\text{in}} \approx 1 \text{arcmin} \leq \theta \leq \theta_{\text{out}} \approx 1.56 \text{arcmin}$, and it starts to deviate beyond $\theta_{\text{out}}$, as we saw for the NIS and 2NIS.

For Hera, we can see in Fig. 20 that the deviation from the true $M$ gets its higher value $|\Delta_\text{r} \kappa| \approx 0.12$ (12%) for $\theta < 0.2 \text{arcmin}$, which approximately coincides with the mass estimation for MD 1; since it is located close to the image center. This behaviour is expected from how the convergence behaves close to the MD, where, as we mentioned above, it presents a underestimation $|\Delta_\text{r} \kappa| \gtrsim 0.1$ (10%), which translates in a underestimation in $M$ with the same range of error. Beyond such underestimation, the difference decreases rapidly, achieving $|\Delta_\text{r} \kappa| \approx 0$ (0%). Between $\theta_{\text{in}} \approx 0.38 \text{arcmin} \leq \theta \leq \theta_{\text{out}} \approx 0.82 \text{arcmin}$ again $M$ starts to deviate from its true value, getting up to $|\Delta_\text{r} M| \approx 0.05$ (5%) beyond $\theta_{\text{out}}$.

Among the reconstructions discussed in Meneghetti et al. (2017), we compare our results to those reconstructions produced with

**Figure 16.** Convergence maps ($\kappa$) for Ares, given for a source with redshift $z_s = 9$. The reconstruction and true map are shown respectively in the left and right panels.

**Figure 17.** Same as in Fig. 16 but for Hera. The white dashed contour encloses the substructure present in the true map of Hera (right panel), and which is not recovered by the reconstruction (left panel).
\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure18.png}
\caption{Relative difference \(\Delta \kappa\) between the reconstructed and true convergence maps for Ares (left panel), as well as for Hera (right panel). Here, the red and green solid contours correspond to \(\Delta \kappa = 0.1\) and \(\Delta \kappa = -0.1\), respectively. The purple dashed contour delimits the region within which we have multiple images available. Additionally, the solid violet contours enclose the main deflectors. Here, the black dashed contour encloses the substructure present in Hera (left panel).}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure19.png}
\caption{Mass for Ares within an angular radius \(\theta\), centered at \(\theta = (0, 0)\) arcmin (upper panel), and the relative difference \(\Delta M\) comparing the reconstructed and true mass (lower panel). The vertical dashed lines represent the inner and outer radius of the inner region (purple dashed contour in Fig. 18 left). The horizontal dashed lines enclose \(|\Delta M| \leq 0.05\).}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure20.png}
\caption{Same as in Fig. 19 but for Hera.}
\end{figure}

\subsection{Magnification}

In Fig. 21 we can see the magnification maps for Ares, where the reconstruction is shown in the left panel, and the true map in the right panel. The external curve of high \(|\mu|\) exhibits an elongated shape, within which there is one internal curve surrounding each MD. The main difference appears close to the MD 2, where we get an overestimation of \(|\mu|\). Such overestimation coincides with a true critical curve, as one cannot notice by direct comparison between the true (red) and reconstructed (blue) critical curves in Fig. 24 (upper left panel). Despite this overestimation, we are not able to retrieve such critical curve. It is worth mentioning that it is possible to reproduce such critical curve by forcing the reconstruction parameters, where its presence is more frequent for Type 2 reconstructions. Nonetheless, in practice we are not aware of these details, so that the best we can do is to use the parameters that provide the best overall outcome.

It is clear that our reconstruction recovers the orientation, shape and size of such critical curves with great accuracy. Additionally, the
Figure 21. Magnification maps ($|\mu|$) for Ares, given for a source with redshift $z_r = 9$. They correspond to the maps shown in Fig. 16. The true map is shown in the left panel.

Figure 22. Same as in Fig. 21 but for Hera. Here, the black dashed contour encloses the critical curve produced by the substructure present in Hera.

critical curves corresponding to the substructure cannot be recovered, since it is difficult to account for such detailed peaks from the limited information provided by strong lensing.

For Hera, its magnification maps are shown in Fig. 22. In general terms, our reconstruction is capable of recovering the main features shown by the true map (Fig. 22 left panel), except for the curve (surrounded by the white dashed contour) produced by the substructure present in Hera, which our reconstruction fail to account for. As we discussed above, it is not possible to account for such substructure, due to the lack of observations there. Besides such curve, the external critical curve exhibits an elongated shape with one inner curve surrounding each MD; similarly to Ares. Fig. 24 (upper right panel) shows the direct comparison between true (red) and reconstructed (blue) critical curves. It is clear that the shape, and size of the main curves are reproduced by our reconstruction.

Since for Hera we have fewer systems/images than for Ares, it is expected for us to obtain less accurate curves. This can be seen in the external curve, where at its ends the reconstructions fail to accurately reproduce its shape. This becomes more evident in the critical curves, as it is depicted in Fig. 24 (upper left panel). Nonetheless, our results are incredibly accurate, considering that the only assumption that we have made with respect to the cluster itself relies on the NIE and 2NIE used as input. Beyond that, the reconstructions just adapt to the observations.

From Fig. 23 it is clear again that the higher difference between the true and reconstructed $|\mu|$ maps is produced close to the true critical curves (solid black lines), where the deviation from the true curves shows the overestimation in $|\mu|$ associated with the presence of the reconstructed critical curves. This is more evident for Hera.

Now, moving to the source plane, in Fig. 24 we have the true (red) and predicted (blue) caustic curves for Ares (lower left panel) and Hera (lower right panel). We successfully recover the overall morphological characteristics present in the caustic curves, except for
those curves associated with substructure. Our curves are deviated from the true ones, which is a consequence of (18).

Now, for the $|\mu|$ maps given by BHm, which are shown for Ares and Hera respectively in Fig. 19 (upper left panel) and Fig. 20 (upper left panel) in Meneghetti et al. (2017), it is clear that the irregularities present in Ares produce equally irregular critical curves. BHm is capable of reproducing the curves around the MD, and the external curve follows the orientation of the true curve. Nonetheless, the external curve produces several regions of high magnification that do not follow the true map. In contrast, our reconstruction has produced smoother curves which present less deviations from true critical curves. On the other hand, the BHm of Hera shows a critical curve smoother curves which present less deviations from true critical curves. Additionally, our reconstruction has produced the input deflection potential has on the reconstructions, where we discuss up to here, we are confident to say that 

![Graph](https://via.placeholder.com/150)

Figure 23. Relative difference $\Delta \mu$ between the reconstructed and true magnification maps for Ares (left panel) and Hera (right panel). The red and green solid contours correspond to $\Delta \mu = 0.2$ and $\Delta \mu = -0.2$, respectively. The black solid lines depict the true critical curves. Additionally, the black dashed contour encloses the critical curve produced by the substructure present in Hera.

Lastly, we consider the root-mean-square (rms) on the lens plane, define as

$$
\Delta_{rms} = \sqrt{\frac{1}{N_{img}} \sum_{i=1}^{N_i} \sum_{n=1}^{N_{ir}} \left( \theta_{in} - \theta_{in}^* \right)^2},
$$

where $N_i$ is the number of multiple images corresponding to the $i$-th multiply imaged system, and $N_{img} = \sum_{i=1}^{N_i} N_i$ is the total number of multiple images recovered from the reconstruction. Recall that $N_i$ is the number of multiple imaged systems used in the reconstruction, as well, $\theta_{in}$ and $\theta_{in}^*$ are the $n$-th reconstructed and true angular position of the $i$-th multiply imaged system, respectively. In general, $N_{ir} \neq N_{img}$, since it is possible to get an excess or lack of multiple images. In case of getting an excess of multiple images, we discard those which are outside a given neighborhood with respect to the closest true image.

A small rms is an indication of how well the reconstruction reproduces the input image positions. However, in practice it is perhaps more important to get a small prediction rms (prms), which is computed for a set of multiply imaged systems which were not considered as input data. Now, take into account that, as it is discussed in e.g Williams et al. (2018), a small rms by itself is a necessary but not sufficient criteria to qualify the overall quality of the reconstruction, particularly since it may be the indication of an overfitting of the intrinsic noise in the data. Thus, it becomes necessary to get small values for both, rms and prms. For further details in this regard see e.g James et al. (2013). We left the study of the prediction power of relensing for another paper.

For Ares we have $\Delta_{rms} \approx 0.17$ arcsec with $N_{img} = 240$ out of the total $N_{img} = 242$. Meanwhile, for Hera we get $\Delta_{rms} \approx 0.16$ arcsec with $N_{img} = 240$. Considering the rms aforementioned, as well as the reproduction of the different properties of Ares and Hera discussed up to here, we are confident to say that relensing is capable of producing accurate reproduction of the mass distributions of galaxy clusters, and their properties as gravitational lens.

6 SUMMARY AND CONCLUSIONS

In this work we describe and test a free-form method which makes use of gravitational lensing in order to produce an estimation of the mass profile of galaxy clusters, as well as its properties as a gravitational lens. In addition, we note that this approach does not consider that mass traces light; commonly assumed in parametric methods. This approach is an extension of the method presented in Bradač et al. (2005a, 2009), which at the same time is an extension of the work discussed in Bartelmann et al. (1996).

Here, we use an irregular and adaptive grid that is intended to produce a higher resolution around those mass peaks (main deflectors) responsible for most of the strong lensing effect, as well as around the multiple images. Additionally, we have opted for an alternative finite difference approach (generalized finite difference). Moreover, we include two different ways of computing the penalty function $\chi^2$, which we have named Type 1 and Type 2.

We start by testing our approach on two simple distributions; a NIS and 2NIS, in order to explore the set of input parameters that take us to reliable results. With them, we also explore the impact that the input deflection potential has on the reconstructions, where we use a blind reconstruction and a NIE. Within the region delimited by the available multiple images (inner region) there are no major differences. However, for blind reconstructions it is difficult to scale up
outside the inner region. The introduction of weak lensing improves the reconstruction on such region, but there is still a lack of mass towards the outskirts. The NIE helps to extend the reconstruction outside the inner region along with weak lensing, where for the 2NIS we get an improvement in the mass estimation of ~ 7%. For $\kappa$ most of the region presents an estimation that lies below ~ 20%, with just a few sectors mainly close to the boundaries where we get higher values. Thus, the choice of $\psi^{(0)}$ does not affect significantly the reconstruction where strong lensing is dominant. However, outside such region the reconstruction is more susceptible to $\psi^{(0)}$. If $\psi^{(0)}$ corresponds to an overestimation of $\kappa$ that moves away from the true map, it becomes difficult for the reconstruction to scale down towards the true value. The opposite effect occurs in blind reconstructions.

We also show that by recomputing the deflection potential from the weight of every node; i.e. by using (12), we are able to smooth the shape of convergence and magnification maps, which otherwise tend to be quite irregular and noisy. For instance, the overall relative difference $\Delta r \kappa$ and $\Delta r \mu$ improves. We refer to this process as smoothing.

This smoothing improves the power of this free-form approach, which we also test on Ares and Hera Meneghetti et al. (2017), two simulated galaxy clusters, which provide a more realistic framework to prove the reliability of our reconstructions. Our results show an improvement with respect to e.g. BHm (Bradac-Hoag models) in Meneghetti et al. (2017), where we get less irregular profiles and magnification maps, which turns into a more accurate reconstruction of these distributions.

Among the reconstructions discussed in this work, we have that within the inner region the reconstructions mainly satisfy $|\Delta r \kappa| < 0.1$ (10%), with 0.1 (10%) < $\Delta r \kappa$ < 0.2 (20%) mostly close to inner region boundary. Higher values are obtained most likely where small substructures are located, like in Ares. When weak lensing is included, the reconstruction provides a more accurate estimation of the mass profile outside the inner region. For $|\mu|$, the higher discrepancies appear close to the critical curves, as expected due to their nature. The morphological properties of the critical and caustic curves do not show significant changes when weak lensing is also used in the reconstruction.

Our reconstructions have shown to be capable of recovering the orientation, shape, and size of critical curves, despite the irregulari-
ties present in the reconstructions. We have found a way of reducing such irregularities by means of the smoothing. Nevertheless, since critical curves are sensitive to variations by nature, they might produce high deviations in $\mu$; even if such variations are small, this is a problem that requires further improvement.

We have seen that Type 1 and Type 2 reconstructions produce similar results, except for some minor differences that are related to the input parameters. Therefore, it is best to use both types of reconstruction for a given galaxy cluster, which is done by taking several realizations of the reconstruction for each cluster. As a result, we got a rms on the lens plane of about $0.17 \text{arcsec}$ and $0.16 \text{arcsec}$ for Ares and Hera. Also, we were able to recover 240 of the 242 images used in Ares, and all the images used in Hera.

relensing has shown an incredible capability in terms of producing an accurate estimation of the mass profile of galaxy clusters, along with their properties as gravitational lens. Therefore, with this work we expect to provide an accessible package written in python, focused on the characterization of galaxy clusters by using gravitational lensing. This provides users with a wider range of options to choose from, which becomes indispensable in order to validate physical results.

In this regard, the application of relensing to real galaxy clusters, such as those from e.g. HFF, is naturally the next step (this is a work in progress). Also, with the expected increase in quantity and quality of observations from upcoming facilities, we plan to explore how relensing behaves depending on the number of multiply imaged systems, and multiple images available. Additionally, as already discussed in Cain et al. (2016), flexion allows us to get information from substructure where strong lensing is not dominant. So that it becomes interesting to explore the impact the smoothing has on reconstructions when flexion is included. Moreover, it is indispensable to test the prediction power of relensing particularly in the context of predicting the existence of multiple images, which is of great importance for strong lensing studies.

DATA AVAILABILITY

The data discussed in this work will be shared upon request to the authors. We plan to make relensing available through GitHub in the near future.

REFERENCES

Abdelsalam H. M., Saha P., Williams L. L. R., 1998, MNRAS, 294, 734
Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Astropy Collaboration et al., 2013, A&A, 558, A33
Astropy Collaboration et al., 2018, AJ, 156, 123
Bartelmann M., 2003, arXiv preprint astro-ph/0304162
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Bartelmann M., Narayan R., Seitz S., Schneider P., 1996, ApJ, 464, L115
Benito J., Urena F., Gavete L., 2001, Applied Mathematical Modelling, 25, 1039
Benito J., Urena F., Gavete L., Alvarez R., 2003, Computer methods in applied mechanics and engineering, 192, 735
Bergamini P., et al., 2021, A&A, 645, A140
Blandford R. D., Narayan R., 1992, ARA&A, 30, 311
Bradač M., Schneider P., Lombardi M., Erben T., 2005a, A&A, 437, 39
Bradač M., et al., 2005b, A&A, 437, 49
Bradač M., et al., 2009, ApJ, 706, 1201
Broadhurst T., et al., 2005, ApJ, 621, 53
Cacciato M., Bartelmann M., Meneghetti M., Moscardini L., 2006, A&A, 458, 349
Cain B., Bradač M., Levinson R., 2016, MNRAS, 463, 4287
Cataneo M., Rapetti D., 2018, International Journal of Modern Physics D, 27, 1848006
Coe D., Fuselier E., Benitez N., Broadhurst T., Frye B., Ford H., 2008, ApJ, 681, 814
Coe D., et al., 2013, ApJ, 762, 32
Coe D., et al., 2019, ApJ, 884, 85
D’Aloisio A., Natarajan P., 2011, MNRAS, 411, 1628
Deb S., Goldberg D. M., Ramdass V. J., 2008, ApJ, 687, 39
Diego J. M., Protopapas P., Sandvik H. B., Tegmark M., 2005, MNRAS, 360, 477
Diego J. M., Tegmark M., Protopapas P., Sandvik H. B., 2007, MNRAS, 375, 958
Diego J. M., et al., 2016, MNRAS, 456, 356
Furtak L. J., Atek H., Lehnert M. D., Chevallard J., Charlot S., 2021, MNRAS, 501, 1568
Gardner J. P., et al., 2006, Space Sci. Rev., 123, 485
Gavete L., Gavete M., Benito J., 2003, Applied Mathematical Modelling, 27, 331
Ghosh A., Williams L. L. R., Liesenborgs J., 2020, MNRAS, 494, 3998
Gilmore J., Natarajan P., 2009, MNRAS, 396, 354
Golse G., Kneib J. P., 2002, A&A, 390, 821
Harris C. R., et al., 2020, Nature, 585, 357
Hattori M., Kneib J., Makino N., 1999, Progress of Theoretical Physics Supplement, 133, 1
Hoekstra H., 2013, arXiv e-prints, p. arXiv:1312.5981
Hoekstra H., Bartelmann M., Dahle H., Israel H., Limousin M., Meneghetti M., 2013, Space Sci. Rev., 177, 75
Huber K., Tchernin C., Merten J., Hilbert S., Bartelmann M., 2019, A&A, 627, A143
Hunter J. D., 2007, Computing in Science & Engineering, 9, 90
Ivezic Z., et al., 2019, ApJ, 873, 111
James G., Witten D., Hastie T., Tibshirani R., 2013, An introduction to statistical learning. Springer
Jauzac M., et al., 2016, MNRAS, 457, 2029
Jee M. J., et al., 2007, ApJ, 661, 728
Julio E., Kneib J. P., 2009, MNRAS, 395, 1319
Julio E., Kneib J. P., Limousin M., Eliasdóttir A., Marshall P. J., Verdugo T., 2007, New Journal of Physics, 9, 447

MNRAS 000, 1–24 (2021)
