Massive elliptical galaxies at $z \sim 0.2$ are well described by stars and a Navarro–Frenk–White dark matter halo

Anowar J. Shajib,[1]* Tommaso Treu[1]†, Simon Birrer[2] and Alessandro Sonnenfeld[3]

[1]*Department of Physics and Astronomy, University of California, Los Angeles, CA 90095-1547, USA*
[2]*Kavli Institute for Particle Astrophysics and Cosmology and Department of Physics, Stanford University, Stanford, CA 94305, USA*
[3]*Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, the Netherlands*

ABSTRACT

We investigate the internal structure of elliptical galaxies at $z \sim 0.2$ from a joint lensing–dynamics analysis. We model *Hubble Space Telescope* images of a sample of 23 galaxy–galaxy lenses selected from the Sloan Lens ACS (SLACS) survey. Whereas the original SLACS analysis estimated the logarithmic slopes by combining the kinematics with the imaging data, we estimate the logarithmic slopes only from the imaging data. We find that the distribution of the lensing-only logarithmic slopes has a median 2.08 ± 0.03 and intrinsic scatter 0.13 ± 0.02, consistent with the original SLACS analysis. We combine the lensing constraints with the stellar kinematics and constrain the amount of adiabatic contraction in the dark matter (DM) halos. We find that the DM halos are well described by a standard Navarro–Frenk–White halo with no contraction on average for both of a constant stellar mass-to-light ratio (M/L) model and a stellar M/L gradient model. For the M/L gradient model, we find that most galaxies are consistent with no M/L gradient. Comparison of our inferred stellar masses with those obtained from the stellar population synthesis method supports a heavy initial mass function (IMF) such as the Salpeter IMF. We discuss our results in the context of previous observations and simulations, and argue that our result is consistent with a scenario in which active galactic nuclei feedback counteracts the baryonic-cooling-driven contraction in the DM halos.

Key words: gravitational lensing: strong – galaxies: elliptical and lenticular, cD

1 INTRODUCTION

Measurements of structural properties of elliptical galaxies can test prediction of galaxy formation theories in the cold dark matter (CDM) paradigm (e.g., Dubinski 1994; Kazantzidis et al. 2004; Debattista et al. 2008; Read 2014). In this paradigm, small haloes merge to hierarchically form larger halos. CDM-only *N*-body simulations predict that the dark matter is universally distributed according to the Navarro–Frenk–White (NFW) profile with a 'cuspy' central slope, i.e., the 3D density scales as $\rho \propto r^{-1}$ in the inner region (Navarro et al. 1996, 1997). Such slopes have been observed in galaxy clusters (e.g., Limousin et al. 2007; Caminha et al. 2017). However, shallower central density slopes also have been observed in some galaxy clusters and in dwarf and low-surface-brightness galaxies (e.g., de Blok et al. 2001; Sand et al. 2008; Oh et al. 2011; Newman et al. 2013). One possible explanation of these shallower slopes is given by alternative dark matter models, e.g., self-interacting dark matter and warm dark matter (e.g., Dodelson & Widrow 1994; Spergel & Steinhardt 2000; Colín et al. 2000). Another possible explanation,

instead, is related to the astrophysics of galaxy formation and evolution. In fact, as massive elliptical galaxies are believed to be the end-product of the hierarchical merging and accretion processes, their mass density profiles are a sensitive probe of the physics of galaxy formation and evolution.

In the galaxy formation process, baryons play an important role that can affect the central density slope of the dark matter distribution. If gas cools and inflows slowly towards the center, then the dark matter distribution can be adiabatically contracted (Blumenthal et al. 1986). In contrast, the dark matter distribution can expand in dissipationless mergers or due to gas outflows driven by supernova feedback, stellar feedback, or active galactic nucleus (AGN) feedback (e.g., El-Zant et al. 2001; Nipoti et al. 2003; Peirani et al. 2008; Pontzen & Governato 2012). Observational avenues to study these processes and their importance in galaxy formation have been limited. The combination of lensing and dynamics has been one of the most informative probes of the dark and luminous matter distributions in the inner region of galaxies and clusters (e.g., Treu & Koopmans 2002; Czoske et al. 2008; Barnabè et al. 2011). The elliptical galaxies in the Sloan Lens ACS (SLACS) survey were found to have no contraction on average (Dutton & Treu 2014; Newman et al. 2015), although individual galaxies can have contracted (or,

* E-mail: ajshajib@astro.ucla.edu
† Packard Fellow

expanded) dark matter distributions (e.g., Sonnenfeld et al. 2012). As many or all of the above mentioned baryonic processes happen at various points of the galaxy formation process, the degree of contraction or expansion depends on the relative importance of these baryonic processes.

The interplay between baryon and dark matter in galaxy formation is also highlighted by the so-called 'bulge-halo conspiracy' (Treu et al. 2006; Humphrey & Buote 2010; Cappellari 2016). This conspiracy refers to the nearly isothermal profiles of the total matter distribution with small scatter (∼0.1–0.2) observed within half of the half-light radii to 100 half-light radii of elliptical galaxies (e.g., from strong and weak lensing: Treu & Koopmans 2004; Gavazzi et al. 2007; Auger et al. 2010b; Ritondale et al. 2019; from stellar dynamics: Thomas et al. 2007; Tortora et al. 2014; Bellstedt et al. 2018). As neither of the baryonic and dark matter distributions follows a power law, fine tuning between these two distributions is required to produce the isothermal distribution for the total mass. To understand the origin of this conspiracy through simulation, we can use two parameters: the dark matter fraction f_{dm} within the inner region and the distribution of logarithmic slope γ for the total mass profile. Near the half-light radius, dark matter has as a logarithmic slope of $\gamma \leq 1.5$ and the baryonic distribution has a logarithmic slope of $\gamma \sim 2.3$ for a de Vaucouleurs profile (de Vaucouleurs 1948). Thus, fine-tuning in f_{dm} is required to achieve $\gamma \sim 2$ for the combination of baryonic and dark matter. However, cosmological hydrodynamic simulations have been unable to match both the observed f_{dm} and slope distribution. The γ distribution can be reproduced in simulations by having no or weak feedback, but this leads to an overestimated galaxy formation efficiency and underestimated f_{dm} (Naab et al. 2007a; Duffy et al. 2010; Johansson et al. 2012). Conversely, reproducing the observed f_{dm} requires strong feedback, but then the predicted γ distribution is too shallow. Dubois et al. (2013) similarly find that simulation with AGN feedback can predict the observed f_{dm} in 0.4–8 ×10^{13} M$_\odot$ halos at $z = 0$, but underestimate the γ distribution. Without the AGN feedback, overestimated galaxy formation efficiency leads to underestimated f_{dm} and overestimated γ distribution. More recently, Xu et al. (2017) studied simulated elliptical galaxies from the IL-LUSTRIS hydrodynamic simulation that incorporates a number of baryonic processes (Vogelsberger et al. 2014). These authors find higher f_{dm} and lower average γ in ILLUSTRIS galaxies than those observed in lens elliptical galaxies (Auger et al. 2009; Oldham & Auger 2018). These mismatches point to either inadequacy in the theoretical model or systematic biases in the observational methods.

Many of the observational constraints on f_{dm} and γ come from strong lensing systems with elliptical galaxies as deflector. Strong gravitational lensing provides a robust probe of the total projected mass within the Einstein radius. Thus, combining the stellar kinematics with the lensing information can constrain the mass distribution in the deflector galaxy (e.g., Auger et al. 2009; Sonnenfeld et al. 2015). To constrain f_{dm}, decoupling the baryonic and dark components in the total mass is necessary. In previous studies, the stellar mass was inferred from the spectral energy distribution to decouple the dark and baryonic components (e.g., Auger et al. 2009; Spiniello et al. 2011). This stellar mass depends on the assumption of the stellar initial mass function (IMF) introducing an uncertainty by a factor of ∼3. However, the IMF can be constrained by making assumption on the mass profile (Treu et al. 2010), or by decoupling the baryonic and dark components with other external constraints (e.g., Spiniello et al. 2012; Barnabè et al. 2013; Sonnenfeld et al. 2019b).

In this paper, we aim to constrain the γ distribution and f_{dm}

of elliptical galaxies from the lensing and kinematics data – independent of the SED-based stellar mass measurements – and to constrain the amount of adiabatic contraction (or expansion) in these galaxies. We model a sample of 23 galaxy–galaxy lenses to study their structural properties. These lenses are assembled from the SLACS survey (Bolton et al. 2006, 2008). Previous SLACS analyses measured only the Einstein radius from the imaging data and then constrained the radially averaged logarithmic slope using stellar kinematics in combination with the imaging data. Due to several improvements in lens modelling techniques in the past decade, we can now constrain the logarithmic slope only from the imaging data, exploiting the richness of pixel-level information in the lensed arcs (e.g., Suyu & Halkola 2010; Birrer et al. 2015). We model the SLACS lenses in our sample using state-of-the-art lens modelling techniques that simultaneously reconstruct the sources to extract the information contained in the lensed arcs. Thus, we measure the local logarithmic slope at the Einstein radius only from the imaging data, independent of the stellar kinematics. We then combine the stellar kinematics with the lensing constraints to individually constrain the stellar and dark matter distributions, and infer the amount of adiabatic contraction in the dark matter distribution. Lensing-only measurement of the mass distribution is prone to the mass-sheet degeneracy (MSD; Falco et al. 1985). We adopt MSD-invariant quantities as the lensing constraints in our joint lensing–dynamics analysis, and combining these constraints with the stellar kinematics allows us to constrain the MSD in the inferred mass distribution.

This paper is organized as follows. In Section 2, we describe our lens sample and the imaging data. Next in Section 3, we describe our uniform modelling procedure for this sample. We report the structural properties of elliptical lens galaxies from the lens models in Section 4. We discuss our results in Section 5 and summarize the paper in Section 6. Additionally in Appendix A, we investigate the alignment between mass and light distributions. We adopt a flat Λ cold dark matter model as the fiducial cosmology with H_0=70 km s^{-1} Mpc^{-1}, and $\Omega_m = 0.3$. The reported uncertainties are obtained from 16th and 84th percentiles of the corresponding posterior probability distributions. We use $\log x$ to express the natural logarithm and $\log_{10} x$ to express the common logarithm.

2 LENS SAMPLE

Our lens sample consists of 23 galaxy–galaxy lenses from the SLACS survey. We first selected 50 galaxies from the full SLACS sample of 85 lenses by visually inspecting the lens images and selecting those: (i) without nearby satellite or line-of-sight galaxies, (ii) without highly complex source morphology, (iii) with *HST* imaging data in the F555W/F606W bands (hereafter, V-band), and (iv) not disc-like. Criteria (i) and (ii) are adopted so that we can uniformly apply our modelling procedure to the whole sample without needing to tweak the lens model settings on lens-by-lens basis. We adopt criterion (iii), because in the V-band the deflector galaxy is relatively fainter in comparison with the lensed arcs than in the F814W band (hereafter, *I*-band), which makes it easier to decouple the deflector light from the lensed arcs during lens modelling. We provide the list of selected galaxies in Appendix C.

2.1 Imaging data

Among the selected galaxies, some have imaging data from Advanced Camera for Surveys (ACS), and the rest from Wide Field and Planetary Camera 2 (WFPC2). The ACS images were taken

with the F555W filter and the WFPC2 images were taken with the F606W filter. The images are obtained under the *HST* GO programs 10494 (PI: Koopmans), 10798 (PI: Bolton), 10886 (PI: Bolton), and 11202 (PI: Koopmans).

The WFPC2 images were reduced for the original SLACS analysis (Auger et al. 2009). We reduce the ACS images using the standard ASTRODRIZZLE software package (Avila et al. 2015). The final pixel scale after drizzling is 0.05 arcsec.

We obtain the point spread function (PSF) for each filter and camera combination using TINYTIM (Krist et al. 2011).

2.2 Stellar kinematics data

We use the line-of-sight velocity dispersions of the lenses in our sample measured from the Sloan Digital Sky Survey (SDSS) spectra. The fibre radius is 1.5 arcsecond and the typical seeing for the observations is 1.4 arcsec. Bolton et al. (2008) first measured the velocity dispersions from the SDSS reduction pipeline. Shu et al. (2015) improved the measurements by updating the set of templates used to fit the spectra. We use this improved kinematics measurements in this study. These measurements are in good agreement with the Very Large Telescope (VLT) X-Shooter measurements of a subsample presented by Spiniello et al. (2015).

Birrer et al. (2020) find a residual scatter in the joint lensing–dynamics analysis using the same kinematics data used in this study, which accounts for ~6 per cent unaccounted systematic uncertainty in the measured kinematics. Therefore, we add 6 per cent uncertainty in quadrature to the measured uncertainties.

2.3 Weak lensing data

We incorporate weak lensing shear measurements of a sample of 33 SLACS lenses. The measurement pipeline is described by Gavazzi et al. (2007), and the sample size of the analyzed SLACS lenses is increased by Auger et al. (2010a). Sonnenfeld et al. (2018) find that the shear measurements of these 33 lenses is consistent with the shear measurement of the Hyper-Suprime Cam (HSC) survey weak lensing measurements after weighting the HSC sample to match the stellar mass distribution of the SLACS lenses.

Out of the 23 SLACS lenses in our sample, 11 have directly measured reduced shear $\gamma_{\rm shear}/(1 - \kappa)$. For the remaining 12 lenses, we adopt the mean and scatter (which includes both the intrinsic scatter and the noise) of the measured reduced shears for the 33 lenses as the measured value and uncertainty, respectively. We adopt binned reduced shears only up to ~100 kpc as the weak lensing constraint in our analysis. We do not use measurements beyond 100 kpc to avoid any potential bias from the 2-halo term as this term is not accounted for in our model. The centres of the adopted four bins are logarithmically spaced at 9.87 kpc, 17.78 kpc, 32.04 kpc, and 57.72 kpc.

3 LENS MODELLING

We model the lenses using the lens modeling software LENSTRONOMY, which is publicly available on GitHub[1] (Birrer et al. 2015; Birrer & Amara 2018). We package our modelling code into the

---

DOLPHIN pipeline[2], which is a wrapper for LENSTRONOMY to uniformly model large lens samples. First in Section 3.1, we describe the components in our uniform lens model. Then in Section 3.2, we describe the optimization procedure for the lens model and Bayesian inference of the model parameters. Next in Section 3.3, we assess the effect of the PSF on the measured logarithmic slopes.

3.1 Model components

We adopt the power-law ellipsoidal mass distribution (PEMD) for the deflector (Barkana 1998). Although we aim to individually constrain the dark matter and stellar distributions from a joint lensing–dynamics analysis, first we only need to constrain MSD-invariant local lensing properties from our lens models, for which the PEMD model is sufficient. We combine these local lensing constraints with the stellar kinematics data in our joint lensing–dynamics analysis to individually constrain the dark matter and stellar distributions in Section 4.2. The convergence for the PEMD is given by

$$\kappa(x, y) = \frac{3 - \gamma}{2} \left(\frac{R_{\rm E}}{\sqrt{q_{\rm m} x^2 + y^2/q_{\rm m}}} \right)^{\gamma - 1}, \qquad (1)$$

where $R_{\rm E}$ is the Einstein radius, q is the axis ratio, and γ is the logarithmic slope for the mass distribution in 3D. For an isothermal profile, the logarithmic slope is $\gamma = 2$. The on-sky coordinates (x, y) are rotated by position angle $\mathrm{PA_m}$ from the (RA, dec) coordinates to align the x–axis with the major axis of the projected mass distribution. We also adopt an external shear profile parametrized with the shear magnitude $\gamma_{\rm ext}$ and the shear angle $\phi_{\rm ext}$.

We adopt a double Sérsic profile for the deflector's light distribution, as a single Sérsic profile leaves significant residual at the galaxy's center (Claeskens et al. 2006; Suyu et al. 2013). The Sérsic profile is given by

$$I(x, y) = I_{\rm e} \exp \left[-k \left\{ \left(\frac{\sqrt{x^2 + y^2/q_{\rm L}^2}}{R_{\rm eff}} \right)^{1/n_{\rm s}} - 1 \right\} \right], \qquad (2)$$

where $R_{\rm eff}$ is the effective radius, $I_{\rm e}$ is the surface brightness at $R_{\rm eff}$, $q_{\rm L}$ is axis ratio, $n_{\rm s}$ is the Sérsic index, and k is a normalizing constant so that $R_{\rm eff}$ becomes the half-light radius (Sérsic 1968). The position angle for the light distribution is $\mathrm{PA_L}$. To constrain the degeneracy between the pairs of $R_{\rm eff}$ and $n_{\rm s}$ in the double Sérsic profile, we fix $n_{\rm s} = 1$ and $n_{\rm s} = 4$, i.e., the exponential profile and the de Vaucouleurs profile respectively (de Vaucouleurs 1948). For simplicity, we also join the ellipticity parameters $q_{\rm L}$ and $\mathrm{PA_L}$ between the two Sérsic profiles.

We reconstruct the source galaxy's light distribution with a basis of shapelets and a Sérsic profile (Refregier 2003; Birrer et al. 2015). The order parameter $n_{\rm max}$ determines the number of shapelets as $N_{\rm shapelets} = (n_{\rm max} + 1)(n_{\rm max} + 2)/2$. The scale size of the shapelets is ruled by the scaling parameter β.

3.2 Optimization and inference

We obtain the posterior probability distributions for our model parameters using the Markov chain Monte Carlo (MCMC) method. If the MCMC sampling is started from a point close to the maxima of the posterior, the chain can converge with relatively less

---

[1] ⟲ https://github.com/sibirrer/lenstronomy

[2] ⟲ https://github.com/ajshajib/dolphin

computational time. Therefore, we first optimize the lens model to get a point close to the maxima of the posterior. We use the particle swarm optimization (PSO) method for this step (Kennedy & Eberhart 1995). To further make this optimization computationally efficient, we adopt the following optimization recipe:

1. Join the deflector mass and light centroids and fix the logarithmic slope $\gamma = 2$ and shear magnitude $\gamma_{ext} = 0$ for all the steps below.

2. Create a mask for the lensed arcs. We provide the algorithm to automatically make the mask for the arcs in Appendix B. Fix all the model parameters except for the deflector's light profile. Optimize the deflector light parameters masking the lensed arcs.

3. To find a good starting point for the source light parameters, fix the lens model parameters and the deflector light parameters. Fix the Einstein radius R_E and ellipticity parameters $\{q_m, PA_m\}$ to the values measured by the SLACS analysis (Auger et al. 2009). If such pre-determined values are not available, R_E can be fixed to an approximate guess and the ellipticity parameters can be set to the values for the circular case. Fix the shapelet scale parameter $\beta = 0.1$ arcsecond. Optimize only the remaining source light parameters to find an approximate position of the source on the source plane. Note, in this step the lensed arcs are not masked.

4. Keep the deflector light parameters fixed and optimize for the source parameters and the PEMD parameters together. Keep $\beta = 0.1$ arcsecond fixed in this step.

5. Free β and optimize all the non-fixed parameters together.

6. Repeat steps 2–5 with the current initial conditions from the previous step.

We notice prominent residuals at the center of the deflector after performing the above automated procedure. This is expected because the centers of elliptical galaxies are not perfectly described by Sérsic profiles and our signal-to-noise ratio is very high in the center. To avoid bias in the model from this poor fitting of the deflector light profile at the center, we mask out the central 0.4 arcsecond and rerun the whole fitting procedure. The deflector light profile parameters for these systems are constrained from the light distribution that falls outside the central masked region. However for four systems – J0252+0039, J1112+0826, J1313+4615, and J1636+4707 – masking the deflector center leads to even poorer quality fits as evaluated with the p-value of the χ^2 statistic. Therefore, we do not mask the deflector centers for these four systems.

We set $n_{max} = 6$ for most of the lens systems in our sample. However, for the following systems we have adopted the following n_{max} values through trial-and-error with the above optimization procedure – J0252+0039: 10, J0959+0410: 15, J1250+0523: 12, J1313+4615: 10, J1630+4520: 15.

After the pre-sampling optimization, we then initiate the MCMC sampling from the optimized lens model after step 6. During the sampling, we free the parameters γ and γ_{ext} that were fixed in the optimization step. We also independently sample the centroids of the deflector mass and light distributions. We perform the MCMC sampling using EMCEE, which is an affine-invariant ensemble sampler (Goodman & Weare 2010; Foreman-Mackey et al. 2013). We assure the convergence of the chain by checking that the median and standard deviation of the EMCEE walkers at each step have reached equilibrium.

After our uniform modelling procedure for the 50 initially selected lenses, we vet for reliability of the lens models based on the following criteria: (i) absence of prominent model residuals indicative of poor source reconstruction, and (ii) the median of the

inferred logarithmic slope has not converged to too low ($\gamma \lesssim 1.4$) or too high ($\gamma \gtrsim 2.8$) values. The γ distribution can converge toward such extreme values if the lensed arcs are faint and thus the lensing information contained in the imaging data is not sufficient to constrain γ. When γ is too small ($\gamma \lesssim 1.4$), a central image will be produced in the lens models, which is not observed in the imaging data by looking at the color distribution. However, since we mask out the central region for most of our lens systems, a central image is not directly penalized in the likelihood term. Moreover, since the lensed arcs are relatively fainter, prominent residuals are not noticeable in the lens models even if the source reconstruction is poor. We treat such extreme values of γ as numerical artifacts due to weak constraining power of the imaging data and remove these systems from our sample. Note that criterion (ii) is practically a uniform prior median(γ) $\sim \mathcal{U}(1.4, 2.8)$. After this vetting procedure, we are left with 23 systems with reliable lens models. We show the lens images and the models in Figures 1, 2, and 3. We tabulate the marginalized posteriors of the lens model parameters in Table 1.

3.3 Effect of PSF on measured logarithmic slope

We check the effect of the PSF choice on the measured logarithmic slopes. We adopt the following five PSF choices for this test:

(i) TINYTIM PSF with G2V star SED as our baseline PSF,

(ii) TINYTIM PSF with Sc galaxy SED at redshift $\langle z_s \rangle = 0.62$, which is the average source redshift in our sample,

(iii) TINYTIM PSF with elliptical galaxy SED at redshift $\langle z_d \rangle = 0.20$, which is the average source redshift in our sample,

(iv) a PSF created from re-centering and stacking star cutouts from the corresponding *HST* image, and

(v) effective PSF (ePSF) created from the stars from the corresponding *HST* image (Anderson & King 2000).

For PSF choices (iv) and (v), reliable PSFs could be extracted only for the ACS images. Therefore, we only check with PSF choice (iv) and (v) for nine systems with ACS images from our sample. We re-optimize our lens models for PSF types (ii) to (v). We compare the measured logarithmic slopes with these PSF choices in Figure 4. The mean deviation $|\Delta| \sim 0.001–0.009$ of the logarithmic slope distribution is negligible relative to the uncertainty of the individual logarithmic slope (typically $\sigma_\gamma \gtrsim 0.04$) when the SED of the PSF is varied. However, a scatter of approximately 0.02–0.03 is introduced in the logarithmic slope distribution from the choice of PSF.

4   STRUCTURAL PROPERTIES OF LENS GALAXIES

In this section, we report our findings on the structural properties of the lens galaxies. In Section 4.1, we present the distribution of the logarithmic slopes constrained from the lensing-only data. Next in Section 4.2, we combine stellar kinematics with the lensing observables to infer the amount of contraction in our sample.

4.1   The distribution of logarithmic slopes

We plot the distribution of the SLACS lenses in our sample on the $\gamma - R_E/R_{eff}$ plane in Figure 14 left-hand panel. For comparison, we also plot the distribution of the 13 strongly lensed quasar systems from Shajib et al. (2019). Strong lensing constrains the slope of the projected mass profile at the Einstein radius. Therefore, if we assume that elliptical galaxies are self-similar, then this plot illustrates the distribution of slopes of the projected mass density at different
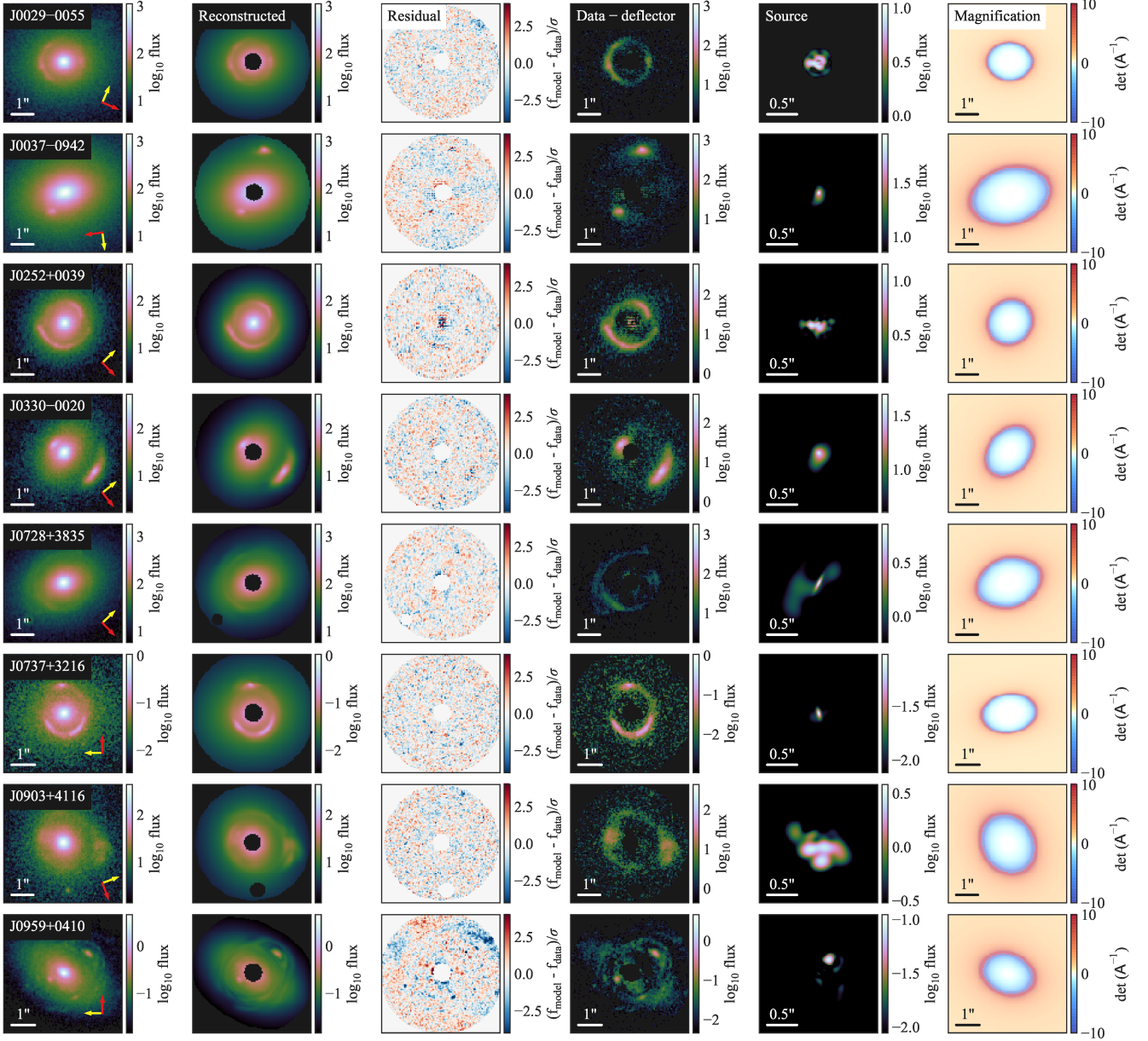
Figure 1. Lens models for the first eight out of the 23 SLACS lenses in our sample. The first column shows the *HST* images in F555W or F606W band. The second column shows the reconstructed image from the best-fit lens model and the third column shows the normalized residual for the reconstruction. In the fourth column, we subtract the modelled deflector light distribution from the observed image to visualize the lensed arcs. The fifth column shows the reconstructed source and the sixth column shows the magnification map of the lens model. The red arrow points to the North and the yellow arrow points to the East. We mask the central 0.4 arcsecond of some deflectors where there would otherwise be large residuals in the deflector's light profile fit without the mask. The lens models for the remaining systems are shown in Figures 2 and 3.

scale sizes. The median of the distribution is $2.09^{+0.03}_{-0.04}$ and the intrinsic scatter of the distribution is 0.13 ± 0.02. We account for the uncertainty in individually measured γ by sampling 1000 sets of γ's assuming Gaussian uncertainty for each measurement and then obtaining the distribution of the medians and scatters from these 1000 sets.

We find no correlation between observed γ distribution and ratio $f_{\mathrm{SIE}} \equiv \sigma_{\mathrm{los}}/\sigma_{\mathrm{SIE}}$ between the observed velocity dispersion and the SIE velocity dispersion (Figure 5). We calculate the SIE

velocity dispersion is given by

$$\sigma_{\mathrm{SIE}} = c\sqrt{\frac{R_{\mathrm{E}}D_{\mathrm{s}}}{4\pi D_{\mathrm{ds}}}}, \tag{3}$$

where c is the speed of light. If the underlying true mass profile follows a power law, then γ should positively correlate with f_{SIE} [see equation (2.3) of Koopmans (2004), cf. Figure 4 of Auger et al. (2010b)]. Therefore, there are two possible explanations for this lack of correlation: (i) the underlying true mass profile deviates from a power-law, and/or (ii) the noise is large enough that any potential correlation is washed away. We perform a linear regression for the
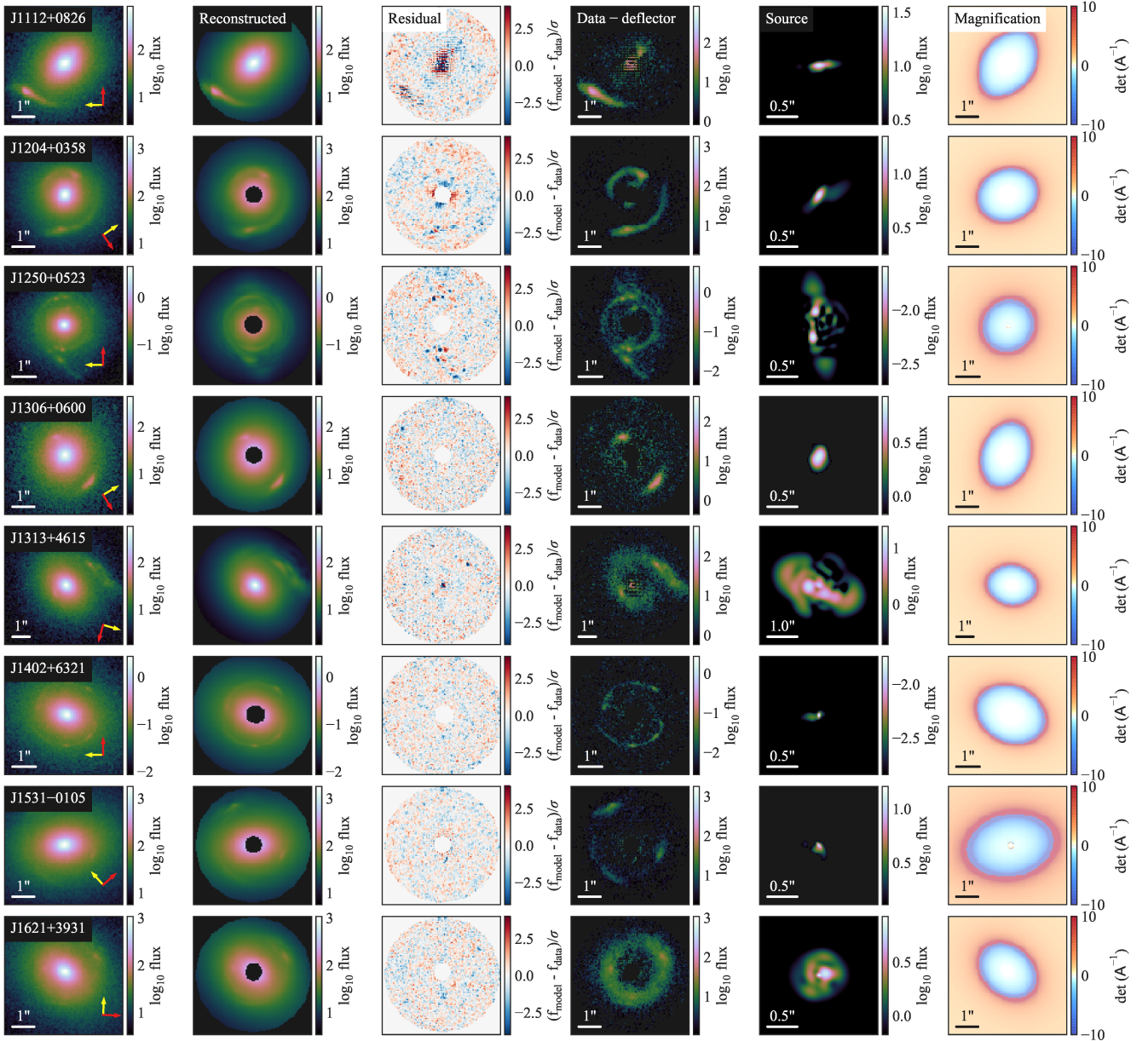
Figure 2. Lens models for the next eight out of the 23 SLACS lenses continuing after Figure 1. The first column shows the *HST* images in F555W or F606W band. The second column shows the reconstructed image from the best-fit lens model and the third column shows the normalized residual for the reconstruction. In the fourth column, we show subtract the modelled deflector light distribution from the observed image to visualized the lensed arcs. The fifth column shows the reconstructed source and the sixth column shows the magnification map of the lens model. The red arrow points to the North and the yellow arrow points to the East. We mask the central 0.4 arcsecond of some deflectors where there would otherwise be large residuals in the deflector's light profile fit without the mask.

$\gamma-f_{\rm SIE}$ distribution with intrinsic scatter as a free parameter. We find that only 3.5 per cent or less intrinsic scatter is necessary with 95 per cent confidence to fit the data. This result indicates that the uncertainties in γ and $f_{\rm SIE}$ are too large to detect the expected correlation for a power-law profile. This is a case of absence of evidence, not evidence of absence.

4.2 Dark matter contraction

When gas inside a dark matter halo cools and condensates at the center as stars begin to form, the dark matter distribution also con-

tracts in response. If the gas infall is slow and smooth, then the dark matter contraction is adiabatic (Blumenthal et al. 1986). Inversely, dark matter halo can also adiabatically expand, if there is gas outflow.

We infer the contraction in the dark matter distribution in our lens galaxies by combining the lensing observables from our models with their measured stellar kinematics. For the dark matter distribution, we allow adiabatic contraction or expansion from the 'pristine' NFW distribution. We follow the contraction formalism of Dutton et al. (2007), which is originally based on the formalism of Blumenthal et al. (1986). In this formalism, the initial radius $r_{\rm i}$
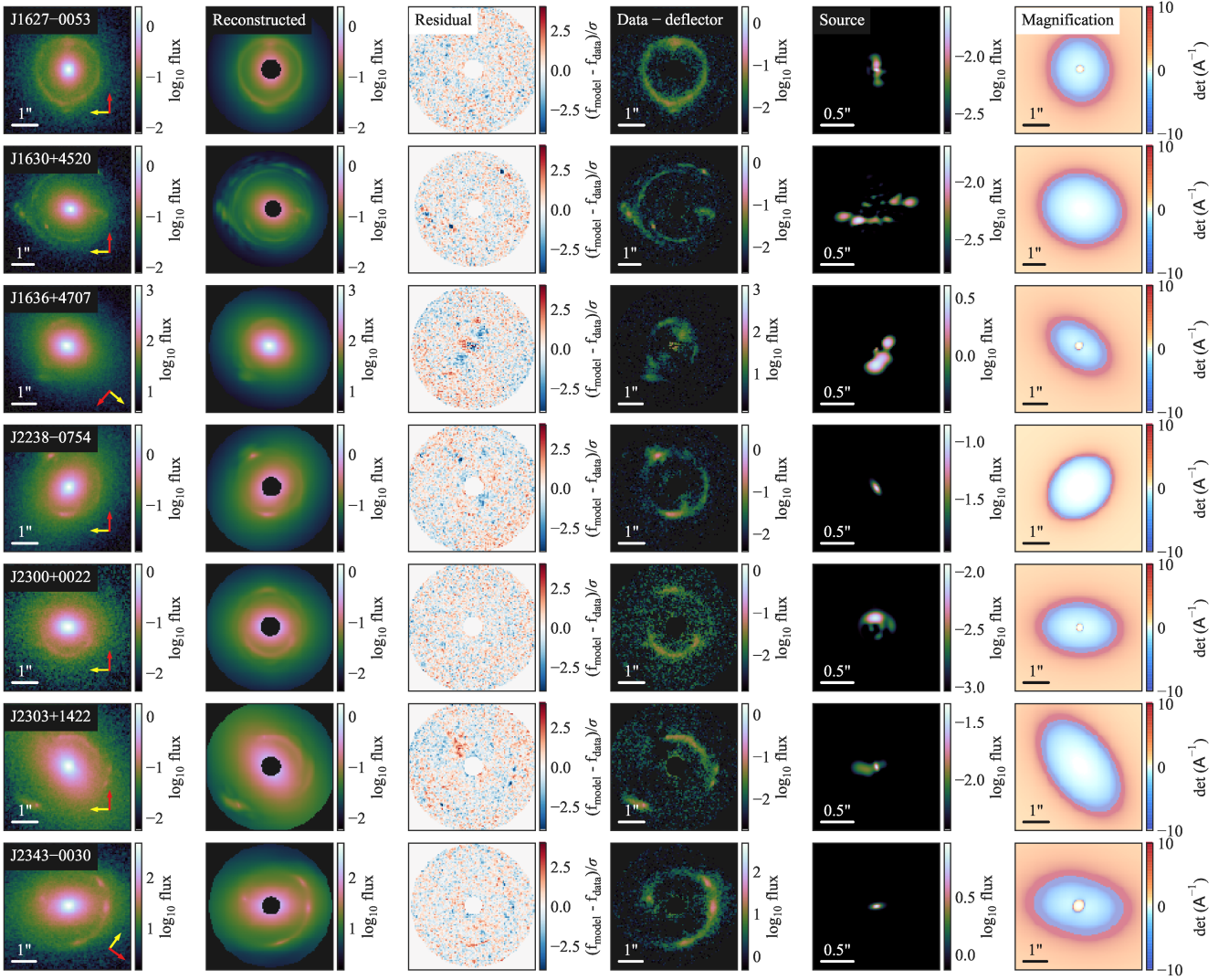
Figure 3. Lens models for the last seven out of the 23 SLACS lenses continuing after Figure 2. The first column shows the *HST* images in F555W or F606W band. The second column shows the reconstructed image from the best-fit lens model and the third column shows the normalized residual for the reconstruction. In the fourth column, we show subtract the modelled deflector light distribution from the observed image to visualized the lensed arcs. The fifth column shows the reconstructed source and the sixth column shows the magnification map of the lens model. The red arrow points to the North and the yellow arrow points to the East. We mask the central 0.4 arcsecond of some deflectors where there would otherwise be large residuals in the deflector's light profile fit without the mask.

of a dark matter particle and its final radius r_f after contraction is related as

$$r_\mathrm{i} M_\mathrm{i}(r_\mathrm{i}) = r_\mathrm{f} M_\mathrm{f}(r_\mathrm{f}), \qquad (4)$$

where $M(r)$ is the total 3D mass within radius r. Then, we have

$$
\begin{aligned}
M_\mathrm{f}(r_\mathrm{f}) &= M_\mathrm{gal,f}(r_\mathrm{f}) + M_\mathrm{dm,f}(r_\mathrm{f}) \\
&= M_\mathrm{gal,f}(r_\mathrm{f}) + (1 - f_\mathrm{gal}) M_\mathrm{i}(r_\mathrm{i}),
\end{aligned}
\qquad (5)
$$

where $f_\mathrm{gal} \equiv M_\mathrm{gal,i}/M_\mathrm{i} = M_\mathrm{gal,i}/(M_\mathrm{dm,i} + M_\mathrm{gal,i})$ is the fraction of the total mass that cools to form stars. We iteratively solve equations (4) and (5) to obtain the contraction factor $\Gamma(r_\mathrm{f}) \equiv r_\mathrm{f}/r_\mathrm{i}$. Following Dutton et al. (2007), we furthermore adopt the halo response parameter ν to modify the relation between r_i and r_f as

$$r_\mathrm{i} \equiv \Gamma^{-\nu}(r_\mathrm{f})\, r_\mathrm{f}. \qquad (6)$$

By varying ν in our models we can adjust the amount of contraction. For example, $\nu = 0$ corresponds to no contraction, $\nu = 1$

corresponds to full contraction as in Blumenthal et al. (1986), and $\nu < 0$ corresponds to the case of adiabatic expansion. Cosmological hydrodynamical simulations have found less contraction than the model of Blumenthal et al. (1986), e.g., the result of Gnedin et al. (2004) corresponds to $\nu = 0.8$ and the result of Abadi et al. (2010) corresponds to $\nu \sim 0.4$ (Dutton & Treu 2014). However, these simulations almost always find contracted dark matter halos to some extent at the masses of interest (e.g., Schaller et al. 2015; Xu et al. 2017; Peirani et al. 2017).

We adopt a composite mass distribution consisting of the contracted dark matter distribution and the stellar distribution to model the stellar kinematics of the lens galaxies in our sample. For computational simplicity, we use spherical cases of these profiles. This assumption is sufficient for the quality of the measured velocity dispersions (Sonnenfeld et al. 2012). The line-of-sight velocity dispersion can be expressed by solving the spherical Jeans equation as
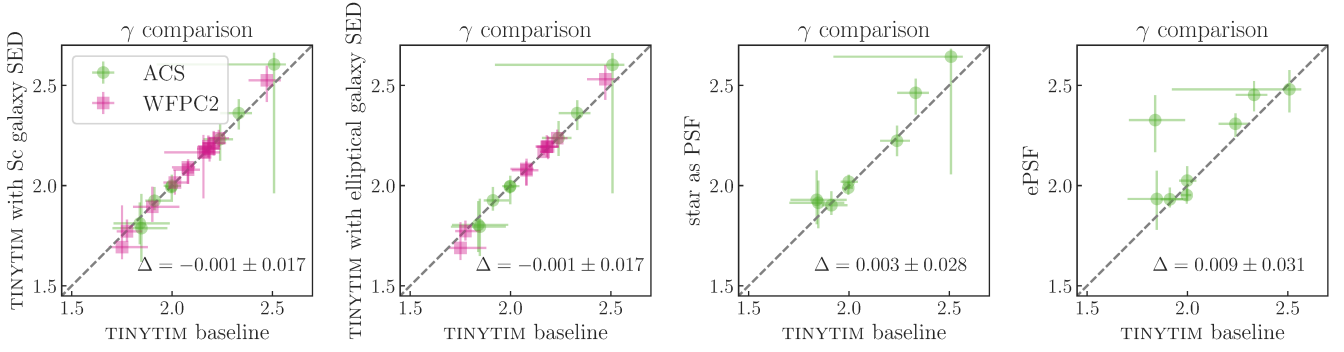
Figure 4. Comparison of measured logarithmic slopes for various choices of PSF. The three choices are: (i) TINYTIM PSF with G2V star SED (baseline), (ii) TINYTIM PSF with Sc galaxy SED at $\langle z_s \rangle = 0.62$, (iii) TINYTIM PSF with elliptical galaxy SED at $\langle z_l \rangle = 0.2$, (iv) centered and stacked stars from the corresponding *HST* image, and (v) ePSF created from the stars in the corresponding *HST* image (Anderson & King 2000). The systems with ACS F555W imaging are shown with green circles and the systems with WFPC2 F606W imaging are shown with pink squares. Reliable PSFs for PSF types (iv) and (v) could be extracted only for the ACS images in our sample. The mean difference Δ between two corresponding PSF types is annotated within each panel. In all cases, the deviation of the sample mean is negligible relative to the uncertainty of individual logarithmic slopes (typically $\gtrsim 0.04$).

Table 1. Point estimates of the parameters in our power-law lens models. Here, R_E is the Einstein radius, $\gamma^{lensing}$ is the logarithmic slope of the mass profile, q_m is the mass axis ratio, PA_m is the mass position angle, γ_{ext} is the external shear magnitude, ϕ_{ext} is the external shear angle, R_{eff} is the effective or half-light radius in V-band, q_L is the light axis ratio, and PA_L is the light position angle.

Name	R_E (arcsec)	$\gamma^{lensing}$	q_m	PA_m (N of E) (deg)	γ_{ext}	ϕ_{ext} (N of E) (deg)	R_{eff} (arcsec)	q_L	PA_L (N of E) (deg)
J0029−0055	$0.951^{+0.003}_{-0.003}$	$2.47^{+0.07}_{-0.09}$	$0.69^{+0.04}_{-0.04}$	66^{+2}_{-2}	$0.002^{+0.002}_{-0.001}$	-22^{+79}_{-63}	2.42 ± 0.05	$0.837^{+0.002}_{-0.002}$	$62.5^{+0.4}_{-0.5}$
J0037−0942	$1.509^{+0.009}_{-0.029}$	$2.18^{+0.05}_{-0.03}$	$0.65^{+0.02}_{-0.03}$	82^{+3}_{-3}	$0.013^{+0.010}_{-0.009}$	21^{+10}_{-13}	2.80 ± 0.06	$0.701^{+0.001}_{-0.005}$	$82.9^{+0.1}_{-0.1}$
J0252+0039	$1.011^{+0.005}_{-0.002}$	$2.16^{+0.08}_{-0.20}$	$0.89^{+0.01}_{-0.03}$	-10^{+8}_{-9}	$0.009^{+0.017}_{-0.006}$	-33^{+16}_{-23}	1.27 ± 0.03	$0.931^{+0.004}_{-0.005}$	$-24.6^{+2.1}_{-2.5}$
J0330−0020	$1.077^{+0.005}_{-0.005}$	$2.16^{+0.02}_{-0.02}$	$0.69^{+0.04}_{-0.04}$	-7^{+3}_{-3}	$0.025^{+0.007}_{-0.008}$	26^{+9}_{-11}	1.45 ± 0.03	$0.738^{+0.004}_{-0.004}$	$-18.7^{+0.4}_{-0.4}$
J0728+3835	$1.283^{+0.004}_{-0.005}$	$2.23^{+0.05}_{-0.05}$	$0.59^{+0.03}_{-0.03}$	24^{+1}_{-1}	$0.061^{+0.007}_{-0.007}$	16^{+3}_{-3}	1.79 ± 0.04	$0.754^{+0.002}_{-0.002}$	$22.6^{+0.3}_{-0.3}$
J0737+3216	$0.973^{+0.015}_{-0.004}$	$2.51^{+0.06}_{-0.58}$	$0.68^{+0.04}_{-0.04}$	-5^{+2}_{-4}	$0.117^{+0.009}_{-0.106}$	78^{+1}_{-31}	3.49 ± 0.07	$0.858^{+0.007}_{-0.006}$	$-11.4^{+1.2}_{-1.4}$
J0903+4116	$1.271^{+0.014}_{-0.007}$	$2.08^{+0.06}_{-0.07}$	$0.72^{+0.12}_{-0.02}$	65^{+3}_{-117}	$0.089^{+0.009}_{-0.088}$	50^{+3}_{-51}	3.53 ± 0.07	$0.855^{+0.006}_{-0.008}$	$-86.8^{+13.5}_{-1.2}$
J0959+0410	$0.962^{+0.042}_{-0.017}$	$1.98^{+0.03}_{-0.05}$	$0.71^{+0.09}_{-0.02}$	13^{+26}_{-2}	$0.058^{+0.010}_{-0.057}$	-11^{+11}_{-3}	0.99 ± 0.02	$0.668^{+0.033}_{-0.042}$	$31.3^{+4.5}_{-4.2}$
J1112+0826	$1.421^{+0.017}_{-0.012}$	$2.21^{+0.06}_{-0.06}$	$0.56^{+0.06}_{-0.03}$	-57^{+2}_{-1}	$0.066^{+0.017}_{-0.037}$	-63^{+173}_{-7}	1.80 ± 0.04	$0.746^{+0.001}_{-0.001}$	$-46.5^{+0.2}_{-0.2}$
J1204+0358	$1.285^{+0.005}_{-0.004}$	$2.19^{+0.05}_{-0.06}$	$0.82^{+0.03}_{-0.03}$	-4^{+3}_{-3}	$0.056^{+0.005}_{-0.006}$	-41^{+5}_{-4}	1.59 ± 0.03	$0.987^{+0.003}_{-0.003}$	$-48.7^{+5.2}_{-4.8}$
J1250+0523	$1.115^{+0.005}_{-0.003}$	$1.91^{+0.08}_{-0.04}$	$0.91^{+0.02}_{-0.04}$	-60^{+23}_{-4}	$0.020^{+0.007}_{-0.019}$	-73^{+74}_{-6}	1.60 ± 0.03	$0.893^{+0.006}_{-0.016}$	$-5.2^{+1.3}_{-1.4}$
J1306+0600	$1.299^{+0.011}_{-0.013}$	$2.18^{+0.06}_{-0.05}$	$0.66^{+0.10}_{-0.07}$	-39^{+3}_{-3}	$0.006^{+0.008}_{-0.004}$	11^{+58}_{-105}	2.30 ± 0.05	$0.909^{+0.003}_{-0.003}$	$-50.2^{+0.8}_{-0.9}$
J1313+4615	$1.302^{+0.011}_{-0.014}$	$2.08^{+0.04}_{-0.08}$	$0.64^{+0.19}_{-0.03}$	-13^{+35}_{-3}	$0.098^{+0.011}_{-0.097}$	-20^{+20}_{-3}	2.15 ± 0.04	$0.806^{+0.004}_{-0.048}$	$30.2^{+2.2}_{-0.6}$
J1402+6321	$1.354^{+0.003}_{-0.003}$	$2.24^{+0.06}_{-0.08}$	$0.70^{+0.04}_{-0.04}$	24^{+2}_{-2}	$0.010^{+0.006}_{-0.005}$	1^{+11}_{-19}	3.02 ± 0.06	$0.758^{+0.002}_{-0.002}$	$17.9^{+0.3}_{-0.3}$
J1531−0105	$1.705^{+0.006}_{-0.011}$	$1.90^{+0.13}_{-0.10}$	$0.68^{+0.02}_{-0.03}$	-61^{+4}_{-3}	$0.046^{+0.012}_{-0.031}$	-69^{+169}_{-8}	3.43 ± 0.07	$0.672^{+0.001}_{-0.001}$	$-55.1^{+0.1}_{-0.1}$
J1621+3931	$1.263^{+0.004}_{-0.005}$	$2.01^{+0.06}_{-0.06}$	$0.77^{+0.02}_{-0.03}$	-50^{+2}_{-3}	$0.007^{+0.007}_{-0.004}$	41^{+38}_{-150}	2.44 ± 0.05	$0.739^{+0.002}_{-0.002}$	$-53.2^{+0.3}_{-0.3}$
J1627−0053	$1.227^{+0.002}_{-0.002}$	$1.84^{+0.15}_{-0.13}$	$0.87^{+0.01}_{-0.01}$	80^{+2}_{-2}	$0.026^{+0.007}_{-0.007}$	-96^{+178}_{-5}	2.76 ± 0.06	$0.836^{+0.004}_{-0.004}$	$85.9^{+0.7}_{-0.6}$
J1630+4520	$1.788^{+0.005}_{-0.006}$	$2.00^{+0.02}_{-0.03}$	$0.83^{+0.02}_{-0.01}$	20^{+2}_{-4}	$0.019^{+0.004}_{-0.018}$	26^{+5}_{-26}	2.11 ± 0.04	$0.825^{+0.003}_{-0.044}$	$19.7^{+0.8}_{-0.6}$
J1636+4707	$1.100^{+0.006}_{-0.007}$	$1.78^{+0.07}_{-0.05}$	$0.66^{+0.03}_{-0.04}$	-4^{+2}_{-3}	$0.055^{+0.014}_{-0.017}$	-3^{+6}_{-7}	1.79 ± 0.04	$0.797^{+0.002}_{-0.001}$	$-14.2^{+0.3}_{-0.3}$
J2238−0754	$1.272^{+0.002}_{-0.003}$	$2.33^{+0.07}_{-0.09}$	$0.70^{+0.03}_{-0.03}$	-44^{+2}_{-1}	$0.003^{+0.003}_{-0.002}$	-19^{+39}_{-39}	2.53 ± 0.05	$0.784^{+0.002}_{-0.002}$	$-50.7^{+0.3}_{-0.3}$
J2300+0022	$1.243^{+0.008}_{-0.006}$	$1.85^{+0.13}_{-0.14}$	$0.69^{+0.04}_{-0.04}$	5^{+2}_{-2}	$0.025^{+0.020}_{-0.018}$	17^{+25}_{-12}	1.86 ± 0.04	$0.790^{+0.003}_{-0.003}$	$4.0^{+0.5}_{-0.5}$
J2303+1422	$1.611^{+0.010}_{-0.005}$	$2.00^{+0.05}_{-0.04}$	$0.59^{+0.03}_{-0.03}$	54^{+2}_{-1}	$0.002^{+0.003}_{-0.002}$	1^{+18}_{-17}	3.48 ± 0.07	$0.659^{+0.002}_{-0.002}$	$52.5^{+0.2}_{-0.2}$
J2343−0030	$1.529^{+0.004}_{-0.005}$	$1.75^{+0.13}_{-0.07}$	$0.65^{+0.02}_{-0.03}$	54^{+1}_{-1}	$0.088^{+0.005}_{-0.007}$	34^{+3}_{-8}	2.46 ± 0.05	$0.662^{+0.003}_{-0.005}$	$51.1^{+0.2}_{-0.2}$

where G is the gravitational constant, $I(R)$ is surface brightness distribution, $M(R)$ is the 3D enclosed mass, $l(r)$ is the 3D luminosity density, and $\mathcal{K}_\beta$ is a function that depends on the parameterization of the anisotropy parameter β (Mamon & Łokas 2005). We adopt

$$\sigma^2_{los}(R) = \frac{2G}{I(R)} \int_R^\infty \mathcal{K}_\beta \left(\frac{r}{R} \right) \frac{l(r)M(r)}{r} \, dr, \qquad (7)$$

the Osipkov–Merritt anisotropy profile given by

$$\beta(r) \equiv 1 - \frac{\sigma^2_t}{\sigma^2_r} = \frac{r^2}{r^2 + a^2_{ani}R^2_{eff}}, \qquad (8)$$

where σ^2_r is the radial velocity dispersion, σ^2_t is the tangential velocity dispersion, and a_{ani} is the anisotropy scaling factor (Osipkov 1979; Merritt 1985a,b). For this parameterization, the function $\mathcal{K}_\beta$
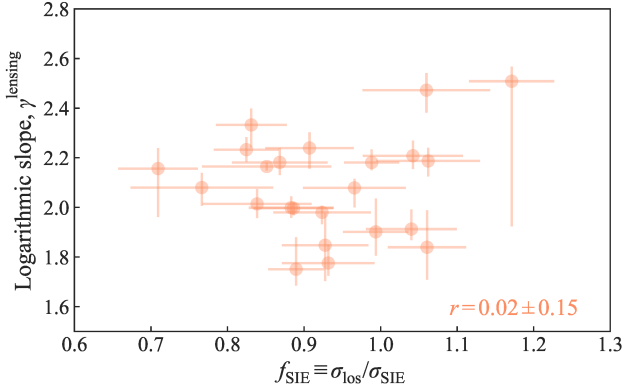
Figure 5. Distribution of $\gamma^{\mathrm{lensing}}$ and $f_{\mathrm{SIE}} \equiv \sigma_{\mathrm{los}}/\sigma_{\mathrm{SIE}}$ for the SLACS sample. The biweight midcorrelation r between these two parameters is annotated in the figure. The measured logarithmic slopes have no correlation with the f_{SIE}. If the global mass distribution follows the power law, then γ should positively correlate with f_{SIE} [see equation (2.3) of Koopmans (2004), cf. Figure 4 of Auger et al. (2010b)]. However, we find that the uncertainty level in the measured γ and f_{SIE} are too large for us to detect a potential correlation. As a result, this comparison is inconclusive, and it cannot confirm or rule out whether the global mass distribution is well approximated by a power law. A more sophisticated analysis, such as that presented later in this paper, or better data are required for this investigation.

has the form

$$
\mathcal{K}_\beta(u) = \frac{u_{\mathrm{ani}}^2 + 1/2}{(u_{\mathrm{ani}} + 1)^{3/2}} \left(\frac{u^2 + u_{\mathrm{ani}}^2}{u} \right) \tan^{-1} \left(\sqrt{\frac{u^2 - 1}{u_{\mathrm{ani}}^2 + 1}} \right) \\
- \frac{1/2}{u_{\mathrm{ani}}^2 + 1} \sqrt{1 - \frac{1}{u^2}}, \tag{9}
$$

where $u_{\mathrm{ani}} \equiv a_{\mathrm{ani}} R_{\mathrm{eff}}/R$ (Mamon & Łokas 2005).

We adopt the NFW profile for the initial distribution of the dark matter, which is given by

$$
\rho_{\mathrm{NFW}}(r) \equiv \frac{\rho_s^{\mathrm{NFW}}}{\left(r/r_s^{\mathrm{NFW}} \right) \left(1 + r/r_s^{\mathrm{NFW}} \right)^2}, \tag{10}
$$

where ρ_s^{NFW} is the density normalization and r_s^{NFW} is the scale radius. For the stellar mass distribution we first adopt a constant mass-to-light ratio (M/L) in Section 4.2.1, and then a M/L gradient in Section 4.2.2. We adopt a double Sérsic profile for the stellar light distribution similar to our lens models. To ensure robustness of the light profile fits, we fit the double Sérsic profile from 20×20 arcsec^2 cutouts around the lens galaxies, which sufficiently contain more than twice the effective radii of these galaxies. Before fitting the light profiles, we subtract the lensed arcs using our best fit lens models. We only obtain the best fit double Sérsic profile and adopt an uncertainty on the fit equivalent to 2 per cent uncertainty for the effective radius. This 2 per cent uncertainty is a conservative estimate, as the 2 per cent corresponds to 95 percentile of the uncertainty distribution of the effective radii in our lens models. We use the concentric Gaussian decomposition method from Shajib (2019) to deproject the 2D stellar light or mass distribution into the corresponding 3D distribution. We take 30 Gaussian components with their standard deviations logarithmically spaced between 0.001 arcsec and 30 arcsec. This decomposition approximates the 2D stellar mass profile shapes within ∼0.2 per cent accuracy between 0.005 arcsec and 6 arcsec.

4.2.1 Constant M/L for stellar mass distribution

We first constrain the model parameters $w \equiv \{\nu,\ a_{\mathrm{ani}},\ f_{\mathrm{gal}},\ r_s^{\mathrm{NFW}},\ \rho_s^{\mathrm{NFW}}\}$ for individual lens systems. Here, r_{NFW} is the scale radius of the NFW profile and $\rho_{0,\mathrm{NFW}}$ is the normalization. We impose a theoretical prior on the $M_{200}-c_{200}$ relation corresponding to $\langle z_{\mathrm{d}} \rangle = 0.20$ from Diemer & Joyce (2019) for the initial NFW halo. To incorporate the lensing constraints into this joint lensing–dynamics analysis, we fold in the posterior distributions of Einstein radius R_{E} and the quantity $R_{\mathrm{E}}\alpha_{\mathrm{E}}''/(1 - \kappa_{\mathrm{E}})$ from our lens models in Section 4.1. Here, α_{E}'' is the double derivative of the deflection angle at R_{E} and the κ_{E} is the convergence at R_{E}. The term $R_{\mathrm{E}}\alpha_{\mathrm{E}}''/(1 - \kappa_{\mathrm{E}})$ is the mass-model-independent observable quantity from imaging data (Kochanek 2020). We also use the weak lensing measurements of the reduced shear, although weak lensing do not provide tight constraints due to large uncertainty in the measured shear of individual systems. We take uniform priors on $\nu \sim \mathcal{U}(-0.5,\ 1)$, $a_{\mathrm{ani}} \sim \mathcal{U}(0.1,\ 5)$, and $f_{\mathrm{gal}} \sim \mathcal{U}(0,\ 0.25)$. Additionally we impose a prior on $\log_{10} M_{200}$ from Sonnenfeld et al. (2018) that depends on the measured stellar mass from the stellar population synthesis method assuming Chabrier IMF. We obtain the Chabrier IMF based stellar masses for the galaxies in our sample from Auger et al. (2009). The form of the prior is

$$
p(\log_{10} M_{200} \mid \log_{10} M_\star^{\mathrm{Chab}}) \\
\equiv \mathcal{N} \left(\mu_{\mathrm{h}} + \beta_{\mathrm{h}} \left[\log_{10} M_\star^{\mathrm{Chab}} - 11.3 \right],\ \sigma_{\mathrm{h}} \right). \tag{11}
$$

The parameters in this prior are obtained by Sonnenfeld et al. (2018) by applying the SLACS selection function on SDSS galaxies that have HSC weak lensing data and assuming a model with M/L gradient for the stellar mass and an NFW profile for the dark matter. We take the uncertainty on these parameters to match with the dispersion observed in the SLACS galaxies by Sonnenfeld et al. (2018) for the same mass model to have $\mu_{\mathrm{h}} = 13.03 \pm 0.25$, $\beta_{\mathrm{h}} = 1.42 \pm 0.92$, and $\sigma_{\mathrm{h}} = 0.18 \pm 0.16$.

We then infer the distribution of population-level parameters τ for our lens sample using hierarchical Bayesian inference. According to the Bayes' theorem, we can express the posterior of the population-level parameters as

$$
p(\tau \mid D) \propto p(D \mid \tau)\, p(\tau) \\
\propto p(\tau) \prod_i^N \int \mathrm{d}w_i\, p(D_i \mid w_i)\, p(w_i \mid \tau) \\
\propto p(\tau) \prod_i^N \int \mathrm{d}w_i\, \frac{p(w_i \mid D_i)}{p(w_i)}\, p(w_i \mid \tau), \tag{12}
$$

Here, $D \equiv \{D_i\}$ refers to the data set containing all the data D_i for individual lenses, $p(D \mid \tau)$ is the likelihood term on the population-level parameters τ, $p(\tau)$ is the prior, $p(w_i \mid \tau)$ governs the distribution of individual lens parameters given the population level parameters, $p(w_i \mid D_i)$ is the posterior of w_i for an individual system given its associated data D_i, and $p(w)$ is the prior for the parameter set w as given above. We assume that a_{ani}, ν, and f_{gal} is Gaussian-distributed for our lens sample with the Gaussians truncated where $p(w) = 0$. Thus, we take $\tau \equiv \{\mu_\nu,\ \sigma_\nu,\ \mu_{a_{\mathrm{ani}}},\ \sigma_{a_{\mathrm{ani}}},\ \mu_{f_{\mathrm{gal}}},\ \sigma_{f_{\mathrm{gal}}}\}$ as the population level parameters. Here, μ and σ refer to the mean and standard deviation of the population level Gaussian distribution for the corresponding parameter. We take uniform priors $\mu_\nu \sim \mathcal{U}(-1,1)$, $\mu_{a_{\mathrm{ani}}} \sim \mathcal{U}(0.5,5)$, and $\mu_{f_{\mathrm{gal}}} \sim \mathcal{U}(0,0.25)$. We take the same uniform prior $\mathcal{U}(-5,1)$ for $\log_{10} \sigma_\nu$, $\log_{10} \sigma_{a_{\mathrm{ani}}}$, and $\log_{10} \sigma_{f_{\mathrm{gal}}}$. We perform the integral in equation (12) using Monte Carlo integration. In the Monte Carlo integration, we sample from $p(w_i \mid \tau)$

and sum over $p(w_i \mid D_i)/p(w_i)$. We approximate $p(w_i \mid D_i)$ with a Gaussian mixture model (GMM) of the corresponding samples drawn using MCMC. Before fitting the GMM, we smooth the distributions of the MCMC samples using the kernel density estimation (KDE) method with Silverman's rule for the bandwidth (Silverman 1986). The number of components in the GMM for each individual lens posterior $p(w_i \mid D_i)$ is selected using Bayesian information criterion (BIC) and it is typically within 5–10.

We illustrate the posterior distributions of the population-level parameters in Figure 6, and tabulate their point estimates in Table 2. In Figure 7, we highlight some of the important results from Figure 6 along with their interpretations for galaxy properties. We infer the mean halo response parameter $\mu_\nu = -0.06 \pm 0.04$ and intrinsic scatter $\sigma_\nu \leq 0.092$ (95 per cent upper limit), which is consistent within 1.5σ with no contraction on average for our lens sample.

We compute the dark matter fraction $f_{\rm dm}$ within $R_{\rm eff}/2$ directly from the decomposed dark and stellar mass distributions. We show the $\gamma - f_{\rm dm}$ distribution in Figure 8. We find moderate anti-correlation between γ and $f_{\rm dm}$ with the biweight midcorrelation $r = -0.39 \pm 0.15$. The anti-correlation is expected as a shallower total mass profile slope will require higher contribution from the dark matter to lower the total mass profile's slope. The distribution of $f_{\rm dm}$ in our sample has a mean of $\langle f_{\rm dm} \rangle = 0.27 \pm 0.02$ and a scatter of 0.14 ± 0.01.

4.2.2 *M/L gradient for stellar mass distribution*

Next, we incorporate a M/L gradient for the stellar mass distribution. We parameterize a power-law M/L gradient as

$$\Upsilon(R) = \Upsilon_{\rm e} \left(\frac{R}{R_{\rm eff}} \right)^{-\eta}, \tag{13}$$

where $\Upsilon_{\rm e}$ is the M/L at $R_{\rm eff}$. However, this parameterization diverges for $\eta > 0$ as $R \to 0$, the total stellar mass computed using the concentric Gaussian decomposition method stays finite. Moreover, the observed lensing properties and the kinematic properties are mostly sensitive to the mass distribution near the Einstein radius and the effective radius, respectively. As a result, the deviation from the exact relation near the center ($\lesssim 0.01$ arcsec) in our Gaussian-decomposed approximation does not noticeably impact our analysis.

We perform the same analysis from Section 4.2.1, but with the M/L gradient implemented in the stellar mass distribution. As a result, the model parameters for individual lenses are extended to $w \equiv \{\nu, a_{\rm ani}, f_{\rm gal}, r_{\rm s}^{\rm NFW}, \rho_{\rm s}^{\rm NFW}, \eta\}$. We only allow positive values for η as previous observations suggest that the M/L is higher at the center than outer region of elliptical galaxies (e.g., Martín-Navarro et al. 2015; van Dokkum et al. 2017). As we want to adopt an uninformative prior on the order of magnitude of η, we take a Jeffrey's prior on η as $p(\eta) \propto 1/\eta$, which is equivalent to a uniform prior on $\log \eta$. We set the bounds of the uniform prior on $\log \eta$ as $p(\log \eta) \sim \mathcal{U}(\log 10^{-3}, 0)$. Furthermore, the population level parameters for the Bayesian hierarchical inference are also extended to $\tau \equiv \{\mu_\nu, \sigma_\nu, \mu_{a_{\rm ani}}, \sigma_{a_{\rm ani}}, \mu_{f_{\rm gal}}, \sigma_{f_{\rm gal}}, \mu_{\log \eta}, \sigma_{\log \eta}\}$.

We infer the 95 per cent upper limit of $\exp(\mu_{\log \eta})$ distribution to be 0.02. The mean halo response parameter for the sample is $\mu_\nu = -0.03^{+0.04}_{-0.05}$ with intrinsic scatter $\sigma_\nu \leq 0.074$ (95 per cent upper limit), which is consistent with no contraction in the NFW halo (Figure 6). We show the distribution of the inferred $f_{\rm dm}$ and the logarithmic slope γ in Figure 8. The distribution of $f_{\rm dm}$ in our sample for the M/L gradient model has a mean $\langle f_{\rm dm} \rangle = 0.28 \pm 0.02$ and an intrinsic scatter of 0.15 ± 0.02.

To check for the impact of the $M–c$ relation prior in this analysis, we perform the joint lensing–dynamics analysis without the $M–c$ relation prior. We find that the uncertainties on the model parameters for each lens system expectedly increase without this prior, and the posteriors from with or without the prior are consistent within 1σ (Figure 6, Table 2).

We also check for the impact of our fiducial cosmology, in particular the Hubble constant, on our inference. We perform the joint lensing–dynamics analysis with $H_0 = 67$ km s^{-1} Mpc^{-1} and with $H_0 = 74$ km s^{-1} Mpc^{-1}. We find all the posterior distributions to be consistent within 1σ. Therefore, our inference is not sensitive to the choice of the Hubble constant.

We show the surface density profiles and the deprojected 3D density profiles of the dark matter and stellar components of the total mass distribution for all the lens galaxies in Figures 9 and 10, respectively. In Figure 11, we illustrate the median deviation of the NFW+stars model from the power law in the total surface density profile and the local logarithmic slope. The total density profile combining dark matter and stars deviates higher by ~5 per cent (median value) at the Einstein radius with a median absolute deviation of ~7 per cent within our sample. Here, the median absolute deviation is a proxy for the intrinsic scatter. The local slope of the density profile at the Einstein radius is shallower by ~2 per cent (median value) with median absolute deviation of ~2 per cent. These deviations are sensitive to our adopted prior on M_{200}. If we choose a prior with lower mean halo mass such as $\log_{10} M_{200} \sim \mathcal{N}(12.9, 0.5)$ – which corresponds to the BOSS constant-mass (CMASS) galaxies with mean redshift $z \approx 0.6$ (Sonnenfeld et al. 2019a) – the resultant profiles would be within 0.8 per cent of the power laws at the Einstein radius (Figure 11).

5 DISCUSSION AND COMPARISON WITH PREVIOUS STUDIES

In this section, we compare our results with previous studies based on observations and simulations, and interpret them in the context of elliptical galaxy evolution. First in Section 5.1, we discuss the central results of this study on the elliptical galaxy structure. We then place these results in the context of massive elliptical galaxy evolution in Section 5.2. We discuss our results on the M/L-gradient and its implication for the stellar IMF in Section 5.3. We discuss the implication of our results for time-delay cosmography in Section 5.4. We state the limitations of this study in Section 5.5.

5.1 Elliptical galaxy structure

The main results of this study on the elliptical galaxy structure are: (i) the dark matter distribution in massive elliptical galaxies at $z \sim 0.2$ are not contracted and close to an NFW profile on average (i.e., $\mu_\nu \sim 0$), and (ii) the total density profile combining the dark matter and stars is close to isothermal (i.e., $\langle \gamma \rangle \sim 2$).

5.1.1 *Dark matter contraction*

Our result on the dark matter contraction agrees with Dutton & Treu (2014), who also find no contraction in SLACS galaxies using the lens models from Auger et al. (2010b). Similarly, Sonnenfeld et al. (2015) and Newman et al. (2015) find that the inner slope of the dark matter distribution is consistent with the NFW profile within the uncertainty for their samples of massive elliptical galaxy halos and group-scale halos, respectively. Therefore, these findings
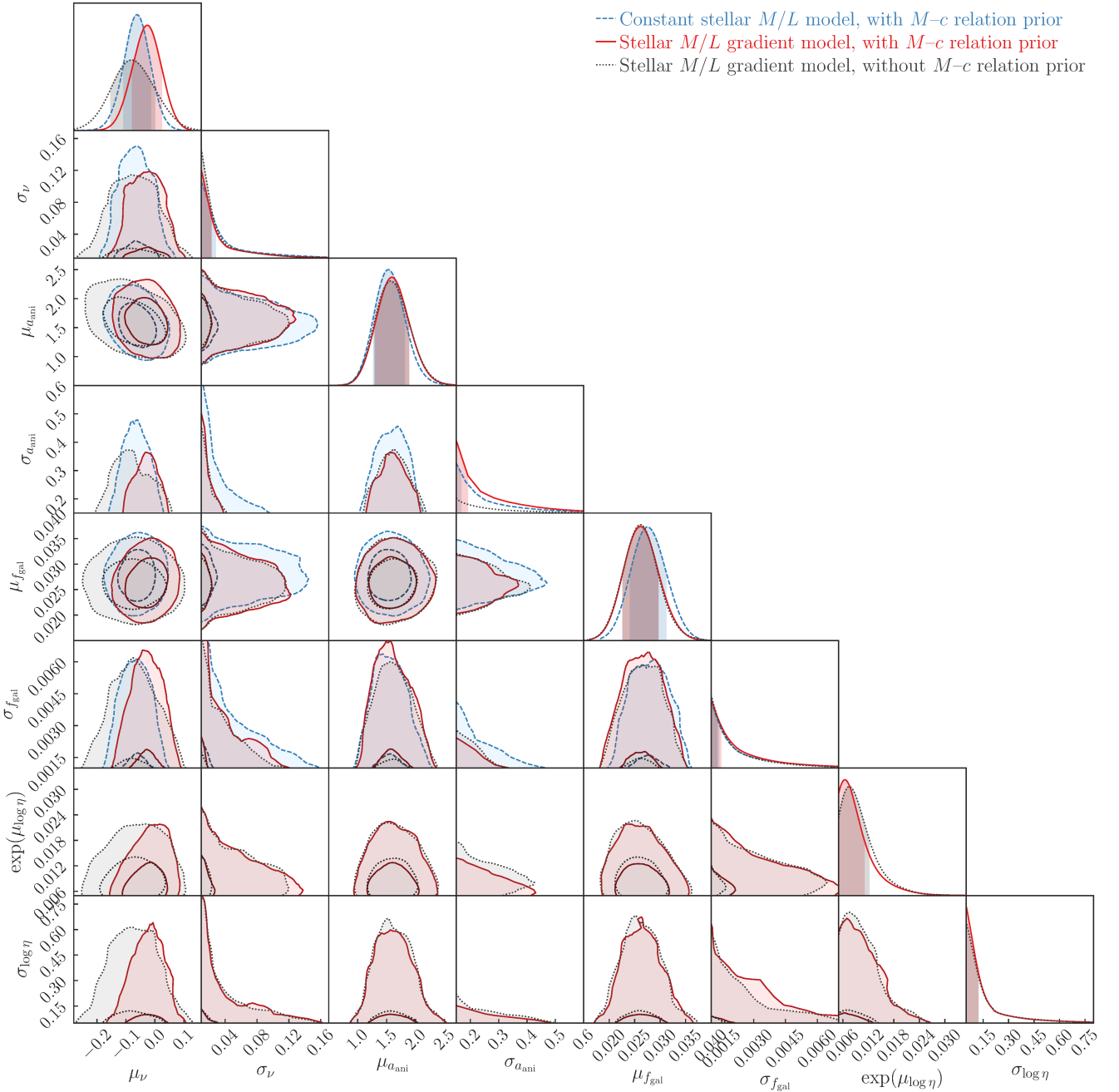
Figure 6. Distribution of population-level parameters assuming Gaussian distribution for the halo response parameter ν, the anisotropy scaling parameter $a_{\rm ani}$, fraction of the total mass that cools to form stars $f_{\rm gal}$, and $\log \eta$ where η is the M/L gradient's exponent. The three model setups illustrated are a constant stellar M/L model with the $M{-}c$ relation prior (blue dashed line), a stellar M/L gradient model with the $M{-}c$ relation prior (red solid line), and a stellar M/L gradient model without the $M{-}c$ relation prior (dotted grey line). The darker and lighter regions represent 68 per cent and 95 per cent credible regions, respectively. The 2D contours and 1D distributions are smoothed with a Gaussian kernel of 1.5σ bandwidth. The point-estimate of these parameters are given in Table 2. Figure 7 highlights the central parameters in our model along with their interpretations.

are consistent with our result. However, our result contradicts the report of steeper central slopes (mean inner logarithmic slope $\langle \beta \rangle = 2.01^{+0.19}_{-0.22}$) than that in the NFW profile ($\beta = 1$) by Oldham & Auger (2018). Interestingly, Oldham & Auger (2018) find their slope distribution to be bimodal, and the steep value given above corresponds to the mode with the larger mean. If Oldham & Auger (2018) adopt a unimodal slope distribution, then the average inner slope is consistent with the vanilla NFW profile within 0.5σ for the

M/L gradient model and within 1.3σ for the constant M/L model. Albeit, the unimodal distribution requires higher intrinsic scatter. Moreover, Oldham & Auger (2018) note that their inferred halo mass is lower than the expectation of abundance matching studies due to the lack of a strong prior to constrain the halo scale size or its mass. Strong-lensing data is sensitive to the mass profile in the central region within the Einstein radius, which is typically $\sim$5–10 times smaller than the NFW scale radius. Thus, additional data or prior
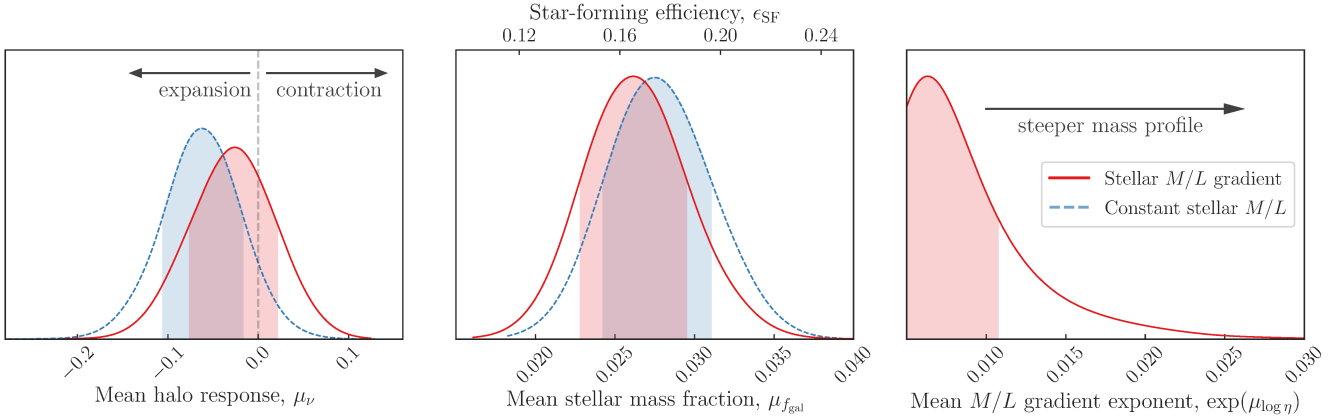
Figure 7. Interpretation of 1D marginalized distributions of some central model parameters for our galaxy sample. **Left-hand panel:** distributions of the sample mean μ_ν of the halo response parameter ν for stellar M/L gradient (red solid line) and constant stellar M/L (blue dashed line) models. Both models are consistent with zero contraction marked by the vertical grey dashed line. The full contraction scenarios in Blumenthal et al. (1986) model ($\nu = 1$), and the simulations of Gnedin et al. (2004, $\nu = 0.8$) and Abadi et al. (2010, $\nu \sim 0.4$) are ruled out. **Middle panel:** distributions of the mean fraction of star-forming baryons $\mu_{f_{gal}}$ that has cooled from the initial distribution. We show along the top axis the corresponding star-forming efficiency $\epsilon_{SF} \equiv f_{gal}\Omega_m/\Omega_b$, where Ω_b/Ω_m is adopted from Planck Collaboration (2018). The $\mu_{f_{gal}}$ distributions peaks at $\epsilon_{SF} \sim 0.17$, which is consistent with moderate to strong feedback mechanisms on the star formation (Hopkins et al. 2014). **Right-hand panel:** The distribution of mean M/L-gradient-exponent in our M/L gradient model. The distribution favors very small gradient, with the 95 per cent upper limit at 0.02.

Table 2. 1D marginalized distributions of the population level parameters from Bayesian hierarchical inference. The corner plot for these parameters are illustrated in Figure 6. The baseline model settings are: $M-c$ relation prior from (Diemer & Joyce 2019). Jeffrey's prior for M/L-gradient exponent η, and $H_0 = 70$ km s^{-1} Mpc^{-1}. "Other settings" column indicates which setting from these baseline settings is varied. The columns for σ_ν, $\sigma_{a_{ani}}$, $\sigma_{f_{gal}}$, $\exp(\mu_{\log \eta})$, and $\sigma_{\log \eta}$ provide the 95 per cent upper limit.

Stellar M/L	Other settings	μ_ν	σ_ν	$\mu_{a_{ani}}$	$\sigma_{a_{ani}}$	$\mu_{f_{gal}}$	$\sigma_{f_{gal}}$	$\exp(\mu_{\log \eta})$	$\sigma_{\log \eta}$
Constant M/L	Baseline	$-0.06^{+0.04}_{-0.04}$	≤ 0.092	$1.54^{+0.25}_{-0.22}$	≤ 0.26	$0.028^{+0.003}_{-0.003}$	≤ 0.0040	–	–
M/L gradient	Baseline	$-0.03^{+0.04}_{-0.05}$	≤ 0.074	$1.60^{+0.27}_{-0.25}$	≤ 0.22	$0.026^{+0.003}_{-0.003}$	≤ 0.0041	≤ 0.017	≤ 0.33
M/L gradient	No $M-c$ prior	$-0.08^{+0.07}_{-0.06}$	≤ 0.069	$1.60^{+0.28}_{-0.25}$	≤ 0.20	$0.026^{+0.003}_{-0.003}$	≤ 0.0036	≤ 0.017	≤ 0.36
M/L gradient	η prior $\sim \mathcal{U}(0, 1)$	$0.02^{+0.04}_{-0.04}$	≤ 0.054	$1.52^{+0.26}_{-0.23}$	≤ 0.25	$0.025^{+0.003}_{-0.003}$	≤ 0.0036	$(\leq 0.061)^a$	$(\leq 0.015)^a$

^{a} For the model with uniform prior on η, the column $\exp(\mu_{\log \eta})$ gives the value for μ_η and the column $\sigma_{\log \eta}$ gives the value for σ_η.

at scales larger than the NFW scale radius is necessary to robustly constrain the halo mass. We find that the degree of contraction depends on the halo mass prior in our analysis with a heavier prior on M_{200} producing shallower inner slopes to fit the joint lensing–kinematics data. The sample of Oldham & Auger (2018) has a similar stellar mass range as our sample and the difference between the mean redshifts of the samples $\Delta\langle z \rangle = 0.15$ does not leave enough room to expand the halos from $\beta \sim 2$ to $\beta \sim 1$ within ~ 1.44 Gyr. Therefore, we conclude that the differences between our result and that from Oldham & Auger (2018) are largely caused by the difference in the adopted priors corresponding to the dark matter halo.

5.1.2 *Slope of the total density profile*

Our lensing-only models provide $\langle \gamma^{lensing} \rangle = 2.08 \pm 0.04$ with a scatter of 0.13 ± 0.02. From the joint lensing–dynamics analysis, we find the total density profile is shallower by approximately 5 per cent, which brings the sample mean of the logarithmic slope at the Einstein radius closer to the isothermal case (Figure 11). This near-isothermality of the total density profile agrees well with a multitude of pervious observations – e.g., based on strong-lensing only or jointly based on lensing and dynamics: Treu & Koopmans

(2004); Gavazzi et al. (2007); Auger et al. (2010b); Ritondale et al. (2019), and based on stellar dynamics: Thomas et al. (2007); Tortora et al. (2014); Bellstedt et al. (2018).

In Figure 12, we compare the distribution of the estimated logarithmic slopes γ constrained from the imaging data only in this study with those estimated by the SLACS analysis from combining stellar kinematics with the imaging data. We find no correlation between the estimated γ distributions from the two analyses with biweight midcorrelation $r = 0.03 \pm 0.17$. However, the SLACS distribution has a mean of $\langle \gamma^{LD} \rangle = 2.078 \pm 0.027$ and intrinsic scatter 0.16 ± 0.02 (Auger et al. 2010b). These values are consistent with our results within 1σ confidence level. For the 21 systems that have measured γ in Auger et al. (2009), we find $\Delta_\gamma^2 \equiv \Sigma_i^{N=21}(\gamma_{i,this study}^{lensing} - \gamma_{i,Auger+(10b)}^{LD})^2/\sigma_{i,total}^2 = 21.76$. The p-value assuming a χ^2-distribution for Δ_γ^2 with 21 degrees of freedom is 0.41. We can see in Figure 9 that the two-component mass profile from lensing–dynamics can deviate from the lensing-only inference of the power-law profile toward either direction. However, the sample mean of the such deviations is smaller than ~ 5 per cent near the Einstein radius, which explains the good agreement for $\langle \gamma \rangle$ between this study and Auger et al. (2010b). Thus, a correlation between the lensing-only local slope and the lensing–dynamics global slope is

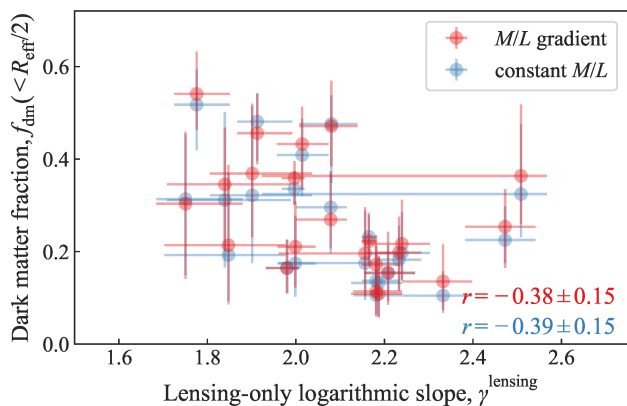**Figure 8.** Distribution of the local logarithmic slope γ and the dark matter fraction $f_{\rm dm}$ within half effective radius. The blue circles are from the constant M/L model for the stellar mass distribution. $f_{\rm dm}$ is moderately anti-correlated with γ, as higher stellar contribution is needed to increase the logarithmic slope, which brings the dark matter fraction $f_{\rm dm}$ down.

expected. However, such correlation cannot be detected within the noise.

In the next subsection, we place these two results in the context of elliptical galaxy formation and evolution.

5.2 Evolution of massive elliptical galaxies

The current paradigm for the formation of elliptical galaxies consists of two phases (e.g., Naab et al. 2007b; Guo & White 2008; Bezanson et al. 2009; Furlong et al. 2015). In the first phase up to $z \approx 2$, gas condensates in a massive halo to form stars, and the dark matter distribution contracts as a result. In the second phase after $z \approx 2$, elliptical galaxies grow in size primarily through multiple major or minor mergers. Most of this later evolution is "dry" involving little gas and star formation, as evidenced by the almost uniformly old stellar populations in elliptical galaxies. If elliptical galaxy halos contract up to $z \approx 2$, our result then raises the question: how does the contracted halo at $z \approx 2$ expands back to resemble the NFW profile at $z \sim 0.2$? Simulations have shown two principle mechanisms that can remove baryons from the central region and expand the dark matter halo – AGN-driven gas outflow, and dynamical heating through accretion of materials from the environment (e.g., Laporte et al. 2012; Martizzi et al. 2012). Although mergers play a crucial rule in the growth and evolution of elliptical galaxies, simulations find that dissipation-less mergers do not change the shape of the dark matter profile (Gnedin et al. 2004; Ma & Boylan-Kolchin 2004). Dissipational gas-rich mergers, in contrast, can make the profile steeper by transferring baryons toward the center and thus further contracting the dark matter halo (Sonnenfeld et al. 2014). The close-to-isothermal nature of the total mass distribution can also be explained as the end result of rearranging the mass distributions through accretion of collision-less materials in gas-poor mergers (Johansson et al. 2009; Remus et al. 2013).

To investigate the driving mechanism behind the halo expansion observed in our sample, we compare the structural properties of our galaxies with those from simulations that adopt varying baryonic physics (Figure 13). The simulation from Oser et al. (2010) includes stellar and supernova feedback, and primordial abundance cooling. In contrast, the Magneticum and IllustrisTNG simulations includes AGN feedback in addition to stellar and supernova feed-

back, and the cooling mechanism includes metals (Remus et al. 2017; Weinberger et al. 2017; Pillepich et al. 2018). For fair comparison with the simulations, we adopt their definition and compute the average logarithmic slope by fitting a power-law to the 3D density profile between $0.4R_{\rm eff}$ and $4R_{\rm eff}$. We also compute the central dark matter fraction within a 3D aperture of radius $R_{\rm eff}$. Here, we treat the half-light radius $R_{\rm eff}$ as the half-mass radius $R_{1/2}$, as the M/L gradient exponent in our sample is very small and consistent with zero. The compared IllustrisTNG simulation corresponds to $z = 0.2$, which is the mean redshift of the galaxies in our sample. The fitting functions from Remus et al. (2017) corresponding to Magneticum and Oser simulations – that are plotted in Figure 13 – are shown to match with the simulated distributions over the redshift range $z = 0$–2. In all of the panels of Figure 13, the general trend found in the simulations matches that in our observations. However, the distributions themselves do no overlap. The mismatch between our observation and IllustrisTNG simulation predominantly arises from the higher $f_{\rm dm}^{\rm 3D}$ found in that simulation. However, the γ distribution of the IllustrisTNG simulation is generally consistent with our sample. Similarly, the $\gamma^{\rm lensing}$ distribution obtained from fitting imaging-only data of simulated lenses in the SEAGLE simulation also reproduces the distribution from our observation (Figure 14). However, the $f_{\rm dm}$ distribution of the SEAGLE simulation is not reported yet. In contrast, the Magneticum simulation produces higher radially averaged γ distribution than our observed one, but reproduces a $f_{\rm dm}$ distribution that is generally consistent with our observation (cf. Fig 8 of Remus et al. 2017). The slope of the correlation in our observed distribution falls between the ones from the Magneticum and Oser simulations, which is consistent with the fact that the AGN feedback is a necessary driver in the evolution of these galaxies. To match the slope of the Magneticum simulation with our observation, weaker AGN and stellar feedback would be necessary. However, that would also push up the simulated γ distribution and increase the discrepancy with our observation. Note that the lower values of $f_{\rm dm}$ than simulations is not unique to our lensing-based analysis, as previous lensing-based studies also find similarly lower $f_{\rm dm}$ distributions (Figure 15; Sonnenfeld et al. 2015; Oldham & Auger 2018). Moreover, the Magneticum and IllustrisTNG simulations themselves have offsets between their predicted $f_{\rm dm}$ distributions due to differences in their implementations of baryonic physics despite the similarities of the adopted feedback mechanisms. The offset between simulations highlights that the effect of AGN feedback depends criticailly on the specific implementation. Thus our observations cannot test AGN feedback in general, but they are a powerful empirical test of specific implementations.

Similarly, the trend in our observed distribution of $f_{\rm dm}^{\rm 3D}$–$R_{\rm eff}$ matches with the ones from the simulations, however the distributions themselves do not match (Figure 13). The moderate correlation ($r = 0.46 \pm 0.07$) between the size of the galaxies and the central dark matter fraction supports that these galaxies have grown predominantly through minor mergers. As minor dissipataion-less mergers mostly deposit material at the outer region of the halo, the galaxy's half-mass radius increases as a result. Thus, the dark matter fraction within the half-mass radius also increases and becomes correlated with the half-mass radius.

The velocity dispersion distribution of the IllustrisTNG simulation is smaller than the measured distribution for our sample. Wang et al. (2020) note that this difference between observation and simulation in velocity dispersion, along with the mismatch in the $f_{\rm dm}$ distribution, point to potentially insufficient implementation of baryonic physics in the IllustrisTNG simulation.

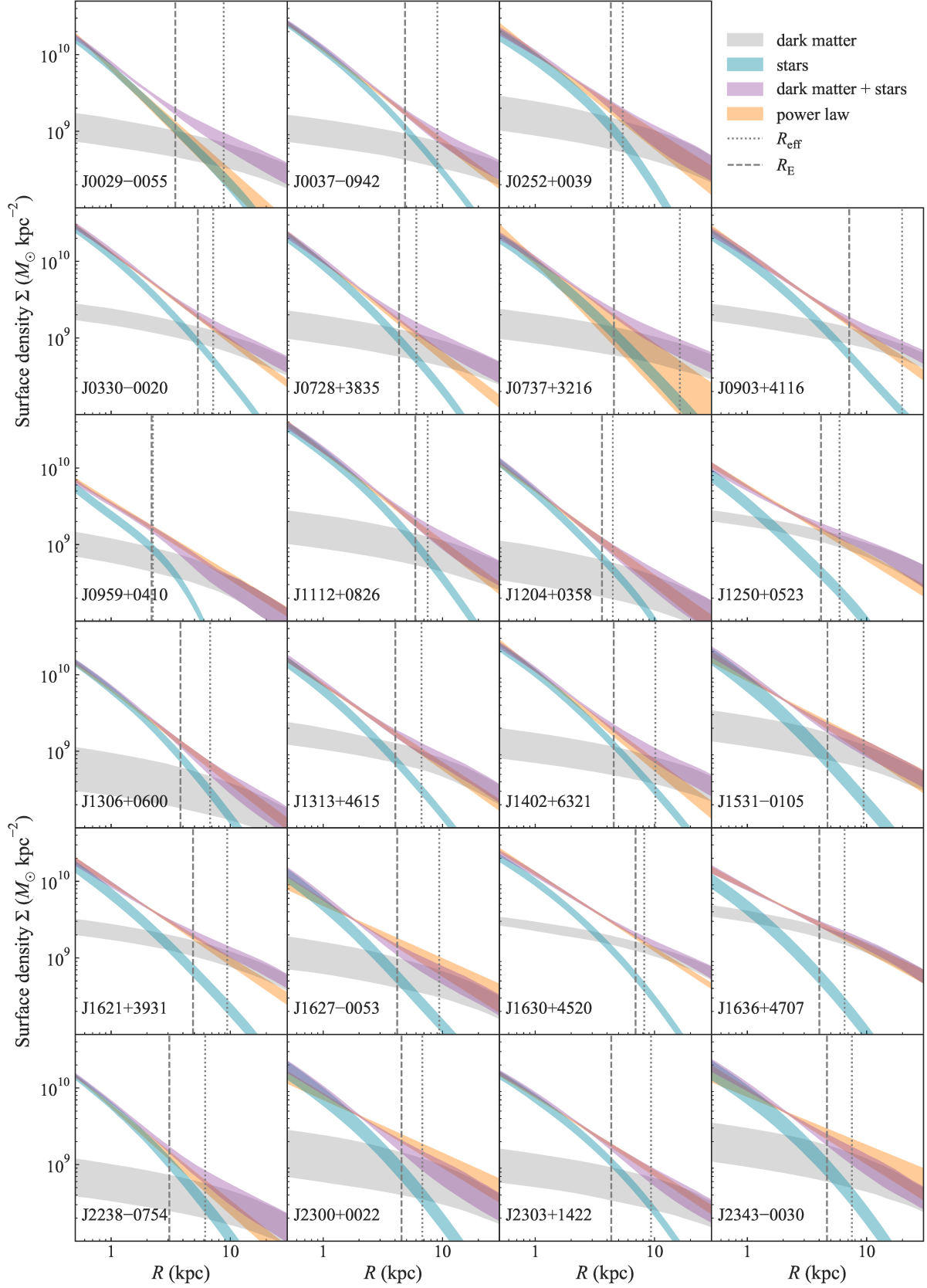In summary, our constraints on the elliptical galaxy structure

Figure 9. Shape of the surface density profile for total and component mass distributions. The purple shaded region shows the 1σ credible region for the total mass distribution from joint strong-lensing–kinematics–weak lensing constraints. The grey and teal blue shaded regions show the dark and stellar component distributions, respectively. The power-law profile from our lens models are shown with the orange shaded region. The Einstein and the effective radii for each lens system are marked with the dashed grey line and the dotted grey line, respectively.
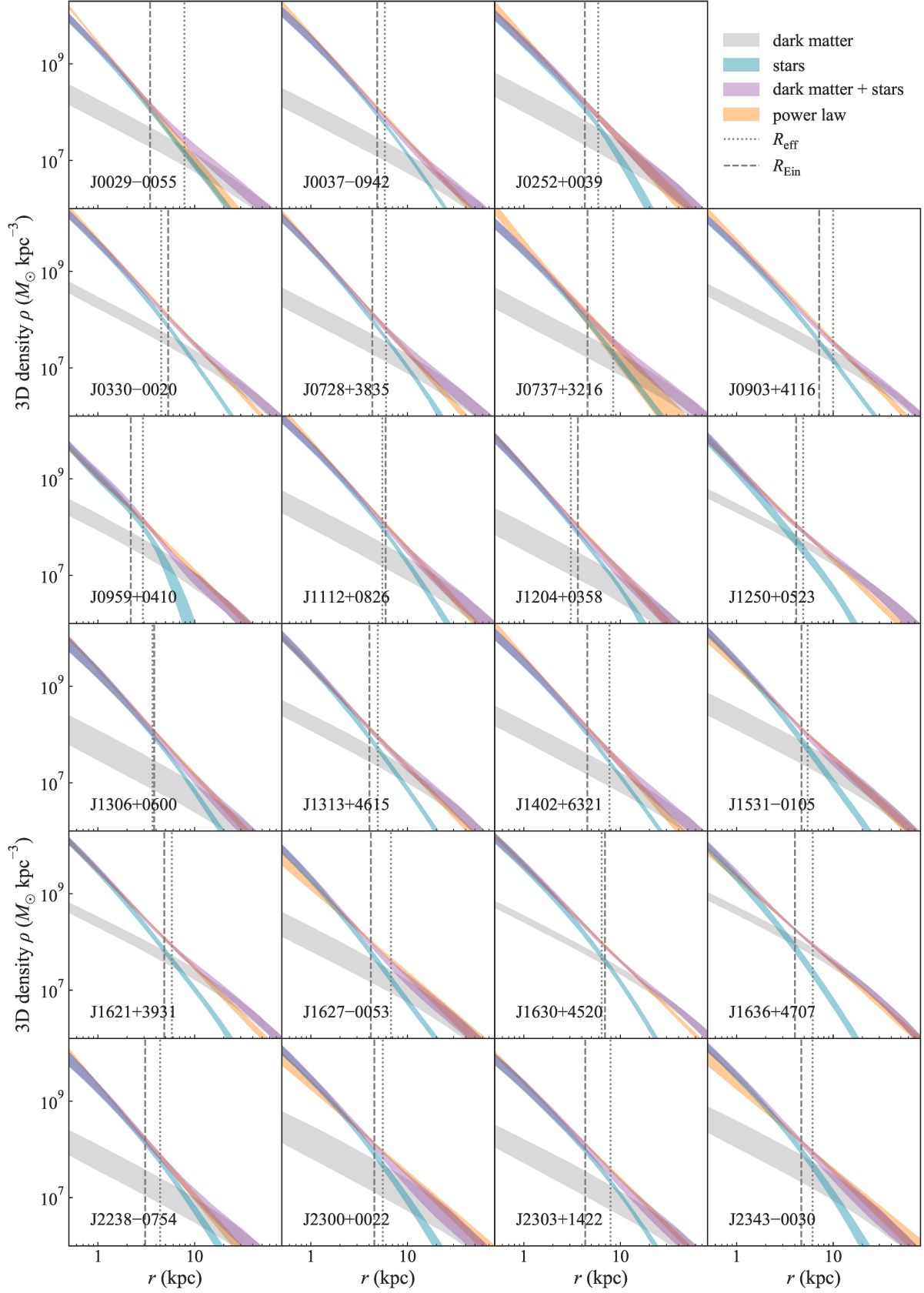
Figure 10. Shape of the 3D density profile for total and component mass distributions from spherical deprojection from the surface density profiles shown in Figure 9. The purple shaded region shows the 1σ credible region for the total mass distribution from joint strong-lensing–kinematics–weak lensing constraints. The grey and teal blue shaded regions show the dark and stellar component distributions, respectively. The power-law profile from our lens models are shown with the orange shaded region. The Einstein and the effective radii for each lens system are marked with the dashed grey line and the dotted grey line, respectively.
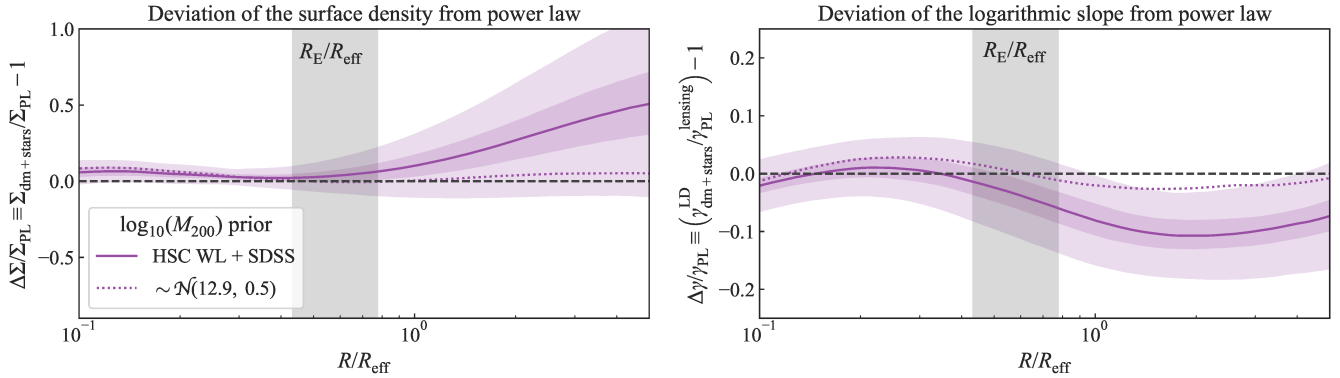
Figure 11. Deviation of the surface density profile (left) and the local logarithmic slope (right) in the NFW+stars mass distribution from the power-law mass distribution. The NFW+stars mass distribution is constrained from joint lensing–dynamics analysis, whereas the power-law mass distribution is constrained from lensing-only data. The solid purple line shows the median deviation for our adopted prior based on HSC weak lensing measurements of SDSS galaxies. The darker and lighter purple shaded regions illustrate the 1σ credible interval of the median and the median absolute deviation, respectively. The grey dashed line represent no deviation from the power-law. The dotted purple line correspond to the median deviation if a prior with a lower mean halo mass is adopted. The vertical grey shaded region shows the 1σ distribution of $R_{\rm E}/R_{\rm eff}$ for our sample. The deviation of the NFW+stars mass distributions from than the power-law estimate at the Einstein radius has a median of ~5 per cent and median absolute deviation of ~7 per cent. The deviation of the logarithmic slopes at the Einstein radius has a median of ~2 per cent and median absolute deviation of ~2 per cent.

in the context of previous observations and simulations are consistent with the following formation and evolution scenario for massive elliptical galaxies. The first stage of the formation of elliptical galaxies is through dissipational processes at $z \gtrsim 2$, when most of their present day stars are formed. The dissipation leads to contraction in the dark matter halo and the resultant total density profile observed in cosmological numerical simulations is steeper than the isothermal case (e.g., Gnedin et al. 2004; Naab et al. 2007b; Duffy et al. 2010). After $z \approx 2$, the growth of the elliptical galaxies is dominated by gas-poor dissipation-less mergers, explaining their growth in size without the addition of younger stellar populations (e.g., Newman et al. 2012; Nipoti et al. 2012). This growth mechanisms is consistent with the $R_{\rm eff} - f_{\rm dm}$ correlation observed in our sample. Multiple dissipation-less mergers decrease the total density profile of the galaxies to bring it close to isothermal, and increase the half-mass radius and the central dark matter fraction with decreasing redshift (Tortora et al. 2014). Furthermore, AGN feedback expands back the contracted dark matter halos (Martizzi et al. 2013; Peirani et al. 2019), which is supported by the slope of our observed $\gamma^{\rm LD}_{[0.4, 4]R_{\rm eff}} - f^{\rm 3D}_{\rm dm}$ distribution. Dynamical heating from accretion may also play a role in expanding the dark matter halos in addition to the AGN feedback, however we do not find any indication either in favor or against the presence of dynamical heating in our sample. Gas-rich dissipational mergers can also happen at some point of the galaxies' growth history. However, the small intrinsic scatter in the halo response parameter ($\sigma_\nu \lesssim 0.1$) around the mean $\mu_\nu \sim 0$ in our sample, and the ages of the stellar populations (Thomas et al. 2005), indicate that such gas-rich mergers are relatively rare, even though their contribution cannot be ruled out or confirmed conclusively given the present precision of numerical simulations and observations (see, e.g., Sonnenfeld et al. 2014; Remus et al. 2017; Xu et al. 2017, for discussion).

5.3 Stellar IMF and M/L gradient

Our inferred stellar mass $M^{\rm LD}_\star$ distribution from the joint lensing–dynamics analysis support a heavier stellar IMF such as the Salpeter IMF. In Figure 16, we illustrate the distribution of the IMF mismatch
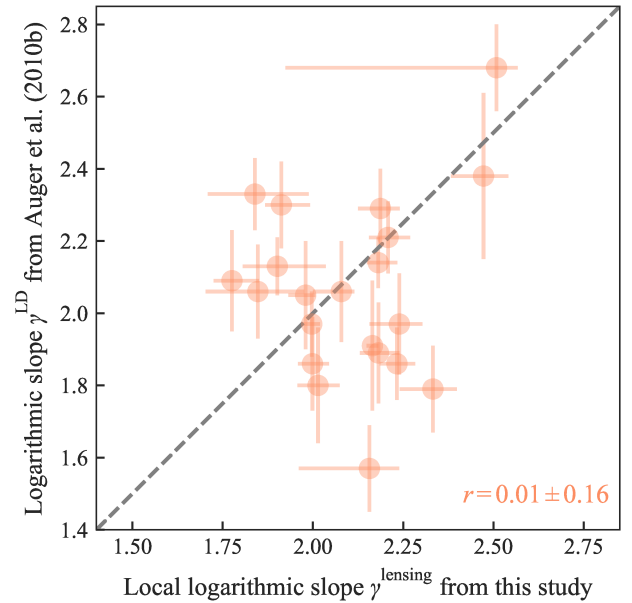

Figure 12. Comparison between the estimated local logarithmic slopes γ in this study, and the radially averaged logarithmic slope estimated by the SLACS analysis (Auger et al. 2010b). Whereas Auger et al. (2010b) constrain only the enclosed mass within the Einstein radius from the imaging data and then obtain the profile slope in combination with the stellar kinematics, we constrain both the enclosed mass and the local slope from the imaging data only. This plot contains 21 systems that have $\gamma_{\rm SLACS}$ measured by Auger et al. (2010b). The dashed grey line illustrates a hypothetical perfect correlation to guide the visualization. We find no correlation between the two measurements with biweight midcorrelation $r = 0.01 \pm 0.16$. Given the measurement uncertainties and an additional 12 per cent uncertainty on $\gamma_{\rm SLACS}$ – which is equivalent to 6 per cent unaccounted systematic uncertainty in the measured kinematics (Birrer et al. 2020) – we cannot conclude based on this comparison whether the expected correlation for a global power law is present under the noise, or not.
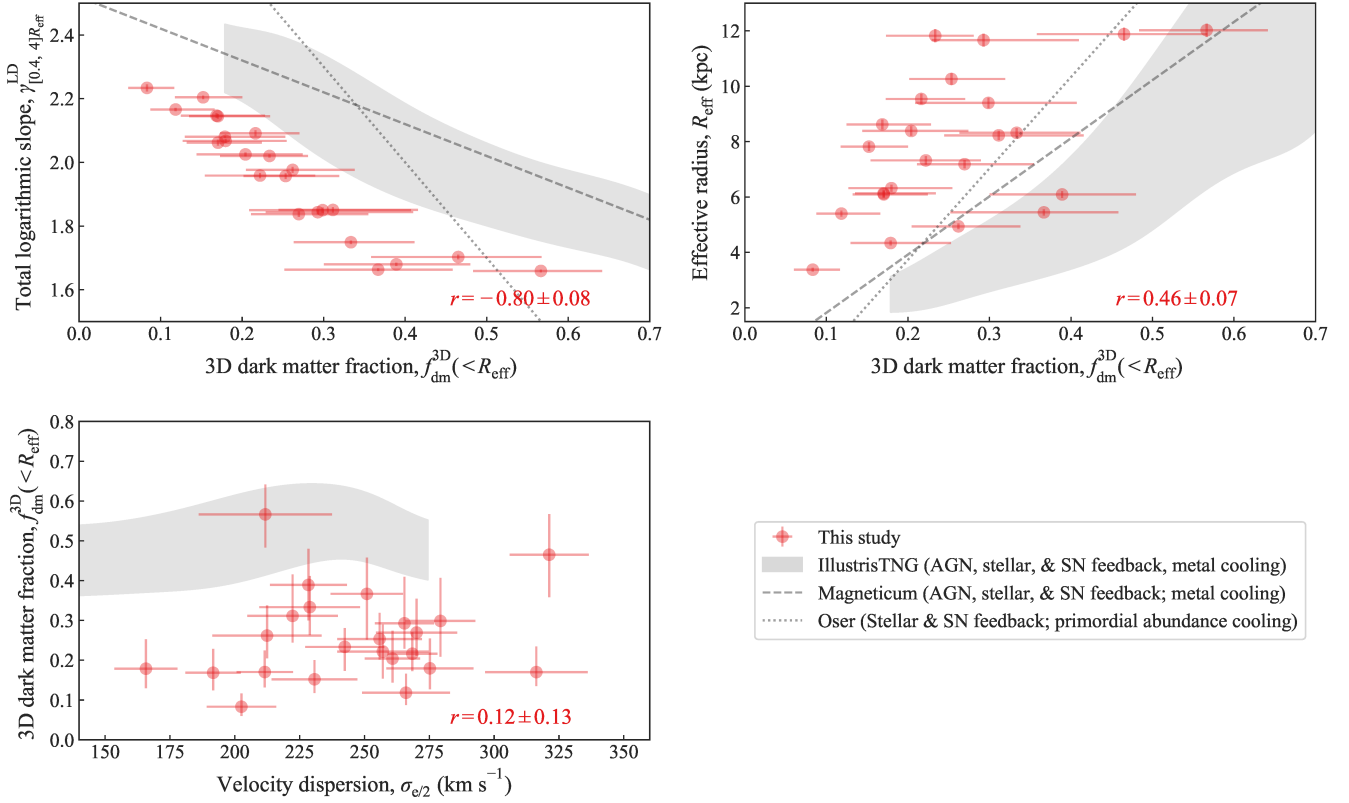
Figure 13. Comparison of structural parameters of our SLACS sample with cosmological hydrodynamical simulations: IllustrisTNG (grey shaded region reprenting 1σ confidence interval; Wang et al. 2020), Magneticum (grey dashed line; Remus et al. 2017), and from Oser et al. (2010, ; grey dashed-dotted line). The 1σ interval for the IllustrisTNG simulation is produced by smoothing the corresponding distributions with a Gaussian kernel of bandwidth 0.05 for f_{dm}^{3D} and 20 km s^{-1} for $\sigma_{e/2}$. **Top left:** distributions of the logarithmic slope γ^{LD} of 3D total mass profile within $0.4R_{eff}$ and $4R_{eff}$ and the 3d dark matter fraction f_{dm}^{3D} within R_{eff}. The distributions from simulations do not match both the slope and the dark matter fraction. The orientation of the SLACS distribution is between the orientations from Magneticum and Oser simulations, which indicates moderate amount AGN and stellar feedback in SLACS galaxies' formation history. **Top right**: distributions of the 3D dark matter fraction f_{dm}^{3D} and the effective radius R_{eff}. These two quantites are moderately correlated, which is consistent with the scenario where elliptical galaxies grow in size through minor mergers and thus the central dark matter fraction within the half-mass radius increases. **Bottom left:** distributions of the velocity dispersion within $R_{eff}/2$ and f_{dm}^{3D}. The apertures for the observed velocity dispersions are adjusted following Jorgensen et al. (1995). The IllustrinsTNG galaxies match neither with the dark matter fraction distribution nor with the velocity dispersion distribution of their observational counterparts at the same stellar mass range, which points to potential inadequacies in the implementation of baryonic physics in the simulation.

parameter as defined by (Treu et al. 2010) as

$$\alpha^{Salp} \equiv \frac{M_\star^{LD}}{M_\star^{SPS,\ Salp}}, \tag{14}$$

where $M_\star^{SPS,\ Salp}$ is the SPS-based measurement of the stellar mass assuming the Salpeter IMF. If the true underlying IMF follows the Salpeter function, then $\alpha^{Salp} = 1$ is expected. For the Chabrier IMF, in comparison, the IMF mismatch parameter should be $\alpha^{Salp} \approx 0.56$ or $\log_{10} \alpha^{Salp} \approx -0.25$. The mean IMF mismatch parameter of our sample is $\langle \log_{10} \alpha^{Salp} \rangle_{median} = 0.00 \pm 0.03$ with an intrinsic scatter of $0.11^{+0.03}_{-0.02}$. Our observed distribution of α^{Salp}–σ_e is in good agreement with the fitting function provided by Posacki et al. (2015)

$$\log_{10} \alpha^{Salp} = (0.38 \pm 0.04) \times \log_{10} \left(\frac{\sigma_e}{200 \text{ km s}^{-1}} \right) + (-0.06 \pm 0.01), \tag{15}$$

which is obtained from fitting SLACS (Auger et al. 2010b) and ATLAS^{3D} (Cappellari et al. 2013) galaxies. This relation has a rms scatter of 0.12 dex.

Several previous strong-lensing studies also found evidence for heavier IMF such as the Salpeter (e.g., Treu et al. 2010; Spiniello et al. 2011, 2012; Sonnenfeld et al. 2012; Oldham & Auger 2018). Multiple non-lensing studies similarly found evidence for a heavy IMF in elliptical galaxies – e.g., from dynamics of local elliptical galaxies (Cappellari et al. 2012), and based on single stellar population (SSP) models (e.g, Conroy & van Dokkum 2012; La Barbera et al. 2013; Spiniello et al. 2014).

Whereas the observed IMF in the Milky Way stellar populations is closer to a Chabrier IMF (e.g., Bastian et al. 2010), our observation of a Salpeter-like IMF in massive elliptical galaxies indicates that stars formed in a different environment in these elliptical galaxies. Star formation in a turbulent environment with high gas density – that can result from gar-rich mergers and accretion events at the early phase of elliptical galaxy formation – can lead to a heavy IMF such as the Salpeter IMF (Hopkins 2013; Chabrier et al. 2014).

We find that the M/L-gradient exponent is very small and consistent with zero, as the 95 per cent upper limit for the sample mean of the M/L-gradient exponent η is 0.02. This small value is consistent with Oldham & Auger (2018), who found that most of
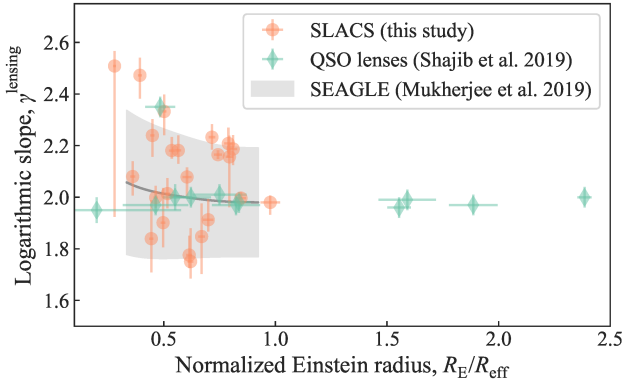
Figure 14. Distribution of the local logarithmic slopes for the SLACS lenses (orange points in the left-hand panel). We take $R_{\rm E}/R_{\rm eff}$ along the x-axis to normalize the radius where the local slope is estimated. For comparison, we show the distribution of slopes for strongly lensed quasars systems (green) from Shajib et al. (2019). The grey solid line shows the mean of the slope distribution of simulated lens systems from SEAGLE and the grey shaded region shows 1 standard deviation around the mean (Mukherjee et al. 2019). The 1σ interval is produced by smoothing the simulated points from SEAGLE AGNdT8 simulation – which includes AGN feedback – with a Gaussian kernel of bandwidth 0.15. Our measured values of the logarithmic slopes from lensing-only data match well with the distribution of similarly measured logarithmic slopes of simulated lens systems that include AGN feedback.
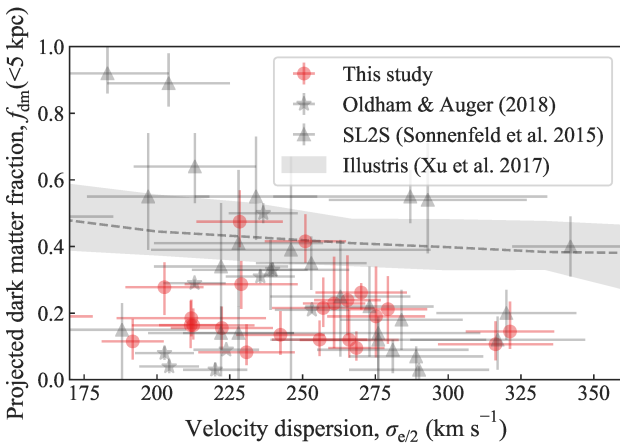
Figure 15. Comparison of our observed distribution of projected dark matter fraction and the stellar velocity dispersion with previous observation and the Illustris simulation (Xu et al. 2017). The grey shaded region shows the 90 per cent confidence interval and the dashed grey line shows the median of the Illustris distribution. Our observed distribution is qualitatively consistent with those from Sonnenfeld et al. (2015) and Oldham & Auger (2018). However, all of these strong-lensing based observations have lower average central dark matter fraction than predicted by the Illustris simulation.

their lens galaxies individually favor $\eta < 0.1$. Our result, however, is in tension with Sonnenfeld et al. (2018), who inferred $\mu_\eta = 0.24 \pm 0.04$ and $\sigma_\eta = 0.09 \pm 0.02$ from a larger SLACS subsample and HSC weak-lensing information. Some fraction of the difference can be explained by the difference in the adopted light profile. Whereas Sonnenfeld et al. (2018) adopt a de Vaucouleurs profile for I-band, we adopt a double Sérsic profile in V-band. The double Sérsic profile fits the light distribution better than the de Vaucouleurs profile. We observe that our double Sérsic profile fits are generally

steeper than the de Vaucouleurs fit of the SLACS galaxies from Auger et al. (2009), which explains the steeper mass profile than the light profile reported by Sonnenfeld et al. (2018). Such steeper mass profile than the light profile can not arise from a color gradient between I and V bands, as observationally I-band light profile is steeper than the V-band as the color gets bluer radially outward (e.g., Tamura et al. 2000). Furthermore, Sonnenfeld et al. (2018) adopt an isotropic profile for orbital anisotropy in their dynamical model, whereas we adopt an Osipkov–Merritt anisotropy profile, which can potentially account for some part of the observed discrepancy by trading radial anisotropy for M/L gradient to reproduce the observed velocity dispersion.

Several recent studies based on SPS modelling have pointed to a radially varying IMF in local elliptical galaxies, where the IMF in the central $\lesssim 2$ kpc region is bottom-heavy – even super-Salpeter – and it gets more bottom-light radially outward to match with the Chabrier IMF in the outer region (e.g., Martín-Navarro et al. 2015; van Dokkum et al. 2017; La Barbera et al. 2019). Such a radially varying IMF is consistent with the scenario where the central stellar population is dominated by *in situ* stars that formed following the Salpeter IMF in a highly turbulent environment at high redshift ($z \gtrsim 2$), and the stellar population in the outer region mostly comes from merged satellites that had a different star-forming environment closer to the Milky Way and thus the Chabrier IMF. However, this radially varying IMF is interpreted from a radially varying M/L ratio based on the SPS models. Thus, at face value our result of very small M/L gradient appears to be in tension with the SPS-based studies mentioned above. However, the M/L gradient observed in the SPS-based studies is dominated by a prominent gradient at $< 0.1 R_{\rm eff}$, whereas our joint lensing–dynamics analysis is sensitive near $0.5 R_{\rm eff} - R_{\rm eff}$ where the SPS-based M/L flattens out. Thus, the apparent difference between our result and SPS-based M/L gradient results can be reconciled. Indeed, van Dokkum et al. (2017) point out that the observed Salpeter IMF in joint lensing–dynamics studies is consistent with the SPS-based result of a radially varying IMF, if the measurement scales are taken into account [cf. Figure 17 of van Dokkum et al. (2017)].

5.4   Implication for time-delay cosmography

In time-delay cosmography, the measurement of the Hubble constant H_0 requires an accurate constraint on the surface density profile near the Einstein radius. Similar to this study, analyses by the Time-Delay COSMOgraphy (TDCOSMO) have adopted a power-law mass distribution – among other choices – for lens modelling (Suyu et al. 2010, 2013; Wong et al. 2017; Birrer et al. 2019; Rusu et al. 2019; Chen et al. 2019; Shajib et al. 2020). A sample of seven strong-lensing systems with measured time delays provide a ~ 2 per cent measurement of the Hubble constant (Millon et al. 2019). However, relaxing the assumption on the power-law profile and thus allowing the full range of mass-sheet degeneracy in the lens models inflates the uncertainty to ~8 per cent (Birrer et al. 2020). As a result, precise knowledge of the internal structure of elliptical galaxies is currently the limiting factor for the time-delay measurement of H_0. Therefore, one option to shrink back the uncertainty on H_0 is to apply an external prior on the internal structure of time-delay lens galaxies, and thus narrow down the allowed range of the MSD. Alternatively, Birrer et al. (2020) constrain the MSD by incorporating the structural information from the power-law lens models presented in this study and their measured kinematics *under the assumption that the TDCOSMO and SLACS lenses are drawn from the the same parent population*. The additional hypothesis and
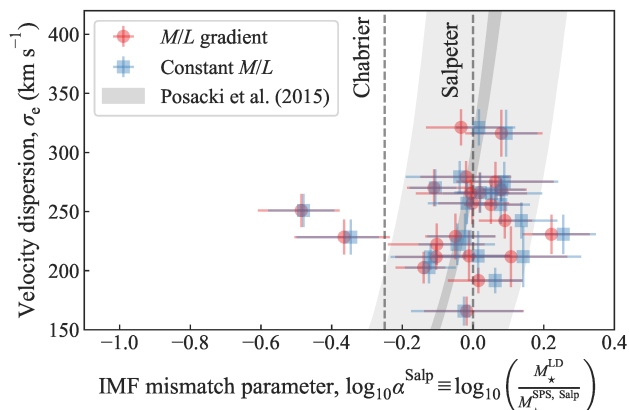
Figure 16. Distribution of the IMF mismatch parameter $\alpha^{\mathrm{Salp}} \equiv M_\star^{\mathrm{LK}}/M_\star^{\mathrm{SPS,\,Salp}}$ and the velocity dispersion σ_{e} within the Effective radius R_{eff}. The vertical dashed grey lines mark the α^{Salp} values expected from the Salpeter or Chabrier IMFs. The distribution of IMF mismatch parameter is consistent with Salpeter IMF with median $\langle \log_{10}\alpha^{\mathrm{Salp}} \rangle_{\mathrm{median}} = 0.00 \pm 0.03$ and intrinsic scatter $0.11^{+0.03}_{-0.02}$. We also show the fitting function from Posacki et al. (2015) obtained from fitting SLACS (Auger et al. 2010b) and atlas$^{\mathrm{3d}}$ galaxies. The deeper grey shaded region represent the mean of the relation and the lighter grey shaded region represents the rms scatter around the relation.

the addition of external information reduces the uncertainty on H_0 to 5 per cent. However, including the structural information from the SLACS galaxies also increases the mean surface density of the time-delay lens galaxies near the Einstein radius by ∼9.5 per cent from the power-law estimate, which also translates to a ∼ 9.5 per cent decrease in the inferred H_0. This increased estimate of the surface density is consistent within the uncertainty with our joint lensing–dynamics analysis, as we find on average ∼5 per cent higher surface density near the Einstein radius for ours stars+NFW model than the power-law model (see Fig 11).

TDCOSMO analyses have also adopted an NFW+stars model in addition to the power-law model and found that the two models are in very good agreement (Millon et al. 2019). In contrast, we observe that the NFW+stars model deviate from the power-law on average by ∼5 per cent at the Einstein radius. We show that the deviation of the NFW+stars model from the power-law model in our analysis is largely driven by the adopted prior on the halo mass. As strong-lensing and kinematics data mostly constrain the mass distribution in the central region, the total halo mass or the dark matter profile is poorly constrained from lensing–dynamics analysis without a prior on the halo mass, or on the mass distribution at scales larger than the NFW scale radius. Our empirical prior on the halo mass M_{200} is obtained from the HSC weak-lensing measurements of the SDSS galaxies, which is the parent sample of the SLACS galaxies. Moreover, the selection function for the SLACS galaxies is accounted for in this prior. If instead, we had adopted a prior on M_{200} with a smaller mean by ∼0.3 dex, the resultant NFW+stars density profile would have matched the power-law model very well on average (Figure 11). A higher M_{200} translates to a higher normalization of the NFW profile, thus the total density profile becomes dark-matter-dominated at a relatively smaller radius and the slope of the total density profile gets shallower. We provide two possible explanations for why our NFW+stars model deviates from the power-law whereas the ones in the TDCOSMO analyses do not.

The first possible explanation is that there may be a bias in the NFW profile normalization in the TDCOSMO models due to

weak prior constraints on the mass distribution at large scales. TD-COSMO analyses have generally adopted a prior only on the NFW scale radius, which only has a weak constraint on M_{200} as the $M-c$ relation is not imposed. As a result, the resultant halo mass may be biased low as lens models tend to produce lower halo masses without strong priors on the mass distribution at scales larger than the NFW scale radius (e.g., Oldham & Auger 2018). A smaller halo mass by ∼0.3 dex may not be noticeable within the measurement uncertainty when looking at individual lens galaxies, but it may be identifiable by investigating the $M_\star - M_{200}$ relation for the TD-COSMO lenses. However, due to the typically large intrinsic scatter in M_{200}, a sample size of seven may not be sufficient to identify such a bias.

The second possible explanation is that the SLACS lens galaxies and the TDCOSMO lens galaxies do not belong to the same galaxy population. Even if the predecessors of the SLACS galaxies at $z = 0.6$ represent the time-delay lens galaxy population, they may have evolved since $z = 0.6$ to have a different internal structure at $z = 0.2$ (see, e.g., Sonnenfeld et al. 2015). A M_{200} prior with larger mean leads to a larger dark matter fraction f_{dm}. As f_{dm} is observed to increase from higher redshift to lower redshift (Tortora et al. 2014), it is possible that a lower f_{dm} in the time-delay lenses with mean deflector redshift $\langle z_{\mathrm{d}} \rangle = 0.6$ makes the power-law model and the NFW+stars model to be in good agreement.

We note that an increasing f_{dm} with decreasing redshift in the TDCOSMO lenses would also be consistent at least qualitatively with the weak trend observed in the individual H_0 measurements from these lenses (Wong et al. 2019). As increasing f_{dm} can create larger positive deviation between the NFW+stars model and the power-law model, the resultant H_0 from the power-law model would shift toward higher values with decreasing redshift. However, this scenario requires a combination of both of the above explanations where high-redshift galaxies are less vulnerable to an NFW normalization bias due to their lower underlying f_{dm}, and lower redshift galaxies are more vulnerable to the normalization bias due to their higher underlying f_{dm}.

Blum et al. (2020) suggest that a hypothetical 10 per cent positive bias in the Hubble constant from time-delay cosmography would point to the presence of a large core in the dark matter distribution. We show that such a shift in the Hubble constant can also be explained by a relatively higher normalization in the vanilla NFW profile, if the SLACS lens galaxies and the TDCOSMO lens galaxies are structurally self-similar.

5.5   Limitations of this study

Although we do not account for external convergence from the line-of-sight structures in our lens models, this simplification has no impact on our results, as we discuss in this paragraph. Birrer et al. (2020) find the mean external convergence for a subsample of 33 SLACS lenses to be $\langle \kappa_{\mathrm{ext}} \rangle = 0.005^{+0.022}_{-0.009}$. However, this subsample is curated to only select systems with small local overdensities [see Birrer et al. (2020) for details]. Our sample of 23 lenses has a large overlap with this sample of 33 lenses from Birrer et al. (2020), thus the overlapping fraction of our sample has negligible mean external convergence. The external convergences for the remaining systems are likely to be not negligible and our joint lensing–dynamics analysis can be suspect of bias due to not correcting for the external convergence in our joint lensing–dynamics analysis. However, we only adopt MSD-invariant quantities that are not sensitive to the external convergence – such as the Einstein radius R_{E} and the local differential term $R_{\mathrm{E}}\alpha''_{\mathrm{E}}/(1 - \kappa_{\mathrm{E}})$ – as the lensing constraints in our

joint lensing–dynamics analysis. Thus, our joint lensing–dynamics analysis is insensitive to the external convergence.

In our adiabatic contraction model, we assume that the stellar mass distribution initially resembled the NFW profile following the adiabatic contraction model of Blumenthal et al. (1986). This is certainly not true as the stellar mass delivered to the elliptical galaxies through mergers did not resemble the NFW profile at the time of the merger. However, cosmological hydrodynamical simulations have shown that such a simple model can reproduce the contracted profile of the dark matter even after gas-rich mergers, albeit with a modification in the adiabatic invariant quantity $rM(r)$ (e.g., Gnedin et al. 2004; Duffy et al. 2010). We adopted the modification of Dutton et al. (2007), which can approximate the contracted profiles of both Gnedin et al. (2004) and Abadi et al. (2010). However, since cosmological hydrodynamical simulations have so far been unable to match all the properties of their observational counterparts, our simple model for adiabatic contraction may not be fully justified to truthfully represent the interplay between the baryonic and the dark matter distributions. Such inconsistencies may be identified by an alternative analysis of the same sample that adopts the generalized NFW profile for the dark matter distribution, and compare the results with the ones from our the adiabatic contraction model. We leave such explorations for future studies.

6 SUMMARY

We uniformly modelled a sample of 23 SLACS lenses and constrained their structural properties from a joint lensing–dynamics analysis. The lens modelling in this study is different from the original SLACS analysis, in which the lens images were modelled by fixing the logarithmic slope to $\gamma = 2$ and then the logarithmic slopes were inferred from the stellar kinematics. In contrast, in this study we first estimate the logarithmic slopes only from the lensing observables, i.e., the lens image. We then combine the stellar kinematics to constrain the amount of contraction in the dark matter distribution for two models of stellar mass distribution: (i) with constant M/L and (ii) with M/L gradient. We summarize the main results of this paper below.

- From the combination of lensing and kinematic observables, we constrain the average halo response parameter $\mu_\nu = -0.06 \pm 0.04$ with intrinsic scatter $\sigma_\nu \leq 0.092$ (95 per cent upper limit) for a constant stellar M/L model. For a stellar M/L gradient model, we find $\mu_\nu = -0.03^{+0.04}_{-0.05}$ and $\sigma_\nu \leq 0.074$. Our results are consistent with a dark matter halo described by an NFW profile with no contraction nor expansion. For comparison, the Blumenthal et al. (1986) model corresponds to $\nu = 1$, the contraction in Gnedin et al. (2004) simulations correspond to $\nu = 0.8$, and the contraction in Abadi et al. (2010) simulations correspond to $\nu \sim 0.4$ – which are all ruled out by our result. Our results are consistent with a scenario in which elliptical galaxies grow by dissipational processes at $z \gtrsim 2$, steepening their dark matter halos. At later times, AGN feedback – with potential additional contributions from dynamical heating through accretion events – expand the dark matter halos back to an NFW profile, on average.

- The distribution of logarithmic slopes γ for the power-law model constrained from the imaging-only data has a median $\langle \gamma_{\rm PL}^{\rm lensing} \rangle = 2.08 \pm 0.03$ and an intrinsic scatter 0.13 ± 0.02. This is consistent with the slope distribution from the SLACS analysis $\langle \gamma_{\rm PL}^{\rm LD} \rangle = 2.078 \pm 0.027$ with an intrinsic scatter of 0.16 ± 0.02. We find that the NFW+stars profile constrained from our joint lensing–

dynamics analysis only deviates by $\lesssim 5$ per cent on average near the Einstein radius ($R_{\rm E} \sim R_{\rm eff}/2$ for our sample), with even smaller deviation at smaller scales. The small deviation in the mean compared to the intrinsic scatter explains the good agreement between the average local logarithmic slope from lensing-only data and the radially averaged logarithmic slopes from Auger et al. (2009).

- For the stellar M/L gradient model, we find that most galaxies do not require a significant M/L gradient. The 95 per cent upper limit for the sample mean of the exponent η in our M/L-gradient model $\Upsilon \propto (R/R_{\rm eff})^{-\eta}$ is 0.02, which corresponds to $\lesssim 5$ per cent decrease in M/L between $0.1R_{\rm eff}$ and $R_{\rm eff}$. Moreover, the inferred stellar masses from joint lensing–dynamics analysis for the galaxies in our sample is consistent with the Salpeter IMF, with the IMF mismatch parameter $\langle \log_{10} \alpha^{\rm Salp} \rangle_{\rm median} = 0.00 \pm 0.03$ with an intrinsic scatter $0.11^{+0.03}_{-0.02}$. Such a heavy IMF in the central regions of the elliptical galaxies can be explained by a different star-forming environment than in the Milky Way, e.g., the presence of turbulence in high gas density.

In the future, larger samples of galaxy–galaxy lenses at different redshifts will be able to further constrain the evolutionary tracks of elliptical galaxies. Such larger samples can be assembled from past surveys with available high-resolution imaging and ancillary data, e.g., the full SLACS sample (Auger et al. 2009), the Strong Lensing Legacy Survey (SL2S) sample (Sonnenfeld et al. 2013), and the SLACS for the MASSES (S4TM) sample (Shu et al. 2017). Current surveys such as the Dark Energy Survey (DES) and the Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Surveys are producing new galaxy–galaxy lens candidates for confirmation and follow-up on the order of hundreds (Jacobs et al. 2019a,b; Huang et al. 2020). Future surveys, e.g., the Vera Rubin Legacy Survey for Space and Time and the Nancy Grace Roman Space Telescope, will increase the number of newly discovered galaxy–galaxy lenses to thousands (Collett 2015). An automated and uniform modelling pipeline will be essential to study such large samples of lenses and this paper has taken the initial steps towards such an automated pipeline.

REFERENCES

Abadi M. G., Navarro J. F., Fardal M., Babul A., Steinmetz M., 2010, MNRAS, 407, 435

Adams F. C., Bloch A. M., Butler S. C., Druce J. M., Ketchum J. A., 2007, ApJ, 670, 1027

Anderson J., King I. R., 2000, PASP, 112, 1360

Astropy Collaboration 2013, A&A, 558, A33

Astropy Collaboration 2018, AJ, 156, 123

Auger M. W., Treu T., Bolton A. S., Gavazzi R., Koopmans L. V. E., Marshall P. J., Bundy K., Moustakas L. A., 2009, ApJ, 705, 1099

Auger M. W., Treu T., Gavazzi R., Bolton A. S., Koopmans L. V. E., Marshall P. J., 2010a, ApJ, 721, L163

Auger M. W., Treu T., Bolton A. S., Gavazzi R., Koopmans L. V. E., Marshall P. J., Moustakas L. A., Burles S., 2010b, ApJ, 724, 511

Avila R. J., Hack W., Cara M., Borncamp D., Mack J., Smith L., Ubeda L., 2015, in Taylor A. R., Rosolowsky E., eds, Astronomical Society of the Pacific Conference Series Vol. 495, Astronomical Data Analysis Software an Systems XXIV (ADASS XXIV). p. 281 (arXiv:1411.5605)

Barkana R., 1998, ApJ, 502, 531

Barkana R., 1999, FASTELL: Fast calculation of a family of elliptical mass gravitational lens models, Astrophysics Source Code Library (ascl:9910.003)

Barnabè M., Czoske O., Koopmans L. V. E., Treu T., Bolton A. S., 2011, MNRAS, 415, 2215

Barnabè M., Spiniello C., Koopmans L. V. E., Trager S. C., Czoske O., Treu T., 2013, MNRAS, 436, 253

Bastian N., Covey K. R., Meyer M. R., 2010, ARA&A, 48, 339

Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32

Bellstedt S., et al., 2018, MNRAS, 476, 4543

Bertin E., Arnouts S., 1996, A&AS, 117, 393

Bezanson R., van Dokkum P. G., Tal T., Marchesini D., Kriek M., Franx M., Coppi P., 2009, ApJ, 697, 1290

Birrer S., Amara A., 2018, Physics of the Dark Universe, 22, 189

Birrer S., Amara A., Refregier A., 2015, ApJ, 813, 102

Birrer S., et al., 2019, MNRAS, 484, 4726

Birrer S., et al., 2020, arXiv e-prints, p. arXiv:2007.02941

Blum K., Castorina E., Simonović M., 2020, arXiv e-prints, p. arXiv:2001.07182

Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27

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

Bruderer C., Read J. I., Coles J. P., Leier D., Falco E. E., Ferreras I., Saha P., 2016, MNRAS, 456, 870

Caminha G. B., et al., 2017, A&A, 607, A93

Cappellari M., 2016, ARA&A, 54, 597

Cappellari M., et al., 2012, Nature, 484, 485

Cappellari M., et al., 2013, MNRAS, 432, 1709

Chabrier G., Hennebelle P., Charlot S., 2014, ApJ, 796, 75

Chen G. C.-F., et al., 2019, MNRAS,

Claeskens J.-F., Sluse D., Riaud P., Surdej J., 2006, A&A, 451, 865

Colín P., Avila-Reese V., Valenzuela O., 2000, ApJ, 542, 622

Collett T. E., 2015, ApJ, 811, 20

Collett T. E., Cunnington S. D., 2016, MNRAS, 462, 3255

Conroy C., van Dokkum P. G., 2012, ApJ, 760, 71

Czoske O., Barnabè M., Koopmans L. V. E., Treu T., Bolton A. S., 2008, MNRAS, 384, 987

Debattista V. P., Moore B., Quinn T., Kazantzidis S., Maas R., Mayer L., Read J., Stadel J., 2008, ApJ, 681, 1076

Debattista V. P., van den Bosch F. C., Roškar R., Quinn T., Moore B., Cole D. R., 2015, MNRAS, 452, 4094

Diemer B., 2018, ApJS, 239, 35

Diemer B., Joyce M., 2019, ApJ, 871, 168

Dobler G., Keeton C. R., Bolton A. S., Burles S., 2008, ApJ, 685, 57

Dodelson S., Widrow L. M., 1994, Phys. Rev. Lett., 72, 17

Dubinski J., 1994, ApJ, 431, 617

Dubois Y., Gavazzi R., Peirani S., Silk J., 2013, MNRAS, 433, 3297

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., Battye R. A., Booth C. M., 2010, MNRAS, 405, 2161

Dutton A. A., Treu T., 2014, MNRAS, 438, 3594

Dutton A. A., van den Bosch F. C., Dekel A., Courteau S., 2007, ApJ, 654, 27

El-Zant A., Shlosman I., Hoffman Y., 2001, ApJ, 560, 636

Falco E. E., Gorenstein M. V., Shapiro I. I., 1985, ApJ, 289, L1

Ferreras I., Saha P., Burles S., 2008, MNRAS, 383, 857

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Furlong M., et al., 2015, MNRAS, 450, 4486

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., Marshall P. J., Brault F., Ruff A., 2012, ApJ, 761, 170

Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, ApJ, 616, 16

Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Guo Q., White S. D. M., 2008, MNRAS, 384, 2

Hopkins P. F., 2013, MNRAS, 433, 170

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581

Huang X., et al., 2020, ApJ, 894, 78

Humphrey P. J., Buote D. A., 2010, MNRAS, 403, 2143

Hunter J. D., 2007, Computing in Science and Engineering, 9, 90

Jacobs C., et al., 2019a, ApJS, 243, 17

Jacobs C., et al., 2019b, MNRAS, 484, 5330

Johansson P. H., Naab T., Burkert A., 2009, ApJ, 690, 802

Johansson P. H., Naab T., Ostriker J. P., 2012, ApJ, 754, 115

Jones E., Oliphant T., Peterson P., Others 2001, SciPy: Open source scientific tools for Python, http://www.scipy.org/

Jorgensen I., Franx M., Kjaergaard P., 1995, MNRAS, 276, 1341

Kazantzidis S., Kravtsov A. V., Zentner A. R., Allgood B., Nagai D., Moore B., 2004, ApJ, 611, L73

Keeton C. R., Kochanek C. S., Falco E. E., 1998, ApJ, 509, 561

Kennedy J., Eberhart R., 1995, in Proceedings of ICNN'95 - International Conference on Neural Networks. IEEE, doi:10.1109/icnn.1995.488968, https://doi.org/10.1109/icnn.1995.488968

Kluyver T., et al., 2016, in Loizides F., Schmidt B., eds, Positioning and Power in Academic Publishing: Players, Agents and Agendas. IOS Press BV, Amsterdam, Netherlands, pp 87 – 90, doi:10.3233/978-1-61499-649-1-87

Kochanek C. S., 2002, in Natarajan P., ed., The Shapes of Galaxies and their Dark Halos. WORLD SCIENTIFIC, pp 62–71 (arXiv:astro-ph/0106495), doi:10.1142/9789812778017_0010

Kochanek C. S., 2020, MNRAS, 493, 1725

Koopmans L. V. E., 2004, arXiv e-prints, pp astro–ph/0412596

Koopmans L. V. E., 2006, in Mamon G. A., Combes F., Deffayet C., Fort B., eds, EAS Publications Series Vol. 20, EAS Publications Series. pp 161–166 (arXiv:astro-ph/0511121), doi:10.1051/eas:2006064

Kostrzewa-Rutkowska Z., Wyrzykowski Ł., Auger M. W., Collett T. E., Belokurov V., 2014, MNRAS, 441, 3238

Krist J. E., Hook R. N., Stoehr F., 2011, 20 years of Hubble Space Telescope optical modeling using Tiny Tim. p. 81270J, doi:10.1117/12.892762

La Barbera F., Ferreras I., Vazdekis A., de la Rosa I. G., de Carvalho R. R., Trevisan M., Falcón-Barroso J., Ricciardelli E., 2013, MNRAS, 433, 3017

La Barbera F., et al., 2019, MNRAS, 489, 4090

Laporte C. F. P., White S. D. M., Naab T., Ruszkowski M., Springel V., 2012, MNRAS, 424, 747

Limousin M., et al., 2007, ApJ, 668, 643

Ma C.-P., Boylan-Kolchin M., 2004, Phys. Rev. Lett., 93, 021301

Mamon G. A., Łokas E. L., 2005, MNRAS, 363, 705

Martín-Navarro I., La Barbera F., Vazdekis A., Falcón-Barroso J., Ferreras I., 2015, MNRAS, 447, 1033

Martizzi D., Teyssier R., Moore B., Wentz T., 2012, MNRAS, 422, 3081

Martizzi D., Teyssier R., Moore B., 2013, MNRAS, 432, 1947

Merritt D., 1985a, AJ, 90, 1027

Merritt D., 1985b, MNRAS, 214, 25P

Millon M., et al., 2019, arXiv e-prints, p. arXiv:1912.08027

Mukherjee S., Koopmans L. V. E., Metcalf R. B., Tortora C., Schaller M., Schaye J., Vernardos G., Bellagamba F., 2019, arXiv e-prints, p. arXiv:1901.01095

Naab T., Johansson P. H., Ostriker J. P., Efstathiou G., 2007a, ApJ, 658, 710

Naab T., Johansson P. H., Ostriker J. P., Efstathiou G., 2007b, ApJ, 658, 710

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Newman A. B., Ellis R. S., Bundy K., Treu T., 2012, ApJ, 746, 162

Newman A. B., Treu T., Ellis R. S., Sand D. J., 2013, ApJ, 765, 25

Newman A. B., Ellis R. S., Treu T., 2015, ApJ, 814, 26

Nipoti C., Stiavelli M., Ciotti L., Treu T., Rosati P., 2003, MNRAS, 344, 748

Nipoti C., Treu T., Leauthaud A., Bundy K., Newman A. B., Auger M. W., 2012, MNRAS, 422, 1714

Oh S.-H., de Blok W. J. G., Brinks E., Walter F., Kennicutt Robert C. J., 2011, AJ, 141, 193

Oldham L. J., Auger M. W., 2018, MNRAS, 476, 133

Oliphant T. E., 2015, Guide to NumPy, 2nd edn. CreateSpace Independent Publishing Platform, USA

Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, ApJ, 725, 2312

Osipkov L. P., 1979, Pisma v Astronomicheskii Zhurnal, 5, 77

Peirani S., Kay S., Silk J., 2008, A&A, 479, 123

Peirani S., et al., 2017, MNRAS, 472, 2153

Peirani S., et al., 2019, MNRAS, 483, 4615

Pillepich A., et al., 2018, MNRAS, 473, 4077

Planck Collaboration 2018, arXiv e-prints, p. arXiv:1807.06209

Pontzen A., Governato F., 2012, MNRAS, 421, 3464

Posacki S., Cappellari M., Treu T., Pellegrini S., Ciotti L., 2015, MNRAS, 446, 493

Read J. I., 2014, Journal of Physics G Nuclear Physics, 41, 063101

Refregier A., 2003, MNRAS, 338, 35

Remus R.-S., Burkert A., Dolag K., Johansson P. H., Naab T., Oser L., Thomas J., 2013, ApJ, 766, 71

Remus R.-S., Dolag K., Naab T., Burkert A., Hirschmann M., Hoffmann T. L., Johansson P. H., 2017, MNRAS, 464, 3742

Ritondale E., Auger M. W., Vegetti S., McKean J. P., 2019, MNRAS, 482, 4744

Rusu C. E., et al., 2016, MNRAS, 458, 2

Rusu C. E., et al., 2019, arXiv e-prints, p. arXiv:1905.09338

Sand D. J., Treu T., Ellis R. S., Smith G. P., Kneib J.-P., 2008, ApJ, 674, 711

Schaller M., et al., 2015, MNRAS, 452, 343

Schreiber J., 2018, Journal of Machine Learning Research, 18, 1

Sérsic J. L., 1968, Atlas de Galaxias Australes. http://adsabs.harvard.edu/abs/1968adga.book.....S

Shajib A. J., 2019, MNRAS, 488, 1387

Shajib A. J., Treu T., Agnello A., 2018, MNRAS, 473, 210

Shajib A. J., et al., 2019, MNRAS, 483, 5649

Shajib A. J., et al., 2020, MNRAS, 494, 6072

Shu Y., et al., 2015, ApJ, 803, 71

Shu Y., et al., 2017, ApJ, 851, 48

Silverman B. W., 1986, Density estimation for statistics and data analysis. Chapman and Hall, London New York

Sluse D., Chantry V., Magain P., Courbin F., Meylan G., 2012, A&A, 538, A99

Sonnenfeld A., Treu T., Gavazzi R., Marshall P. J., Auger M. W., Suyu S. H., Koopmans L. V. E., Bolton A. S., 2012, ApJ, 752, 163

Sonnenfeld A., Gavazzi R., Suyu S. H., Treu T., Marshall P. J., 2013, ApJ, 777, 97

Sonnenfeld A., Nipoti C., Treu T., 2014, ApJ, 786, 89

Sonnenfeld A., Treu T., Marshall P. J., Suyu S. H., Gavazzi R., Auger M. W., Nipoti C., 2015, ApJ, 800, 94

Sonnenfeld A., Leauthaud A., Auger M. W., Gavazzi R., Treu T., More S., Komiyama Y., 2018, MNRAS, 481, 164

Sonnenfeld A., Wang W., Bahcall N., 2019a, A&A, 622, A30

Sonnenfeld A., Jaelani A. T., Chan J., More A., Suyu S. H., Wong K. C., Oguri M., Lee C.-H., 2019b, A&A, 630, A71

Spergel D. N., Steinhardt P. J., 2000, Phys. Rev. Lett., 84, 3760

Spiniello C., Koopmans L. V. E., Trager S. C., Czoske O., Treu T., 2011, MNRAS, 417, 3000

Spiniello C., Trager S. C., Koopmans L. V. E., Chen Y. P., 2012, ApJ, 753, L32

Spiniello C., Trager S., Koopmans L. V. E., Conroy C., 2014, MNRAS, 438, 1483

Spiniello C., Trager S. C., Koopmans L. V. E., 2015, ApJ, 803, 87

Suyu S. H., Halkola A., 2010, A&A, 524, A94

Suyu S. H., Marshall P. J., Auger M. W., Hilbert S., Blandford R. D., Koopmans L. V. E., Fassnacht C. D., Treu T., 2010, ApJ, 711, 201

Suyu S. H., et al., 2013, ApJ, 766, 70

Tamura N., Kobayashi C., Arimoto N., Kodama T., Ohta K., 2000, AJ, 119, 2134

Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, ApJ, 621, 673

Thomas J., Saglia R. P., Bender R., Thomas D., Gebhardt K., Magorrian J., Corsini E. M., Wegner G., 2007, MNRAS, 382, 657

Tortora C., Napolitano N. R., Saglia R. P., Romanowsky A. J., Covone G., Capaccioli M., 2014, MNRAS, 445, 162

Treu T., Koopmans L. V. E., 2002, MNRAS, 337, L6

Treu T., Koopmans L. V. E., 2004, ApJ, 611, 739

Treu T., Koopmans L. V., Bolton A. S., Burles S., Moustakas L. A., 2006, ApJ, 640, 662

Treu T., Gavazzi R., Gorecki A., Marshall P. J., Koopmans L. V. E., Bolton A. S., Moustakas L. A., Burles S., 2009, ApJ, 690, 670

Treu T., Auger M. W., Koopmans L. V. E., Gavazzi R., Marshall P. J., Bolton A. S., 2010, ApJ, 709, 1195

Vogelsberger M., et al., 2014, MNRAS, 444, 1518

Wang Y., et al., 2020, MNRAS, 491, 5188

Waskom M., et al., 2014, seaborn: v0.5.0 (November 2014), doi:10.5281/zenodo.12710, https://doi.org/10.5281/zenodo.12710

Weinberger R., et al., 2017, MNRAS, 465, 3291

Wong K. C., et al., 2017, MNRAS, 465, 4895

Wong K. C., et al., 2019, arXiv e-prints, p. arXiv:1907.04869

Xu D., Springel V., Sluse D., Schneider P., Sonnenfeld A., Nelson D., Vogelsberger M., Hernquist L., 2017, MNRAS, 469, 1824

de Blok W. J. G., McGaugh S. S., Rubin V. C., 2001, AJ, 122, 2396

de Vaucouleurs G., 1948, Annales d'Astrophysique, 11, 247

van Dokkum P., Conroy C., Villaume A., Brodie J., Romanowsky A. J., 2017, ApJ, 841, 68

APPENDIX A: MASS AND LIGHT ALIGNMENTS

Alignment between the dark and baryonic components of mass distribution can be used to validate the predictions from simulations. In elliptical lens galaxies, the misalignment between the mass and light is observed to be within $\pm10°$ in the absence of large external shear effects ($\gtrsim0.1$), whereas larger misalignments are usually

accompanied with large external shear (e.g., Keeton et al. 1998; Koopmans 2006; Treu et al. 2009; Sluse et al. 2012; Shajib et al. 2019). This observation agrees well with the ILLUSTRIS simulation (Xu et al. 2017). However, there have been mismatching reports in the literature on the correlation between the light and mass ellipticities or axes ratios, which can most likely be attributed to different selection functions. The differences in the selection functions can arise from different lens-finding methods and from different population of lenses – e.g., quads or doubles, galaxy–galaxy lenses or lensed quasars. For example, Koopmans (2006); Sluse et al. (2012); Gavazzi et al. (2012) report strong correlation between mass and light ellipticities. In comparison, Keeton et al. (1998); Ferreras et al. (2008); Rusu et al. (2016); Shajib et al. (2019) find weak to no correlation between the mass and light ellipticities.

In this appendix, we investigate the alignment between mass and ellipticities and their correlation with some other model parameters. We present the results in Section A1 and discuss them in Section A2.

A1   Results

We illustrate the distributions of these quantities in Figure A1. In the following subsections, we present the offset between the centroids and the misalignments between ellipticity magnitudes and orientations.

A1.1   Centroid

We find a root-mean-square (rms) scatter of 67 ± 5 mas in the centroid offset. For our fiducial cosmology, the 68 per cent and 95 per cent upper limits of the absolute offsets are 218 ± 19 pc and 426 ± 68 pc, respectively (Figure A1 first panel).

A1.2   Ellipticity

We find a weak correlation between the axis ratios of the mass and light distributions with a biweight midcorrelation of $r = 0.54 \pm 0.08$ (Figure A1 middle panel). Biweight midcorrelation is similar to the Pearson's r coefficient. However, it depends on the median instead of the mean, thus it is more robust against outliers (e.g., Beers et al. 1990). We get the uncertainty in the midcorrelation by sampling the axis ratios of all the lenses 1000 times from their posterior distributions.

We check if the lens systems are double-image or quadruple-image systems by checking the number of images produced by the lens models for a hypothetical point source located at the source centers. We find 19 out of 23 lenses to be double-image systems and the remaining 4 to be quadruple-image systems. The majority of doubles explain the lower ellipticity and shear distributions observed in the SLACS sample compared to that of the quadruply lensed quasars (Figure A1).

A1.3   Position angle

While comparing the position angles of the mass and light, we ignore 4 lenses with $q_m > 0.9$ and $q_L > 0.9$, as the estimate of the position angle in low-ellipticity cases can be unreliable. Out of the remaining 19 lenses, we find 16 to have misalignment angle $\Delta PA \leq 12$ deg. The three systems with $\Delta PA > 12$ deg have relatively higher external shear $\gamma_{ext} \gtrsim 0.05$ (Figure A1 right panel).

A2   Discussion

We find moderate correlation between axis ratios of the mass and light (biweight midcorrelation $r = 0.55 \pm 0.09$). Although Gavazzi et al. (2012) and Sluse et al. (2012) both report strong correlation between mass and light ellipticities, we compute the biweight midcorrelation from the values reported by these authors to find $r = 0.26 \pm 0.07$ (weak correlation) from the galaxy–galaxy lenses of Gavazzi et al. (2012) and $r = 0.91 \pm 0.02$ (very strong correlation) from the lensed quasars of Sluse et al. (2012). Additionally, Kostrzewa-Rutkowska et al. (2014) find strong correlation ($r = 0.77 \pm 0.16$) from a sample of 9 galaxy–galaxy lenses from the Cambridge And Sloan Survey Of Wide ARcs in the skY (CASSOWARY). In contrast, Shajib et al. (2018) find very weak to no correlation between mass and light for a sample of 13 quadruply lensed quasars ($r = 0.13 \pm 0.10$). Similarly, Keeton et al. (1998), Ferreras et al. (2008), and Rusu et al. (2016) also find no correlation between the mass and light ellipticities. Some of the differences between these studies, including ours, can be due to the differences in the data quality, analysis techniques, and adopted models. Some of these differences can also be due to the selection function, for example between the galaxy–galaxy lenses and lensed quasars. SLACS is deflector-selected and mostly comprises doubly lensed objects, whereas quadruply lensed quasars are expected to be strongly selected to favor big inner caustics (Dobler et al. 2008; Collett & Cunnington 2016). Therefore, they favor high ellipticity and shear. Moreover, the SLACS sample has a smaller average redshift than the sample of quasar lenses, e.g., from Shajib et al. (2019). Thus, the line-of-sight effect is much less important for SLACS than for the lensed quasars. As Shajib et al. (2019) model their quasar lens sample using *HST* imaging of comparable quality, similar models, and the same modeling software as in this paper, the observed differences between the SLACS sample and the quasar lens sample in Figure A1 have to arise from the differences in the selection functions.

We find that in most galaxies with moderate ellipticity ($q \leq 0.9$), the major axes of the mass and light distributions are well aligned within ± 12 deg. The lenses with large misalignment ($\Delta PA > 12$ deg) also have relatively larger ($\gamma_{ext} \gtrsim 0.05$) external shear within the sample. This result is consistent with previous studies (Kochanek 2002; Ferreras et al. 2008; Treu et al. 2009; Gavazzi et al. 2012; Sluse et al. 2012; Bruderer et al. 2016; Shajib et al. 2018). The absence of systems with large misalignment angle and low external shear is consistent with the prediction of galaxy formation models that highly misaligned orbits in isolated galaxies are unstable and thus rare (e.g., Adams et al. 2007; Debattista et al. 2015). Only in blue starburst galaxies – unlike the galaxies in our sample – constant gas-flow can sustain highly misaligned orbits (Debattista et al. 2015).

APPENDIX B: ALGORITHM TO MASK THE LENSED ARCS

We illustrate the algorithm for creating the mask for the lensed arcs in Figure B1.

APPENDIX C: SELECTED SLACS GALAXIES FOR MODELLING

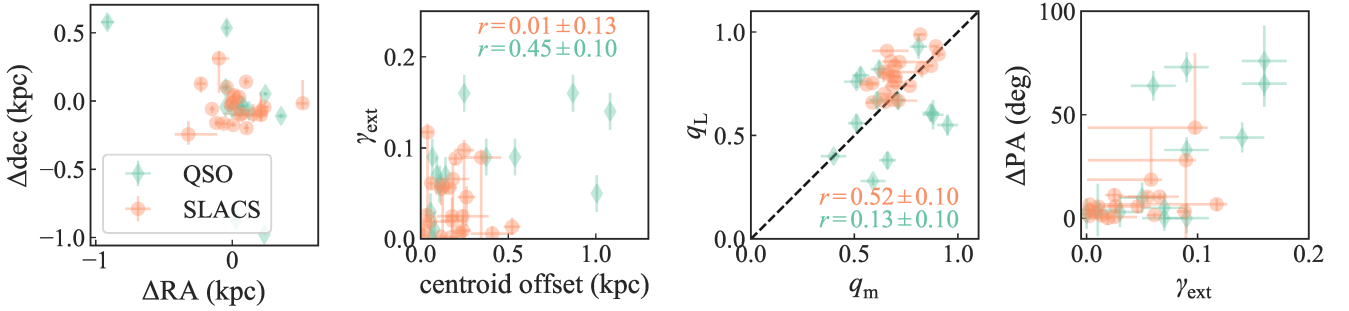We provide the list of selected galaxies in Table C1.

Figure A1. The alignment between mass and light distribution for SLACS lenses in our sample (orange) and quasar lenses from Shajib et al. (2019, green). **First panel:** The angular offset between the centroids of mass and light. For our fiducial cosmology, the offset distribution for the SLACS lenses has a 95 per cent upper limit of 426±68 pc. **Second panel:** Distribution of the absolute centroid offsets and external shear. These two quantities are moderately correlated for the quasar lenses from Shajib et al. (2019) with biweight midcorrelation $r = 0.45 \pm 0.10$, which indicates that the presence of nearby perturbers may cause an offset between mass and light centroids in the quasar lenses. However, there is no such correlation for the SLACS lenses. Although the shear distribution for the SLACS lenses is at a lower range than that of the quasar lenses, thus any such potential correlation is hard to identify given the measurement uncertainties. **Third panel:** The distribution of mass and light axis ratios. The axis ratios are weakly correlated with biweight midcorrelation $r = 0.54 \pm 0.08$. **Fourth panel:** The distribution of major axis misalignment between mass and light and external shear. The 19 lenses from our sample of 23 that have $q_{\rm L} < 0.9$ and $q_{\rm m} < 0.9$ – and thus reliable position angle estimates – are plotted in this panel. For these SLACS lenses, misalignment angle ΔPA> 12 deg are the systems with $\gamma_{\rm ext} \gtrsim 0.05$.

This paper has been typeset from a TeX/LaTeX file prepared by the author.

(1) Take pixel-wise radial gradient outward from the center of the deflector. The pixels with positive gradient (red) mark the inner edge of the arcs.

(2) Make a binary image of the gradient with positive values set to 1 (white) and negative values set to 0 (black). Set all the pixels within the inner 0.4 arcsecond to 0.

(3) Set connected white regions with area below 5 pixels to zero. This procedure removes small white regions near the arc that were created due to noise.

(4) Dilate the binary image separately in the four quadrants with these structural elements. The size of the structural element matrices is 7×7.

(5) The dilation grows the white regions radially outward. Thus, the white regions can span over the whole width of the lensed arcs.

(6) Invert the binary image to create the mask for the lensed arcs. The figure on the left shows the lens image with the lensed arcs masked.
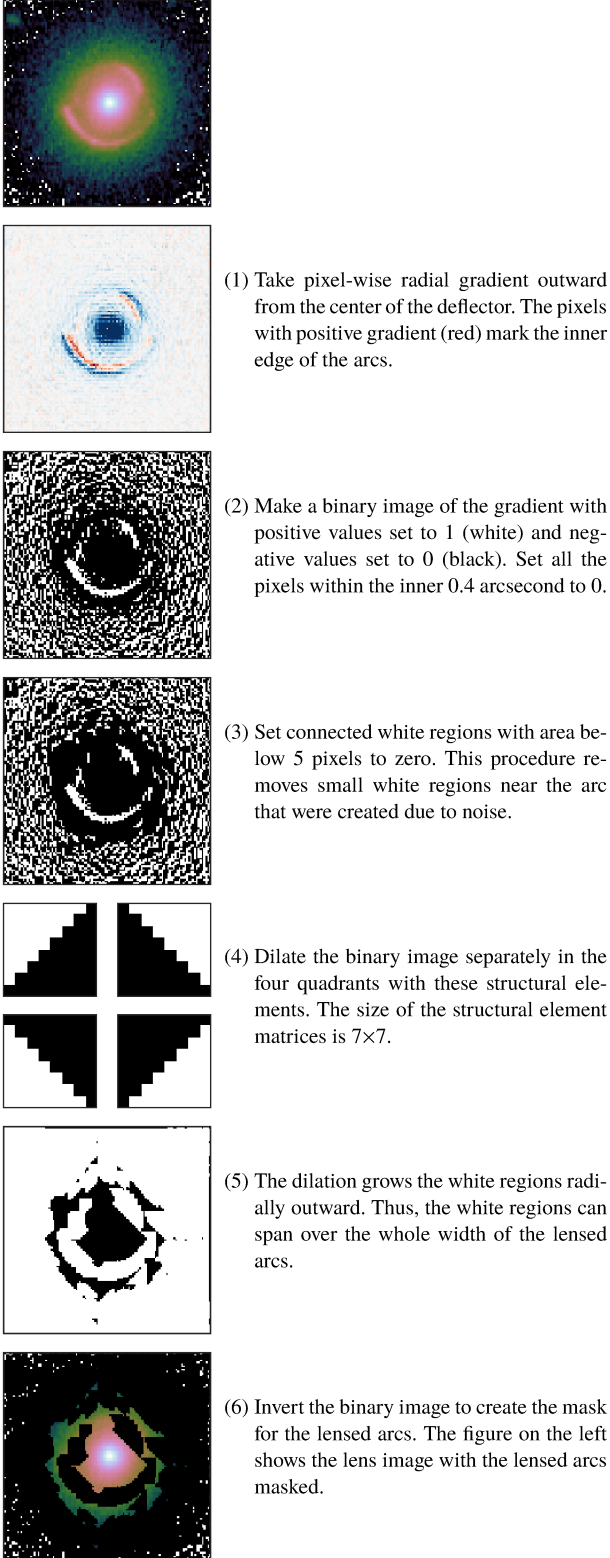
Figure B1. The algorithm to mask the lensed arcs. We use this mask to robustly fit the lens light profile only at one particular step within our model fitting procedure. The top figure shows an example lens to illustrate the procedure. All the subsequent figures illustrate different steps of the algorithm. The bottom figure shows the same lens image from the top figure with the lensed arcs masked. We used this mask to only fit the deflector's light distribution during an initial step of the optimization procedure.

Table C1. List of SLACS lenses selected for modelling following the criteria described in Section 2. The 'Modelling success' column states whether we achieved a good quality model or not from our automated and uniform modelling approach.

Name	*HST* camera and filter	Modelling success
J0008−0004	WFPC2 F606W	No
J0029−0055	WFPC2 F606W	Yes
J0037−0942	WFPC2 F606W	Yes
J0044+0113	WFPC2 F606W	No
J0252+0039	WFPC2 F606W	Yes
J0330−0020	WFPC2 F606W	Yes
J0728+3835	WFPC2 F606W	Yes
J0737+3216	ACS F555W	Yes
J0819+4534	WFPC2 F606W	No
J0903+4116	WFPC2 F606W	Yes
J0912+0029	ACS F555W	No
J0935−0003	WFPC2 F606W	No
J0936+0913	WFPC2 F606W	No
J0959+0410	ACS F555W	Yes
J0959+4416	WFPC2 F606W	No
J1016+3859	WFPC2 F606W	No
J1020+1122	WFPC2 F606W	No
J1023+4230	WFPC2 F606W	No
J1100+5329	WFPC2 F606W	No
J1112+0826	WFPC2 F606W	Yes
J1134+6027	WFPC2 F606W	No
J1142+1001	WFPC2 F606W	No
J1143−0144	ACS F555W	No
J1153+4612	WFPC2 F606W	No
J1204+0358	WFPC2 F606W	Yes
J1213+6708	WFPC2 F606W	No
J1218+0830	WFPC2 F606W	No
J1250+0523	ACS F555W	Yes
J1306+0600	WFPC2 F606W	Yes
J1313+4615	WFPC2 F606W	Yes
J1319+1504	WFPC2 F606W	No
J1402+6321	ACS F555W	Yes
J1403+0006	WFPC2 F606W	No
J1432+6317	WFPC2 F606W	No
J1531−0105	WFPC2 F606W	Yes
J1538+5817	WFPC2 F606W	No
J1614+4522	WFPC2 F606W	No
J1621+3931	WFPC2 F606W	Yes
J1627−0053	ACS F555W	Yes
J1630+4520	ACS F555W	Yes
J1636+4707	WFPC2 F606W	Yes
J1644+2625	WFPC2 F606W	No
J2238−0754	ACS F555W	Yes
J2300+0022	ACS F555W	Yes
J2302−0840	WFPC2 F606W	No
J2303+1422	ACS F555W	Yes
J2321−0939	WFPC2 F606W	No
J2341+0000	WFPC2 F606W	No
J2343−0030	WFPC2 F606W	Yes
J2347−0005	WFPC2 F606W	No