

# *AutoLens*: automated modeling of a strong lens's light, mass, and source

J. W. Nightingale,<sup>1,2★</sup> S. Dye<sup>2</sup> and Richard J. Massey<sup>1</sup>

<sup>1</sup>Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK

<sup>2</sup>School of Physics and Astronomy, Nottingham University, University Park, Nottingham, NG7 2RD, UK

Accepted 2018 May 3. Received 2017 August 24; in original form 2018 May 3

## ABSTRACT

This work presents *AutoLens*, the first entirely automated modeling suite for the analysis of galaxy-scale strong gravitational lenses. *AutoLens* simultaneously models the lens galaxy's light and mass whilst reconstructing the extended source galaxy on an adaptive pixel-grid. The method's approach to source-plane discretization is amorphous, adapting its clustering and regularization to the intrinsic properties of the lensed source. The lens's light is fitted using a superposition of Sersic functions, allowing *AutoLens* to cleanly deblend its light from the source. Single-component mass models representing the lens's total mass density profile are demonstrated, which in conjunction with light modeling can detect central images using a centrally cored profile. Decomposed mass modeling is also shown, which can fully decouple a lens's light and dark matter and determine whether the two components are geometrically aligned. The complexity of the light and mass models is automatically chosen via Bayesian model comparison. These steps form *AutoLens*'s automated analysis pipeline, such that all results in this work are generated without any user intervention. This is rigorously tested on a large suite of simulated images, assessing its performance on a broad range of lens profiles, source morphologies, and lensing geometries. The method's performance is excellent, with accurate light, mass, and source profiles inferred for data sets representative of both existing Hubble imaging and future Euclid wide-field observations.

**Key words:** gravitational lensing – galaxy: structure – methods: data analysis.

## 1 INTRODUCTION

Strong gravitational lensing offers a unique means of measuring the mass distribution and composition of galaxies within our Universe. Through the intricate analysis of a lensed source's extended light profile, one can robustly infer the lens galaxy's density profile, a technique that has been exploited to provide observations in the fields of dark matter substructure (Vegetti & Koopmans 2009a, b; Vegetti et al. 2012, 2014), stellar dynamics (Barnabè & Koopmans 2007; Barnabè et al. 2009, 2011), and cosmology (Suyu et al. 2013; Collett & Auger 2014; Suyu et al. 2016; Wong et al. 2017). Equally, this analysis provides a full reconstruction of the highly magnified source galaxy and therefore offers an unprecedented view of the high redshift Universe (Shirazi et al. 2014; Dye et al. 2014, 2015; Rybak et al. 2015; Swinbank et al. 2015).

However, unlike the works above, the majority of strong lensing studies exploit just one lensing observable, the Einstein Mass,  $M_{\text{Ein}}$ , which is widely accepted as a robust mass estimator that is essentially independent of the density profile assumed for the lens.  $M_{\text{Ein}}$  is constrained by the first derivative of the lens's potential, therefore it is the *position* of the lensed source in the image-plane

that is key. Measuring  $M_{\text{Ein}}$  therefore requires relatively simple lens modeling methodology (e.g. Bolton et al. 2008; Sonnenfeld et al. 2013a) and has already been performed on the majority of known strong lenses over the past decade (Bolton et al. 2008; Koopmans et al. 2009; Bolton et al. 2012; Brewer et al. 2012; Dutton et al. 2013; Sonnenfeld et al. 2013b, 2015).

When an extended source is lensed, light rays emanating from different regions of the source trace through different regions of the lens galaxy. Therefore, the lensed source's *extended surface-brightness profile* contains a wealth of additional information, that if exploited can be used to measure the lens potential's second derivative, its density profile. Extracting this signal requires more sophisticated lens modeling capable of both reconstructing the source's intrinsic light profile and modeling the lens galaxy's mass distribution (Warren & Dye 2003; Dye & Warren 2005; Suyu et al. 2006; Vegetti & Koopmans 2009a; Tagore & Keeton 2014; Birrer, Amara & Refregier 2015a; Tessore, Bellagamba & Metcalf 2016). Unfortunately, the involved nature of extended source modeling has seen it struggle to scale up to large samples, with most analyses focusing on samples of just 1–10 objects (e.g. Dye et al. 2014; Vegetti et al. 2014; Birrer, Amara & Refregier 2016; Suyu et al. 2016; Dye et al. 2017).

\* E-mail: james.w.nightingale@durham.ac.uk

The aim of this paper is to rectify this, by demonstrating a fully automated approach to extended source modeling that performs all analysis and generates all results without any user intervention after a brief initial set-up. This is well motivated, given archival lens data sets have several hundred *HST*-quality images warranting such an analysis (Bolton et al. 2006; Auger et al. 2010; Sonnenfeld et al. 2013a). Furthermore, with ongoing and future surveys such as the Large Synoptic Survey Telescope and Euclid set to find of order one hundred thousand strong lenses (Oguri et al. 2012; Collett 2015), an automated pipeline is paramount to fully exploit the expansive incoming data sets.

This paper builds upon the adaptive semilinear inversion method developed by Nightingale & Dye (2015, N15 hereafter) with a new comprehensive and automated modeling process that we have named AutoLens. We test this on an extensive suite of simulated imaging, which is paramount given we are in a regime where the detailed inspection of results on an individual case-by-case basis is not feasible. This also includes data representative of Euclid imaging, thus giving first insights into the type of modeling and observations that may (and may not) be possible with direct analysis of wide-field imaging.

This work is performed using a Fortran version of AutoLens. A project is now under way to redevelop AutoLens in and make it publicly available and open-source software for the community. The latest status of this project can be found at <https://github.com/Jammy2211/PyAutoLens>.

Table D2 at the end of the script lists parameters and symbols used in this work.

This paper is structured as follows. Section 2 gives an overview of AutoLens's key features. Section 3 describes the light profiles, mass profiles, and simulated images used to test AutoLens. Section 4 presents in detail AutoLens's lens and source analysis, including the method's adaptive source analysis and variance scaling, which Section 5 demonstrates. Section 6 describes the method's development into an automated analysis pipeline. Section 7 demonstrates this on the simulated image suite and Section 8 discusses the results and summarizes the paper.

## 2 OVERVIEW OF AUTOLENS

AutoLens is described fully in Section 4. Here, an overview of the method's key features is given.

AutoLens brings about a number of improvements over the source analysis of N15, who demonstrated the use of an adaptive pixel-grid to reconstruct the source galaxy. This computed a unique source pixelization in a completely stochastic manner for every lens model, a feature that was key to removing previously unknown systematics associated with the discrete nature of source reconstruction. AutoLens now adapts its pixel-grid and regularization scheme to the morphology of the lensed source galaxy, in a manner that significantly improves lens modeling within the Bayesian framework of Suyu et al. (2006). This ensures that the method can handle the diverse range of strongly lensed sources that are in existing lens samples (e.g. Newton et al. 2011; Dye et al. 2015; Shu et al. 2016; Oldham et al. 2017; Enia et al. 2018).

AutoLens now fits the lens galaxy's light, thereby unifying the modeling of the lens's mass and light and the reconstruction of the source galaxy into one coherent framework. This is in contrast to other methods in the literature, which typically subtract the lens galaxy's light before performing lens modeling (e.g. Bolton et al. 2006), therefore discarding the information that it contains. Light profile fitting with AutoLens supports both single- and multi-

component models and the complexity of the light model is chosen within the framework of Bayesian model comparison, ensuring that an appropriate light profile is fitted for lenses of different morphological classes. This achieves a clean separation of the lens and source light (which is not possible when modeled independently; Marshall et al. 2007; Biernaux et al. 2016) and measures the lens galaxy's light profile, a quantity routinely measured to study the structure of large samples of galaxies (Hoyos et al. 2011; Vika et al. 2013; Bluck et al. 2014; Bruce et al. 2014a,b; Nikutta et al. 2014; Vika et al. 2014; Vulcani et al. 2014).

Three approaches to mass modeling are demonstrated, the first invoking the same mass model as N15, a power-law density profile representing the lens's total mass distribution. This is the model assumed in most strong lensing works (e.g. Dye et al. 2014; Vegetti et al. 2014) and has been fitted to over one hundred Early-Type Galaxy (ETG) lenses from surveys such as the Sloan Lens ACS Survey (SLACS) (Bolton et al. 2006; Auger et al. 2010), the Strong Lensing in the Legacy Survey (SL2S) (Sonnenfeld et al. 2013a) and the BOSS Emission-Line Lens Survey (BELLS) (Brownstein et al. 2012). This has revealed that the inner mass distribution of ETGs are accurately approximated by a nearly isothermal density profile (Gavazzi et al. 2007; Koopmans et al. 2009; Barnabè et al. 2011; Bolton et al. 2012; Sonnenfeld et al. 2013b; Sonnenfeld et al. 2015) and has been termed the 'bulge-halo conspiracy', given that neither the light matter component, the bulge, nor the dark component, the halo, has this profile and yet their combination conspires to produce one. However, these large samples measure only  $M_{\text{Ein}}$  from the lensing data, using it as an additional constraint on stellar dynamical modeling. The extended source modeling used in this work can infer the lens's density profile without any kinematic data, offering a complementary measurement to these previous studies. AutoLens also models the contribution of large-scale structure via an external shear term, the inclusion of which is subject to Bayesian model comparison given that incorrectly assuming a shear can potentially bias the lens model inferred (Balmès & Corasani 2013).

By simultaneously fitting and subtracting the lens's light, AutoLens can potentially reveal faint features in a lensed source that previous approaches to lens analysis may have missed, due to, for example, falsely oversubtracting these features before lens modeling or masking them in the subsequent lens analysis. Therefore, AutoLens's second approach to mass modeling attempts to detect a source's central features, like a third or fifth image (Mao, Witt & Koopmans 2001; Rusin & Ma 2001; Keeton 2003) or radial arcs, by invoking the cored power-law density profile. AutoLens is able to detect these features provided they are present in the image and sufficiently extended and again uses Bayesian model comparison to ascertain whether such features are residuals resulting from the subtraction of an overly simplistic lens light profile or whether they are genuine lensed image components. AutoLens thus brings a new capability of searching for central images at optical and UV wavelengths to compete with existing efforts in the submm and radio where the lens light is typically not detected (Winn, Rusin & Kochanek 2004; Hezaveh, Marshall & Blandford 2015; Quinn et al. 2016). A promising aspect of searching at shorter wavelengths is the possibility that sources possess a flatter and more extended light profile, lessening the central image's demagnification.

Fitting of the lens's light plays another crucial role, allowing AutoLens to advocate decomposed mass profiles that separately treat the lens's light and dark matter. The final approach to mass modeling thus incorporates the light profile into the mass model, exploiting the fact that by tracing the lens's underlying stellar mass distribution it offers additional information about approximately half of

the lens’s overall density profile, information that other methods omit. In doing so, `AutoLens` is able to make unique measurements about both components, for example the light component’s mass-to-light ratio (independent of stellar population synthesis) and the dark matter’s ellipticity, as well as comparing how the two are distributed relative to one another. `AutoLens` makes no assumptions about the geometric alignment of the light and dark matter, determining whether there is a positional and/or rotational offset between the components via Bayesian Model comparison, allowing `AutoLens` to offer a first observational insight into the geometry of light and dark matter (Navarro & Benz 1991; Bett et al. 2010; Piontek & Steinmetz 2011; Sales et al. 2012; Schaller et al. 2015; Velliscig et al. 2015; Liao et al. 2017).

`AutoLens` is tested on a large suite of simulated images chosen to rigorously test `AutoLens` on a range of lens light and mass profiles, source morphologies, and strong lens geometries representative of forthcoming large lens samples. Each simulated lens has two images generated, one at the resolution and signal-to-noise level of currently available *HST* imaging data and one where these properties are in line with what can be anticipated from Euclid imaging. Finally, the reader should note that, although not shown in this paper, `AutoLens`’s development was performed in conjunction with testing on images of real strong lenses and some of the method’s design choices reflect circumstances not tested by the simulated data set. These design choices are discussed in this work whenever relevant.

### 3 SIMULATED DATA

#### 3.1 Light profiles

Light profiles for the lens galaxy are computed using Sersic functions, which have elliptical coordinates  $\xi_1 = \sqrt{x_1^2 + y_1^2/q_1^2}$ , such that the intensity at a given coordinate is given by

$$I_{\text{Ser}}(\xi_1) = I_1 \exp \left\{ -k_1 \left[ \left( \frac{\xi_1}{R_1} \right)^{n_1} - 1 \right] \right\}, \quad (1)$$

which has seven parameters:  $(x_1, y_1)$ , the light centre,  $q_1$ , the axis ratio,  $\theta_1$ , the orientation angle (defined counter-clockwise from the positive x-axis),  $I_1$ , the intensity at the effective radius  $R_1$  and  $n_1$ , the Sersic index.  $k_1$  is a function of  $n_1$ . In general, a subscript ‘l’ signifies that a parameter belongs to the light model. The de Vaucouleurs light profile  $I_{\text{Dev}}(\xi_1)$  corresponds to  $n_1 = 4$  and the exponential light profile  $I_{\text{Exp}}(\xi_1)$  corresponds to  $n_1 = 1$ , respectively. The resulting 2-D light profile is then convolved with the instrumental point spread function (PSF).

Composite light models are calculated by summing individual component intensity maps. When multiple light components are used, each component’s parameters are labeled with an additional numeric subscript (e.g.  $n_{11}, n_{12}$ , etc.). All multicomponent light models are assumed to share the same center and rotation angle.

An adaptive oversampling routine is used to ensure that light profiles are computed in both an accurate and efficient manner. This is described in Appendix A.

#### 3.2 Mass profiles

To generate a lensed source, each image pixel must be traced from the image-plane to the source-plane via the lens equation. This is performed using the deflection angles computed from the lens convergence profile (see below),  $\kappa(\xi)$ . Like the light profile,

this is a function of the elliptical radius  $\xi = \sqrt{x^2 + y^2/q^2}$ . This has center  $(x, y)$ , projected axis ratio  $q$  and is rotated by an angle  $\theta$  defined counter-clockwise from the positive x-axis. In general, all parameters associated with the lens’s total mass profile have no subscript, whereas those associated with a dark matter component have subscript ‘d’ and a light matter component a subscript ‘l’.

**N15** used a Singular Power-Law Ellipsoid (*SPLE*) lens model with volume mass density profile of the form  $\rho(r) = \rho_o(r/r_o)^{-\alpha} \cdot \alpha_{x,y}$  was computed following **K01**, where the lens mass normalization was given by the equivalent velocity dispersion  $\sigma$ . However, this parameterization is uncommon within the literature and to ease future comparison the formalism of Suyu (2012) is used hereafter, where the elliptical power-law surface density is given by

$$\kappa_{\text{pl}}(\xi) = \frac{(3 - \alpha)}{1 + q} \left( \frac{\theta_E}{\xi + S^2} \right)^{\alpha-1}. \quad (2)$$

Here  $\theta_E$  is the model Einstein radius in arc seconds. The core radius is given by  $S$ , which is set to zero for singular power-law models. Profiles that include a core are referred to as ‘*PL<sub>Core</sub>*’. The factors  $(3 - \alpha)$  and  $1 + q$  rescale  $\theta_E$  to give the same mass normalization for a changing density slope  $\alpha$  or axis ratio  $q$ . The potential and deflection angles are computed from equation (2) using the method of Barkana (1998). The case  $\alpha = 2$  again corresponds to the Singular Isothermal Ellipsoid (SIE) lens profile. Although parameterized differently, the  $\kappa_{\text{pl}}(\xi)$  profile given in (2) and used in **N15** are identical, therefore the  $\sigma$ - $q$ - $\alpha$  degeneracy described in **N15** is again present, however now between  $\theta_{\text{Ein}}$ ,  $q$  and  $\alpha$  (although the rescalings by  $(3 - \alpha)$  and  $1 + q$  make these degeneracies appear more orthogonal).

The inclusion of an external shear field is supported, which introduces two additional parameters with subscript ‘sh’, the shear strength  $\gamma_{\text{sh}}$ , and orientation of the semi-major axis measured counter-clockwise from east,  $\theta_{\text{sh}}$ . To numerically compute its deflection angles, this shear requires a center, which assumes either the mass profile’s center  $(x$  and  $y)$  or dark matter’s center  $(x_d$  and  $y_d)$  for a decomposed model.

Decomposed mass profiles assume separate density profiles for the light and dark matter components. The light component uses the elliptical Sersic profile given by equation (3), converting it to a mass profile as

$$\kappa_{\text{Ser}}(\xi_1) = \Psi_1 I_{\text{Ser}}(\xi_1), \quad (3)$$

therefore sharing the same parameters as the light profile, but with an additional parameter  $\Psi_1$ , the mass-to-light ratio. Simulated lenses with multiple light components are generated using the same  $\Psi_1$  for each component.

The dark matter component is given by an elliptical Navarro–Frenk–White (*NFW*) profile, which represents the universal density profile predicted for dark matter halos by cosmological  $N$ -body simulations (Navarro, Frenk & White 1996a, b; Zhao 1996). This has volume mass density given by

$$\rho = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2}, \quad (4)$$

where  $\rho_s$  gives the halo normalization and  $r_s$  the scale radius, which is fixed throughout this work to the value  $r_s = 30$  kpc (Bullock et al. 2001). Coordinates for the *NFW* profile are scaled by  $r_s$ , giving the scaled elliptical coordinate  $\eta_d = \xi_d/r_s$ .

Analytic solutions for the *NFW* model are given in (Golse & Kneib 2002) and are given by

$$\kappa_{\text{NFW}}(\eta_d) = 2\kappa_d \frac{1 - \mathcal{F}(\eta_d)}{\eta_d^2 - 1}, \quad (5)$$

where

$$\mathcal{F}(\eta_d) = \begin{cases} \frac{1}{\sqrt{\eta_d^2 - 1}} \arctan \sqrt{\eta_d^2 - 1} & : \eta_d > 0 \\ \frac{1}{\sqrt{1 - \eta_d^2}} \operatorname{arctanh} \sqrt{1 - \eta_d^2} & : \eta_d < 0 \\ 1 & : \eta_d = 1 \end{cases} \quad (6)$$

where  $\kappa_d$  is related to the lens halo normalization by  $\kappa_d = \rho_s r_s / \Sigma_{\text{cr}}$  and  $\Sigma_{\text{cr}}$  is the critical surface density. A spherical NFW profile may also be used, which removes the axis ratio  $q_d$  and rotation angle  $\theta_d$  as free parameters and has the title *NFWSpH*.

Unlike the *PL<sub>Core</sub>* profile, the light and dark matter profiles above do not have a prescription to include a centrally cored profile. Therefore, the decomposed mass profiles used in this work are not equipped to produce the type of source features indicative of a cored central density, like a central image or radial arc. Cored models for both the light (e.g. cored-Sersic Dullo & Graham 2013, 2014 or Nuker Faber et al. 1997 profiles) and dark matter (e.g. a generalized NFW profile; Zhao 1996) will be considered if and when objects with central source features are detected first using the *PL<sub>Core</sub>* profile.

The deflection of light at the center of each image pixel is computed by integrating the lens's convergence profile  $\kappa$  using the equation

$$\alpha_{x,y}(\mathbf{x}) = \frac{1}{\pi} \int \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^2} \kappa(\mathbf{x}') d\mathbf{x}', \quad (7)$$

where  $\mathbf{x}$  is the image-plane coordinate. The method of Keeton (2001, K01 hereafter) is followed to compute the 2-D deflection angle map  $\alpha_{x,y}$  from equation (7). An adaptive numerical integrator following the method of K01 is used to compute deflection angles and is described in Appendix A.

### 3.3 Source profiles

The intrinsic source surface brightness profiles used in this work follow one or a summation of several elliptical Sersic functions using the elliptical radius  $\xi_s = \sqrt{x_s^2 + y_s^2/q_s^2}$ . All parameters associated with the source have the subscript 's'. The lensed image of the source is convolved with the instrumental PSF.

### 3.4 Simulation suite

AutoLens is tested using a suite of 54 simulated images that are generated using the light, mass, and source profiles described above. Such an extensive library of images is necessary to explore the diverse range of lensing geometries, mass profiles, and lens and source morphologies that are possible in any strong lens sample, as well as for ensuring that AutoLens's Bayesian model comparison features (see Section 6) correctly choose the lens model complexity. The images have been chosen to span a broad range of image resolutions, signal-to-noise ratios, and source morphologies, the three key attributes in determining how accurately a lens model can be constrained for a given image (Lagattuta et al. 2012; Vegetti et al. 2014). The highest quality images simulated in this work are of a comparatively lower S/N and resolution than currently available (e.g. SLACS, SL2S) strong lens observations, thus the precision of the results may be viewed as conservative.

Images are representative of either *Hubble Space Telescope* (*HST*) strong lens imaging or that which is anticipated from Euclid optical imaging. *HST* simulated images are generated with a pixel scale of 0.06 arcmin, are convolved with a circularly symmetric Gaussian PSF of size 0.085 arcmin, include read noise of 4 e-

a flat background sky of 1500 e- and Poisson noise (including the background sky). Euclid images have a pixel scale 0.1 arcmin, a PSF of size 0.125 arcmin, read noise of 4 e-, a background sky of 300 e-, and Poisson noise. Details of how sky subtraction is performed are given in the next section. The signal-to-noise (S/N) ratio of the lensed source component in each image is summarized using the S/N value of its brightest pixel, which is located by scanning the model lensed source during the image simulation (after the PSF convolution step but before noise is added). The S/N of the lens's light profile is computed in an analogous way using its model light profile. In line with current lens data sets, these S/N values range between 10 and 50 for the source and between 40 and 80 for the lens. By using just the brightest pixel in each component, more concentrated source and lens profiles have fewer high S/N pixels compared to flatter profiles.

Lens profiles are chosen using parameters consistent with fiducial redshifts of  $z_{\text{lens}} = 0.5$  for the lens and  $z_{\text{src}} = 1.0$  for the source. The lens light profile is generated using the adaptive oversampling routine to a fractional accuracy of  $10^{-7}$ . Each lensed source pixel is computed with oversampling of degree  $20 \times 20$ . Images are representative of a single exposure, therefore omitting dithered observing strategies and multidrizzling of images to a common frame. Thus, effects such as correlated noise are not present in the simulated data set and are not considered in this work.

To generate the 54 simulated images, 19 different combinations of light, mass, and source profiles are used, which are shown in Tables 1 and 2. These lens and source models have been chosen to test specific aspects of AutoLens. Their model names reflect these chosen aspects (e.g. the **LensMassShear** model tests mass modeling with an external shear). Each model is then used to generate multiple images at different image resolutions and S/N ratios, with Tables 1 and 2 also listing the images generated from each model. An image is then referred to by its model name and a subscript describing the image properties as follows; an 'H' or 'E' for *HST* or Euclid resolution, 'S##' for the source S/N, 'L##' for the lens S/N and either 'Disk', 'Bulge', 'BD' (Bulge-Disk), 'Cusp' or 'Multi' (multiple sources) to describe the source morphology. The suite of images can be characterized in more detail as follows:

(i) 24 images generated without a lens light component ('NL' in tag replacing 'L##'), aimed primarily at testing AutoLens's source analysis. Of these 24 images, 4 unique *SPLE* lens models are used. Therefore, each lens model comprises six images with a S/N of 50, 30, 10 and representing either *HST* of Euclid imaging. Example image names are **SrcBD**<sub>HS30NLBD</sub> (tests modeling a source with a bulge-disc morphology) or **SrcDisk**<sub>ES10NLDisk</sub> (tests a source with only a disc).

(ii) 14 images generated with a lens light component and an *SPLE* mass profile, aimed at testing lens light modeling for different lens and source morphologies. Seven unique lens models are used, using sources with S/N ranging from 20 to 30, lenses with S/N ranging from 40 to 80 and with every model generating a Hubble and Euclid resolution image. Included in this set are images that test a shear component in the mass model, multicomponent light profiles and how light modeling fares with either cuspy or flat source morphologies. Example image names are **LensSrcBulge**<sub>HS25L50BD</sub> (tests light modeling with a bulge-like source) and **LensMassShear**<sub>ES50L80Flat</sub> (tests shear modeling).

(iii) Six images generated with a lens light component and a *PL<sub>Core</sub>* mass profile, aimed at testing cored-mass modeling and central image detection. Three unique lens models are used, using sources with S/N ranging from 20 to 30 and lenses ranging from 40

**Table 1.** The lens light, mass, and source profiles used to create each image of the simulation suite generated using a total-mass profile (e.g. an *SPLE* or *PL<sub>Core</sub>*). The left-hand column of the table gives the title of each model, the name of which signifies the aspect of AutoLens that model has been made under test (e.g. the **SrcBulge** model tests modeling sources with a bulge morphology). Each model is used to generate multiple images and in the first column underneath each model name are the tags describing those images, which can be read as follows: ‘H’ or ‘E’ for Hubble or Euclid resolution, ‘S##’ the source S/N, ‘L##’ the lens S/N (NL for no lens), and ‘Bulge’, ‘Disk’, ‘BD’, ‘Cusp’ or ‘Multi’ to describe the source morphology. The second and third columns list each component and its corresponding profile. The remaining columns show the input lens parameters of each lens model.

Model Title	Component	Model	Parameters							
SrcBulge	Light	None								
HS50NLBulge, HS30NLBulge	Mass	<i>SPLE</i>	$x = 0.00''$	$y = 0.00''$	$\theta = 127^\circ$	$\theta_E = 1.2$	$q = 0.8$	$\alpha = 2.0$		
HS10NLBulge, ES50NLBulge	Source	<i>Sersic</i>	$x_s = 0.06''$	$y_s = -0.03''$	$\theta_s = 90^\circ$	$I_s = 0.015$	$R_s = 0.4''$	$n_s = 2.5$	$q_s = 0.7$	
ES30NLBulge, ES10NLBulge	Source	<i>Sersic</i>								
SrcDisk	Light	None								
HS50NLDisk, HS30NLDisk	Mass	<i>SPLE</i>	$x = 0.00''$	$y = 0.00''$	$\theta = 75^\circ$	$\theta_E = 1.2$	$q = 0.75$	$\alpha = 2.3$		
HS10NLDisk, ES50NLDisk	Source	<i>Sersic</i>	$x_s = 0.02''$	$y_s = 0.6''$	$\theta_s = 10^\circ$	$I_s = 0.033$	$R_s = 0.9''$	$n_s = 1.0$	$q_s = 0.80$	
ES30NLDisk, ES10NLDisk	Source	<i>Sersic</i>								
SrcBD	Light	None								
HS50NLBD, HS30NLBD	Mass	<i>SPLE</i>	$x = 0.00''$	$y = 0.00''$	$\theta = 45^\circ$	$\theta_E = 1.0$	$q = 0.7$	$\alpha = 1.7$		
HS10NLBD, ES50NLBD	Source 1	<i>Sersic</i>	$x_{s1} = 0.06''$	$y_{s1} = -0.03''$	$\theta_{s1} = 0^\circ$	$I_{s1} = 0.006$	$R_{s1} = 0.4''$	$n_{s1} = 2.5$	$q_{s1} = 0.7$	
ES30NLBD, ES10NLBD	Source 2	<i>Sersic</i>	$x_{s2} = 0.06''$	$y_{s2} = -0.03''$	$\theta_{s2} = 0^\circ$	$I_{s2} = 0.0003$	$R_{s2} = 0.9''$	$n_{s2} = 1.0$	$q_{s2} = 0.80$	
SrcMulti	Light	None								
HS50NLMulti, HS30NLMulti	Mass	<i>SPLE</i>	$x = 0.0''$	$y = 0.0''$	$\theta = 160^\circ$	$\theta_E = 1.0$	$q = 0.75$	$\alpha = 2.1$		
HS10NLMulti, ES50NLMulti	Source 1	<i>Sersic</i>	$x_{s1} = 0.06''$	$y_{s1} = -0.03''$	$\theta_{s1} = 0^\circ$	$I_{s1} = 0.0013$	$R_{s1} = 0.3''$	$n_{s1} = 4.0$	$q_{s1} = 0.9$	
ES30NLMulti, ES10NLMulti	Source 2	<i>Sersic</i>	$x_{s2} = 0.06''$	$y_{s2} = -0.03''$	$\theta_{s2} = 0^\circ$	$I_{s2} = 0.0003$	$R_{s2} = 0.9''$	$n_{s2} = 1.0$	$q_{s2} = 0.9$	
	Source 3	<i>Sersic</i>	$x_{s3} = 0.15''$	$y_{s3} = 0.12''$	$\theta_{s3} = 0^\circ$	$I_{s3} = 0.0036$	$R_{s3} = 0.3''$	$n_{s3} = 3.0$	$q_{s3} = 0.9$	
	Source 4	<i>Sersic</i>	$x_{s4} = -0.07''$	$y_{s4} = -0.11''$	$\theta_{s4} = 0^\circ$	$I_{s4} = 0.0036$	$R_{s4} = 0.3''$	$n_{s4} = 3.0$	$q_{s4} = 0.9$	
LensSrcBulge	Light	<i>Sersic</i>	$x_1 = 0.00''$	$y_1 = 0.00''$	$\theta_1 = 127^\circ$	$I_1 = 0.0085$	$R_1 = 0.6''$	$n_1 = 4.0$	$q_1 = 0.72$	
HS30L50Bulge	Mass	<i>SPLE</i>	$x = 0.00''$	$y = 0.00''$	$\theta = 127^\circ$	$\theta_E = 1.2$	$q = 0.8$	$\alpha = 2.0$		
ES30L50Bulge	Source 1 & 2	<i>Sersic</i>	Identical to <b>SrcBulge</b>							
LensSrcDisk	Light	<i>Sersic</i>	$x_1 = 0.00''$	$y_1 = 0.00''$	$\theta_1 = 127^\circ$	$I_1 = 0.0085$	$R_1 = 0.6''$	$n_1 = 4.0$	$q_1 = 0.72$	
HS30L50Disk	Mass	<i>SPLE</i>	$x = 0.00''$	$y = 0.00''$	$\theta = 127^\circ$	$\theta_E = 1.2$	$q = 0.75$	$\alpha = 2.3$		
ES30L50Disk	Source 1 & 2	<i>Sersic</i>	Identical to <b>SrcDisk</b>							
LensSrcCusp	Light	<i>Sersic</i>	$x_1 = -0.03''$	$y_1 = 0.04''$	$\theta_1 = 30^\circ$	$I_1 = 0.1377$	$R_1 = 1.2''$	$n_1 = 1.25$	$q_1 = 0.6$	
HS20L60Cusp	Mass	<i>SPLE</i>	$x = -0.03''$	$y = 0.04''$	$\theta = 30^\circ$	$\theta_E = 1.4$	$q = 0.7$	$\alpha = 2.35$		
ES20L60Cusp	Source 1	<i>Sersic</i>	$x_{s1} = 0.07''$	$y_{s1} = 0.04''$	$\theta_{s1} = 125^\circ$	$I_{s1} = 0.001$	$R_{s1} = 0.4''$	$n_{s1} = 4.0$	$q_{s1} = 0.8$	
	Source 2	<i>Sersic</i>	$x_{s2} = 0.07''$	$y_{s2} = 0.04''$	$\theta_{s2} = 125^\circ$	$I_{s2} = 0.00004$	$R_{s2} = 0.8''$	$n_{s2} = 1.0$	$q_{s2} = 0.7$	
LensSrcDouble	Light	<i>Sersic</i>	$x_1 = -0.03''$	$y_1 = -0.08''$	$\theta_1 = 127^\circ$	$I_1 = 0.024$	$R_1 = 0.9''$	$n_1 = 2.0$	$q_1 = 0.8$	
HS25L60BD	Mass	<i>SPLE</i>	$x = 0.03''$	$y = -0.08''$	$\theta = 127^\circ$	$\theta_E = 1.2$	$q = 0.75$	$\alpha = 2.1$		
ES25L60BD	Source 1 & 2	<i>Sersic</i>	Identical to <b>SrcBD</b> except $x_{s1} = x_{s2} = 0.25''$ and $y_{s1} = y_{s2} = 0.15''$							
LensSrcMulti	Light	<i>Sersic</i>	$x_1 = 0.04''$	$y_1 = -0.03''$	$\theta_1 = 45^\circ$	$I_1 = 0.017$	$R_1 = 0.8''$	$n_1 = 3.0$	$q_1 = 0.7$	
HS25L75BD	Mass	<i>SPLE</i>	$x = 0.0''$	$y = 0.0''$	$\theta = 160^\circ$	$\theta_E = 1.0$	$q = 0.75$	$\alpha = 2.1$		
ES25L75BD	Source 1, 2, 3 & 4	<i>Sersic</i>	Identical to <b>SrcMulti</b>							
LensMassShear	Light	<i>Sersic</i>	$x_1 = 0.03''$	$y_1 = 0.05''$	$\theta_1 = 60^\circ$	$I_1 = 0.02$	$R_1 = 1.5''$	$n_1 = 2.5$	$q_1 = 0.6$	
HS40L80Disk	Mass	<i>SPLE</i>	$x = 0.03''$	$y = 0.05''$	$\theta = 60^\circ$	$\theta_E = 1.15$	$q = 0.95$	$\alpha = 1.92$		
ES40L80Disk	Mass	<i>Shear</i>	$x_{sh} = 0.03''$	$y_{sh} = 0.05''$	$\theta_{sh} = 40^\circ$	$\gamma_{sh} = 0.03$				
	Source	<i>Sersic</i>	$x_s = 0.06''$	$y_s = -0.07''$	$\theta_s = 30^\circ$	$I_s = 0.016$	$R_s = 0.5''$	$n_s = 1.0$	$q_s = 0.6$	
LensLightBD	Light 1	<i>Dev</i>	$x_{11} = 0.00''$	$y_{11} = 0.00''$	$\theta_{11} = 90^\circ$	$I_{11} = 0.012$	$R_{11} = 0.4''$	$n_{11} = 3.0$	$q_{11} = 0.74$	
HS25L50BD	Light 2	<i>Exp</i>	$x_{12} = 0.00''$	$y_{12} = 0.00''$	$\theta_{12} = 90^\circ$	$I_{12} = 0.026$	$R_{12} = 1.15''$	$n_{12} = 1.0$	$q_{12} = 0.8$	
ES25L50BD	Mass	<i>SPLE</i>	$x = 0.00''$	$y = 0.00''$	$\theta = 90^\circ$	$\theta_E = 1.2$	$q = 0.8$	$\alpha = 2.05$		
	Source 1	<i>Sersic</i>	$x_{s1} = -0.04''$	$y_{s1} = -0.07''$	$\theta_{s1} = 90^\circ$	$I_{s1} = 0.0036$	$R_{s1} = 0.4''$	$n_{s1} = 2.5$	$q_{s1} = 0.7$	
	Source 2	<i>Sersic</i>	$x_{s2} = 0.1''$	$y_{s2} = -0.1''$	$\theta_{s2} = 0^\circ$	$I_{s2} = 0.00036$	$R_{s2} = 0.9''$	$n_{s2} = 1.0$	$q_{s2} = 0.8$	
CoreSrcDisk	Light	<i>Sersic</i>	$x_1 = 0.00''$	$y_1 = 0.00''$	$\theta_1 = 40^\circ$	$I_1 = 0.027$	$R_1 = 0.48''$	$n_1 = 2.5$	$q_1 = 0.6$	
HS35L70Disk	Mass	<i>PL<sub>Core</sub></i>	$x = 0.00''$	$y = 0.00''$	$\theta = 40^\circ$	$\theta_E = 1.4$	$q = 0.8$	$\alpha = 1.85$	$s = 0.2''$	
ES35L70Disk	Source	<i>Sersic</i>	$x_s = 0.12''$	$y_s = 0.25''$	$\theta_s = 10^\circ$	$I_s = 0.04$	$R_s = 0.4''$	$n_s = 1.0$	$q_s = 0.8$	
CoreSrcQuad	Light 1	<i>Sersic</i>	$x_{11} = 0.00''$	$y_{11} = 0.00''$	$\theta_{11} = 110^\circ$	$I_{11} = 0.045$	$R_{11} = 0.25''$	$n_{11} = 2.5$	$q_{11} = 0.77$	
HS40L60BD	Light 2	<i>Exp</i>	$x_{12} = 0.00''$	$y_{12} = 0.00''$	$\theta_{12} = 110^\circ$	$I_{12} = 0.03$	$R_{12} = 1.35''$	$n_{12} = 1.0$	$q_{12} = 0.6$	
ES40L60BD	Mass	<i>PL<sub>Core</sub></i>	$x = 0.00''$	$y = 0.00''$	$\theta = 110^\circ$	$\theta_E = 1.0$	$q = 0.7$	$\alpha = 1.75$	$s = 0.3''$	
	Source 1 & 2	<i>Sersic</i>	Identical to <b>SrcBD</b> except $x_{s1} = x_{s2} = 0.01''$ and $y_{s1} = y_{s2} = -0.01''$							
CoreSrcDouble	Light	<i>Sersic</i>	$x_1 = 0.05''$	$y_1 = -0.06''$	$\theta_1 = 170^\circ$	$I_1 = 0.0067$	$R_1 = 1.5''$	$n_1 = 3.5$	$q_1 = 0.7$	
HS25L50BD	Mass	<i>PL<sub>Core</sub></i>	$x = 0.05''$	$y = -0.06''$	$\theta = 170^\circ$	$\theta_E = 1.3$	$q = 0.8$	$\alpha = 1.65$	$s = 0.25''$	
ES25L50BD	Source 1	<i>Sersic</i>	$x_{s1} = 0.2''$	$y_{s1} = -0.1''$	$\theta_{s1} = 90^\circ$	$I_{s1} = 0.0075$	$R_{s1} = 0.4''$	$n_{s1} = 2.0$	$q_{s1} = 0.9$	
	Source 2	<i>Sersic</i>	$x_{s2} = 0.2''$	$y_{s2} = -0.1''$	$\theta_{s2} = 90^\circ$	$I_{s2} = 0.0025$	$R_{s2} = 0.8''$	$n_{s2} = 1.0$	$q_{s2} = 0.8$	

to 80, with every model again generating a Hubble and Euclid image. Example image names are **CoreSrcDisk**<sub>HS35L70Disk</sub> (tests cored modeling for a flat source profile) and **CoreSrcDouble**<sub>ES25L50BD</sub> (tests cored modeling for a doubly imaged source).

(iv) 10 images generated using a decomposed mass model, aimed at testing light and dark matter modeling as well as the detection of light/dark matter alignments. Five unique lens models are used, generating sources and lenses with the same S/N ranges as before and with every model again being used to generate a Hubble and Euclid image. Included in this set are images that test a small rotational offset and aligned light and dark components. Example image names are **LMDMRot**<sub>HS30L40BD</sub> (tests the detection of a rotational offset) and **LMDMShear**<sub>ES25L80Cusp</sub> (tests decomposed modeling with an external shear).

Fig. 1 shows postage-stamp cut-outs of a small sub-set of images and their source-plane configurations. Comparison between the different cut-outs shows the broad range in image resolution, S/N, and source and lens morphologies the simulation suite covers.

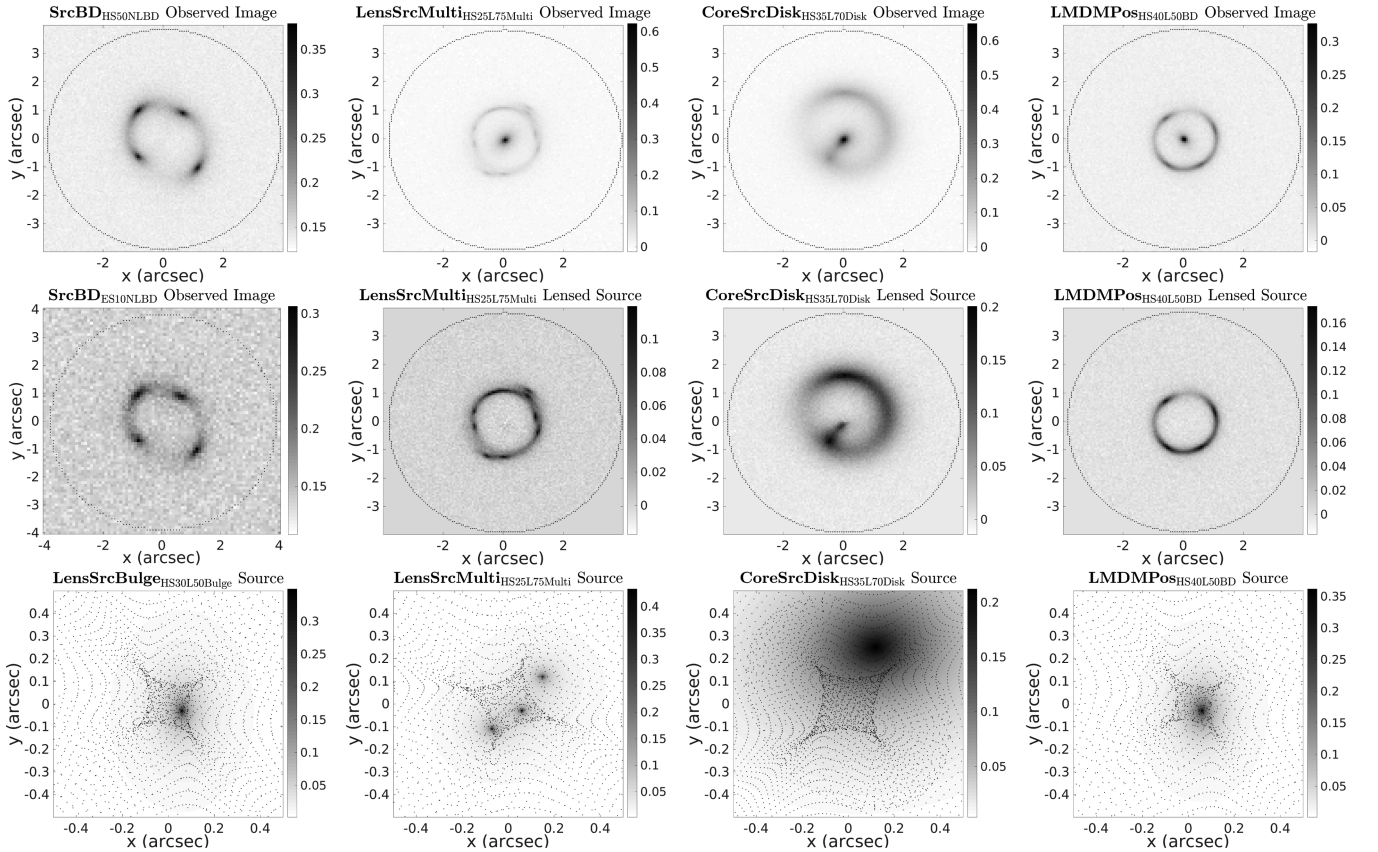
## 4 IMAGE ANALYSIS, LENS MODELING, AND SOURCE RECONSTRUCTION

### 4.1 Extended source modeling

This section gives a brief overview of the theory relevant for modeling strongly lensed extended sources. A more detailed description of this overview can be found in Schneider & Weiss (1992) and Keeton (2003).

**Table 2.** The lens light, mass, and source profiles used to create each image of the simulation suite generated using a decomposed mass profile (e.g. a *Sersic* + *NFW*). The table follows the exact same lay-out as Table 1. The values of  $I_1$ ,  $I_s$ , and  $\Psi_1$  correspond to the values used for the highest S/N Hubble resolution image of each lens model and their values are reduced for generating each model's Euclid resolution images.

Model Title	Component	Model	Parameters							
<b>LMDMalign</b>	Light	<i>Sersic</i>	$x_1 = 0.00''$	$y_1 = 0.00''$	$\theta_1 = 127^\circ$	$I_1 = 0.0085$	$R_1 = 0.6''$	$m_1 = 4.0$	$q_1 = 0.72$	
HS50L40BD	Mass	<i>NFW</i> + $\Psi_1$	$x_d = 0.00''$	$y_d = 0.00''$	$\theta_d = 127^\circ$	$\kappa_d = 0.13''$	$q = 0.82$	$\Psi_1 = 25.0$	$\Psi_1 = 6.73$	
ES50L40BD	Source 1 & 2	<i>Sersic</i>	Identical to <b>SrcBD</b>							
<b>LMDMRot</b>	Light	<i>Sersic</i>	Identical to <b>LMDMalign</b>							
HS50L40BD	Mass	<i>NFW</i> + $\Psi_1$	Identical to <b>LMDMalign</b> except $\theta_d = 132^\circ$							
ES50L40BD	Source 1 & 2	<i>Sersic</i>	Identical to <b>SrcBD</b>							
<b>LMDMPos</b>	Light	<i>Sersic</i>	Identical to <b>LMDMalign</b>							
HS50L40BD	Mass	<i>NFW</i> + $\Psi_1$	Identical to <b>LMDMalign</b> except $x_d = 0.05''$							
ES50L40BD	Source 1 & 2	<i>Sersic</i>	Identical to <b>SrcBD</b>							
<b>LMDMRot90</b>	Light	<i>Sersic</i>	Identical to <b>LMDMalign</b>							
ES30L65Multi	Mass	<i>NFW</i> + $\Psi_1$	$x_d = 0.00''$	$y_d = 0.00''$	$\theta_d = 17^\circ$	$\kappa_d = 0.13''$	$q = 0.82$	$\Psi_1 = 42.0$	$\Psi_1 = 6.73$	
ES50L40BD	Source 1 & 2	<i>Sersic</i>	Identical to <b>SrcBD</b>							
<b>LMDMShear</b>	Light	<i>Sersic</i>	$x_1 = 0.00''$	$y_1 = 0.00''$	$\theta_1 = 100^\circ$	$I_1 = 0.033$	$R_1 = 0.75''$	$m_1 = 2.5$	$q_1 = 0.75$	
HS35L80Cusp	Mass	<i>NFW</i> + $\Psi_1$	$x_d = 0.03''$	$y_d = 0.03''$	$\theta_d = 90^\circ$	$\kappa_d = 0.13''$	$q = 0.8$	$\Psi_1 = 11.5$	$\Psi_1 = 2.0$	
ES35L80Cusp	Mass	<i>Shear</i>	$x_{sh} = 0.03''$	$y_{sh} = 0.03''$	$\theta_{sh} = 150^\circ$	$\gamma_{sh} = 0.03$				
	Source 1 & 2	<i>Sersic</i>	Identical to <b>LensSrcCusp</b> except $R_{s2} = 0.3''$							



**Figure 1.** A sub-set of simulated images corresponding to the images **SrcBD**<sub>HS50NLBulge</sub> (top left panel), **SrcBD**<sub>ES10NLBulge</sub> (middle left panel), **LensSrcMulti**<sub>HS25L75Multi</sub> (second column), **CoreSrcDisk**<sub>HS35L70Disk</sub> (third column), and **LMDMPos**<sub>HS40L50BD</sub> (fourth column), which are given by the 3rd, 9th, 14th, and 17th models listed in Table 1, respectively. The top row shows the simulated images, the middle row lens subtracted images (or a second simulated image at lower resolution and S/N in the left column), and third row the simulated parametric source of each image, where black points are the traced image pixel coordinates.

As discussed in the Introduction, extended source modeling offers information about the second derivative of the lens's potential. However, this signal is encoded into the lensed source's extended surface-brightness distribution and is therefore only available whenever the lensed source is actually observed, around  $R_{Ein}$ , the Einstein

radius. The extension of this measurement to smaller radii (where there is typically no source light) is therefore something of an extrapolation (Sonnenfeld et al. 2012; Schneider & Sluse 2013b; Xu et al. 2016), albeit one aided by how the mass model's overall normalization must still give an accurate  $M_{Ein}$ . The constraints that a

lens offers therefore varies from system to system, depending on the source size, lensing geometry and source and lens redshifts, with the most exceptional examples spanning over 15 kpc in extent (Gavazzi et al. 2008; Eichner, Seitz & Bauer 2012; Sonnenfeld et al. 2012;). Thus, for many lenses, this measurement does not require large extrapolations, except in the very central regions. By fitting the lens’s light profile, `AutoLens` partly constrains these central regions, both through the detection or absence of the lensed source’s faint central features and by incorporating the lens’s light profile into the mass model.

**N15** illustrated the nature of extended source analysis. First, a clear degeneracy emerges between the parameters governing the lens’s mass distribution, which for the power-law density profile used in **N15** was its mass normalization, ellipticity, and density slope (see also Suyu 2012; Suyu et al. 2013). This degenerate subset of mass models all integrate to give approximately the same  $M_{\text{Ein}}$ , with the different models varying only how they distribute this mass. The favoured model from this sub-set is then whichever best reconstructs the extended source. **N15** also demonstrated how these degenerate lens models are fully degenerate with the source-plane magnification (see also Birrer et al. 2016), such that more centrally concentrated mass profiles result in a more spatially expanded source reconstruction (i.e. lower total magnification; see fig. 4 of **N15**). This requires specific care to ensure that the inferred lens model is not biased (Section 4.5) and as such, `AutoLens` adapts to and scales with this phenomenon (Section 4.7).

There is an important caveat to lens modeling of this nature, associated with the form of  $\kappa$  assumed for the lens. If the allowed (parametric) form of  $\kappa$  is unable to accurately follow the actual mass distribution, the sub-set of lens models that integrate to give the correct  $M_{\text{Ein}}$  will offer only an approximate match to the lens’s actual mass profile. They may still provide a good fit to the lensing data, but can misestimate a number of the lens’s properties, like the lens’s true slope at  $R_{\text{Ein}}$ . This is a manifestation of the much-studied mass-sheet transformation (MST) and source position transformation (Falco, Gorenstein & Shapiro 1985; Schneider & Sluse 2013a, b; Schneider 2014a, b; Xu et al. 2016; Tagore et al. 2018). This work circumvents this issue by using the same density profile for both the modelling and creation of each simulated image, as was performed in **N15**. Use of the lens’s light profile to trace its underlying stellar matter profile may reduce the freedom of the MST; however, a more detailed investigation of this is beyond this paper’s scope.

## 4.2 Semilinear inversion

The semilinear inversion (SLI) method simultaneously reconstructs the surface brightness distribution of a strongly lensed source and models the lens galaxy mass distribution. It was first presented in Warren & Dye (2003, **WD03** hereafter), placed within a Bayesian framework by Suyu et al. (2006, **S06** hereafter) and developed into adaptive SLI in Nightingale & Dye (2015, **N15** hereafter). An outline of the SLI method is given here, but readers are referred to these publications for comprehensive details.

The SLI method assumes a pixelized source plane, computing the linear superposition of PSF-smear source pixel images that best fits the observed image, for a given lens model. This is done via the matrix  $f_{ij}$ , which maps the  $j$ th pixel of each lensed image to each source pixel  $i$  and produces the source pixel surface brightness vector  $s$ . Finally, the  $j$ th pixel of the model image is computed as  $\sum_i s_i f_{ij}$ , which is subtracted from the observed image with flux values  $d_j$  and statistical uncertainties  $\sigma_j$ . In the original implementation of the SLI method, the values  $d_j$  have had a pre-computed

foreground light model subtracted. The sum of the squared significances of the residuals between the observed and model images then gives a  $\chi^2$  statistic.

The pixelization used by the SLI method may be discretized into pixels of arbitrary shape or tessellation. In **N15**, the source-plane pixelization was derived using an  $h$ -means clustering algorithm, which defined source pixels as clusters of traced image pixels. The same clustering methodology is used here to compute source pixels, however switching instead to a weighted  $k$ -means clustering algorithm (see Hartigan & Wong 1979). This allows clustering to be weighted, thus enabling the source pixelization to adapt to the source’s surface brightness (see Section 4.7), unlike **N15**, which adapted to the mass model magnification.  $K$ -means clustering also produces more uniform and regular source-plane pixelizations (albeit still stochastic enough to sample and overcome discretization biases). The randomization of the clustering that **N15** showed to remove discreteness biases has also been slightly modified to ensure that even the exact same lens model parametrization gives a different source-plane pixelization (the reason for this is described in Section 4.6).

Due to the ill-posed nature of the matrix inversion used by the SLI method the solution must be regularized using a linear regularization matrix, which is described in **WD03** and Appendix B. Regularization acts as a prior on the source reconstruction, imposing a smooth source solution. `AutoLens` follows a Voronoi regularization scheme that is scale-independent, such that regularization is the same for a larger or smaller source, a property key to handling the source rescaling that emerges during lens modeling. This Voronoi grid is also used to visualize source reconstructions. In **N15**, regularization was controlled by the hyper-parameter  $\lambda$ , which set the degree to which smoothness is imposed on the solution following the Bayesian framework of **S06**. Section 4.6 presents `AutoLens`’s new approach to source-plane regularization.

## 4.3 Lens light and mass modeling

Fitting and subtraction of the lens’s light is fully integrated into `AutoLens`, with all parameters associated with the lens’s light model sampled within the same non-linear parameter space as those governing the mass model. Therefore, for each iteration of the method, before reconstruction of the lensed source, `AutoLens` first computes a model 2-D light distribution using one or more elliptical Sersic functions. The resulting 2-D light model is then convolved with the instrumental PSF and subtracted from the observed image.

The mass model is then used to compute the deflection angle map  $\alpha_{x,y}$  and trace image-pixel to the source plane. The source reconstruction outlined above is then performed. **N15** showed that, due to aliasing effects, the source reconstruction benefits from oversampling (termed subgridding in **N15**), which splits each image-pixel into a set of square sub-pixels, which are each individually traced to the source plane and used by the inversion. Appendix A describes a bilinear interpolation scheme used to speed this calculation up, allowing higher levels of oversampling ( $8 \times 8$ ) to be used in this work. Appendix A also describes how the positions of the image’s brightest pixels are used to speed up mass modeling, by discarding models where they do not trace close to one another.

The incorporation of lens light fitting into `AutoLens` only slightly changes the modeling formalism given in **N15** and the previous section. All pixels within the masked region retain the subscript  $j$ , with the definition of terms  $f_{ij}$  and  $\sigma_j$  unchanged. However,  $d_j$  is now defined to be the observed flux in pixel  $j$  including the lens flux contribution that is denoted as  $b_j$ . The quantity  $D_j$  used

in [WD03](#) and [N15](#) must therefore also change to

$$D_i = \sum_{j=1}^J f_{ij}(d_j - b_j)/\sigma_j^2. \quad (8)$$

$\chi^2$  is therefore given by

$$\chi^2 = \sum_{j=1}^J \left[ \frac{(\sum_{i=1}^I s_i f_{ij}) + b_j - d_j}{\sigma_j} \right]^2. \quad (9)$$

This is identical to before, except for the change in the definition of  $d_j$  and inclusion of the  $b_j$  term. The overall likelihood function follows the same Bayesian framework used in [N15](#) and is given in Section 4.6.

The determination of the lens model parameters is a standard non-linear search problem, performed using the `MultiNest` algorithm (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009), based on the nested sampling Monte Carlo technique of Skilling (2006). As described in [N15](#), the random nature of `AutoLens`'s source-plane discretization results in a noisy likelihood function that can rapidly fluctuate over small scales in parameter space, ill suited to Markov chain Monte Carlo analysis. `MultiNest`'s approach of first mapping out parameter space over large scales, followed by convergence toward the more noisy, higher evidence small scales, is therefore well suited. The implementation of `MultiNest` in `AutoLens` uses constant efficiency sampling mode. This tunes the reduction of `MultiNest`'s elliptical contours such that the acceptance ratio is kept at a target level, which is set to 10 per cent for the final analysis of each image. Importance sampling is also employed, which as discussed in Feroz et al. (2013) improves the accuracy of sampling in constant efficiency mode, especially the estimation of the Bayesian evidence.

#### 4.4 Masking

Before performing the lens analysis, the image is masked, removing the regions in the image that only contain background sky (or unwanted contaminants like foreground stars). In the early stages of development, `AutoLens` used a dual-masking scheme. The first mask encapsulated both the lens and source and had only the lens's light profile subtracted from it. The second was then tailored to contain only the lensed source galaxy, with the source reconstruction performed only within this smaller second mask. The motivation behind this was that the source reconstruction is the most computationally demanding aspect of the analysis, thus a much faster runtime is possible by performing it exclusively on a smaller masked region. Unfortunately, testing of this masking scheme found that it biased the lens's light model, as the omission of the source reconstruction in the first mask meant it dominated the overall  $\chi^2$  value. Attempts to circumvent this by, for example, weighting the likelihood of each mask never led to satisfactory results. Therefore, it was concluded that the lens and source must be analysed within the same masked region and that the approach used in [N15](#) of tailoring a hand-drawn mask around the lensed source was no longer viable. Thus, comparatively wide and extensive masking possibly extending well beyond the lensed source as well as encompassing the entire region within the  $R_{\text{Ein}}$  (where typically no source light is present) is now necessary.

This masking scheme offers a number of benefits to lens modeling. For example, it ensures that if a lens model incorrectly places extraneous images within the image reconstruction, they are not masked out and ignored. Equally, faint source features that may

have been masked previously will now be detected and modeled. The drawback (and reason why source-only masks are generally used in other studies, e.g. Dye et al. 2014; Vegetti et al. 2014) is that the overall run-time of a lens analysis scales directly with the number of image pixels. For this more extensive masking scheme, the number of image pixels increases by a factor of 2–4, leading to an increase of `AutoLens`'s overall run time by the same factor or more. This provides a significant computational challenge and motivates the new source-plane analysis features described in Section 4.7.

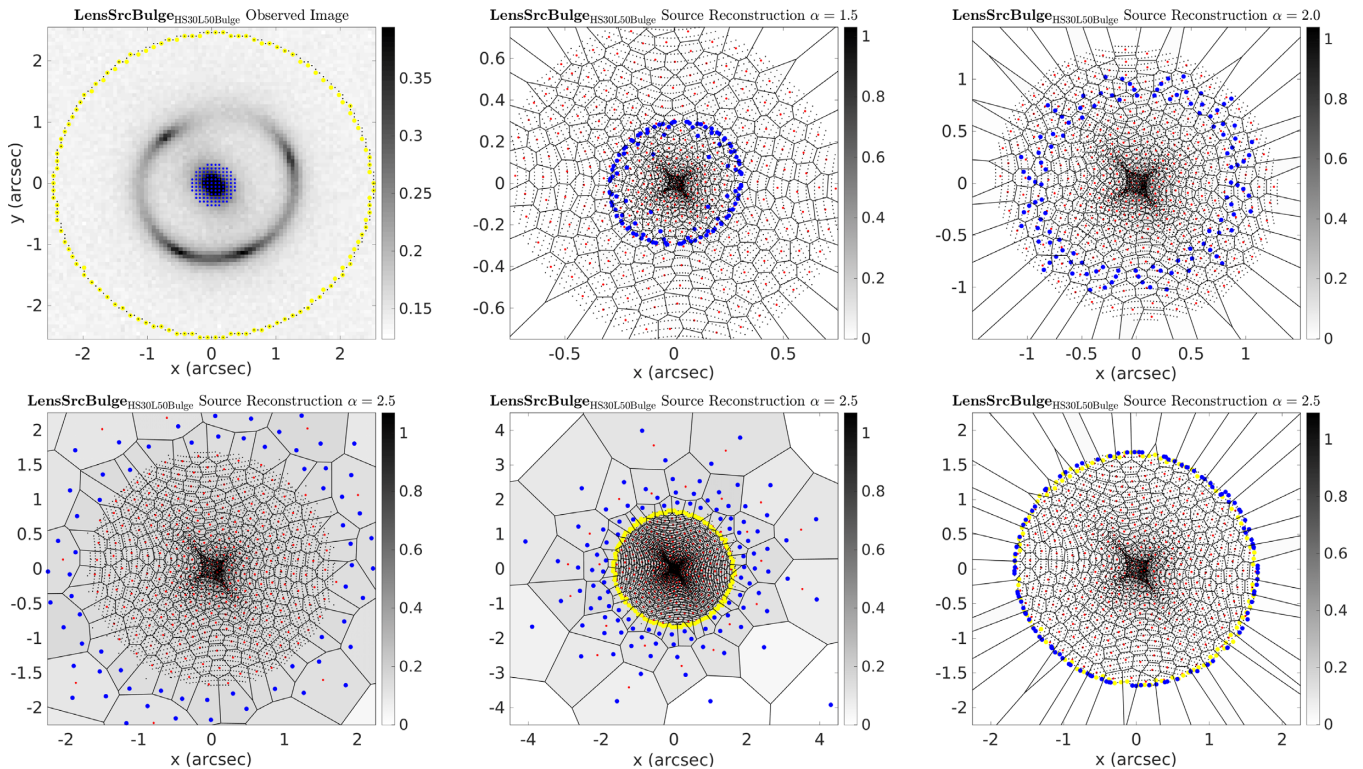
In this work, a circular mask of radius 3.9 arcsec is used to model all lenses, which is sufficiently large to fully capture the source and lens of every simulated image. The use of a circular mask provides a regular and symmetric source-plane pixelization. This gives the adaptive source-plane features described in Section 4.7 better control of the source reconstruction and reduces the discretization effects discussed in [N15](#). In contrast, masks tailored to the lensed source produce irregular edges in the source pixelization that, as discussed next, have the potential to bias the analysis. During an early initialization phase of the pipeline, an annulus mask is used instead of a circular mask.

#### 4.5 Central image pixels

A consequence of the masking scheme above is that the lens's central image pixels are now traced to the source-plane and included as part of the image and source reconstruction. This is acceptable for modeling a cored mass profile (like the  $PL_{\text{Core}}$  profile), as these pixels are expected to trace near the source. However, for singular mass profiles these central image pixels may be significantly demagnified and trace to exterior regions of the source plane that negatively impact the resulting source pixelization. This is illustrated in Fig. 2 for the image `LensSrcSersicCH50L40BD`, where central image pixels are marked as blue dots and in the top-left panel can be seen to correspond to the regions where the lens's light is brightest. The remaining panels show source reconstructions using three different *SPL*E mass models with density slopes ( $\alpha$ ) of 1.5, 2.0 (this image's input value), and 2.5 (see the figure's caption for how the overall lens model is computed). This figure reaffirms the source-plane scaling discussed in [N15](#), noting that the source-plane axis increases from 1.5 arcsec  $\times$  1.5 arcsec for the  $\alpha = 1.5$  mass model to 4.5 arcsec  $\times$  4.5 arcsec for  $\alpha = 2.5$ . For  $\alpha = 2.0$  (the top-right panel) the lens model matches the image's input model. Thus, the lens's light is subtracted perfectly (not shown) and the central pixels trace to regions where the source is very faint, therefore having no impact on the source and image reconstruction.

During testing, it emerged that this ideal scenario was not always reached and central pixels could bias lens modelling in two different ways. The first is due to the interplay between the location to which central image pixels trace in the source plane and the mass model's density slope  $\alpha$ , illustrated in Fig. 2. For mass profiles with a lower value of  $\alpha$ , central image pixels are less demagnified and thus trace closer to the source, giving them the potential to impact the source reconstruction. Indeed, the central image pixels shown in the top-centre panel (for  $\alpha = 1.5$ ) trace within the source's faint extended envelope and are therefore allocated a low level of extraneous flux by the source reconstruction. Thus, it is possible that the source reconstruction wrongly places extraneous flux in the image reconstruction's central regions, which can potentially bias lens modeling in two different ways:





**Figure 2.** A demonstration of where central image pixels trace in the source plane for mass models with varying density-slope, using the image `LensSrcBulgeHS30L50Bulge`, which is pictured in the top left-hand panel. The remaining panels show source reconstructions for this image, using three different lens models, which all use the image’s input light model (see Table 1) and input mass model geometric parameters ( $x = 0.0$  arcsec,  $y = 0.0$  arcsec, and  $\theta = 127^\circ$ ). Three different sets of mass profile parameters are used: *top centre* ( $\alpha = 1.5$ ,  $\theta_E = 1.16$ ,  $q = 0.93$ ); *top right* ( $\alpha = 2.0$ ,  $\theta_E = 1.2$ ,  $q = 0.73$ ); *bottom row* ( $\alpha = 2.5$ ,  $\theta_E = 1.36$ ,  $q = 0.56$ ). These parameters were determined by fixing  $\alpha$  to 1.5, 2.0, and 2.5 and computing  $\theta_{Ein}$  and  $q$  via a full `AutoLens` analysis. Black dots depict the locations of traced image pixels and red dots are the centres of each Voronoi source pixel. Central image pixels are defined as pixels whose image-plane coordinates are within 6 pixels (0.36 arcsec) of the mass model centre and are marked using blue dots, depicted in both the image plane (top left) and source plane (remaining panels). Their source-plane locations can be seen to depend critically on  $\alpha$ , tracing further away from the source as  $\alpha$  increases. The overall size of the source plane increases from  $1.5 \text{ arcsec} \times 1.5 \text{ arcsec}$  for  $\alpha = 1.5$  to  $4.5 \text{ arcsec} \times 4.5 \text{ arcsec}$  for  $\alpha = 2.5$ , demonstrating the source-plane scaling effect discussed in N15. The image-plane mask and border are marked with yellow dots in the top-left panel and the location this border traces to in the source plane is shown in the bottom-centre and bottom-right panels (for the  $\alpha = 2.5$  mass model), forming a ring of pixels outside of which only demagnified central image pixels trace. Outside this border, central image pixels can be seen to form their own source pixels, an effect that can prove problematic for the source reconstruction. Therefore, the scheme shown in the bottom-right panel is used, which relocates central image pixels that trace beyond this border to its edge.

- (i) When the lens light subtraction leaves significant residuals. In this instance, the mass model may be biased towards lower  $\alpha$  solutions that allow the source reconstruction to fit these residuals.
- (ii) When the value of  $\alpha$  assumed for the mass model is lower than the true value. In this case, the source reconstruction may fit some of the flux in the central image pixels, leading to an inaccurate lens light model.

At the beginning of a lens analysis, care must therefore be taken to ensure that these biases are circumvented. `AutoLens` achieves this by assuming an *SPLE* mass model with a fixed value of  $\alpha = 2.2$  early in the analysis. Later in the analysis, once the lens subtraction is accurate,  $\alpha$  can safely be treated as free. The second problem is counteracted because a slope  $\alpha = 2.2$  is steeper than most strong lenses (Koopmans et al. 2009), thus ensuring that central image pixels trace well away from the source. Whilst this may not be sufficient for all lenses (e.g. those with a very steep density profile or very extended source) it has proven adequate for all test cases thus far. Sanity checks flag up when central image pixels receive

extraneous flux, ensuring that this bias will be spotted on large lens samples.

The second problem is also illustrated in Fig. 2, particularly the bottom-centre and bottom-right panels, which depict where central image pixels trace relative to the ‘image border’ (yellow dots), the ring of image pixels located at the edge of the image-plane mask. These panels show that for the  $\alpha = 2.5$  mass model (chosen to exaggerate this effect) central image pixels trace well beyond this image-plane border (the yellow ring of dots) in the source plane, forming their own source-pixels and offering the reconstruction an unphysical means by which to fit the lens subtraction’s residuals or noise. Thus, the mass model may be biased to high  $\alpha$  solutions that allow these exterior source pixels to form. To counteract this, `AutoLens` relocates all central image pixels that trace beyond the image border in the source-plane to its edge, as shown in the bottom-right panel of Fig. 2. This prevents central image pixels forming their own source pixels and therefore removes their potential to bias the lens model in a computationally efficient manner.

#### 4.6 Bayesian framework

AutoLens's WQsource and image analysis is based on the Bayesian framework for interpolation, model comparison, and regularization presented in MacKay (1992) (in particular chapters 2 and 6), which S06 generalized to lens modeling. Many other methods in the literature are also based on this (e.g. Dye et al. 2008; Vegetti & Koopmans 2009a; Collett & Auger 2014; Tagore & Keeton 2014). This framework objectively ranks every image and source reconstruction that is produced by AutoLens's linear inversion step. For every light model, mass model, and source reconstruction, the overall probability is given by the Bayesian evidence,  $\epsilon$ ,

$$-2 \ln \epsilon = \chi^2 + \mathbf{s}^T \mathbf{H}_\Lambda \mathbf{s} + \ln [\det(\mathbf{F} + \mathbf{H}_\Lambda)] - \ln [\det(\mathbf{H}_\Lambda)] + \sum_{j=1}^J \ln [2\pi(\sigma_j)^2]. \quad (10)$$

This expression was derived in Dye et al. (2008) from S06, is used for all modeling presented in this work, and is equivalent to the expression used in N15. However, the regularization matrix  $\mathbf{H}_\Lambda$  has been redefined such that AutoLens can now also apply a non-constant regularization scheme to the source reconstruction, as described next. The mathematical formalism for this non-constant regularization is given in Appendix B.

Equation (10) quantifies three aspects of the image and source reconstruction, the first being the quality of the image reconstruction. Because the source reconstruction is a linear inversion that takes as an input the image-data when reconstructing it, it is in principle able to perfectly reconstruct the image regardless of the image's noise or the accuracy of the lens model (e.g. at infinite source resolution without regularization). This is why the problem is 'ill-posed' and why regularization is necessary. However, this still raises the question of what constitutes a 'good' solution? Equation (10) defines this by assuming that the image data consist of independent Gaussian noise in each image pixel, defining a 'good' solution as one whose  $\chi^2$  residuals are consistent with Gaussian noise, therefore producing a reduced  $\chi^2 \sim 1$ . Solutions that give a reduced  $\chi^2 < 1$  are penalized for being overly complex and fitting the image's noise, whereas those with a reduced  $\chi^2 > 1$  are penalized for not invoking a more complex source model when the data support it. In both circumstances, these penalties lead to a reduction in  $\ln \epsilon$ .

The second aspect of the analysis that equation (10) quantifies is the complexity of the source reconstruction. This uses terms 2, 3, and 4 of this expression (those containing the regularization matrix  $\mathbf{H}_\Lambda$ ), which from here on are collectively referred to as the 'regularization terms'. These terms estimate the number of source pixels that are used to reconstruct the image, after accounting for their correlation with one another due to regularization. Solutions that require fewer correlated source pixels collectively decrease the total value of these regularization terms, increasing the value of  $\ln \epsilon$ . Thus, simpler and less complex source reconstructions are favoured by this expression.

Finally, equation (10) favours models that fit higher S/N realizations of the observed imaging data (where the S/N is determined using the image-pixel variances, the  $\sigma_j$ 's found in the  $\chi^2$ ,  $\mathbf{F}$ , and  $\sum_{j=1}^J \ln [2\pi(\sigma_j)^2]$  terms of equation (10)). If fixed variances are assumed throughout the analysis, this aspect of equation (10) has no impact on modeling. However, a number of methods have invoked scaling the image pixel variances wherever the image reconstruction fits the data poorly (e.g. Suyu 2012), an approach AutoLens follows.

The premise is that whilst increasing the variances of image pixels lowers their S/N values and therefore also decreases  $\ln \epsilon$  by increasing  $\sum_{j=1}^J \ln [2\pi(\sigma_j)^2]$ , doing so may produce a net increase in  $\ln \epsilon$  by decreasing  $\chi^2$  and  $\mathbf{F}$ . This occurs when the  $\chi^2$  values of the image pixels whose variances are increased were initially very high and therefore fit poorly by the lens model. Conversely, variances cannot be reduced to arbitrarily low values, as doing so will inflate their  $\chi^2$  contribution (again decreasing  $\ln \epsilon$ ). In fact, AutoLens does not allow a pixel's variance to be scaled below its 'baseline' value, the value that is expected from a consideration of instrumental noise sources like Poisson counts and read noise.

In summary,  $\ln \epsilon$  is maximized for solutions that most accurately reconstruct the highest S/N realization of the observed image, without overfitting its noise and using the fewest correlated source pixels. By employing this framework throughout, AutoLens objectively determines the final lens model following the principles of Bayesian analysis and Occam's Razor.

The simplest application of the Bayesian evidence was shown in N15, where it was used to set the regularization coefficient  $\lambda$ , a hyperparameter that controls the degree of smoothing applied to the source reconstruction ( $\lambda$  is included in equation (10) via the matrix  $\mathbf{H}_\Lambda$ ). This amounted to fixing the lens model (and the source pixelization, regularization scheme, etc.) and iterating over the value of  $\lambda$  until the peak value of  $\ln \epsilon$  is reached. This peak value strikes a balance: too high values of  $\lambda$  oversmooth the source reconstruction and thus give a poor overall fit to the data (decreasing  $\ln \epsilon$  by increasing  $\chi^2$ ), whereas too low values give a source reconstruction that not only accurately reconstructs the image but also fits large portions of its noise (decreasing  $\ln \epsilon$  by increasing the regularization terms). The optimum value of  $\lambda$  therefore again corresponded to the solution that gives an overall reduced  $\chi^2$  of approximately 1. Section 5 demonstrates that there are many scenarios where this simple scheme does not produce a satisfactory fit to the image data, motivating the features introduced below.

#### 4.7 Adaptive image and source reconstruction

In addition to the lens model, the source and image analysis therefore determine the value of  $\ln \epsilon$ . For instance,  $\ln \epsilon$  depends on the source-plane pixelization (see N15 and also Tagore & Keeton 2014), the degree of regularization applied to it and the regularization scheme that is applied (e.g. zeroth-order, gradient, curvature; see WD03). The observed image's variances (which can now be scaled) also determine  $\ln \epsilon$ . Thus, the set-up of the source and image analysis will determine the lens model that is inferred. To determine what is objectively the most probable lens model, one must therefore find the model that maximizes  $\ln \epsilon$  including all these aspects of the analysis. AutoLens achieves this by changing its source pixelization, regularization, and image variances, in conjunction with the lens model, throughout the analysis. Other methods follow a similar approach for choosing the source-plane regularization scheme or resolution (e.g. Suyu 2012; Vegetti et al. 2014), but do not do so in a fully automated or self-consistent way.

To adapt the source reconstruction and scale the image's variances, pre-computed model images of the lens galaxy's light profile and reconstructed lensed source (in the image-plane) are used (using lens light and mass models that have already been estimated in earlier phases of the automated analysis pipeline described in Section 6). These model images are stored in vectors of  $J$  image pixels, where the lensed source model in each image pixel is given by  $\Xi_j = \sum_i s_i f_{ij}$  and the lens light model in each image pixel by

$\mathbf{L}_j = b_j$ . These vectors are updated throughout the automated analysis pipeline and correspond to the highest likelihood model that has previously been estimated.

In the following sub-sections, we describe each adaptive feature of the source and image analysis, alongside their associated hyperparameters.

#### 4.7.1 Source pixelization

Three new hyperparameters are associated with the source pixelization. The first is simply the number of source-plane pixels,  $N_s$ . The second and third control the source plane clustering. The weighted  $k$ -means clustering algorithm used for source plane pixels minimizes the sum of cluster ‘energies’,  $E$ , given by

$$E = \sum_{i=1}^I e_i = \sum_{i=1}^I \sum_{k=1}^K \left( \frac{r_k}{W_k} \right)^2, \quad (11)$$

where a cluster energy  $e_i$  is the quadrature sum of the distances  $r_k$  of its  $K$  associated traced image-plane pixels to its center divided by each pixel’s weight  $W_k$ . In N15, all traced image pixels were given unit weighting, leading the method to adapt to the mass model’s magnification pattern. Instead, the method now adapts to the surface brightness of the lensed source, using the weight vector  $\mathbf{W}$ , which calculates the weight  $W_j$  of each image pixel using the pre-computed source model vector as

$$W_j = \left[ \left( \frac{\Xi_j - \Xi_{\min}}{\Xi_{\max} - \Xi_{\min}} \right) + L_{\text{Clust1}} \right]^{L_{\text{Clust2}}}, \quad (12)$$

where  $\Xi_{\min}$  and  $\Xi_{\max}$  are the maximum and minimum values of  $\Xi$  such that the first term of the right-hand side of this equation ranges between 0 and 1. As the hyperparameter  $L_{\text{Clust2}}$  increases, the separation between the lowest and highest  $W_j$  values increases, such that minimization of the statistic  $E$  prefers a source-plane clustering that places a greater number of smaller source pixels within the source’s brightest regions. Conversely, as larger values of the hyperparameter  $L_{\text{Clust1}}$  are added, the resulting distribution of  $W_j$  values is flattened, such that minimization of the statistic  $E$  places more source pixels away from the source. Together,  $L_{\text{Clust1}}$  and  $L_{\text{Clust2}}$  give AutoLens complete control of its source pixelization. For  $L_{\text{Clust2}} = 0$ , all  $W_j = 1$  and source-plane adaptation reverts to pure magnification scaling as in N15. Negative values of  $L_{\text{Clust1}}$  and  $L_{\text{Clust2}}$  are not permitted, which would lead the source-plane to adapt to regions of background sky.

Fig. 3 illustrates surface brightness adaptation for the simulated images **LensSrcDisk**<sub>HS30L50Disk</sub> and **LensSrcCusp**<sub>HS20L60Cusp</sub>. The left-hand panels show source reconstructions not using this feature, equivalent to using the analysis of N15 or a value  $L_{\text{Clust2}} = 0$ . The right-hand panels show the result of including  $N_s$ ,  $L_{\text{Clust1}}$ , and  $L_{\text{Clust2}}$  as free parameters in the hyperparameter optimization. For the image **LensSrcCusp**<sub>HS20L60Cusp</sub> surface-brightness adaptation can be seen to have a significant effect, congregating a large number of source pixels around the cuspy source’s bright central regions. For the image **LensSersicDisk**<sub>HS30N50Flat</sub>, it plays a lesser role, owing to the source’s flatter light profile.

#### 4.7.2 Source and lens contribution maps

The remaining adaptive image and source features require an estimate of how much of the flux in each image pixel can be attributed to the source and lens. To achieve this, two ‘flux contribution maps’ are generated,  $\Omega_{\text{Src}}$  and  $\Omega_{\text{Lens}}$ . To compute these vectors, the total flux

in each image pixel that can be accounted for by the pre-computed source and lens light models is first computed as

$$T_j = \Xi_j + \mathbf{L}_j. \quad (13)$$

The contribution of flux that can be attributed to the source light in each image pixel is then estimated as

$$\Omega_{\text{Src},j} = \frac{\Xi_j}{T_j + \omega_{\text{SrcFrac}}}, \quad (14)$$

where values of  $\Omega_{\text{Src}}$  below 0.02 are set to 0 to remove residual features in the source reconstruction. The contribution of flux from the lens is then given as

$$\Omega_{\text{Lens},j} = \frac{\mathbf{L}_j}{T_j + \omega_{\text{LensFrac}}}. \quad (15)$$

Both vectors are then divided by their maximum values, such that they range between values just above 0 and 1.  $\Omega_{\text{Src}}$  will therefore contain values close to 1 where only the source is present and close to 0 where it is not, whereas  $\Omega_{\text{Lens}}$  will behave analogously for the lens. The above expressions also include the hyperparameters  $\omega_{\text{SrcFrac}}$  and  $\omega_{\text{LensFrac}}$ , the practical role of which is to allow the source and lens contribution maps to attribute more pixels to values closer to 1. Without these hyperparameters only the brightest pixels are able to obtain a value near 1, limiting the applicability of the contribution maps for the features discussed next.

Fig. 4 shows the flux contribution maps of the images **LensSrcCusp**<sub>HS20L60Cusp</sub> and **LensSrcDisk**<sub>HS50L100Flat</sub>, where both, as expected, correctly trace either the source or lens.

#### 4.7.3 Luminosity-weighted regularization

The next three hyperparameters introduce a luminosity-weighted regularization scheme, using the redefined regularization matrix  $\mathbf{H}_\Lambda$  given in equation (10) and described in Appendix B. A similar scheme is employed by Suyu et al. (2013) and Vegetti et al. (2014).

To weight regularization by the lensed source’s flux, each source pixel requires some measure of how much of the source’s flux it contains before the actual source reconstruction is performed. To do this,  $\Omega_{\text{Src}}$  is used, summing over the  $K$  image pixels allocated to each source pixel to compute the vector  $\mathbf{v}$  as

$$v_i = \frac{\sum_{k=1}^K \Omega_{\text{Src},k}}{K}, \quad (16)$$

where  $i$  is again the source pixel number. Each element in  $\mathbf{v}$  is divided by  $K$  to normalize for the number of allocated image pixels, which can vary due to the  $k$ -means algorithm. The vector  $\mathbf{V}$  is then computed, where each element is given by

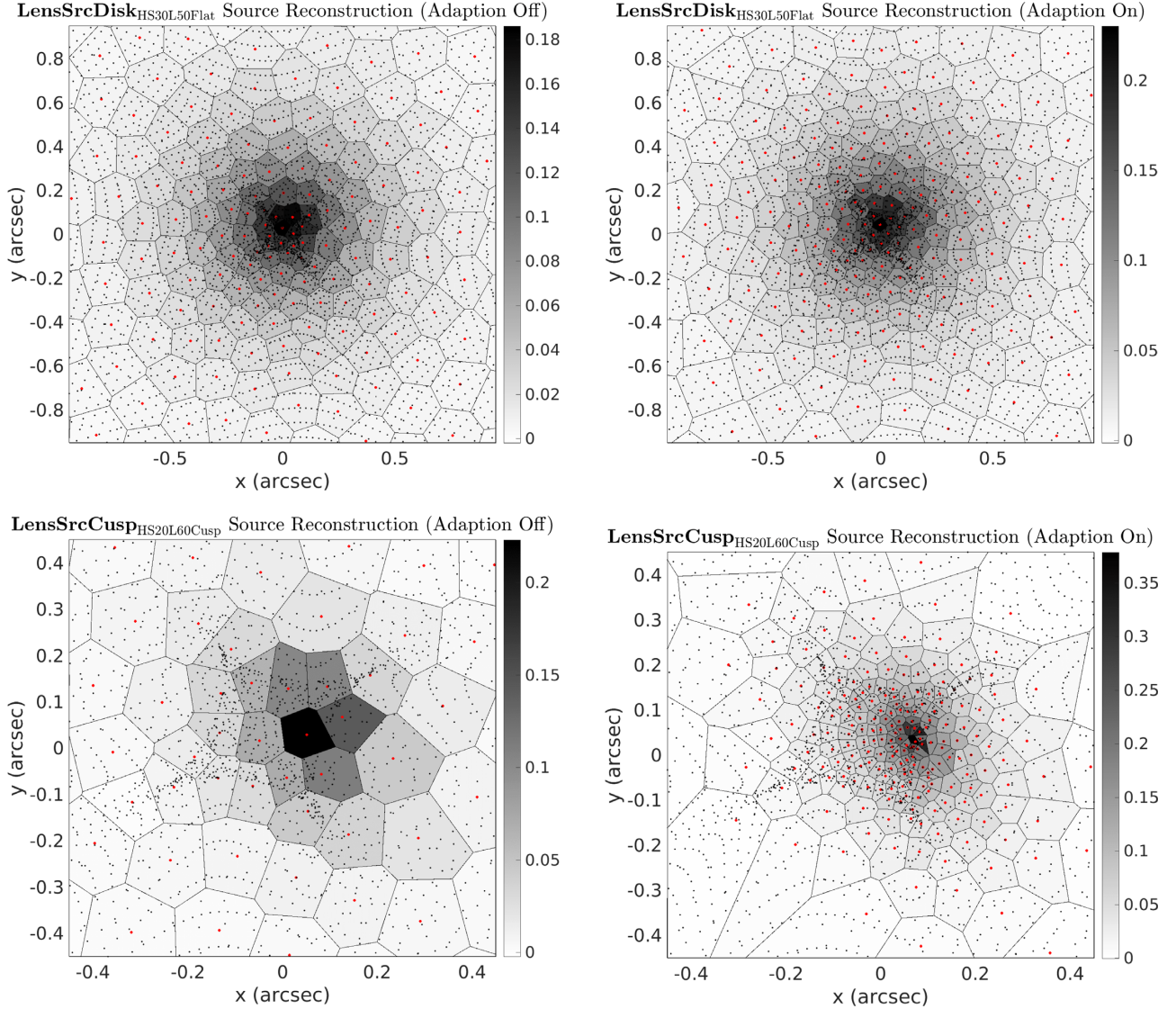
$$V_i = \left[ \frac{v_i}{v_{\max}} \right]^{L_{\text{Lum}}}. \quad (17)$$

Once again, each element is divided by the maximum value of  $\mathbf{v}$  to scale all values between 0 and 1 and raised to the power of the hyperparameter  $L_{\text{Lum}}$ .  $\mathbf{V}$  is then used to compute the luminosity-weighted regularization value of each source pixel (see Appendix B) as

$$\Lambda_i = \lambda_{\text{Src}} V_i + \lambda_{\text{BG}} (1 - V_i), \quad (18)$$

therefore leading to two regularization coefficients  $\lambda_{\text{Src}}$  and  $\lambda_{\text{BG}}$ .

The importance of luminosity-weighted regularization is that it divides source-plane regularization into two regions: (i) pixels that map to the lensed source and (ii) pixels that map to the background sky or central regions of the lens galaxy. The hyperparameter  $L_{\text{Lum}}$



**Figure 3.** An illustration of source surface-brightness adaptation (see Section 4.7.1) using the images **LensSersicDisk**<sub>HS30L50Disk</sub> (top row) and **LensCuspySrc**<sub>HS20L60Cusp</sub> (bottom row). All panels depict the result of using each image's input lens model to optimize the hyperparameters, with the left-hand panels adapting to the mass model's magnification and optimizing just  $\lambda$  ( $N_s = 500$ ) and right-hand panels adapting to the source's surface brightness by optimizing all hyperparameters described in this section, including  $N_s$ ,  $L_{\text{Clust}1}$ , and  $L_{\text{Clust}2}$ . Black dots depict the locations of traced image-pixels and red dots are the centres of each Voronoi source pixel. For both images, surface-brightness adaptation congregates more source pixels in the source's brighter central regions, leading to an increase in  $\ln c$ .

controls the smoothness of the transition between these two regions, whereby higher values give a sharper transition and  $L_{\text{Lum}} = 0$  reverts to the constant regularization scheme of N15. By using two regularization coefficients ( $\lambda_{\text{Src}}$  and  $\lambda_{\text{BG}}$ ), each region therefore receives its own level of regularization.

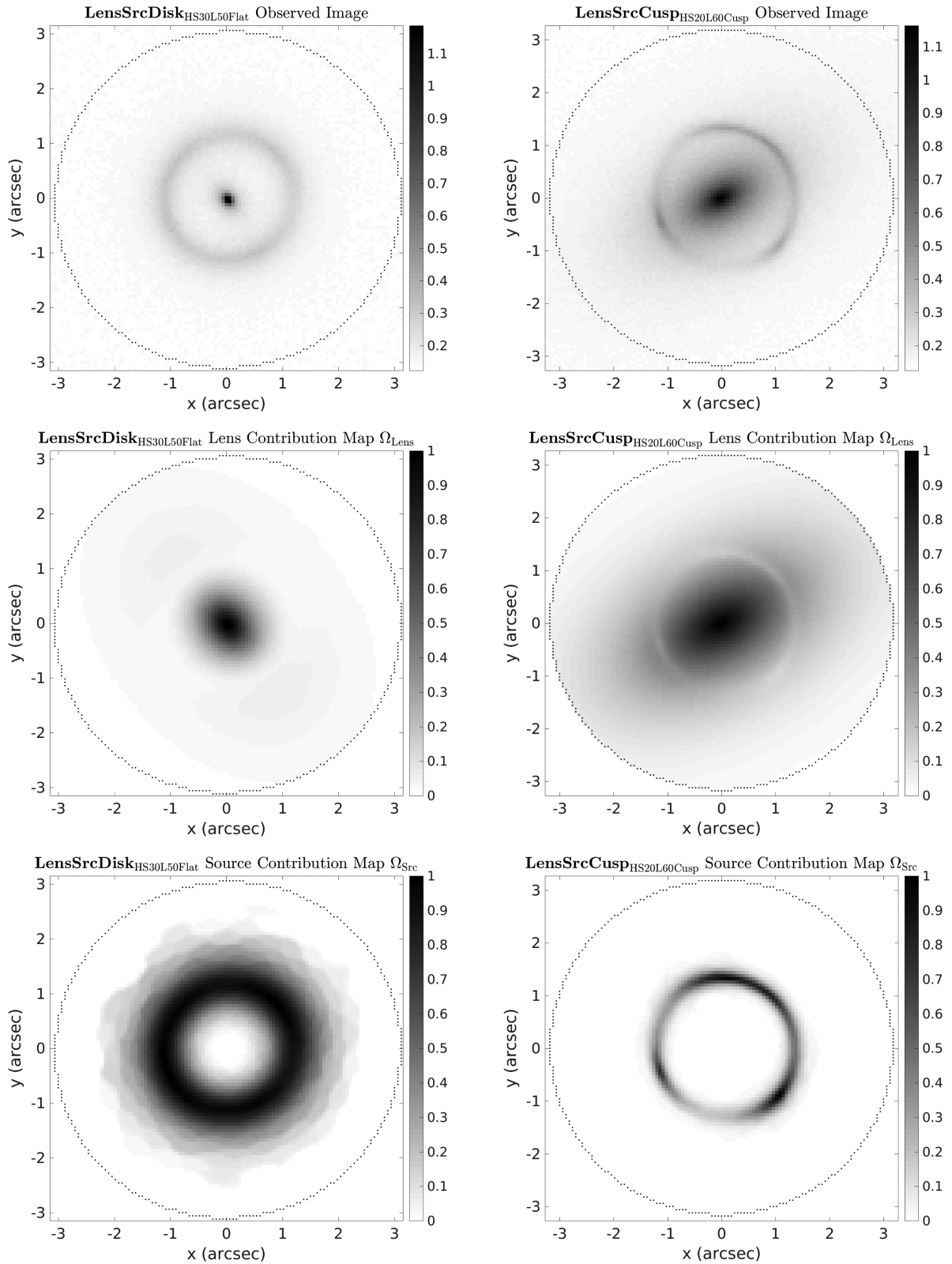
Fig. 5 illustrates this, by showing the effective regularization coefficient  $\lambda_{\text{eff}}$  applied to each source pixel (see Appendix B for the exact definition of this quantity). The constant regularization scheme, which gives each source pixel the same value of  $\lambda_{\text{eff}}$ , is not shown, but computes values of  $\lambda_{\text{eff}} = 610$  and  $\lambda_{\text{eff}} = 230$  for **LensSersicDisk**<sub>HS30L50Disk</sub> and **LensCuspySrc**<sub>HS20L60Cusp</sub>, respectively. The luminosity-weighted scheme reduces regularization in the central regions of the source, facilitating a more detailed reconstruction of its light. Simultaneously, it increases the regularization of the source-plane's exterior pixels (where the source isn't located),

allowing the method to fully correlate the exterior source pixels that only reconstruct the background sky. The constant regularization scheme is used in the early stages of the automated pipeline, before non-constant regularization has been appropriately set up.

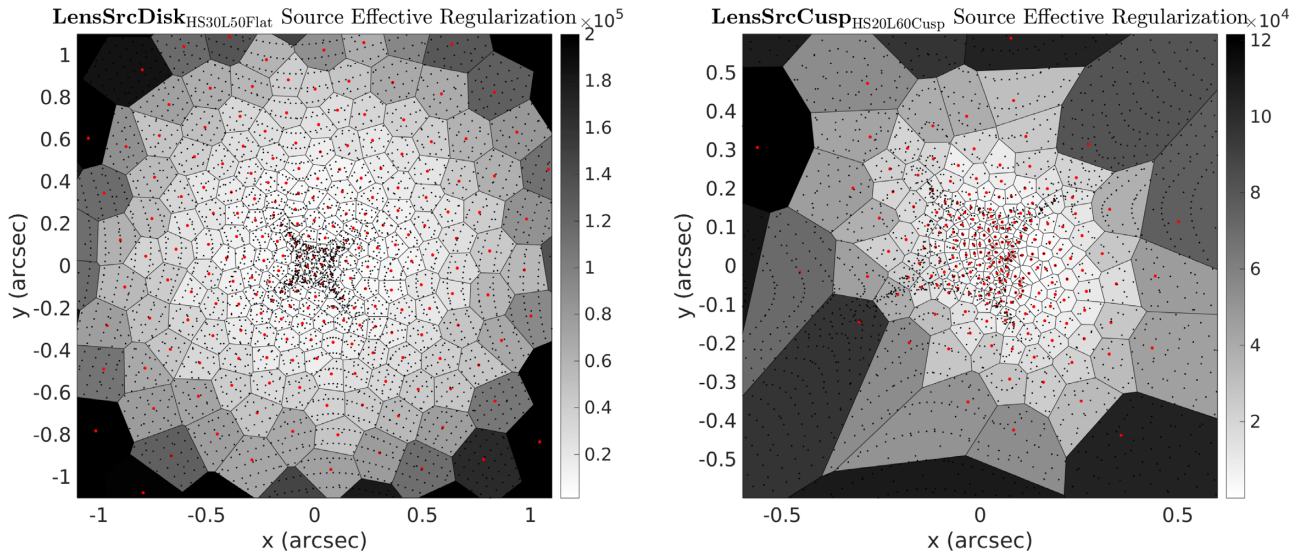
#### 4.7.4 Sky background

The data-vector  $\mathbf{d}$  does not include the background sky, which is subtracted using each image's background-sky flux  $f_{\text{BG}}$ . However, for real data,  $f_{\text{BG}}$  is estimated using the image itself (e.g. by taking a median value of a region of sky), which introduces uncertainty in  $f_{\text{BG}}$ . Therefore, to include this uncertainty, sky subtraction is controlled by the hyperparameter  $\omega_{\text{sky}}$ , where

$$\mathbf{d} = \mathbf{d}' + \omega_{\text{sky}} f_{\text{BG}}, \quad (19)$$



**Figure 4.** The observed images (top row), lens flux contribution maps  $\Omega_{\text{Lens}}$  (middle row), and source flux contribution maps  $\Omega_{\text{Src}}$  (bottom row) described in Section 4.7.5 for the images **LensSrcDisk<sub>HS30L50Disk</sub>** (left-hand column) and **LensSrcCusp<sub>HS20L60Cusp</sub>** (right-hand column). Figures show the result of using each image's input lens model to optimize all of the hyperparameters described in this section, including  $\omega_{\text{SrcFrac}}$  and  $\omega_{\text{LensFrac}}$ . The contribution maps successfully split each pixel's flux contribution between the source and lens.



**Figure 5.** An illustration of the luminosity-weighted regularization scheme described in Section 4.7.3 using the images **LensSersicDisk**<sub>HS30L50Disk</sub> (left-hand panel) and **LensCuspySrc**<sub>HS20L60Cusp</sub> (right-hand panel). Black dots depict the locations of traced image-pixels and red dots are the centres of each Voronoi source pixel. Both panels correspond to the result of optimizing the hyperparameters, including  $\lambda_{\text{Src}}$ ,  $\lambda_{\text{Bg}}$ , and  $L_{\text{Lum}}$ , with the input lens model, thereby producing a nonconstant regularization scheme. Grey scaling shows the effective regularization coefficient  $\lambda_{\text{eff}}$  of each source pixel, which as described in Appendix B represents the effective degree of regularization applied to that pixel. The constant regularization scheme that is produced by using only the hyperparameter  $\lambda$  is not shown, but gives all pixels the same value of  $\lambda_{\text{eff}} = 610$  for the image **LensSersicDisk**<sub>HS30L50Disk</sub> and  $\lambda_{\text{eff}} = 230$  for **LensCuspySrc**<sub>HS20L60Cusp</sub>. Luminosity-weighted regularization can be seen to reduce the degree of regularization applied to the source's brightest regions and increase it in its exterior regions, both contributing to an increase in  $\ln \epsilon$ .

noting that  $\mathbf{d}$  and  $\mathbf{d}'$  must therefore be in units of electrons per second. Early in the analysis  $\omega_{\text{Sky}}$  is fixed to zero. The prior assigned to  $\omega_{\text{Sky}}$  can be chosen to match the uncertainties found by the sky estimation. The motivation behind incorporating the sky subtraction into the analysis is well documented, whereby a poor or uncertain sky subtraction can make it difficult to quantify the faint regions of a galaxy's extended light profile (Häußler et al. 2013).

#### 4.7.5 Variance scaling

A baseline variance is assigned to each pixel as the quadrature sum of Gaussian background noise and Poisson photon error

$$\sigma_{j,\text{base}} = \sqrt{\sigma_{\text{BG}}^2 + d_j}, \quad (20)$$

where  $\sigma_{\text{BG}}$  is the value of the overall background noise in counts (in this work the background sky and read noise) and where  $d_j$  has been converted from electrons per second to counts. This gives the image's 'baseline variance-map', with its corresponding  $\chi^2$  values termed  $\chi_{\text{base}}^2$ . For real imaging data, equation (20) could contain additional terms due to other aspects of the data reduction (e.g. hot pixels, cosmic rays, dithering etc.). The initial stages of an AutoLens analysis use the baseline variance map. However, after these initial stages, the variances may be increased in regions of the image where a poor fit is obtained.

The next set of hyperparameters thus offers AutoLens the ability to perform this variance scaling, therefore producing a 'scaled variance-map' with corresponding scaled  $\chi_{\text{scale}}^2$  values. The variances are scaled separately for the background, source, and lens, by using the flux contributions maps and the expression

$$\sigma_{j,\text{scale}} = \sigma_{j,\text{base}} + \omega_{\text{BG}}\sigma_{\text{BG}} + \omega_{\text{Src}}\sqrt{d_j}(\Omega_{\text{Src},j})^{\omega_{\text{Src}2}} + \omega_{\text{Lens}}\sqrt{d_j}(\Omega_{\text{Lens},j})^{\omega_{\text{Lens}2}}, \quad (21)$$

where  $\sigma_{\text{BG}}$  and  $d_j$  are again in counts. The  $\omega$  terms are all hyperparameters that scale the variances of image pixels in different regions of the observed image. If the lens's light is not modeled, the corresponding  $\omega_{\text{Lens}}$  terms are not included in equation (21) or the hyperparameter optimization. The scaled variances cannot go below their baseline values because the method requires that  $\omega_{\text{BG}} \geq 0$ ,  $\omega_{\text{Src}} \geq 0$ , and  $\omega_{\text{Lens}} \geq 0$ . The background sky variances are scaled because there is uncertainty in the background sky subtraction.

During testing, it emerged that the method must not be allowed to scale variances to arbitrarily high values. This prevents the S/N of image pixels that actually contain significant flux from being unrealistically small. To implement this, first, the maximum values of  $\chi_{\text{base}}^2$  are determined using the cuts  $\Omega_{\text{Src}} > 0.75$  and  $\Omega_{\text{Lens}} > 0.75$ . These provide an indication of how well the lens model currently fits the source and lens galaxies. The highest values allowed for the hyperparameters  $\omega_{\text{Src}}$  and  $\omega_{\text{Lens}}$  are then set such that their corresponding  $\chi_{\text{scale}}^2$  values cannot be scaled below a target value  $\chi^2$  value, which we set to 10. If either  $\chi_{\text{base}}^2$  value is already below 10 (because the lens model is already fitting the data accurately), variance scaling for that component is switched off and its corresponding hyperparameters are omitted. For this reason, variance scaling is not required in cases where the lens model that created the simulated image matches the model used to fit it. We therefore defer its demonstration until the next subsection when this is not the case.

#### 4.7.6 Implementation

Whereas N15 used just one hyperparameter, AutoLens now uses up to 14 simultaneously. Setting these parameters by maximizing  $\ln \epsilon$  can therefore no longer rely on simple iteration as it did in N15. Instead, a fully non-linear MultiNest search is performed, which treats every hyperparameter as a free parameter. This again uses

constant efficiency mode and importance sampling and the priors used for each hyperparameter are discussed in Appendix C. The figures shown in the previous section were generated by using this non-linear optimization.

In general, the lens model and hyperparameters are sampled separately from one another. That is, a lens model is estimated, which is used to optimize the hyperparameters, which are next fixed to improve the lens model, and so on. Bayesian inference therefore retains the three-level structure described in Dye et al. (2008), where the linear source inversion forms the inner-most level. However, `AutoLens` can include the hyperparameters in the non-linear search of the lens model, thus sampling any of the parameters described above alongside the mass and light models. This feature is used at various points throughout the analysis pipeline and the importance of this will be discussed in Section 6.

It is here that the fully randomized source-plane discretization discussed in Section 4.2 is important, where the random seed of source-plane clustering was updated to always produce a different source-plane discretization. When the lens’s mass model is fixed (as it is for a hyperparameter optimization), its corresponding deflection angles are also fixed, therefore also fixing the source-plane coordinates of the traced image pixels. Therefore, for the implementation of clustering used in `N15` the exact same source-plane discretization would have been used throughout the entire hyperparameter optimization. This is problematic, as a particular source-plane discretization favors a particular combination of hyperparameters, which in turn favor a particular lens model, leading the overall analysis to end up biased towards a specific parameter set. Thus, by fully randomizing the source-plane discretization, this biasing is removed.

## 5 DEMONSTRATION

The importance of these features is now demonstrated by analysing two simulated images with two different hyperparameter strategies: (i) the ‘basic implementation’, which uses only  $N_s$  and  $\lambda$  as hyperparameters and therefore omits the adaptive image and source features above. This closely resembles `N15`, with the only difference being that `N15` used a fixed value of  $N_s$ . (ii) The ‘adaptive implementation’, which uses all of the features above, giving up to 14 free hyperparameters if variance scaling is activated for both components. Both implementations use a `MultiNest` search to set the hyperparameters and the analysis is performed using the automated analysis pipeline introduced in the next section. Whilst readers are not yet aware of how this operates, the specific details are not important for providing a simple demonstration of the adaptive image and source features. However, it is worth noting that (i) this analysis does not use a fixed lens model, but determines the lens model via multiple `MultiNest` runs, and (ii) initial runs assume a simplified lens model (a *Sersic* light profile and *SIE* mass profile). The second point is of particular importance for demonstrating variance scaling.

Figs 6, 7, and 8 show the results of this analysis. These figures follow the same format, showing the reconstructed model image, residuals,  $\chi^2$  image (residuals divided by the baseline or scaled variances squared, equation (9)), and reconstructed source (Fig. 6) or baseline/scaled variance maps (Figs 7 and 8) for the basic implementation (top rows) and adaptive implementation (middle and bottom rows). The input lens models for both images are given in Table 1.

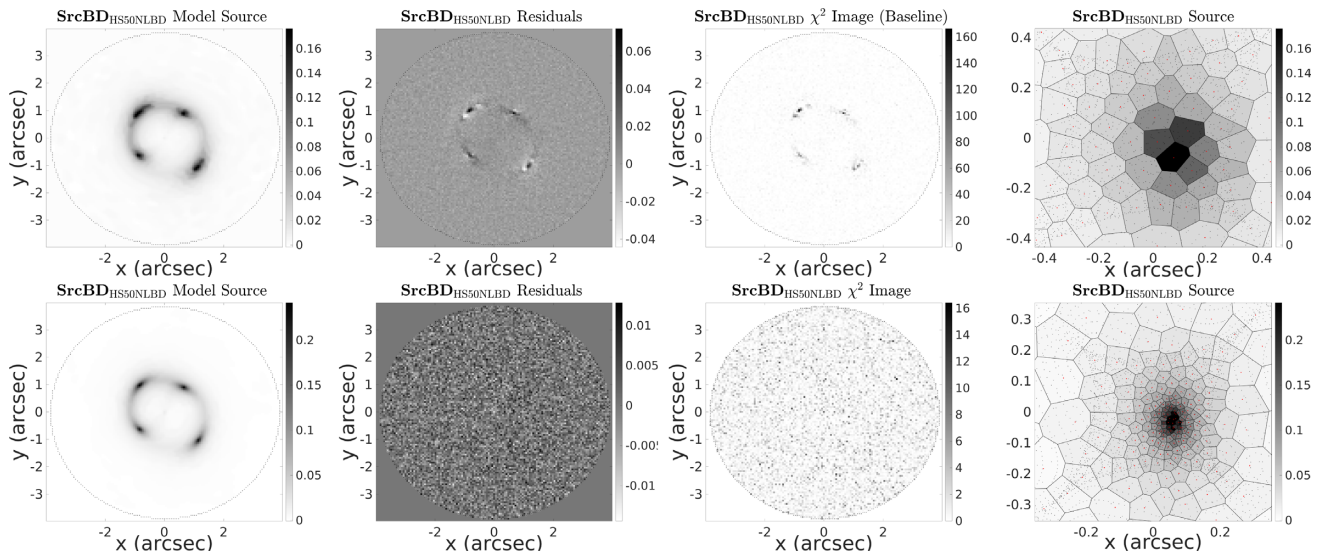
### 5.1 Source modeling – correct mass model

The first issue arises when modeling sources with a cuspy and rapidly changing light profile. This is illustrated in Fig. 6 using the image `SrcBDHS50NLBD`. Significant residuals can be seen in the basic implementation’s reconstruction of the image’s bright, high S/N pixels, which causes a small sub-set of image pixels to obtain large  $\chi_{\text{base}}^2$  values. This means that the overall  $\chi^2$  is constrained by only a small portion of data; approximately 5 per cent of image pixels contribute to over 90 per cent of the overall  $\chi^2$  value. Whilst this is clearly not ideal for any form of data analysis, Table 3 shows that the basic implementation still computes the correct lens model, suggesting that the poor residuals and skewed  $\chi^2$  distribution do not bias lens modeling. However, as discussed next, this holds only for simulated images where the mass model matches exactly the lens’s true mass profile. For real imaging this is not the case, thus these issues must be corrected.

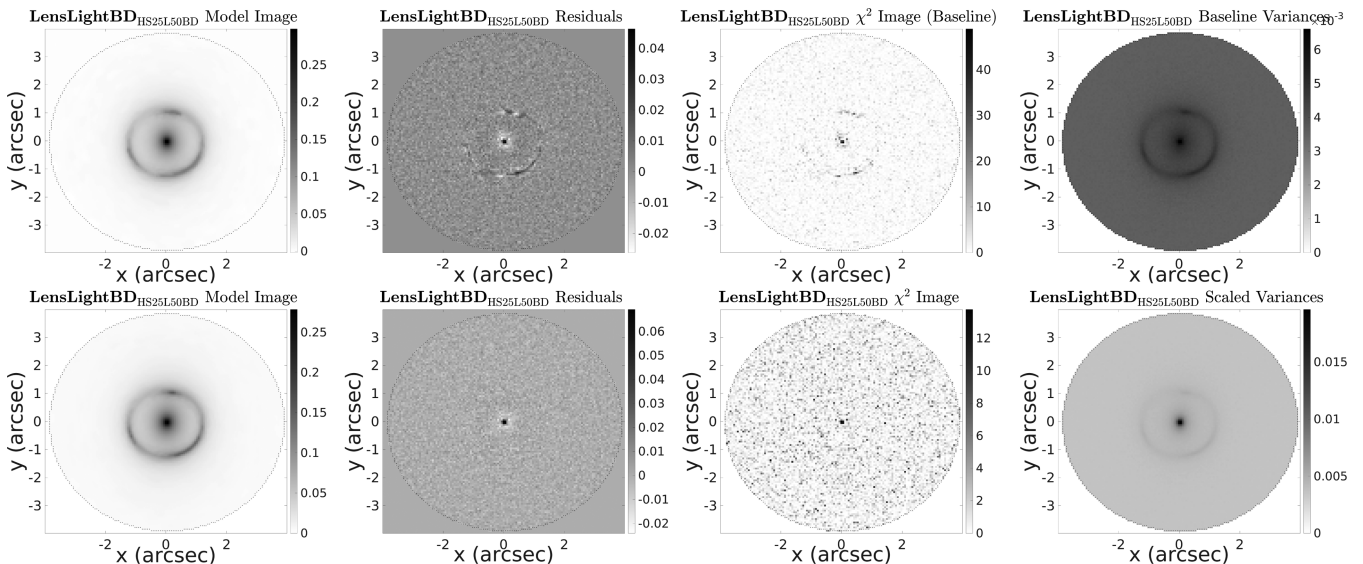
The issue arises because of the basic implementation’s source-plane pixelization and regularization scheme. By adapting to the magnification, pixels in the source’s central regions (where its intrinsic light profile is most rapidly changing) cover roughly the same area as those further out (where its light profile is flatter). However, although the pixels in these central regions are reconstructing a more rapidly declining light profile, they are regularized with the same  $\lambda$  as those further out. Therefore, when setting  $\lambda$ , the method has to compromise between a  $\lambda$  low enough to accurately reconstruct the source’s central regions but also high enough to correlate the source pixels further out. A compromised and intermediate value of  $\lambda$  is ultimately calculated. This ‘oversmooths’ the reconstructed source’s central light, producing the residuals seen in Fig. 6 (where the fact these are the highest S/N image pixels inflates their  $\chi^2$  contribution). The exterior regions of the source-plane (which map to background sky in the image) are simultaneously ‘underregularized’, in the sense that an unnecessarily high number of correlated source pixels are used to fit the regions of the image where the source is not present. Altogether, a reduced value of  $\ln\epsilon$  is inferred. The fact  $N_s$  is a free parameter for the basic implementation demonstrates that a higher source-plane resolution by itself cannot alleviate these problems.

The bottom row of Fig. 6 shows that the adaptive implementation removes this problem, producing nearly featureless residuals and a  $\chi^2$  image fully consistent with Gaussian noise. Table 3 confirms that this comes with an increase in  $\ln\epsilon$  and shows a reduction in the lens model parameter errors, suggesting the basic implementation’s skewed  $\chi^2$  distribution produced overestimated errors because it constrained the model with a small subset of the available image data. For this analysis, variance scaling was switched off, as the baseline  $\chi^2$  values in the lensed source model were all below 10.

In this example, both the improved pixelization and regularization scheme contribute to this. The congregation of smaller source pixels around the source’s central cusp of light provides a better spatial sampling of its rapidly changing surface brightness, meaning it reconstructs the source more accurately. Non-constant regularization ensures that each source pixel is subject to an appropriate level of regularization, regardless of its location in the source plane. The interplay between both of these features gives `AutoLens` complete freedom in how it reconstructs the source and ensures it finds the simplest solution possible (in a Bayesian sense). These solutions, by definition, use the fewest number of correlated source pixels and the adaptive implementation is indeed found to assume lower  $N_s$  values, offering significant gains in run-time efficiency. Model comparison with `AutoLens` therefore has no bias or preference to



**Figure 6.** The model source, residuals,  $\chi^2$  images, and source reconstructions for the analysis of the image  $\text{SrcBD}_{\text{HS50NLBulge}}$ , using either the basic implementation (top row, adaptive image and source analysis switched off) or adaptive implementation (bottom row, adaptive image and source analysis turned on). The resulting lens models are given by the first two rows of Table 3. The basic implementation can be seen to fit the image poorly, leaving noticeable residuals in each multiple image, which dominate the model's overall  $\chi^2$  value. The adaptive implementation gives featureless residuals and a Gaussian  $\chi^2$  image.



**Figure 7.** The model source, residuals,  $\chi^2$  images, and variance maps for the analysis of the image  $\text{SrcBD}_{\text{HS50NLBulge}}$ , using either the basic implementation (top row, adaptive image and source analysis switched off) or adaptive implementation with variance scaling turned off (middle row) or turned on (bottom row) and a lens model which incorrectly assumes that  $\alpha = 2.0$  (the true value is  $\alpha = 1.7$ ). The resulting lens models are given by the third, fourth, and fifth rows of Table 3. The basic implementation can be seen to fit the image poorly, leaving noticeable residuals in each multiple image, which dominate the model's overall  $\chi^2$  value. The adaptive implementation also gives poor residuals, given that the lens model is incorrect, however the  $\chi^2$  image shows a far less skewed distribution, which the right-hand panels show is the result of variance map scaling.

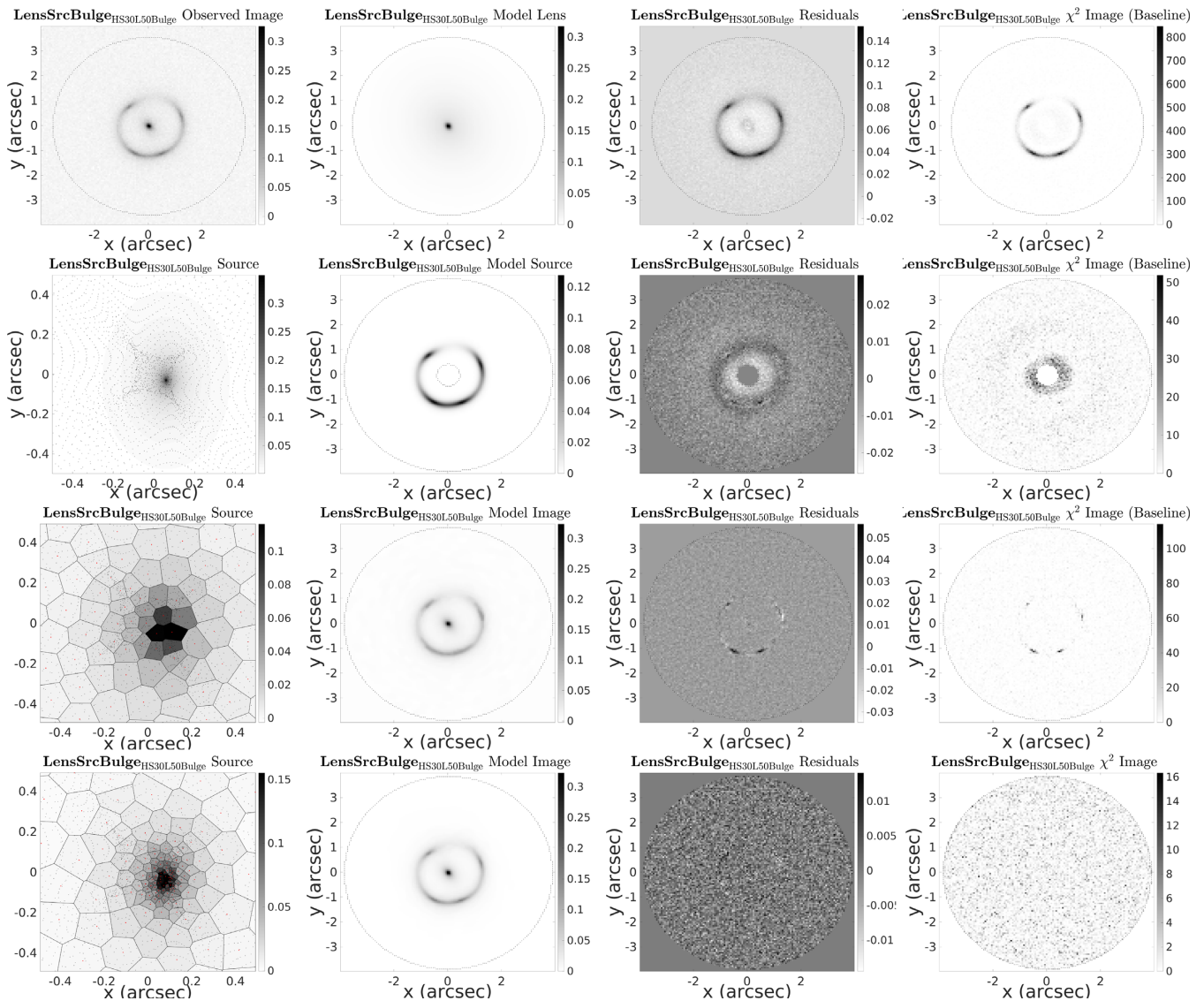
sources of a specific morphology or smoothness profile, which is not necessarily the case for approaches using a fixed pixelization (see S06).

## 5.2 Source modeling – incorrect mass model

Fig. 7 shows three independent analyses of the same image above,  $\text{SrcBD}_{\text{HS50NLBD}}$ , but where its *SPLE* mass model has been fixed to an incorrect power-law slope  $\alpha = 2.0$  (its input value is  $\alpha = 1.7$ ).

Three analyses are performed using the basic implementation (top row), the adaptive implementation with variance-scaling manually switched off (middle row), and with variance scaling on (bottom row). The basic implementation suffers the same issues as before; noticeable residuals and a skewed  $\chi^2$  distribution. However, the same is also now true for the adaptive implementation. This is because, even with the improved source pixelization and regularization, the mismatch between the assumed mass model and true lens profile means only a poor fit is obtainable. When variance scal-





**Figure 8.** The model images, residuals,  $\chi^2$  images, and variance maps for the analysis of the image **LensLightBD<sub>HS25L50BD</sub>**, fitted using a *Sersic* light model even though its true underlying profile is a *Sersic + Exp*. Two fits are shown, using the basic implementation (top row, adaptive image and source analysis switched off) and adaptive implementation (middle row, adaptive image and source analysis turned on). The adaptive implementation can be seen to increase the variances where the lens and source are located. The resulting lens models are given by the sixth to ninth rows of Table 3. The basic implementation can be seen to fit the image poorly, leaving residuals around both the lens and source, which dominate the model’s overall  $\chi^2$  value. The adaptive implementation gives poor residuals around the lens (as expected because the *Sersic* model cannot provide an accurate fit), however the source residuals are mostly removed and its  $\chi^2$  image shows a significantly less skewed distribution.

ing is turned on, the residuals are equally poor. After all, scaling the variances can’t change the fact that this mass model simply does not provide a good fit to the observed image. However, the  $\chi^2$  image shows fewer pixels with high  $\chi^2$  values and  $\chi^2$  values that are lower. As discussed for the previous issue, this is the more desirable solution, as it uses all of the data that is available to constrain the lens model. Table 3 shows this in turn provides the smallest errors and highest Bayesian evidence. However, the scaled  $\chi^2$  values shown in the bottom row of Fig. 7 are not fully consistent with Gaussian noise, because the limits of variance scaling prevent  $\chi^2_{\text{sca}}$  values to go below 10. The reasoning behind this set-up is discussed next in Section 5.4.

In this example, the issue could easily be fixed by allowing  $\alpha$  to be a free parameter. For real lenses, however, the mass model will

always (to some degree) be ‘incorrect’, because the mass models assumed during a lens analysis are a simplified representation of any galaxy’s true underlying mass distribution (see discussions by Brewer et al. 2012; Suyu 2012). Noticeable residuals and non-uniformly distributed  $\chi^2$  images (without variance scaling) are therefore commonplace when analysing real strong lenses, which will negatively impact lens modeling by overfitting a small fraction of the available imaging data. This can lead to overestimated parameter errors or biased parameter estimates. Instead, it is more desirable that the lens model is constrained using all of the data that is available by fitting the image in an equally weighted manner, especially once other uncertainties like a poor PSF-sampling are considered. Variance scaling ensures that this is the case and in conjunction with the adaptive source features offers a natural

**Table 3.** A sub-set of lens model parameters calculated for the images **SrcBD**<sub>HS50NLBD</sub> (top five rows) and **LensLightBD**<sub>HS40L60BD</sub> (bottom four rows), using either the ‘basic implementation’ (which omits the adaptive image and source analysis features described above) or the ‘adaptive implementation’ (which uses them). The input mass model parameters for the image **SrcBD**<sub>HS50NLBD</sub> are  $\theta_{\text{Ein}} = 1.2$  arcsec,  $q = 0.7$ , and  $\alpha = 1.7$  and for **LensLightBD**<sub>HS40L60BD</sub> are  $\theta_{\text{Ein}} = 1.0$  arcsec,  $q = 0.8$ , and  $\alpha = 2.05$ . In the top two rows, **SrcBD**<sub>HS50NLBD</sub> is fitted with a *SPLF* profile. The next three rows show the results of fitting it with an *S/E* profile, to demonstrate the effect of assuming an approximate mass model, with the fourth row showing results where variance scaling is manually switched off. The final four rows show fitting the image **LensLightBD**<sub>HS40L60BD</sub>, generated using a *Sersic + Exp + SPLF* model (see Table 1), with a *Sersic + SPLF* model, to demonstrate the effects of fitting a mismatched light profile. The second column displays whether the basic or adaptive implementation was used. Columns 3–6 give parameter estimates. The final column gives the value of Bayesian evidence that results from the analysis, which can be seen to increase for the adaptive implementation in all cases. These models are visualized in Figs 6, 7, and 8.

Image	Implementation	Parameters ( $3\sigma$ )	Parameters ( $1\sigma$ )	ln $\epsilon$
<b>SrcBD</b> <sub>HS50NLBD</sub>	<b>Basic</b>	$\theta_E = 1.2184^{+0.0672}_{-0.1011}$	$q = 0.6802^{+0.1296}_{-0.1527}$	$\alpha = 1.7342^{+0.0468}_{-0.0725}$
<b>SrcBD</b> <sub>HS50NLBD</sub>	<b>Adaptive</b>	$\theta_E = 1.2047^{+0.01520}_{-0.01700}$	$q = 0.7000^{+0.0279}_{-0.0261}$	$\alpha = 1.7082^{+0.0386}_{-0.0442}$
<b>SrcBD</b> <sub>HS50NLBD</sub>	<b>Basic</b>	$\theta_E = 1.3785^{+0.0408}_{-0.0456}$	$q = 0.4925^{+0.0535}_{-0.0575}$	$\alpha = 2.0$
<b>SrcBD</b> <sub>HS50NLBD</sub>	<b>Adaptive</b>	$\theta_E = 1.3899^{+0.0201}_{-0.0211}$	$q = 0.4818^{+0.0181}_{-0.0203}$	$\alpha = 2.0$
<b>SrcBD</b> <sub>HS50NLBD</sub>	<b>Adaptive</b>	$\theta_E = 1.3899^{+0.0201}_{-0.0211}$	$q = 0.4818^{+0.0181}_{-0.0203}$	$\alpha = 2.0$
<b>LensLightBD</b> <sub>HS50NLBD</sub>	<b>Basic</b>	$\eta_1 = 3.7567^{+0.0225}_{-0.0222}$	$R_1 = 4.5132^{+0.0021}_{-0.0083}$	$R_1 = 4.5132^{+0.0001}_{-0.0003}$
<b>LensLightBD</b> <sub>HS50NLBD</sub>	<b>Basic</b>	$\theta_E = 1.2275^{+0.0053}_{-0.0051}$	$q = 0.7333^{+0.0081}_{-0.0074}$	$\alpha = 2.0$
<b>LensLightBD</b> <sub>HS50NLBD</sub>	<b>Adaptive</b>	$\eta_1 = 2.3453^{+0.0169}_{-0.0169}$	$R_1 = 3.3554^{+0.1522}_{-0.1522}$	$R_1 = 3.3554^{+0.0784}_{-0.1193}$
<b>LensLightBD</b> <sub>HS50NLBD</sub>	<b>Adaptive</b>	$\theta_E = 1.2298^{+0.0021}_{-0.0013}$	$q = 0.7411^{+0.0031}_{-0.0064}$	$\alpha = 2.0$

means to test more complex mass models using Bayesian model comparison.

### 5.3 Simultaneous source and lens modeling – incorrect light model

Further issues arise when the lens and source are modeled simultaneously, in particular when the lens light subtraction leaves residuals. Fig. 8 illustrates this particular circumstance, showing the results of fitting the image **LensLightBD**<sub>HS25L50BD</sub> with a *Sersic* light profile, even though the simulated lens was generated using the more complex *Sersic + Exp* profile. As expected, the lens subtraction for both implementations leaves significant residuals, as the *Sersic* profile is simply unable to provide a good fit to the lens's more complex morphology. However, comparison of the residuals and  $\chi^2$  image of the two implementations show they differ in two ways: (i) the basic implementation leaves residuals after fitting the lensed source, whereas the adaptive implementation does not (the much smoother source morphology for this object means this is not due to the issue above); (ii) the  $\chi^2$  image for the adaptive implementation almost fully realizes the image's Gaussian noise, except for a few pixels in the centre of the image, whereas the basic implementation again suffers the skewed  $\chi^2$  distribution discussed above, however now also in the central pixels where the lens is located.

The fundamental problem here is that when simultaneously modeling both the lens and the source reconstruction cannot distinguish between lensed source flux and residual lens light. Problematically, it treats the latter as if they are part of the source, corrupting the image reconstruction and ultimately biasing the inferred lens model, as shown in Table 3. This issue impacts the image reconstruction in two ways, both of which the adaptive implementation was developed specifically to tackle.

The first is the impact of central residual light on the source reconstruction. The linear inversion will attempt to fit these pixels like any other, but fail to do so, given that they map to the exterior regions of the source-plane where all of the other traced image pixels map to the background sky (see Fig. 2). This by itself is acceptable, as the method shouldn't reconstruct these pixels as if they are part of the source. The problem, however, is regularization, as (in an analogous manner to the cuspy source above) these pixels lead the basic implementation to set a compromised higher value of  $\lambda$ , which leads the source reconstruction to be oversmoothed, producing the source residuals seen in Fig. 8 and inflating the lens model parameter errors as shown in Table 3.

The adaptive implementation does not suffer this issue because of luminosity-weighted regularization, which allows the source reconstruction to simultaneously smooth over the exterior regions of the source plane that map to the lens subtraction residuals (and background sky) whilst simultaneously fitting the detailed structure of the source with an appropriate and reduced level of regularization. Therefore, even in the presence of a poor lens subtraction the adaptive implementation can still fit the source accurately, as shown by the removal of source residuals in Fig. 8. When the method is able to smooth over the source plane's exterior regions with very high levels of regularization, the issues discussed in Section 4.5 related to the source reconstruction fitting residual flux in central image pixels are circumvented.

The second problem is also due to central image pixels, but instead how the lens light model fits and subtracts them. These pixels are the highest S/N pixels in the data (typically by a large margin) and therefore have the potential to overwhelm the model's  $\chi^2$  contribution if the lens subtraction is not perfect. When this occurs, the

light model concentrates its flux into these central regions so as to accurately fit only these high S/N image pixels, overconcentrating the inferred light profile and failing to give a global representation of the lens's morphology. This is shown by Table 3, where the basic implementation can be seen to go to a much higher value of  $n_1$  compared to the adaptive implementation.

For real lenses, no light subtraction will ever be perfect and many lenses will possess detailed structures (e.g. bars, dust lanes) a smoothly parametrized profile cannot fit completely. Thus, it is paramount this issue is removed from the lens analysis, which the bottom row of Fig. 8 shows is exactly what variance scaling achieves, by increasing the variances around the lens galaxy such that the  $\chi^2$  image reverts to being (almost) Gaussian, thereby again giving a global fit to the imaging data.

#### 5.4 Variance overscaling

The  $\chi^2$  images shown in Figs 7 and 8 for the adaptive implementation (with variance scaling on) are not fully consistent with the image's Gaussian noise. A small subset of pixels retaining values of  $\chi^2 \sim 10$  can be seen. This is because of the upper limits placed on the values of  $\omega_{\text{Src}}$  and  $\omega_{\text{Lens}}$ , which restricted their maximum values such that  $\chi_{\text{sca}}^2$  could not be scaled below 10. When these limits are not imposed, the  $\chi^2$  image becomes fully consistent with Gaussian noise. However, during testing of AutoLens, it emerged that giving these hyperparameters the freedom to go to arbitrarily high values lead to 'variance overscaling'.

Here, the scaled variances were increased to such large values that the effective S/N (and  $\chi_{\text{sca}}^2$  values) of their corresponding image pixels was  $\sim 0$ . This decreased  $\ln\epsilon$ , as the overall S/N of the observed image was reduced (recall the description of the Bayesian evidence in Section 4.6). However, a net increase in  $\ln\epsilon$  was still possible, because the source reconstruction and lens model could change so as to fit other regions of the image better. These solutions are not desirable, as the method is essentially ignoring the central regions of the lens and source galaxies in order to better fit more exterior regions of each. The limits on  $\omega_{\text{Src}}$  and  $\omega_{\text{Lens}}$  are therefore imposed to prevent this from happening, by ensuring the  $\chi_{\text{sca}}^2$  values of these image pixels cannot be reduced to 0. The target values of  $\chi_{\text{base}}^2 = 10$  are chosen to give a balance between reducing the  $\chi^2$  values such that the source reconstruction and lens light profiles do not overfit the central regions, whilst also ensuring that variance overscaling does not occur. For real lenses, it will be important to investigate the impact of changing these target scalings on the inferred lens model.

#### 5.5 Decomposed mass modeling

Equation (3), used for decomposed mass modeling, assumes that light perfectly traces mass, an assumption that will hold only approximately for real lenses. However, the lens's light profile is constrained by two aspects of a decomposed analysis: (i) the quality of the light subtraction and (ii) the decomposed mass model's source and image reconstructions. For decomposed mass modeling it is therefore desirable to give the lens's light matter distribution the freedom to deviate from the lens's light profile, if doing so improves the mass model. Variance scaling facilitates this, such that the  $\chi^2$  contribution of central image pixels can be downweighted to allow the light profile to deviate from its true profile. Doing so decreases  $\ln\epsilon$ , but could potentially produce a net  $\ln\epsilon$  increase by improving the mass model. Thus, AutoLens does not strictly assume that light traces mass and is able to deviate from this assumption within a Bayesian context. Comparison to models assuming a total-mass

profile, like the *SPLE*, can offer insight into whether this is occurring and to what degree. However, the simulated images used in this work assume that light traces mass, thus this is not tested explicitly here.

## 6 PIPELINE AUTOMATION

AutoLens is a multiphase automated analysis pipeline, designed with scalability to very large lens data sets in mind. Each phase involves a separate MultiNest search, but generates initial points from priors derived from the highest likelihood regions of the previous phase's posterior distributions. Many tasks required to set up AutoLens are performed automatically between phases, most notably optimizing the hyperparameters of the adaptive image and source reconstruction features. Fig. 9 provides a flow diagram of AutoLens, showing the different phases used throughout the automated analysis framework. The figure shows that the pipeline incorporates three parallel routes, a particular route being chosen depending on whether a singular total-mass profile, cored total-mass profile or decomposed mass profile is being fitted. Also shown is the primary aim of each phase and lens model that is fitted.

### 6.1 Pipeline phase linking

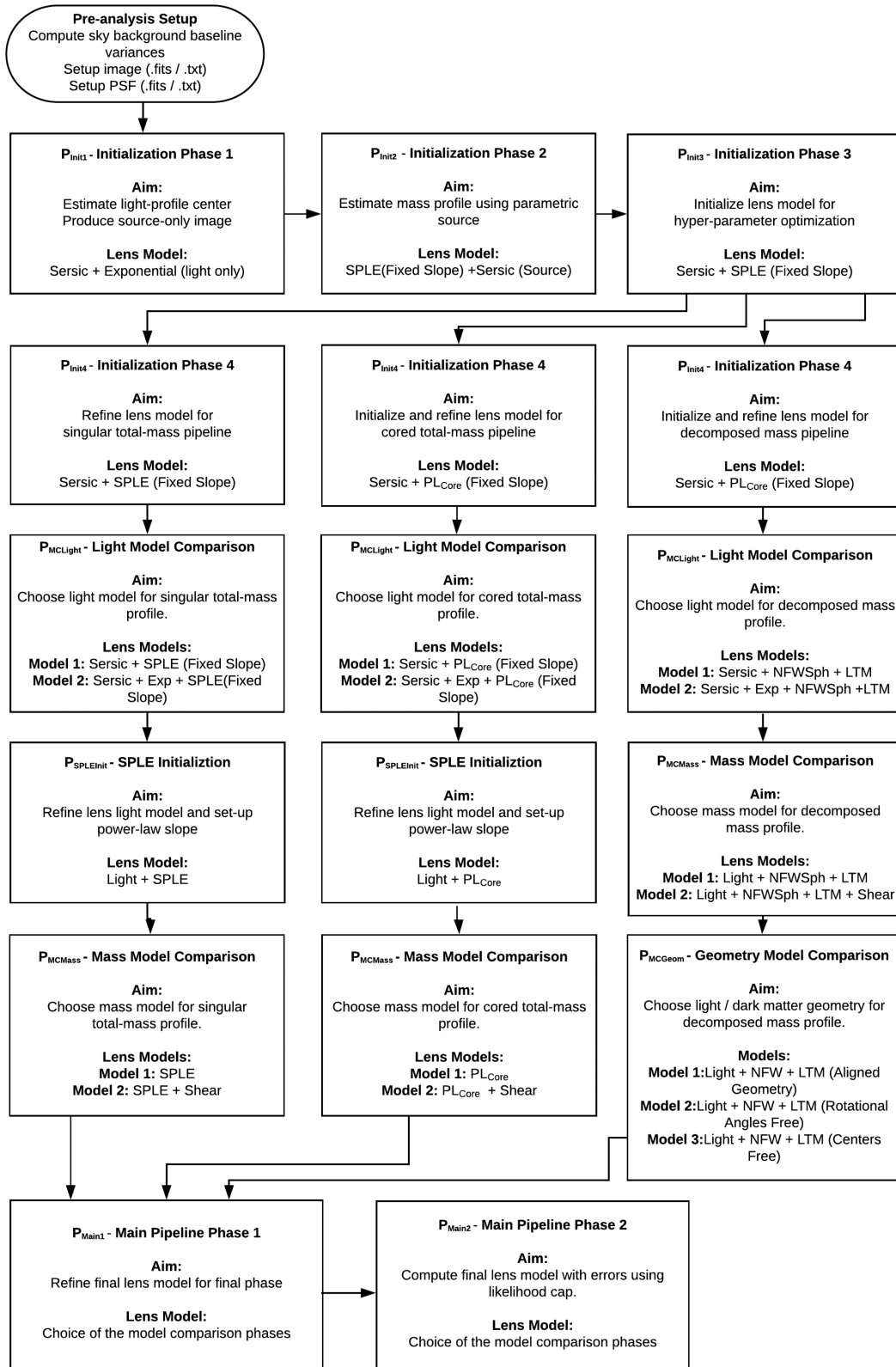
In the initial phases of the automated analysis pipeline broad uniform priors are assigned to all lens model parameters, since they have no expectation values computed for them. However, once estimated, this information is used to set that parameter's priors in the subsequent phase of the pipeline. The motivation behind this is that the more complex lens models used by AutoLens have a large and highly degenerate non-linear parameter space within which accurate sampling and location of the global maximum is unattainable if broad priors are assumed on all parameters simultaneously. Therefore, the initial phases of AutoLens accurately estimate a less complex lens model, with later phases using these results to gradually increase the lens model complexity whilst ensuring the non-linear parameter space is sampled accurately. To accompany this, the image and source reconstructions also gradually adapt to the properties of the lens and source being analyzed, facilitating further the fitting of more complex lens models.

The lens models used in different phases of the pipeline are linked via Gaussian priors centered on each parameter's high-likelihood regions, as estimated in the previous phase. Although it is possible to choose narrow priors to expedite the exploration of parameter space, the prior scaling values chosen for this work sample very broad regions, ensuring that no results are simply a consequence of overly restrictive priors (but offering enough information to ensure parameter space is sampled robustly). Nevertheless, the freedom offered by the ability to scale the degree of sampling will be key to scaling the method up to large lens samples in the future. Appendix C gives a full description of how each phase is linked, along with the priors used to link every mass and light model from one phase to the next.

### 6.2 Pipeline initialization

Initialization involves four automated tasks, the aim of which is to compute an accurate light (*Sersic*) and mass (*SPLE*  $\alpha = 2.2$ ) model alongside a robust initialization of the hyperparameters.

(i) **P<sub>init1</sub>** – Lens Light Subtraction – This phase fits a *Sersic* + *Exponential* light profile to the observed strong lens with the lensing



**Figure 9.** A flow diagram of the AutoLens automated analysis framework, which is described in detail throughout Section 6. The method begins with four initialization phases, which aim to accurately compute an accurate lens model alongside a robust initialization of the hyperparameters. Before the fourth initialization phase, the pipeline splits into one of three routes, depending on whether a singular total-mass, cored total-mass, or decomposed mass profile is desired. Each of these pipelines then runs a set of model comparison phases that set the complexity of the lens light profile, mass profile, and light and dark matter geometry. The main analysis pipeline then begins which computes high precision estimates of every lens model parameter. Images without a lens light component use a reduced version of this pipeline described in Section 6.6.

analysis turned off. Shown in Fig. 10 (top row), the resulting light model and subtraction are poor. However, they are sufficient for the centre of the lens’s light profile to be estimated and to provide a lens-subtracted image for the next phase. No other information is used from this phase.

(ii) **P<sub>Init2</sub>** – Parametric Source Model – Initializing a lens model with a pixelized source reconstruction is a surprisingly non-trivial task. This is because of the over-/underfit solutions discussed in N15, which reconstruct the source as a demagnified version of the lensed image. In general, these solutions have a lower  $\ln\epsilon$  than the input model, but occupy much larger volumes of non-linear parameter, causing `MultiNest` to get stuck in their local maxima. To circumvent this issue, a parametric source profile is used first, for which these unwanted and unphysical solutions do not exist. This allows an initial estimate of the mass profile to be computed in a completely general way, which can then be used in the next phase to prevent `MultiNest` from sampling these over-/underfit solutions. Thus, this phase fits the *SPLE* ( $\alpha = 2.2$ ) + *Sersic* (source) model to the image, omitting lens light modeling and masking the central regions where the poor lens subtraction leaves residuals. This is illustrated in Fig. 10 (top-middle row), where the lensed source models can be seen to fit the observed image well enough to ensure the mass model has been estimated reliably. Finally, the positional image pixels and threshold value are updated (see Appendix A).

(iii) **P<sub>Init3</sub>** – Initial Lens Model and Hyper-Parameters – This phase now uses a pixelized source-plane, fitting the *Sersic* + *SPLE* ( $\alpha = 2.2$ ) model to the image ( $\alpha = 2.2$  to remove the biases described in Section 4.5). Restrictive priors are placed on the mass model parameters, ensuring that the method does not sample the unphysical solutions corresponding to a demagnified image reconstruction. The baseline variance-map is used, source-light adaptation is turned off, and a constant regularization scheme is applied with  $\lambda$  included as a free parameter in the non-linear search. This is illustrated in Fig. 10 (middle row), where both the light and source models can be seen to fit the observed image reasonably well, but with the residuals and  $\chi^2$  image showing the issues discussed previously (as is expected given the adaptive source/image features are not implemented yet). However, the model is of sufficient accuracy to initialize the hyperparameters, which is performed before phase **P<sub>Init4</sub>**. The positional image pixels and threshold are again updated.

(iv) **P<sub>Init4</sub>** – Model Refinement – This phase fits one of three models: (i) the *Sersic* + *SPLE* ( $\alpha = 2.2$ ) model (with priors relaxed compared to the previous phase); (ii) the *Sersic* + *PL<sub>Core</sub>* ( $\alpha = 2.0$ ) model (using priors from the previous phase and a broad prior on the core radius) or (iii) the *Light* + *NFWSph* model (with broad uniform priors). The models correspond to the singular total mass pipeline, cored total mass pipeline, or decomposed mass pipeline, respectively. This phase benefits from the adaptive image and source features following the hyperparameter initialization of the previous phase. The aim of this phase is to refine the lens model and ensure an excellent optimization of the hyperparameters for the model comparison phases. This is illustrated in Fig. 10 (bottom row), where an accurate model for both the light and mass components is shown, alongside much improved residuals and  $\chi^2$  values. Following this phase, hyperparameters are re-optimized and positional image pixels are recomputed.

### 6.3 Bayesian model comparison

The next stage of the pipeline ‘builds’ the lens model, by performing Bayesian model comparison. A number of publications have already detailed the hierarchical Bayesian formalism of pixel-based lens

analysis methods like `AutoLens` (e.g. S06, Vegetti & Koopmans 2009a; Tagore & Keeton 2014). Hence, only a brief overview is given here. Bayes’s theorem is given by

$$P(m|d, M) = P(m|M) \frac{P(d|m, M)}{P(d|M)}, \quad (22)$$

where  $d$  is the data and  $m$  is a particular realization of the overall model  $M$ , which comprises all linear source parameters, hyperparameters, lens model parameters, and  $\mathbf{H}_\Lambda$ .  $P(d|m, M)$  gives the likelihood,  $P(m|M)$  the priors on model parameters, and  $P(m|d, M)$  the posterior probability. The Bayesian evidence is given by  $P(d|M)$  (which readers should note is different to the Bayesian evidence given by equation (10), which ranks the source reconstruction) and can be obtained by integrating over all possible models  $m$  in the set of models  $M$  as

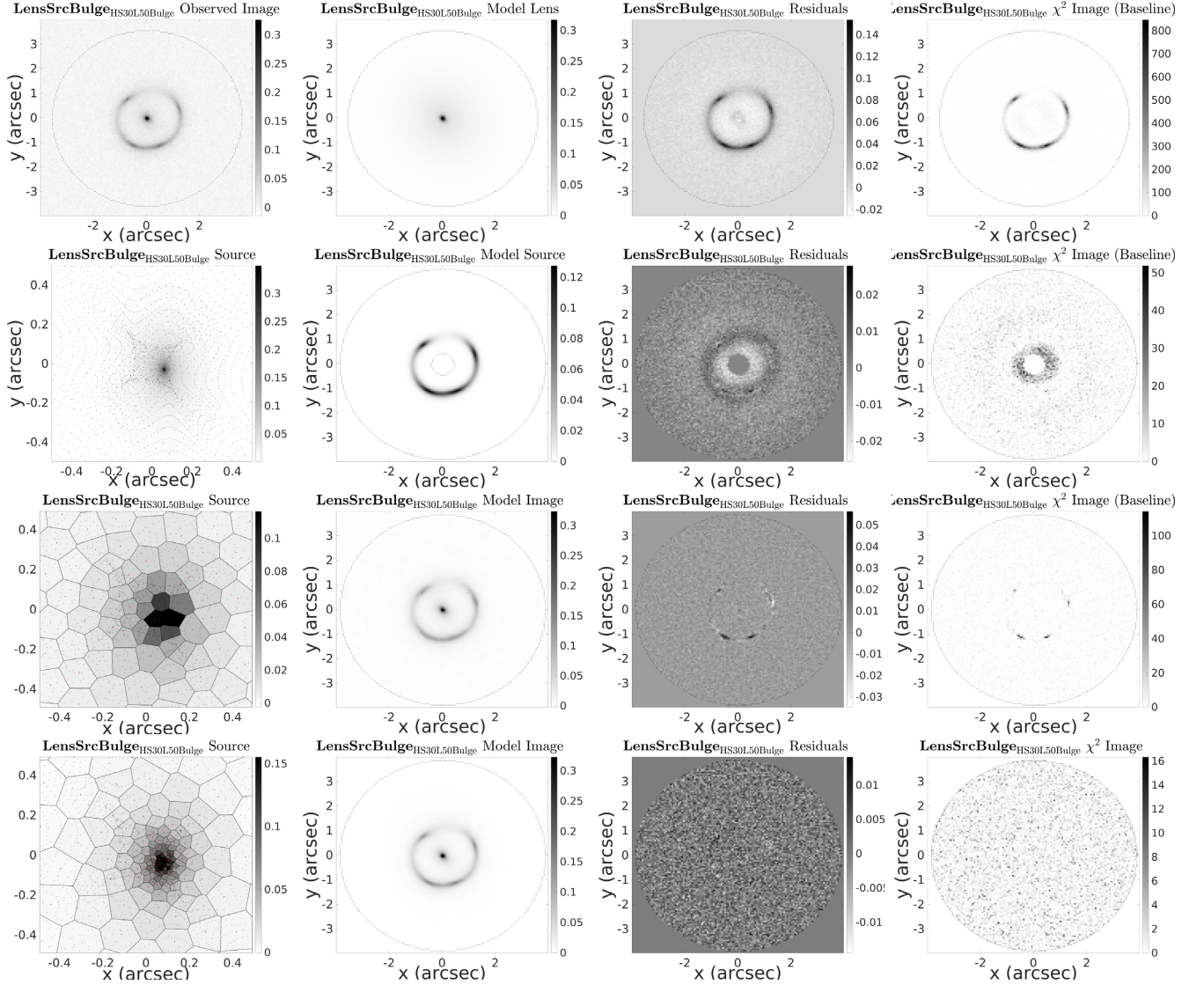
$$P(d|M) = \int P(m|M)P(d|m, M)dm. \quad (23)$$

This expression has the principle of Occam’s razor built into it, whereby overly complex models are penalized if they do not give a justifiably improved fit to the observed data. Thus, maximizing this quantity objectively chooses the model which best fits the data without being overly complex. The ratio of the evidence of two models (e.g.  $P(d|M_1)/P(d|M_2)$ ) gives their Bayes factor and to accept a more complex model, a Bayes factor greater than twenty is required (considered ‘strong’ evidence in Bayesian statistics). The integral given in equation (23) is estimated by `MultiNest` and therefore is a natural by-product of `AutoLens`’s analysis. It should be noted that during the model comparison phases, `MultiNest`’s non-linear parameter space comprises a subset of  $M$ ’s parameters (because certain hyperparameters are left fixed) and the evidence is estimated for only these parameters. The parameters that are omitted have no impact on the value of evidence and are omitted for efficiency. In practice, model comparison with `AutoLens` simply amounts to fitting different light or mass models at various stages of the pipeline and choosing a more complex model when the evidence increases over the previous (simpler) model by a threshold value, which is set to 20.

When transitioning to a more complex lens model, the set-up of the adaptive image and source features may be problematic. These features adapt to a specific source morphology and suppress the  $\chi^2$  contribution of poorly fitted image pixels by increasing their variances. If a more complex model changes the reconstructed source’s morphology or accurately fits pixels that previously had their variances increased, there is a risk that using the hyper-parameters of the simpler model may prevent the more complex model from making a sufficiently high gain in  $\ln\epsilon$  to be correctly favoured by model comparison. Therefore, the relevant hyperparameters are included as free parameters in each model comparison’s non-linear search. This allows the model to change the source reconstruction and undo the suppression of image pixels  $\chi^2$  values if and when the new lens model begins to accurately fit them, which in turn allows the correct  $\ln\epsilon$  values to be sampled.

The model comparison phases (and an intermediate linking phase) follow the initialization phases above and are (noting that following all the phases below is a hyperparameter re-optimization):

(i) **P<sub>MCLight</sub>** – Light Model – This phase chooses the light model. For the singular total mass pipeline an *SPLE* ( $\alpha = 2.2$ ) mass model is used, for a cored total-mass pipeline a *PL<sub>Core</sub>* ( $\alpha = 2.0$ ) model and decomposed pipeline an *NFWSph* model. The mass model’s parameters are initialized using the results of phase **P<sub>Init4</sub>**. The hyperparameters  $\lambda_{\text{Src}}$ ,  $\omega_{\text{Lens}}$ , and  $\omega_{\text{Src}}$  are not fixed. First, the *Sersic* +



**Figure 10.** A demonstration of the analysis performed in phases  $\mathbf{P}_{\text{Init}1}$  (top row),  $\mathbf{P}_{\text{Init}2}$  (top-middle row),  $\mathbf{P}_{\text{Init}3}$  (bottom-middle row), and  $\mathbf{P}_{\text{Init}4}$  (bottom row) using the image  $\text{LensSrcBulge}_{\text{HS30L50Bulge}}$ . The left-hand panels show the observed image and mask (top-row) and source reconstructions of each phase (remaining rows), the middle-left panels the reconstructed model images, the middle-right panels the residuals and right most panels the  $\chi^2$  images (equation 9). *top row* ( $\mathbf{P}_{\text{Init}1}$ ) – A *Sersic + Exp* light profile is fitted to the observed image with lensing analysis turned off. The light model gives a poor overall fit to the observed image but gives an accurate estimate of  $x_1$  and  $y_1$  and reveals the lensed source for the next phase. *top-middle row* ( $\mathbf{P}_{\text{Init}2}$ ) – A *SPLE* ( $\alpha = 2.2$ ) + *Sersic* model is fitted to the lens subtracted image generated in the previous phase, where the source is modeled using a smoothly parametrized profile which is sampled simultaneously with the mass model. This gives a robust initialization of the mass model’s parameters for the next phase and initializes the positional image pixels. *bottom-middle row* ( $\mathbf{P}_{\text{Init}3}$ ) – A *Sersic + SPLE* ( $\alpha = 2.2$ ) model is fitted to the observed image with the adaptive image and source analysis features switched off and  $\lambda$  included as a free parameter. The light and source models give a reasonable fit to the image, but the residuals and  $\chi^2$  image show the issues demonstrated in Section 5 are present. Nevertheless, the model is sufficient to optimize the hyperparameters so that the next phase can use the adaptive image and source features. *bottom row* ( $\mathbf{P}_{\text{Init}4}$ ) – A *Sersic + SPLE* ( $\alpha = 2.2$ ) model is again fitted to the observed image but now with the adaptive image and source features. The light and source models give an accurate fit and the residuals and  $\chi^2$  are improved from the previous phase (and will be further improved after this phase’s hyperparameter optimization).

*Mass* model is refitted, to compute the Bayesian evidence now the hyperparameters have been re-optimized. This is compared to the *Sersic + Exp + Mass* model, thus determining whether a two-component light profile is required. For the total-mass pipelines the light model does not contribute to the mass model, whereas for the decomposed pipeline it does.

The simplified mass profiles used during this phase often leave residuals in the lensed source. The two-component light model was found to make gains in Bayesian evidence by subtracting these

source residuals. This behaviour is undesirable, therefore the upper limit on the hyperparameter  $\omega_{\text{Src}}$  is increased to a target value  $\chi_{\text{sca}}^2 = 1.0$  (as opposed to the value  $\chi_{\text{sca}}^2 = 10.0$  used everywhere else in the pipeline). This ensures these residuals are not fitted by the light profile, as they are down-weighted by variance scaling. After this phase, the method reverts to a target value  $\chi_{\text{sca}}^2 = 10.0$ .

(ii)  $\mathbf{P}_{\text{SPLEinit}}$  – *SPLE Initialization* – For the total-mass pipelines, if the lens’s slope deviates from the fixed value previously assumed for  $\alpha$  it is beneficial to refine the mass model and adaptive source

features to reflect. This phase does exactly this, by fitting the *SPL*E or *PL*Core mass model with free  $\alpha$ , alongside the light profile just chosen. If variance scaling is on,  $\omega_{\text{Lens}}$  and  $\omega_{\text{Src}}$  are included as free parameters, to ensure the new light profile (if chosen) and mass profile are able to fit regions of the image that may have previously had their  $\chi^2$  suppressed.

(iii) **P<sub>MCMass</sub> – Mass Model** – For the total-mass pipelines, the most probable light model computed in **P<sub>SPL</sub>E<sub>Init</sub>** is subtracted from the observed image to create a source-only image. This image is then fitted to choose the mass model with light modeling turned off for computational speed. For the decomposed mass pipeline, the original image is used with light modeling turned on. The hyperparameters  $\lambda_{\text{Src}}$  and  $\omega_{\text{Src}}$  (and  $\omega_{\text{Lens}}$  for the decomposed pipeline) are not fixed. First, the *SPL*E, *PL*Core, or *Light + NFW*Sph model is fitted, which is subsequently compared to the same model with the inclusion of a *Shear* term, thus determining if an external shear component is necessary.

(iv) **P<sub>MCGeom</sub> – Light/Dark Matter Geometry** – For the decomposed model pipeline this phase determines the light and dark matter geometries, where the *Light* model is the light model chosen in the phase **P<sub>MCL</sub>ight**. The hyperparameters  $\lambda_{\text{Src}}$ ,  $\omega_{\text{Src}}$ , and  $\omega_{\text{Lens}}$  are not fixed. First, the *Light + NFW* model is fitted assuming geometric alignment ( $\theta_1 = \theta_d$ ,  $x_1 = x_d$ ,  $y_1 = y_d$ ). This is compared to a *Light + NFW* model that allows rotational misalignment ( $\theta_1$  and  $\theta_d$  both free) but retains the assumption of a common center. A third comparison is then performed, which allows the centers to vary ( $x_1, y_1, x_d$ , and  $y_d$  all free) and assuming the rotational alignment determined from the previous result.

## 6.4 Separate pipelines

It is worth noting the importance of using separate pipelines for different mass profiles. For example, attempts to fit images generated using a cored mass profile with a singular mass profile were found to give poor results, because lensed source features specific to a cored mass model (radial arcs and a central image) cannot be replicated accurately by a singular model. On the other hand, if a cored model is wrongly assumed, it will wrongly include some of the lens’s light in the source reconstruction, biasing the lens model. Singular total-mass and decomposed mass models were generally found to produce lensed sources with the same overall structure. However, geometric offsets between a decomposed model’s light and dark matter components produce unique features a singular model cannot replicate (see Section 7).

Therefore, it is important to use separate pipelines for models that produce different and unique lensed source features. To choose between these models, Bayesian model comparison is again used, now using the final results of each pipeline. Whilst this is computationally expensive, the splitting of each pipeline means they can run in parallel.

## 6.5 Main pipeline

Following the initialization and model comparison, phase 1 of the main pipeline begins, using the lens model previously chosen. The lens’s light and mass are modeled simultaneously, with all initial parameters sampled via Gaussian priors initialized using the previous phase’s results. The main pipeline comprises two phases: (i) the lens model is estimated and the hyperparameters are re-optimized; (ii) the same model is estimated again, but using more computationally intensive settings providing more accurate parameter estimates and errors. A third phase is possible, which includes all hyperparameters

in the non-linear search. However, this final phase is computationally expensive and was found to have negligible influence on the inferred lens model.

## 6.6 No lens light

**AutoLens** uses a simplified pipeline for modeling images where there is no lens light component. This applies the phases **P<sub>Init</sub>2**, **P<sub>Init</sub>3**, and **P<sub>MCMass</sub>** without lens light modeling. The main pipeline then runs with only the mass model, using only phase 1.

## 6.7 Pipeline settings

The analysis has a number of settings that are changed throughout the pipeline. In the early phases these are chosen to give a fast run-time, since only an estimate of which models reasonably fit the data is necessary. More computationally intensive settings are used later on once high accuracy parameter, error, and Bayesian evidence estimation is required. These are shown in Table 4, where settings like the PSF trimming and image subgridding are altered to give fast computation early on and high-accuracy later. Early hyperparameter initialization is also restricted to lower values of  $N_s$ , given how this drives the computational run-time (provided doing so does not decrease  $\ln e$  significantly). The set-up of **MultiNest** is also changed, such that earlier pipeline phases converge more quickly towards a solution, with more thorough sampling employed in later phases. These settings can be altered to allow **AutoLens** to scale up to larger lens samples whilst keeping the overall run-time feasible.

## 6.8 Parameter estimation

Unless otherwise stated, estimates for each parameter are the median of their 2-D marginalized posterior probability distribution, which is calculated by weighting each accepted sample in **MultiNest** by its sampling probability.<sup>1</sup> This set of parameters then constitutes what is referred to as the ‘most probable’ lens model and the set that corresponds to the maximum overall likelihood gives the ‘most likely’ lens model. Errors correspond to the  $3\sigma$  confidence bounds on each parameter’s marginalized 1-D posterior distribution function (PDF) unless otherwise stated and 2D PDF contours are calculated by marginalizing over all other parameters. The results presented in this work use only the **MultiNest** samples generated from the second phase of the main pipeline unless stated otherwise (or the final phase of the no lens light pipeline).

## 6.9 Stopping criteria

A **MultiNest** search stops when its estimate of the global posterior log-evidence exceeds a user-defined threshold accuracy, which corresponds to the point where all active points have roughly the same likelihood values. However, as shown in N15, the changing discretization of the source pixelization leads to a noisy and non-smooth likelihood function in non-linear parameter space. Therefore, whilst **MultiNest** does an adequate job sampling this, its stopping criterion is ill-defined, as it tries to fully map out all of the noise in parameter space. In practice this means that once

<sup>1</sup>The sampling probability is a quantity output by **MultiNest** corresponding to the sample prior mass multiplied by likelihood and normalized by the evidence.

**Table 4.** Settings used in each phase of the AutoLens pipeline described in Section 6. Column 1 shows the pipeline phase. Column 2 shows whether the light and mass are modeled independently or simultaneously. Column 3 shows the likelihood tolerance on the selection of  $N_s$ , whereby the hyperparameter with the lowest value of  $N_s$  is chosen within this likelihood range to give a faster overall run-time. Column 4 shows the degree of image subgridding used by the analysis. Column 5 shows the fraction of the PSF that is trimmed about its centre. Columns 6–8 show the MultiNest settings for the number of live points, tolerance, and reduction factor. Column 9 shows the stopping factor discussed in Section 6.9, which determines at what acceptance rate MultiNest terminates. Settings are changed to trade off fast run time in the early phases to high precision accuracy in the later phases. The final row shows the settings used when optimizing the hyperparameters between each phase.

Phase	Method	$N_s$ likelihood tolerance	$\alpha_{x,y}$ subgrid degree	PSF trim	MultiNest live points	MultiNest tolerance	MultiNest reduction factor	Stopping factor
<b>PL</b> <sub>Init1</sub>	Light Only	N/A	N/A	10 per cent	50	1000.	0.2	N/A
<b>PL</b> <sub>Init2</sub>	Mass + Source	N/A	$2 \times 2$	10 per cent	200	100.0	0.2	5.0
<b>PL</b> <sub>Init3</sub>	Light + Mass	15.0	$4 \times 4$	10 per cent	125	50.	0.2	2.0
<b>PL</b> <sub>Init4</sub>	Light + Mass	15.0	$4 \times 4$	10 per cent	125	50.	0.2	2.0
<b>PL</b> <sub>MCLight</sub>	Light + Mass	8.0	$8 \times 8$	0 per cent	150	1.0	0.2	2.0
<b>PL</b> <sub>SPLEInit</sub>	Light + Mass	8.0	$8 \times 8$	0 per cent	150	1.0	0.2	2.0
<b>PL</b> <sub>MCMass</sub>	Mass Only	5.0	$8 \times 8$	0 per cent	125	1.0	0.2	2.0
<b>PL</b> <sub>MCGeom</sub>	Light + Mass	8.0	$8 \times 8$	0 per cent	175	1.0	0.2	2.0
Phase One	Light + Mass	2.0	$8 \times 8$	0 per cent	150	0.8	0.15	2.0
Phase Two	Light + Mass	0.5	$8 \times 8$	0 per cent	400	0.8	0.1	4.0
Hyper	Fixed	Next	Next	Next	150	100.	0.2	2.0

the lens model is estimated accurately there are one or two active MultiNest points with anomalously high likelihood values (due to discretization noise) that prevent MultiNest from stopping. This leads MultiNest's acceptance rate to plummet, as it can no longer maintain a high acceptance rate by further reducing the lens model's iso-density contours around the high-likelihood regions. At this point, any further increase in likelihood (or the Bayesian evidence) comes from randomly producing a 'good' source-plane discretization, information which is of no practical use in terms of actually constraining the lens model.

Therefore, to circumvent this issue and offer a meaningful stopping criterion, MultiNest is automatically terminated once its acceptance rate falls below the target sampling rate divided by a user-specified value, which are both given in Table 4. This division value is never below 2, ensuring all phases end only when noise in the parameter space is all that is left being fitted (MultiNest consistently maintains its target sampling efficiency otherwise).

In the final analysis phase a different approach is used. Instead, a 'likelihood cap' is imposed, such that any samples with a higher likelihood are reduced to this cap's value. This cap is calculated by taking the previous phase's most likely lens model and hyperparameter set and computing the mean likelihood of 100 different source reconstructions, corresponding to the value above which MultiNest begins fitting noise (relying on the fact that the lens model is unchanged from the previous phase and already estimated accurately). MultiNest then runs until all active points hit this value, thus preventing it from fitting the parameter space's noise. This is important for ensuring the errors of the most probable lens model are estimated accurately, as noise-fitting can bias this towards a few points that gain anomalously high likelihoods due to favourable discretizations. A low value for the likelihood cap will only lead the method to overestimate parameter errors, given that it exclusively trims the highest likelihood regions of parameter space.

## 7 RESULTS

This section presents AutoLens's automated analysis of the full suite of simulated images. Given the large library of results, this section focuses on only a subset of lens model parameters that best summarize the accuracy of each analysis. For the *Sersic* light

component of a lens model, the effective radius  $R_1$ , sersic index  $n_1$  and axis ratio  $q_1$  are used, with multicomponent light models using  $n_{11}$ ,  $R_{11}$ ,  $q_{11}$ ,  $R_{12}$ , and  $q_{12}$ . For *SPLE* mass components the Einstein radius  $\theta_{\text{Ein}}$ , axis ratio  $q$  and density slope  $\alpha$  are used, with the core radius  $S$  included for the *PL*<sub>Core</sub> model and shear orientation  $\theta_{\text{sh}}$  and magnitude  $\gamma_{\text{sh}}$  for a *Shear* component. The *NFW* model is summarized with its normalization  $\kappa_d$  and axis-ratio  $q_d$  whereas a light profile's mass component uses instead its mass-to-light ratio  $\Psi_1$ . The geometry of light and dark matter is also investigated using their centroids and rotation angles  $x_1$ ,  $y_1$ ,  $\theta_1$ ,  $x_d$ ,  $y_d$ , and  $\theta_d$ .

To ease the reader's comparison to the input values, all results are presented as the difference between the estimated value and simulated lens's true input value,  $\Delta P = P_{\text{true}} - P_{\text{model}}$ , where  $P$  is a given parameter. The parameters that have been omitted are those that are generally 'easy' to estimate and share no degeneracies with the other parameters (e.g.  $x_1$ ,  $x$ ,  $\theta$ ). Again, for brevity, only a sub-set of images for each analysis is presented, choosing results that show the general trends and exceptions.

### 7.1 Source-only

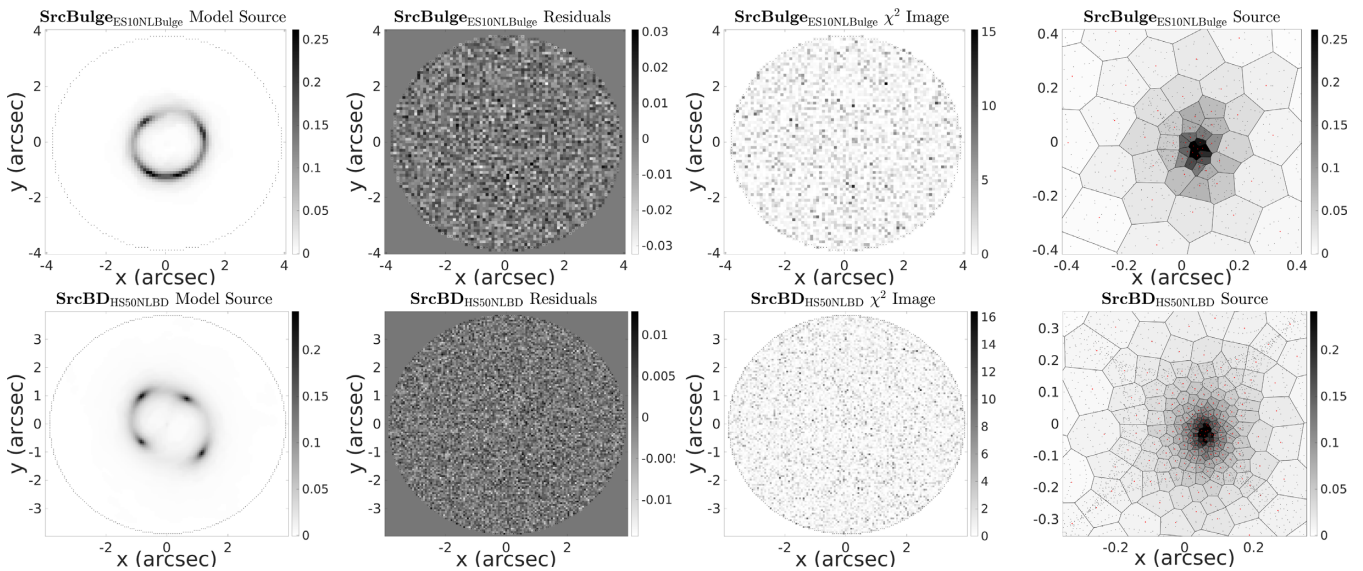
The source-only simulation suite consists of four unique lens and source models, each of which is used to generate six images, three at Hubble resolution with a source  $S/N = 50, 30, 10$  and three at Euclid resolution at  $S/N = 50, 30, 10$ , giving a total of 24 images. The analysis of each image uses the reduced pipeline for objects without a lens light component, therefore also omitting the  $\omega_{\text{Lens}}$  hyperparameters. The shear model comparison phase is also omitted for brevity. The results are summarized using the mismatch parameters  $\Delta\theta_{\text{Ein}}$ ,  $\Delta q$  and  $\Delta\alpha$  in Table 5, where all parameters are correctly estimated within  $3\sigma$  confidence. Fig. 11 shows the observed image, model image, residuals,  $\chi^2$  images and source reconstructions for the images **SrcBulge**<sub>ES10NLBulge</sub> and **SrcBD**<sub>HS50NLFlat</sub>, which cover different lens models, source morphologies and image resolutions and  $S/N$  ratios. The residuals are nearly featureless and  $\chi^2$  image realizes the image's noise, as Section 5 discussed is the desired solution.

Fig. 12 shows the 2-D PDFs of  $\Delta\alpha$  against  $\Delta\theta_{\text{Ein}}$  for the Hubble resolution (first panel) and Euclid resolution (third panel) images of the  $S/N = 30$  images of the **SrcBulge**, **SrcDisk**, and **SrcMulti**



**Table 5.** Results of fitting the **SreBulge**, **SreBD**, **SreDisk**, and **SreMulti** images using the (no lens light) automated analysis pipeline, corresponding to results generated at end of phase one of the main pipeline. Each image's name is given in the first column and the mass model fitted in the second column. The third to sixth columns show parameter estimates, where each parameter is offset by  $\Delta P = P_{\text{true}} - P_{\text{model}}$ , such that zero corresponds to the input lens model. The input lens model values are given in brackets to the right of each parameter estimate. Columns three to five show  $\Delta\theta_{\text{Ein}}$ ,  $\Delta q$  and  $\Delta\alpha$  within  $3\sigma$  confidence and column six  $\Delta\alpha$  at  $1\sigma$ . Parameter estimates in bold text are inconsistent with the input lens model at their stated error estimates. The other parameters not shown ( $x$ ,  $y$  and  $\theta$ ) are all estimated accurately within  $3\sigma$ .

Image	Component	Parameters ( $3\sigma$ )			Parameters ( $1\sigma$ )		
<b>SreBulge</b> <sub>HS50NLSBulge</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0043^{+0.0125}_{-0.0136}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0094^{+0.0329}_{-0.0300}$ ( $q = 0.80$ )	$\Delta\alpha = 0.0227^{+0.0692}_{-0.0789}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0227^{+0.0290}_{-0.0248}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0227^{+0.0290}_{-0.0248}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0227^{+0.0290}_{-0.0248}$ ( $\alpha = 2.00$ )
<b>SreBulge</b> <sub>HS30NLSBulge</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0087^{+0.0248}_{-0.0223}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0194^{+0.0499}_{-0.0551}$ ( $q = 0.80$ )	$\Delta\alpha = 0.0456^{+0.0422}_{-0.0427}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0456^{+0.0422}_{-0.0427}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0456^{+0.0422}_{-0.0427}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0456^{+0.0422}_{-0.0427}$ ( $\alpha = 2.00$ )
<b>SreBulge</b> <sub>HS10NLSBulge</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0228^{+0.0337}_{-0.0320}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0495^{+0.0713}_{-0.0702}$ ( $q = 0.80$ )	$\Delta\alpha = 0.1092^{+0.1378}_{-0.1524}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.1092^{+0.1378}_{-0.1524}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.1092^{+0.1378}_{-0.1524}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.1092^{+0.1378}_{-0.1524}$ ( $\alpha = 2.00$ )
<b>SreBulge</b> <sub>ES50NLSBulge</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0096^{+0.0171}_{-0.0165}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0217^{+0.0427}_{-0.0423}$ ( $q = 0.80$ )	$\Delta\alpha = 0.0468^{+0.0852}_{-0.0891}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0468^{+0.0852}_{-0.0891}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0468^{+0.0852}_{-0.0891}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0468^{+0.0852}_{-0.0891}$ ( $\alpha = 2.00$ )
<b>SreBulge</b> <sub>ES30NLSBulge</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0028^{+0.0253}_{-0.0247}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0072^{+0.0655}_{-0.0663}$ ( $q = 0.80$ )	$\Delta\alpha = 0.0040^{+0.1425}_{-0.1559}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0040^{+0.0564}_{-0.0507}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0040^{+0.0564}_{-0.0507}$ ( $\alpha = 2.00$ )	$\Delta\alpha = 0.0040^{+0.0564}_{-0.0507}$ ( $\alpha = 2.00$ )
<b>SreBulge</b> <sub>ES10NLSBulge</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = -0.0020^{+0.0308}_{-0.0275}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = 0.0131^{+0.0705}_{-0.0720}$ ( $q = 0.80$ )	$\Delta\alpha = -0.0259^{+0.1810}_{-0.2001}$ ( $\alpha = 2.00$ )	$\Delta\alpha = -0.0259^{+0.0688}_{-0.0576}$ ( $\alpha = 2.00$ )	$\Delta\alpha = -0.0259^{+0.0688}_{-0.0576}$ ( $\alpha = 2.00$ )	$\Delta\alpha = -0.0259^{+0.0688}_{-0.0576}$ ( $\alpha = 2.00$ )
<b>SreBD</b> <sub>HS50NLBD</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0047^{+0.0196}_{-0.0198}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0038^{+0.0335}_{-0.0332}$ ( $q = 0.70$ )	$\Delta\alpha = 0.0082^{+0.0505}_{-0.0527}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0082^{+0.0183}_{-0.0178}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0082^{+0.0183}_{-0.0178}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0082^{+0.0183}_{-0.0178}$ ( $\alpha = 1.70$ )
<b>SreBD</b> <sub>HS30NLBD</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0015^{+0.0255}_{-0.0242}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = 0.0012^{+0.0425}_{-0.0440}$ ( $q = 0.70$ )	$\Delta\alpha = -0.0005^{+0.0675}_{-0.0673}$ ( $\alpha = 1.70$ )	$\Delta\alpha = -0.0005^{+0.0239}_{-0.0238}$ ( $\alpha = 1.70$ )	$\Delta\alpha = -0.0005^{+0.0239}_{-0.0238}$ ( $\alpha = 1.70$ )	$\Delta\alpha = -0.0005^{+0.0239}_{-0.0238}$ ( $\alpha = 1.70$ )
<b>SreBD</b> <sub>HS10NLBD</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0106^{+0.0586}_{-0.0568}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0101^{+0.0962}_{-0.0927}$ ( $q = 0.70$ )	$\Delta\alpha = 0.0203^{+0.1422}_{-0.1537}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0203^{+0.0561}_{-0.0514}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0203^{+0.0561}_{-0.0514}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0203^{+0.0561}_{-0.0514}$ ( $\alpha = 1.70$ )
<b>SreBD</b> <sub>ES50NLBD</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0038^{+0.0210}_{-0.0214}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0001^{+0.0407}_{-0.0393}$ ( $q = 0.70$ )	$\Delta\alpha = 0.0041^{+0.0586}_{-0.0615}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0041^{+0.0221}_{-0.0211}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0041^{+0.0221}_{-0.0211}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0041^{+0.0221}_{-0.0211}$ ( $\alpha = 1.70$ )
<b>SreBD</b> <sub>ES30NLBD</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0109^{+0.0336}_{-0.0295}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0147^{+0.0479}_{-0.0530}$ ( $q = 0.70$ )	$\Delta\alpha = 0.0234^{+0.0802}_{-0.0751}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0234^{+0.0269}_{-0.0288}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0234^{+0.0269}_{-0.0288}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0234^{+0.0269}_{-0.0288}$ ( $\alpha = 1.70$ )
<b>SreBD</b> <sub>ES10NLBD</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0144^{+0.0720}_{-0.0534}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0113^{+0.0920}_{-0.1058}$ ( $q = 0.70$ )	$\Delta\alpha = 0.0159^{+0.1504}_{-0.1397}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0159^{+0.0492}_{-0.0506}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0159^{+0.0492}_{-0.0506}$ ( $\alpha = 1.70$ )	$\Delta\alpha = 0.0159^{+0.0492}_{-0.0506}$ ( $\alpha = 1.70$ )
<b>SreDisk</b> <sub>HS50NLDisk</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0020^{+0.0204}_{-0.0191}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0071^{+0.0555}_{-0.0552}$ ( $q = 0.75$ )	$\Delta\alpha = 0.0062^{+0.0541}_{-0.0537}$ ( $\alpha = 2.30$ )	$\Delta\alpha = 0.0062^{+0.0187}_{-0.0196}$ ( $\alpha = 2.30$ )	$\Delta\alpha = 0.0062^{+0.0187}_{-0.0196}$ ( $\alpha = 2.30$ )	$\Delta\alpha = 0.0062^{+0.0187}_{-0.0196}$ ( $\alpha = 2.30$ )
<b>SreDisk</b> <sub>HS30NLDisk</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = -0.0045^{+0.0286}_{-0.0218}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = 0.0158^{+0.0674}_{-0.0809}$ ( $q = 0.75$ )	$\Delta\alpha = -0.0186^{+0.0785}_{-0.0762}$ ( $\alpha = 2.30$ )	$\Delta\alpha = -0.0186^{+0.0262}_{-0.0271}$ ( $\alpha = 2.30$ )	$\Delta\alpha = -0.0186^{+0.0262}_{-0.0271}$ ( $\alpha = 2.30$ )	$\Delta\alpha = -0.0186^{+0.0262}_{-0.0271}$ ( $\alpha = 2.30$ )
<b>SreDisk</b> <sub>HS10NLDisk</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = -0.0100^{+0.0546}_{-0.0369}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = 0.0205^{+0.1263}_{-0.1440}$ ( $q = 0.75$ )	$\Delta\alpha = -0.0249^{+0.1521}_{-0.1536}$ ( $\alpha = 2.30$ )	$\Delta\alpha = -0.0249^{+0.0561}_{-0.0547}$ ( $\alpha = 2.30$ )	$\Delta\alpha = -0.0249^{+0.0561}_{-0.0547}$ ( $\alpha = 2.30$ )	$\Delta\alpha = -0.0249^{+0.0561}_{-0.0547}$ ( $\alpha = 2.30$ )
<b>SreDisk</b> <sub>ES50NLDisk</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0029^{+0.0352}_{-0.0305}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0046^{+0.0926}_{-0.0951}$ ( $q = 0.75$ )	$\Delta\alpha = 0.0118^{+0.0732}_{-0.0819}$ ( $\alpha = 2.30$ )	$\Delta\alpha = 0.0118^{+0.0285}_{-0.0262}$ ( $\alpha = 2.30$ )	$\Delta\alpha = 0.0118^{+0.0285}_{-0.0262}$ ( $\alpha = 2.30$ )	$\Delta\alpha = 0.0118^{+0.0285}_{-0.0262}$ ( $\alpha = 2.30$ )
<b>SreDisk</b> <sub>ES30NLDisk</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = -0.0042^{+0.0335}_{-0.0291}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = 0.0228^{+0.0944}_{-0.1029}$ ( $q = 0.75$ )	$\Delta\alpha = -0.0144^{+0.1032}_{-0.1045}$ ( $\alpha = 2.30$ )	$\Delta\alpha = -0.0144^{+0.0375}_{-0.0367}$ ( $\alpha = 2.30$ )	$\Delta\alpha = -0.0144^{+0.0375}_{-0.0367}$ ( $\alpha = 2.30$ )	$\Delta\alpha = -0.0144^{+0.0375}_{-0.0367}$ ( $\alpha = 2.30$ )
<b>SreDisk</b> <sub>ES10NLDisk</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0128^{+0.1009}_{-0.0140}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0229^{+0.2111}_{-0.2175}$ ( $q = 0.75$ )	$\Delta\alpha = 0.0653^{+0.1885}_{-0.2548}$ ( $\alpha = 2.30$ )	$\Delta\alpha = 0.0653^{+0.0840}_{-0.0646}$ ( $\alpha = 2.30$ )	$\Delta\alpha = 0.0653^{+0.0840}_{-0.0646}$ ( $\alpha = 2.30$ )	$\Delta\alpha = 0.0653^{+0.0840}_{-0.0646}$ ( $\alpha = 2.30$ )
<b>SreMulti</b> <sub>HS50NLMulti</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0007^{+0.0140}_{-0.0137}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0010^{+0.0281}_{-0.0282}$ ( $q = 0.75$ )	$\Delta\alpha = 0.0006^{+0.0640}_{-0.0658}$ ( $\alpha = 2.10$ )	$\Delta\alpha = 0.0006^{+0.0236}_{-0.0230}$ ( $\alpha = 2.10$ )	$\Delta\alpha = 0.0006^{+0.0236}_{-0.0230}$ ( $\alpha = 2.10$ )	$\Delta\alpha = 0.0006^{+0.0236}_{-0.0230}$ ( $\alpha = 2.10$ )
<b>SreMulti</b> <sub>HS30NLMulti</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = -0.0044^{+0.0192}_{-0.0193}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = 0.0092^{+0.0391}_{-0.0397}$ ( $q = 0.75$ )	$\Delta\alpha = -0.0251^{+0.0892}_{-0.0949}$ ( $\alpha = 2.10$ )	$\Delta\alpha = -0.0251^{+0.0345}_{-0.0316}$ ( $\alpha = 2.10$ )	$\Delta\alpha = -0.0251^{+0.0345}_{-0.0316}$ ( $\alpha = 2.10$ )	$\Delta\alpha = -0.0251^{+0.0345}_{-0.0316}$ ( $\alpha = 2.10$ )
<b>SreMulti</b> <sub>HS10NLMulti</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = -0.0049^{+0.0397}_{-0.0337}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = 0.0065^{+0.0667}_{-0.0736}$ ( $q = 0.75$ )	$\Delta\alpha = -0.0284^{+0.1697}_{-0.1700}$ ( $\alpha = 2.10$ )	$\Delta\alpha = -0.0284^{+0.0612}_{-0.0611}$ ( $\alpha = 2.10$ )	$\Delta\alpha = -0.0284^{+0.0612}_{-0.0611}$ ( $\alpha = 2.10$ )	$\Delta\alpha = -0.0284^{+0.0612}_{-0.0611}$ ( $\alpha = 2.10$ )
<b>SreMulti</b> <sub>ES50NLMulti</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0028^{+0.0158}_{-0.0149}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0073^{+0.0311}_{-0.0309}$ ( $q = 0.75$ )	$\Delta\alpha = 0.0143^{+0.0679}_{-0.0715}$ ( $\alpha = 2.10$ )	$\Delta\alpha = 0.0143^{+0.0230}_{-0.0219}$ ( $\alpha = 2.10$ )	$\Delta\alpha = 0.0143^{+0.0230}_{-0.0219}$ ( $\alpha = 2.10$ )	$\Delta\alpha = 0.0143^{+0.0230}_{-0.0219}$ ( $\alpha = 2.10$ )
<b>SreMulti</b> <sub>ES30NLMulti</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0162^{+0.0358}_{-0.0346}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0315^{+0.0667}_{-0.0671}$ ( $q = 0.75$ )	$\Delta\alpha = 0.0624^{+0.1246}_{-0.1330}$ ( $\alpha = 2.10$ )	$\Delta\alpha = 0.0624^{+0.0453}_{-0.0453}$ ( $\alpha = 2.10$ )	$\Delta\alpha = 0.0624^{+0.0453}_{-0.0453}$ ( $\alpha = 2.10$ )	$\Delta\alpha = 0.0624^{+0.0453}_{-0.0453}$ ( $\alpha = 2.10$ )
<b>SreMulti</b> <sub>ES10NLMulti</sub>	Mass (SPL)	$\Delta\theta_{\text{Ein}} = 0.0008^{+0.1147}_{-0.0831}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0047^{+0.1650}_{-0.1834}$ ( $q = 0.75$ )	$\Delta\alpha = -0.0191^{+0.3940}_{-0.4142}$ ( $\alpha = 2.10$ )	$\Delta\alpha = -0.0191^{+0.1479}_{-0.1413}$ ( $\alpha = 2.10$ )	$\Delta\alpha = -0.0191^{+0.1479}_{-0.1413}$ ( $\alpha = 2.10$ )	$\Delta\alpha = -0.0191^{+0.1479}_{-0.1413}$ ( $\alpha = 2.10$ )



**Figure 11.** The simulated model images, residuals,  $\chi^2$  images and source reconstructions for the no lens light analysis of **SrcBulge**<sub>ES10NLBulge</sub> (top row) and **SrcBD**<sub>HS50NLBD</sub> (bottom row). Images correspond to the most likely model at the end of the main pipeline, corresponding to the models given by rows six, seven, thirteen and twenty-one of Table 5 respectively.

models. The same degeneracy discussed in N15 is seen between the parameters governing the lens's mass distribution, where the degenerate models shown by the contours each integrate to give approximately the same  $M_{\text{Ein}}$ . Accompanying this (but not shown) is the source-plane scaling effect, demonstrated in fig. 4 of N15, whereby steeper mass profiles lead to a more expanded source reconstruction. As expected, the posterior probability distribution function broadens for lower resolution imaging.

Fig. 13 shows  $\Delta\alpha$ 's 1-D PDF for each image of every lens model. There is no systematic deviation of  $\Delta\alpha$  with varying image resolution, S/N ratio, mass model or source morphology, confirming AutoLens's source-only analysis is free of systematic bias. As expected, images at higher spatial resolution or S/N give tighter lens model constraints. This figure also gives a sense of how precisely images of different image resolution or S/N constrain  $\alpha$ , suggesting that Euclid wide-field imaging will be able to estimate  $\alpha$  to a precision  $\pm 0.1$ , or better, at  $3\sigma$  confidence.

By comparing each panel, one can also see how the lens model's precision depends on the source morphology. The bulge-disk morphology is marginally the most-well constrained, benefiting from how its source has both a smooth extended disk component and cuspy central light profile. In contrast, the bulge-only morphology offers the loosest constraints, suggesting that an extended envelope of source light is important to reducing errors. However, for images with the same resolution and S/N, the differences in error estimates are marginal, thus for smoothly parametrized source morphologies the profile shape and number of components appears to play no major role in determining how precisely the mass model is constrained. This contradicts discussions by the authors (Vegetti et al. 2012; Lagattuta et al. 2012), who argue that multiple sources with non-symmetric morphologies offer much tighter constraints, as they produce a less degenerate set of possible image reconstructions from which there is a smaller sub-set of mass models that are able to reconstruct them accurately. Such a trend is not seen for the analysis of the **SrcMulti** images. However, this is most likely a reflection of the fact that these simulated images are modeled with their input mass profile and the tighter constraints offered by complex sources

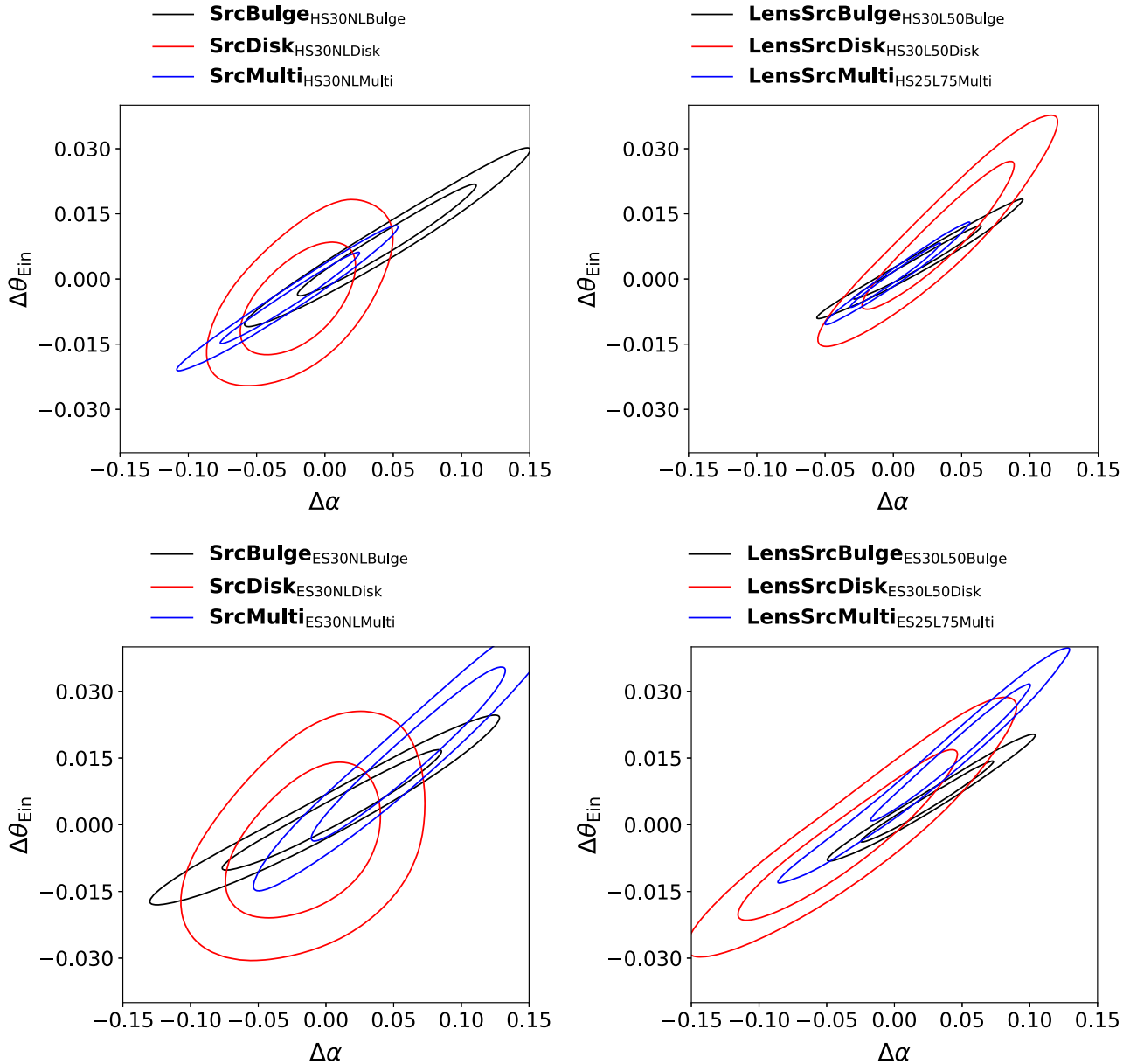
are more readily observed on real strong lens imaging, where the mass model offers only an approximate fit.

## 7.2 Non-cored Lens and Source

The non-cored lens simulation suite consists of seven unique lens and source models, which are each used to generate two images at Hubble and Euclid resolution with a range of source and lens S/N ratios, giving a total of fourteen images. The analysis of each image uses the singular total-mass pipeline, with parameter estimates and results corresponding to the end of phase two. The results of using Bayesian model comparison to choose the light and mass profiles are shown in Table 6, with the input model correctly chosen for all images. The *Sersic* light model is summarized using the mismatch parameters  $\Delta R_1$ ,  $\Delta n_1$  and  $\Delta q_1$ , whereas the *Exp* model use  $\Delta R_{11}$ ,  $\Delta q_{11}$ ,  $\Delta R_{12}$  and  $\Delta q_{12}$ . These are listed, alongside  $\Delta\theta_{\text{Ein}}$ ,  $\Delta q$  and  $\Delta\alpha$ , in Table 7, showing the majority of parameter estimates are accurate within  $3\sigma$  confidence. Fig. 14 shows the observed images, model images, model sources, residuals,  $\chi^2$  images and source reconstructions for three images. The image residuals can be seen to realize each image's noise whereas the  $\chi^2$  images are Gaussian, as desired. These images are indicative of the analysis of all non-cored images.

### 7.2.1 Model comparison

The results of the model comparison phases  $\mathbf{P}_{\text{LightMC}}$  and  $\mathbf{P}_{\text{MassMC}}$  for all images are given in Table 6. For all images, model comparison correctly chooses the input light and mass models. Therefore, at Euclid resolution or higher, multi-component light profiles and detection of an external shear are possible. Model comparison also never wrongly favours a more complex model, reaffirming that model comparison functions exactly as expected. For many of these comparisons the more complex model is an extension of the true model (e.g. all *SPLE* + *Shear* models can reproduce the input *SPLE* model if  $\gamma_{\text{sh}} = 0$ ), the scenario which acts as the most stringent test of model comparison. For many model comparisons, the high-



**Figure 12.** Marginalized 2-D PDFs of the mismatch parameters  $\Delta P = P_{\text{True}} - P_{\text{Model}}$  for the mass model Einstein radius  $\Delta\theta_{\text{Ein}}$  and power-law density slope  $\Delta\alpha$  for images of the **SrcBulge**, **SrcDisk** and **SrcMulti** models (left panels) and **LensSrcBulge**, **LensSrcDisk** and **LensSrcMulti** models (right panels). The legend at the top of each panel indicates the image each contour corresponds to, where the **Src** images have been chosen to closely match the S/N of the **Lens** images, such that the left two panels compare the PDFs of similar lensed sources at Hubble resolution and right panels at Euclid resolution. Contours give the  $1\sigma$  (interior) and  $3\sigma$  (exterior) confidence regions. For **SrcBulge** and **LensSrcBulge** the input values of each parameter are  $\theta_{\text{Ein}} = 1.2''$ ,  $q = 0.8$  and  $\alpha = 2.0$ , for **SrcDisk** and **LensSrcDisk**  $\theta_{\text{Ein}} = 1.2$ ,  $q = 0.75$  and  $\alpha = 2.1$ . Panels one and two, or three and four, therefore compare the analysis of nearly identical lensed sources but for a model which either does or does not include lens light modeling. With lens light modeling included, the PDFs do not appear broader.

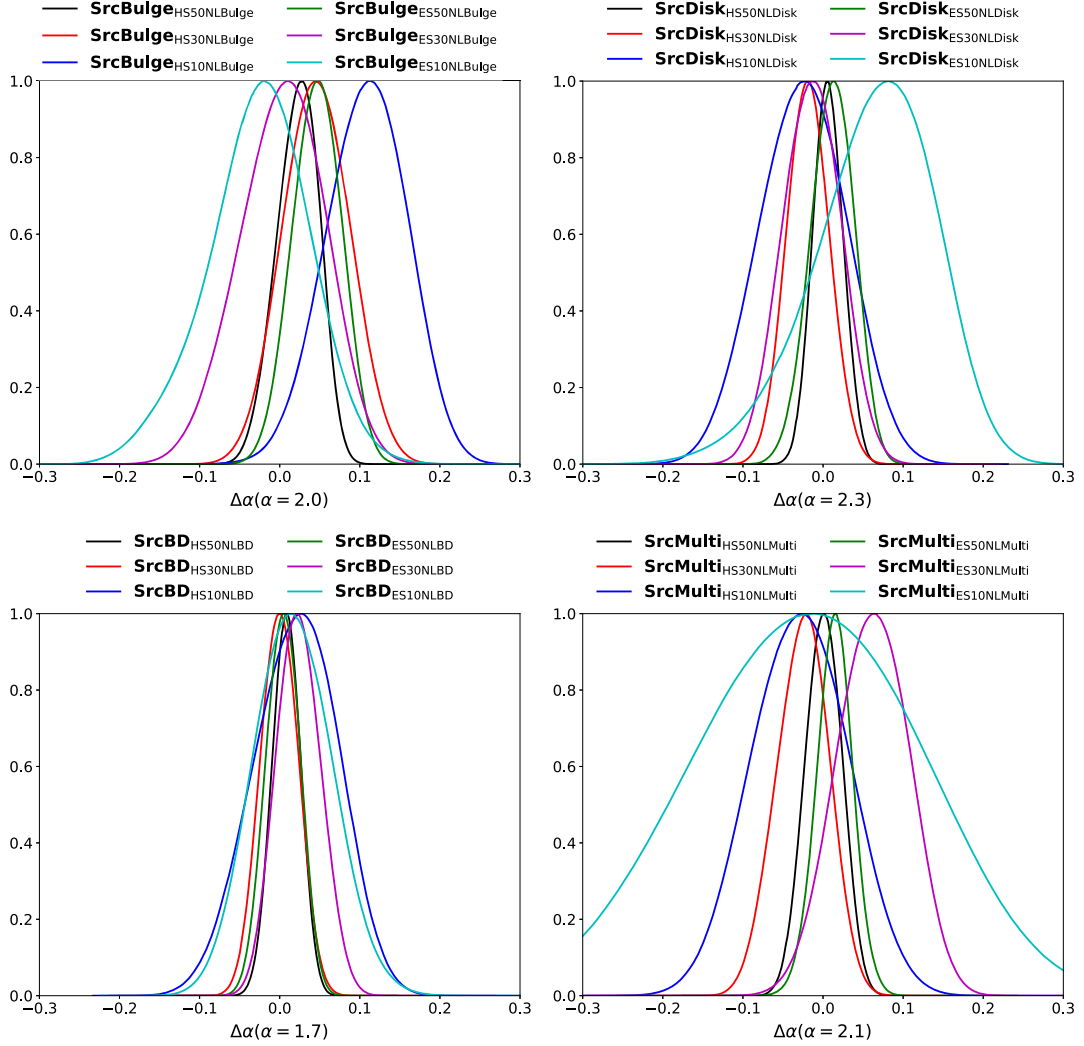
est likelihood values found in the (rejected) more complex models were higher than the highest likelihoods of the simpler model, by values of approximately  $\ln\epsilon = 5 - 15$ . This shows a maximum likelihood based approach is not well suited to determining the lens model complexity and demonstrates **AutoLens**'s use of Occam's Razor.

### 7.2.2 Modeling results

The modeling results for the non-cored simulation suite, given in Table 7, are positive, with all but one mass model parameter esti-

mated incorrectly within  $3\sigma$  confidence and their estimates at  $1\sigma$  consistent with expectations. The incorrect mass model parameter is the density-slope  $\alpha$  for the image **LensMassShear**<sub>ES40L80Disk</sub>. Given the accurate parameter estimates for the equivalent high-resolution image, it appears that this image is simply too low resolution to accurately constrain the density-slope simultaneously with an external shear.

The majority of light model parameters are also estimated correctly within  $3\sigma$  confidence and estimates at  $1\sigma$  are again consistent with Gaussian expectations. However, there is a trend throughout the parameter estimates whereby the light model's  $R_1$  and  $n_1$  val-



**Figure 13.** Marginalized 1-D PDF's of the mass-profile density mismatch  $\Delta\alpha = \alpha_{\text{True}} - \alpha_{\text{Model}}$  for the twenty-four mass models given in Table 5. The top-left panel corresponds to the six images of the **SrcBulge** model, the top-right **SrcDisk**, bottom-left **SrcBD** and bottom-right **SrcMulti**. Each graph's legend indicates the image that each line corresponds to, where black / red / blue lines give Hubble resolution images at  $S/N = 50 / 30 / 10$  and green / purple / cyan give Euclid resolution images at  $S/N = 50 / 30 / 10$ , respectively. The input value of  $\alpha$  for each model is given in brackets by the x-axis label. All PDF's are consistent with the input lens model ( $\Delta\alpha = 0.0$ ).

ues are over-estimated compared to their true input value. In most cases, this is consistent within their errors (at  $1 - 2\sigma$  confidence), however there are five cases which are incorrect within  $3\sigma$ . These offsets are most prevalent for simulated images which are: (i) generated with a low  $R_1$ ; (ii) generated with a high  $n_1$  and; (iii) generated at Euclid resolution. Thus, these offsets are most significant in cases where there is more blending between the source and lens light, especially when this occurs towards its central regions where the strongest constraints on  $R_1$  and  $n_1$  are placed. This is a limitation of any analysis which tries to deblend the lens from the source using single-waveband imaging without a sufficiently high resolution.

Fig. 15 shows the 1-D PDF of  $\Delta n_1$  for the Hubble resolution (left-hand panel) and Euclid resolution (right-hand panel) images of each lens model, corresponding to the models given by Table 7. The majority of  $n_1$  values are estimated correctly, with those that are not discussed above. The tendency to be shifted above the lens's input value of  $n_1$  is also visible in this figure. The size of each PDF and therefore the precision inferred on the value of  $n_1$  can be seen to vary greatly, but without an obvious dependence on the

image resolution or  $S/N$ . Thus, these factors do not appear to be the most important for constraining the lens's light profile. Instead, the model's precision is most heavily dependent on the lens's input values of  $R_1$  and  $n_1$ , where larger effective radii and lower Sersic indexes provide a more precise light model. In these cases, a greater amount of lens light (that is not obscured by the source's light) is visible and available to constrain the lens's light profile. Therefore, it is the degree of lens and source light blending that drives how well the light profile can be measured.

Now the lens's light and mass are modeled simultaneously, it is interesting to investigate what interplay, if any, there is between the two components. Fig. 16 shows the marginalized 2-D PDF's between  $\Delta\alpha$  and  $\Delta n_1$  for the Hubble resolution (left panel) and Euclid resolution (right panel) images. Contours are orthogonal, demonstrating there is no noticeable degeneracy between the mass and light models, where visual inspection of other parameter pairs (e.g.  $I_1$ ,  $R_L$ ,  $\theta_{\text{Ein}}$ , etc.) reveals this holds in general.

The precision of the mass model also appears to show no dependence on whether the lens's light is included, as can be inferred

**Table 6.** The results of Bayesian model comparison in the phases  $\mathbf{P}_{\text{LightMC}}$  and  $\mathbf{P}_{\text{MassMC}}$  for the fourteen **Lens** images, where the image's listed in the first column are generated using a variety of source morphologies, lens profile and mass models. The third and fourth columns show the Bayesian evidence values (equation 23) computed for the *Sersic* and *Sersic + Exp* light models and the fifth and sixth columns the values for the *SPLE* and *SPLE + Shear* mass models. Values in bold correspond to those chosen by the pipeline, noting that a threshold of twenty must be exceeded to favour a more complex model.

Image	Sersic ( $\mathbf{P}_{\text{MCLight}}$ )	Sersic + Exp ( $\mathbf{P}_{\text{MCLight}}$ )	SPLE ( $\mathbf{P}_{\text{MCMass}}$ )	SPLE + Shear ( $\mathbf{P}_{\text{MCMass}}$ )
<b>LensSrcBulge</b> <sub>HS30L50Bulge</sub>	<b>55549.5448</b>	55552.1930	<b>55602.9633</b>	55604.7013
<b>LensSrcBulge</b> <sub>ES30L50Bulge</sub>	<b>15895.6885</b>	15894.4006	<b>15915.2358</b>	15917.5771
<b>LensSrcDisk</b> <sub>HS30L50Disk</sub>	<b>55113.8123</b>	55122.6322	<b>55154.4289</b>	55153.1476
<b>LensSrcDisk</b> <sub>ES30L50Disk</sub>	<b>15567.8761</b>	15571.8895	<b>15580.8484</b>	15573.1699
<b>LensSrcCusp</b> <sub>HS20L60Cusp</sub>	<b>54989.4958</b>	54981.9373	<b>55025.2533</b>	55025.8970
<b>LensSrcCusp</b> <sub>ES20L60Cusp</sub>	<b>15111.1705</b>	15101.0015	<b>15137.4186</b>	15137.7187
<b>LensSrcDouble</b> <sub>HS25L60BD</sub>	<b>55450.3420</b>	55426.1254	<b>55478.9227</b>	55480.2748
<b>LensSrcDouble</b> <sub>ES25L60BD</sub>	<b>15588.8998</b>	15600.0188	<b>15623.4948</b>	15621.5500
<b>LensSrcMulti</b> <sub>HS25L75Multi</sub>	<b>55487.6488</b>	55491.7482	<b>55522.6763</b>	55522.9063
<b>LensSrcMulti</b> <sub>ES25L75Multi</sub>	<b>15609.1851</b>	15576.4738	<b>15629.4569</b>	15633.4754
<b>LensMassShear</b> <sub>HS40L80Disk</sub>	<b>54300.2109</b>	54314.5680	54528.1859	<b>54722.6101</b>
<b>LensMassShear</b> <sub>ES40L80Disk</sub>	<b>14787.7869</b>	14797.0800	14841.6874	<b>14951.8195</b>
<b>LensLightBD</b> <sub>HS25L50BD</sub>	55177.7344	<b>55327.3654</b>	<b>55402.3462</b>	55403.6771
<b>LensLightBD</b> <sub>ES25L50BD</sub>	15319.9874	<b>15407.3859</b>	<b>15487.9352</b>	15488.5821

by contrasting the errors on  $\theta_{\text{Ein}}$ ,  $q$ , and  $\alpha$  for the  $S/N = 30$  images of the **LensSrcBulge/SrcBulge**, **LensSrcDisk/SrcDisk** and **LensSrcMulti/SrcMulti** models. Each pair of models shares the same spatial resolution,  $S/N$  ratio, mass, and source profiles, with the only difference between them being the inclusion of the lens's light. The magnitude of errors for the source-only case and lens and source case shows no systematic increase across all images, demonstrating that the light subtraction (provided it is accurate) does not impact the mass model precision. Fig. 12 reinforces this further, showing that the mass-profile degeneracies between  $\theta_{\text{Ein}}$  and  $\alpha$  are similar for these images, regardless of whether the lens is included or not.

Whilst the lens's light has no impact on the mass model precision, the reverse is not true. That is, the presence of the source's light has a huge impact on the precision of the inferred light profile. This was confirmed by comparing the errors on the light profile parameters shown in Fig. 7 to the errors computed by fitting lens-only variants of each image. For example, for the image **LensSrcBulge**<sub>HS30L50Bulge</sub>, errors on  $n_1$  were approximately  $\pm 0.05$  when no source was present and  $\pm 0.7$  when it was (see Table 7). This occurs because of the smooth and symmetric nature of the lens's light profile, which has a sizable fraction of its light obscured by the ring-like source, such that much looser constraints are possible. Furthermore, for the simulated lenses used in this work, the source obscures the lens in and around its half-light radius, where the largest impact on the estimation of  $R_1$  (and  $n_1$  due to their degeneracy) can be expected. In contrast, the source morphology is asymmetric and irregular and therefore looks similar, regardless of the smooth lens profile used for the light subtraction, such that similar constraints are offered across a range of lens light models.

### 7.3 Cored lens and source

The cored lens simulation suite consists of three unique lens and source models, each of which is again used to generate an image at Hubble resolution and Euclid resolution both with a source  $S/N = 35-50$  and lens  $S/N$  spanning  $50-70$ , giving a total of six images. Each image is analyzed using AutoLens's singular total-mass profile and cored total-mass profile pipelines. The results of model comparison are shown in Table 8, showing that the light profiles and cored models are correctly chosen for all images.

Parameter estimates for each image are summarized in Table 9, using  $\Delta S$  and the same mismatch parameters as before, where many parameters are estimated correctly within  $3\sigma$  confidence but one image, discussed next, has clear problems. Fig. 17 shows the observed images, model images, model sources, residuals  $\chi^2$  images, and source reconstruction for the high-resolution images, showing that the image and source reconstructions successfully reproduce the features of a cored mass profile, like radial arcs or a central image, and again gives residuals and  $\chi^2$  images consistent with the noise.

#### 7.3.1 Model comparison

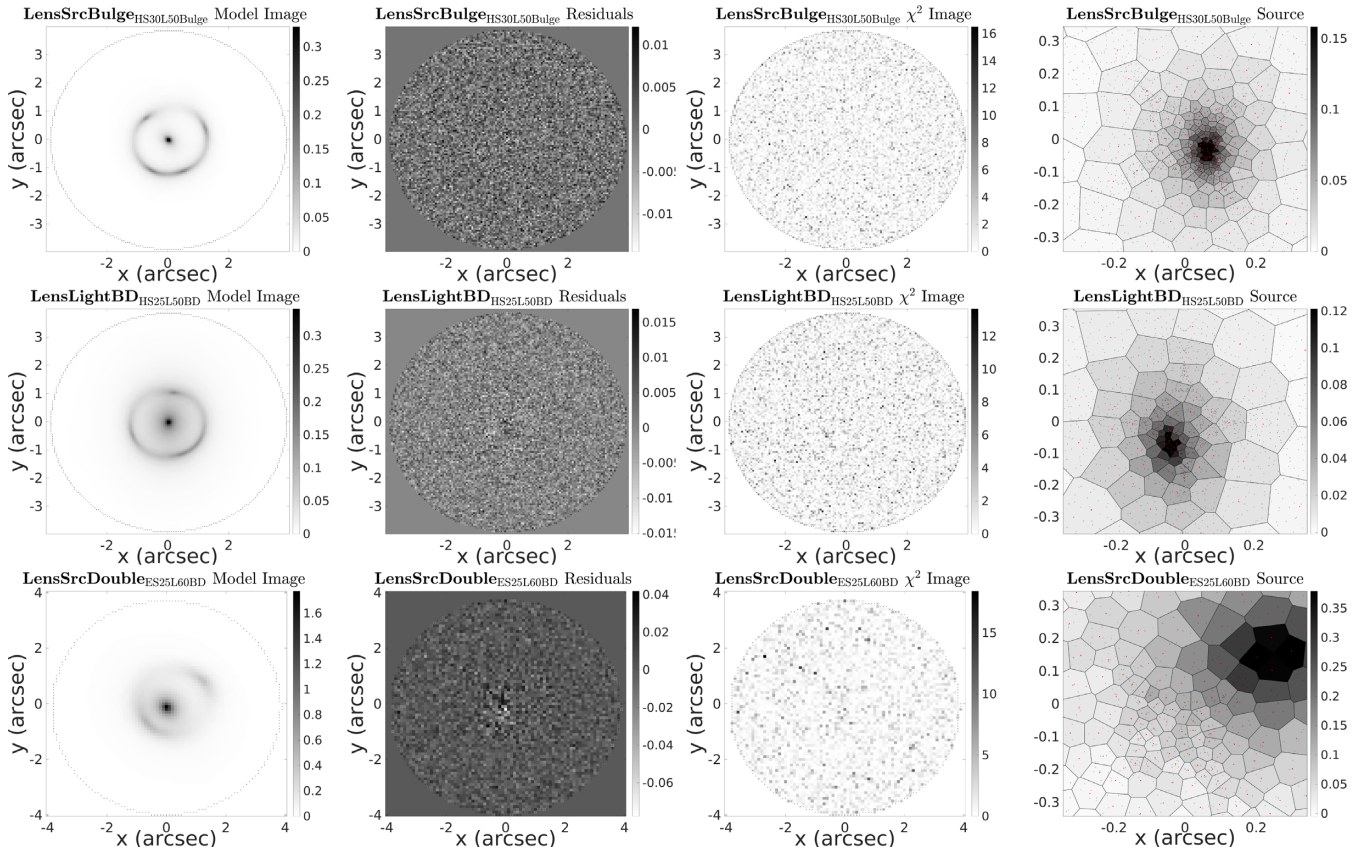
The results of the model comparison between the *SPLE* and *PL*<sub>Core</sub> models, using a full AutoLens analysis for each, are given for all six images in Table 8. For all images model comparison correctly chooses the cored model, demonstrating the method's success at modeling a cored profile. The correct lens light profile is also chosen for all images. There is one case, given by the bottom row of Table 8, which is incorrect, where the image **CoreSrcDouble**<sub>ES25L50BD</sub> includes a *Shear* component. There is no obvious explanation for this occurrence and the shear magnitude  $\gamma_{\text{sh}}$  reverts to approximately 0 in the main pipeline. Thus, its inclusion is disregarded as non-consequential.

#### 7.3.2 Modeling results

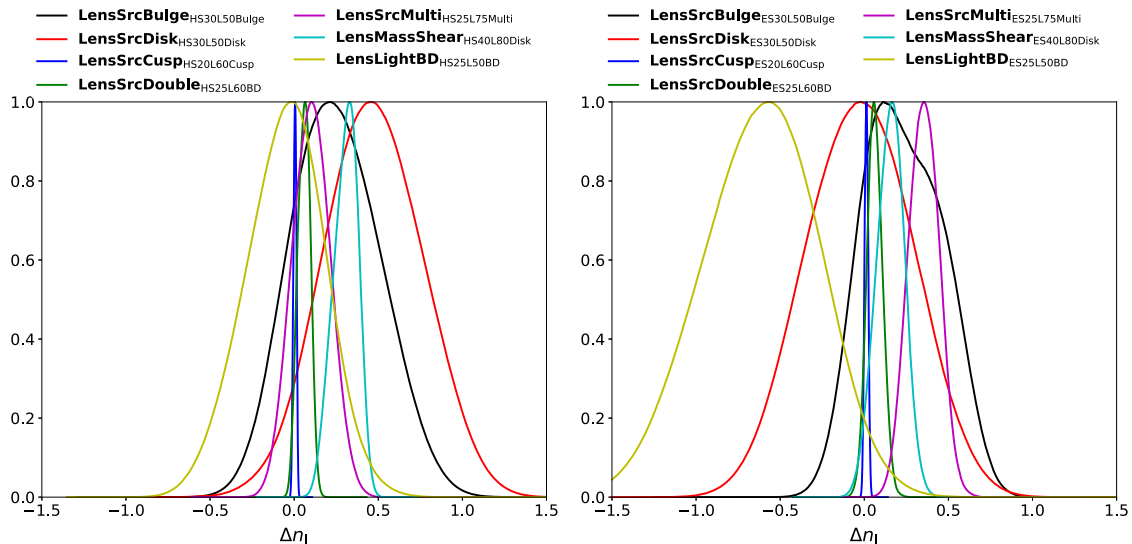
The parameter estimates for the cored images are given in Table 9, where parameter estimates for four out of six images are all accurate within  $3\sigma$  confidence. However, for images of the **CoreSrcQuad** model, the light profile parameter estimates are poor, with the values of  $n_1$  significantly offset from the input value. The middle row of Fig. 17 shows the image **CoreSrcQuad**<sub>HS40L60BD</sub>, where it can be seen that this is an example of an image where the source's central image perfectly overlaps the centre of the lens's light profile. This is the case of maximum blending and it is no surprise that the light model is inaccurate, as the lens's central light profile is completely obscured. This means that if lenses of this configuration are found in nature, care must be taken in ensuring their lens subtraction is accurate and their inferred light profiles should be viewed with

**Table 7.** Results of fitting the fourteen **Lens** model images using the automated analysis pipeline, corresponding to results generated at end of phase two of the main pipeline. Each image's name is given in the first column and the light or mass model component in the second column. The third to sixth columns show parameter estimates, where each parameter is offset by  $\Delta P = P_{\text{true}} - P_{\text{model}}$ , such that zero corresponds to the input lens model. The input lens model values are given in brackets to the right of each parameter estimate. Parameters estimates are shown using  $\Delta R_1$ ,  $\Delta q$  and  $\Delta \gamma_{\text{sh}}$  for *Sersic* light models,  $\Delta R_1$ ,  $\Delta q_1$ ,  $\Delta \theta_{\text{Ein}}$ ,  $\Delta R_2$  and  $\Delta q_2$  for *Sersic + Exp* light models,  $\Delta \theta_{\text{Ein}}$ ,  $\Delta q$  and  $\Delta \alpha$  for *SPL* mass models and  $\Delta \theta_{\text{sh}}$  and  $\Delta \gamma_{\text{sh}}$  for a *Shear* component. Columns three to five show parameter estimates within  $3\sigma$  confidence and column six at  $1\sigma$ . Parameter estimates in bold text are inconsistent with the input lens model at their stated error estimates. The other parameters not shown (e.g.  $x_1$ ,  $\theta_1$ ,  $\theta$ ) are all estimated accurately within  $3\sigma$  or above.

Image	Component	Parameters ( $3\sigma$ )			Parameters ( $1\sigma$ )		
<b>LensSrcBulge</b> <sub>HS30L50Bulge</sub>	Light (Sersic)	$\Delta R_1 = 0.1133^{+0.2837}_{-0.2232}$ ( $R_1 = 0.60$ )	$\Delta q_1 = -0.0039^{+0.0465}_{-0.0442}$ ( $q_1 = 0.72$ )	$\Delta \theta_{\text{Ein}} = 0.0030^{+0.0020}_{-0.0151}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta R_1 = 0.2463^{+0.7624}_{-0.6917}$ ( $R_1 = 4.00$ )	$\Delta \alpha = 0.0193^{+0.2534}_{-0.2852}$ ( $\alpha = 4.00$ )	$\Delta \gamma_{\text{sh}} = 0.2463^{+0.2534}_{-0.2852}$ ( $\gamma_{\text{sh}} = 4.00$ )
<b>LensSrcBulge</b> <sub>HS30L50Bulge</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0039^{+0.0020}_{-0.0151}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0098^{+0.0353}_{-0.0433}$ ( $q = 0.80$ )		$\Delta \alpha = 0.0193^{+0.1005}_{-0.0874}$ ( $\alpha = 2.00$ )	$\Delta \alpha = 0.0193^{+0.0300}_{-0.0312}$ ( $\alpha = 2.00$ )	
<b>LensSrcBulge</b> <sub>ES30L50Bulge</sub>	Light (Sersic)	$\Delta R_1 = 0.0786^{+0.1810}_{-0.1523}$ ( $R_1 = 0.60$ )	$\Delta q_1 = 0.0060^{+0.0364}_{-0.0387}$ ( $q_1 = 0.72$ )		$\Delta R_1 = 0.2296^{+0.6116}_{-0.5765}$ ( $R_1 = 4.00$ )	$\Delta \alpha = 0.2296^{+0.2992}_{-0.2637}$ ( $\alpha = 4.00$ )	
<b>LensSrcBulge</b> <sub>ES30L50Bulge</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0052^{+0.0175}_{-0.0151}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0078^{+0.0361}_{-0.0417}$ ( $q = 0.80$ )		$\Delta \alpha = 0.0248^{+0.0917}_{-0.0863}$ ( $\alpha = 2.00$ )	$\Delta \alpha = 0.0248^{+0.0302}_{-0.0341}$ ( $\alpha = 2.00$ )	
<b>LensSrcDisk</b> <sub>HS30L50Disk</sub>	Light (Sersic)	$\Delta R_1 = 0.1939^{+0.3672}_{-0.3250}$ ( $R_1 = 0.60$ )	$\Delta q_1 = -0.0114^{+0.0563}_{-0.0572}$ ( $q_1 = 0.72$ )		$\Delta R_1 = 0.4707^{+0.8360}_{-0.8643}$ ( $R_1 = 4.00$ )	$\Delta \alpha = 0.4707^{+0.2931}_{-0.3025}$ ( $\alpha = 4.00$ )	
<b>LensSrcDisk</b> <sub>HS30L50Disk</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0097^{+0.0323}_{-0.0280}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0189^{+0.0645}_{-0.0685}$ ( $q = 0.75$ )		$\Delta \alpha = 0.0347^{+0.0988}_{-0.1009}$ ( $\alpha = 2.30$ )	$\Delta \alpha = 0.0347^{+0.0368}_{-0.0356}$ ( $\alpha = 2.30$ )	
<b>LensSrcDisk</b> <sub>ES30L50Disk</sub>	Light (Sersic)	$\Delta R_1 = -0.0029^{+0.2580}_{-0.2306}$ ( $R_1 = 0.60$ )	$\Delta q_1 = -0.0064^{+0.0519}_{-0.0541}$ ( $q_1 = 0.72$ )		$\Delta R_1 = -0.0257^{+0.8801}_{-0.8826}$ ( $R_1 = 4.00$ )	$\Delta \alpha = -0.0257^{+0.3230}_{-0.3230}$ ( $\alpha = 4.00$ )	
<b>LensSrcDisk</b> <sub>ES30L50Disk</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = -0.0030^{+0.0355}_{-0.0284}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = 0.0108^{+0.0745}_{-0.0827}$ ( $q = 0.75$ )		$\Delta \alpha = -0.0331^{+0.1406}_{-0.1345}$ ( $\alpha = 2.30$ )	$\Delta \alpha = -0.0331^{+0.0507}_{-0.0559}$ ( $\alpha = 2.30$ )	
<b>LensSrcCusp</b> <sub>HS20L60Cusp</sub>	Light (Sersic)	$\Delta R_1 = 0.0151^{+0.0246}_{-0.0244}$ ( $R_1 = 1.20$ )	$\Delta q_1 = 0.0009^{+0.0061}_{-0.0061}$ ( $q_1 = 0.60$ )		$\Delta R_1 = 0.0043^{+0.0233}_{-0.0253}$ ( $R_1 = 1.25$ )	$\Delta \alpha = 0.0043^{+0.0096}_{-0.0086}$ ( $\alpha = 1.25$ )	
<b>LensSrcCusp</b> <sub>HS20L60Cusp</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0279^{+0.0477}_{-0.0403}$ ( $\theta_{\text{Ein}} = 1.40$ )	$\Delta q = -0.0409^{+0.0599}_{-0.0664}$ ( $q = 0.70$ )		$\Delta \alpha = 0.0694^{+0.0883}_{-0.0884}$ ( $\alpha = 2.35$ )	$\Delta \alpha = 0.0694^{+0.0324}_{-0.0302}$ ( $\alpha = 2.35$ )	
<b>LensSrcCusp</b> <sub>ES20L60Cusp</sub>	Light (Sersic)	$\Delta R_1 = 0.0240^{+0.0281}_{-0.0312}$ ( $R_1 = 1.20$ )	$\Delta q_1 = 0.0113^{+0.0084}_{-0.0145}$ ( $q_1 = 0.60$ )		$\Delta R_1 = 0.0125^{+0.0297}_{-0.0300}$ ( $R_1 = 1.25$ )	$\Delta \alpha = 0.0125^{+0.0110}_{-0.0108}$ ( $\alpha = 1.25$ )	
<b>LensSrcCusp</b> <sub>ES20L60Cusp</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = -0.0137^{+0.0617}_{-0.0422}$ ( $\theta_{\text{Ein}} = 1.40$ )	$\Delta q = 0.0298^{+0.0771}_{-0.0956}$ ( $q = 0.70$ )		$\Delta \alpha = -0.0217^{+0.1690}_{-0.1687}$ ( $\alpha = 2.35$ )	$\Delta \alpha = -0.0217^{+0.0555}_{-0.0590}$ ( $\alpha = 2.35$ )	
<b>LensSrcDouble</b> <sub>HS25L60BD</sub>	Light (Sersic)	$\Delta R_1 = 0.0811^{+0.0776}_{-0.0741}$ ( $R_1 = 0.90$ )	$\Delta q_1 = 0.0081^{+0.0225}_{-0.0237}$ ( $q_1 = 0.80$ )		$\Delta R_1 = 0.0589^{+0.0997}_{-0.1002}$ ( $R_1 = 2.00$ )	$\Delta \alpha = 0.0589^{+0.0389}_{-0.0364}$ ( $\alpha = 2.00$ )	
<b>LensSrcDouble</b> <sub>HS25L60BD</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0119^{+0.0220}_{-0.0226}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0228^{+0.0540}_{-0.0498}$ ( $q = 0.75$ )		$\Delta \alpha = 0.0576^{+0.0937}_{-0.1001}$ ( $\alpha = 2.10$ )	$\Delta \alpha = 0.0576^{+0.0349}_{-0.0328}$ ( $\alpha = 2.10$ )	
<b>LensSrcDouble</b> <sub>ES25L60BD</sub>	Light (Sersic)	$\Delta R_1 = 0.0828^{+0.1065}_{-0.0947}$ ( $R_1 = 0.90$ )	$\Delta q_1 = 0.0122^{+0.0261}_{-0.0250}$ ( $q_1 = 0.80$ )		$\Delta R_1 = 0.0624^{+0.1483}_{-0.1398}$ ( $R_1 = 2.00$ )	$\Delta \alpha = 0.0624^{+0.0444}_{-0.0481}$ ( $\alpha = 2.00$ )	
<b>LensSrcDouble</b> <sub>ES25L60BD</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0160^{+0.0319}_{-0.0253}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0303^{+0.0629}_{-0.0668}$ ( $q = 0.75$ )		$\Delta \alpha = 0.0676^{+0.1244}_{-0.1332}$ ( $\alpha = 2.10$ )	$\Delta \alpha = 0.0676^{+0.0378}_{-0.0421}$ ( $\alpha = 2.10$ )	
<b>LensSrcMulti</b> <sub>HS25L75Multi</sub>	Light (Sersic)	$\Delta R_1 = 0.0633^{+0.1790}_{-0.1578}$ ( $R_1 = 0.80$ )	$\Delta q_1 = 0.0086^{+0.0220}_{-0.0227}$ ( $q_1 = 0.70$ )		$\Delta R_1 = 0.1005^{+0.3237}_{-0.3123}$ ( $R_1 = 3.00$ )	$\Delta \alpha = 0.1005^{+0.1137}_{-0.1136}$ ( $\alpha = 3.00$ )	
<b>LensSrcMulti</b> <sub>HS25L75Multi</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0009^{+0.0124}_{-0.0128}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = 0.0011^{+0.0262}_{-0.0278}$ ( $q = 0.75$ )		$\Delta \alpha = 0.0022^{+0.0626}_{-0.0595}$ ( $\alpha = 2.10$ )	$\Delta \alpha = 0.0022^{+0.0210}_{-0.0218}$ ( $\alpha = 2.10$ )	
<b>LensSrcMulti</b> <sub>ES25L75Multi</sub>	Light (Sersic)	$\Delta R_1 = 0.1998^{+0.1452}_{-0.1379}$ ( $R_1 = 0.80$ )	$\Delta q_1 = -0.0135^{+0.0183}_{-0.0191}$ ( $q_1 = 0.70$ )		$\Delta R_1 = 0.3531^{+0.2701}_{-0.2601}$ ( $R_1 = 3.00$ )	$\Delta \alpha = 0.3531^{+0.0937}_{-0.0915}$ ( $\alpha = 3.00$ )	
<b>LensSrcMulti</b> <sub>ES25L75Multi</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0153^{+0.0330}_{-0.0375}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0278^{+0.0744}_{-0.0588}$ ( $q = 0.75$ )		$\Delta \alpha = 0.0389^{+0.1184}_{-0.1595}$ ( $\alpha = 2.10$ )	$\Delta \alpha = 0.0389^{+0.0476}_{-0.0310}$ ( $\alpha = 2.10$ )	
<b>LensMassShear</b> <sub>HS40L80Disk</sub>	Light (Sersic)	$\Delta R_1 = 0.0253^{+0.0433}_{-0.0674}$ ( $R_1 = 1.50$ )	$\Delta q_1 = -0.0034^{+0.0083}_{-0.0073}$ ( $q_1 = 0.60$ )		$\Delta R_1 = 0.0321^{+0.0773}_{-0.0633}$ ( $R_1 = 2.50$ )	$\Delta \alpha = 0.0321^{+0.0437}_{-0.0412}$ ( $\alpha = 2.50$ )	
<b>LensMassShear</b> <sub>HS40L80Disk</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0031^{+0.0050}_{-0.0040}$ ( $\theta_{\text{Ein}} = 1.15$ )	$\Delta q = -0.0093^{+0.0156}_{-0.0160}$ ( $q = 0.95$ )		$\Delta \alpha = 0.0467^{+0.0754}_{-0.0732}$ ( $\alpha = 1.92$ )	$\Delta \alpha = 0.0467^{+0.0165}_{-0.0197}$ ( $\alpha = 1.92$ )	
<b>LensMassShear</b> <sub>HS40L80Disk</sub>	Mass (Shear)	$\Delta \theta_{\text{sh}} = -1.3846^{+3.1279}_{-2.4625}$ ( $\theta_{\text{sh}} = 40.00$ )	$\Delta \gamma_{\text{sh}} = 0.0019^{+0.0044}_{-0.0055}$ ( $\gamma_{\text{sh}} = 0.03$ )		$\Delta \gamma_{\text{sh}} = 0.0019^{+0.0025}_{-0.0018}$ ( $\gamma_{\text{sh}} = 0.03$ )		
<b>LensMassShear</b> <sub>ES40L80Disk</sub>	Light (Sersic)	$\Delta R_1 = 0.1845^{+0.1433}_{-0.1416}$ ( $R_1 = 1.50$ )	$\Delta q_1 = -0.0029^{+0.0175}_{-0.0177}$ ( $q_1 = 0.60$ )		$\Delta R_1 = 0.0777^{+0.1237}_{-0.1254}$ ( $R_1 = 2.50$ )	$\Delta \alpha = 0.0777^{+0.0437}_{-0.0412}$ ( $\alpha = 2.50$ )	
<b>LensMassShear</b> <sub>ES40L80Disk</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0040^{+0.0060}_{-0.0045}$ ( $\theta_{\text{Ein}} = 1.15$ )	$\Delta q = -0.0101^{+0.0262}_{-0.0270}$ ( $q = 0.95$ )		$\Delta \alpha = 0.2562^{+0.1829}_{-0.1761}$ ( $\alpha = 1.92$ )	$\Delta \alpha = 0.2562^{+0.0573}_{-0.0589}$ ( $\alpha = 1.92$ )	
<b>LensMassShear</b> <sub>ES40L80Disk</sub>	Mass (Shear)	$\Delta \theta_{\text{sh}} = -3.8668^{+5.0033}_{-4.3041}$ ( $\theta_{\text{sh}} = 40.00$ )	$\Delta \gamma_{\text{sh}} = 0.0049^{+0.0064}_{-0.0062}$ ( $\gamma_{\text{sh}} = 0.03$ )		$\Delta \gamma_{\text{sh}} = 0.0049^{+0.0025}_{-0.0023}$ ( $\gamma_{\text{sh}} = 0.03$ )		
<b>LensLightBD</b> <sub>HS25L50BD</sub>	Light (Sersic)	$\Delta R_1 = 0.0202^{+0.2174}_{-0.2309}$ ( $R_1 = 0.40$ )	$\Delta q_1 = 0.0064^{+0.0902}_{-0.0827}$ ( $q_1 = 0.74$ )		$\Delta R_1 = -0.0535^{+0.6778}_{-0.6973}$ ( $R_1 = 3.00$ )	$\Delta \alpha = -0.0535^{+0.2493}_{-0.2215}$ ( $\alpha = 3.00$ )	
<b>LensLightBD</b> <sub>HS25L50BD</sub>	Light (Exp)	$\Delta R_2 = 0.0849^{+0.0783}_{-0.0681}$ ( $R_2 = 1.15$ )	$\Delta q_2 = -0.0112^{+0.0360}_{-0.0343}$ ( $q_2 = 0.80$ )		$\Delta R_2 = 0.0849^{+0.0218}_{-0.0235}$ ( $R_2 = 1.15$ )	$\Delta \alpha = 0.0849^{+0.0270}_{-0.0347}$ ( $\alpha = 2.05$ )	
<b>LensLightBD</b> <sub>HS25L50BD</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0092^{+0.0190}_{-0.0148}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0173^{+0.0397}_{-0.0408}$ ( $q = 0.80$ )		$\Delta \alpha = 0.0448^{+0.0980}_{-0.0850}$ ( $\alpha = 2.05$ )	$\Delta \alpha = 0.0448^{+0.0270}_{-0.0347}$ ( $\alpha = 2.05$ )	
<b>LensLightBD</b> <sub>ES25L50BD</sub>	Light (Sersic)	$\Delta R_1 = -0.1312^{+0.1404}_{-0.1266}$ ( $R_1 = 0.40$ )	$\Delta q_1 = -0.0197^{+0.1134}_{-0.1045}$ ( $q_1 = 0.74$ )		$\Delta R_1 = -0.6174^{+0.9480}_{-1.0253}$ ( $R_1 = 3.00$ )	$\Delta \alpha = -0.6174^{+0.3687}_{-0.3285}$ ( $\alpha = 3.00$ )	
<b>LensLightBD</b> <sub>ES25L50BD</sub>	Light (Exp)	$\Delta R_2 = 0.0618^{+0.0713}_{-0.0636}$ ( $R_2 = 1.15$ )	$\Delta q_2 = 0.0002^{+0.0563}_{-0.0307}$ ( $q_2 = 0.80$ )		$\Delta R_2 = 0.0618^{+0.0219}_{-0.0225}$ ( $R_2 = 1.15$ )	$\Delta \alpha = 0.0618^{+0.0219}_{-0.0225}$ ( $\alpha = 1.15$ )	
<b>LensLightBD</b> <sub>ES25L50BD</sub>	Mass (SPL)	$\Delta \theta_{\text{Ein}} = 0.0092^{+0.0330}_{-0.0280}$ ( $\theta_{\text{Ein}} = 1.20$ )	$\Delta q = -0.0190^{+0.0654}_{-0.0744}$ ( $q = 0.80$ )		$\Delta \alpha = 0.0372^{+0.1628}_{-0.1654}$ ( $\alpha = 2.05$ )	$\Delta \alpha = 0.0372^{+0.0595}_{-0.0646}$ ( $\alpha = 2.05$ )	



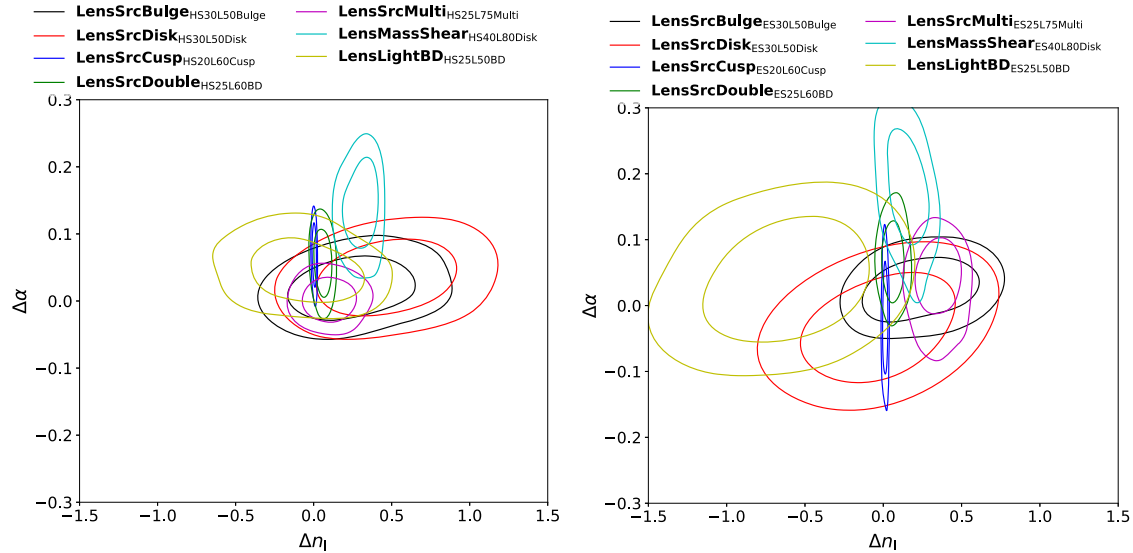
**Figure 14.** The model images, residual,  $\chi^2$  images and source reconstructions for the images **LensSrcBulge<sub>HS30L50Bulge</sub>** (top row), **LensLightBD<sub>HS25L50BD</sub>** (middle row) and **LensSrcDouble<sub>ES25L60BD</sub>** (bottom row). Images correspond to the most likely model at the end of phase 2 of the main pipeline, corresponding to the models given by rows 1 – 2, 27 – 29, 15 – 16 and 11 – 12 of Table 7 respectively.



**Figure 15.** Marginalized 1-D PDF's of  $\Delta n_1$  (or  $\Delta n_{11}$  for images of the model **LensLightBD**) for the seven Hubble resolution (left) and seven Euclid resolution (right) images of the **Lens** model images, corresponding to the results given in Table 7. Each graph's legend indicates the colour each line corresponds to. The input value of  $n_1$  for each model are give in Table 1. The majority of PDF's are consistent with the input lens model ( $\Delta n_1 = 0.0$ ). The width of PDFs shows no trend with image S/N or resolution, instead it is the size of the lens galaxy which drives the precision of parameter estimates.

caution. Multiwavelength imaging may be able to decouple the lens and source. The mass model for this configuration is still estimated accurately.

Fig. 18 shows the 1-D and 2-D PDFs of  $\Delta\theta_{\text{Ein}}$ ,  $\Delta q$ ,  $\Delta\alpha$ , and  $\Delta S$  for all of the **Core** images. In general, lower-resolution imaging gives wider parameter estimates compared to their higher resolution



**Figure 16.** Marginalized 2-D PDF's of  $\Delta n_1$  versus  $\Delta\alpha$  for the Hubble resolution (left panels) and Euclid resolution (right panels) models of the **Lens** images. Legends at the top of each panel indicate the image that each line corresponds too. Contours give the  $1\sigma$  (interior) and  $3\sigma$  (exterior) confidence regions. The input value of  $n_1$  and  $\alpha$  for each model are give in Table 1. Little to no degeneracy is observed in any of the PDFs, demonstrating that the mass and light models are essentially independent of one another.

**Table 8.** The results of Bayesian model comparison in the phases  $\mathbf{P}_{\text{LightMC}}$  and  $\mathbf{P}_{\text{MassMC}}$  for the six **Core** images, using both the non-cored and cored mass model pipelines. Image names are listed in the first column, with the second column listing whether each row corresponds to a non-cored or cored pipeline run. The third and fourth columns show the Bayesian evidence values (equation 23) computed for each light model and the fifth and sixth columns evidences for comparison of the *SPL*E and *SPL*E + *Shear* mass model. These are then used to determine whether the cored or non-cored model is favoured, as shown by the final column. Values in bold correspond to those chosen by the pipeline, noting that a threshold of 20 must be exceeded to favour a more complex model.

Image	Pipeline Run	Sersic ( $\mathbf{P}_{\text{MCLight}}$ )	Sersic + Exp ( $\mathbf{P}_{\text{MCLight}}$ )	SPL ( $\mathbf{P}_{\text{MCMass}}$ )	SPL + Shear ( $\mathbf{P}_{\text{MCMass}}$ )	Core model chosen?
<b>CoreSrcDisk</b> <sub>HS35L70Disk</sub>	Non-cored	<b>53835.2452</b>	53849.6432	54459.7417	54482.4440	
<b>CoreSrcDisk</b> <sub>HS35L70Disk</sub>	Cored	<b>54505.0945</b>	54500.4490	<b>54549.9967</b>	54554.9393	
<b>CoreSrcDisk</b> <sub>ES35L70Disk</sub>	Non-cored	<b>14791.5954</b>	14769.7912	15013.6825	15021.6143	Yes
<b>CoreSrcDisk</b> <sub>ES35L70Disk</sub>	Cored	<b>15025.1912</b>	15032.7236	<b>15048.2671</b>	15047.9044	Yes
<b>CoreSrcQuad</b> <sub>HS40L60BD</sub>	Non-cored	54019.3266	<b>54125.3480</b>	54551.2353	54554.2885	
<b>CoreSrcQuad</b> <sub>HS40L60BD</sub>	Cored	54376.6692	<b>54632.3307</b>	<b>54706.6518</b>	54709.8142	
<b>CoreSrcQuad</b> <sub>ES40L60BD</sub>	Non-cored	14697.6901	<b>14834.0991</b>	14937.6529	14938.3421	Yes
<b>CoreSrcQuad</b> <sub>ES40L60BD</sub>	Cored	14801.0601	<b>14998.6605</b>	<b>15105.6929</b>	15110.8025	Yes
<b>CoreSrcDouble</b> <sub>HS25L50BD</sub>	Non-cored	53375.7084	<b>53425.0099</b>	54429.5815	54431.3582	
<b>CoreSrcDouble</b> <sub>HS25L50BD</sub>	Cored	<b>54236.5839</b>	54234.0987	<b>54563.2273</b>	54564.5934	
<b>CoreSrcDouble</b> <sub>ES25L50BD</sub>	Non-cored	<b>14454.5533</b>	14458.9280	14867.2487	14868.3826	Yes
<b>CoreSrcDouble</b> <sub>ES25L50BD</sub>	Cored	<b>14795.4639</b>	14809.6472	14903.2335	<b>14931.0893</b>	Yes

counterparts and the **CoreSrcQuad** images are less precisely constrained, due to the lens light blending discussed above.  $S$  becomes another parameter in the mass-profile degeneracy discussed in Section 4 and in N15, offering the mass model an additional means by which to change its mass distribution whilst still integrating to give an accurate  $M_{\text{Ein}}$ . The additional freedom introduced by  $S$  is constrained by the very central regions of the lensed source, either its central image or radial arcs.

Having established that the lens's light and mass profiles are separable for non-cored models, it is interesting to ask whether a degeneracy emerges between the light profile and cored mass model. Fig. 19 shows 2-D PDFs between  $n_1$ ,  $\Delta\alpha$  and  $\Delta S$ . Once again, no degeneracy is observed between the mass and light model parameters (with inspection of other parameter pairs confirming this trend is general). This is initially surprising, but builds on the discussion above that because the lensed source's appearance is non-

symmetric and irregular it shares no degeneracy with the subtraction of a smooth light profile.

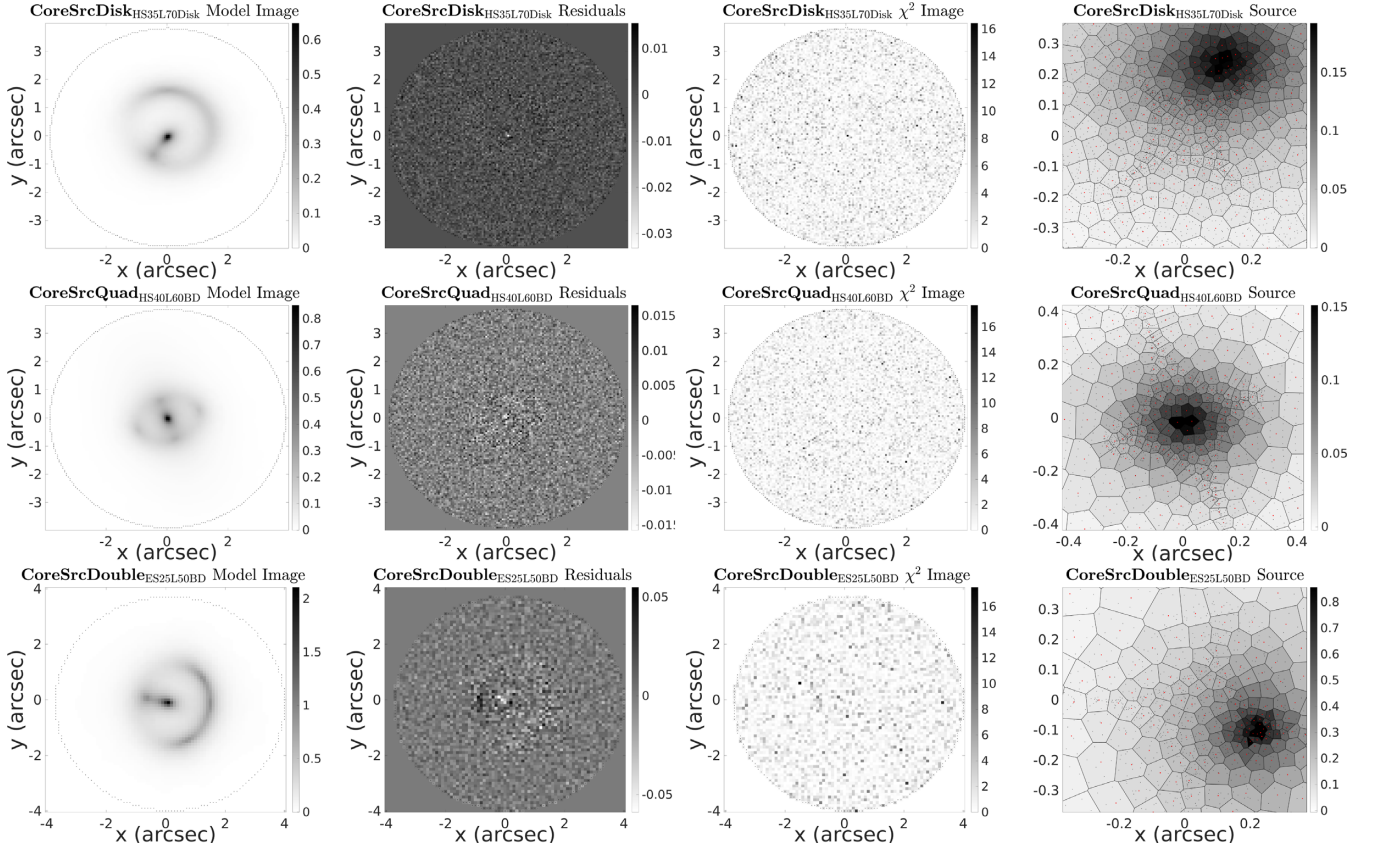
#### 7.4 Decomposed mass models

The decomposed lens simulation suite consists of five unique lens and source models, each of which is used to generate an image at Hubble and Euclid resolution, with source  $S/N=30-40$  and lens  $S/N=50-80$ , giving a total of 10 images. Each image is analyzed using the decomposed profile pipeline, producing a *Light + NFW* model. The results of model comparison are shown in Table 10. The correct light model and inclusion of an external shear are correct for all models, however a number of light and dark matter geometries are not consistent with their input models. Parameter estimates for each image are summarized in Table 11, using  $\Delta R_1$ ,  $\Delta n_1$ ,  $\Delta q_1$ ,  $\Delta\kappa_d$ ,  $\Delta q_d$ , and  $\Delta\Psi_1$ . Fig. 20 shows the observed images, model images,



**Table 9.** Results of fitting the six **Core** model images using cored mass profile automated analysis pipeline, corresponding to results generated at end of phase 2. Each image's name is given in the first column and the light or mass model component in the second column. The third to sixth columns show parameter estimates, where each parameter is offset by  $\Delta P = P_{\text{true}} - P_{\text{model}}$ , such that zero corresponds to the input lens model. The input lens model values are given in brackets to the right of each parameter estimate. Parameters estimates are shown using  $\Delta R_1$ ,  $\Delta q_1$ , and  $\Delta n_1$  for *Sersic* light models,  $\Delta R_{11}$ ,  $\Delta q_{11}$ ,  $\Delta n_{11}$ ,  $\Delta R_{12}$ , and  $\Delta q_{12}$  for *Sersic* + *Exp* light models and  $\Delta S$  and  $\Delta \alpha$  for *P*  $L_{\text{Core}}$  mass models. Columns 3–5 show parameter estimates within  $3\sigma$  confidence and column 6 at  $1\sigma$ . Parameter estimates in bold text are inconsistent with the input lens model at their stated error estimates. The other parameters not shown (e.g.  $\chi_1$ ,  $\theta_1$ ,  $\phi$ ) are all estimated accurately within  $3\sigma$ .

Image	Component	Parameters ( $3\sigma$ )			Parameters ( $1\sigma$ )		
<b>CoreSrcDisk</b> <sub>HS35L70Disk</sub>	Light ( <b>Sersic</b> )	$\Delta R_1 = 0.0647^{+0.2325}_{-0.1977}$ ( $R_1 = 0.7$ )	$\Delta q_1 = 0.0085^{+0.0699}_{-0.0642}$ ( $q_1 = 0.80$ )	$\Delta n_1 = 0.1021^{+0.4380}_{-0.4239}$ ( $n_1 = 2.50$ )	$\Delta R_1 = 0.1021^{+0.4380}_{-0.4239}$ ( $n_1 = 2.50$ )	$\Delta q_1 = 0.1021^{+0.4380}_{-0.4239}$ ( $n_1 = 2.50$ )	$\Delta n_1 = 0.1021^{+0.4380}_{-0.4239}$ ( $n_1 = 2.50$ )
<b>CoreSrcDisk</b> <sub>HS35L70Disk</sub>	Mass ( <b>SPL</b> )	$\Delta S = -0.0191^{+0.0787}_{-0.0809}$ ( $S = 0.20$ )		$\Delta \alpha = -0.0379^{+0.1781}_{-0.1497}$ ( $\alpha = 1.85$ )	$\Delta S = -0.0191^{+0.0787}_{-0.0809}$ ( $S = 0.20$ )	$\Delta \alpha = -0.0379^{+0.1781}_{-0.1497}$ ( $\alpha = 1.85$ )	$\Delta S = -0.0191^{+0.0787}_{-0.0809}$ ( $S = 0.20$ )
<b>CoreSrcDisk</b> <sub>ES35L70Disk</sub>	Light ( <b>Sersic</b> )	$\Delta R_1 = 0.1761^{+0.3080}_{-0.2814}$ ( $R_1 = 0.7$ )	$\Delta q_1 = -0.0079^{+0.0652}_{-0.0640}$ ( $q_1 = 0.80$ )	$\Delta n_1 = 0.3441^{+0.5634}_{-0.5754}$ ( $n_1 = 2.50$ )	$\Delta R_1 = 0.3441^{+0.5634}_{-0.5754}$ ( $n_1 = 2.50$ )	$\Delta q_1 = -0.0079^{+0.0652}_{-0.0640}$ ( $q_1 = 0.80$ )	$\Delta n_1 = 0.3441^{+0.5634}_{-0.5754}$ ( $n_1 = 2.50$ )
<b>CoreSrcDisk</b> <sub>ES35L70Disk</sub>	Mass ( <b>SPL</b> )	$\Delta S = 0.0319^{+0.1282}_{-0.1335}$ ( $S = 0.20$ )		$\Delta \alpha = 0.0556^{+0.2383}_{-0.2510}$ ( $\alpha = 1.85$ )	$\Delta S = 0.0319^{+0.1282}_{-0.1335}$ ( $S = 0.20$ )	$\Delta \alpha = 0.0556^{+0.2383}_{-0.2510}$ ( $\alpha = 1.85$ )	$\Delta S = 0.0319^{+0.1282}_{-0.1335}$ ( $S = 0.20$ )
<b>CoreSrcQuad</b> <sub>HS40L60BD</sub>	Light ( <b>Sersic</b> )	$\Delta R_{11} = -0.0682^{+0.0520}_{-0.0435}$ ( $R_{11} = 0.25$ )	$\Delta q_{11} = 0.0116^{+0.0583}_{-0.0522}$ ( $q_{11} = 0.77$ )	$\Delta n_{11} = -0.7565^{+0.4355}_{-0.3844}$ ( $n_{11} = 2.50$ )	$\Delta R_{11} = -0.0682^{+0.0520}_{-0.0435}$ ( $R_{11} = 0.25$ )	$\Delta q_{11} = 0.0116^{+0.0583}_{-0.0522}$ ( $q_{11} = 0.77$ )	$\Delta n_{11} = -0.7565^{+0.4355}_{-0.3844}$ ( $n_{11} = 2.50$ )
<b>CoreSrcQuad</b> <sub>HS40L60BD</sub>	Light ( <b>Exp</b> )	$\Delta R_{12} = -0.0307^{+0.0990}_{-0.0830}$ ( $R_{12} = 1.35$ )	$\Delta q_{12} = 0.0802^{+0.0278}_{-0.0335}$ ( $q_{12} = 0.60$ )		$\Delta R_{12} = -0.0307^{+0.0990}_{-0.0830}$ ( $R_{12} = 1.35$ )	$\Delta q_{12} = 0.0802^{+0.0278}_{-0.0335}$ ( $q_{12} = 0.60$ )	$\Delta R_{12} = -0.0307^{+0.0990}_{-0.0830}$ ( $R_{12} = 1.35$ )
<b>CoreSrcQuad</b> <sub>HS40L60BD</sub>	Mass ( <b>SPL</b> )	$\Delta S = -0.0060^{+0.1476}_{-0.1412}$ ( $S = 0.30$ )		$\Delta \alpha = 0.0487^{+0.1450}_{-0.1007}$ ( $\alpha = 1.75$ )	$\Delta S = -0.0060^{+0.1476}_{-0.1412}$ ( $S = 0.30$ )	$\Delta \alpha = 0.0487^{+0.1450}_{-0.1007}$ ( $\alpha = 1.75$ )	$\Delta S = -0.0060^{+0.1476}_{-0.1412}$ ( $S = 0.30$ )
<b>CoreSrcQuad</b> <sub>ES40L60BD</sub>	Light ( <b>Sersic</b> )	$\Delta R_{11} = -0.0941^{+0.0372}_{-0.0236}$ ( $R_{11} = 0.25$ )	$\Delta q_{11} = -0.0125^{+0.0663}_{-0.0576}$ ( $q_{11} = 0.77$ )	$\Delta n_{11} = -1.1441^{+0.3980}_{-0.3459}$ ( $n_{11} = 2.50$ )	$\Delta R_{11} = -0.0941^{+0.0372}_{-0.0236}$ ( $R_{11} = 0.25$ )	$\Delta q_{11} = -0.0125^{+0.0663}_{-0.0576}$ ( $q_{11} = 0.77$ )	$\Delta n_{11} = -1.1441^{+0.3980}_{-0.3459}$ ( $n_{11} = 2.50$ )
<b>CoreSrcQuad</b> <sub>ES40L60BD</sub>	Light ( <b>Exp</b> )	$\Delta R_{12} = -0.0934^{+0.0482}_{-0.0489}$ ( $R_{12} = 1.35$ )	$\Delta q_{12} = 0.0673^{+0.0163}_{-0.0176}$ ( $q_{12} = 0.60$ )		$\Delta R_{12} = -0.0934^{+0.0482}_{-0.0489}$ ( $R_{12} = 1.35$ )	$\Delta q_{12} = 0.0673^{+0.0163}_{-0.0176}$ ( $q_{12} = 0.60$ )	$\Delta R_{12} = -0.0934^{+0.0482}_{-0.0489}$ ( $R_{12} = 1.35$ )
<b>CoreSrcQuad</b> <sub>ES40L60BD</sub>	Mass ( <b>SPL</b> )	$\Delta S = -0.1168^{+0.1323}_{-0.1468}$ ( $S = 0.30$ )		$\Delta \alpha = -0.0457^{+0.1072}_{-0.1067}$ ( $\alpha = 1.75$ )	$\Delta S = -0.1168^{+0.1323}_{-0.1468}$ ( $S = 0.30$ )	$\Delta \alpha = -0.0457^{+0.1072}_{-0.1067}$ ( $\alpha = 1.75$ )	$\Delta S = -0.1168^{+0.1323}_{-0.1468}$ ( $S = 0.30$ )
<b>CoreSrcDouble</b> <sub>HS25L50BD</sub>	Light ( <b>Sersic</b> )	$\Delta R_1 = 0.1938^{+0.7768}_{-0.6668}$ ( $R_1 = 1.50$ )	$\Delta q_1 = 0.0006^{+0.0461}_{-0.0482}$ ( $q_1 = 0.70$ )		$\Delta R_1 = 0.1938^{+0.7768}_{-0.6668}$ ( $R_1 = 1.50$ )	$\Delta q_1 = 0.0006^{+0.0461}_{-0.0482}$ ( $q_1 = 0.70$ )	$\Delta R_1 = 0.1938^{+0.7768}_{-0.6668}$ ( $R_1 = 1.50$ )
<b>CoreSrcDouble</b> <sub>HS25L50BD</sub>	Mass ( <b>SPL</b> )	$\Delta S = -0.0099^{+0.0349}_{-0.0350}$ ( $S = 0.25$ )		$\Delta \alpha = -0.0096^{+0.0631}_{-0.0691}$ ( $\alpha = 1.65$ )	$\Delta S = -0.0099^{+0.0349}_{-0.0350}$ ( $S = 0.25$ )	$\Delta \alpha = -0.0096^{+0.0631}_{-0.0691}$ ( $\alpha = 1.65$ )	$\Delta S = -0.0099^{+0.0349}_{-0.0350}$ ( $S = 0.25$ )
<b>CoreSrcDouble</b> <sub>ES25L50BD</sub>	Light ( <b>Sersic</b> )	$\Delta R_1 = 0.1850^{+0.8773}_{-0.8788}$ ( $R_1 = 1.50$ )	$\Delta q_1 = 0.0057^{+0.0754}_{-0.0716}$ ( $q_1 = 0.70$ )		$\Delta R_1 = 0.1850^{+0.8773}_{-0.8788}$ ( $R_1 = 1.50$ )	$\Delta q_1 = 0.0057^{+0.0754}_{-0.0716}$ ( $q_1 = 0.70$ )	$\Delta R_1 = 0.1850^{+0.8773}_{-0.8788}$ ( $R_1 = 1.50$ )
<b>CoreSrcDouble</b> <sub>ES25L50BD</sub>	Mass ( <b>SPL</b> )	$\Delta S = -0.0106^{+0.0979}_{-0.1035}$ ( $S = 0.25$ )		$\Delta \alpha = -0.0071^{+0.1498}_{-0.1654}$ ( $\alpha = 1.65$ )	$\Delta S = -0.0106^{+0.0979}_{-0.1035}$ ( $S = 0.25$ )	$\Delta \alpha = -0.0071^{+0.1498}_{-0.1654}$ ( $\alpha = 1.65$ )	$\Delta S = -0.0106^{+0.0979}_{-0.1035}$ ( $S = 0.25$ )



**Figure 17.** The model image, residuals,  $\chi^2$  image, and source reconstructions for the analysis of the images **CoreSrcDisk**<sub>HS35L70Disk</sub> (top row), **CoreSrcQuad**<sub>HS40L60BD</sub> (middle row), and **CoreSrcDouble**<sub>ES25L50BD</sub> (bottom row). Images correspond to the images of the main pipeline, corresponding to the models given by rows 1–2, 5–8, and 13–14 of Table 9, respectively.

model sources, residuals,  $\chi^2$  images, and source reconstructions for three images, showing residuals and  $\chi^2$  images that realize the image's noise.

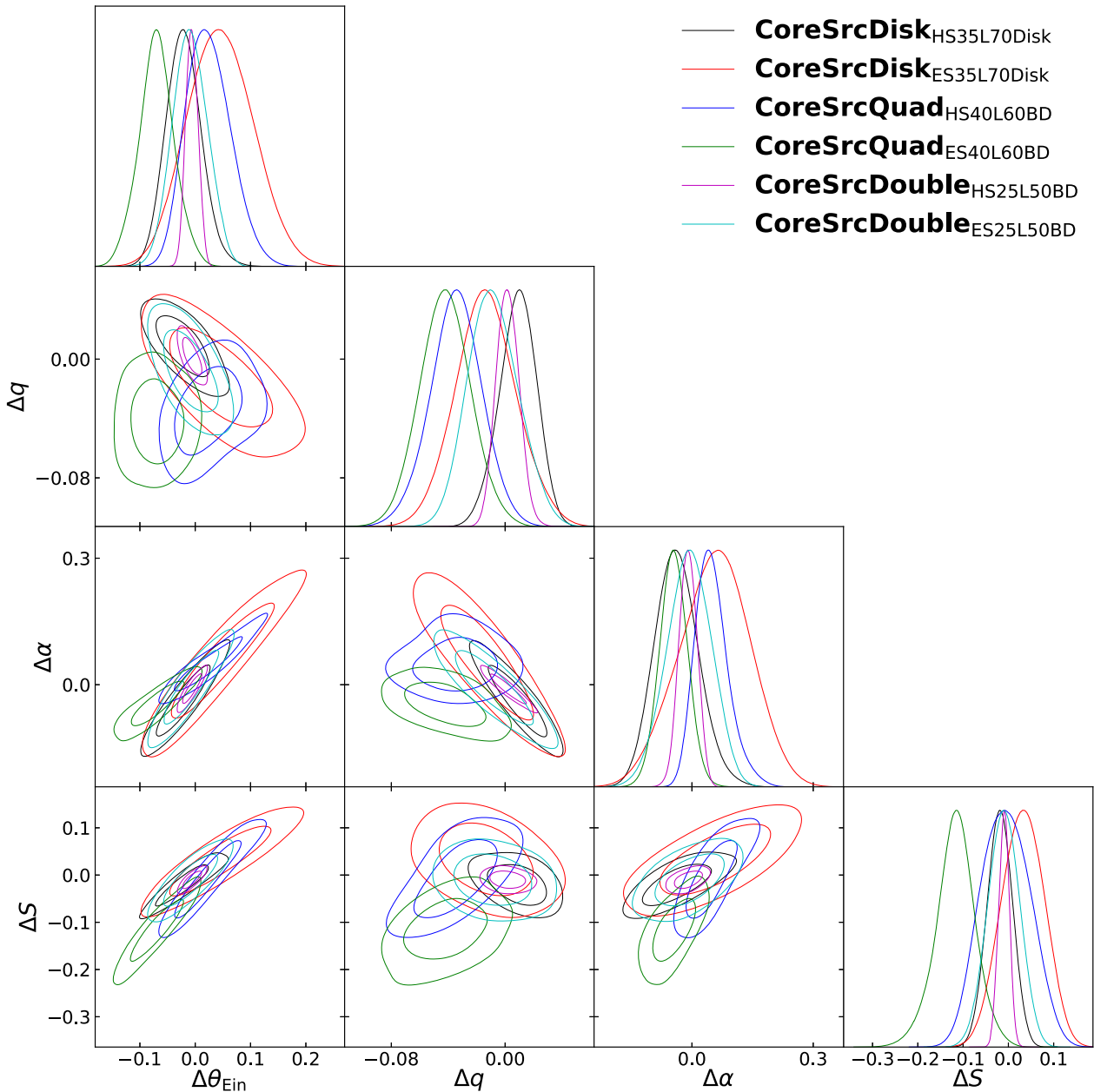
#### 7.4.1 Model comparison

The results of the model comparison phase  $\mathbf{P}_{\text{MCGeom}}$  for all images is shown in Table 10. The input light profiles are chosen for every image, and the shear is chosen correctly for 9 out of 10 images. However, for five images, an offset in their light and dark matter geometry is not chosen, even though they are present in their lens's input profiles. In all cases, model comparison opts to choose the simpler (aligned) model over the input geometrically offset models. Therefore, the non-inclusion of these components can simply be attributed to the observed image's resolution and S/N being insufficient to offer a large enough increase in evidence to favour the more complex model. In many cases, the input model does increase the Bayesian evidence, but does not meet the threshold value of 20. Nevertheless, detections are made for the high-resolution images of the **LMDMRot90** and **LMDMPos** images, demonstrating that geometric offsets can be detected in strong lens imaging, but that such a detection requires higher quality data than the models discussed previously.

Whilst image quality is a driving factor in detecting geometric offsets, another important aspect is the presence of an external shear. In the absence of an external shear, no degeneracy is observed between the geometric parameters  $\Delta x_d$ ,  $\Delta y_d$ , and  $\Delta \theta_d$  and

those governing the lens's light or mass profiles. This is important, because it suggests that when inferring a geometric misalignment any assumptions related to the lens's mass distribution (e.g. the form of the dark matter profile or use of a constant mass-to-light ratio) may not be very important. However, when a shear is present, this is found to no longer be the case and a degeneracy emerges between the dark matter geometry parameters, the shear parameters  $\gamma_{\text{sh}}$  and  $\theta_{\text{sh}}$ , and the mass profile's other parameters. Degeneracies between ellipticity, rotational misalignments, and an external shear have been long established (Keeton, Kochanek & Seljak 1997) and they are particularly problematic for decomposed mass modeling as it becomes the dark matter profile that makes up a smaller fraction of overall mass where one is trying to detect them.

The degeneracy between dark matter geometry and an external shear also explains why analysis of the high-resolution **LMDMRot90** image incorrectly includes an external shear in the mass model. The *Shear* component is chosen earlier in the pipeline using the axisymmetric *NFWSpH* profile, thus the *Shear* is included as it mimics the effect of the rotational misalignment that the *NFWSpH* cannot capture. Later in the pipeline, when the *NFW + Light* model with a rotational offset is chosen, the shear magnitude  $\gamma_{\text{sh}}$  reduces to nearly zero, effectively removing the shear and giving an accurate lens model. In the future, it may prove beneficial to choose an external shear via model comparison independently in the  $\mathbf{P}_{\text{MCGeom}}$  phase. This strategy will be considered in the future, where independent constraints on the shear from weak-lensing will also be considered.



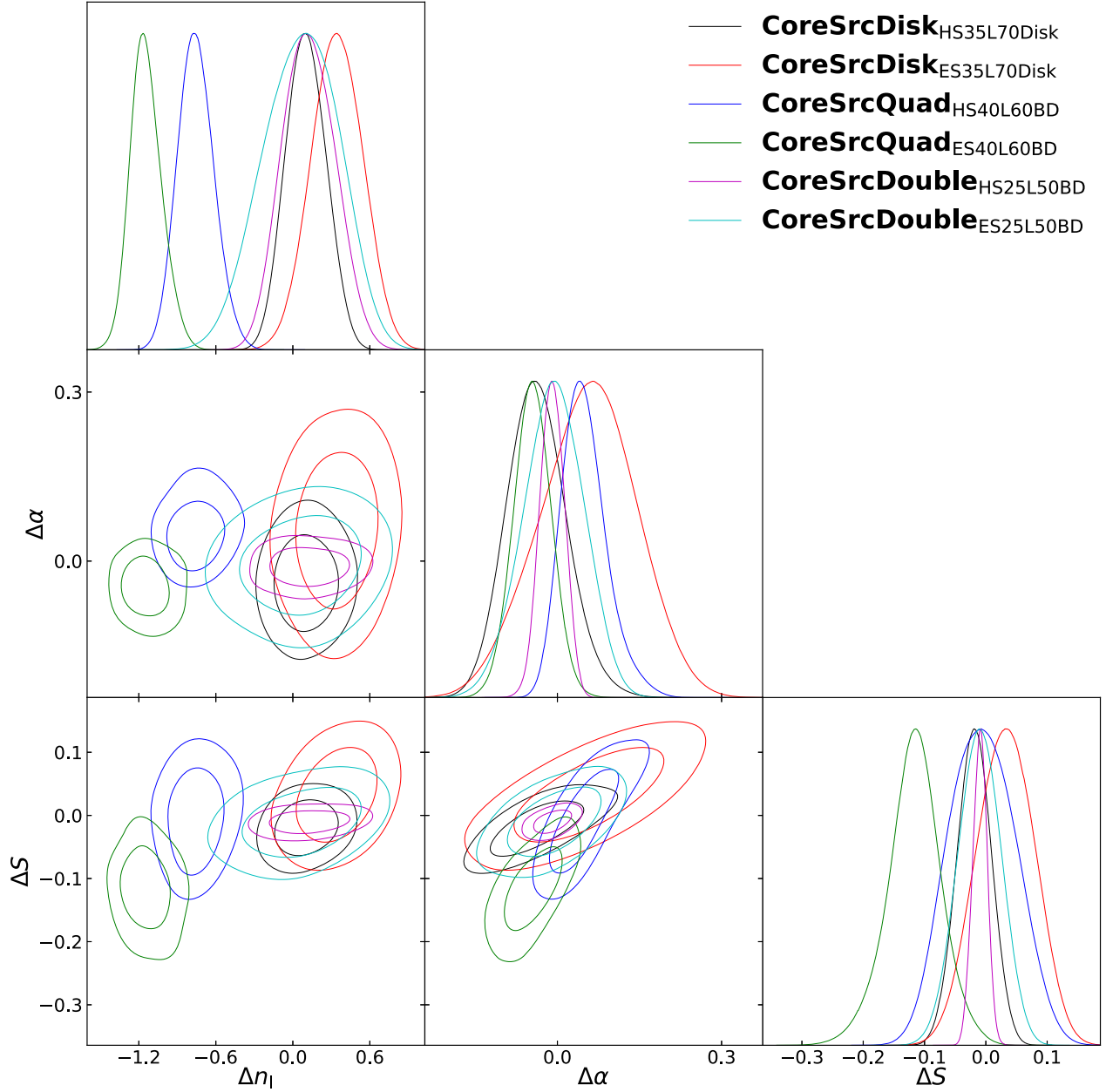
**Figure 18.** Marginalized 2-D PDF's of  $\Delta\theta_{\text{Ein}}$ ,  $\Delta q$ ,  $\Delta\alpha$ , and  $\Delta S$  for the images of the **Core** models. The top-right legend indicates the image that each coloured line corresponds to. Contours give the  $1\sigma$  (interior) and  $3\sigma$  (exterior) confidence regions. For **CoreSrcDisk** the input values of each parameter are  $\theta_{\text{Ein}} = 1.4$  arcsec,  $q = 0.8$ ,  $\alpha = 1.85$ , and  $S = 0.2$  arcsec, for **CoreSrcQuad**  $\theta_{\text{Ein}} = 1.0$  arcsec,  $q = 0.7$ ,  $\alpha = 1.75$ , and  $S = 0.3$  arcsec and for **CoreSrcDouble**  $\theta_{\text{Ein}} = 1.3$  arcsec,  $q = 0.8$ ,  $\alpha = 1.65$ , and  $S = 0.25$  arcsec. A degeneracy can be seen between all parameters, which is an extension of the degeneracy between mass, ellipticity and slope described in N15, but including also the core radius  $S$ .

In summary, **AutoLens** can successfully detect geometrically misaligned light and dark matter profiles. This is because, in an analogous fashion to the cored profile earlier, geometrically misaligned mass components impart unique features into the lensed source's extended light profile that an axisymmetric model cannot fit, especially given the constraints placed on the light profile's geometry due to lens light fitting. However, based on these results, it is clear that the prospects for detecting geometric offsets (if present in nature) are heavily dependent on the quality of the imaging data, the size of the offset, and the presence and magnitude of an external shear. Such an analysis may be beyond the reach of Euclid wide-field imaging, but within the realms of possibility for cur-

rently available Hubble imaging. Additional information from an independent measurement, such as weak-lensing, may be crucial, as it can offer independent constraints on the shear magnitude and direction.

#### 7.4.2 Modeling results

The results of the parameter estimates for the decomposed mass models are given in Table 11. For the first seven images, the results are positive, with the majority of lens model parameters estimated correctly within  $3\sigma$ . However, the image **LMDMRot90**<sub>ES30L75Multi</sub>, and both of the **LMDMShear** images, are poorly estimated. In



**Figure 19.** Marginalized 2-D PDFs of  $\Delta n_1$ ,  $\Delta\alpha$ , and  $\Delta S$  for the images of the **Core** models. The top-right legend indicates the image that each coloured line corresponds to. Contours give the  $1\sigma$  (interior) and  $3\sigma$  (exterior) confidence regions. For **CoreSrcDisk** the input values of each parameter are  $n_1 = 2.5$ ,  $\alpha = 1.85$ , and  $S = 0.2$  arcsec, for **CoreSrcQuad**  $n_1 = 2.5$ ,  $\alpha = 1.75$ , and  $S = 0.3$  arcsec and for **CoreSrcDouble**  $n_1 = 3.5$ ,  $\alpha = 1.65$ , and  $S = 0.25$  arcsec. No degeneracy can be seen between  $n_1$  and the mass-profile parameters, suggesting that there is no degeneracy between a light model and cored mass model.

all cases, a rotational misalignment was not chosen when present in the input image, which likely contributes to this result. The large and inaccurate values of  $x_d$  and  $y_d$  for the high-resolution **LMDMShear** image also suggest that the analysis is unable to accurately resolve the degeneracy between geometric offset and external shear. This result reaffirms the caution that must be taken when attempting to model an external shear and geometric offset.

The most significant parameter degeneracies of the previous mass models were found between the parameters governing their mass distributions, which consisted of either three or four parameters. For

a decomposed mass model, seven (or more for a multicomponent light profile) parameters determine the lens's mass distribution;  $I_1$ ,  $R_1$ ,  $n_1$ ,  $q_1$ ,  $\kappa_d$ ,  $q_d$ , and  $\Psi_1$ . This would create an extremely complex and degenerate non-linear parameter space, from which lens models constraints are not possible, if it were not for the constraints placed on the light model parameters by the light profile's fit to the lens galaxy. Shown by Fig. 21 for the low-resolution images, this results in essentially no degeneracy emerging between the light profile parameters ( $I_1$ ,  $R_1$ ,  $n_1$ ) and dark matter parameters ( $\kappa_d$ ,  $q_d$ ), with their degeneracy instead folded into the mass-to-light ratio  $\Psi_1$ .

**Table 10.** The results of Bayesian model comparison in the phases  $\mathbf{P}_{\text{LightMC}}$ ,  $\mathbf{P}_{\text{MassMC}}$ , and  $\mathbf{P}_{\text{GeomMC}}$  for the 10 LMDM images, which test decomposed mass modeling. Image names are listed in the first column. The second and third columns show the Bayesian evidence values (equation 23) computed for the *Sersic* and *Sersic + Exp* light models, the fourth and fifth columns evidences for comparison of the *SPLE* and *SPLE + Shear* mass models and final three columns the evidences for the geometrically aligned, rotationally offset and/or positionally offset *Light + NFW* model. Values in bold correspond to those chosen by the pipeline, noting that a threshold of 20 must be exceeded to favour a more complex model.

Image	Sersic ( $\mathbf{P}_{\text{MCLight}}$ )	Sersic + Exp ( $\mathbf{P}_{\text{MCLight}}$ )	SPLE ( $\mathbf{P}_{\text{MCMass}}$ )	SPLE + Shear ( $\mathbf{P}_{\text{MCMass}}$ )	NFW (Align) ( $\mathbf{P}_{\text{MCGeom}}$ )	NFW (Rot) ( $\mathbf{P}_{\text{MCGeom}}$ )	NFW (Pos) ( $\mathbf{P}_{\text{MCGeom}}$ )
LMDMAlign <sub>HS40L50BD</sub>	<b>55633.3189</b>	55640.7631	<b>55647.3234</b>	55651.7835	<b>55643.8825</b>	55644.4506	55651.3071
LMDMAlign <sub>ES40L50BD</sub>	<b>15820.3812</b>	15824.0408	<b>15843.2371</b>	15841.4428	<b>15831.0632</b>	15833.0704	15829.2484
LMDMPos <sub>HS40L50BD</sub>	<b>55628.4640</b>	55632.5936	<b>55655.0241</b>	55657.4802	55612.5394	55612.7578	<b>55659.5510</b>
LMDMPos <sub>ES40L50BD</sub>	<b>15919.0649</b>	15920.2272	<b>15944.3584</b>	15948.5535	<b>15913.5722</b>	15910.0842	15928.2960
LMDMRot <sub>HS40L50BD</sub>	<b>55413.7926</b>	55424.4787	<b>55469.5119</b>	55467.4304	<b>55449.3334</b>	55456.2422	55443.2313
LMDMRot <sub>ES40L50BD</sub>	<b>15840.9690</b>	15842.9032	<b>15851.6601</b>	15852.1060	<b>15840.3070</b>	15842.4563	15841.4788
LMDMRot90 <sub>HS30L75Multi</sub>	<b>54794.3025</b>	54795.2551	55195.8635	<b>55265.1315</b>	55263.0546	<b>55287.5303</b>	55262.7623
LMDMRot90 <sub>ES30L75Multi</sub>	<b>15070.9989</b>	15066.1369	<b>15367.6957</b>	15379.6356	<b>15372.5886</b>	15371.7799	15377.4355
LMDMShear <sub>HS35L80Cusp</sub>	<b>55363.3473</b>	55341.4098	55462.1340	<b>55593.4590</b>	55538.4963	55548.8675	<b>55577.3864</b>
LMDMShear <sub>ES35L80Cusp</sub>	<b>15621.6493</b>	15620.9123	15654.5684	<b>15695.2625</b>	<b>15686.0568</b>	15690.4766	15695.1630

## 8 DISCUSSION AND SUMMARY

### 8.1 Automated analysis

All results were generated in this work without user intervention, demonstrating that *AutoLens* successfully automates the lens modeling process. High-quality imaging of many hundreds of strong lenses exist in the *HST* archive, a data set that has not been fully exploited due to the time overheads historically associated with lens modeling. Therefore, in the short term, *AutoLens* can significantly increase the total number of strong lenses with a complete lens model and source reconstruction. In the long term, consideration must be given to how this automated framework will be expanded to samples of lenses in the tens of thousands, which will be provided by surveys such as *Euclid* in less than 5 yr (Collett & Auger 2014).

Of course, one can envisage scenarios where the automated analysis, as presented here, breaks down. These include the presence of luminous or dark high-mass substructures in the lens galaxy or lens morphologies whose light profile cannot be fitted accurately with the models used here (e.g. late-type lenses with complicated structure due to star formation and dust lanes). Therefore, as *AutoLens* is expanded to larger samples, it will inevitably require new functionality and the design philosophy, so far, has been to always develop automated solutions that offer a choice between precision and run-time. *AutoLens*'s modular structure facilitates this, by breaking down the complex and iterative nature of lens modeling into a simple set of self-contained phases ensuring that new functionality can be introduced. Equally, the model comparison framework means that adding more complex light and mass profiles does not require a large reworking of each *AutoLens* pipeline. However, it remains to be seen whether the most complex lens modeling tasks can be performed in a truly automated fashion.

The total run-time for a full *AutoLens* analysis depends critically on the total number of image pixels in the analysis and is therefore driven by two factors: (i) the image resolution and (ii) the overall size of the lens and lensed source. The total run time for a unified lens and source analysis in this work ranged from 15 to 300 CPU hours, which, when run in parallel across eight cores, translates to 2–40+ real-time hours. The large samples that will be provided by *Euclid* will fall in the faster range of run times, whereas currently available high-resolution *HST* imaging could exceed these figures if drizzled to resolutions higher than those used in this work.

Thus, processing large lens data sets is already feasible on modern high-performance computing facilities, even for samples in the tens of thousands. Increasing the analysis speed further would be beneficial and is possible, through either advances in computational processing like graphical processing units or by using a reduced and simpler analysis pipeline with more restrictive lens model priors.

Another pressing issue is continued testing of *AutoLens*. The simulated images used throughout this work were generated using the same mass models that were then ultimately fitted, circumventing issues like the MST and not testing assumptions like a constant mass-to-light ratio. Thus, they are somewhat uninformative in revealing what a strong lens analysis actually measures for a real lens galaxy. The assumptions underlying lens modeling with *AutoLens* will be discussed in future publications, where they can be considered in more detail by comparing and contrasting a range of different lens models with different underlying assumptions, as well as looking for additional guidance from independent mass probes. However, the ideal means of testing this analysis will use simulated lenses generated via ray-tracing through realistic cosmological hydrodynamical simulations (Schaye et al. 2015) and it is anticipated *AutoLens* will take part in such work in the future. This work also bypassed a number of instrumental effects that could have the potential to bias the lens model, like a poor PSF sampling or image artefacts.

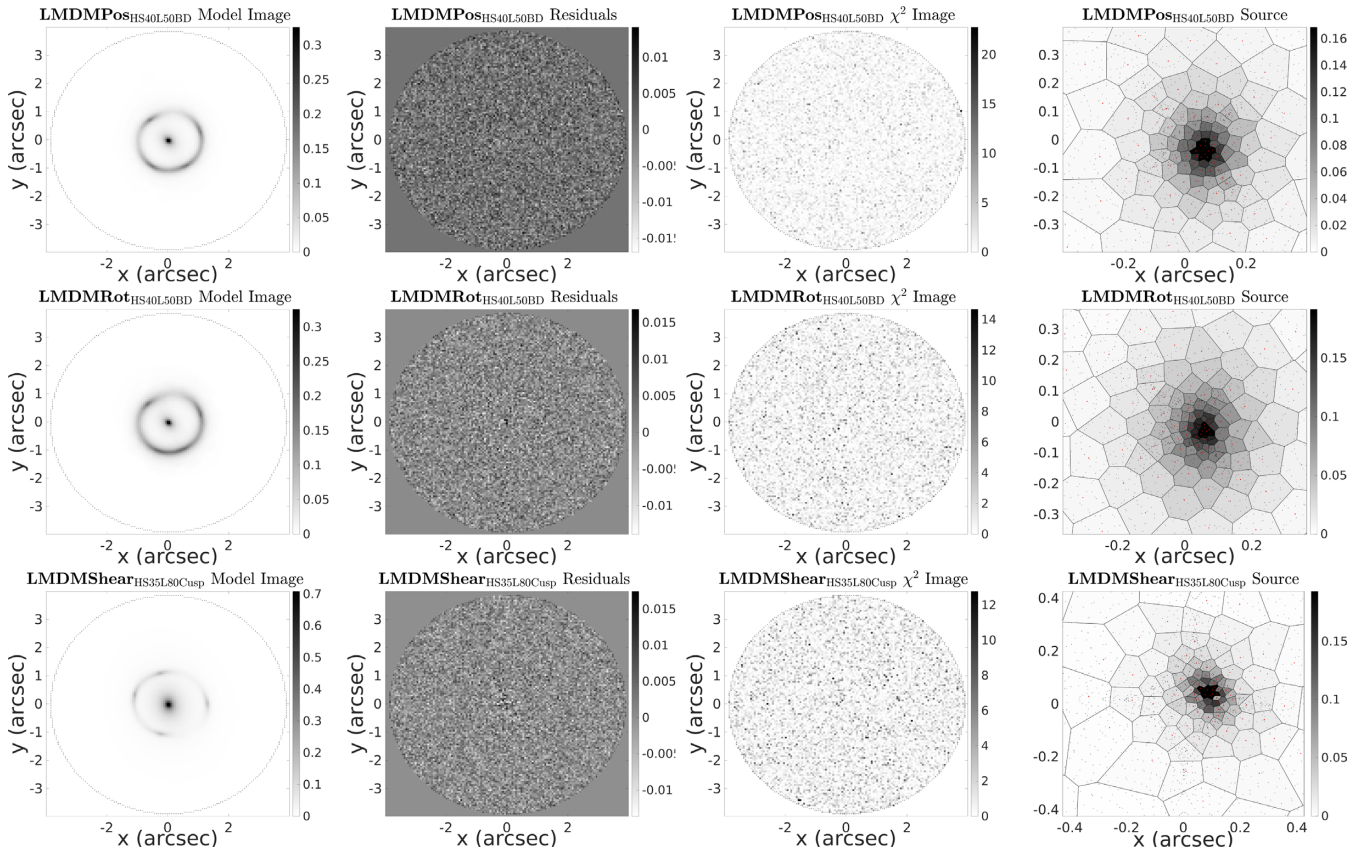
### 8.2 Unified modeling

The results of unified modeling were highly successful and motivate existing high-resolution lens samples for *AutoLens*'s first analyses of known strong lenses. These samples have superior spatial resolution, S/N (e.g. Bolton et al. 2008) and more source complexity (Newton et al. 2011) than the images simulated in this work, as well as average lens and source redshifts below this work's fiducial values of  $z_{\text{lens}} = 0.5$  and  $z_{\text{src}} = 1.0$ . Altogether, this should improve the precision of each lens model compared to the values quoted in this work and may even allow for more complex lens models to be invoked, for example assuming a non-constant mass-to-light ratio. The lower redshifts also reduce the average external shear across the sample (Jaroszynski & Kostrzewa-Rutkowska 2012), thus reducing its degeneracy with the mass profile ellipticity and, if present, any geometric offsets.

For the *SPLE* profile, the mass and light models were independent of one another, confirming that the lens light subtraction has

**Table 11.** Results of fitting the 10 LMDM model images using the automated analysis pipeline, corresponding to results generated at the end of phase 2. Each image's name is given in the first column and the light or mass model component in the second column. The third to sixth columns show parameter estimates, where each parameter is offset by  $\Delta P = P_{\text{true}} - P_{\text{model}}$ , such that zero corresponds to the input lens model. The input lens model values are given in brackets to the right of each parameter estimate. Parameters estimates are shown using  $\Delta R_i$ ,  $\Delta q_i$ , and  $\Delta \eta_i$  for *Sersic* light models,  $\Delta R_{1i}$ ,  $\Delta q_{1i}$ , and  $\Delta \eta_{1i}$  for *Sersic* light models,  $\Delta R_i$ ,  $\Delta q_i$ , and  $\Delta \eta_i$  for *NFW* mass models,  $\Delta \kappa_d$  and  $\Delta q_d$  for *NFW* mass models,  $\Psi_1$  for light mass models, and  $\gamma_{\text{sh}}$  and  $\theta_{\text{sh}}$  for a *Shear* component. Columns 3–5 show parameter estimates within  $3\sigma$  confidence and column 6 at  $1\sigma$ . Parameter estimates in bold text are inconsistent with the input lens model at their stated error estimates. The other parameters not shown (e.g.  $x_i$ ,  $\theta_i$ ) are all estimated accurately within  $3\sigma$ .

Image	Component	Parameters ( $3\sigma$ )				Parameters ( $1\sigma$ )
LMDMalign <sub>HS40L50BD</sub>	Light (Sersic)	$\Delta R_1 = 0.0675^{+0.1676}_{-0.1667}$ ( $R_1 = 0.60$ )	$\Delta q_1 = -0.0033^{+0.0396}_{-0.0357}$ ( $q_1 = 0.72$ )	$\Delta \eta_1 = 0.0587^{+0.5377}_{-0.5343}$ ( $\eta_1 = 4.00$ )	$\Delta \eta_1 = 0.0587^{+0.2059}_{-0.2118}$ ( $\eta_1 = 4.00$ )	
LMDMalign <sub>HS40L50BD</sub>	Mass (NFW + $\Psi_1$ )	$\Delta \kappa_d = -0.0271^{+0.0294}_{-0.0297}$ ( $\kappa_d = 0.13$ )	$\Delta q_d = -0.0523^{+0.0821}_{-0.1042}$ ( $q_d = 0.82$ )	$\Delta \Psi_1 = 5.3773^{+6.5113}_{-6.6825}$ ( $\Psi_1 = 42.00$ )	$\Delta \Psi_1 = 5.3773^{+3.0125}_{-2.3066}$ ( $\Psi_1 = 42.00$ )	
LMDMalign <sub>ES40L50BD</sub>	Light (Sersic)	$\Delta R_1 = 0.1544^{+0.202}_{-0.2259}$ ( $R_1 = 0.60$ )	$\Delta q_1 = -0.0167^{+0.0049}_{-0.0370}$ ( $q_1 = 0.72$ )	$\Delta \eta_1 = 0.3100^{+0.7137}_{-0.7622}$ ( $\eta_1 = 4.00$ )	$\Delta \eta_1 = 0.3100^{+0.2798}_{-0.2705}$ ( $\eta_1 = 4.00$ )	
LMDMalign <sub>ES40L50BD</sub>	Mass (NFW + $\Psi_1$ )	$\Delta \kappa_d = -0.0368^{+0.0568}_{-0.0512}$ ( $\kappa_d = 0.13$ )	$\Delta q = -0.0843^{+0.1523}_{-0.2575}$ ( $q = 0.82$ )	$\Delta \Psi_1 = 1.3238^{+2.0777}_{-2.2659}$ ( $\Psi_1 = 6.74$ )	$\Delta \Psi_1 = 1.3238^{+0.8504}_{-0.7660}$ ( $\Psi_1 = 6.74$ )	
LMDMRot <sub>HS40L50BD</sub>	Light (Sersic)	$\Delta R_1 = 0.0833^{+0.1777}_{-0.1639}$ ( $R_1 = 0.60$ )	$\Delta q_1 = -0.0090^{+0.0418}_{-0.0396}$ ( $q_1 = 0.72$ )	$\Delta \eta_1 = 0.1747^{+0.5522}_{-0.6237}$ ( $\eta_1 = 4.00$ )	$\Delta \eta_1 = 0.1747^{+0.2314}_{-0.1820}$ ( $\eta_1 = 4.00$ )	
LMDMRot <sub>HS40L50BD</sub>	Mass (NFW + $\Psi_1$ )	$\Delta \kappa_d = 0.0014^{+0.0334}_{-0.0408}$ ( $\kappa_d = 0.13$ )	$\Delta q_{d1} = 0.0151^{+0.0531}_{-0.0784}$ ( $q_{d1} = 0.82$ )	$\Delta \Psi_{12} = -1.6112^{+9.5929}_{-8.1477}$ ( $\Psi_{12} = 42.00$ )	$\Delta \Psi_{12} = -1.6112^{+2.6403}_{-2.8220}$ ( $\Psi_{12} = 42.00$ )	
LMDMRot <sub>ES40L50BD</sub>	Light (Sersic)	$\Delta R_1 = 0.0371^{+0.1695}_{-0.1512}$ ( $R_1 = 0.60$ )	$\Delta q_1 = -0.0135^{+0.0307}_{-0.0309}$ ( $q_1 = 0.72$ )	$\Delta \eta_1 = 0.1219^{+0.5845}_{-0.5974}$ ( $\eta_1 = 4.00$ )	$\Delta \eta_1 = 0.1219^{+0.2239}_{-0.2083}$ ( $\eta_1 = 4.00$ )	
LMDMRot <sub>ES40L50BD</sub>	Mass (NFW + $\Psi_1$ )	$\Delta \kappa_d = -0.0257^{+0.0410}_{-0.0433}$ ( $\kappa_d = 0.13$ )	$\Delta q = -0.0572^{+0.0935}_{-0.1600}$ ( $q = 0.82$ )	$\Delta \Psi_1 = 1.0416^{+1.7427}_{-1.6689}$ ( $\Psi_1 = 6.74$ )	$\Delta \Psi_1 = 1.0416^{+0.5385}_{-0.5935}$ ( $\Psi_1 = 6.74$ )	
LMDMPos <sub>HS40L50BD</sub>	Light (Sersic)	$\Delta R_1 = 0.0142^{+0.1021}_{-0.1114}$ ( $R_1 = 0.60$ )	$\Delta q_1 = 0.0013^{+0.0302}_{-0.0532}$ ( $q_1 = 0.72$ )	$\Delta \eta_1 = 0.0451^{+0.4004}_{-0.4506}$ ( $\eta_1 = 4.00$ )	$\Delta \eta_1 = 0.0451^{+0.2154}_{-0.1641}$ ( $\eta_1 = 4.00$ )	
LMDMPos <sub>HS40L50BD</sub>	Mass (NFW + $\Psi_1$ )	$\Delta \kappa_d = -0.0318^{+0.0251}_{-0.0251}$ ( $\kappa_d = 0.13$ )	$\Delta q = -0.0859^{+0.0827}_{-0.1089}$ ( $q = 0.82$ )	$\Delta \Psi_1 = 8.6911^{+6.7850}_{-6.8201}$ ( $\Psi_1 = 42.00$ )	$\Delta \Psi_1 = 8.6911^{+2.5766}_{-2.4224}$ ( $\Psi_1 = 42.00$ )	
LMDMPos <sub>HS40L50BD</sub>	Geometry	$\Delta x_1 = 0.0005^{+0.0023}_{-0.0022}$ ( $x_1 = 0.00$ )		$\Delta x = 0.0105^{+0.0292}_{-0.0247}$ ( $x = 0.05$ )	$\Delta x = 0.0105^{+0.0088}_{-0.0104}$ ( $x = 0.05$ )	
LMDMPos <sub>HS40L50BD</sub>	Geometry	$\Delta y_1 = 0.0004^{+0.0023}_{-0.0021}$ ( $y_1 = 0.00$ )		$\Delta y = -0.0086^{+0.0271}_{-0.0322}$ ( $y = 0.00$ )	$\Delta y = -0.0086^{+0.0099}_{-0.0083}$ ( $y = 0.00$ )	
LMDMPos <sub>ES40L50BD</sub>	Light (Sersic)	$\Delta R_1 = 0.1853^{+0.2576}_{-0.2174}$ ( $R_1 = 0.60$ )	$\Delta q_1 = 0.0097^{+0.0410}_{-0.0386}$ ( $q_1 = 0.72$ )	$\Delta \eta_1 = 0.6120^{+0.7859}_{-0.7706}$ ( $\eta_1 = 4.00$ )	$\Delta \eta_1 = 0.6120^{+0.2903}_{-0.3108}$ ( $\eta_1 = 4.00$ )	
LMDMPos <sub>ES40L50BD</sub>	Mass (NFW + $\Psi_1$ )	$\Delta \kappa_d = -0.0041^{+0.0479}_{-0.0555}$ ( $\kappa_d = 0.13$ )	$\Delta q = 0.0010^{+0.0888}_{-0.1307}$ ( $q = 0.82$ )	$\Delta \Psi_1 = 0.0345^{+2.1667}_{-1.9210}$ ( $\Psi_1 = 6.74$ )	$\Delta \Psi_1 = 0.0345^{+0.7048}_{-0.7664}$ ( $\Psi_1 = 6.74$ )	
LMDMRot <sub>90<sub>HS30L75Multi</sub></sub>	Light (Sersic)	$\Delta R_1 = 0.0346^{+0.1457}_{-0.1774}$ ( $R_1 = 0.60$ )	$\Delta q_1 = 0.0037^{+0.0242}_{-0.0286}$ ( $q_1 = 0.72$ )	$\Delta \eta_1 = 0.1452^{+0.3286}_{-0.3463}$ ( $\eta_1 = 4.00$ )	$\Delta \eta_1 = 0.1452^{+0.1807}_{-0.1424}$ ( $\eta_1 = 4.00$ )	
LMDMRot <sub>90<sub>HS30L75Multi</sub></sub>	Mass (NFW + $\Psi_1$ )	$\Delta \kappa_d = 0.0456^{+0.0543}_{-0.0589}$ ( $\kappa_d = 0.13$ )	$\Delta q_{d1} = 0.0873^{+0.0678}_{-0.0937}$ ( $q_{d1} = 0.82$ )	$\Delta \Psi_{12} = -6.743^{+8.2456}_{-9.456}$ ( $\Psi_{12} = 42.0$ )	$\Delta \Psi_{12} = -6.743^{+4.1238}_{-4.4573}$ ( $\Psi_{12} = 42.0$ )	
LMDMRot <sub>90<sub>HS30L75Multi</sub></sub>	Mass (Shear)	$\Delta \gamma_{\text{sh}} = 0.0022^{+0.0020}_{-0.0021}$		$\Delta \theta_{\text{sh}} = 111.3445^{+27.5354}_{-23.4699}$	$\Delta \theta_{\text{sh}} = 111.3445^{+7.3562}_{-7.2346}$	
LMDMRot <sub>90<sub>HS30L75Multi</sub></sub>	Geometry	$\Delta \theta_1 = 0.2368^{+0.6853}_{-0.3861}$ ( $\theta_1 = 127.0$ )		$\Delta \theta_4 = 1.253^{+1.3373}_{-3.3842}$ ( $\theta_4 = 17.0$ )	$\Delta \theta_4 = 1.253^{+1.4686}_{-1.4921}$ ( $\theta_4 = 17.0$ )	
LMDMRot <sub>90<sub>ES30L75Multi</sub></sub>	Light (Sersic)	$\Delta R_1 = 0.0665^{+0.0920}_{-0.1174}$ ( $R_1 = 0.60$ )	$\Delta q_1 = 0.0043^{+0.0340}_{-0.0380}$ ( $q_1 = 0.72$ )	$\Delta \eta_1 = 0.1952^{+0.4069}_{-0.4900}$ ( $\eta_1 = 4.00$ )	$\Delta \eta_1 = 0.1952^{+0.1807}_{-0.1424}$ ( $\eta_1 = 4.00$ )	
LMDMRot <sub>90<sub>ES30L75Multi</sub></sub>	Mass (NFW + $\Psi_1$ )	$\Delta \kappa_d = 0.0508^{+0.0289}_{-0.0342}$ ( $\kappa_d = 0.13$ )	$\Delta q_{d1} = 0.1691^{+0.0152}_{-0.0539}$ ( $q_{d1} = 0.82$ )	$\Delta \Psi_{12} = -2.4666^{+1.3591}_{-1.1391}$ ( $\Psi_{12} = 6.74$ )	$\Delta \Psi_{12} = -2.4666^{+0.3988}_{-0.3666}$ ( $\Psi_{12} = 6.74$ )	
LMDMRot <sub>90<sub>ES30L75Multi</sub></sub>	Mass (Shear)	$\Delta \gamma_{\text{sh}} = 0.0239^{+0.0070}_{-0.0049}$		$\Delta \theta_{\text{sh}} = 108.0721^{+5.6003}_{-4.6170}$	$\Delta \theta_{\text{sh}} = 108.0721^{+1.5561}_{-1.9700}$	
LMDMShear <sub>HS35L80Cusp</sub>	Light (Sersic)	$\Delta R_1 = 0.0572^{+0.0565}_{-0.0518}$ ( $R_1 = 0.75$ )	$\Delta q_1 = -0.0020^{+0.0118}_{-0.0105}$ ( $q_1 = 0.75$ )	$\Delta \eta_1 = 0.1083^{+0.1030}_{-0.0996}$ ( $\eta_1 = 2.50$ )	$\Delta \eta_1 = 0.1083^{+0.0366}_{-0.0376}$ ( $\eta_1 = 2.50$ )	
LMDMShear <sub>HS35L80Cusp</sub>	Mass (NFW + $\Psi_1$ )	$\Delta \kappa_d = -0.0989^{+0.0157}_{-0.0217}$ ( $\kappa_d = 0.13$ )	$\Delta q = -0.0733^{+0.2089}_{-0.2052}$ ( $q = 0.80$ )	$\Delta \Psi_1 = 5.8475^{+0.9455}_{-2.747}$ ( $\Psi_1 = 11.52$ )	$\Delta \Psi_1 = 5.8475^{+0.4979}_{-0.3113}$ ( $\Psi_1 = 11.52$ )	
LMDMShear <sub>HS35L80Cusp</sub>	Geometry	$\Delta x_1 = 0.0001^{+0.0014}_{-0.0014}$ ( $x_1 = 0.00$ )		$\Delta x = 0.1648^{+0.1530}_{-0.1101}$ ( $x = 0.03$ )	$\Delta x = 0.1648^{+0.0399}_{-0.0522}$ ( $x = 0.03$ )	
LMDMShear <sub>HS35L80Cusp</sub>	Geometry	$\Delta y_1 = 0.0010^{+0.0019}_{-0.0019}$ ( $y_1 = 0.00$ )		$\Delta y = 0.1721^{+0.1918}_{-0.1400}$ ( $y = 0.03$ )	$\Delta y = 0.1721^{+0.0485}_{-0.0630}$ ( $y = 0.03$ )	
LMDMShear <sub>ES35L80Cusp</sub>	Light (Sersic)	$\Delta R_1 = 0.0454^{+0.0673}_{-0.0678}$ ( $R_1 = 0.75$ )	$\Delta q_1 = -0.0064^{+0.0186}_{-0.0184}$ ( $q_1 = 0.75$ )	$\Delta \eta_1 = 0.0839^{+0.1495}_{-0.1597}$ ( $\eta_1 = 2.50$ )	$\Delta \eta_1 = 0.0839^{+0.0564}_{-0.0533}$ ( $\eta_1 = 2.50$ )	
LMDMShear <sub>ES35L80Cusp</sub>	Mass (NFW + $\Psi_1$ )	$\Delta \kappa_d = -0.0687^{+0.0604}_{-0.0630}$ ( $\kappa_d = 0.13$ )	$\Delta q = -0.1507^{+0.2686}_{-0.3081}$ ( $q = 0.80$ )	$\Delta \Psi_1 = 0.7474^{+0.6308}_{-0.6199}$ ( $\Psi_1 = 2.00$ )	$\Delta \Psi_1 = 0.7474^{+0.2185}_{-0.2259}$ ( $\Psi_1 = 2.00$ )	
LMDMShear <sub>ES35L80Cusp</sub>	Mass (Shear)	$\Delta \gamma_{\text{sh}} = 0.0185^{+0.0147}_{-0.0121}$ ( $\gamma_{\text{sh}} = 0.03$ )		$\Delta \theta_{\text{sh}} = 0.5427^{+13.9894}_{-15.3828}$ ( $\theta_{\text{sh}} = 150.00$ )	$\Delta \theta_{\text{sh}} = 0.5427^{+5.5488}_{-5.1392}$ ( $\theta_{\text{sh}} = 150.00$ )	



**Figure 20.** The model image, residuals,  $\chi^2$  image, and source reconstructions for analysis of the images **LMDMPos<sub>HS40L50BD</sub>**, **LMDMRot<sub>ES40L50Multi</sub>**, and **LMDMShear<sub>HS35L80Cusp</sub>**. Images correspond to the most probable model at the end of the main pipeline, corresponding to the models given by Table 11.

little impact, if any, on the accuracy of the mass model inferred alongside only a minor impact on its precision. This validates approaches in the literature that infer a mass model from an image whose lens is subtracted before lens modeling (e.g. via a B-Splines interpolation). However, the reverse is not true and this approach cannot yield an accurate model for the lens’s light, as there is no analogous way to accurately subtract the source before fitting the lens. Therefore, a selling point of unified modeling is its ability to accurately characterize the lens’s visual morphology, which underpins the method’s ability to perform decomposed mass modeling.

Features such as radial arcs or a central image were successfully detected, demonstrating that the method is able to diagnose a cored mass profile. Therefore, `AutoLens` will also begin searching for cored mass profiles in existing lens data sets. In the future, this effort will benefit from data-sets which resolve the source and lens at an increased spatial resolution as well as multiwavelength imaging that contrast the relative brightnesses of the lens and source relative. It remains to be seen whether dust in the lens galaxy will prevent any such effort.

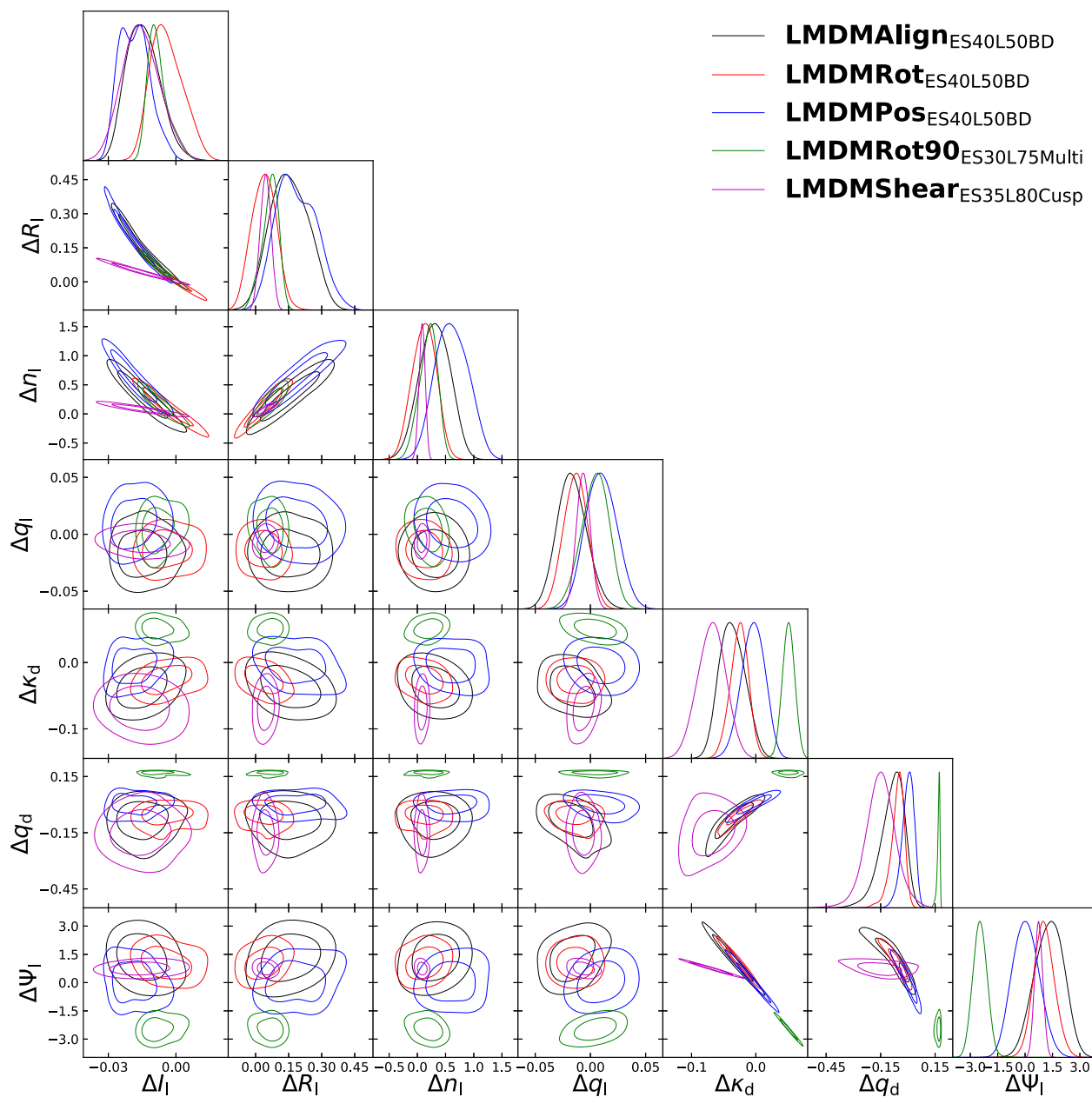
Although the mass model was not degenerate with the light subtraction, there are cases where the improved light subtraction offered by unified modeling could improve or even change the mass model altogether, when it reveals faint features in the lensed source that other approaches wrongly subtracted. This happened recently when a number of teams reassessed the light subtraction of *HST* imaging of the lens ID81 (Dye et al. 2014, 2015; Rybak et al. 2015; Swinbank et al. 2015), finding additional structure in the lensed source that changed the mass model and led to a more complete source

reconstruction. Following the same argument, the non-detection of source light is an equally powerful means by which to constrain the mass model and it is not uncommon for the image reconstruction to place extraneous flux where it isn’t observed, information that a poor lens subtraction and restrictive source-only masking will struggle to exploit. An extreme of this arises when a mass model goes to particularly shallow density profiles and begins to predict a central image in the region other approaches will have most likely masked before the analysis began.

Finally, and most importantly, unified modeling enables use of decomposed mass models that allow a number of unique measurements to be made for both a lens’s stellar and dark components. These models offer a significant improvement to lens modeling in general, because fitting of the lens’s light profile is exploited to place constraints on its underlying stellar mass distribution. For example, in the majority of SLACS lenses, the stellar component is dominant, making up over 90 per cent of the total mass within  $R_{\text{Ein}}$  (Barnabè et al. 2011). Therefore, lenses with a fainter, less extended or doubly imaged source, which offer loose constraints on an *SPL* mass model, will no doubt benefit from the additional information extracting by a decomposed mass model. Care must, of course, be taken to understand the impact of assuming a constant mass-to-light ratio.

### 8.3 Comparison to other methods

Other methods in the literature (Dye & Warren 2005; Vegetti & Koopmans 2009a; Suyu 2012; Collett & Auger 2014; Tagore & Kee-



**Figure 21.** Marginalized 1-D and 2-D PDFs of the parameters governing the lens's mass distribution for decomposed mass modeling,  $\Delta I_l$ ,  $\Delta R_l$ ,  $\Delta n_l$ ,  $\Delta q_l$ ,  $\Delta \kappa_d$ ,  $\Delta q_d$ , and  $\Delta \Psi_l$ , for the low-resolution image's of the **LMDM** models. The legend at the top indicates the image that each coloured line corresponds to. Contours give the  $1\sigma$  (interior) and  $3\sigma$  (exterior) confidence regions. The input values for each model can be found in Table 1. A degeneracy can be seen between all parameters, an extension of the degeneracy between mass, ellipticity, and slope described in N15, but now including components for both the light and dark matter profiles. However, fitting of the lens's light reduces this degeneracy, thereby permitting accurate sampling.

ton 2014; Birrer, Amara & Refregier 2015b) use approaches similar to AutoLens's image and source analysis. The key differences are: (i) the amorphous nature of AutoLens's source pixelization, which ensures the method can truly achieve the simplest solution using the fewest correlated source pixels; (ii) the freedom given to AutoLens's variances and source regularization, which are key to correctly fitting the lens's light profile, and (iii) the method's removal of the discreteness biases described in N15, which (although not shown explicitly shown in this work) continued to have a significant impact on both mass and lens light modeling if not handled using the approach advocated in N15. Furthermore, Au-

toLens changes its image and source analysis in an automated and fully self-consistent manner, in contrast to other methods that follow a more ad hoc approach. This makes the results of AutoLens reproducible.

## 8.4 Summary

This paper presents AutoLens, the first automated modelling suite for strong gravitational lenses. Our key results are as follows:

- (i) The image and source analysis adapt automatically to the properties of the strong lens being analysed. This includes an



amorphous source-plane pixelization that adapts to the source’s unlensed surface-brightness profile, a source regularization scheme that adapts to the source’s intrinsic morphology and scaling of the observed image’s variance-map to ensure the method provides a global fit to the strong lens imaging. These are all performed self-consistently within the Bayesian framework of MacKay (1992); Suyu et al. (2006).

(ii) Fitting of the lens’s light profile is fully integrated into `AutoLens` and performed simultaneously with the mass and source. Light profiles comprise both single- and multicomponent models, allowing the method to perform a bulge-disc decomposition and model lens galaxies of arbitrary morphology. The adaptive image and source features above are key to ensuring that the lens’s light profile is inferred accurately, something other methods are not well suited to.

(iii) The improved lens subtraction allows `AutoLens` to model and detect faint features in a lensed source that previous methods may omit. This includes, but is not limited to, features indicative of a cored density profile, such as radial arcs or the source’s third or fifth central image.

(iv) Lens light fitting allows the method to invoke decomposed mass models that model separately the lens’s light and dark matter. These offer a significant improvement to the inferred mass model because they exploit how the lens galaxy’s light traces its underlying stellar mass distribution. Thus, by incorporating its light profile into the mass model new information is exploited about the lens’s mass distribution; information that previous approaches to lens modeling omit completely.

(v) The complexity of the light and mass models, including the detection of radial arcs, a central image or a geometrically offset light/dark matter profile, is decided objectively via Bayesian model comparison. This is performed by estimating the Bayesian evidence of each unique lens model, by running a non-linear search using the nested sampling algorithm `MultiNest` (Skilling 2006; Feroz et al. 2009) for each.

(vi) A single lens therefore provides a diverse range of observations. Mass models that assume a total (light and dark summed) profile infer its inner density slope, can detect a central core, and fully characterize the lens galaxy’s light profile. Decomposed mass models offer a stellar mass-to-light ratio, dark matter halo ellipticity and dark matter fraction as a function of radius, as well as determining if the light and dark matter are geometrically aligned with or offset from one another. The highly magnified source galaxy is also fully reconstructed.

(vii) `AutoLens` is fully automated, such that all results presented in this work are generated without any user intervention. This successfully automates the lens modeling process.

(viii) `AutoLens` is demonstrated on a suite of 54 simulated images which span a variety of lens and source morphologies, mass profiles, lensing geometries and imaging data resolutions, and signal-to-noise ratios. The method performs well for all of the observables listed above, choosing the correct model complexity in the majority of cases and inferring most parameters accurately within  $3\sigma$  confidence.

## 8.5 Concluding remarks

Over the past two decades an outstanding and diverse multiwavelength data set of high-quality strong lens images has been amassed throughout the literature. However, the complex and iterative nature of lens modeling has historically restricted their analysis to small samples, simplified mass models, and little to no considera-

tion of how the lens galaxy’s light profile can benefit their analysis. `AutoLens` addresses these issues, enabling the application of decomposed mass profiles that fully exploit the information contained within the lens’s light on large lens samples. Future work will see `AutoLens` applied to these data sets, with the lens galaxy morphologies, density profiles, and dark matter geometries key topics of interest. This work will lay the foundations for study of the anticipated large lens data sets comprising tens of thousands of strong lenses to ultimately give an unprecedented understanding of the nature of galaxy formation, dark matter, and the Universe in general.

## ACKNOWLEDGEMENTS

JN acknowledges support from STFC grant ST/N00149/1 and the University of Nottingham. SD acknowledges support from the Midlands Physics Alliance and STFC Ernest Rutherford Fellowship Scheme. RJM acknowledges support from the Royal Society University Research Fellowship. We are grateful for access to the University of Nottingham High Performance Computing Facility. This paper used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility ([www.dirac.ac.uk](http://www.dirac.ac.uk)). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, and STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National e-Infrastructure. We acknowledge use of the software program ([getdist.readthedocs.io/en/latest/](http://getdist.readthedocs.io/en/latest/)).

## REFERENCES

- Auger M. W., Treu T., Bolton A. S., Gavazzi R., Koopmans L. V., Marshall P. J., Moustakas L. A., Burles S., 2010, *ApJ*, 724, 511
- Balmès I., Corasaniti P. S., 2013, *MNRAS*, 431, 1528
- Barkana R., 1998, *ApJ*, 502, 531
- Barnabe M., Koopmans L. V. E., 2007, *ApJ*, 666, 726
- Barnabè M., Czoske O., Koopmans L. V., Treu T., Bolton A. S., Gavazzi R., 2009, *MNRAS*, 399, 21
- Barnabè M., Czoske O., Koopmans L. V., Treu T., Bolton A. S., 2011, *MNRAS*, 415, 2215
- Bett P., Eke V., Frenk C. S., Jenkins A., Okamoto T., 2010, *MNRAS*, 404, 1137
- Biernaux J., Magain P., Sluse D., Chantry V., 2016, *A&A*, 585, A84
- Birrer S., Amara A., Refregier A., 2015a, *ApJ*, 813, 102
- Birrer S., Amara A., Refregier A., 2015b, *ApJ*, 813, 102
- Birrer S., Amara A., Refregier A., 2016, *J. Cosmol. Astropart. Phys.*, 2016, 020
- Bluck A. F., Trevor Mendel J., Ellison S. L., Moreno J., Simard L., Patton D. R., Starkenburg E., 2014, *MNRAS*, 441, 599
- Bolton A. S. et al., 2012, *ApJ*, 757, 82
- Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Moustakas L. A., 2006, *ApJ*, 638, 703
- Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Gavazzi R., Moustakas L. A., Wayth R., Schlegel D. J., 2008, *ApJ*, 682, 964
- Brewer B. J. et al., 2012, *MNRAS*, 422, 3574
- Brewer B. J., Lewis G. F., 2008, *MNRAS*, 390, 39
- Brownstein J. R. et al., 2012, *ApJ*, 744, 41
- Bruce V. A. et al., 2014a, *MNRAS*, 444, 1001
- Bruce V. A. et al., 2014b, *MNRAS*, 444, 1660
- Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, *MNRAS*, 321, 559
- Collett T. E., 2015, *ApJ*, 811, 20
- Collett T. E., Auger M. W., 2014, *MNRAS*, 443, 969
- Dullo B. T., Graham A. W., 2013, *ApJ*, 768, 36
- Dullo B. T., Graham A. W., 2014, *MNRAS*, 444, 2700
- Dutton A. A. et al., 2013, *MNRAS*, 428, 3183

- Dye S. et al., 2014, *MNRAS*, 440, 2013  
Dye S. et al., 2015, *MNRAS*, 452, 2258  
Dye S., Warren S. J., 2005, *ApJ*, 623, 31  
Dye S., Evans N. W., Belokurov V., Warren S. J., Hewett P., 2008, *MNRAS*, 388, 384  
Dye H. S. et al., 2017, *MNRAS*, 000, 1  
Eichner T., Seitz S., Bauer A., 2012, *MNRAS*, 427, 1918  
Enia A. et al., 2018, *MNRAS*, 475, 3467  
Faber S. M. et al., 1997, *AJ*, 114, 1771  
Falco E. E., Gorenstein M. V., Shapiro I. I., 1985, *ApJ*, 289, L1  
Feroz F., Hobson M. P., 2008, *MNRAS*, 384, 449  
Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601  
Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, preprint (arXiv: 1306.2144)  
Gavazzi R., Treu T., Rhodes J. D., Koopmans L. V. E., Bolton A. S., Burles S., Massey R. J., Moustakas L. A., 2007, *ApJ*, 667, 176  
Gavazzi R., Treu T., Koopmans L. V. E., Bolton A. S., Moustakas L. A., Burles S., Marshall P. J., 2008, *ApJ*, 677, 1046  
Golse G., Kneib J.-P., 2002, *A&A*, 390, 821  
Hartigan J. A., Wong M. A., 1979, *Appl. Stat.*, 28, 100  
Häußler B. et al., 2013, *MNRAS*, 430, 330  
Hezaveh Y. D., Marshall P. J., Blandford R. D., 2015, *ApJ Lett.*, 799, 1  
Hoyos C. et al., 2011, *MNRAS*, 411, 2439  
Jaroszynski M., Kostrzewa-Rutkowska Z., 2012, *MNRAS*, 424, 325  
Keeton C. R., 2001, *J. Astron. Astrophys.*, 17, preprint (arXiv: 0102.0341)  
Keeton C. R., 2003, *ApJ*, 582, 17  
Keeton C. R., Kochanek C. S., Seljak U., 1997, *ApJ*, 482, 604  
Koopmans L. V. et al., 2009, *ApJ*, 703, L51  
Lagattuta D. J., Vegetti S., Fassnacht C. D., Auger M. W., Koopmans L. V., Mckean J. P., 2012, *MNRAS*, 424, 2800  
Liao S., Gao L., Frenk C. S., Guo Q., Wang J., 2017, *MNRAS*, 470, 2262  
MacKay D. J. C., 1992, thesis, CIT, 4, 98  
Mao S., Witt H. J., Koopmans L. V. E., 2001, *MNRAS*, 323, 301  
Marshall P. J. et al., 2007, *ApJ*, 671, 1196  
Navarro J. F., Benz W., 1991, *ApJ*, 380, 320  
Navarro J. F., Frenk C. S., White S. D. M., 1996a, *ApJ*, 490, 493  
Navarro J. F., Frenk C. S., White S. D. M., 1996b, *ApJ*, 462, 563  
Newton E. R., Marshall P. J., Treu T., Auger M. W., Gavazzi R., Bolton A. S., Koopmans L. V., Moustakas L. A., 2011, *ApJ*, 734, 104  
Nightingale J. W., Dye S., 2015, *MNRAS*, 452, 2940  
Nikutta R., Nenkova M., Ivezi?, Hunt-Walker N., Elitzur M., 2014, Rev. Mex. Astron. Astrofisica Ser. Conf., 44, 100  
Oguri M. et al., 2012, *AJ*, 143, 120  
Oldham L. et al., 2017, *MNRAS*, 465, 3185  
Piontek F., Steinmetz M., 2011, *MNRAS*, 410, 2625  
Quinn J. et al., 2016, *MNRAS*, 459, 2394  
Rusin D., Ma C.-P., 2001, *ApJ*, 549, L33  
Rybak M., Vegetti S., McKean J. P., Andreani P., White S. D., 2015, *MNRAS*, 453, L26  
Sales L. V., Navarro J. F., Theuns T., Schaye J., White S. D., Frenk C. S., Crain R. A., Dalla Vecchia C., 2012, *MNRAS*, 423, 1544  
Schaller M., Robertson A., Massey R., Bower R. G., Eke V. R., 2015, *MNRAS*, 453, L58  
Schaye J. et al., 2015, *MNRAS*, 446, 521  
Schneider P., 2014a, *A&A*, 568, L2  
Schneider P., 2014b, preprint (arXiv:1409.0015)  
Schneider P., Sluse D., 2013a, *A&A*, 559, A37  
Schneider P., Sluse D., 2013b, *A&A*, 564, A103  
Schneider P., Weiss A., 1992, *A&A*, 260, 1  
Shirazi M., Vegetti S., Nesvadba N., Allam S., Brinchmann J., Tucker D., 2014, *MNRAS*, 440, 2201  
Shu Y. et al., 2016, *ApJ*, 824, 86  
Skilling J., 2006, *Bay. Anal.*, 1, 833  
Sonnenfeld A., Treu T., Gavazzi R., Marshall P. J., Auger M. W., Suyu S. H., Koopmans L. V., Bolton A. S., 2012, *ApJ*, 752, 163  
Sonnenfeld A., Gavazzi R., Suyu S. H., Treu T., Marshall P. J., 2013a, *ApJ*, 777, 97  
Sonnenfeld A., Treu T., Gavazzi R., Suyu S. H., Marshall P. J., Auger M. W., Nipoti C., 2013b, *ApJ*, 777, 98  
Sonnenfeld A., Treu T., Marshall P. J., Suyu S. H., Gavazzi R., Auger M. W., Nipoti C., 2015, *ApJ*, 800, 94  
Suyu S. H. et al., 2013, *ApJ*, 766, 70  
Suyu S. H. et al., 2016, *MNRAS*, 468, 2590  
Suyu S. H., 2012, *MNRAS*, 426, 868  
Suyu S. H., Marshall P. J., Hobson M. P., Blandford R. D., 2006, *MNRAS*, 371, 983  
Swinbank A. M. et al., 2015, *ApJ*, 806, L17  
Tagore A. S., Keeton C. R., 2014, *MNRAS*, 445, 694  
Tagore A. S., Barnes D. J., Jackson N., Kay S. T., Schaller M., Schaye J., Theuns T., 2018, *MNRAS*, 474, 3403  
Tessore N., Bellagamba F., Metcalf R. B., 2016, *MNRAS*, 463, 3115  
Vegetti S., Koopmans L. V., 2009a, *MNRAS*, 392, 945  
Vegetti S., Koopmans L. V., 2009b, *MNRAS*, 400, 1583  
Vegetti S., Lagattuta D. J., McKean J. P., Auger M. W., Fassnacht C. D., Koopmans L. V., 2012, *Nature*, 481, 341  
Vegetti S., Koopmans L. V., Auger M. W., Treu T., Bolton A. S., 2014, *MNRAS*, 442, 2017  
Velliscig M. et al., 2015, *MNRAS*, 453, 721  
Vika M., Bamford S. P., Häußler B., Rojas A. L., Borch A., Nichol R. C., 2013, *MNRAS*, 435, 623  
Vika M., Bamford S. P., Häußler B., Rojas A. L., 2014, *MNRAS*, 444, 3603  
Vulcani B. et al., 2014, *MNRAS*, 441, 1340  
Warren S., Dye S., 2003, *ApJ*, 590, 673  
Winn J. H., Rusin D., Kochanek C. S., 2004, *Nature*, 427, 613  
Wong K. C. et al., 2017, *MNRAS*, 465, 4895  
Xu D., Sluse D., Schneider P., Springel V., Vogelsberger M., Nelson D., Hernquist L., 2016, *MNRAS*, 456, 739  
Zhao H., 1996, *MNRAS*, 278, 488  
Ziegel E., Press W., Flannery B., Teukolsky S., Vetterling W., 1987, *Technometrics*, 29, 501

## APPENDIX A: EFFICIENCY TRICKS AND ALGORITHMS

### A1 Light profile

High levels of oversampling are required to accurately compute a 2-D Sersic light profile. This is especially true when evaluated at low radii, where it diverges. Therefore, adaptive oversampling is applied to ensure that the light profile is computed fast and accurately. This routine first acts on all pixels within  $R_1$ . The intensity of the pixel,  $I(\xi_1)$ , is computed at its center. This pixel is then divided into a  $2 \times 2$  sub-grid and the mean of intensities at the sub-pixel centres is computed. If the fractional change in  $I(\xi_1)$  is  $< 0.0001$ , then sufficient accuracy has been achieved and the value just computed is used. However, if the fractional change is  $> 0.0001$ , this process is iteratively repeated for higher levels of oversampling, up to a sub-grid of degree  $2000 \times 2000$ . Finally, if all pixels within  $R_1$  required oversampling, then pixels outside this radius are re-evaluated until the accuracy threshold is met at least once without oversampling. If the initial  $I(\xi_1)$  value is below the numerical precision of `AutoLens` ( $10^{-16}$ ), it is not oversampled, as rounding errors prevent convergence (and the flux is negligible anyway). These occurrences are rare and happen when  $\xi_1$  is large and  $R_1$  is small.

### A2 Deflection angles

Numerical integration is used to compute  $\alpha_{x,y}$  from the above  $\kappa(x)$  profiles. `AutoLens` uses an adaptive numerical integration technique, which iteratively refines the subintervals over which the integral is evaluated until a threshold accuracy is achieved. The expressions for  $\phi_r$  can be found in [K01](#), with equation (55) giving the NFW

profile, equation (45) the de Vaucouleurs profile, and equation (74) the exponential profile. For every profile the  $\kappa(\mathbf{x})$  and  $\alpha_{x,y}$  maps generated by `AutoLens` have been compared with those given by the lensing software `gravlens` (Keeton 2003), ensuring that all are implemented in `AutoLens` correctly.

**N15** showed that image oversampling (termed subgridding in **N15**) is required to remove aliasing effects that lead to inaccurate lens modeling. Oversampling splits each image pixel into a set of square sub-pixels, the centers of which are all traced to the source-plane and used by the inversion. However, high levels of oversampling requires  $\alpha_{x,y}$  to be computed for each additional sub-pixel, which can prove computationally expensive.

A bi-linear interpolation scheme is therefore now applied to greatly increase efficiency, whereby deflection angles are computed at the center of image pixels and interpolated to give the sub-pixel deflection angles. In the central regions of the mass profile (where the density profile is rapidly increasing), this interpolation scheme becomes inaccurate, thus in these central regions each sub-pixel deflection angle is computed explicitly. Further out (where the density profile is much flatter) deflection angles are interpolated from a grid of computed deflection angles, thus reducing the number of overall deflection angle calculations. This grid becomes more coarse as one reaches the flatter regions of the mass profile's density. This interpolation scheme calculates  $\alpha_{x,y}$  at sub-pixels to a fractional accuracy of  $10^{-4}$ , which is more than sufficient, given the systematics associated with source-plane discretization. Whilst a  $4 \times 4$  sub-pixel grid was found to be sufficient in **N15**, higher levels of oversampling ( $8 \times 8$ ) are used in this work, given that it is now computationally feasible.

### A3 Positional information

The initial calculation of lens models that accurately fit the image data involves searching large portions of non-linear parameter space, which is prohibitively expensive. Positional information is therefore used to increase speed, by requiring that any lens model must first meet the requirement that four image pixels in the lensed source must trace to within a threshold value of one another in the source-plane. If this criterion fails, a new lens model is sampled. This approach was introduced by Brewer & Lewis (2008) for a strongly lensed quasar, the point source nature of which allowed them to impose that image pixels trace to within  $10^{-5}$  arcsec of one another. Here, a much larger threshold is used throughout, because unlike Brewer & Lewis (2008), this is not imposed to constrain the lens model but simply to improve the speed of the initial non-linear sampling.

However, due to the complex source morphologies of real strong lens imaging, one can never be sure which image-plane pixels actually neighbour one another in the source plane. Therefore, positional information is only exploited once an accurate model for the source has been computed, which in the automated analysis pipeline is after the phase  $\mathbf{P}_{\text{Init}2}$  where a parametric source is fitted. The positional image pixels are calculated as the four image pixels that trace closest to the parametric source's center  $(x_s, y_s)$  and also (i) are separated by over 20 per cent of the lens's  $R_{\text{Ein}}$  value in the image-plane and (ii) have 1 pixel rotationally offset from the others in the image-plane by at least  $120^\circ - 240^\circ$  degrees around the lens center (to ensure multiple images are sampled as opposed to just one image's extended arc). To compute positional image pixels for a pixelated source-plane, these requirements are followed using image pixels that trace to the brightest pixel in the source reconstruction.

If they are not met using only this source pixel, its closest traced image pixels are iteratively used until four image pixels are chosen.

The threshold value is then reduced to a value of  $3 \times$  the maximum source-plane separation of these newly allocated image pixels or 0.3 arcsec, whichever is smaller, thereby giving significant efficiency gains whilst ensuring no feasible lens models are wrongly discarded. Positional information is also key to removing the unwanted over-/underfit solutions described in **N15** and Section 6.

## APPENDIX B: CONSTANT REGULARIZATION

The linear regularization matrix  $\mathbf{H}$  used in Warren & Dye (2003) and **N15** is derived following the formalism given in Ziegel et al. (1987). This computes  $\mathbf{H}$  as  $\mathbf{H} = \mathbf{B}^T \mathbf{B}$ , where the matrix  $\mathbf{B}$  stores the regularization pattern of source pixels with one another. For example, to regularize each source pixel with its neighbour, assuming the numbering scheme is such that pixel 1 is a neighbour of pixel 2, and 2 of 3, etc., the matrix  $\mathbf{B}_x$  is given as

$$\begin{bmatrix} -1 & 1 & 0 & 0 & \dots \\ 0 & -1 & 1 & 0 & \dots \\ 0 & 0 & -1 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}. \quad (\text{B1})$$

For gradient regularization on an  $N \times N$  square grid, this matrix gives the regularization of source pixels across the  $x$ -direction, where every  $N$  elements will be a row of zeros. This matrix then gives a regularization matrix  $\mathbf{H}_x = \mathbf{B}_x^T \mathbf{B}_x$ . For regularization in the  $y$ -direction, a second  $\mathbf{B}_y$  matrix is generated, where the negative ones are again across the diagonal and the positive ones every  $N$  elements across from this, with the final  $N$  rows all zeros.  $\mathbf{B}_y$  is then used to compute a second regularization matrix  $\mathbf{H}_y = \mathbf{B}_y^T \mathbf{B}_y$ , which is added to the first to give the overall regularization matrix  $\mathbf{H} = \mathbf{H}_x + \mathbf{H}_y$ . For `AutoLens`'s Voronoi regularization scheme the same pattern is followed, using around 5–10  $\mathbf{H}$  matrices corresponding to regularization across all of the Voronoi vertex indices.

### B1 Non-constant regularization

Formally,  $\lambda$  can be included in the  $\mathbf{B}$  matrices above. However, because it is a fixed single value, it is convention to take it outside  $\mathbf{B}$ . For example, in **N15**, this saw  $\lambda$  included in three terms in the expression for the Bayesian evidence (e.g.  $\lambda s^T \mathbf{H} s$ ). For the non-constant regularization scheme used to weight regularization by the source's luminosity, a 1D vector of regularization coefficients  $\Lambda$  must be employed and incorporated into the computation of  $\mathbf{H}$ .

Therefore, the  $\mathbf{B}$  matrices above are redefined to include each pixel's effective regularization coefficient,  $\lambda_{\text{eff}}$ , as  $\mathbf{B}_\Lambda = \Lambda \mathbf{B}$ , where  $\Lambda$  is computed as described in Section 4.7.3. The corresponding regularization matrix is then  $\mathbf{H}_\Lambda = \mathbf{B}_\Lambda^T \mathbf{B}_\Lambda$ .

## APPENDIX C: PIPELINE PRIORS AND LINKING

This appendix presents a more detailed overview of pipeline phase linking, describing the priors given to the different light and mass profiles used to initialize that profile at different points in the pipeline.

For each phase, initial parameter sampling is performed using one of two priors:

- (i) **Uniform Prior (UP)** - -- Draws points randomly from a uniform distribution defined by a maximum and minimum value.

**Table C1.** Priors used to initialize the Sersic + Exponential fit used in  $\mathbf{P}_{\text{Init1}}$ 

Model	Component	Prior load	Parameters						
<i>Sersic</i>	Light	None	$x_{11}$ UP $_{-0.5}^{0.5}$	$y_{11}$ UP $_{-0.5}^{0.5}$	$I_{11}$ UP $_0^2$	$R_{11}$ UP $_0^4$	$n_{11}$ UP $_{0.7}^{9.5}$	$q_{11}$ UP $_{0.3}^{1.0}$	$\theta_{11}$ UP $_0^{180}$
<i>Exponential</i>	Light	None	$x_{12} = x_{11}$	$y_{12} = y_{11}$	$I_{12}$ UP $_0^2$	$R_{12}$ UP $_0^4$	$n_{12} = 1$	$q_{12}$ UP $_{0.3}^{1.0}$	$\theta_{12} = \theta_{11}$

**Table C2.** Priors used to initialize the *SPLE* + *Sersic* (Source) or *PL*<sub>Core</sub> + *Sersic* (Source) profile fit used in  $\mathbf{P}_{\text{Init2}}$ . The Sersic and *SPLE* and *PL*<sub>Core</sub> parameters are initialized from the results of  $\mathbf{P}_{\text{Init1}}$ .

Model	Component	Prior load	Parameters						
<i>SPLE</i>	Mass	$\mathbf{P}_{\text{Init1}}$	$x$ GP 0.1''	$y$ GP 0.1''	$\theta_E$ UP $_0^4$	$q$ UP $_{0.3}^{1.0}$	$\theta$ UP $_0^{180}$		
							$\alpha = 2.2$		
<i>Sersic</i>	Source	None	$x_s$ UP $_{-0.5}^{0.5}$	$y_s$ UP $_{-0.5}^{0.5}$	$I_s$ UP $_0^2$	$R_s$ UP $_0^4$	$n_s$ UP $_{0.7}^{9.5}$	$q_s$ UP $_{0.3}^{1.0}$	$\theta_s$ UP $_0^{180}$

(ii) **Gaussian Prior (GP)** – – Draws points randomly from a normal distribution defined by a mean  $\mu_g$  and width  $\sigma_g$ . The value for  $\mu_g$  is estimated from the previous phase, using the median of the same parameter or a related parameter's marginalized 1-D is estimated from the previous phase, using the median of the same parameter or a related parameter's marginalized 1-D

The value of  $\sigma_g$  is defined such that 68.2 per cent of points sampled (on average) are between that parameter's previously estimated  $3\sigma$  confidence bounds. These confidence bounds are generally not symmetric and the value furthest from the mean is used. This ensures that when linking phases the wider area of the previous phase's posterior is sampled. For example, if a parameter is estimated as  $n = 4_{-0.3}^{+0.2}$  at  $3\sigma$  confidence, 68.2 per cent of samples in the next phase will (on average) lie between 3.7 and 4.3. However, for high-quality imaging data, parameters can be estimated to a very high accuracy and their errors could therefore be very small. Using these errors to set  $\sigma_g$  therefore runs a risk of biasing an analysis by placing overly restrictive priors. To overcome this, each parameter has a minimum  $\sigma_g$  value, which replaces the previous analysis's error estimate if it is below this minimum. Extending the previous example, if this minimum value were 0.6, then 68.2 per cent of samples in the next phase will be between 3.4 and 4.6 despite its errors only corresponding to a size of 0.3. These minimum values are given for each phase in the tables described next.

For the first optimization of the hyperparameters uniform priors are assumed on all parameters, except  $\lambda$  which uses a broad Gaussian prior centered on phase  $\mathbf{P}_{\text{Init3}}$ 's initial estimate of  $\lambda$ . These uniform priors are broad, but may not be sufficient to capture the optimum value of all of the hyperparameters. However, re-optimization of the hyperparameters uses Gaussian priors centered on their previous phase's most probable values with  $\sigma_g = \mu_g/2$ , thus the optimum values will be reached after two or three hyperparameter optimizations.

The following tables give the priors used in each pipeline phase for the different light and mass models. It should be noted that some quantities that depend on the image properties, like  $I_1$  and  $\Psi_1$ , have priors that depend on initial estimates of their values from the image data. All tables follow the same notation, where  $\mathbf{UP}_b^a$  corresponds to a uniform prior between the values  $a$  and  $b$  and  $\mathbf{GP}_b^a$  corresponds to a Gaussian prior with minimum value  $\sigma_g$  given by  $a$ . Initialization of parameters in the main pipeline uses exclusively Gaussian priors, retaining the minimum  $\sigma_g$  values given for the model comparison phases.

**Table C3.** Priors used to initialize the *Sersic* light model and *SPLE* mass model used in  $\mathbf{P}_{\text{Init3}}$ . The *SPLE* and  $PL_{\text{Core}}$  parameters  $x$ ,  $y$ , and  $\theta$  are initialized from the results of  $\mathbf{P}_{\text{Init2}}$ .

Model	Component	Prior load	Parameters						
<i>Sersic</i>	Light	None	$x_1$ UP <sub>-0.5</sub> <sup>0.5</sup>	$y_1$ UP <sub>-0.5</sub> <sup>0.5</sup>	$I_{11}$ UP <sub>0</sub> <sup>2</sup>	$R_{11}$ UP <sub>0</sub> <sup>4</sup>	$m_{11}$ UP <sub>0.7</sub> <sup>9.5</sup>	$q_{11}$ UP <sub>0.3</sub> <sup>1.0</sup>	$\theta_{11}$ UP <sub>0</sub> <sup>180</sup>
<i>SPLE</i>	Mass	$\mathbf{P}_{\text{Init2}}$	$x$ (x <sub>1</sub> ) GP 0.02	$y$ (y <sub>1</sub> ) GP 0.02	$\theta_E$ GP $\theta_{\text{Ein}}/4$	$q$ GP 0.05	$\theta$ GP 20.0	$\alpha = 2.2$	
Hyper Params	Hyper	None	$\lambda$ UP <sub>0</sub> <sup>1000</sup>						

**Table C4.** Priors used to initialize the *Sersic* light profile and mass profile used in  $\mathbf{P}_{\text{Init4}}$ . The *Sersic*, *SPLE*,  $PL_{\text{Core}}$ , and *NFWsPh* parameters are initialized from the results of  $\mathbf{P}_{\text{Init3}}$ .

Model	Component	Prior load	Parameters						
<i>Sersic</i>	Light	$\mathbf{P}_{\text{Init3}}$	$x_1$ GP 0.02''	$y_1$ GP 0.02''	$I_1$ GP $I_1/2$	$R_1$ GP $R_1/2$	$m_1$ GP 0.8	$q_1$ GP 0.1	$\theta_1$ GP 20.0
<i>SPLE</i>	Mass	$\mathbf{P}_{\text{Init3}}$	$x$ GP 0.1''	$y$ GP 0.1''	$\theta_E$ GP $\theta_E/2$	$q$ GP 0.1	$\theta$ GP 20.0	$\alpha = 2.2$	
$PL_{\text{Core}}$	Mass	$\mathbf{P}_{\text{Init3}}$	$x$ GP 0.1''	$y$ GP 0.1''	$\theta_E$ GP $\theta_E/2$	$q$ GP 0.1	$\theta$ GP 20.0	$\alpha = 2.0$	$s$ GP 0.1''
<i>NFWsPh</i>	Mass	$\mathbf{P}_{\text{MCMass}}$	$x_d$ (x) GP 0.2''	$y$ (y) GP 0.2''	$\kappa_s$ UP <sub>0</sub> <sup>3</sup>	$q = 1.0$			
$\Psi_1$	Mass		$\Psi_1$ UP <sub>0.0</sub> <sup>5.0</sup>						

**Table C5.** Priors used to initialize the Mass + Light profile fits used for model comparison in phase  $\mathbf{P}_{\text{MCLight}}$ . Parameter initializations are derived from the results of  $\mathbf{P}_{\text{Init4}}$ . The hyperparameters  $\omega_{\text{Lens}}$  and  $\omega_{\text{Src}}$  are initialized using the  $\chi_{\text{Base}}^2$  values of the observed image.

Model	Component	Prior load	Parameters						
<i>Sersic</i>	Light	$\mathbf{P}_{\text{Init4}}$	$x_1$ GP 0.02''	$y_1$ GP 0.02''	$I_1$ UP <sub>0</sub> <sup>2</sup>	$R_1$ UP <sub>0</sub> <sup>4</sup>	$m_1$ UP <sub>0.3</sub> <sup>9.5</sup>	$q_1$ UP <sub>0.3</sub> <sup>1.0</sup>	$\theta_1$ GP 20.0
<i>Sersic + Exp</i>	Light	$\mathbf{P}_{\text{Init4}}$	$x_1$ GP 0.02''	$y_1$ GP 0.02''	$I_{11}$ UP <sub>0</sub> <sup>2</sup>	$R_{11}$ UP <sub>0</sub> <sup>4</sup>	$m_{11}$ UP <sub>0.7</sub> <sup>9.5</sup>	$q_{11}$ UP <sub>0.3</sub> <sup>1.0</sup>	$\theta_1$ GP 20.0
			$x_{12} = x_{11}$	$y_{12} = y_{11}$	$I_{12}$ UP <sub>0</sub> <sup>2</sup>	$R_{12}$ UP <sub>0</sub> <sup>4</sup>	$m_{12} = 1.0$	$q_{12}$ UP <sub>0.3</sub> <sup>1.0</sup>	$\theta_{12} = \theta_{11}$
<i>SPLE</i>	Mass	$\mathbf{P}_{\text{Init4}}$	$x$ GP 0.01''	$y$ GP 0.01''	$\theta_E$ GP 0.05''	$q$ GP 0.01	$\theta$ GP 2.0	$\alpha = 2.2$	
$PL_{\text{Core}}$	Mass	$\mathbf{P}_{\text{Init4}}$	$x$ GP 0.05''	$y$ GP 0.05''	$\theta_E$ GP $\theta_E/2$	$q$ GP 0.1	$\theta$ GP 10.0	$\alpha = 2.0$	$S$ UP <sub>0.0</sub> <sup>0.2''</sup>
<i>NFWsPh</i>	Mass	$\mathbf{P}_{\text{MCMass}}$	$x_d$ (x) GP 0.2''	$y$ (y) GP 0.2''	$\kappa_s$ GP $\kappa_s/2$	$q = 1.0$			
$\Psi_1$	Mass		$\Psi_1$ GP $\Psi_1/2$						
Hyper Params	Hyper	Prev	$\lambda_{\text{Src}}$ GP	$\omega_{\text{Lens}}$ UP	$\omega_{\text{Src}}$ UP				
			$\lambda_{\text{Src}}/2$	$\omega_{\text{Lens}}/2$	$\omega_{\text{Src}}/2$				

**Table C6.** Priors used to initialize the light and mass profile fits used for model comparison phase  $\mathbf{P}_{\text{SPLEInit}}$ . All components of the light model use the same priors given in the row labeled the *Light* model, regardless of whether it is a single *Sersic* or multiple *Exp* and/or *Sersic* profile. These are derived from the results of the  $\mathbf{P}_{\text{MCLight}}$  phase, as are the *SPLE*,  $PL_{\text{Core}}$ , and *Shear* parameter initializations.

Model	Component	Prior load	Parameters						
<i>Light</i>	Light	$\mathbf{P}_{\text{MCLight}}$	$x_l$ GP 0.02''	$y_l$ GP 0.02''	$I_l$ GP $I_l/2$	$R_l$ GP $R_l/2$	$m_l$ GP 0.8	$q_l$ GP 0.1	$\theta_l$ GP 20.0
<i>SPLE</i>	Mass	$\mathbf{P}_{\text{MCLight}}$	$x$ GP 0.05''	$y$ GP 0.05''	$\theta_E$ UP <sub>0</sub> <sup>4</sup>	$q$ UP <sub>0.3</sub> <sup>1.0</sup>	$\theta$ GP 60.0	$\alpha$ UP <sub>1.5</sub> <sup>2.5</sup>	
$PL_{\text{Core}}$	Mass	$\mathbf{P}_{\text{Init4}}$	$x$ GP 0.05''	$y$ GP 0.05''	$\theta_E$ GP $\theta_E/2$	$q$ GP 0.1	$\theta$ GP 10.0	$\alpha = 2.0$	$S$ UP <sub>0.0</sub> <sup>0.2''</sup>
Hyper Params	Hyper	Prev	$\omega_{\text{Lens}}$ GP	$\lambda_{\text{Src}}$ GP					
			$\omega_{\text{Lens}}/2$	$\lambda_{\text{Src}}/2$					

**Table C7.** Priors used to initialize the mass profile fits used for model comparison phase  $\mathbf{P}_{\text{MCMass}}$ . Parameter initializations are derived from the results of  $\mathbf{P}_{\text{MCLight}}$ .

Model	Component	Prior Load	Parameters						
<i>SPLE</i>	Mass	$\mathbf{P}_{\text{MCLight}}$	$x$ GP 0.05''	$y$ GP 0.05''	$\theta_E$ UP <sub>0</sub> <sup>4</sup>	$q$ UP <sub>0.3</sub> <sup>1.0</sup>	$\theta$ GP 60.0	$\alpha$ UP <sub>1.5</sub> <sup>2.5</sup>	
$PL_{\text{Core}}$	Mass	$\mathbf{P}_{\text{MCLight}}$	$x$ GP 0.05''	$y$ GP 0.05''	$\theta_E$ UP <sub>0</sub> <sup>4</sup>	$q$ UP <sub>0.3</sub> <sup>1.0</sup>	$\theta$ GP 60.0	$\alpha$ UP <sub>1.5</sub> <sup>2.5</sup>	$s$ UP <sub>0.0</sub> <sup>2.0</sup>
$PL_{\text{Core}}$	Mass	$\mathbf{P}_{\text{Init4}}$	$x$ GP 0.05''	$y$ GP 0.05''	$\theta_E$ GP $\theta_E/2$	$q$ GP 0.1	$\theta$ GP 10.0	$\alpha = 2.0$	$S$ UP <sub>0.0</sub> <sup>0.2''</sup>
<i>NFWsPh</i>	Mass	$\mathbf{P}_{\text{MCMass}}$	$x_d$ (x) GP 0.2''	$y$ (y) GP 0.2''	$\kappa_s$ GP $\kappa_s/2$	$q = 1.0$			
$\Psi_1$	Mass		$\Psi_1$ GP $\Psi_1/2$						
<i>Shear</i>	Mass	$\mathbf{P}_{\text{MCLight}}$	$x_{\text{sh}} = x$	$y_{\text{sh}} = y$	$\gamma_{\text{sh}}$ UP <sub>0.0</sub> <sup>0.4</sup>	$\theta_{\text{sh}}$ UP <sub>0</sub> <sup>180</sup>			
Hyper Params	Hyper	Prev	$\lambda_{\text{Src}}$ GP	$\omega_{\text{Src}}$ UP	$\omega_{\text{Lens}}$ UP				
			$\lambda_{\text{Src}}/2$	$\omega_{\text{Src}}/2$	$\omega_{\text{Lens}}/2$				

**Table C8.** Priors used to initialize the light and mass profile fits used for model comparison phase  $\mathbf{P}_{\text{LMCGeom}}$ . All components of the light model use the same priors given in the row labeled the *Light* model, regardless of whether it is a single *Sersic* or multiple *Dev*, *Exp* and/or *Sersic* profile. These are derived from the results of the  $\mathbf{P}_{\text{MCLight}}$  phase, whereas the *Shear* parameter initializations are derived from the results of  $\mathbf{P}_{\text{MCMass}}$ .

Model	Component	Prior load	Parameters						
<i>Light</i>	Light	$\mathbf{P}_{\text{MCLight}}$	$x_l$ GP 0.02''	$y_l$ GP 0.02''	$I_l$ GP $I_l/2$	$R_l$ GP $R_l/2$	$n_l$ GP 0.8	$q_l$ GP 0.1	$\theta_l$ GP 20.0
<i>NFW</i>	Mass	$\mathbf{P}_{\text{MCMass}}$	$x_d(x)$ GP 0.2''	$y(y)$ GP 0.2''	$\kappa_s$ UP $\frac{3}{0}$	$q$ UP $\frac{1.0}{0.2}$	$\theta$ UP $\frac{180.0}{0.0}$		
$\Psi_1$	Mass		$\Psi_l$ UP $\frac{5.0}{0.0}$						
<i>Shear</i>	Mass	$\mathbf{P}_{\text{MCLight}}$	$x_{\text{sh}} = x_d$	$y_{\text{sh}} = y_d$	$\gamma_{\text{sh}}$ GP 0.1	$\theta_{\text{sh}}$ GP 90.0			
Hyper	Hyper	Prev	$\lambda_{\text{Src}}$ GP	$\omega_{\text{Lens}}$ UP	$\omega_{\text{Src}}$ UP				
Params			$\lambda_{\text{Src}}/2$	$\omega_{\text{Lens}}/2$	$\omega_{\text{Src}}/2$				

**Table C9.** Priors used to initialize the hyperparameter optimization. After the first hyperparameter initialization, hyperparameters are initialized using Gaussian priors with a width half their median value.

Model	Feature	Parameters						
<i>Hyper</i>	Source Adaption	$L_{\text{Clust1}}$ UP $\frac{0.0}{10.0}$	$L_{\text{Clust2}}$ UP $\frac{0.0}{10.0}$	$N_s$ UP $\frac{80.0}{800.0}$				
	Contribution Maps	$\omega_{\text{Frac}}$ UP $\frac{0.0}{10.0}$						
	Variance Scaling	$\omega_{\text{BG}}$ UP $\frac{0.0}{2.0}$	$\omega_{\text{Lens}}$ UP $\frac{0.0}{6.0}$	$\omega_{\text{Lens2}}$ UP $\frac{0.0}{6.0}$	$\omega_{\text{Src}}$ UP $\frac{0.0}{6.0}$	$\omega_{\text{Src2}}$ UP $\frac{0.0}{6.0}$		
	Regularization	$\lambda_{\text{Src}}$ UP $\frac{0.0}{1000.0}$	$\lambda_{\text{Bg}}$ UP $\frac{0.0}{1000.0}$	$L_{\text{Lum}}$ UP $\frac{0.0}{5.0}$				
	Sky Subtraction	$\omega_{\text{Sky}}$ UP $\frac{-1.0}{1.0}$						

## APPENDIX D: PARAMETER TABLE

**Table D1.** Parameter symbols and descriptions for all parameters used in this work.

Lensing quantities	
$M_{\text{Ein}}$	Einstein mass
$R_{\text{Ein}}$	Einstein radius
$\mathbf{x}$	Image-plane coordinate (image-plane reference frame)
$\mathbf{x}'$	Image-plane coordinate (lens reference frame)
$\kappa(\mathbf{x})$	Lens convergence profile
$\phi$	Lens deflection potential
$\alpha_{x,y}$	Deflection angle map ( $x$ and $y$ dimensions)
$\Sigma_{\text{cr}}$	Critical surface mass density
Light profiles	
$x_l, y_l$	Centers (arcseconds) [ <i>Exp, Sersic</i> ]
$\theta_l$	Rotation angle (clockwise from north) [ <i>Exp, Sersic</i> ]
$q_l$	Axis ratio [ <i>Exp, Sersic</i> ]
$\xi_l$	Elliptical coordinate ( $\xi_l = \sqrt{x_l^2 + y_l^2/q_l^2}$ ) [ <i>Exp, Sersic</i> ]
$I_l$	Intensity (electrons per second) [ <i>Exp, Sersic</i> ]
$R_l$	Effective radius (circular) [ <i>Exp, Sersic</i> ]
$n_l$	Sersic index [ <i>Exp</i> ( $n_l = 1.0$ ), <i>Sersic</i> ]
$k_l$	Function of Sersic index [ <i>Exp, Sersic</i> ]
Mass profiles	
$x, y$	Centers (arcseconds) [ <i>SIE, SPLE, PLCore</i> ]
$\theta$	Rotation angle (clockwise from north) [ <i>SIE, SPLE, PLCore</i> ]
$q$	Axis ratio [ <i>SIE, SPLE, PLCore</i> ]
$\xi$	Elliptical coordinate ( $\xi = \sqrt{x^2 + y^2/q^2}$ ) [ <i>SIE, SPLE, PLCore</i> ]
$\theta_{\text{Ein}}$	Einstein radius (arcseconds) [ <i>SIE, SPLE, PLCore</i> ]
$\alpha$	Power-law density slope ( $\rho(r) = \rho_o(r/r_o)^{-\alpha}$ , $\alpha_{x,y}$ ) [ <i>SIE, SPLE, PLCore</i> ]
$S$	Core radius (arcseconds) [ <i>PLCore</i> ]
$x_d, y_d$	Centers (arcseconds) [ <i>NFW</i> ]
$\theta_d$	Rotation angle (clockwise from north) [ <i>NFW</i> ]
$q_d$	Axis ratio [ <i>NFW</i> ]
$\xi_d$	Elliptical coordinate ( $\xi_d = \sqrt{x_d^2 + y_d^2/q_d^2}$ ) [ <i>NFW</i> ]
$\rho_s$	Halo scale normalization [ <i>NFW</i> ]
$r_s$	Halo scale radius ( $r_s = 30$ kpc) [ <i>NFW</i> ]
$\kappa_d$	Halo normalization $\kappa_d = \rho_s r_s / \Sigma_{\text{cr}}$ [ <i>NFW</i> ]
$\eta_d$	Scaled elliptical coordinate ( $\eta_d = \xi_d / r_s$ ) [ <i>NFW</i> ]
$\Psi_l$	Mass-to-light ratio (electrons per second) [ <i>Exp, Sersic</i> ]
$x_{\text{sh}}, y_{\text{sh}}$	Centers (arcseconds) [ <i>Shear</i> ]
$\gamma_{\text{sh}}$	Magnitude [ <i>Shear</i> ]
$\theta_{\text{sh}}$	Rotation angle (clockwise from north) [ <i>Shear</i> ]
Source profiles	
$x_s, y_s$	Centers (arcseconds) [ <i>Exp, Sersic</i> ]
$\theta_s$	Rotational angle (clockwise from north) [ <i>Exp, Sersic</i> ]
$q_s$	Axis ratio [ <i>Exp, Sersic</i> ]
$\xi_s$	Elliptical coordinate ( $\xi_s = \sqrt{x_s^2 + y_s^2/q_s^2}$ ) [ <i>Exp, Sersic</i> ]
$I_s$	Intensity (electrons per second) [ <i>Exp, Sersic</i> ]
$R_s$	Effective radius (circular) [ <i>Exp, Sersic</i> ]
$n_s$	Sersic index [ <i>Exp</i> ( $n_s = 1.0$ ), <i>Sersic</i> ]
$k_s$	Function of Sersic index [ <i>Exp, Sersic</i> ]

**Table D2.** Parameter symbols and descriptions for all parameters used in this work.

Semi-linear inversion	
$I$	Total source pixels
$i$	Source pixel number
$J$	Total image pixels
$j$	Image pixel number
$f_{i,j}$	Matrix mapping image pixels to source pixels
$d$	Observed image values in (electrons per second)
$\sigma$	Observed image statistical uncertainties (electrons per second)
$\mathbf{b}$	Model lens light profile values
$\mathbf{s}$	Model reconstructed source surface-brightness values
$\mathbf{D}$	Observed image mapping vector used for linear inversion (see Warren & Dye (2003))
$\mathbf{F}$	Image-source plane mapping matrix used for linear inversion (see Warren & Dye (2003))
$\mathbf{H}_\Lambda$	Regularization matrix
$\chi^2$	Residuals over uncertainties squared
$\epsilon$	Linear inversion bayesian evidence
Hyper parameters	
$N_s$	Source-plane resolution (number of source pixels)
$L_{\text{clust1}}, L_{\text{clust2}}$	Control the source pixelization
$\omega_{\text{LensFrac}}, \omega_{\text{SrcFrac}}$	Control the lens and source contribution maps
$\omega_{\text{BG}}$	Scale the background sky variances
$\omega_{\text{Lens1}}, \omega_{\text{Lens2}}$	Scale the lens light variances
$\omega_{\text{Src1}}, \omega_{\text{Src2}}$	Scale the lensed source variances
$\omega_{\text{Sky}}$	Scale the background sky subtraction
$\lambda$	Regularization coefficient (constant regularization scheme)
$\lambda_{\text{Src}}$	Source regularization coefficient (non-constant scheme)
$\lambda_{\text{Bg}}$	Background regularization coefficient (non-constant scheme)
$L_{\text{Lum}}$	Controls transition of non-constant regularization
Adaptive image/source vectors	
$\Xi$	Preloaded model of lensed source (from previous pipeline phase)
$\mathcal{L}$	Preloaded model of lens galaxy (from previous pipeline phase)
$K$	Number of image pixels allocated to a given source pixel
$\mathbf{E}$	Cluster energies used for $k$ -means clustering
$\mathbf{r}$	Distances of $K$ traced image pixels to allocated source pixel
$\mathbf{W}$	Weights of each source pixel, used for surface-brightness adaption
$\mathbf{T}$	Preloaded source and lens image $\mathbf{T} = \Xi + \mathcal{L}$
$\Omega_{\text{Src}}$	Lensed source flux contribution map
$\Omega_{\text{Lens}}$	Lens light flux contribution map
$\mathbf{v}$	Source flux contribution of each source-pixel (computed from $\Omega_{\text{Src}}$ )
$\mathbf{V}$	Weights of each source pixel, used for luminosity-weighted regularization
$\Lambda, (\lambda_{\text{eff}})$	Effective regularization coefficients
$f_{\text{Bg}}$	Background sky flux used for sky subtraction
$\sigma_{\text{base}}$	Observed image variances without scaling (counts)
$\chi^2_{\text{base}}$	$\chi^2$ values generated using unscaled baseline variances
$\sigma_{\text{scale}}$	Observed image variances including scaling (counts)
$\chi^2_{\text{scale}}$	$\chi^2$ Values generated using scaled variances

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.