

Sparse Lens Inversion Technique (SLIT): lens and source separability from linear inversion of the source reconstruction problem[★]

R. Joseph¹, F. Courbin¹, J.-L. Starck², and S. Birrer³

¹ Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland

e-mail: remy.joseph@epfl.ch

² Laboratoire AIM, CEA/DSM-CNRS-Université Paris Diderot, Irfu, Service d'Astrophysique, CEA Saclay, Orme des Merisiers, 91191 Gif-sur-Yvette, France

³ Department of Physics and Astronomy, University of California, Los Angeles, 475 Portola Plaza, Los Angeles, CA 90095-1547, USA

Received 25 April 2017 / Accepted 27 November 2018

ABSTRACT

Strong gravitational lensing offers a wealth of astrophysical information on the background source it affects, provided the lensed source can be reconstructed as if it was seen in the absence of lensing. In the present work, we illustrate how sparse optimisation can address the problem. As a first step towards a full free-form-lens-modelling technique, we consider linear inversion of the lensed source under sparse regularisation and joint deblending from the lens light profile. The method is based on morphological component analysis, assuming a known mass model. We show with numerical experiments that representing the lens and source light using an undecimated wavelet basis allows us to reconstruct the source and to separate it from the foreground lens at the same time. Both the source and lens light have a non-analytic form, allowing for the flexibility needed in the inversion to represent arbitrarily small and complex luminous structures in the lens and source. In addition, sparse regularisation avoids over-fitting the data and does not require the use of an adaptive mesh or pixel grid. As a consequence, our reconstructed sources can be represented on a grid of very small pixels. Sparse regularisation in the wavelet domain also allows for automated computation of the regularisation parameter, thus minimising the impact of the arbitrary choice of initial parameters. Our inversion technique for a fixed mass distribution can be incorporated into future lens-modelling techniques iterating over the lens mass parameters.

Key words. gravitational lensing: strong – methods: data analysis – techniques: image processing – galaxies: high-redshift

1. Introduction

Strong gravitational lensing is a powerful tool to study astrophysics and cosmology, from structures on a wide range of scales to probing dark matter and dark energy. At cluster scale, strong gravitational lensing allows us to map the distribution of dark matter and to compare it with the visible mass (e.g. see recent work in the Hubble Frontier Fields; [Piewe et al. 2017](#); [Harvey et al. 2016](#); [Sebesta et al. 2016](#); [Diego et al. 2016](#)). In merging or interacting galaxy clusters, it is used to infer a limit on the dark matter cross-section (e.g. [Harvey et al. 2015](#)). At galaxy scale, strong lensing is a powerful way of studying the interplay between visible and dark matter, and therefore also galaxy formation and evolution. The growing sample of known galaxy-scale strong lenses includes early-type galaxies (SLACS; [Bolton et al. 2008](#)), spiral galaxies (SWELLS; [Treu et al. 2011](#)), emission-line galaxies (BELLS; [Bolton et al. 2012](#)), groups of galaxies ([More et al. 2012](#)) and even lensing by AGNs ([Courbin et al. 2012](#)). Strong lensing is very sensitive to small substructures in the lensing galaxy and consists in a unique

way of detecting such small structures, hinting at the nature of dark matter and the cosmological model (e.g. [Vegetti et al. 2010](#)). Finally, the measurement of the so-called time delay between the images of strongly lensed quasars ([Refsdal 1964](#)) allows to constrain cosmology independently of any other cosmological probes (for a review see [Treu & Marshall 2016](#)), with high sensitivity to H_0 and little dependence on the other cosmological parameters (e.g. [Suyu et al. 2010, 2014](#); [Tewes et al. 2013](#); [Bonvin et al. 2017](#)).

Most of these applications require reliable mass models for the lens, which in turn requires robust methods to “unlens” the original source, whatever the level of complexity in the shape may be. They also require the light of the foreground lens to be deblended from the background source either prior to the mass modelling or together with it, in a single-step process. Different methods exist to carry out these tasks, all with their own assumptions, advantages, and drawbacks. In the case of cluster-scale lensing, the positions of multiply imaged background sources give sufficient constraints on the mass macro-models. However, additional information about the distortion of extended images might allow to break certain degeneracies in the cluster lens models while allowing us to recover the morphologies of highly magnified sources to probe galaxy evolution at high redshift, as recently illustrated in [Cava et al. \(2018\)](#).

[★] The python package corresponding to the algorithms described in this article can be downloaded via the github platform at <https://github.com/herjy/SLIT>.

With the quality of modern lens observations, taken with the HST or ground-based adaptive optics, it becomes increasingly important to design modelling techniques, both for galaxies and clusters, that are flexible and robust: they must simultaneously capture all possible observational constraints and remain robust with respect to noise. This is particularly important to explore the degeneracies inherent to the lensing phenomenon (e.g. [Schneider & Sluse 2013, 2014](#)) without applying strong priors both on the source and lens shape.

Some of the modelling techniques currently in use consider a full analytical lens mass and light distribution ([Kneib et al. 2011](#); [Bellagamba et al. 2017](#); [Oguri 2010](#)). Others use a semi-analytic approach, where the source is pixelated and regularised but where the lens has an analytical representation ([Dye & Warren 2005](#); [Warren & Dye 2003](#); [Suyu et al. 2006](#); [Vegetti & Koopmans 2009](#)) or where the lens is represented on a pixelated grid with regularisation or assumptions on its symmetry (e.g. [Coles et al. 2014](#); [Nightingale et al. 2018](#)). Further studies in this direction involve adaptive pixel grids to represent the source ([Nightingale et al. 2018](#); [Nightingale & Dye 2015](#)) or an analytic decomposition of the source on a predefined dictionary as was done in [Birrer et al. \(2015\)](#), where the authors used shapelets ([Refregier 2003](#)). In the following, a dictionary is a collection of atoms (vectors) that together form a generative set of \mathbb{R}^n , with n being the number of samples in the data.

In the present work we address the problems of source reconstruction and deblending as a single linear inverse problem. By using a family of functions called starlet (or undecimated isotropic wavelet transform [Starck et al. 2005](#)), we are able to use sparse regularisation over the lens and source light profiles. Sparsity with starlets has the advantage of performing model-independent reconstructions of smooth profiles and allows for deterministic expression of the regularisation parameter. As the lensed source can be represented using only a limited number of starlet coefficients, the pixel grid can be almost as fine as desired and the reconstructed source is denoised and deconvolved from the instrumental point spread function (PSF).

The choice of wavelets is motivated by the successful use of this decomposition in recent years to model astronomical objects ([Ngolè Mboula et al. 2015](#); [Garsden et al. 2015](#); [Lanusse et al. 2016](#); [Farrens et al. 2017](#); [Joseph et al. 2016](#); [Livermore et al. 2017](#); [Pratley et al. 2018](#)). In the present application, the separation between lens and source light is performed through morphological component analysis ([Starck et al. 2005, 2015](#)). The method relies on the sparsity of the lens and source light profiles in their respective dictionaries to perform the separation. We apply sparse regularisation to the source and lens light distributions and illustrate with numerical simulations the effectiveness of the method at reconstructing the source and deblending lens and source.

This paper is intended as a proof of concept to show how convex optimisation under a sparsity prior for the source light profile can be used as an adequate minimisation technique for lens modelling. The scope of this paper remains limited to the modelling of light distribution alone and to the potential of using morphological component analysis to provide a new framework for lens modelling. Our ultimate goal is to come up with a full lens-modelling technique that would perform joint free-form modelling of the light and mass density profiles. The latter being a non-linear problem and the problem of free-form lens modelling as a whole being largely underconstrained and degenerate, this will be the subject of another paper.

The paper is organised as follows: Sect. 2 introduces the basics of strong gravitational lensing. In Sect. 3 we describe

the problem of lens/source light profile estimation as a linear inverse problem. In Sect. 4 we present the detail of our method and Sect. 5 describes the testing of our method on simulated images and the comparison with a state-of-the-art technique: Lenstronomy ([Birrer & Amara 2018](#)). Finally, Sect. 6 details the content of the public package we produced to release our code.

2. Source reconstruction given a known lens mass

In this section, we give the basics of the gravitational lensing formalism we use to back-project lensed images to the source plane.

We note θ the angular position on the sky of an object seen through a gravitational lens (image plane coordinates), with intrinsic angular position β (source plane coordinates). The mapping from source to image plane is described by the lens equation:

$$\beta = \theta - \alpha(\theta), \quad (1)$$

where

$$\alpha(\theta) = \frac{1}{\pi} \int_{\mathbb{R}^2} \kappa(\theta') \frac{\theta - \theta'}{|\theta - \theta'|^2} d^2\theta', \quad (2)$$

which is the deflection angle in the lens plane. The value $\kappa(\theta)$ is the dimensionless convergence of the gravitational lens at position θ . Convergence κ is defined as:

$$\kappa = \frac{\Sigma(D_L\theta)}{\Sigma_c}, \quad (3)$$

where Σ_c is the critical surface mass density defined as

$$\Sigma_c = \frac{c^2}{4\pi G} \frac{D_S}{D_L D_{LS}}, \quad (4)$$

and $\Sigma(D_L\theta)$ is the surface mass density of the lens defined as the projected volume density of the lens on a plane perpendicular to the line of sight. Parameters D_S , D_L , and D_{LS} are the angular diameter distances between the observer and the source, between the observer and the lens, and between the lens and the source, respectively.

The problem of inverting Eq. (1) from photometric observations only (meaning only θ is known) is a non-linear and highly under-constrained problem with two unknowns: The position of the source, β , and the convergence map of the lens. In the case of extended sources, the goal of lens inversion is to recover the light profile of a lensed galaxy as seen in the source plane, which implies being able to calculate the flux at each position β knowing the flux at position θ .

In practice, current techniques for lens inversion rely on an iterative process that consists in successively reconstructing the source profile brightness and the κ map. In an effort to develop an automated, model-independent method for lens inversion, we choose to decompose the problem. In this paper, we address the problem of source light profile reconstruction given a known convergence map, in which case the problem is linear. In [Warren & Dye \(2003\)](#), the authors express the mapping between source and image surface brightness using an operator F_κ , such that the observed surface brightness of a lensed galaxy can be written

$$Y = F_\kappa S + Z, \quad (5)$$

where Y is the observed surface brightness, flattened as a vector with length N_{pix} , vector S is the unknown source surface brightness vector in the source plane with length N_{ps} , F_κ is a $N_{\text{ps}} \times N_{\text{pix}}$ matrix where, following the formalism of Warren & Dye (2003), element $f_{i,j}$ is the j th pixel in image plane of the mapping of a source that has only its i th pixel set to one. In other words, F_κ indicates which pixels from the source plane have to be combined to predict the value of a pixel in the image plane. The elements of F_κ are entirely determined by the mass density distribution κ . Vector Z is an additive noise map. In this work, we consider Z as a white Gaussian noise with standard deviation σ , but the method can easily be extended to Poisson statistics, or more generally to noise with known root mean square.

2.1. Pixel-to-pixel mapping

As illustrated in Fig. 1 of Wayth & Webster (2006), a square pixel in the image plane is a diamond-shaped pixel in the source plane, with a total area depending on the magnification at the pixel location. Although this phenomenon should strictly be accounted for, we choose to make the approximation that each photon hitting a pixel whether in the source or image plane hits at the centre of the pixel. This way, we lose part of the information provided by the distribution of photons over the whole surface of a pixel, but importantly avoid correlating the noise when back projecting from the image plane to the source plane. Furthermore, using a small pixel size limits this imperfect modelling and allows to compensate for the variation in the light profile. This inverse problem being ill-posed, it admits no unique and stable solution, hence calling for a regularisation to solve it. Increasing the pixel size, and therefore losing resolution can be seen as a naive regularisation. We will show how advanced regularisation techniques can be used efficiently.

2.2. Projection and back-projection between source and lens planes

To compute the elements of F_κ for each of the N_{ps} pixels in the source plane, we associate a pixel in the image plane by shooting a photon from the centre, β , of a source plane pixel and recording the position(s), θ , given by Eq. (1), of the pixel(s) where the deflected photon hits the image plane. Summarising, we record the positions, θ , where $\beta + \alpha(\theta) - \theta = 0$. The element of F_κ at position(s) (β, θ) is (-are in the case of a multiply imaged pixel) set to one to indicate the mapping. Since F_κ is a sparse matrix with only very few non-zero coefficients, we choose to only store the positions (β, θ) that map into one another in order to save memory and thus computation time in the following steps.

The projection of a source light profile into a lensed light profile in the image plane is then performed by allocating to each image pixel the sum of the intensities of the corresponding source pixels according to F_κ . Conversely, back-projection is performed by allocating to each source pixel the average value of all its lensed counterparts according to F_κ . This ensures conservation of surface brightness between the source and lens planes.

3. Linear development of strong gravitational lens imaging

In real imaging data of strong gravitational lenses, the problem of finding the delensed light profile of a lensed galaxy is harder than solving Eq. (5), which is already non-trivial. First, one has to include the impulse response of the instrument that acquired

the image. This effect corresponds to a convolution of the images described by $F_\kappa S$, by the point spread function (PSF). Letting the linear operator H account for the convolution by a known PSF, Eq. (5) becomes

$$Y = HF_\kappa S + Z, \quad (6)$$

which is the problem one has to solve when dealing only with the lensed light profile of a source, assuming that the light profile from the foreground lens galaxy has been perfectly removed prior to the analysis.

In practice, images of strongly lensed galaxies are contaminated by light from a foreground lens galaxy, G . Taking this into account, Eq. (6) can then be written as

$$Y = H(F_\kappa S + G) + Z. \quad (7)$$

When Z is a white Gaussian noise, solving Eq. (7) reduces to finding S and G such that

$$\|Y - H(F_\kappa S + G)\|_2^2 < \epsilon, \quad (8)$$

where ϵ accounts for the precision of the reconstruction and depends on the noise level. Given that we have an N_{pix} -sized image and aim at finding an N_{pix} -sized galaxy light profile and an N_{ps} -sized source light profile, we need to impose further constraints on these unknowns. Classically, the light distribution of the lens is approximated by an analytic profile such as a Sérsic or deVaucouleur profile.

As reconstructing the source light profile is an ill-posed problem, where unknowns largely outnumber the number of observables, several strategies have been investigated in the literature; for example, adaptive pixel grids (Dye & Warren 2005), negentropy minimisation (Warren & Dye 2003; Wayth & Webster 2006), Bayesian inference over the regularisation parameters of the source (Suyu et al. 2006), perturbative theory (Alard 2009), and model profile fitting (Bellagamba et al. 2017). Although these methods have their own advantages and disadvantages, only a few of them are able to reconstruct complex sources without degrading the resolution of the output. In Birrer et al. (2015), the authors used a family of functions to reconstruct the source light profile with promising results. Here we propose to push this idea further by exploiting a family of functions that is well suited to representing galaxies, and that possesses properties of redundancy allowing for the use of sparse regularisation.

3.1. Sparse source reconstruction in the absence of light from the lens

We propose a new approach to solving Eq. (6). Given that galaxies are compact and smooth objects, their decomposition over the starlet dictionary (Starck et al. 2007) will be sparse, meaning that only a small number of non-zero starlet coefficients will contain all the information in a galaxy image. This property allows us to constrain the number of coefficients used in starlet space to reconstruct galaxy profiles, therefore offering a powerful regularisation to our problem. Starlets are a class of discrete undecimated, isotropic wavelets, formally introduced in Starck et al. (2007) and extensively described with regard to computation algorithms in Starck et al. (2015).

Assuming a signal is sparse in a dictionary Φ , the solution to an inverse problem like in Eq. (6) is the solution that uses the least number of coefficients in the Φ dictionary while minimising the square error between the observables and the reconstruction. In a more formal way, sparsity is enforced by minimising the ℓ_1 -norm of the decomposition over Φ of a signal known to

be sparse in this dictionary. In addition, because the mapping of an image from lens plane to source plane does not conserve shapes, the edges of the image in the lens plane does not match the borders of the image in the source plane, leaving parts of the source image unconstrained as they map into pixels outside the field of view of the lens plane. Let us call \mathbb{S} the set of pixels in the source plane that have an image in the lens plane. We impose that the coefficients of the solution outside set \mathbb{S} be set to zero. This allows us to write the problem of finding S as an optimisation problem of the form

$$\operatorname{argmin}_{\alpha_S} \|Y - HF_k \Phi \alpha_S\|_2^2 + \lambda \|W \odot \alpha_S\|_1, \quad (9)$$

where Φ is the starlet dictionary, and α_S are the starlet coefficients of S such that $\alpha_S = \Phi^T S$. The operator \odot is the term-by-term multiplication operator, and W is a vector of weights that serves the purpose of setting all coefficients outside \mathbb{S} to zero, while keeping the ℓ_1 -norm constraint from biasing the results (discussed further in the following paragraphs). In practice, minimising the ℓ_1 -norm of a vector is done by soft-thresholding the vector. This consists in decreasing the absolute value of the vector's coefficients by a positive value λ , and by setting the coefficients smaller than λ to zero, as shown in the following equation:

$$\operatorname{ST}_\lambda(x) = \begin{cases} \operatorname{sign}(x) \times (|x| - \lambda) & \text{if } |x| > \lambda \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

The regularisation parameter λ controls the trade-off between fitting the observed data and enforcing sparse solutions. From the definition of Eq. (10), it appears that solutions derived with soft-thresholding will present a bias due to the subtraction by λ . In order to mitigate this effect, we use the reweighting scheme from Candes et al. (2008). In order to prevent the most significant coefficients from being truncated, we multiply the regularisation parameter λ by

$$W = \frac{2}{1 + \exp(-10(\lambda - \alpha_0))}, \quad (11)$$

where α_0 is the solution of Eq. (9) with $W = 1$. With this definition for W , the coefficients that are much larger than λ are less affected by soft thresholding than others. Values of W for coefficients outside \mathbb{S} are set to infinity, hence naturally ensuring that the corresponding wavelet coefficients are set to zero.

This approach can be used to recover the source light profile in systems involving a faint foreground lens galaxy, a large Einstein radius, or when a reliable deblending of the lens and source light profiles is available prior to the source reconstruction scheme presented here.

3.2. Source reconstruction and deblending of the foreground lens light profile

In a more general case, one has to deal with the separation between the lens and source light profiles. Although several techniques allow for their separation prior to the analysis of the lens system (e.g. Joseph et al. 2014, 2016; Brault & Gavazzi 2015), each of these methods has limitations in the sense that they require specific inputs (field of view or multiband images) or do not take into account the lensed source profile when fitting the lens, resulting in potential biases. Another approach consists in fitting an analytic lens light profile while reconstructing the lens mass density profile and the source (Birrer et al. 2015; Tessore et al. 2016). Here, we propose a solution to reconstruct

and separate the lens and source light profiles using the fully linear framework provided by morphological component analysis (MCA; Starck et al. 2005).

Very importantly, a galaxy is sparse in starlets in its own plane (source or image), meaning that, given a mapping F_k with $\kappa(\theta) > 1$ at several positions θ between source and lens plane, a galaxy in the source plane and a galaxy in the lens plane are both sparse in starlets. We justify and illustrate this statement with simulations in Sect. 5 using simulated lenses.

Morphological component analysis allows for separation of two mixed components in a signal based on the fact that each component can be sparsely represented in its own dictionary but not in others. In the context of lens-source separation, the explicit dictionaries are the starlet transform of a back-projection in the source plane on one hand, and the starlet transform for the lens in the lens plane on the other.

We can therefore iteratively project the mixed signals in their own respective dictionaries, impose a sparsity constraint on each projection and therefore reconstruct the corresponding components separately. As seen in Sect. 3.1, sparsity is imposed by minimising the ℓ_1 -norm of both decompositions. Due to the problem of reconstructing S being ill-posed, a sparse version of S has to be computed at each iteration, which requires subiterations.

3.3. Optimisation problem

In mathematical terms, the aforementioned MCA problem boils down to finding the model $\{\hat{S}, \hat{G}\}$ that provides the best approximation of the data set Y according to Eq. (7), while minimising the ℓ_1 -norms of the starlet coefficients α_S and α_G , with $\alpha_G = \Phi^T G$. This is written as

$$\hat{\alpha}_S, \hat{\alpha}_G = \operatorname{argmin}_{\alpha_S, \alpha_G} \|Y - H(F\Phi\alpha_S - \Phi\alpha_G)\|_2^2 + \lambda_S \|W_S \odot \alpha_S\|_1 + \lambda_G \|W_G \odot \alpha_G\|_1, \quad (12)$$

where λ_S and λ_G account for the sparsity of α_S and α_G , respectively. Similarly to Eq. (9), W_S and W_G are weights that play the same role as in Eq. (11). We describe the calculation of these values in Sect. 4.1.

Since, our main interest is to fully reconstruct the source S , it is not necessary to compute the fully deconvolved vector G . Instead, we limit ourselves to estimating the convolved vector $G_H = HG$ so that we extract the convolved foreground lens galaxy G_H and decrease the computational time by avoiding several convolution steps of G when solving Eq. (12), which becomes

$$\hat{\alpha}_S, \hat{\alpha}_{G_H} = \operatorname{argmin}_{\alpha_S, \alpha_{G_H}} \|Y - HF\Phi\alpha_S - \Phi\alpha_{G_H}\|_2^2 + \lambda_S \|\alpha_S\|_1 + \lambda_{G_H} \|\alpha_{G_H}\|_1. \quad (13)$$

4. Method: the SLIT algorithms

In this section, we describe the two algorithms, SLIT and SLIT_MCA that we implemented to solve Eq. (9) (no lens light) and Eq. (13) (full light reconstruction problem), respectively.

4.1. Source delensing: SLIT algorithm

Starting with the simpler case of solving Eq. (6), we made use of the fast iterative soft thresholding algorithm (FISTA; Beck & Teboulle 2009). This iterative algorithm is similar to a forward backward (Gabay 1983) algorithm with an inertial step

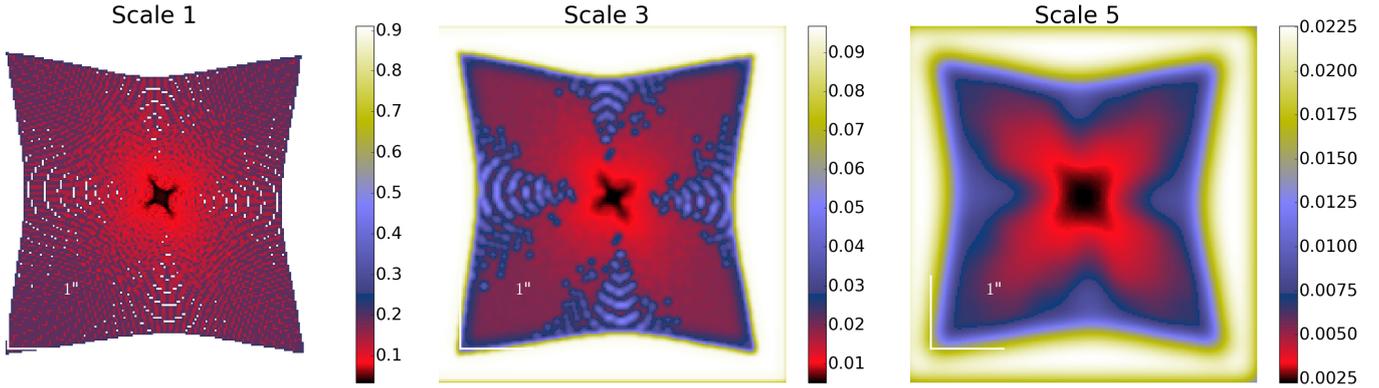


Fig. 1. Noise levels in the source plane (λ_S) for three starlet scales (scale 1, 3 and 5) out of the five computed for 100×100 pixel images with noise standard deviation $\sigma = 1$.

that accelerates the convergence. We show the pseudo-algorithm for one iteration in algorithm 1. It consists in a step of gradient descent (steps 2 and 3) followed by a soft-thresholding of the starlet coefficients of the source (step 4), which acts as a sparse regularisation. Step 6 aims at pushing forward the solution in the direction of smaller error, which accelerates the convergence. The process is repeated until convergence, as shown in algorithm 2.

Algorithm 1 FISTA iteration

```

1: procedure FISTA( $Y, \alpha_i, H, F, \xi_i, t_i, \lambda, W$ )
2:    $R \leftarrow F^T H^T (Y - HF\Phi\alpha_i)$ 
3:    $\gamma \leftarrow \xi_i + \mu\Phi^T R$ 
4:    $\alpha_{i+1} \leftarrow ST_{\lambda\otimes W}(\gamma)$ 
5:    $t_{i+1} \leftarrow \frac{1 + \sqrt{1 + 4t_i^2}}{2}$ 
6:    $\xi_{i+1} \leftarrow \alpha_i + \frac{t_i - 1}{t_{i+1}}(\alpha_{i+1} - \alpha_i)$ 
7:   return  $\xi_{i+1}, \alpha_{i+1}, t_{i+1}$ 
8: end procedure
    
```

In this algorithm, Y is the original image of a lensed galaxy, ξ_i and γ are local variables used to carry local estimates from one iteration to another, α_i is the starlet decomposition of the estimated source at iteration i and t_i gives the size of the inertial step. This sequence has been chosen to ensure that the cost function convergence is bounded by the Euclidean distance between the starting point for S and a minimum of the cost function. This is explained in more detail in [Chambolle & Dossal \(2015\)](#) (theorem 1) and [Beck & Teboulle \(2009\)](#) (theorem 4.1). The gradient step μ is chosen to be $\mu = (\|HF\Phi\|_s^2)^{-1}$, with $\|\cdot\|_s$ being the spectral norm of a matrix, defined by

$$\|M\|_s = \max_{x \neq 0} \frac{\|Mx\|_2}{\|x\|_2}. \quad (14)$$

The function ST_λ in algorithm 1 is the soft-thresholding operator described by Eq. (10).

Parameter λ has to be chosen with care as it accounts for the sparsity of the solution. In practice, λ is a threshold that is applied to each starlet coefficient of the solution in order to reduce its ℓ_1 -norm in starlet space. In the present case, given the presence of noise in the input data Y , it is important to choose a threshold above the noise levels. This is done by propagating the noise levels in image Y to the starlet coefficients α_i . The starlet transform being an undecimated multi-scale transform, coefficients α_i can be ordered as a set of images, each image representing the variations in the data at different scales. Therefore,

Algorithm 2 SLIT

```

1: procedure SLIT( $Y, H, F, \lambda, Niter, W$ )
2:    $\xi_0, \alpha_0, t_0 \leftarrow 0, 0, 1$ 
3:   for  $0 < i \leq Niter$  do
4:      $\xi_{i+1}, \alpha_{i+1}, t_{i+1} \leftarrow FISTA(Y, \alpha_i, H, F, \xi_i, t_i, \lambda, W)$ 
5:   end for
6:    $S \leftarrow \Phi\alpha_{Niter}$ 
7:    $W \leftarrow \frac{2}{1 + \exp(-10(\lambda - \alpha_0))}$ 
8:   return  $S, W$ 
9: end procedure
    
```

we have to estimate how noise levels translate from the data to each scale of the starlet transform. In the current implementation, the noise from image Y also has to be propagated through the H^T and F^T operators as shown in step 2 of algorithm 1. Because the convolution H^T correlates the noise in the data and the back-projection to source plane induces varying multiplicity of the delensed pixels across the field of view, it is necessary to estimate a different threshold λ at each pixel location in each scale of the starlet transform of the source. In practice, for measurements affected by noise with known covariance Σ , noise standard deviation in the starlet domain of the source plane are given by the square root of the diagonal elements of

$$\Sigma_S = \Phi^T F_\kappa^T H^T \Sigma H F_\kappa \Phi. \quad (15)$$

In the case of the ‘‘a trou’’ (french for ‘‘with whole’’, [Holschneider et al. 1989](#); [Shensa 1992](#)) algorithm, which relies on filter bank convolution to perform the starlet transform, the elements of Φ are never explicitly calculated. Instead, the noise standard deviation at scale s and pixel p in the starlet domain of the source plane is given by

$$\Xi_{s,p}^2 = \Delta_s^2 * \Gamma_{i,j(i=j)}^2, \quad (16)$$

where Δ_s is the starlet transform of a dirac function at scale s , $*$ is the convolution operator, and $\Gamma_{i,j(i=j)}$ is the vector containing the diagonal elements of $F_\kappa^T H^T \Sigma H F_\kappa$.

The result is the noise level in source space at each location and each scale of the starlet transform. By construction, the last scale contains the coarse details in the image and is left untouched in the thresholding process. Figure 1 shows the noise levels in the source plane calculated from a simulation where the surface mass density is a singular isothermal ellipsoid (SIE). The PSF is a simulated Tiny Tim PSF ([Krist et al. 2011](#)) for the F814W filter of the ACS/WFC instrument on the HST and

the noise standard deviation is set to 1. The original image is a 100×100 pixel image which is decomposed into six starlet scales, that is, the maximum number of scales that we can possibly compute, given the size of the image.

4.2. Deblending and source delensing: SLIT_MCA algorithm

In real data, the source and lens light profiles are blended, that is, the light of the lens impacts the quality of the source reconstruction. Handling the deblending and delensing simultaneously can be done using MCA.

In classical source separation problems where two components are to be separated, solutions are obtained through MCA by performing a gradient step and alternatively regularising over each component in its own dictionary. In the present case, solving Eq. (7) requires that we solve an inverse problem each time we aim at reconstructing the quantity $HF_\kappa S$. This inverse problem corresponds to solving Eq. (6) for which we already presented a solution in algorithm 2.

Our MCA algorithm is therefore an iterative process that consists in alternatively subtracting a previous estimation of G_H and S from the data:

$$D_S = Y - G_H \quad (17)$$

and

$$D_G = Y - HF_\kappa S, \quad (18)$$

as detailed in algorithm 3. At each iteration, the previous subtractions D_S and D_G are used to estimate S and G_H , respectively. Estimating S requires running the full SLIT algorithm on D_S until convergence. Estimating G_H at a given iteration is simply done by running one single iteration of the FISTA algorithm on D_G with inputs F_κ and H being identity matrices. We found empirically that using the projections of S and G_H on the subsets of vectors with positive values (meaning that we set all negative coefficients to zero), we could achieve faster convergence towards more realistic solutions. Although this is not a formal positivity constraint, since we do not apply positivity on the solutions themselves, we found that in practice this leads to galaxy profiles with less negative structures, which is not a physical feature we find in galaxy light profiles.

Estimating λ_G is as crucial as estimating λ_S but is much simpler given that there is no inverse problem to solve in this case. The threshold λ_G only depends on the noise level in the image. Given that we impose sparsity in starlet space, we still have to evaluate noise levels at each scale of the starlet transform. To do so, we simply compute how a unitary signal in direct space translates into starlet space and multiply it by the noise standard deviation. In other words, we take the starlet transform of a Dirac function and compute the 2-norm of each scale of the starlet transform. This tells us how energy is distributed into starlet space. For a decomposition over six starlet scales, the values we obtain for the first five scales in order of increasing scale are: $\lambda_G = [0.891, 0.200, 0.086, 0.041, 0.020]$. As in the previous section, the last scale is left untouched. The obtained values are then multiplied by a scalar that accounts for the desired detection level in units of noise. The scalar is often chosen to be between 3 and 5σ as seen previously. A detection at 3σ will produce very complete but noisy reconstruction of the signal, whereas a detection at 5σ will lead to a more conservative reconstruction of the highest-S/N features only. The obtained thresholds are applied uniformly across each scale.

Algorithm 3 SLIT MCA algorithm

```

1: procedure SLIT_MCA( $Y, H, F_\kappa, N_{iter}, N_{subiter}, \lambda_S, \lambda_G$ )
2:    $\mathcal{S} \leftarrow 0$ 
3:    $\tilde{G}_H \leftarrow 0$ 
4:    $[\xi_{S0}, \xi_{G0}] = [0, 0]$ 
5:    $[\alpha_{S0}, \alpha_{G0}] = [0, 0]$ 
6:    $[t_{S0}, t_{G0}] = [1, 1]$ 
7:   for  $0 < i \leq N_{iter}$  do
8:      $D_S \leftarrow Y - \tilde{G}_H$ 
9:      $S \leftarrow SLIT(D_S, H, F_\kappa, \lambda_S, N_{subiter})$ 
10:     $D_G \leftarrow Y - HF_\kappa S$ 
11:     $\xi_{Gi}, \alpha_{Gi}, t_{Gi} \leftarrow FISTA(D_G, \alpha_{Gi-1}, I_d, I_d, \xi_{Gi-1}, t_{Gi-1}, \lambda_G, 1)$ 
12:     $G_H \leftarrow \Phi^T \alpha_{Gi}$ 
13:  end for
14:  return  $\mathcal{S}, G$ 
15: end procedure

```

5. Numerical experiments with simulations

In the following we illustrate the performances of our algorithms with numerical experiments that mimic different observational situations. We also apply our algorithms to a set of simulated images and show comparisons of reconstructions with a state-of-the-art method: lenstronomy (Birrer & Amara 2018).

5.1. Creating realistic simulated lenses

In order to make the simulations as realistic as possible, we use galaxy light profiles extracted from deep HST/ACS images taken in the F814W filter. The images are part of the Hubble Frontier Fields program and the specific data we used¹ were taken from the galaxy cluster Abell 2744 (Lotz et al. 2017). We selected various patches, each containing a galaxy that we use to represent a lens or a source. The HFF images were cleaned using starlet filtering with a $5\text{-}\sigma$ threshold. Source galaxies were chosen to display visually apparent substructures with several modes, which we aim at recovering with our lens inversion methods. The lens light profile was chosen to present a smooth monomodal distribution, as expected for a typical massive early-type galaxy, such as for example in the SLACS samples (Bolton et al. 2008).

To generate our simulations we then lens the sources following the recipe in Sect. 2.2, using various lens mass profiles. We then add the lens light and convolve it with a PSF created with the Tiny Tim software (Krist et al. 2011) for the ACS/WFC and the F814W filter.

The images shown in this section were created from images taken with the ACS/WFC instrument on HST. Flux units are shown in e^- and pixels in the image plane are 0.05 arcseconds in width.

5.2. Plane-wise sparsity of galaxy light profiles

The MCA-SLIT algorithm consists in projecting the mixed image of the lens galaxy and the lensed galaxy back and forth between the source plane and the lens plane, and thresholding the starlet coefficients of each projection. Our hypothesis is that the starlet thresholding favours, in each plane, the corresponding galaxy: source galaxy in source plane and lens galaxy in lens plane. This assumes that a lens galaxy in the lens plane is well reconstructed with only a few starlet coefficients, while a source galaxy projected to the lens plane is not. Conversely, it implies that a source

¹ The frames were recovered from the HFF site at <http://www.stsci.edu/hst/campaigns/frontier-fields/FF-Data>.

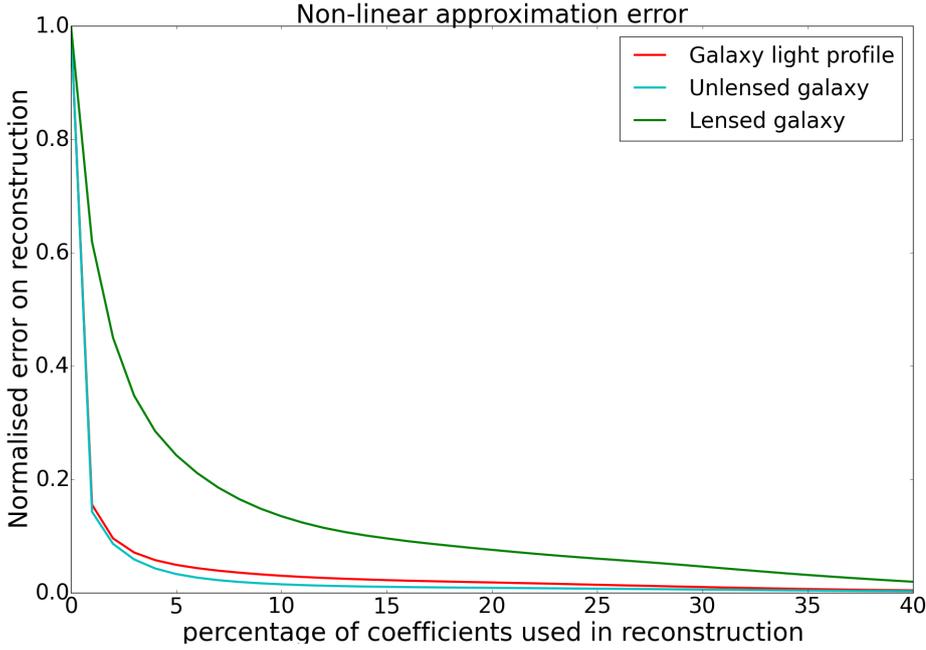


Fig. 2. Normalised non-linear approximation (NLA; Sect. 5.2) of galaxies projected in the source and lens planes. The red curve illustrates the average NLA of galaxy images that can be seen as source or lens galaxies, the cyan curve shows the average NLA of the same galaxies once projected from lens to source plane, and the green curve shows the NLA of the same galaxies projected to lens plane.

galaxy in the source plane is sparser in the starlet domain than a lens galaxy projected to the source plane.

To verify this hypothesis, we selected 167 images of galaxies from cluster MACS J0717 from the HFF survey in the F814W filter of the ACS/WFC camera of HST. We used the HFF-DeepSpace catalogue by Shipley et al. (2018) to select galaxies with a semi-major axis of at least five pixels and a flag at 0 in the F814W filter to insure that the galaxies are isolated in their stamp. We performed the starlet decomposition of these images along with their projections to source and lens plane using three different mass profiles (SIE, SIS and elliptical power law) with realistic draws of the lens parameters. For each of these starlet decompositions we set the $p\%$ smallest coefficients to zero and reconstructed the image in pixel space. We then computed the error on the reconstruction as a function of p . The resulting curve is the non-linear approximation error (NLA; see Starck et al. 2015), shown in Fig. 2. The NLA can be used as a metric for the sparsity of a galaxy profile, in the sense that a sparse galaxy sees its NLA decreasing rapidly with p .

In Fig. 2, we see in particular that the NLA of galaxies (in red) decreases faster than the NLA of the lensed source (in green). This means that keeping only a small percentage of the highest coefficients in starlets (e.g. 10%) of the decomposition of an image Y will reconstruct the lens galaxy well, but the same is not true for the lensed source. When comparing the NLA of a galaxy (red curve in Fig. 2) to that of its projection to the source plane (cyan curve in Fig. 2), we see that the NLAs of both profiles are very similar, making it difficult to disentangle between them. In practice, the reconstructed source images can be contaminated with features belonging to projections of the lens galaxy if not converged properly. One positive aspect, however, is that since lens galaxy light profiles are being very efficiently reconstructed in the lens plane, the signal from a lens galaxy in the source plane decreases very rapidly with iterations.

5.3. Testing SLIT and SLIT_MCA with simulations

In the present work, our goal is to show the potential of MCA-based algorithms as a simultaneous source reconstruction and source-lens deblending technique. All our tests therefore assume

that the mass density profile of the lens is known, as well as the PSF. Unless stated otherwise, the following images were simulated with white Gaussian noise with standard deviation σ . We define the S/N of an image I with N_{pix} pixels as:

$$S/N = \frac{1}{N_{\text{pix}} \sigma^2} \sum_{N_{\text{pix}}} I^2. \quad (19)$$

5.3.1. Case 1: simulation with no lens light

We first reconstruct an image of a lensed galaxy with no foreground light. The simulation contains Gaussian white noise with $S/N = 50$. We used 50 iterations of algorithm 2. The results are presented in Figs. 3 and 4, illustrating the quality of the reconstruction. In Fig. 4 in particular, we show that the central region of the source galaxy, where the flux is larger, is reconstructed with less than 10% error. The error increases in the outer parts of the galaxy where the flux is smaller. In this figure, the relative error is set to zero at locations where the flux in the source is zero, which does not account for false detection.

In this simulation, there are four times more pixels in the source plane image than in the lens plane, that is, in the image plane the observable is an image of 100×100 pixels, while the source is reconstructed using an area of 200×200 pixels.

5.3.2. Case 2: simulation with both lens and source light

We then applied algorithm 3 to the simulated images of a full lens system. In this case we recover both the convolved lens light profile and the source light profile (Fig 5). We enforce the sparsity of each solution by using enough iterations of the algorithm to perform an efficient separation. One difficulty here is in choosing the numbers of iterations and sub-iterations such that both components converge to a sparse solution. In our experiments, $N_{\text{iter}} = N_{\text{subiter}}$ ensures similar quality in the reconstructions of both components.

The results show no structure in the residuals and visually good separation between the lens and the source as well as a good reconstruction of the source without significant leakage between the two. However, the residuals in the first three lines of

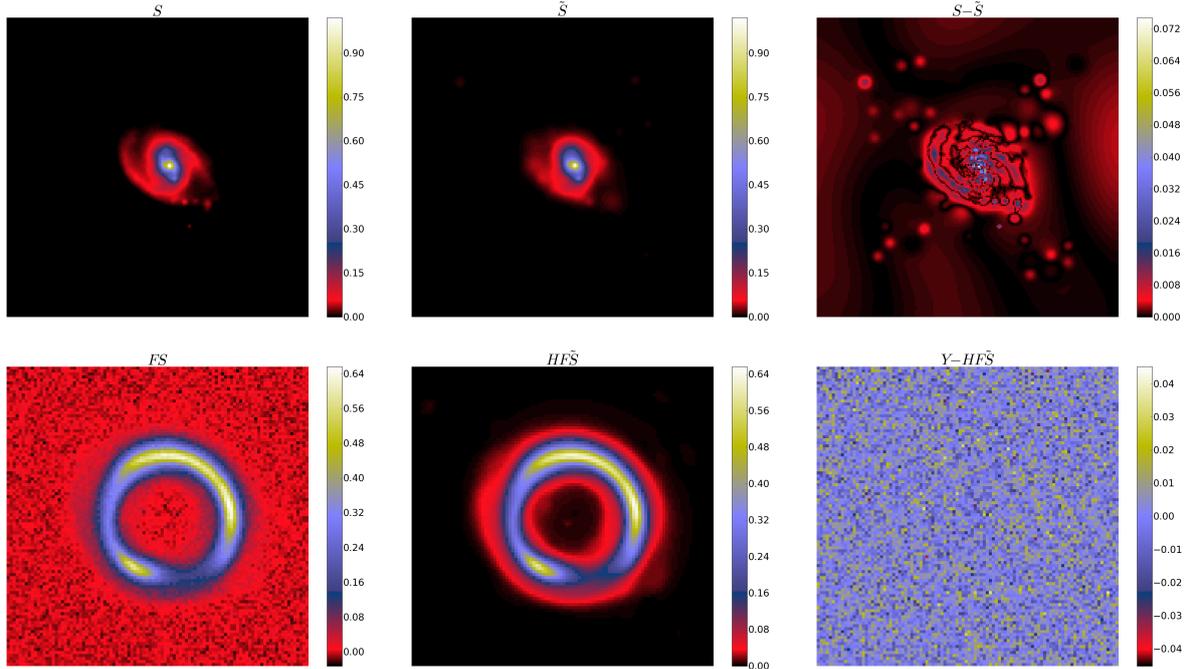


Fig. 3. Application of the SLIT algorithm to a simulated lensing system in the simple case where there is no light from the lensing galaxy. *Left:* simulated source is shown on the top while its lensed and noisy counterpart is shown on the bottom. Both include a PSF convolution. *Middle:* source recovered with the SLIT algorithm and its lensed counterpart. We note that both are still convolved with the PSF. *Right:* difference with the true source (*top*) and the residuals in the lens plane (*bottom*). The original and reconstructed images are displayed with the same colour cuts. The residuals in the *bottom right panel* are shown with $\pm 5\sigma$ cuts.

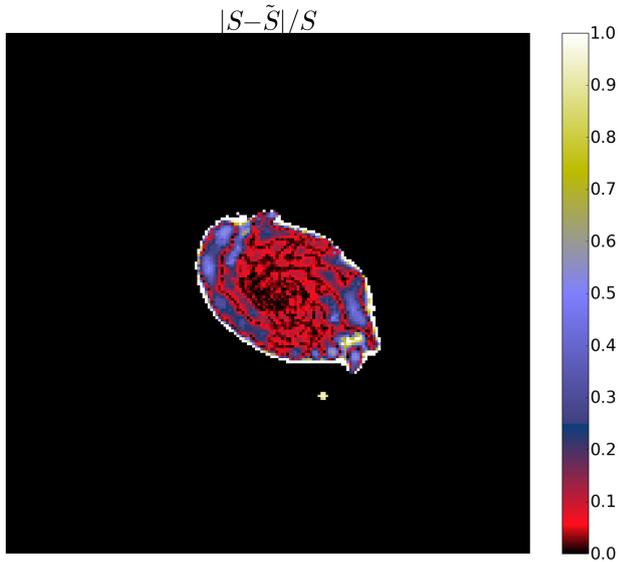


Fig. 4. Residuals of the reconstruction from SLIT relative to the true light profile of the source.

the Figure show that the source flux was slightly overestimated at larger scales, while the lens galaxy was slightly underestimated. The amplitude of the phenomenon reaches no more than 10σ of the noise level given the amplitudes displayed in the last column.

5.4. Comparison with lenstronomy

We tested our source reconstruction technique on three other simulations with various source morphologies and lens mass profiles, including one generated with the `lenstronomy` package. We compared our reconstructions of three lensed sources with the ones computed by S. Birrer, the author of

`lenstronomy`. In order to avoid favouring one method over the other, for reconstructing sources, we tried as hard as we could to choose representations for true sources that do not correspond to the decomposition of either code. Sources generated with SLIT were extracted from HFF images and convolved with a Gaussian kernel with a full width at half maximum (FWHM) of five pixels. This produces a smooth version of the noisy HFF images from which we then subtract the median value of the image in order to set the sky background to zero. All remaining negative values in the image are set to zero. The image generated with `lenstronomy` uses a source from a jpeg image of NGC 1300 from NASA, ESA. The image resolution is degraded by a factor 25 and decomposed over the shapelet dictionary (Refregier 2003) using enough coefficients (11 476) to accurately recover the morphology of the image. Despite `lenstronomy` relying on shapelets to solve the source inversion problem, the number of coefficients that it is possible to recover in the reconstruction is much smaller than the number of coefficients used in generating the true source. Therefore, the basis set of the reconstruction is different from the one used in generating the true source. The three systems tested here were made from sources with different morphologies and different lens profiles; Table 1. In this exercise we test our methods on simulated images that were generated with different procedures. This comparison therefore allows us to show how robust these techniques are to the underlying mapping between source and lens plane.

In order to compare the results of both methods we show the resulting reconstructions of the runs in Figs. 6 and 7 and we use three metrics:

- Quality of the residuals (QoR), given by the relative standard deviation of the residuals for a model of the source, \tilde{S} :

$$\text{QoR}(\tilde{S}) = \text{std}\left(\frac{Y - H\tilde{S}}{\sigma}\right). \quad (20)$$

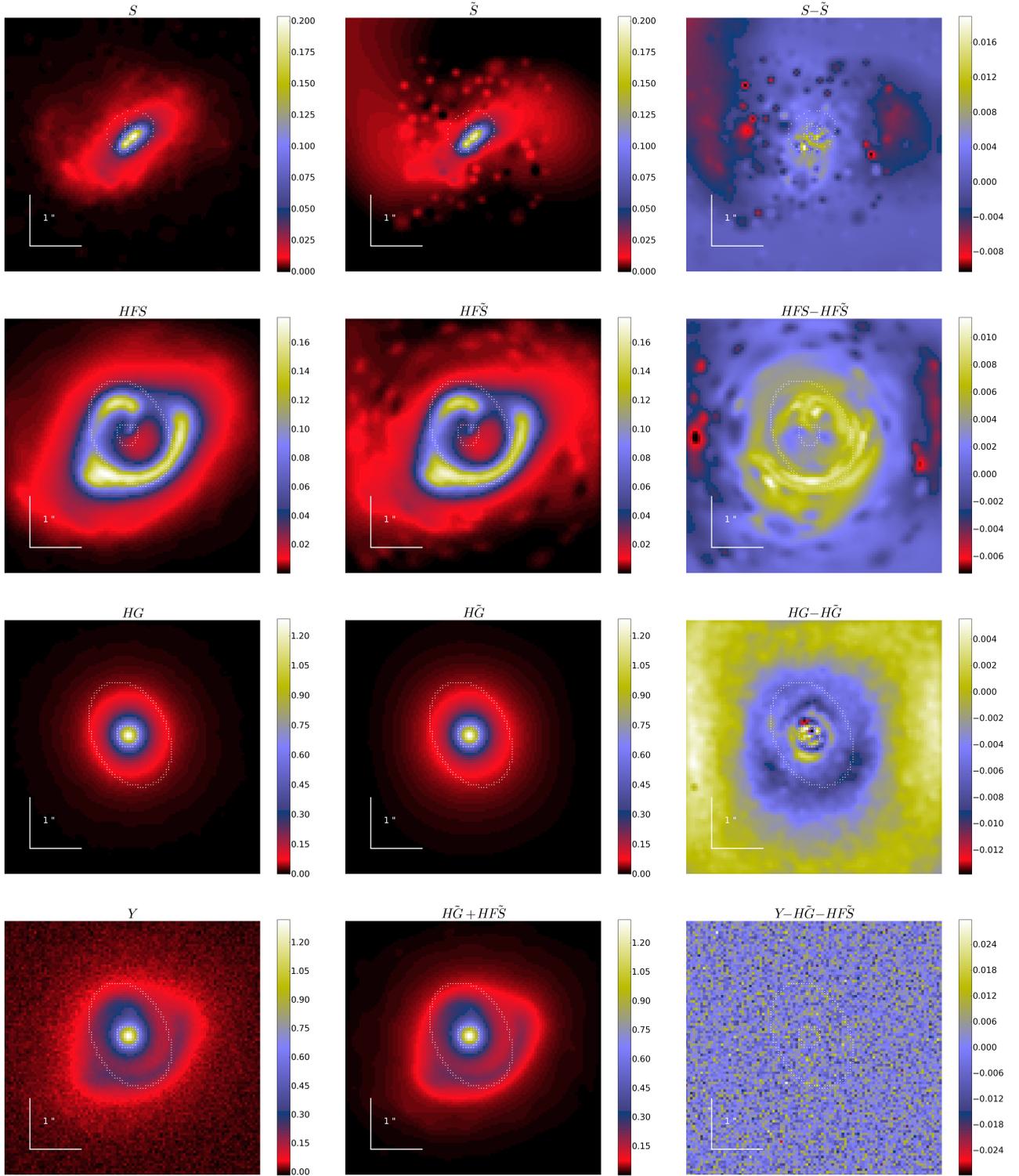


Fig. 5. Illustration of the SLIT_MCA algorithm with simulated data. *Left:* simulated ground truths. *From top to bottom* are shown the original un-lensed source, its lensed version convolved with the PSF, the lensing galaxy (convolved with the PSF), and the full simulated system with noise. *Middle:* output of the SLIT_MCA algorithm. *Right:* differences between the left and middle panels. The original and reconstructed images are displayed with the same colour cuts. The residuals in the *bottom right panel* are shown with cuts set to $\pm 5\sigma$. White dots show the positions of pixels crossed by critical lines in the lens plane and by caustics in the source plane.

In cases of Gaussian and Poisson mixture noise, $\sigma = \sqrt{\sigma_G^2 + f}$, where σ_G is the standard deviation of the Gaussian component and f is a 2D map of the flux in the true noiseless model for image Y . Given that definition,

- the closer the value of QoR is to one, the better the reconstruction.
- Quality of the source reconstruction relative to the true source, S_{true} , estimated with the source distortion ratio (SDR, Vincent et al. 2006). The SDR is the logarithm of the inverse

Table 1. Description of the simulated images.

Image number	Source origin	Source processing	Factor	PSF	Noise	Lens model
1	NASA, ESA jpeg	Shapelets	5	Gaussian	P+G	SPEP
2	HFF fields	$\star G_{-}$	4	Gaussian	SNR = 10	SIE
3	HFF fields	$\star G_{-}$	2	Gaussian	SNR = 100	SIS

Notes. Symbol $\star G_{-}$ stands for convolution by a Gaussian kernel and subtraction of the median (see text). Column factor stands for the resolution factor between source and lens plane. For instance, in image 1, the source has five times more pixels on the side than the image. Gaussian PSFs were used in all three images with a FWHM of 2 pixels, P+G stands for Gaussian poisson mixture. Gaussian noise with $\sigma = 2$ was used. The poisson noise value at pixel i is drawn from a Gaussian distribution with $\sigma = \sqrt{f_i}$, f_i being the flux in pixel i .

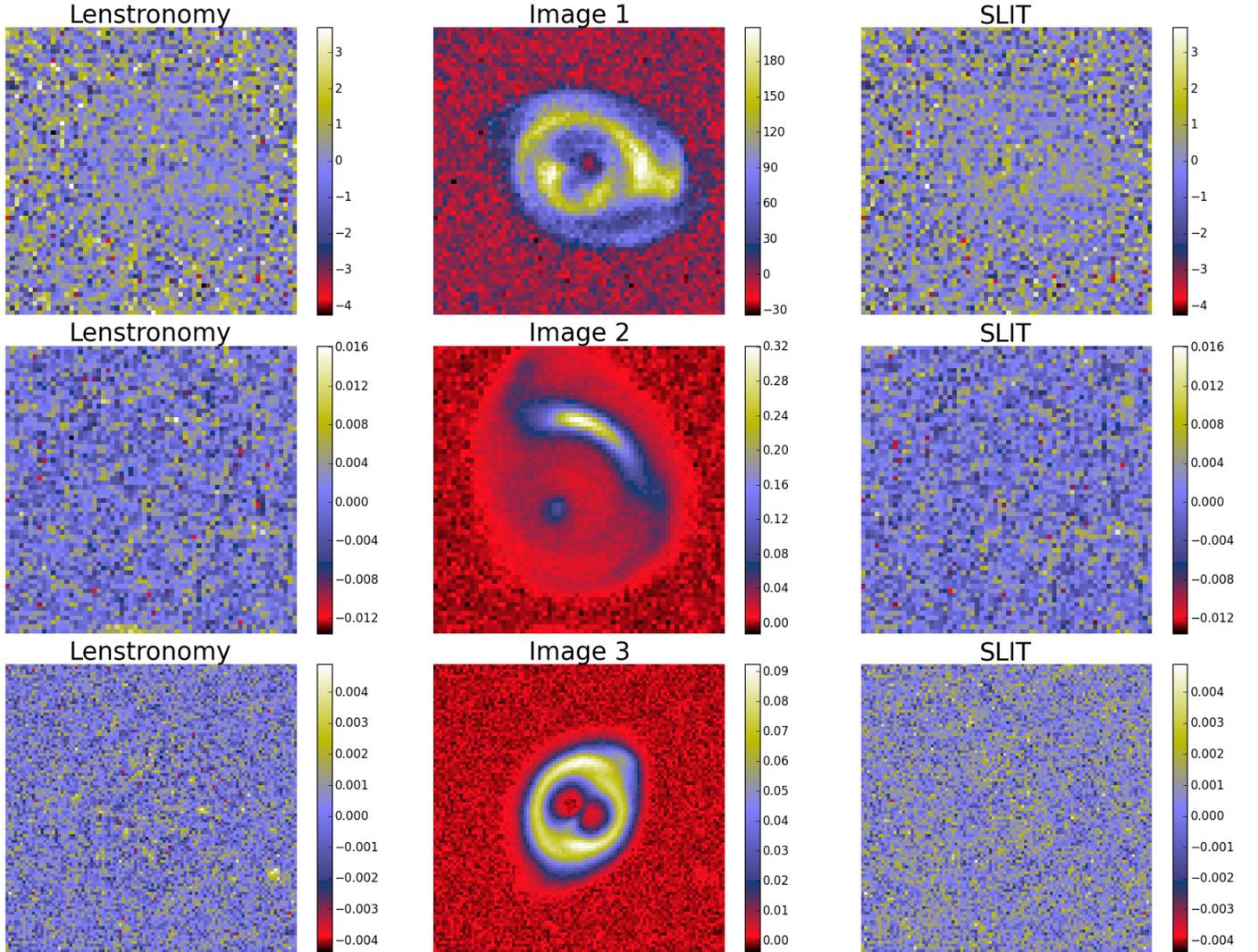


Fig. 6. Reconstructions with Lenstronomy and SLIT in image plane. *Middle panels from top to bottom: simulated images 1, 2, and 3. Left panels: corresponding residuals after reconstruction with lenstronomy, right panels: residuals obtained with SLIT.*

of the error on the source light profile, weighted by true flux of the source. As a result, the higher the SDR, the better the reconstruction of the source. We compute the SDR as

$$\text{SDR}(\tilde{S}) = 10 \log_{10} \frac{\|S_{\text{true}}\|}{\|S_{\text{true}} - \tilde{S}\|}. \quad (21)$$

– Computation time.

The two metrics of quality and source residuals were chosen to provide a measure of the quality of the reconstruction in

both source and lens planes. While the SDR of the sources is the most informative metric with regards to the quality of the reconstruction of the source, it is also necessary to ensure that both methods are able to reconstruct the observables similarly well, hence the role of metric QoR. The evaluation of these metrics for both methods are given in Table 2.

The residuals in Fig. 6, as well as the results for QoR(\tilde{S}) in Table 2, show that both codes achieve similar quality of reconstruction. While lenstronomy leaves a little bit more signal in the residuals, in particular in cases of smooth sources generated

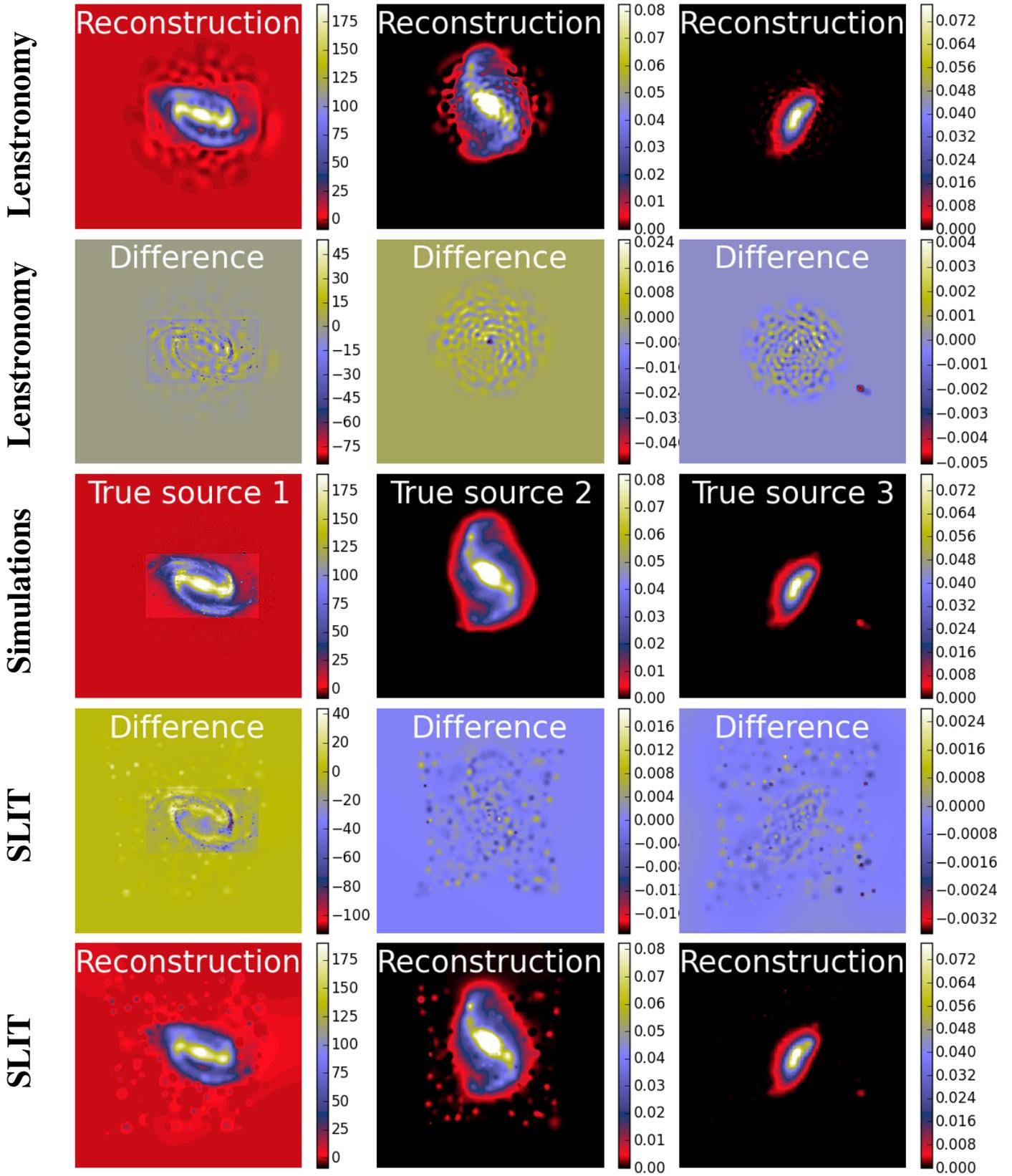


Fig. 7. Reconstructions with `Lenstronomy` and `SLIT` in the source plane. Panels from the *middle row* show the true sources used to generate simulated images 1, 2 and 3, respectively. *First row*: source reconstruction from `lenstronomy`. *Second row*: difference between the true sources and the sources reconstructed by `lenstronomy`. *Last row*: source reconstruction from `SLIT`. *Penultimate row*: difference between true sources and sources reconstructed with `SLIT`. Panels between reconstructed and true images show the difference between the two for the corresponding technique.

Table 2. Comparison of reconstructions of three lensed sources by SLIT and lenstronomy.

	SDR		QoR		T (s)	
	Lenstronomy	SLIT	Lenstronomy	SLIT	Lenstronomy	SLIT
Image 1	8.31	6.39	1.10	1.05	~0.1	~100
Image 2	9.30	13.06	1.26	1.21	~0.1	~400
Image 3	13.69	16.09	1.24	1.20	~0.1	~2000

with SLIT, on the contrary SLIT tends to create false detections at noise level at locations where the actual signal is zero, resulting in a slight over-fitting. In the case of `lenstronomy`, over-fitting of outer regions with no signal is prevented by the fact that the method relies on shapelets, which are localised around the centre of the images provided that the number of coefficients used in the reconstruction is kept small. The downside of that strategy is that shapelets do not accurately represent companion galaxies in the source such as the one on the right side of image 3 in Fig. 7. To circumvent this problem, one needs to use a second set of shapelets positioned at a different location to represent a second displaced light component. This has deliberately not been done in this comparison. From looking at the reconstructions of image 1 and the corresponding SDRs, it appears that `lenstronomy` performs slightly better at reconstructing images generated with high resolution and detailed features. Despite both reconstructions showing the same level of detail, false detections in SLIT at locations where the actual signal is zero contribute to diminishing its SDR. Also, the large number of pixels in the source compared to the number of pixels in the observations (25 times more pixels in the source than in the image) makes the problem highly under-constrained. In `lenstronomy`, this is overcome by computing a small number of shapelet coefficients and displaying them on an arbitrarily fine grid. With SLIT, we optimise for each coefficient in the starlet dictionary, which means more pixels in the source plane equals more unknowns as the number of pixels increases.

Regarding computational time, while `lenstronomy` runs in less than a second, a typical SLIT run such as those provided above will last between ~100 and ~1000 s. For a full run of `SLIT_MCA`, this number is multiplied by a factor of between five and ten. While this is a current weakness of our algorithm, we are confident that optimised packages for starlet transforms and linear optimisation as well as parallel computing will allow us to lower these numbers by at least a factor of ten.

5.5. Lens-parameter optimisation

In order to assess whether our technique holds the potential to be used in a full lens-modelling context, we test its sensitivity to the density slope γ . In [Koopmans et al. \(2006, 2009\)](#), the authors showed that real lens galaxies have, on average, a total mass density profile with a density slope $\gamma \sim 2$, that is, isothermal. In Fig. 1 of [Koopmans et al. \(2009\)](#), the authors show that the posterior probability distributions of a sample of 58 strong lenses are maximised for a density slope between 1.5 and 2.5, which sets the limits for the values of γ investigated here.

In our analysis, we generate a lens system (shown in Fig. 9) with a power-law mass density profile with $\gamma = 2$. The light profiles for the lens and source galaxies were drawn from HFF images. The PSF is a Gaussian profile with a FWHM of 2 pixels. We create 100 realisations of this system with additive Gaussian white noise at a S/N of 100. Each realisation is then reconstructed with `SLIT_MCA`, using mass models with density

slopes, $\tilde{\gamma}$, ranging from 1.5 to 2.5. The results are shown in Fig. 8. The true and the reconstructed profiles are shown in Appendix A. This appendix serves to give a visual impression of how sparsity allows to discriminate between lens mass models by discarding unphysical solutions that have a high ℓ_1 norm in Starlets.

Figure 8 shows that the actual morphologies of S and G_H are recovered very accurately (SDR ~ 20) for $\gamma = 2$. In a real case study, the truth for S and G_H light distributions is not known, and therefore it is impossible to use the SDR to discriminate between lens-mass model parameters. Instead, we have to rely on quantities derived from the observations or on properties of the reconstructed profiles. The top panel of Fig. 8 shows the likelihood defined as:

$$\mathcal{L}(\gamma) = \exp(-\|Y - HF_{\kappa(\tilde{\gamma})}\Phi\alpha_S - \Phi\alpha_{G_H}\|_2^2), \quad (22)$$

of a lens model with slope $\tilde{\gamma}$ to be the right model. Because of the strategy we choose for `SLIT_MCA`, which consists in optimising alternatively over the source and lens light profiles, the algorithm is very likely to estimate light profile models that minimise the residuals in the image, even in cases where the mass model does not correspond to the truth. Hence the relative flatness, compared to the error bars, of the likelihood profile from Fig. 8. In particular, we observe that, for $\tilde{\gamma} > 2.2$, we achieve a likelihood as high or higher than the likelihood at $\gamma = 2$, despite using the wrong lens model. This is caused by the extreme steepness of the mass density profile, which causes `SLIT_MCA` to model the source as its lensed version back in source plane. With these results, one would think `SLIT_MCA` unfit to be used in a full lens modelling framework. The strength of the algorithm lies in its potential to find the sparsest solution to a problem of lens light modelling. Since incorrect mass models introduce artefacts in the reconstructions of light models, their ℓ_1 -norm is significantly higher than in the case of a reconstruction with a true model where light profiles are smooth. The middle panel of Fig. 8 shows the cumulative ℓ_1 -norms of α_S and α_{G_H} as an argument of the likelihood. In this case, the metric is maximised around the true value for $\tilde{\gamma}$. Despite the sparsest solution being found for $\tilde{\gamma} = 2.01$, while the true value for γ sits at 2.00, this result is still in the error bars estimated in [Koopmans et al. \(2009\)](#) over a sample of 58 lenses.

6. Reproducible research

In order to allow for reproducibility of these results and also to provide a tool that can be used by the community, we provide the `python` code that was used to implement the method we describe here. We also provide the routines used to create all plots shown in this paper as well as a `readme` file that should allow users to run the method easily on their own data. The repository is available via the `github` platform². The only required libraries to run the code are the standard `numpy`, `matplotlib`, `scipy`, and `pyfits` libraries. All functions and routines described in this paper

² <https://github.com/herjy/SLIT>

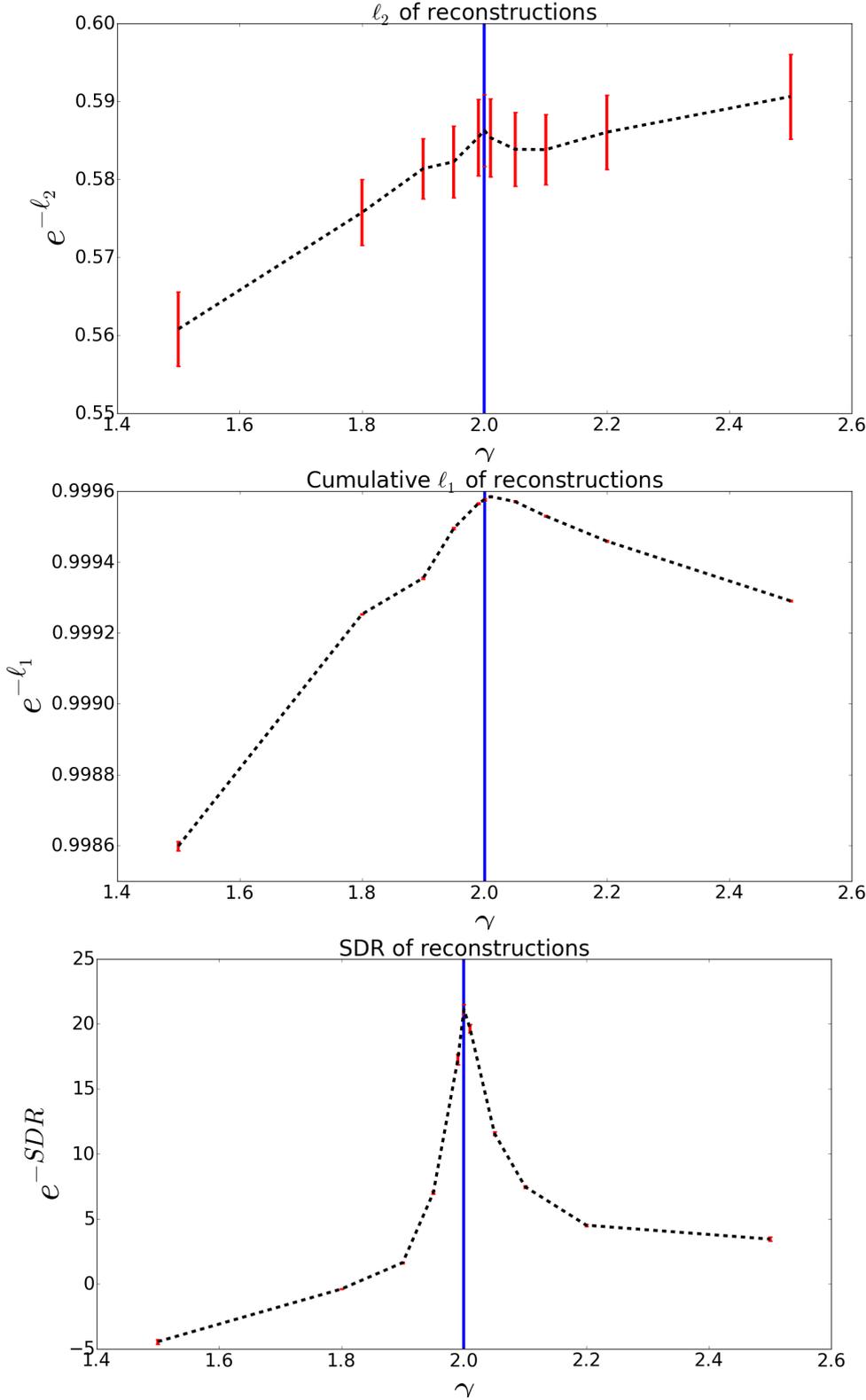


Fig. 8. Metrics of the reconstructions of the system in Fig. 9 as a function of mass density slope. *Top-left panel:* cumulative SDR of the source and galaxy light profile reconstruction. *Top-right panel:* average of the residuals as $\exp(-\|Y - HF_{\kappa(\bar{\gamma})}\Phi\alpha_S - \Phi\alpha_{G_H}\|_2^2)$ over 100 noise realisations. *Bottom panels:* cumulative ℓ_1 -norm of α_S and α_{G_H} as $\exp(-(\lambda_S\|\alpha_S\|_1 + \lambda_G\|\alpha_{G_H}\|_1))$. The error bars show the standard deviation of these metrics over 100 noise realisations. The blue line shows the true value of γ .

were implemented by the authors. Table 3 lists the products we make available in our public repository.

7. Conclusion

We have developed a fully linear framework to separate the lens and source light profiles in strong lensing systems. We also are able to reconstruct the source shape as it was prior to the lensing

effect at fixed lens mass model. As the problem is linear we were able to apply sparsity-based optimisation techniques to solve it, leading to two algorithms.

The first algorithm, SLIT, decomposes the source plane on the basis of un-decimated wavelets (starlets). This allows us to represent the source in a non-analytical way, hence providing a sufficiently large number of degrees of freedom to capture small structures in the data. Using a sparse regularisation allows us

Table 3. List of products made available in this paper in the spirit of reproducible research.

Product name	Type	Description
Software products:		
SLIT	python package	Includes SLIT implementation and visualisation tools.
Lens.py	python code	Toolbox for lensing.
Solve.py	python code	Python implementation of algorithm 2 and algorithm 3 and related tools.
tools.py	python code	Code for various functions used in the minimisation process. Such as the starlet transform or the FISTA iteration. Used to compute Fig. 1.
Routines:		
Test_SLIT.py	code (python)	Routines to reproduce Fig. 3.
Test_SLIT_MCA.py	code (python)	Routines to reproduce Fig. 5.
Quality.py	code (python)	Routines to reproduce Figs. 6, 7 and Table 2.
Test_sparsity.py	code (python)	Routines to reproduce Fig. 2.
Results_slope_MCA.py	code (python)	Routines to reproduce Figs. 9, 8 and A.1.

Notes. All above material is available at the following url: <https://github.com/herjy/SLIT>.

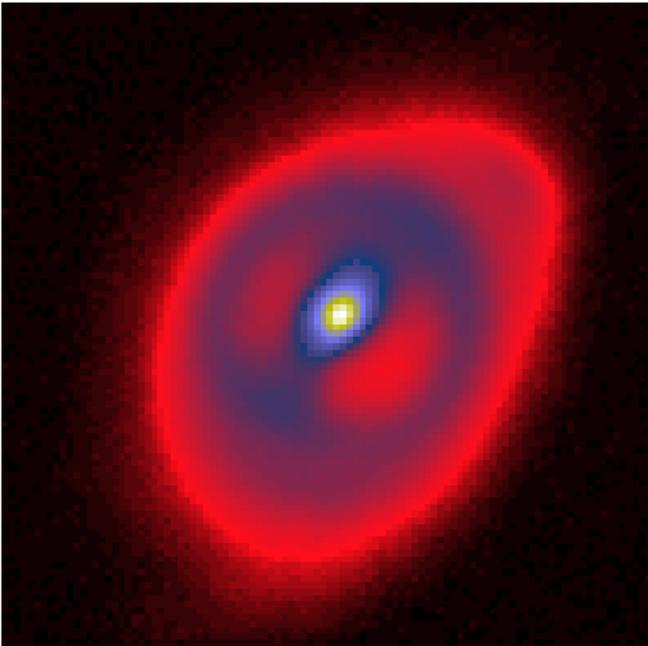


Fig. 9. Light profile of a simulated lens system (lens and lensed source light profiles) generated with a power-law mass profile with $\gamma = 2$.

to reduce the effect of noise and artefacts on the reconstruction. SLIT applies to lensed systems that have a fixed mass profile and where the lens light has been removed.

The second algorithm, SLIT_MCA, also applies in the case of a fixed mass model but is able to deblend the lens light from the source as it reconstructs it in the source plane and deconvolves it from the instrumental PSF. The separation of lens and source light profiles in SLIT_MCA relies on the principle of morphological component analysis, but uses the distortion introduced by lensing itself as a way of discriminating between lens and source features. As is the case for SLIT, SLIT_MCA does not use any analytical form for either the source or the lens representation. Both algorithms account for the instrumental PSF.

We tested our algorithms with simulated images, showing that sources can be reconstructed while separating lens and source light profiles. We identify several advantages of our approach:

- The lens and source light profiles are pixelated numerical profiles, allowing a large number of degrees of freedom in their reconstruction.
- The code implementing the algorithms is fully automated due to a careful automated computation of the regularisation parameter. It does not require any prior or assumption about the light profiles of the source or the lens.
- The performances of the algorithms are robust against pixel size in the sense that arbitrarily small pixel sizes can be used without leading to noise amplification or artefacts. The pixel size of both the source and lens can be an order of magnitude smaller than the PSF without negative impact on the results. On the contrary, adopting very small pixel sizes allows for detailed reconstruction of the source.
- The python code is public.

Furthermore, the linear framework of the algorithms presented here allows for the development of a lens modelling technique based on non-negative matrix factorisation (NMF, Lee & Seung 1999; Pentti & Unto 1994) for the estimate of F_κ . This technique could allow us to solve Eq. (7) in F_κ , G , and S simultaneously. Knowledge of F_κ would in turn allow us to recover free-form solutions for the density mass profile κ . We showed in Fig. 8 how our strategy favoured the reconstruction of the sparsest solution, hinting at the possibility to be able to find a model for F_κ that favours the sparsest set of lens and source profiles. Because this problem introduces another degree of complexity due to the number of unknowns and the increasing degeneracies of the solutions, we delay this study to a later publication.

The main limitation of the algorithm at this time is its very high computational cost, which scales to dozens of minutes if not an hour when estimating the light profiles of very large sources with SLIT_MCA. Although this renders the method inefficient as a possible minimiser in a Monte-Carlo Markov Chain sampling strategy, we are confident that upcoming optimised packages for linear optimisation will increase the speed of the algorithm. Also, the motivation behind the development of this technique is to be able to estimate free-form lens-mass models from NMF for instance, which requires fewer evaluations of SLIT_MCA.

To our knowledge, this is the first time that an inverse problem has been used to perform component separation inside a blind source reconstruction problem. In this sense, the present work may extend beyond astrophysics and address a more general class of deblending and inversion problems.

Acknowledgements. We would like to thank Danuta Paraficz, Markus Rexroth, Christoph Schäfer, Sherry Suyu and Johan Richard for useful discussions on gravitational lensing and lens/source projection as well as François Lanusse for his help on sparse regularisation and related algorithms. This research is supported by the Swiss National Science Foundation (SNSF) and the European Community through the grant DEDALE (contract no. 665044). RJ also acknowledges support from the Swiss Society for Astronomy and Astrophysics (SSAA). S. Birrer acknowledges support by NASA through grants STSCI-GO-14630, STSCI-GO-15320, and STSCI-GO-15254 and by NSF through grant 1714953.

References

- Alard, C. 2009, *A&A*, **506**, 609
- Beck, A., & Teboulle, M. 2009, *SIAM J. Imaging Sci.*, **2**, 183
- Bellagamba, F., Tessore, N., & Metcalf, R. B. 2017, *MNRAS*, **464**, 4823
- Birrer, S., & Amara, A. 2018, *Phys. Dark Universe*, **22**, 189
- Birrer, S., Amara, A., & Refregier, A. 2015, *ApJ*, **813**, 102
- Bolton, A. S., Burles, S., Koopmans, L. V. E., et al. 2008, *ApJ*, **682**, 964
- Bolton, A. S., Brownstein, J. R., Kochanek, C. S., et al. 2012, *ApJ*, **757**, 82
- Bonvin, V., Courbin, F., Suyu, S. H., et al. 2017, *MNRAS*, **465**, 4914
- Braut, F., & Gavazzi, R. 2015, *A&A*, **577**, A85
- Candes, E. J., Wakin, M. B., & Boyd, S. P. 2008, *J. Fourier Anal. Appl.*, **14**, 877
- Cava, A., Schaefer, D., Richard, J., et al. 2018, *Nat. Astron.*, **2**, 76
- Chambolle, A., & Dossal, C. 2015, *J. Optim. Theory Appl.*, **166**, 968
- Coles, J. P., Read, J. I., & Saha, P. 2014, *MNRAS*, **445**, 2181
- Courbin, F., Faure, C., Djorgovski, S. G., et al. 2012, *A&A*, **540**, A36
- Diego, J. M., Broadhurst, T., Wong, J., et al. 2016, *MNRAS*, **459**, 3447
- Dye, S., & Warren, S. J. 2005, *ApJ*, **623**, 31
- Farrens, S., Starck, J. L., & Ngolè Mboula, F. M. 2017, *A&A*, **601**, A66
- Gabay, D. 1983, *Stud. Math. Appl.*, **15**, 299
- Garsden, H., Girard, J. N., Starck, J. L., et al. 2015, *A&A*, **575**, A90
- Harvey, D., Massey, R., Kitching, T., Taylor, A., & Tittley, E. 2015, *Science*, **347**, 1462
- Harvey, D., Kneib, J. P., & Jauzac, M. 2016, *MNRAS*, **458**, 660
- Holschneider, M., Kronland-Martinet, R., Morlet, J., & Tchamitchian, P. 1989, in *Wavelets: Time-Frequency Methods and Phase Space*, eds. J.-M. Combes, A. Grossmann, & P. Tchamitchian, 286
- Joseph, R., Courbin, F., Metcalf, R. B., et al. 2014, *A&A*, **566**, A63
- Joseph, R., Courbin, F., & Starck, J.-L. 2016, *A&A*, **589**, A2
- Kneib, J.-P., Bonnet, H., Golse, G., et al. 2011, *LENSTOOL: A Gravitational Lensing Software for Modeling Mass Distribution of Galaxies and Clusters (strong and weak regime)*, Astrophysics Source Code Library [record ascl:1102.004]
- Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, *ApJ*, **649**, 599
- Koopmans, L. V. E., Bolton, A., Treu, T., et al. 2009, *ApJ*, **703**, L51
- Krist, J. E., Hook, R. N., & Stoehr, F. 2011, in *Optical Modeling and Performance Predictions V*, Proc. SPIE, 8127, 81270J
- Lanusse, F., Starck, J.-L., Leonard, A., & Pires, S. 2016, *A&A*, **591**, A2
- Lee, D. D., & Seung, H. S. 1999, *Nature*, **401**, 788 EP
- Livemore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, *ApJ*, **835**, 113
- Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, *ApJ*, **837**, 97
- More, A., Cabanac, R., More, S., et al. 2012, *ApJ*, **749**, 38
- Ngolè Mboula, F. M., Starck, J.-L., Ronayette, S., Okumura, K., & Amiaux, J. 2015, *A&A*, **575**, A86
- Nightingale, J. W., & Dye, S. 2015, *MNRAS*, **452**, 2940
- Nightingale, J. W., Dye, S., & Massey, R. J. 2018, *MNRAS*, **478**, 4738
- Oguri, M. 2010, *PASJ*, **62**, 1017
- Pentti, P., & Unto, T. 1994, *Environmetrics*, **5**, 111
- Pratley, L., McEwen, J. D., d’Avezac, M., et al. 2018, *MNRAS*, **473**, 1038
- Priewe, J., Williams, L. L. R., Liesenborgs, J., Coe, D., & Rodney, S. A. 2017, *MNRAS*, **465**, 1030
- Refregier, A. 2003, *MNRAS*, **338**, 35
- Refsdal, S. 1964, *MNRAS*, **128**, 307
- Schneider, P., & Sluse, D. 2013, *A&A*, **559**, A37
- Schneider, P., & Sluse, D. 2014, *A&A*, **564**, A103
- Sebesta, K., Williams, L. L. R., Mohammed, I., Saha, P., & Liesenborgs, J. 2016, *MNRAS*, **461**, 2126
- Shensa, M. J. 1992, *IEEE Trans. Signal Process.*, **40**, 2464
- Shipley, H. V., Lange-Vagle, D., Marchesini, D., et al. 2018, *ApJS*, **235**, 14
- Starck, J. L., Elad, M., & Donoho, D. L. 2005, *IEEE Trans. Image Process.*, **14**, 1570
- Starck, J.-L., Fadili, J., & Murtagh, F. 2007, *IEEE Trans. Image Process.*, **16**, 297
- Starck, J., Murtagh, F., & Fadili, J. 2015, *Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis* (Cambridge: Cambridge University Press)
- Suyu, S. H., Marshall, P. J., Hobson, M. P., & Blandford, R. D. 2006, *MNRAS*, **371**, 983
- Suyu, S. H., Marshall, P. J., Auger, M. W., et al. 2010, *ApJ*, **711**, 201
- Suyu, S. H., Treu, T., Hilbert, S., et al. 2014, *ApJ*, **788**, L35
- Tessore, N., Bellagamba, F., & Metcalf, R. B. 2016, *MNRAS*, **463**, 3115
- Tewes, M., Courbin, F., Meylan, G., et al. 2013, *A&A*, **556**, A22
- Treu, T., & Marshall, P. J. 2016, *ARA&A*, **24**, 11
- Treu, T., Dutton, A. A., Auger, M. W., et al. 2011, *MNRAS*, **417**, 1601
- Vegetti, S., & Koopmans, L. V. E. 2009, *MNRAS*, **392**, 945
- Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., & Gavazzi, R. 2010, *MNRAS*, **408**, 1969
- Vincent, E., Gribonval, R., & Fevotte, C. 2006, *IEEE Trans. Audio Lang. Process.*, **14**, 1462
- Warren, S. J., & Dye, S. 2003, *ApJ*, **590**, 673
- Wayth, R. B., & Webster, R. L. 2006, *MNRAS*, **372**, 1187

Appendix A: Reconstructions for lens mass model optimisation

In this appendix, we show the average reconstructed lens and source light profiles as well as the corresponding truth profiles for reconstructions of a lens system with different slopes for a power law mass density profile. The profiles are shown in Fig. A.1. The light profile of lens and source galaxies reconstructed with incorrect mass models show features and modes

that are poorly represented by starlets and that visually do not correspond to the shape of the galaxy light profile. This is particularly visible in the most extreme cases: for $\gamma = 2.5$, the steepness of the slope leads to a mass model localised at the centre of the image, such that the reconstructed source appears as a lensed galaxy; for $\gamma = 1.5$ and $\gamma = 1.8$, the light profile of the source is scattered across the source plane and the lens light profile contains most of the signal from both source and galaxy light profiles.

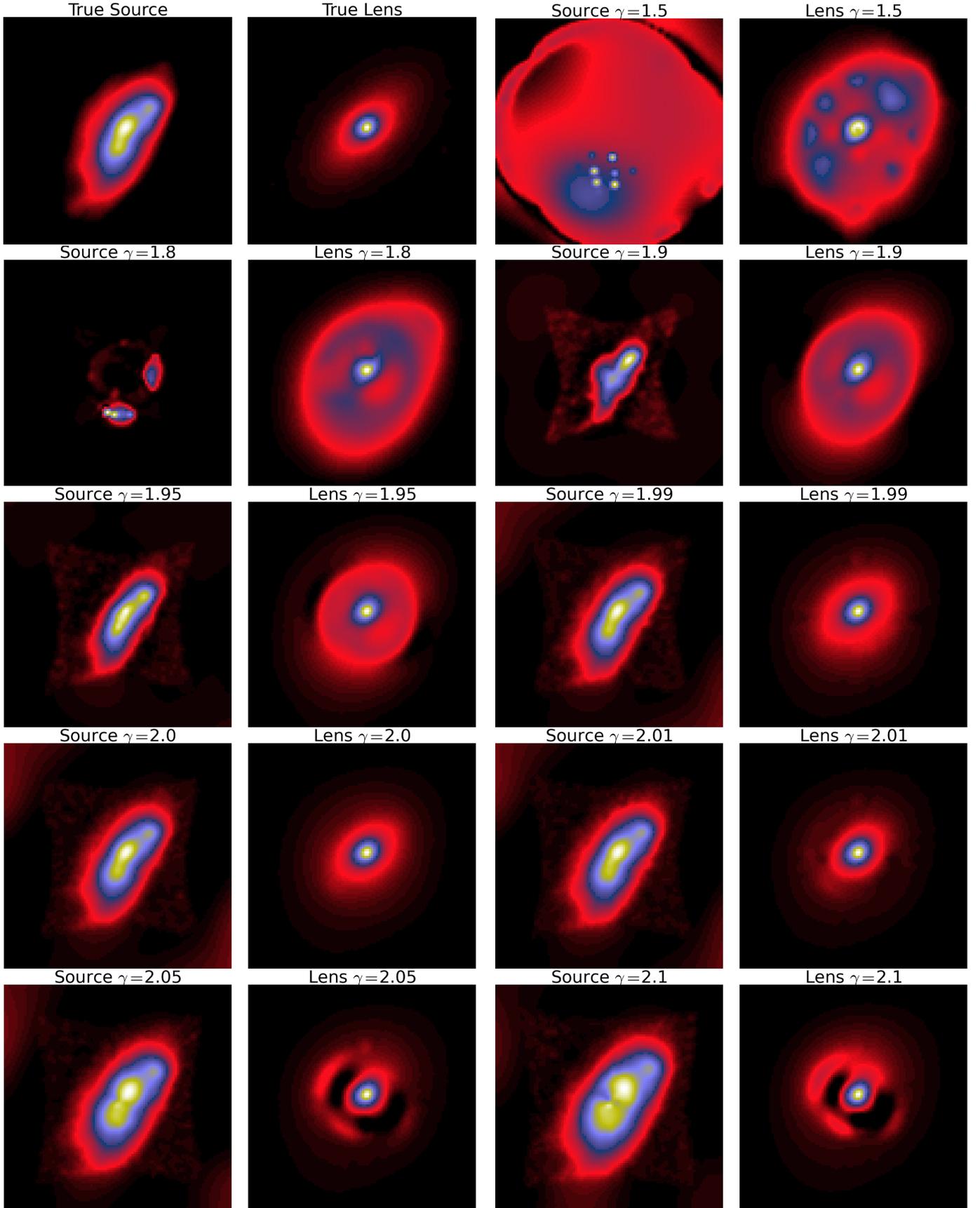


Fig. A.1. Reconstructions of lens and source light profiles for various values of mass density slope of a lens generated with a mass density slope of 2. The *first two panels* show the true source (*left-hand side*) and lens (*right-hand side*) light profiles used to generate the simulated images. The other pairs of panels *from left to right and from top to bottom* show the source and lens reconstructions for increasing values of $\tilde{\gamma}$.

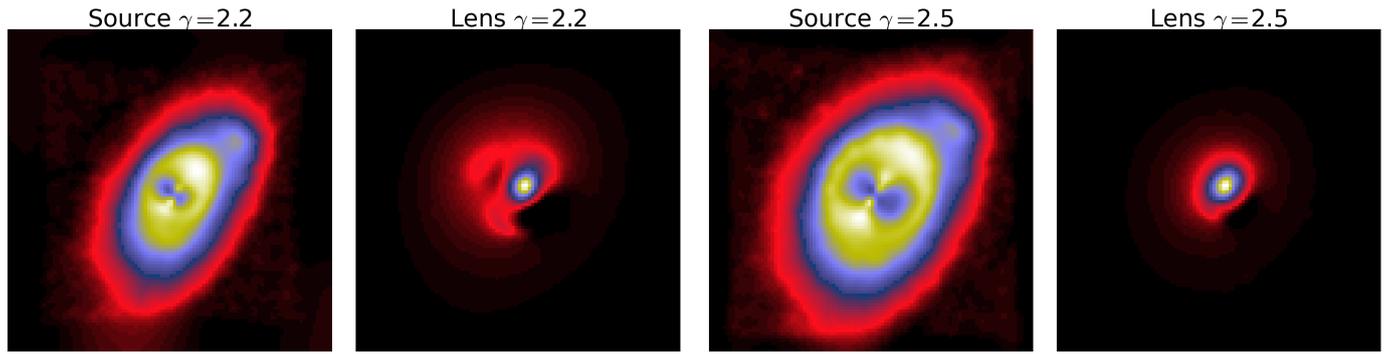


Fig. A.1. continued.