SF2A 2013
L. Cambrésy, F. Martins, E. Nuss and A. Palacios (eds)

SEARCH AND ANALYSIS OF GALAXY-SCALE STRONG GRAVITATIONAL LENSES IN COSMOLOGICAL SURVEYS

F. Brault[1]

Abstract. This article focuses on the development of a novel detector of strong galaxy-galaxy lenses based on the massive modelling of candidates in wide-field ground-based imaging data. Indeed, not only are these events rare in the Universe, but they are at the same time very valuable to understand galaxy formation and evolution in a cosmological context. We use parametric models, which are optimized by MCMC in a bayesian framework, so that we know the distribution of errors. We first generate several training samples : a hundred lenses simulated in HST and CFHT conditions, along with 325 observed lens candidates resulting from a series of preselections on the CFHTLS-Wide galaxies, and that we classify according to their credibility. The whole challenge in designing this detector lies in a subtle balance between the quality of models and the execution time. We massively run the modelling on our samples, beginning with ideal application conditions that we make more complex by stages so as to get closer to the observation conditions and save time. We show that a 7-parameter model assuming a spherical source can recover the Einstein radius from the CFHT simulations with a precision of 7%. We apply a mask to the input data that noticeably enhances the robustness of the models facing environment problems, with a median convergence time of 4 minutes that could be easily reduced by a factor of 10 with more direct optimization techniques. From our results, we define selection contours in the parameter space, resulting in a completeness of 38% and a purity of 55% for the sample of 51 candidates accepted by our robot among the 325 preselected systems.

Keywords: gravitational lenses - galaxies : elliptical - dark matter - observational cosmology - large surveys - methods : statistics.

1 Introduction

Large surveys enable us to tackle the question of galaxy evolution in a cosmological context. In the current state of research, the merger scenario is still unclear but tends to show a dichotomy : cuspy inner profils, as seen in simulations, would result from dissipational mergers (wet mergers), whereas observed cores would result from dissipationless merger (dry mergers). The actual scenario is most probably a mix of those two kinds of mergers, but we need more probes to reveal the density profile of galaxies we observed, and more especially their dark matter (DM) density profile.
Strong lensing modelling gives an accurate mesaure of the total mass (baryons and DM) projected in a characteristic radius called the Einstein radius R_{Ein}. Under the assumption that the lens potential is close to a SIS (Singular Isothermal Sphere), this relation is given by the following formula :

$$R_{Ein} = 4\pi(\frac{\sigma_c}{c})^2\frac{D_{ds}}{D_s} \tag{1.1}$$

where σ_c is the central velocity dispersion of the lens galaxy, D_{ds} the distance between the deflector and the source, and D_s the distance to the source. Nevertheless, there are some drawbacks in contraining the density profile of a galaxy via the strong lensing tool : the mass-sheet degeneracy specific to lensing, the need to know the lens and the source redshifts, along with the fact that strong lenses at galaxy scale are very rare objects.

[1] Institut d'Astrophysique de Paris

Stellar dynamics break the mass-sheet degeneracy by delivering the total mass of the lens galaxy in another radius close to its effective radius. This mass results from the inversion of the Jeans equation. Knowing the mass from the lens modelling and the mass from the stellar dynamics, we get the mass gradient in the inner profile of the lens galaxy.
Ruff et al. 2011 Ruff et al. (2011) and Sonnenfeld and al. 2013 b) Sonnenfeld et al. (2013a) have carried out a joint strong lensing and stellar dynamics analysis that works on the following principles : strong lens modelling gives the Einstein radius, the position angle and the ellipticity of the deflector (see Gavazzi et al. 2012 Gavazzi et al. (2012) and Sonnenfeld et al. 2013 a) Sonnenfeld et al. (2013b)) ; spectroscopic measurements from Keck give the velocity dispersion as well as the deflector and source redshifts z_d and z_s; and the photometry, with the assumption of a Salpeter IMF (Initial Mass Function), gives the fraction of DM in half the effective radius of the deflector. Applying this methodology to a deep survey or to several surveys complementary in redshifts, one gets the cosmic evolution of the density profile inner slope in the lens galaxies and of their dark matter fraction. Figure1 from Ruff et al. 2011 Ruff et al. (2011) shows such evolution trends for a pilot analysis of 11 lenses from the SLACS (Sloan Lens ACS Survey, median $z_d \sim 0.2$), the SL2S (Strong Lensing Legacy Survey, median $z_d \sim 0.5$), and the LSD survey (Lenses Structure and Dynamics survey, median $z_d \sim 0.8$). We will talk in more detail about the SL2S later.

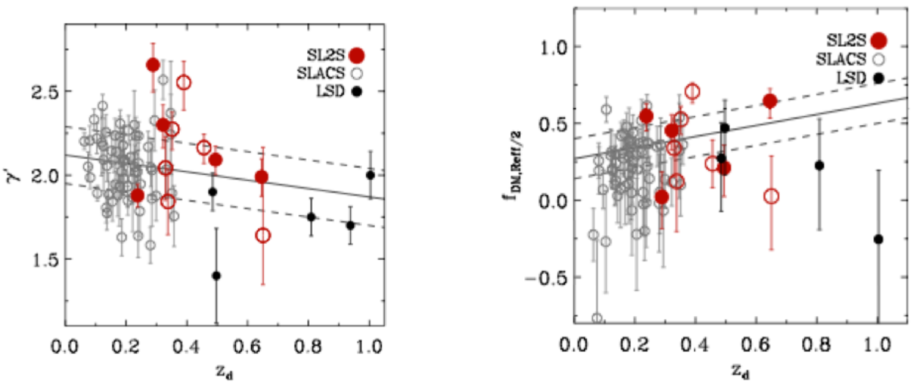

Fig. 1. *Cosmic evolution of the density profile inner slope (on the left) and of the dark matter fraction (on the right) for a pilot analysis of 11 lenses (Ruff et al. 2011 Ruff et al. (2011))*

As one can observe on figure 1, the results obtained so far are quite marginal and not convincing enough. Nevertheless, galaxy surveys are getting larger and larger : although we have only detected few hundreds of galaxy-scale lenses so far with spectroscopic and photometric methods, about 300 000 such lenses are expected in Euclid's 15 000 deg^2. In this context, we need a lens finder that gives a final sample as pure as possible, in other words, a lens finder that knows the lens equation, able to model each input candidate and to assign it a realistic lensing status ("yes it is a lens", "no it is not a lens", or even "there is some probability that it is a lens"). This automated modelling tool could also slighly increase the completeness of actual samples by finding original lenses like, for example, red gravitational arcs.

2 Lens modelling

The lens modelling is carried out by a program called SL-Fit (Gavazzi et al, 2012 Gavazzi et al. (2012)) that fits gravitational arcs from galaxy-galaxy lensing after subtraction of the deflector. It relies on parametric models optimized by MCMC (Monte Carlo Markov Chain) in a bayesian context, which might take more time than other optimization techniques but enables us to know the distribution of errors. A strong assumption is made concerning the potential of the lenses, that we model by a SIE (Singular Isothermal Ellipsoid). Furthermore, to mimic the kind of lenses which have most commonly been observed to date, made of blue arcs around a red galaxy, we assume a De Vaucouleurs luminosity profile for the deflector and an exponential luminosity profile for the source. In the finest application of SL-Fit, the lens models are then described by a set of 9 parameters, 3 for the lens potential and 6 for the source (the lens position is fixed at the center of the lens plane), which are : the Einstein radius R_{Ein}, the potential ellipticity, the potential orientation, the 2 coordinates of the source in the source plane, the intrinsic ellipticity of the source, the intrinsic orientation of the source, the intrinsic

flux of the source F_S, the effective radius of the source R_{effS}. Note that I chose to include the external shear ellipticity and orientation in the lens potential ellipticity and orientation respectively since those parameters are degenerated : it helps minimizing the number of parameters to fit.
Moreover, the MCMC is boosted by a process of simulated annealing (SA) : I have spent time optimizing the influence of the SA on the chain by determining the successive SA steps along with the typical number of iterations that secures the convergence of the chain at each step. This operation provides a Markov chain which is mainly driven by the priors in the early stages of its walk, and that will therefore move faster towards the distribution we are interested in; in parallel to this evolution, more and more weight is regiven to the likelihood as the latter is better known. Because of this modulation of the likelihood, the chain constituents in the SA phase are not considered for the final estimation of the parameters : only are retained the N' last MCMC elements which we think compose the convergence phase. The estimator we take for each parameter is then the median of its marginal distribution on these elements. I have found that we needed $N_{tot} \sim 2700$ iterations to ensure the convergence, along with $N' \sim 90$.
At this point, it appears that the whole challenge in designing this lens finder lies in a subtle balance between speed and robustness.

3 Training samples

3.1 Simulations

I have made ∼100 simulated lenses in HST filters (ACS F814W and F555W) and CFHT filters (Megacam i and g), using the GRAVLENS software (Keeton et al. 2001 Keeton (2001)) with the same assumptions than in section 5. for the MCMC modelling. The main interest of the simulated lenses sample is that, by construction, one knows the input values of each system parameters, to which one will be able to compare its output MCMC model values. This comparison allows to determine the maximum modelling accuracy of the lens finder.

3.2 CFHT data

The real data sample is made of 325 lens candidates found in the CFHTLS Wide (4 independant fields, a total of 155 deg^2, $i < 24.5$) in the framework of the SL2S, a survey that aims at providing a homogeneous sample of strong lenses at several scales from the CFHTLS. Those 325 galaxy-scale lenses result from a set of preselections involving :

- from photometric redshift catalogs containing 2.10^5 objects/deg^2 for T06 (Ilbert et al. 2006 Ilbert et al. (2006), we selected about 3000 ETG candidates/deg^2, that is 413000 Early-Type Galaxies (ETG) candidates in all.

- we run Raphael Gavazzi's Ringfinder Cabanac et al. (2007) (Gavazzi et al., in prep.) : a photometric lens finder for which the selection consists in looking for faint blue background galaxies embedded in bright red foreground galaxies, by substracting images observed in a redder band from their counterpart in a bluer band. We end up with about 14000 candidates in all.

- the remaining objects are then visually inspected and classified to yield a final sample of 325 candidates to be modelled by our new lens finder).

This sample firstly allows to improve the modelling by several methods trying to make it efficient in realistic observing conditions. Moreover, once these candidates are classified with the naked eye and with extra information (space-based images, spectra etc), one is able to encode the finder selection process in the modelling parameter space and derive from it the completeness and purity of the final sample (see the subsection about the selection step).
One can notice that in the first place, this lens finder is designed to work on ground-based data. Indeed, to date, ground-based images have been far more accessible than space-based data; furthermore, the seeing makes ground-based data much more difficult to model, so if the lens finder is able to work on this kind of data, it should be all the more efficient when adapted to a space-based survey like, for example, Euclid (see the part about further research).

4 Development of the lens finder

4.1 Optimizing the modelling step

I massively run SL-Fit modelling on my samples, beginning with ideal application conditions that I made more complex by stages so as to get closer to the observing conditions while saving time. Five progressive stages can be distinguished as follows :

- **(i) the fine modelling of lenses simulated in HST filters** :
  this is the zero complexity level for SL-Fit that fitted all the 9 MCMC parameters described above from clear space-based images;
- **(ii) the fine modelling of lenses simulated in CFHT filters** :
  I fitted all the 9 model parameters again but this time on grouned-based images blurred by seeing, which made a first complication;
- **(iii) the reduced modelling of lenses simulated in CFHT/g** :
  Focussing on the bluer band (g) of the ground-based simulations, where most of the gravitational arcs should more clearly appear, I tried to save time by reducing the number of parameters fitted in the MCMC process : this was made possible through additional simplifying assumptions on the source;
- **(iv) the reduced modelling of SL2S lens candidates observed in CFHT/g**:
  as might be expected, new difficulties arose while modelling the 325 preselected lens candidates, which are mostly due to :
  - the modelling assumptions being oversimple compared to the kinds of lenses that can be observed in the sky and the ones to discover;
  - the fact that, by construction, a perfect subtraction of the deflector can be performed on simulated lenses, whereas concerning the observed candidates, the subtraction is carried out by Galfit (Peng et al. 2002 Peng et al. (2002)) which will necessarily leave a smaller or larger residual at the center of the image : this residual may then be fitted instead of the counter image in the frequent case of double lenses, or will be added to the MCMC modelling residual;
  - the natural environment of observed candidates, which has not been mimicked in the simulations : this environment is made of galaxies that might again be wrongly fitted as a gravitational arc, misleading the Markov chain by making the parameter space exploration chaotic.
- **(v) the masked modelling of SL2S lens candidates observed in CFHT/g**:
  the first difficulty mentionned above while dealing with real data cannot be solved easily since it is not profitable to make the model assumptions more complex in the framework of an automated lens finder dedicated to massive modelling; concerning the environment and the central residual left by the deflector subtraction, I tried to tackle both problems by applying on the input image a uniform mask bounded by a small radius of 0.3" hiding the possible residual inside, and a large radius of 2.5" covering the galaxies beyond; the MCMC fit is still done on all the pixels in the image so that models predicting arcs in the corners are excluded : this disfavours the detection of lenses with large R_{Ein}, which is not a real problem since the latters are the easiest to find by other means.

Figure 2 shows a successful case of modelling by the lens finder on a candidate for which the lensing status has been confirmed by HST imaging.

4.2 Optimizing the selection step

Once a candidate has been modelled by the lens finder, the latter has to decide wether or not it deserves a lensing status. In order to train the finder to make a decision, R. gavazzi and I first classified the real data candidates in five categories, from "excellent candidates" to "bad candidates" and "undetermined candidates". This classification has been carried out with the naked eye but also with extra information from HST and Keck high resolution images and from Keck and VLT spectra. Afterwards I cut through the space of MCMC parameters and other related quantities following these kinds of justifications :

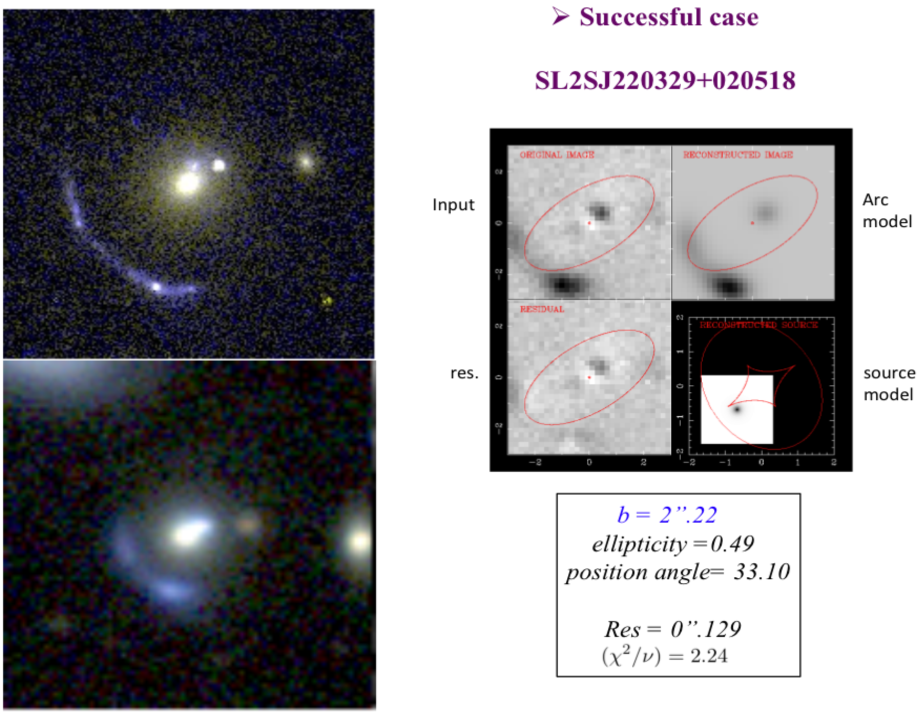

Fig. 2. *Modelling of the confirmed lens SL2SJ220329+020518, showing the HST image (top left), the CFHT image (bottom left), and SL-Fit output (top right) with (in the direction of reading) the input image, the arc model, the residual resulting from the subtraction of the arc model from the input image, and the source model (Brault et al, in prep).*

- physical justifications, like a reasonable range of values for the Einstein radius and the effective radius of the source : $0.''75 < R_{Ein} < 1.''19$ and $0.''015 < R_{effS} < 0.''13$, a lower limit on the amplification : $\mu > 1.5$ (I believe this limit could be easily set higher with a larger training sample), or an upper limit on the distance of the source from the projected center of the lens potential in the source place : $ds < 1.''02$;
- statistical justifications like constraints on the reduced chi-squared to keep a good fit to the data : $1. < (\chi^2/\nu) < 2.5$;
- practical justifications like an upper limit on the time of convergence of the Markov chain in the context of automated massive modelling : $t_{conv} < 7$ min;
- empirical justifications from the repartition of the classified candidates in the parameter space : figure 3 is an example of this repartition in the 2-D space showing the amplification as a function of the Einstein radius plus the error on the Einstein radius; of course the precise contraints given on some parameters above are tuned to converge with those empirical findings.

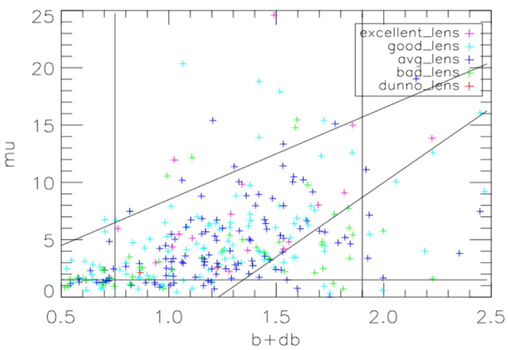

Fig. 3. *Distribution of the 325 SL2S lens candidates in the amplification–{Enstein radius+error} plane with scientifically and empirically justified cuts*

5 Conclusions

As things stand, the main results concerning my lens finder can be summed up as follows :

- I showed that a 7-parameter model assuming a spherical source (which amounts to fixing the the ellipticity and orientation of the source) can recover the Einstein radius from the CFHT simulations with a precision of $\sim 7\%$.
- The median MCMC convergence time is ~ 4 minutes on real data, and could be easily reduced by a factor of 10 using more direct optimization techniques.
- With the selection contours I defined in the parameter space, the final sample selected by the lens finder is $\sim 55\%$ pure and $\sim 37\%$ complete (if one considers the "good" and "excellent" candidates) with the time constrain $t_{conv} < 7$ min. Relaxing that latter contraint, one now gets a purity of $\sim 50\%$ along with a completeness of $\sim 64\%$: there will always be a trade-off to deal with between completeness and purity, but this shows that a high completeness should be reachable by saving even more time in the modelling process.Those are quite encouraging results even though one should keep in mind that this selection is made downstream from the set of preselections detailed above (see section 6, about the training samples) : in particular, the purity should be lower for the new lens finder running alone.

  The automated lens finder described above is being improved in several aspects like its model quality, its execution time, its lens selection, its autonomy and its portability in order to be adaptable and applicable to current and future, ground-based and space-based, wide-field imaging surveys.

References

Cabanac, R. A., Alard, C., Dantel-Fort, M., et al. 2007, A&A, 461, 813

Gavazzi, R., Treu, T., Marshall, P. J., Brault, F., & Ruff, A. 2012, ApJ, 761, 170

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Keeton, C. R. 2001, ArXiv 0102340

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H. 2002, AJ, 124, 266

Ruff, A. J., Gavazzi, R., Marshall, P. J., et al. 2011, ApJ, 727, 96

Sonnenfeld, A., Gavazzi, R., Suyu, S. H., Treu, T., & Marshall, P. J. 2013a, ArXiv 1307.4764

Sonnenfeld, A., Treu, T., Gavazzi, R., et al. 2013b, ArXiv 1307.4759