

A compact group lens modeled with GIGA-Lens: Enhanced inference for complex systems

F. Urcelay^{1,*}, E. Jullo², L. F. Barrientos¹, X. Huang³, and J. Hernandez¹

¹ Institute of Astrophysics, Pontificia Universidad Católica de Chile, Santiago, Chile

² Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France

³ Department of Physics & Astronomy, University of San Francisco, San Francisco, USA

Received 17 January 2024 / Accepted 14 December 2024

ABSTRACT

Context. In the era of large-scale astronomical surveys, the fast modeling of strong lens systems has become increasingly vital. While significant progress has been made for galaxy-scale lenses, the development of automated methods for modeling larger systems, such as groups and clusters, is not as extensive.

Aims. Our study aims to extend the capabilities of the GIGA-Lens code, enhancing its efficiency in modeling multi-galaxy strong lens systems. We focus on demonstrating the potential of GPU-accelerated Bayesian inference in handling complex lensing scenarios with a high number of free parameters.

Methods. We employed an improved inference approach that combines image position and pixelated data with an annealing sampling technique to obtain the posterior distribution of complex models. This method allowed us to overcome the challenges of limited prior information, a high number of parameters, and memory usage. We validated our process through the analysis of the compact group lens system DES J0248-3955 and we present the relevant VLT/X-shooter spectra.

Results. We measured a redshift of $z = 0.69 \pm 0.04$ for the group, and $z = 1.2722 \pm 0.0005$ for one of the extended arcs. Our enhanced method successfully constrained a lens model with 29 free parameters and lax priors in a remarkably short time. The mass of the lens is well described by a single dark-matter halo with a velocity dispersion of $\sigma_v = (690 \pm 30) \text{ km s}^{-1}$. The model predicts the presence of a second source at the same redshift and a third source at approximately $z \sim 2.7$.

Conclusions. Our study demonstrates the effectiveness of our lens modeling technique for dealing with a complex system in a short time using ground-based data. This presents a considerable prospect within the context of large surveys, such as LSST, in the future.

Key words. gravitational lensing: strong – methods: data analysis – galaxies: groups: individual: DES J0248-3955

1. Introduction

Strong gravitational lenses are a valuable tool in astrophysics, with diverse applications. They allow us to measure the mass distribution of galaxies and clusters (Mellier et al. 1993; Limousin et al. 2005, 2022; Auger et al. 2010; Sharon et al. 2020), often providing enough sensitivity to constrain the nature of dark matter (Dalal & Kochanek 2002; Koopmans et al. 2009; Newman et al. 2013; Vegetti et al. 2024). Additionally, strong lenses offer unique opportunities to study distant, magnified galaxies (Lotz et al. 2017; Atek et al. 2015) and constrain the expansion rate of the universe through time-delay Hubble constant measurements of multiply imaged supernovae and quasars (Refsdal 1964; Wong et al. 2020; Kelly et al. 2023; Acebron et al. 2023; Pascale et al. 2024).

Existing surveys have already revealed a substantial number of lensing systems and this count is expected to increase significantly with upcoming imaging surveys. The forthcoming LSST and Euclid surveys are set to usher in a new era in lensing studies; they are poised to discover 120 000 and 170 000 strong galaxy-galaxy lenses, respectively (Collett 2015). Given the assumption of the lens distribution delineated in Oguri (2006), this projection would correspond to approximately 20 500 (29 000) group-scale lenses and 5 900 (8 300) cluster-scale lenses within the

scope of LSST (Euclid). Coupled with machine learning detection methods (Jacobs et al. 2019; Magro et al. 2021; Huang et al. 2021; Rojas et al. 2022; Shu et al. 2022; Zaborowski et al. 2023), these surveys will provide an extensive sample of galaxy-to-cluster scale lenses, facilitating statistical applications.

For these applications, it is essential to carry out the accurate, fast, and automated modeling of strong gravitational lenses. The lens model parameterizes the lens mass distribution and source properties, and constraining these parameters to observations is a crucial step. Traditional modeling techniques can be time-consuming, taking hours and even days for cluster-scale lenses to explore the parameter space. This makes the large number of lenses delivered by forthcoming surveys challenging to handle efficiently.

This paper addresses the specific challenges of achieving fast modeling for group- and cluster-scale lenses. These lenses are commonly modeled using the positions of strongly lensed images, usually knots of star formation in extended arcs (Cerny et al. 2018; Sharon et al. 2020; Mahler et al. 2023; Bergamini et al. 2023). Identifying and pairing multiple image sets requires a level of expertise that is generally challenging for automated modeling. Furthermore, the identification of point sources requires high-resolution imaging, thereby limiting the datasets to those from space-based telescopes. Despite these drawbacks, using the positions of these lensed images is fast, making it a preferred method for modeling complex mass distributions.

* Corresponding author; fjurcelay@uc.cl

While the use of extended images instead of positions would be ideal for clusters thanks to the detailed information they provide, the computational time required becomes prohibitive. This is because the cluster mass model involves numerous parameters, and each source requires shape model parameters.

Pioneering studies in automated modeling, such as the works of Zitrin et al. (2012) and Stapelberg et al. (2019), have successfully used the light of cluster members to model cluster-scale lenses without relying on the identification of lensed images. While fast, this approach assumes that light “traces mass” and requires a calibration sample, making it dependent on the quality of the selected sample. It also makes it less sensitive to the systemic peculiarities, such as variations on the mass-to-light ratio and unrelaxed mass distributions.

On the other hand, galaxy-scale lenses are modeled through simulations, reconstructing the surface mass of the lens and the surface brightness of the source and lens galaxy, which are then compared with observations (Birrer et al. 2015; Nightingale et al. 2018). Although this approach can be computationally expensive, significant progress has been made in achieving the modeling speed needed for large surveys, employing machine learning aided by graphics processing units (GPU) (Hezaveh et al. 2017; Morningstar et al. 2019; Pearson et al. 2019; Chianese et al. 2020; Pearson et al. 2021; Schuldt et al. 2021; Biggio et al. 2022; Adam et al. 2023; Schuldt et al. 2023). Among them, GIGA-Lens (Gu et al. 2022), a promising new software for fast and automated galaxy-galaxy lens modeling, leverages GPU acceleration to achieve comprehensive Bayesian posterior estimates in a few minutes for each system.

Although this method has been successful for galaxy-scale systems, scaling this method to group or cluster lenses is challenging. Cluster lenses often exhibit complex, unrelaxed, and asymmetrical mass distributions (Merten et al. 2011; Limousin et al. 2012; Jauzac et al. 2016), significantly influenced by subhalos (Meneghetti et al. 2007), often resulting in limited prior knowledge of the specific system. This, combined with the large number of parameters needed to describe both the lens and the sources may lead to the “curse of dimensionality.” In other words, the exploration space grows exponentially with the number of free parameters and when combined with the weak prior, this would require an increasing number of samples to obtain the posterior distribution.

We propose adapting the established methods from galaxy-scale lens modeling to overcome the challenges in modeling lenses at the group and cluster scales, by modifying the GIGA-Lens code with a novel inference approach to handle the increased model complexity. This enhancement involves a hybrid approach that integrates the approximate positions of multiple images with their surface brightness to efficiently estimate the lens mass and source brightness parameters.

The design is primarily focused on achieving fast inference for these systems using ground-based data as constraints. This is of special interest for lens systems on the LSST survey, as it will have a high depth but low resolution compared to space-based surveys like Euclid. This resolution constraint might hinder the detection of galaxy-scale lenses, but larger-scale lenses remain unaffected. We illustrate the method by applying it to a group-scale strong lens candidate DES J0248-3955 (Jacobs et al. 2019) to demonstrate its capability to provide accurate lens models, even with the limited resolution challenges posed by surveys such as LSST.

The system under study is relevant on its own as a potential candidate for having two source planes. Collett et al. (2012) has shown that it is possible to constrain the cosmological

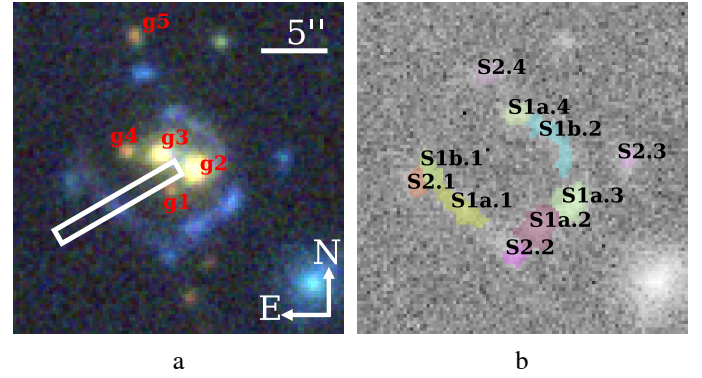


Fig. 1. Identification of the group members and multiple images, along with the X-shooter slit configuration. (a) *grz* color composite image of the system, the red labels indicate possible group members, and the white box represents the X-shooter slit passing through image S1a.1 and between members g2 and g3. (b) *g*-band image with Sérsic surface brightness model of g1 to g5 subtracted, overlaid with the segmentation of the multiple images labeled in black.

parameters, including the density of matter, Ω_m , and the equation of state of dark energy, w , independently of H_0 , with this kind of lens system. While cluster-scale lenses frequently feature multiple source systems, allowing for robust cosmological constraints (Jullo et al. 2010; Caminha et al. 2022), their complex mass distribution limits their applicability to a few well-studied lenses. Galaxy-scale strong lenses with multiple sources are rarer, but more suitable for statistical application (Sharma et al. 2023). As a galaxy group lens, DES J0248-3955 is a mid-step between the two: more massive than galaxies, but simpler than clusters. Thus, it is an interesting candidate for probing cosmology (in future observations) and testing the GIGA-Lens code.

The remainder of the paper is organized as follows. Section 2 presents the details of the data. In Section 3, our novel hybrid method is described. Section 4 introduces the lens model for the system. Section 5 presents the resulting lens model and possible sources of errors. Section 6 discusses the advantages and disadvantages of the algorithm. Magnitudes are reported in the AB system. Our analysis is based on the assumption of a flat Λ CDM model with $\Omega_{m,0} = 0.3$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

2. Data

The system under study is a compact group lens candidate listed in *NeuraLens* (Storfer et al. 2024), a catalog of 3057 strong lens candidates, including galaxies, groups, and clusters. The network assigned a probability of 1.0 that it was a lens, and a human assessment gave it the maximum rating. This system was first discovered by a systematic search with neural networks in DES imaging data by Jacobs et al. (2019). O’Donnell et al. (2022) confirmed it with a rank of 6/10 and measured an Einstein radius of $\theta_E = 3.95'' \pm 0.40''$ as the average distance between each image and the brightest lens. A photometric redshift $z_{\text{phot}} = 0.67 \pm 0.02$ was derived from the DES Y6 photometry for the brightest group galaxy (BGG) (g2 in our labeling) located at 02:48:09.54, −39:55:48.4.

We used the *grz*-band Dark Energy Camera (DECam) images from the DESI Legacy Imaging Survey data release 9 (Dey et al. 2019) to produce a composite color image shown in Fig. 1a. We identified the central group members (g1–g5) based on the *g* – *r* and *r* – *z* colors shown in Table 1. We modeled their surface brightness with the same code used for modeling the

Table 1. Photometric redshifts, colors, and magnitudes of group members.

Member	Photo z	$g - r$	$r - z$	z
g1	0.6 ± 0.2	1.63	1.69	21.80
g2	0.68 ± 0.06	1.80	1.74	18.79
g3	0.66 ± 0.06	1.58	1.64	19.11
g4	0.8 ± 0.3	1.88	1.69	21.18
g5	0.66 ± 0.07	1.74	1.76	21.04

Notes. Photometric redshifts are from Zhou et al. (2021). The spectroscopic redshift of the BGG (g2), $z_{BGG,spec} = 0.685 \pm 0.002$, is consistent with its photometric redshift $z_{BGG,phot} = 0.68 \pm 0.06$. Colors were obtained using a $1''$ diameter circular aperture, and the z -band magnitudes were computed using a Sérsic model.

lens, but without incorporating lens deflection. Each galaxy was modeled with a single Sérsic profile and the photometric magnitudes derived from the model are reported in Table 1. We also provide the photometric redshifts from Zhou et al. (2021). These were calculated using a random forest regression algorithm that incorporates photometric data in the grz bands, as well as $W1$ and $W2$ bands, in addition to information on the morphology of the galaxies.

The g -band image after g1–g5 surface brightness subtraction is displayed in Fig. 1b, featuring the segmentation of multiple-image candidates obtained with Photutils software (Bradley et al. 2023). The image labeling has a format $Sij.k$ being i the source plane, j the identifier of a family of images from the same source in that plane, and k the identifier of the image within the family.

Spectroscopic data were acquired from the X-shooter instrument mounted on the Very Large Telescope (VLT) Unit Telescope 3 (Vernet et al. 2011), through the observation program ID 110.23U2.001 (PI: Jullo). This program aims to characterize the Neuralens selection function by measuring the redshift of 100 lens systems (Jullo et al., in prep.). The observation was made with an air mass of 1.041 and a seeing of about $1.75''$. The time spent on target was divided into exposures of 2×1468 s, 2×1384 s, and 5×300 s each in the UBV, VIS, and NIR channels respectively. Each exposure was reduced independently with the ESO/REFLEX pipeline in STARE mode. Sky lines were template subtracted and the sky residual present in a fraction of the slit was subtracted to produce the spectra presented in Fig. 2.

As depicted in Figure 1a, the slit passes through the east arc S1a and close to the group members g2 and g3. As shown in Figure 2, the source spectrum presents a clear H_α emission line at redshift $z_{S1a} = 1.2722 \pm 0.0005$. An [O II] doublet is also detected at this redshift but is highly affected by skylines. The 1D spectrum of the source was obtained by optimal spectral extraction (Horne 1986), with a spatial profile obtained with an aperture of 12 \AA around the H_α emission.

The one group member with a spectroscopic redshift has a spectrum that does not show any emission lines. A redshift of $z_{BGG,spec} = 0.685 \pm 0.002$ was deduced from the K and H absorption lines. This spectrum is associated with members g2 and g3, likely dominated by g2 as the BGG. The offset between the dark matter halo and the brightest galaxy is expected to be in the order of a few kpc, according to simulations (Roche et al. 2024) and observations (Harvey et al. 2017), meaning that the BGG is representative of the group redshift. This is supported by the photometric redshifts of the group members, as shown in Table 1, with the variance-weighted average redshift of the

central galaxies g1–g4 being $z_{L-phot} = 0.67 \pm 0.04$. The 4000 \AA break falls within the r band, allowing the g , r , and z bands to bracket the break and determine their redshift.

For the redshift of the group, we adopt the spectroscopic redshift of the BGG as its centroid. Given that this redshift is based on a single galaxy, we account for the uncertainty in the group redshift by incorporating the standard error of the variance-weighted average of the photometric redshifts of the central members. This approach yields an estimated group redshift of $z_L = 0.69 \pm 0.04$. While this approach incorporates potential variations within the group, additional observations are required to confirm the accuracy of the group redshift, given the limited number of spectroscopic measurements currently available.

3. Method

Gravitational lensing arises from the bending of the light path in curved spacetime, a phenomenon described by the lensing equation (Schneider et al. 1992):

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

where β is the angular position in the source plane, θ in the image plane, and α is the deflection angle due to the lens at the image position. In the context of strong lensing, this equation yields multiple solutions that account for multiple images. The deflection is related to the convergence $\kappa(\theta)$ through:

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

where $\kappa(\theta)$ is the projected mass density of the lens $\Sigma(\theta)$ over the critical density Σ_{crit} , while the latter is dependent on the angular diameter distances between the observer and the lens (D_L), the observer and source (D_S), and the lens and source (D_{LS}):

$$\Sigma_{crit} \equiv \frac{c^2 D_S}{4\pi G D_L D_{LS}}. \quad (3)$$

This means that two sources at different distances are deflected differently by the same lens. To be able to model sources at different redshifts z_1 and z_2 , we convert α from one source plane to another by rescaling it by the distance ratio η , defined as:

$$\frac{\alpha_1}{\alpha_2} = \frac{D_{LS}(z_1) D_S(z_2)}{D_S(z_1) D_{LS}(z_2)} \equiv \eta. \quad (4)$$

We used a parametric modeling approach, that is, the lens potential is described by an analytic profile dependent on a set of parameters, Θ . These parameters can be inferred by constraining them to observations. In cluster lens modeling, the main constraint is the position of the multiple images. The model parameters can be inferred by minimizing the χ^2 between the observed image positions and the ones predicted by the model (Schneider et al. 1988):

$$\chi_\theta^2 = \sum_{i=1}^N \sum_{j=1}^{n_i} \frac{(\theta_{obs,j} - \theta_{model,j}(\Theta))^2}{\sigma_{\theta,j}^2}, \quad (5)$$

for a system comprising N sources, each source i exhibits n_i images with positions, θ_j , and associated errors, $\sigma_{\theta,j}^2$.

This approach faces two primary challenges: matching observed and predicted images, along with the computationally

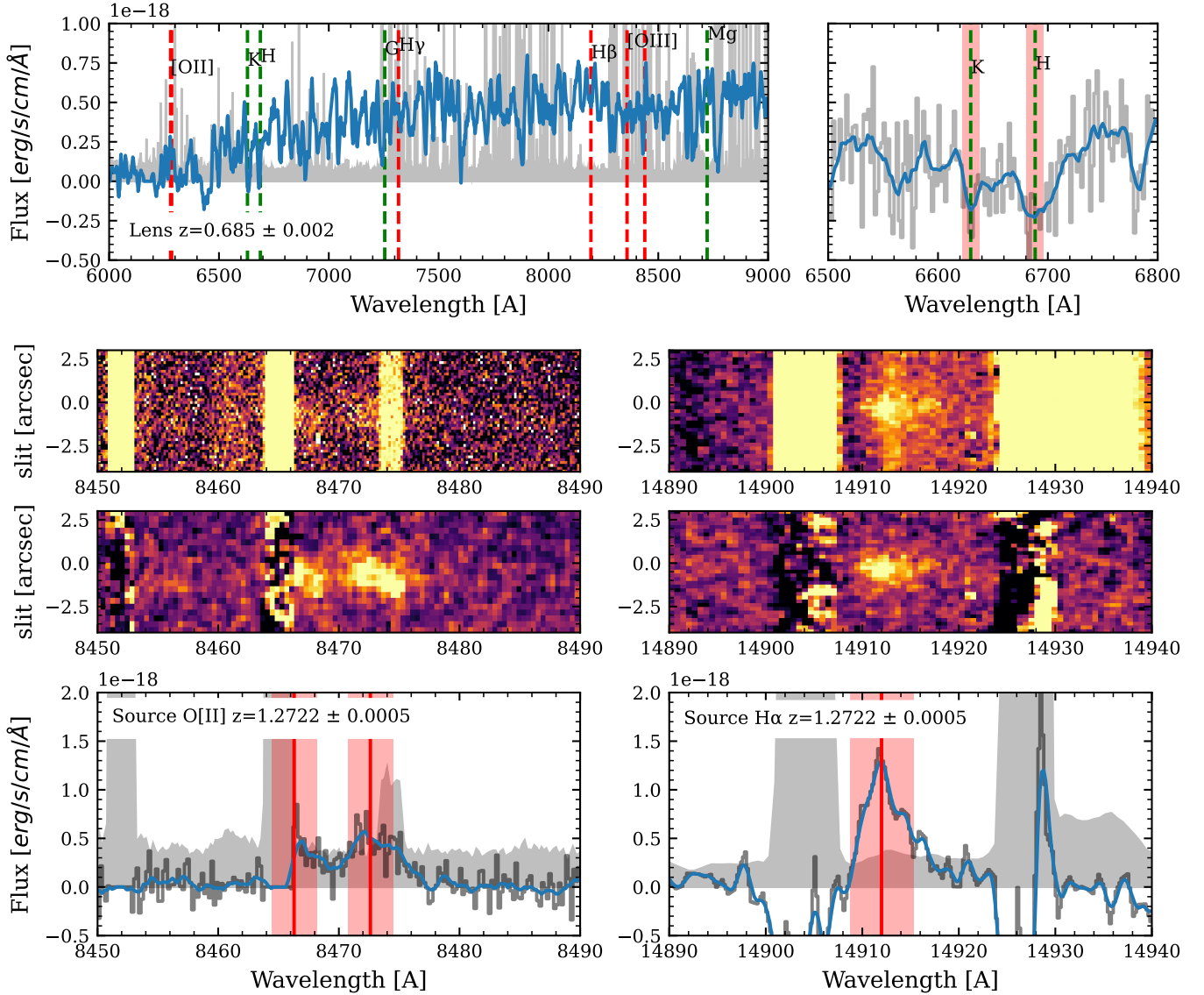


Fig. 2. X-shooter spectra of the BGG and the source S1a.1. *Top panels:* X-shooter 1D spectrum of the BGG at redshift $z = 0.685 \pm 0.002$. The redshift is driven by the K and H absorption lines (right panel), as well as the shape of the continuum. The red-shaded region around the K and H lines represents the associated uncertainty. *Second-row panels:* 2D spectrum of the source S1a.1 at redshift $z = 1.2722 \pm 0.0005$ without sky subtraction, centered around the $[O II]$ (left) and $H\alpha$ (right). *Third row panels:* 2D spectrum with sky subtraction applied. The $[O II]$ spectrum has been rebinned by a factor of 2×2 to enhance the signal-to-noise ratio. *Bottom panels:* 1D sky-subtracted spectra of both $[O II]$ and $H\alpha$ emission lines at redshift $z = 1.2722 \pm 0.0005$, the red line shows the expected position of the line and the red shaded region shows the uncertainty. The red vertical line marks the expected position of each line, and the red shaded region represents the associated uncertainty. In all 1D spectra, the gray histogram displays the unbinned spectrum, the blue curve shows the same spectrum with a binning of 3 pixels and the gray shaded region represents the rescaled sky level for visualization purposes.

expensive task of inverting the lens equation. A viable alternative is to employ a first-order approximation (Kochanek 1991), $\delta\beta \approx \mu^{-1} \delta\theta$, which allows us to compute χ^2 in the source plane:

$$\chi^2_\beta = \sum_{i=1}^N \sum_{j=1}^{n_i} \frac{(\beta(\theta_{obs,j}; \Theta) - \beta_{model}(\Theta))^2}{\mu^{-2}(\theta_{obs,j}; \Theta) \sigma_{\theta,j}^2}. \quad (6)$$

Here, $\beta(\theta_{obs,j}; \Theta)$ is the position of the observed image, j , mapped to the source plane, and μ is its magnification. This is straightforward to compute, as it avoids the need to invert Eq. (1). While most elliptical mass profiles lack an analytical expression for μ , our code efficiently calculates it using automatic differentiation, ensuring accurate gradients and simplifying the computation.

As the χ^2 value given by Eqs. (5) and (6) depend on the image positions, they only give a good constraint to the model when point-like features can be identified in the images. Thus, this calculation is only possible when high-resolution data are available.

In the case of galaxy-scale lens modeling, χ^2 is computed with the pixel-to-pixel comparison between the surface brightness of the observed image, \mathcal{I}_{obs} , (in units of counts/s) and a simulation, \mathcal{I}_{model} . For this, the model needs to include not only the source position but also its surface brightness, and the χ^2 is defined as (Birrer et al. 2015):

$$\chi^2_{pix} = \sum_{x,y} \frac{(\mathcal{I}_{obs}(x,y) - \mathcal{I}_{model}(x,y; \Theta))^2}{\sigma_{pix}^2(x,y; \Theta)}. \quad (7)$$

The error, σ_{pix}^2 , includes contributions from background noise and Poisson noise, related to exposure time, t_{exp} , and gain, \mathcal{G} :

$$\sigma_{\text{pix}}^2 = \sigma_{\text{bkg}}^2 + \sum_{x,y} \frac{\mathcal{I}_{\text{model}}(x,y;\Theta)}{\mathcal{G} t_{\text{exp}}}. \quad (8)$$

This method offers robust constraints on extended images such as arcs and Einstein rings in galaxy-galaxy lensing, and has proven to be reliable on ground-based data (Knabel et al. 2023; Schuldt et al. 2023).

Parameter estimation is accomplished through a Bayesian method, assuming a normal distribution for the likelihood:

$$\log \mathcal{L} = -\frac{1}{2} [\chi^2 + \log(2\pi\sigma^2)]. \quad (9)$$

This requires the choice of a prior on the parameters, $\pi(\Theta)$, and exploring the posterior distribution, $\mathcal{P}(\Theta; \mathcal{D}) = \pi(\Theta) \mathcal{L}$ (possibly in multiple steps).

The Giga-Lens code (Gu et al. 2022) relies on Eq. (7) to obtain the posterior in three steps. First, the parameters that yield the maximum probability (maximum a posteriori, MAP) are obtained by a multi-start gradient descent. This method involves executing multiple optimization trials from different initial parameter values to reduce the chance of getting stuck in local minima. Then, the posterior distribution is approximated with a Normal distribution centered on the MAP estimate using variational inference. This method approximates complex probability distributions by optimizing a more straightforward distribution, thereby improving computational efficiency. Finally, the posterior distribution is obtained by the Markov chain Monte Carlo (MCMC) method using the Hamiltonian Monte Carlo (HMC) kernel, initialized by the distribution obtained in the previous step. The three steps use gradient information obtained by automatic differentiation to guide the exploration efficiently. Although this method has proven to be fast in galaxy-galaxy scenarios, the multi-start gradient descent stage may fail to obtain the MAP in the case of a large exploration space, leading to a biased result.

In this paper, we propose an alternative exploration method to account for the complexity of multiple-galaxy lens systems. We primarily rely on Eq. (7) as the main constraint for our model. However, instead of simulating the entire pixel matrix, we limited it to a region of interest, typically defined by a mask that encompasses the multiple images (as depicted in Fig. 1a). During the model evaluation, the pixel grid is represented by a sparse matrix instead of a dense one. This approach is more memory-efficient and avoids the necessity of simulating the surface brightness of cluster members or field galaxies distant from the images.

The minimization of the Eq. (7) can be slow initially, as the simulated images can deviate significantly from the observed ones, resulting in a flat χ^2 surface. To address this, we initially employed Eq. (6) to approximate the high-probability region and subsequently refine the solution using Eq. (7). This exploration step is much faster as the χ^2 decreases when the images are mapped closer in the source plane, leading to a convex surface in the same situation when the other is flat. Additionally, it does not require computing the surface brightness of the deflected source, this leads to a much faster likelihood evaluation.

To join both steps we employ a simulated annealing approach by weighting each likelihood term by a power λ and gradually

transitioning from the prior $\pi(\Theta)$ to the posterior $\mathcal{P}(\Theta; \mathcal{D})$ by varying λ in two stages:

$$\mathcal{P}_{\lambda_1, \lambda_2}(\Theta; \mathcal{D}) = \pi(\Theta) \mathcal{L}_{\beta}^{\lambda_1} \mathcal{L}_{\text{pix}}^{\lambda_2}. \quad (10)$$

Here, \mathcal{L}_{β} is the likelihood given by the source plane positions, as per Eq. (6), and \mathcal{L}_{pix} by pixels, as per Eq. (7). Initially, we set $\lambda_2 = 0$ and increase λ_1 from 0 to 1, this way, we start by sampling the prior and increase the weight of \mathcal{L}_{β} to approach the posterior constrained by the position of the images. Subsequently, we increase λ_2 to 1 and set $\lambda_1 = 1 - \lambda_2$ to slowly mutate the previous posterior to the one constrained by pixels. Finally, we keep sampling with fixed $\lambda_1 = 0$ and $\lambda_2 = 1$ to obtain accurate probability contours. The prior remains unchanged throughout the sampling process.

The sampling is performed using a sequential Monte Carlo (SMC) method available through TensorFlow Probability (Dillon et al. 2017), heavily based on Del Moral et al. (2012). The posterior is modeled with a set of weighted “particles” that are updated using an MCMC transition kernel and resampled to generate a new set of particles. For the MCMC step, we employed the HMC kernel, which uses gradient-informed evolution through an analog of Hamiltonian dynamics, enabling the exploration of complex posteriors with minimal likelihood evaluations. In the first phase, the particles are initialized from a random sample from the prior of the lens mass. In the second one, we took a subsample of the particles resulting from the previous step and extended their dimension with a random sample from the source light prior.

The code is implemented in JAX (Bradbury et al. 2018), a library for high-performance machine learning that enables GPU parallelization and efficient computation of gradients through automatic differentiation. Additionally, we employ TensorFlow Probability (Dillon et al. 2017), a JAX-compatible package for probabilistic programming. This allows us to compute multiple simulated images simultaneously, while handling the complex optimization and sampling processes effectively.

4. Lens model

The adopted lens model corresponds to a singular isothermal ellipsoid (SIE) mass profile that describes the mass of the group. The radial SIE convergence is given by:

$$\kappa(r) = \frac{1}{2} \frac{\theta_E}{r} \quad (11)$$

where θ_E is the Einstein radius of the lens relative to a source at a distance D_S , which relates to the velocity dispersion of the deflector as:

$$\sigma_v = \sqrt{\frac{\theta_E c^2}{4\pi} \frac{D_S}{D_{LS}}}. \quad (12)$$

We also included a shear component that describes the effect of external structures. Each source galaxy’s surface brightness is described with a Sérsic profile. Consequently, the model encompasses a total of seven free parameters for the lens mass and seven parameters per source galaxy.

The lens model is constrained using the foreground-light subtracted g -band image, limited to the pixels within the segmentation image depicted in Fig. 1b, with an additional extension of 3 pixels. This ensures that we include adjacent pixels and prevents the model images from extending further than the

Table 2. Prior distribution for the lens parameters.

Lens mass:	θ_E	$\sim \mathcal{U}(3, 8)$
	x, y	$\sim \mathcal{U}(-2, 2)$
	e_1, e_2	$\sim \mathcal{TN}(0.0, 0.2; -0.5, 0.5)$
Source light:	R	$\sim \mathcal{N}(0.4, 0.1)$
	n	$\sim \mathcal{TN}(3.0, 2.0; 0.5, 5.0)$
	I	$\sim \mathcal{N}(0.25, 0.07)$
	x, y	$\sim \mathcal{U}(-3, 3)$
	e_1, e_2	$\sim \mathcal{TN}(0.0, 0.2(0.05); -0.5, 0.5)$
Source distance:	η	$\sim \mathcal{U}(0.6, 0.7)$

Notes. The lens mass profile is parametrized by the Einstein radius θ_E , and the source surface brightness by the Sérsic radius R , index n , and half-light I . For both cases, the elliptical geometry is described by its center (x, y) and eccentricities e_1, e_2 . $\mathcal{U}(a, b)$ denotes a uniform distribution, $\mathcal{N}(\mu, \sigma)$ a normal distribution, and $\mathcal{TN}(\mu, \sigma; a, b)$ a truncated normal distribution. The value in parentheses is a different prior for S2 eccentricities as is expected to have low elongation. The source distance prior limits z_{S2} to (2.1, 3.2). All positions and radius $(x, y, \theta_E \text{ and } R)$ are expressed in arcseconds, and the positions are relative to the central point of the pixelated image. I is in micro-Jansky per arcsecond squared. A normal prior on e_1 and e_2 is needed to reflect a rotational symmetry (see Gu et al. 2022).

observed ones. Regarding the image position, we use the flux-weighted centroid of the brightest 20% pixels within the segment and its uncertainty.

Due to large errors in the colors, which hindered an accurate match of the multiple images, we employed the model to predict the image families. For this, we explored various models considering a single source and different image combinations until we obtained a model that successfully reproduces images S1a.1-4, from a single source, denoted S1a, located at the spectroscopic redshift $z_{S1} = 1.2722$.

We then employed this model to predict the counterimages of the remaining one to two image systems by mapping each image to the source plane and then back to the image plane, assuming an arbitrary redshift. With this method, we were able to predict the presence of a second source, denoted S1b, also at z_{S1} , which contributes to the images S1b.1-2, and a third source, labeled S2, at a higher redshift ($z_{S2} \approx 3$), generating a quadruple image pattern comprising images S2.1-4.

The conclusive model incorporates all three aforementioned sources and introduces an additional parameter η that accounts for the distance to source S2 (see Eq. (4)). This yields a total of 29 independent parameters, which collectively define the model's complexity and flexibility. We explored the parameter space with the pipeline described in Sect. 3 using 1000 particles for the SMC, then making 100 sampling steps to comprise a total of 100 000 samples.

We assumed non-informative uniform priors for all the parameters describing the lens mass distribution, the positions of the sources and η ; along with normal priors for the parameters describing the Sérsic profiles. The complete prior is shown in Table 2. This prior is highly flexible, covering a broad spectrum of lenses and sources. It is important to note that the same prior applies to all three sources, except for the ellipticity of S2, which is more tightly constrained due to the minimal elongation observed in the four images.

The final model is shown in Fig. 3 and is discussed in further detail in the following section.

5. Results

5.1. Model constraints and possible sources of errors

Despite the simplicity of the mass profile, the model shown in Fig. 3 produces a suitable reproduction of the observed configuration of image systems S1a, S1b, and S2 within the errors. The reduced χ^2 is $\chi^2_{v,pix} = 1.0$ within the simulated region (enclosed in green in Fig. 3), and $\chi^2_{v,pix} = 1.1$ when only considering pixels with a signal-to-noise ratio greater than 1 (the same as in Fig. 1b). The normalized residual shown in Fig. 3 indicates that the reproduction of the S1a images is at the noise level. In contrast, the images of S1b are not as accurately reproduced as those of S1a. Specifically, S1b.2 displays three bright clumps that exceed the model predictions. This discrepancy could be due to the magnification effects of adjacent group members g1 and g2, or to a source complexity that a smooth Sérsic profile does not capture. However, the model replicates the shape of the arc. The four images of S2 are well reproduced by the model, favoring the hypothesis that it is a strongly lensed source at higher redshift. The residual slightly exceeds that of S1a, attributed to a positional offset of the images, with a root mean square error of $0.22''$ in the image plane. The offset may be due to the influence of the source being doubly lensed, first by S1 and then by the group (see below).

As indicated in Table 3, the model appears highly constrained and the uncertainties in the lens parameters are likely underestimated. For example, the Einstein radius, θ_E , has an uncertainty of only 0.1%, which is extremely low considering the low signal-to-noise ratio and resolution of the data. Although the simple SIE model for the halo and Sérsic profiles for the sources effectively mimic the data, their rigidity may cause only a limited portion of the parameter space to replicate the observations. This problem might be mitigated by adopting a more sophisticated model for the lens mass, such as an elliptical power law with a free slope, or by including a mass model for the group members.

The model has a small but non-negligible shear of $\gamma_{ext} = 0.015$, which may be attributed to the faint group members that extend further than the center described by the model, as well as a richer galaxy group located approximately $2'$ to the south.

5.2. Lens and sources analysis

As seen in Table 3, our measurement of the Einstein radius for the source S1 is considerably larger than that of O'Donnell et al. (2022) $\theta_E = 3.95'' \pm 0.40''$ measured with the positions of the images. The difference is probably due to the high ellipticity of the halo in our model, namely, $\epsilon = 0.203 \pm 0.002$ (defined as $\epsilon = (1 - q)/(1 + q)$ with q the axis ratio), which deviates largely from a spherical case ($\epsilon = 0$) assumed by O'Donnell et al. (2022). Using this Einstein radius and Eq. (12), we measured a velocity dispersion for the galaxy group of $\sigma_v = (690 \pm 30) \text{ km s}^{-1}$, considering the uncertainties associated with the model as well as the redshifts of the source and lens.

The distance ratio between the source planes S1 and S2 deviates from unity with a discrepancy greater than 5σ . This deviation persists even when accounting for a more conservative uncertainty of 10%, yielding $\eta = 0.63 \pm 0.06$. This provides strong evidence for the system having two distinct source planes, thereby making it capable of constraining the intrinsic cosmology. According to Collett et al. (2012), the ratio of the deflection angles, η , is a measurement of the distance ratios (see Eq. (4)), which can be compared with a redshift measure to determine the

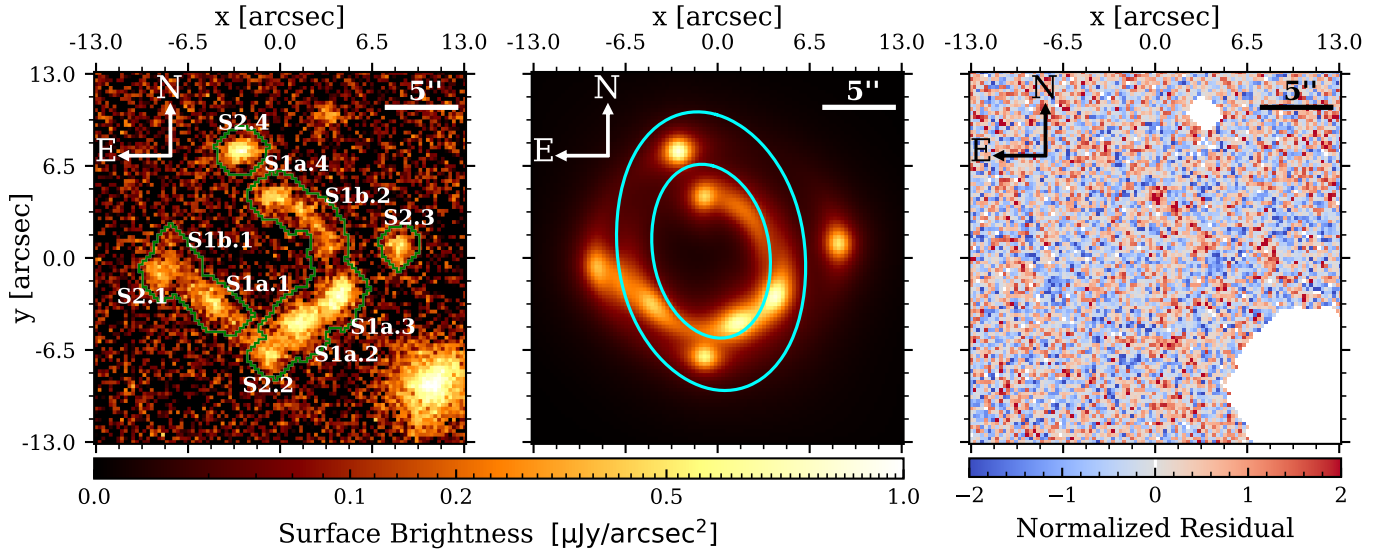


Fig. 3. Model reproduction of the observed g-band surface brightness. *Left panel:* g-band image with g1–g5 foreground-light subtracted. The green regions show the pixels included in the simulation. Image systems are labeled as S1a.1–4, S1b.1–2, and S2.1–4. *Middle panel:* model reproduction of the image based on the median of the marginalized parameters. The tangential critical lines are shown in cyan for the source plane S1 (inner) and S2 (outer). *Right panel:* normalized residual between the model and the g-band image. The remaining foreground galaxies are masked; the residual is below 2σ for most of the model.

Table 3. Lens model parameters and their uncertainties.

θ_E (arcsec)	RA	Halo Dec	ϕ (deg)	ϵ	γ_1	Shear γ_2	S2 η
$5.052^{+0.005}_{-0.005}$	2h48m10s	–39d55m48s	$12.1^{+0.3}_{-0.2}$	$0.203^{+0.002}_{-0.002}$	$0.003^{+0.001}_{-0.001}$	$-0.015^{+0.001}_{-0.001}$	$0.626^{+0.001}_{-0.001}$

Notes. The table presents median values and 1σ confidence intervals for the lens model parameters. These include the Einstein radius (θ_E), right ascension (RA), declination (Dec), position angle (ϕ), and ellipticity (ϵ); the external shear components (γ_1 and γ_2); and the distance ratio between sources S1 and S2 (η) as defined in Eq. (4). The parameters θ_E , γ_1 , and γ_2 are relative to source S1 at redshift $z = 1.2722$.

density of matter, Ω_m , and the equation of state of dark energy, w . To do this, spectroscopic observations of S2 are necessary, which may be the focus of a future study.

Furthermore, this would require a double-lens-plane model; otherwise, the single-plane model would overestimate the distance to S2 by not accounting for the additional deflection induced by S1 (see Collett & Auger 2014). In the single-plane model, the light is deflected only once, projecting the source further away than in the double-plane case, where the second deflection allows the source to be closer, while still matching the observed image separation. Thus, with our cosmological assumptions and disregarding the mass of S1a, the model estimates an upper limit for the redshift of S2 as $z_{S2} \leq 2.7$. Given that the mass of S1 is significantly lower than that of the group, the redshift of S2 is anticipated to approach this upper boundary.

The source surface brightness models for both source planes are shown in Fig. 4. There are two sources in the left panel: S1a, which has a spectroscopic redshift of $z = 1.2722$, and S1b, which the model predicts to be at a similar redshift; thus, we fixed its redshift to $z = 1.2722$ during the modeling process. The difference in color between S1a ($g - r = 0.19 \pm 0.01$) and S1b ($g - r = 0.31 \pm 0.24$) suggests they are different galaxies, although the spectra of S1b and higher-quality imaging would be needed to confirm this. The right panel shows the third source, S2, which the model predicts to be at approximately $z \sim 2.7$.

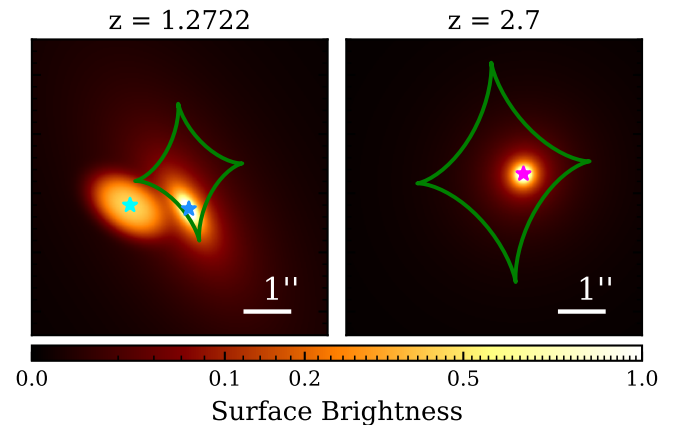


Fig. 4. Surface brightness model for the sources S1a and S1b (left) and S2 (right) in micro-Jansky per arcsecond squared. The green lines show the caustic. The blue, cyan, and magenta stars indicate the point source at the center of S1a, S1b, and S2 respectively. Source S2 is predicted to be at $z = 2.7$ by the single-lens plane model.

5.3. Impact of S1a mass on S2 images

Our algorithm currently does not support multiple lens planes, but we have explored the impact of S1 on S2 using Lenstronomy (Birrer & Amara 2018; Birrer et al. 2021). We performed a

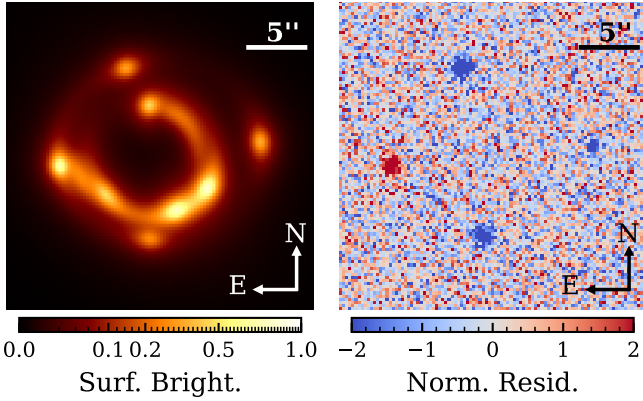


Fig. 5. Simulated impact of the mass of S1 on S2 images. *Left:* model consisting of two lens planes. The plane at z_L has the same mass model as the one shown in Fig. 3, while the plane at z_{S1} is characterized by a SIE model with a velocity dispersion of $\sigma_{v,S1} = 100 \text{ km/s}$ and the same geometry as the surface brightness profile of S1a. *Right:* normalized residual between the dual and single plane models, along with the noise. The S2 images exhibit a noticeable difference between the two models, indicating that the mass of S1a has a significant impact.

double lens plane simulation with an added SIE mass profile for S1a, mirroring its light profile shape. We calculated S1a's velocity dispersion using a mass-to-light ratio (M/L), assuming the lowest reported ratio ($M/L \sim 0.3$) by van de Sande et al. (2015). This yields a dynamical mass of about $6.4 \times 10^9 M_\odot$ for S1a, which leads to a velocity dispersion of $\sigma_v \sim 100 \text{ km s}^{-1}$ when assuming an isothermal profile.

As shown in Fig. 5, the mass of S1a changes the magnification of the S2 images and produces a small offset (about $0.12''$). This suggests a substantial lensing impact of S1a on S2 when the dispersion is at least 100 km s^{-1} , detectable with a 5σ significance. Thus, accurate modeling of S2 requires a multiplane lens approach, which we aim to incorporate in our forthcoming research.

6. Discussion

Compared to the original GIGA-Lens modeling (Gu et al. 2022), our algorithm has several advantages when dealing with complex lens models or large pixelated images. The memory requirements are greatly reduced when modeling only the region of interest. In our case, this reduces the number of simulated pixels from 10 000 to 1506 during the profile evaluation. With this number of pixels, we can simulate $\sim 10\,000$ models and their gradients simultaneously on an NVIDIA A100 GPU. Furthermore, while for this system we previously subtracted the foreground light, this step may be omitted when the images are far from the deflector light, simplifying the modeling pipeline.

We decided to use a multistep annealing algorithm instead of the standard GIGA-Lens inference pipeline. The motivation for this is the following: the GIGA-Lens convergence is highly dependent on the first step, namely: a multi-start gradient descent. If the MAP is not found in this step, the following exploration steps will be biased towards a local maximum. Typically, this is not a problem for galaxy-scale lenses: the model can be described by a smaller set of parameters and tighter priors. Gu et al. (2022) showed that 300 samples are sufficient for consistent MAP identification. For more complex models, this is not the case, as the number of needed samples increases exponentially with the number of parameters. For example, in our system,

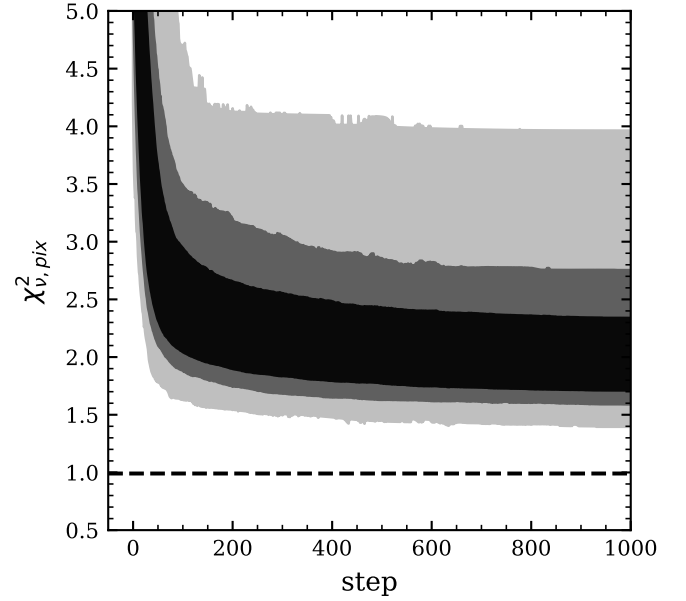


Fig. 6. Evolution of the reduced $\chi^2_{v,pix}$ along multi-start gradient descent steps. The regions account for the 1σ , 2σ , and 3σ percentiles. The dashed line shows the $\chi^2_{v,pix}$ achieved with SMC for comparison.

we were unable to obtain a consistent result with this algorithm when using 10 000 samples (the memory limit with our GPU).

Instead, with the annealing algorithm in SMC, the whole prior can be explored if λ in Eq. (10) is adjusted slowly enough, as the exploring space is reduced in a manner similar to nested sampling (Skilling 2004). We found that we can achieve consistent results in different runs when using SMC with 1000 particles. Furthermore, as demonstrated in Figure 6, most of the gradient descent samples become stuck after 100 steps and the lower $\chi^2_{v,pix}$ value is reached at 250 steps. This implies that they quickly settle into the closest local $\chi^2_{v,pix}$ minimum, which has a higher $\chi^2_{v,pix}$ than the one from SMC.

An additional advantage of our method is the use of positional information to effectively shrink the parameter space. By leveraging the image positions, we significantly reduced the prior volume. For example, this approach allows us to constrain the Einstein radius to $\theta_E = 4.9'' \pm 0.2''$, effectively decreasing the prior variance of $2.1''$ (uniform between $3''$ and $8''$) by 98% within a few seconds. Consequently, pixelated modeling can then be focused around this refined value. This targeted approach improves the efficiency of our algorithm, particularly in complex scenarios where the parameter space can be large.

Our exploration method not only offers depth but is also fast and efficient. The sampling takes 4.5 minutes, which is comparable to the GIGA-Lens time for a galaxy-scale lens of about 6 minutes using a much stricter prior and half the number of parameters (Gu et al. 2022; Cikota et al. 2023). As shown by Gu et al. (2022), this time may be reduced by a factor of ~ 3.5 when using four A100 GPUs instead of only one.

In this paper, we present the modeling of a group-scale lens, but larger-scale systems can be modeled using scaling relations for group and cluster members. As a result, the number of parameters would not depend on the number of deflectors, which is commonly assumed for cluster lenses (Brainerd et al. 1996; Limousin et al. 2005; Jullo et al. 2007). We intend to incorporate this feature in the near future.

7. Conclusions

We present a novel lens modeling software, which stands as an improvement over the original GIGA-Lens for simulating complex systems. This enhanced version employs an improved two-step inference method, using image positions and efficient SMC sampling for comprehensive parameter space exploration, even with limited prior knowledge. We verified our method by modeling the DES J0248-3955 lens system. The primary findings related to this system are as given below:

1. Based on VLT X-shooter spectra, we determined the redshift for the group to be $z_L = 0.69 \pm 0.04$. This value is derived from the spectroscopic redshift of the BGG and the photometric redshift estimates of four additional group members. For one of the sources, we measured a redshift of $z_{S1} = 1.2722 \pm 0.0005$;
2. The observed multiple-image configuration was successfully reproduced using a singular isothermal ellipsoid profile and external shear;
3. With our modeling technique, we obtained the posterior for the model with 29 parameters within a few minutes, with even faster processing possible for simpler models or using multiple GPUs;
4. This system is consistent with being a double source plane lens; however, to fully describe the second source, it is necessary to take into account the mass effect of the first source;
5. Using the lens model, we measured a velocity dispersion of $\sigma_v = (690 \pm 30) \text{ km s}^{-1}$ for the galaxy group.

For future studies, this lens system holds the potential to constrain the nature of dark energy, which is pending further observational data as well as the formulation of a suitable double lens plane model. In conclusion, our software addresses the challenge of modeling strong lenses at group scales using ground-based data by enabling fast modeling and providing strong model constraints, even with low-resolution imaging and a broad prior, on the lens parameters.

Our software not only improves the modeling of individual systems, such as DES J0248-3955, but also holds significant potential for broader applications in future large-scale surveys, including LSST. These investigations are poised to discover numerous such systems, although follow-up observations for all of them will not be feasible. The use of a broad prior enables the modeling of a diverse range of systems, making it suitable for automated applications. We plan a future study focused on exploring the scalability and automation potential of our software, further establishing its utility for high-throughput lens modeling.

Acknowledgements. Based on data obtained from the ESO Science Archive Facility with DOI(s): <https://doi.eso.org/10.18727/archive/71>. The authors acknowledge funding from ANID through BECAS/MAGISTER NACIONAL 22232142, ANID BASAL project FB210003, and FONDECYT project 23050313. We also acknowledge funding from the Centre National de Recherche Scientifique (CNRS) and the Franco-Chilean Laboratory of Astronomy (FCLA), IRL 3386. The DESI Legacy Imaging Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS), the Beijing-Arizona Sky Survey (BASS), and the Mayall z-band Legacy Survey (MzLS). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. Pipeline processing and analyses of the data were supported by

NOIRLab and the Lawrence Berkeley National Laboratory (LBNL). Legacy Surveys also uses data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. Legacy Surveys was supported by: the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility; the U.S. National Science Foundation, Division of Astronomical Sciences; the National Astronomical Observatories of China, the Chinese Academy of Sciences and the Chinese National Natural Science Foundation. LBNL is managed by the Regents of the University of California under contract to the U.S. Department of Energy. The complete acknowledgments can be found at <https://www.legacysurvey.org/acknowledgment/>. The Photometric Redshifts for the Legacy Surveys (PRLS) catalog used in this paper was produced thanks to funding from the U.S. Department of Energy Office of Science, Office of High Energy Physics via grant DE-SC0007914.

References

- Acebron, A., Schuldt, S., Grillo, C., et al. 2023, *A&A*, **680**, L9
- Adam, A., Perreault-Levasseur, L., Hezaveh, Y., & Welling, M. 2023, *ApJ*, **951**, 6
- Atek, H., Richard, J., Jauzac, M., et al. 2015, *ApJ*, **814**, 69
- Auger, M. W., Treu, T., Bolton, A. S., et al. 2010, *ApJ*, **724**, 511
- Bergamini, P., Acebron, A., Grillo, C., et al. 2023, *ApJ*, **952**, 84
- Biggio, L., Vernardos, G., Galan, A., & Peel, A. 2022, *A&A*, **675**, A125
- Birrer, S., & Amara, A. 2018, *PDU*, **22**, 189
- Birrer, S., Amara, A., & Refregier, A. 2015, *ApJ*, **813**, 102
- Birrer, S., Shajib, A., Gilman, D., et al. 2021, *JOSS*, **6**, 3283
- Bradbury, J., Frostig, R., Hawkins, P., et al. 2018, *JAX: composable transformations of Python+NumPy programs*
- Bradley, L., Sipőcz, B., Robitaille, T., et al. 2023, <https://doi.org/10.5281/zenodo.7946442>
- Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, *ApJ*, **466**, 623
- Caminha, G. B., Suyu, S. H., Grillo, C., & Rosati, P. 2022, *A&A*, **657**, A83
- Cerny, C., Sharon, K., Andrade-Santos, F., et al. 2018, *ApJ*, **859**, 159
- Chianese, M., Coogan, A., Hofma, P., Otten, S., & Weniger, C. 2020, *MNRAS*, **496**, 381
- Cikota, A., Bertolla, I. T., Huang, X., et al. 2023, *ApJ*, **953**, L5
- Collett, T. E. 2015, *ApJ*, **811**, 20
- Collett, T. E., & Auger, M. W. 2014, *MNRAS*, **443**, 969
- Collett, T. E., Auger, M. W., Belokurov, V., Marshall, P. J., & Hall, A. C. 2012, *MNRAS*, **424**, 2864
- Dalal, N., & Kochanek, C. S. 2002, *ApJ*, **572**, 25
- Del Moral, P., Doucet, A., & Jasra, A. 2012, *Statist. Comput.*, **22**, 1009
- Dey, A., Schlegel, D. J., Lang, D., et al. 2019, *AJ*, **157**, 168
- Dillon, J. V., Langmore, I., Tran, D., et al. 2017, arXiv e-prints [arXiv:1711.10604]
- Gu, A., Huang, X., Sheu, W., et al. 2022, *ApJ*, **935**, 17
- Harvey, D., Courbin, F., Kneib, J. P., & McCarthy, I. G. 2017, *MNRAS*, **472**, 1972
- Hezaveh, Y. D., Levasseur, L. P., & Marshall, P. J. 2017, *Nature*, **548**, 555
- Horne, K. 1986, *PASP*, **98**, 609
- Huang, X., Storfer, C., Gu, A., et al. 2021, *ApJ*, **909**, 27
- Jacobs, C., Collett, T., Glazebrook, K., et al. 2019, *ApJS*, **243**, 17
- Jauzac, M., Eckert, D., Schwinn, J., et al. 2016, *MNRAS*, **463**, 3876
- Jullo, E., Kneib, J.-P., Limousin, M., et al. 2007, *New J. Phys.*, **9**, 447
- Jullo, E., Natarajan, P., Kneib, J.-P., et al. 2010, *Science*, **329**, 924
- Kelly, P. L., Rodney, S., Treu, T., et al. 2023, *Science*, **380**, abh1322
- Knabel, S., Holwerda, B. W., Nightingale, J., et al. 2023, *MNRAS*, **520**, 804
- Kochanek, C. S. 1991, *ApJ*, **373**, 354
- Koopmans, L. V. E., Bolton, A., Treu, T., et al. 2009, *ApJ*, **703**, L51
- Limousin, M., Kneib, J. P., & Natarajan, P. 2005, *MNRAS*, **356**, 309
- Limousin, M., Ebeling, H., Richard, J., et al. 2012, *A&A*, **544**, A71
- Limousin, M., Beauchesne, B., & Jullo, E. 2022, *A&A*, **664**, A90
- Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, *ApJ*, **837**, 97
- Magro, D., Zarb Adam, K., DeMarco, A., Riggi, S., & Sciacca, E. 2021, *MNRAS*, **505**, 6155
- Mahler, G., Jauzac, M., Richard, J., et al. 2023, *ApJ*, **945**, 49
- Mellier, Y., Fort, B., & Kneib, J.-P. 1993, *ApJ*, **407**, 33
- Meneghetti, M., Argazzi, R., Pace, F., et al. 2007, *A&A*, **461**, 25
- Merten, J., Coe, D., Dupke, R., et al. 2011, *MNRAS*, **417**, 333
- Morningstar, W. R., Levasseur, L. P., Hezaveh, Y. D., et al. 2019, *ApJ*, **883**, 14
- Newman, A. B., Treu, T., Ellis, R. S., & Sand, D. J. 2013, *ApJ*, **765**, 25
- Nightingale, J. W., Dye, S., & Massey, R. J. 2018, *MNRAS*, **478**, 4738
- O'Donnell, J. H., Wilkinson, R. D., Diehl, H. T., et al. 2022, *ApJS*, **259**, 27

- Oguri, M. 2006, [MNRAS](#), **367**, 1241
- Pascale, M., Frye, B. L., Pierel, J. D. R., et al. 2024, *ApJ*, accepted, [arXiv:[2403.18902](#)]
- Pearson, J., Li, N., & Dye, S. 2019, [MNRAS](#), **488**, 991
- Pearson, J., Maresca, J., Li, N., & Dye, S. 2021, [MNRAS](#), **505**, 4362
- Refsdal, S. 1964, [MNRAS](#), **128**, 307
- Roche, C., McDonald, M., Borrow, J., et al. 2024, [Open J. Astrophys.](#), **7**, 65
- Rojas, K., Savary, E., Clément, B., et al. 2022, [A&A](#), **668**, A73
- Schneider, D. P., Turner, E. L., Gunn, J. E., et al. 1988, [AJ](#), **95**, 1619
- Schneider, P., Ehlers, J., & Falco, E. E. 1992, [Gravitational Lenses](#), Astronomy and Astrophysics Library (Berlin, Heidelberg: Springer)
- Schuldt, S., Suyu, S. H., Meinhardt, T., et al. 2021, [A&A](#), **646**, A126
- Schuldt, S., Cañameras, R., Shu, Y., et al. 2023, [A&A](#), **671**, A147
- Sharma, D., Collett, T. E., & Linder, E. V. 2023, [JCAP](#), **2023**, 001
- Sharon, K., Bayliss, M. B., Dahle, H., et al. 2020, [ApJS](#), **247**, 12
- Shu, Y., Cañameras, R., Schuldt, S., et al. 2022, [A&A](#), **662**, A4
- Skilling, J. 2004, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, eds. R. Fischer, R. Preuss, & U. V. Toussaint, [American Institute of Physics Conference Series](#), **735**, 395
- Stapelberg, S., Carrasco, M., & Maturi, M. 2019, [MNRAS](#), **482**, 1824
- Storfer, C., Huang, X., Gu, A., et al. 2024, [ApJS](#), **274**, 16
- van de Sande, J., Kriek, M., Franx, M., Bezanson, R., & van Dokkum, P. G. 2015, [ApJ](#), **799**, 125
- Vegetti, S., Birrer, S., Despali, G., et al. 2024, [Space Sci. Rev.](#), **220**, 58
- Vernet, J., Dekker, H., D’Odorico, S., et al. 2011, [A&A](#), **536**, A105
- Wong, K. C., Suyu, S. H., Chen, G. C. F., et al. 2020, [MNRAS](#), **498**, 1420
- Zaborowski, E. A., Drlica-Wagner, A., Ashmead, F., et al. 2023, [ApJ](#), **954**, 68
- Zhou, R., Newman, J. A., Mao, Y.-Y., et al. 2021, [MNRAS](#), **501**, 3309
- Zitrin, A., Broadhurst, T., Bartelmann, M., et al. 2012, [MNRAS](#), **423**, 2308