Modelling strong lenses from wide-field ground-based observations in KiDS and GAMA

Knabel, S.; Holwerda, B.W.; Nightingale, J.; Treu, T.; Bilicki, M.; Brough, S.; ... ; Wright, A.H.

Modelling strong lenses from wide-field ground-based observations in KiDS and GAMA

Shawn Knabel [1,2]★ B. W. Holwerda [1] J. Nightingale [3] T. Treu [2] M. Bilicki [4] S. Brough [5]
S. Driver [6] L. Finnerty [2] L. Haberzettl [1] S. Hegde [2] A. M. Hopkins [7] K. Kuijken [8] J. Liske [9]
A. K. Pimblett [10] R. C. Steele [1] and A. H. Wright [11]

[1] *University of Louisville, Department of Physics and Astronomy, 102 Natural Science Building, 40292 KY Louisville, USA*
[2] *University of California-Los Angeles, Department of Physics and Astronomy, PAB, 430 Portola Plaza, Box 951547, Los Angeles, CA 90095-1547, USA*
[3] *Durham University, Institute for Computational Cosmology, Stockton Rd, Durham DH1 3LE, UK*
[4] *Center for Theoretical Physics, Polish Academy of Sciences, al. Lotników 32/46, 02-668 Warsaw, Poland*
[5] *School of Physics, University of New South Wales, NSW 2052, Australia*
[6] *International Centre for Radio Astronomy Research (ICRAR), University of Western Australia, Crawley, WA 6009, Australia*
[7] *Australian Astronomical Optics, Macquarie University, 105 Delhi Rd, North Ryde, NSW 2113, Australia*
[8] *Leiden Observatory, Leiden University, PO Box 9513, NL-2300RA Leiden, the Netherlands*
[9] *Universität Hamburg, Hamburg Sternwarte, Gojenbergsweg 112, D-21029 Hamburg, Germany*
[10] *E. A. Milne Centre for Astrophysics, University of Hull, Cottingham Road, Kingston-upon-Hull HU6 7RX, UK*
[11] *German Center for Cosmological Lensing (GCCL), Astronomisches Institut, Ruhr-Universität Bochum, Universitätsstraße 150, D-44780 Bochum, Germany*

ABSTRACT

Despite the success of galaxy-scale strong gravitational lens studies with Hubble-quality imaging, a number of well-studied strong lenses remains small. As a result, robust comparisons of the lens models to theoretical predictions are difficult. This motivates our application of automated Bayesian lens modelling methods to observations from public data releases of overlapping large ground-based imaging and spectroscopic surveys: Kilo-Degree Survey (KiDS) and Galaxy and Mass Assembly (GAMA), respectively. We use the open-source lens modelling software PYAUTOLENS to perform our analysis. We demonstrate the feasibility of strong lens modelling with large-survey data at lower resolution as a complementary avenue to studies that utilize more time-consuming and expensive observations of individual lenses at higher resolution. We discuss advantages and challenges, with special consideration given to determining background source redshifts from single-aperture spectra and to disentangling foreground lens and background source light. High uncertainties in the best-fitting parameters for the models due to the limits of optical resolution in ground-based observatories and the small sample size can be improved with future study. We give broadly applicable recommendations for future efforts, and with proper application, this approach could yield measurements in the quantities needed for robust statistical inference.

Key words: gravitational lensing: strong – methods: observational – galaxies: elliptical and lenticular, cD – galaxies: evolution – galaxies: structure.

1 INTRODUCTION

Strong gravitational lensing is an essential probe of galaxy structure, enabling mass measurements in the centre-most regions of the foreground lensing galaxy without assumptions about stellar populations. Numerous studies have shown that lensing galaxies are, in every other respect, indistinguishable from other galaxies in the observed mass range; therefore, their study offers insight into the larger global population of galaxies at similar mass and redshift (Auger et al. 2009). Complementary to kinematic and stellar population synthesis measurements, strong lensing allows the decoupling of internal mass components (dark and baryonic) and accurate central stellar population mass-to-light ratios (Auger et al. 2010b; Hopkins 2018).

A fundamental issue in astronomy is relating the predicted dark matter halo mass function to the observed galaxy mass function. The masses of dark matter haloes are not well constrained by the amount of visible matter in their constituent galaxies. Few single galaxies corresponding to the lowest and highest halo masses are found, due to feedback processes stopping star formation (Behroozi , Conroy & Wechsler 2010; Behroozi et al 2020). The galaxy stellar-to-halo mass relation represents a fundamental barometer for accretion and feedback processes in galaxy formation. Subhalo abundance matching assigns a galaxy with a specific stellar mass to a specific subhalo but does not consider (i) the enveloping host halo mass or (ii) whether the galaxy is a central or a satellite; thus it suffers from *assembly bias* (Zentner, Hearin & van den Bosch 2014; Chaves-Montero et al. 2016).

Assembly bias is a secondary halo property that is related to the clustering strength of haloes (Matthee et al. 2017; Zehavi et al. 2018),

★ E-mail: shawnknabel@gmail.com

where the clustering of dark matter haloes depends on their mass and formation epoch. Investigations with cosmological simulations have revealed that dark matter halo concentration, formation time, and environment all play a role in the relation between the central galaxy's stellar mass and the mass of the dark matter halo it occupies. Assembly bias appears to be mostly independent of the cosmological parameters assumed (Contreras et al. 2021). At high masses typical of lensing elliptical galaxies, inefficiency in the stellar occupancy of dark matter haloes is ascribed to the effects of active galactic nuclei (Somerville, Popping & Trager 2015). Weak-lensing studies (Mandelbaum et al. 2013; Velander et al. 2014; Lin et al. 2016) and velocity field studies (McCarthy et al. 2021; Posti & Fall 2021) appear to show the effects of assembly bias on the scales of galaxy populations (Cui et al. 2021); assembly bias is especially noticeable at ∼1 Mpc scales (Hearin et al. 2016): i.e. groups of galaxies. While well-explored with hydrodynamical simulations (e.g. Hearin, Watson & van den Bosch 2015; Hearin et al. 2016; Matthee et al. 2017; Artale et al. 2018; Zehavi et al. 2018, 2019; Contreras et al. 2019), observational studies have thus far been limited by the need to average over large numbers of similar-mass central elliptical galaxies to obtain a weak lensing or velocity signal.

With strong lensing, one has the opportunity to directly measure stellar and halo masses in elliptical galaxies. Relations between the environment and internal structure of elliptical galaxies have been explored using Sloan Lens ACS (SLACS; Bolton et al. 2006) by Treu et al. (2010). They find that the SLACS lenses are slightly biased towards overdense environments (12 of 70 are associated with known groups or clusters), which is consistent with the expectation for the most massive of elliptical galaxies. They find this result to be unbiased when compared to similar massive galaxies from SDSS, again showing lens galaxies to be representative of the overall elliptical galaxy population. They find the contribution of the external environment to have little effect on the local potential (except in extreme overdensities) and the internal structure of lens galaxies. SLACS and other lens studies have been conducted using detailed observations of individual lenses with HST-quality data. The application of lens modelling methods to larger wide-field surveys offers an alternative avenue with advantages for conducting experiments relating galaxy properties to environment and group properties. This motivates the need for exploring lens modelling methods in the context of large surveys.

In this paper, we explore what can be done with ground-based imaging and spectroscopy to model lens candidates after they have been identified in imaging surveys. We discuss strategies for ensuring quality control at each stage while extracting meaningful measurements from ground-based data. Using Galaxy and Mass Assembly (GAMA) (Driver et al. 2009, 2011; Liske et al. 2015) survey single-aperture spectroscopy, we explore the utility of automated redshift determination as a tool for identifying the background-source redshifts of strong lenses by applying this technique to lens candidates that were identified in the ground-based imaging of the Kilo-Degree Survey (KiDS) (de Jong et al. 2013, 2015, 2017; Kuijken et al. 2019) using machine learning techniques (Petrillo et al. 2019b; Li et al. 2020). With GAMA spectroscopic redshifts and other measurements in conjuction with KiDS cutout images, we construct lens models utilizing an automated lens modelling program called PYAUTOLENS (Nightingale et al. 2021b).

Our paper is organized as follows: Section 2 describes the KiDS and GAMA data used, as well as the parent samples used in our selection. Section 3 describes how background-source redshifts are identified in single-aperture spectra from GAMA AUTOZ catalogs to create a subsample for modelling. Section 4 describes the PYAU-

TOLENS software and the lens modelling methods we used to perform our analysis. Section 5 outlines the assessment of quality of the models and redshift determinations. Section 6 presents results for the four highest-quality models. Section 7 discusses some challenges that our prescription of second-redshift determination introduced, as well as recommendations for improving that method. Section 8 discusses galaxy environment and potential applications of a refined method to future studies. Section 9 lists our conclusions. Throughout this paper, we adopt a Planck Collaboration 2015 (2016) cosmology ($H_0 = 67.7 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$, $\Omega_m = 0.307$).

2 DATA

2.1 GAMA spectroscopy and AUTOZ redshifts

GAMA (Driver et al. 2009, 2011; Liske et al. 2015) is a multiwavelength survey built around a deep and highly complete redshift survey of five fields with the Anglo-Australian Telescope. GAMA has three major advantages over SDSS in the identification of blended spectra: (i) the spectroscopic limiting depth is 2 mag deeper ($m_r < 19.8$ mag compared with SDSS main survey depth $m_r < 17.7$ (Eisenstein et al. 2001)[1]), (ii) the completeness is close to 98 per cent (Liske et al. 2015), and (iii) the AUTOZ redshift algorithm can identify spectral template matches with signal from two different redshifts (Baldry et al. 2014). These properties and the overlap of GAMA and KiDS fields make these two surveys exceptionally well suited to provide the data required for our study of lens modelling.

The AUTOZ (Baldry et al. 2014) cross-correlation redshift software has been uniformly applied to the GAMA (Liske et al. 2015) spectroscopic data, resulting in a public data base that can be found in GAMA-DR3 AATSPECAUTOZALL V27 (hereafter AUTOZ catalog) and SPECALL V27 tables (Liske et al. 2015, http://www.gama-survey.org/dr3/). The AUTOZ algorithm outputs four flux-weighted cross-correlation peaks (denoted σ) of redshift matches to template spectra of emission line and passive galaxies (denoted ELG and PG, respectively) from SDSS-DR5. σ_1 corresponds to the highest cross-correlation or 'best-fitting' redshift match, σ_2 the second-best, etc. These matches have proven to be highly successful and are the base redshift measurement for GAMA objects. GAMA-DR3 also compiled SDSS-BOSS spectra for overlapping targets that are included in table SPECALL, but these spectra did not utilize AUTOZ for redshift determination.

Holwerda et al. (2015) analysed AUTOZ catalog cross-correlation outputs and identified 104 strong lensing candidates from their blended spectra, all of which showed a PG with an ELG at higher redshift between cross-correlation σ_1 and σ_2. This identification selected candidates from a two-dimensional parameter space defined by the second cross-correlation peak σ_2 and the parameter R, which describes the significance of σ_2 compared to the following 'poorer' matches:

$$R = \frac{\sigma_2}{\sqrt{\frac{\sigma_3^2}{2} + \frac{\sigma_4^2}{2}}} \tag{1}$$

Candidates with second cross-correlation peak, $\sigma_2 \geq 4.5$ and $R \geq 1.85$, were considered likely candidates for strong lensing. Knabel et al. (2020) further analysed and made a cleaner selection of 47 candidates.

---

[1]The spectroscopic LRG sample used to select SLACS lenses is limited to $m_r < 19.5$ thanks to the 4000 Å break.

The completeness of GAMA allows detailed environment measures, including population density and separation (Brough 2011; Alpaslan et al. 2014, 2015), and the GAMA team internal GROUPFINDING catalogs include the total mass and placement of the galaxies in an identified group via a friends-of-friends algorithm (Robotham et al. 2011). In fact, GAMA was conceived to probe the effects of group environment on galaxy properties. In this context, we describe galaxies either as 'group member' galaxies or as those not in galaxy groups, which we designate as 'isolated galaxies'. Stellar masses are taken from the GAMA-DR3 STELLARMASSESLAMBDAR v20 catalog (Taylor et al. 2011).

2.2 Kilo-Degree Survey and machine learning strong lens samples

The KiDS (de Jong et al. 2013, 2015, 2017; Kuijken et al. 2019) is a *VLT Survey Telescope (VST)* program of medium-deep imaging in SDSS-*ugri* filters primarily to identify weak lensing. The deep imaging, high resolution (0.65 arcsec seeing in SDSS r-band), and wide sky-coverage (1350 deg^2) also make this survey ideal for efforts to identify strong lens candidates from imaging. Image-based deep learning efforts have been the most promising of recent developments in automated lens-finding algorithms. Their efficiency and versatility make them ideal for astronomical classification problems involving large data sets, including the detection of strong gravitational lenses [e.g. in Subaru Hyper-Supreme Cam (Speagle et al. 2019), DESI DECam Legacy Survey (Huang et al. 2020), and Dark Energy Survey data (Jacobs et al. 2019)]. Petrillo et al. (2017) developed a machine learning method to identify strong lenses in KiDS using a convolutional neural network (CNN) with artificially constructed lens images as the training set. Training and target catalogs were intentionally constructed utilizing SDSS-luminous red galaxy (LRG; Petrillo et al. 2017) colour-magnitude selection cuts to return the largest of known strong lenses that result in the most readily identifiable lens features (Einstein radii close to and greater than 1 arcsec). The result is the Lenses in KiDS sample (LinKS; Petrillo et al. 2019a, b)^2 of some 1300 strong lensing candidates, 421 of which overlap with the equatorial regions of the GAMA survey. Knabel et al. (2020) compared data from LinKS objects with the AUTOZ spectroscopic identifications in GAMA (Holwerda et al. 2015) as well as with KiDS-GalaxyZoo (Holwerda et al. 2019, Kelvin et al. in preparation) citizen science identifications in the overlapping equatorial fields. A disparity between the candidate samples in terms of stellar mass and redshift is attributed to selection effects. Of the subsample of 421 LinKS candidates in GAMA (hereafter referred to as 'LinKS' or 'LinKS in GAMA' sample), there was no overlap with the subsample of 47 GAMA spectroscopic candidates. Knabel et al. (2020) further subselected 47 LinKS candidates to represent the highest quality of the sample (hereafter referred to as 'LinKS from Knabel-2020' sample).

Li et al. (2020) followed a slightly modified approach from the lens-search prescription utilized by the LinKS team to search for strong lens candidates in KiDS-DR4. They included several more LRGs and applied their CNN to a sample of 'bright galaxies' (BG) that did not undergo LRG colour-magnitude cuts. Their search returned some LinKS candidates and resulted in 286 new candidates within the KiDS survey, 48 of which were identified in the GAMA equatorial regions. 39 of those have matches in the STELLARMASSESLAMBDAR mass catalog, and there are no overlaps
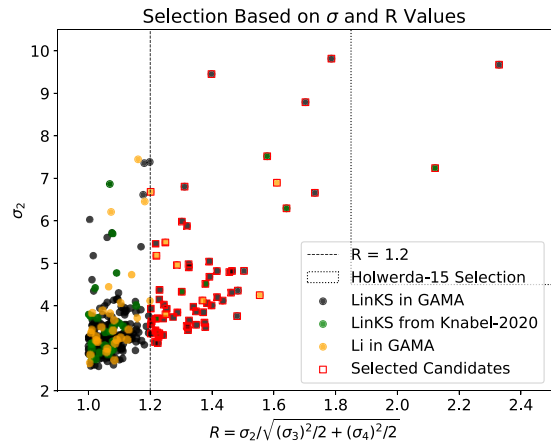
^{2}https://www.astro.rug.nl/lensesinkids/

Figure 1. Initial selection of candidates with AUTOZ second-redshift determinations. The *y*-axis shows the σ_2 second cross-correlation peak, with higher values indicating a stronger match to the second galaxy template. The *x*-axis is the parameter *R*, given by equation (1). Black markers indicate 348 LinKS AUTOZ entries (300 unique candidates), 51 of which are overplotted with green markers to indicate AUTOZ entries of the 47 unique LinKS candidates as selected in Knabel et al. (2020). Orange markers indicate the 53 AUTOZ entries for the 32 unique Li-BG candidates. The dashed vertical line shows $R = 1.2$, and the dotted box encloses the area of parameter space used by Holwerda et al. (2015).
Red squares surround 59 LinKS and 8 Li-BG entries that satisfy $R \geq 1.2$ and are followed with additional selection criteria. See Section 3.1.

between this sample and that of GAMA spectroscopy or GalaxyZoo. This candidate sample, which we will refer to as 'Li-BG', shares essentially the same parameter space as LinKS candidates, even with the exclusion of the LRG selection (Knabel et al. 2020). This is not necessarily surprising considering Petrillo et al. (2019b) report no significant advantage to the inclusion of colour images in the CNNs, as the networks appear to focus more on morphological features and brightness than colour separation. Still, the extension beyond the typical red elliptical galaxy as candidate objects suggests the potential for more variability in candidate characteristics.

3 AUTOZ SECOND REDSHIFT SELECTION AND QUALITY CONTROL

We examine the AUTOZ cross-correlation values for each LinKS and Li-BG lens candidate. Each candidate has been matched to the closest GAMA object by right-ascension and declination within a positional tolerance of 2 arcsec. Not every object in the equatorial fields is featured in the AUTOZ catalog; these candidates are removed from this study. Some of the objects feature duplicated entries, some of which have conflicting AUTOZ outputs, which we retain for examination and selection.

3.1 Selection criteria

Since the candidates that remain have already been identified and vetted through machine learning methods, we adopt a more lenient selection criterion from the same $\sigma_2 - R$ parameter space as that utilized in Holwerda et al. (2015). From a first look at the data, we select candidates with $R \geq 1.2$. This is sufficient for a first selection and for characterizing the output of the AUTOZ algorithm from already positively identified candidates. We show the selection in Fig. 1.

From the 67 AUTOZ entries that pass the *R* selection, we remove those with stellar template matches (i.e. not a galaxy
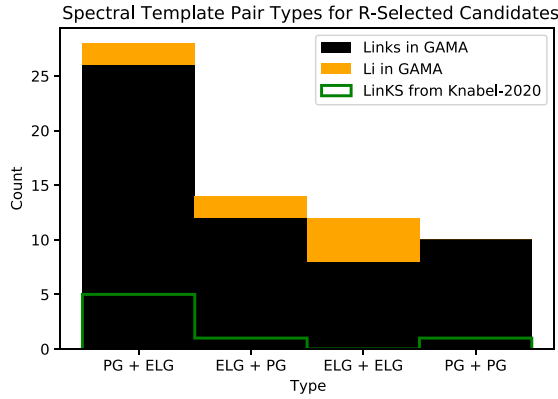
Figure 2. AUTOZ galaxy template combinations of candidate subsamples, listed as foreground + background. Black and orange refer to LinKS and Li-BG candidates, respectively, and are stacked to show total counts of each template combination. The green outline shows the subset of LinKS candidates that were examined in Knabel et al. (2020). As expected, the majority of lens foreground objects match to PG templates, while the majority of background objects match to ELG templates.

spectrum) and retain all those with galaxy–galaxy template match configurations regardless of the galaxy type. The distribution of foreground+background (lens+source) template type (PG + ELG, ELG + PG, ELG + ELG, PG + PG) is shown in the histogram of Fig. 2. Note that the majority of lens foreground objects match to PG templates, with the majority of background objects matching to ELG templates. This is expected. Massive elliptical galaxies tend to be the most readily observable strong lensing foreground objects, and bright emission lines from the background source are the most easily detected behind a PG continuum. This selection bias is further enhanced by the fact that the parent LinKS and Li-BG samples were both identified by CNNs trained with large elliptical galaxies. Configurations with foreground lens ELGs are possible, though much more difficult to detect. The reasons are: (i) ELGs are typically lower in mass, so the lensing is less pronounced, (ii) lensed background sources are often bright ELGs, so the blue light of each can blur together, and (iii) ELGs can include complex morphologies that can be mistaken for lens features. The majority of Li-BG candidates in the AUTOZ catalog were matched to ELGs in the foreground, but further selection and assessment of the spectra showed this to be a false trend. As in Knabel et al. (2020), we select candidates with background source redshifts that are reasonably far away from the foreground lens redshifts, as well as those whose AUTOZ foreground redshifts are greater than 0.05. AUTOZ estimates the probability of success of the primary redshift match, and one candidate is removed due to its very low probability.

Described in more detail and discussed in the context of Knabel et al. (2020) in Appendix E, our selection results in 42 lens candidates (39 from LinKS and 3 from Li-BG) with second redshift matches, which constitute the initial AUTOZ sample. In later work, it may be worth exploring machine learning algorithms to find better use of the parameter space for classification than our naive selection criteria.

3.2 AUTOZ selection quality control

An initial modelling and analysis strategy utilizing the AUTOZ sample of 42 candidates and PYAUTOLENS revealed the need for further quality control. This assessment addresses the validity of our application of AUTOZ output parameters to select second-redshift determinations for the use in strong lens modelling. We examined

the 42 candidate spectra to understand the evidence for a second redshift match within each. We superimposed upon the candidate spectra important emission and absorption features redshifted by the lens and source redshifts determined by AUTOZ. For ELG template matches, we inspected $H\beta$, [O II] $\lambda 3727$, and [O III] $\lambda\lambda 4959,5007$ emission lines. For PG template matches, we looked at Calcium H and K, $H\beta$, Mg, and Na absorption lines. This illuminated some specific cases that pointed to dubious redshift matches. We identify four such cases:

3.2.1 When there are overlapping emission or absorption lines...

AUTOZ template matches utilize strong absorption or emission features to identify passive and star-forming galaxies. We discovered a significant failure condition of our method that occurs when the second redshift match is identified by an emission or absorption line that is also attributed to the foreground lens galaxy redshift. This is particularly obvious in cases where AUTOZ determined ELG + ELG configurations (i.e. both redshifts are identified by emission lines), in which many of the higher-redshift matches were determined by an [O II] $\lambda 3727$ line that overlapped with the [O III] $\lambda\lambda 4959,5007$ lines of the foreground lens galaxy. For configurations where both the lens and source are the same template type (i.e. PG+PG or ELG + ELG), overlapping line features usually indicate a poor second-redshift determination.

However, emission and absorption features are present in the spectra of both PGs and ELGs. For example, Treu et al. (2002) found that ~ 10 per cent of elliptical galaxies in the intermediate redshift range where our lens candidates lie have strong [O II] $\lambda 3727$ emission; in fact, the emission of [O II] $\lambda 3727$ is detectable in the spectra 14 of the PG matches in our sample, and absorption of Calcium H and K lines (as well as some Na and Mg) are identifiable in 15 of the ELG galaxy spectrum matches. In most cases, the overlap of lines draws suspicion of the background source spectrum match, not the foreground lens match, which is typically the primary template match (σ_1). In our sample, emission line overlaps occur exclusively between the background source [O II] $\lambda 3727$ and one of the [O III] $\lambda\lambda 4959,5007$ couplet of the foreground lens PG. Absorption line overlaps are all between source H or K lines and lens $H\beta$, Na, or Mg lines.

3.2.2 When the lens is described as an emission line galaxy...

As shown in Fig. 2, several of the template matches selected by the strategy described in Section 3.1 are configurations with a foreground lens ELG (ELG + ELG or ELG + PG). We reiterate that there is no physical reason to mistrust these redshift matches on this fact alone. In fact, given that several of the ELG matches have strong absorption features at the same redshift, these may very well be elliptical galaxies with strong oxygen emission lines. However, upon examination of candidate spectra and the quality of initial models for these ELG + lens configurations, we found that several of these candidates included dubious template matches and were removed.

The overlapping ELG+ELG emission lines have been discussed. In addition, some of the spectra of ELG+ configurations included key emission line features that would have been observed in wavelength ranges of high noise. Many of the GAMA spectra have their most significant noise at the extremes of the wavelength range, e.g. at wavelengths shorter than ~ 4500 Å. For lenses at redshifts lower than ~ 0.2, the [O II] $\lambda 3727$ Å emission line lies in this region, which removes an important identifying emission line feature from consideration.

3.2.3 *When source emission lines are redshifted beyond observed wavelength range...*

The cases where a background source emission line overlaps with a foreground lens emission line often correlate with another case: some of the background source emission lines have redshifted to wavelengths beyond the observed wavelength range of the spectroscopic survey. For our purposes, this translates to upper limits of background source redshifts beyond which the emission line will not be present in the spectrum. For GAMA, which has an upper limit of 8850 Å, the emission lines we have discussed begin to disappear around $z \sim 0.77$. SDSS-BOSS has an extended upper limit of wavelength to 10 400 Å ($\sim$J-band), which corresponds to upper redshift limits of around $z \sim 1$ where we begin to see missing emission lines. GAMA-DR3 SPECALL table contains SDSS-BOSS spectra for several of the candidates, so we can look to these spectra for evidence of lines that have redshifted beyond the GAMA upper limit. The AUTOZ catalog includes only GAMA spectra, so the outputs we use for selection would not benefit from the extended range of the SDSS-BOSS spectra.

3.2.4 *When primary redshift is background source...*

The primary redshift match corresponding to σ_1 is typically (but not always) the foreground lens galaxy. The background source redshift is the primary redshift match for 10 of the 42 AUTOZ sample candidates. The results of the AUTOZ algorithm are largely flux-weighted, and an especially bright background source (e.g. a strong emission line) could be interpreted by the algorithm as the primary redshift solution instead of the lower-redshift lens object. However, 8 of these 10 are ELG + PG template configurations, where a PG template gives the primary redshift solution at higher redshift. The case of a bright continuum at higher redshift overshadowing the foreground ELG is unlikely.

All candidates in the AUTOZ sample were modelled and examined in the manner described in Sections 4–6 before these four cases were identified. In Section 7, we discuss the results of modelling and assessment in the context of the cases described in this section and recommend alterations to the initial selection scheme. We note that a poor redshift match does not remove/negate the validity of the lens identification, nor does it question the accuracy of the uniform application of AUTOZ to GAMA spectroscopic targets. These additional selection decisions were instituted following critical assessment of problems with our initial strategy that required time with human eyes on the spectra. With reasonable background source redshift determinations, modelling of the imaging data can yield meaningful physical measurements.

4 PYAUTOLENS

We use the open-source lens modelling software PYAUTOLENS (Nightingale et al. 2021b)[3] to perform our analysis. The software is described in Nightingale, Dye & Massey (2018) and Nightingale et al. (2021b), building on the works of Warren & Dye (2003), Suyu et al. (2006), and Nightingale & Dye (2015). We refer readers to these works for a full description of our approach to lens modelling. Section 4.1 broadly describes the method as we apply it here, and a more technical description of the specifics of the implementation is given in Appendices A and B.

4.1 Lens modelling with pyautolens

PYAUTOLENS models the foreground lens galaxy's light and mass as well as the background source galaxy's light simultaneously. First, PYAUTOLENS assumes a profile for the foreground lens's light (e.g. a Sérsic profile), producing a model image of the lens galaxy. A mass model then ray-traces a grid of image-pixels from the image-plane to the source-plane, with the source's light evaluated on this deflected grid via another light profile. This creates an image of the lensed source, which is added to the lens galaxy image to create an overall model image of the strong lens. This image is convolved with the instrument point spread function (PSF) and compared to the data to evaluate the residuals and likelihood of that lens model. To fit a lens model to imaging data, PYAUTOLENS searches an N-dimensional parameter space so as to minimize the residuals (and therefore maximize the likelihood) between the model image and the observed image. The lens models fitted in this work consist of $N = 7$–14 parameters, corresponding to the parameters of the light and mass profiles that represent the lens and source galaxies. To sample parameter space, we use the nested sampling algorithm DYNESTY (Speagle et al. 2019), and we detail its specific implementation below.

The parameter spaces of a strong lens model are challenging to sample, and local maxima and unphysical lens models are often inferred. Automating the model-fitting procedure is therefore difficult, and PYAUTOLENS approaches automation via a technique called non-linear search chaining. Here, a sequence of DYNESTY model-fits are performed that fit lens models of gradually increasing complexity, whereby the results of the initial searches are used to inform the search of more complex parameter spaces in the later searches. Through experimentation, we have designed a pipeline composed of a chain of three DYNESTY searches that we use as a template for fitting each lens. Our three-step pipeline consists of three sequential model-fits:

(i) *Search 1 - Lens Light*: models only the foreground lens elliptical light profile.

(ii) *Search 2 - Lens Mass and Source Light*: focuses on the background source light profile and lensing deflections.

(iii) *Search 3 - Combined Lens and Source Models*: models each component in the system for parameter inference.

Between each search, various aspects of the fit can be altered (e.g. a mask applied to the data can be customized to show only the specific features of interest to each fit). This offers a more efficient lens modelling procedure overall, as the parameter spaces of reduced complexity are sampled faster.

Search chaining uses a technique called 'prior passing' to initialize the regions of parameter space that are searched later in the chain. Here, the models inferred in earlier non-linear searches initialize the priors of the more complex models fitted by the searches later on. This ensures the non-linear search samples only the higher likelihood regions of parameter space [see Nightingale et al. (2018)] and therefore reduces the probability that a local maximum is inferred. Prior passing sets the prior of each parameter as a Gaussian. The mean is that parameter's previous inferred median PDF value, and the width is a value specific to each lens model and parameter. Prior widths have been carefully chosen to ensure they are broad enough not to omit lens model solutions by trimming valid solutions but sufficiently narrow to ensure the lens model does not inadvertently infer local maxima.

The DYNESTY nested sampling algorithm (Speagle et al. 2019) can also balance efficiency in computation with how thoroughly it explores parameter space. Initial model fits require only a rough

---

[3]https://github.com/Jammy2211/PyAutoLens

estimate of the lens model that provides a reasonably approximate fit to the data. These searches therefore use faster DYNESTY settings that give a less thorough sampling of parameter space. Our final results require accurate and robust parameter estimates with precise and well-quantified errors. By Search 3, the priors are initialized such that a deeper exploration of the parameter space can be performed more efficiently, ensuring that DYNESTY does not spend considerable time in regions of parameter space that previous searches tell us do not give a physical lens model. Some basic settings that can be varied to affect the performance of the non-linear search are: (i) number of live points, (ii) number of steps of random walks per iteration, (iii) target acceptance fraction for random walks, (iv) Bayesian evidence tolerance, (v) positions threshold, and (vi) sub-grid size.

The technical details of our modelling method, including data preparation and pipeline, are described more fully for the interested reader in Appendices A and B. The sequence of chained searches and specific parameters that are set via prior passing are listed in Table B1, and DYNESTY settings are tabulated in Table B2. More details on PYAUTOLENS's use of DYNESTY are provided in (Nightingale et al. in preparation).

4.2 Physically motivated priors

Where possible, we calculate priors using photometric observations from GAMA-DR3 preferentially over a universally applied 'typical' value. For either case, it is important not to fix the parameters too restrictively to values from observations.

4.3 Effective radius

In search 1, we initialize the effective radius parameter of the foreground lens light profile with a Gaussian centred at the lesser of two possible radii: (i) effective radius determined from photometric observations from GAMA-DR3 SERSICCAT SDSS v09 catalog (Kelvin et al. 2012), or (ii) the median SLACS lens effective radius and standard deviation from Auger et al. (2010a; 7 ± 3.3 kpc). We expect the GAMA-DR3 observation to include extended blended light from the source feature, which may result in a higher measured effective radius than would be measured from the foreground galaxy if it were not lensing. In order to assist the search in the task of deblending the lens and source light, we ensure that the prior is not predisposed to unusually large effective radii. Another failure state of early models resulted in unrealistically large source galaxy effective radii. Instead of attributing the extended lens features to lensing of a compact background object (which is most often the case for strong lensing), the model makes up for that extra flux as the physical extent of an extremely large, bright background source galaxy at high redshift. This motivated an upper limit to the effective radius of the source galaxy based on typical disc galaxy properties. We take a rough value of 7.5 ± 2.5 kpc and upper limit of about 15 kpc.

4.3.1 Lens mass-to-light ratio

Certain critical parameters are not easily approximated with typical observations (and as such are the goal of the search), such as the stellar mass-to-light ratio. This quantity can be inferred from stellar population studies, but one of our goals is to illuminate mass relations without the dependence on these assumptions. We want the model to tell us about the stellar population as opposed

to the inverse. We want the algorithm to have the maximum freedom to determine the best combination of stellar and dark mass profiles to account for a gravitational potential that can describe the observed lensing deflections. Our first attempts allowed for a wide uniform prior distribution for the mass-to-light ratio of the stellar light-mass profile. The resulting models showed higher values than expected, some of which were unphysical in the context of predicted M_*/L from stellar population models evolving with age. The population would have to be older than the age of the Universe at the given lens redshift for the model's M_*/L to reconcile with our current models of stellar populations and evolution. To approach this problem, we impose a minimum M_*/L of 1 (M/L)$_\odot$ and a maximum determined as a function of lens redshift. We assume a Salpeter IMF and utilize stellar evolution models from (Bruzual & Charlot 2003; updated 2016) based on the STELIB spectral library (Le Borgne 2003). We determine the maximum possible rest-frame bandpass-specific M_*/L corresponding to a formation time close to the beginning of the Universe. Other libraries, BaSeL (Lastennet et al. 2002) and MILES (Sanchez-Blazquez et al. 2006, Falcon-Barroso et al. 2011), give almost identical values. Given a simple stellar population forming from a single starburst at time $t = 0$, Salpeter IMF, and solar metallicity (Z = Z$_\odot$ = 0.02, X = 0.7000, Y = 0.2800, [Fe/H] = + 0.0932), we examine the evolution of the stellar mass-to-light ratio with population age in r- and g-bands sampled at unequally spaced time-steps over 20 Gyr of stellar evolution. In our adopted cosmology, the age of the Universe in the redshift range of our sample (z $\sim$ 0.07–0.45) is about 9–13 Gyrs. At this late stage in stellar evolution, the stellar mass-to-light ratio varies slowly and is reasonably approximated as a linear relation. On the domain [9, 13 Gyrs], the constraint is a linear relation:

$$M_*/L_r[\mathrm{M_\odot/L_\odot}] < 0.466t + 0.719 \qquad (2)$$

$$M_*/L_g[\mathrm{M_\odot/L_\odot}] < 0.717t + 0.380 \qquad (3)$$

where t is the age of the Universe at the lens redshift in the adopted model cosmology.

In order to be implemented as priors in the lens models, these rest-frame constraints must be k-corrected, calibrated to the flux units in which the observed data are given and rewritten in the model mass and intensity units. We use SED-calculated k-corrections from GAMA-DR3 KCORR_AUTO_Z00 v05 (Loveday et al. 2012) for each lensing galaxy to constrain the prior in the observed bandpass. These constraints are converted to angular mass units per eps (electrons per second) with the gain and exposure time of the KiDS observation. These constraints ensure that the model does not attribute mass to a stellar population that is impossible within current stellar evolution models. In cases where the maximum possible M_*/L is fit, the maximum possible stellar mass has also been attributed. In these cases, the model may end up having to compensate with very high amounts of dark matter to account for the lensing potential.

4.3.2 NFW profile scale radius

The scale radius of a dark matter halo is one of the key parameters of the NFW mass density profile for dark matter halos. Gavazzi et al. (2007) modelled 22 SLACS lenses with strong and weak lensing constraints and a two-component mass profile consisting of a de Vaucouleurs stellar component and spherical NFW dark matter component. We adopt their resulting mean scale radius of $r_s = 58 \pm 8h^{-1}$ kpc for our dark matter profile Gaussian prior

distributions. These values are converted to arcseconds from angular diameter distances in our assumed cosmology.

4.4 Choice of image bandpass

KiDS observations of each object in different bandpasses are not equal in exposure time, signal-to-noise quality, or PSF. KiDS r-band imaging is the highest quality of the bandpasses, with an exposure time of 1800 s and a mean PSF of 0.65 arcsec. g-band exposure times are 900 s. For each candidate and for each model search, we select the better of r- or g-band images. Search 1 fits the r-band image because it most clearly shows the foreground lens light. If the lens and source are clearly distinguishable in the r-band, then the same image is used for searches 2 and 3. However, the g-band image often most clearly shows the lensed features of the background source; in these cases, the g-band is preferable for search 2. Since search 3 models the entire system, the image that most clearly shows both profiles is used.

Each search assists the subsequent searches to distinguish between the lens and source light, which is one of the more difficult challenges of modelling lenses from images of the resolution and S/N attainable by ground-based observatories. Two effective first solutions are (i) separating the initial search of foreground lens and background source light profiles by colour-band and (ii) masking specific regions of the image. In this case, the sacrifice in quality between the r- and g-bands as a consequence of survey design presents an additional challenge to the fitting process as well as in later analysis. PYAUTOLENS uses units of electrons per second, so the effect of the difference in exposure times and S/N between observations with the two band-passes is minimized. However, measurements taken from models that fit images of the same bandpass are much easier to compare.

5 MODEL QUALITY ASSESSMENT AND GRADING

With models complete, we assess the highest-likelihood models for each candidate. Ideally, a reliable objective figure of merit, such as image χ^2 or Bayesian evidence, would sufficiently quantify the quality of each lens model fit. Forming robust quantitative goodness-of-fit metrics is currently an open problem in automated lens modelling. Etherington et al. (2022) explored this problem using PYAUTOLENS with much higher resolution images and found that none did a particularly satisfactory job. We take the Bayesian evidence to be the reference figure of merit and follow this with a blind visual inspection of the image, fit, and spectrum of each modelled candidate. We inspect the images and spectrum separately in order to isolate some of the failure states that occur for each and have a clear picture of the factors limiting the precision of the models. Three collaborators give a separate score between 0 and 4 for each of the image, fit, and spectrum for each candidate, so that each of the candidates has three scores out of 12 and a total possible score of 36.

The procedure for assigning quality scores is as follows: The collaborator (the 'scorer') is shown via a randomized selection either the spectrum or a set of images (observed and model-fit) of a randomly selected candidate. The spectrum and image set are not shown sequentially in order to keep the scores unbiased by each. The spectrum score is based on the detection of redshifted line features that correspond to both the foreground and background redshifts. Wavelength accuracy, strength, and number of detectable line features are considered, in addition to template type and the presence of overlapping line features. The image and fit are scored simultaneously from the set of four observed and model-fit images because the fit score is informed by the image score. The image score

Table 1. Grading scheme based on the total score and spectrum score for each candidate as described in Section 5. We give greater weight to spectrum score because the quality of the AUTOZ redshift determination is essential to deriving meaningful physical results. All graded models have received no '0' scores for any individual quality by any scorer, ensuring that the final set is clean.

Grade	Total score $\geq$	Spectrum score $\geq$	# Models with grade
A	30	9	2
B	20	6	3
C	16	5	9
D	12	4	5

is based on two images – (i) the observed image and (ii) the observed image with the model's lens light Sérsic profile subtracted. The scorer considers how well the two images appear to show a well-defined structure outside the central foreground lens light-profile that could be reasonably described as a lens feature. For the fit score, the scorer compares the lens-subtracted model image to the lens-subtracted observed image and examines model background source-plane image. The fit score is influenced by the image score; the fit score cannot be higher than the image score +1. This means that a poor image that is fit perfectly should not get a high fit score, and a good image that is fit poorly should reflect the failure of the model to adequately attribute the image features to the lensing of a background source.

Following the scoring exercise, we remove any candidate that received a '0' for any of the image, fit, or spectrum scores by any scorer. This removes catastrophic failures and ensures that the final set is reliable for follow-up analysis. The 19 candidate models that remain are assigned a letter grade of A, B, C, or D according to the structure outlined in Table 1. There are 2 A, 3 B, 9 C, and 5 D grades in the scored subsample, described in Table 2. 17 of the graded models are candidates from the LinKS subsample. Two of the three candidates from the 'Li-BG' sample were modelled to a level of success that justified presentation alongside the others, though both models are given grades of D. One D-grade model (G419067) with a negative likelihood was a result of high image residuals in the very centre of the lens light profile. Note the asterisk in the *ln*(evidence) column of Table 2. This model scored well enough for inclusion (spectrum score 4, total score 22) by the blind visual inspection. However, the visual inspection may have removed this candidate with the inclusion of a residual or χ^2 map in addition to the model images. The Bayesian evidence is therefore a prudent first cut of extremely poor models. Otherwise, as shown in Fig. 3, we find that the quality of fit determined by careful visual inspection is not correlated to the reported Bayesian evidence.

To the authors' knowledge, none of these lens candidates have been previously confirmed with high-resolution (HST-quality) imaging or spectroscopy. G250289 was identified in HSC as J083726 + 015639 by Sonnenfeld et al. (2019). Spectrum scores of six or better can be considered to be probable spectroscopic evidence for the lens candidate, and the highest spectrum scores for the two A grades can be considered spectroscopic confirmations. No additional extensive efforts were made to identify another possible background source redshift if the one determined by AUTOZ was deemed unreliable. All scores can be considered useful follow-up evidence for the quality of the candidates, in that the success of a model lends additional confidence to the identifications. However, Petrillo et al. (2019a) and Li et al. (2020) have already provided extensive studies of the quality of their lens identifications, and our image modelling is conducted on

Table 2. The 19 models with letter grades as selected in Section 5. The other 24 models were considered too poor for consideration here. ID refers to internal the LinKS sample identifier or our labelling of Li-BG candidates that were modelled. Type refers to the configurations of foreground + background galaxy template type. Scores are the sums of scores given by three individual scorers. Grades classify the quality according to the grading scheme shown in Table 1. *ln*(evidence) is the log of the Bayesian evidence reported by PYAUTOLENS. *G419067 had a negative evidence as a result of high image residuals in the centre of the lens light profile.

| GAMA ID	ID	RA	Dec.	z_{lens}	z_{source}	Type	*Scores*: Spectrum	Total	Grade	*ln*(evidence)
323152	2967	130.546	1.643	0.353	0.722	PG + ELG	12	33	A	7.10
138582	2828	183.140	− 1.827	0.325	0.433	ELG + ELG	11	32	A	7.47
250289	2730	214.367	1.993	0.401	0.720	PG + ELG	8	27	B	6.28
62734	539	213.562	− 0.242	0.274	0.597	PG + ELG	6	26	B	4.50
513159	2123	221.917	− 0.999	0.289	0.701	PG + ELG	7	23	B	7.59
3891172	3056	139.227	− 1.545	0.340	0.609	PG + PG	5	24	C	6.43
373093	2897	139.306	1.198	0.384	0.837	PG + ELG	5	23	C	7.31
559216	2507	176.116	− 0.619	0.250	0.714	PG + ELG	7	19	C	7.77
3629152	1933	135.889	− 0.975	0.407	0.787	PG + PG	5	19	C	7.36
3896212	1483	129.806	− 0.830	0.382	0.848	PG + PG	6	18	C	6.38
342310	2163	215.081	2.171	0.380	0.693	PG + ELG	5	18	C	5.79
272448	2541	179.420	1.423	0.272	0.889	PG + ELG	5	17	C	7.07
262874	26	221.611	2.224	0.386	0.859	PG + PG	6	16	C	6.00
387244	1819	135.569	2.365	0.218	0.712	PG + ELG	5	16	C	7.37
569641	BG3	219.730	− 0.597	0.360	0.826	ELG + ELG	4	25	D	7.27
419067	1179	138.620	2.635	0.188	0.764	PG + ELG	4	22	D	*
16104	BG1	217.678	0.745	0.287	0.849	PG + ELG	4	19	D	7.08
561058	3349	182.560	− 0.495	0.320	0.856	PG + ELG	6	14	D	6.96
262836	1953	221.405	2.314	0.144	0.418	ELG + PG	5	13	D	7.80
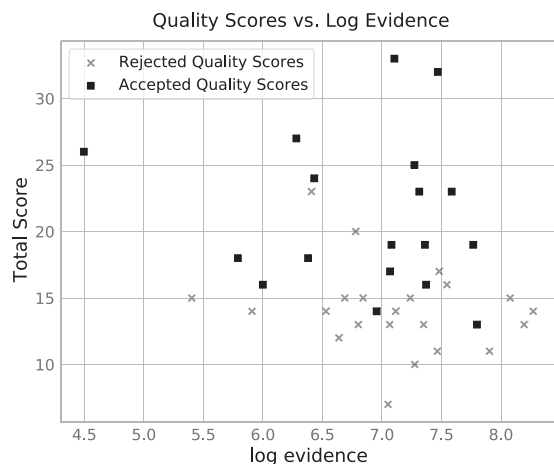
Figure 3. Total score from visual quality inspection versus the natural log of the Bayesian evidence from model fitting. X-markers are rejected based on visual quality inspection. Square markers are accepted. There is very little correlation between the visual inspection results and the objective quality-of-fit metric.

the same observations as their analysis. Therefore, any poor model performance here does not contradict a positive identification.

6 MODEL RESULTS

6.1 Extracting best-fitting parameters

For each lens model, the parameter space is sampled over tens of thousands of iterations, estimating the log-likelihood for each sample fit and constructing a probability density function (PDF) for each free parameter listed in Table B1 of the appendix. Mass and light profiles are fully described by the model-fit parameters. The Einstein radius,

total lensing 'Einstein' mass, mass fractions, luminosity, and mass-to-light ratios are calculated for each sample from the model grid and integrated over the angular area enclosed by the Einstein radius.

Parameters involving the luminosity require k-corrections to rest-frame (see Section 4.3.1). Three of the four models used the g-band image for Search 3, so they are easy to compare. Special attention should be given to G250289, which was instead modelled from its r-band image. In attempting to give the models the best chance to succeed, we were inconsistent in the choice of bandpass for modelling (see Section 4.4). Luminosity and mass-to-light ratios for this model are corrected to the g-band after r-band rest-frame k-correction. This is done by multiplying (or dividing) by an additional factor $10^{-0.4(g - r)}$, where $(g - r) \sim 0.285$ is the colour difference calculated by integrating the product of each bandpass response function and a template elliptical galaxy spectrum from Kinney et al. (1996) over the bandpass range. This has the effect of scaling luminosity down and *M/L* up. Given our goal here, which is to explore the methods, this is sufficient for characterizing the differences between models in a consistent parameter space. However, future efforts that intend to approach these measurements more rigorously should approach the initial modelling with more consistency.

We briefly present best-estimate results from the highest-likelihood model fits for the four highest quality models in Table 3, selected primarily by the blind quality scoring described in Section 5. Bayesian evidence reported by PYAUTOLENS and the subjective reasonableness of the inferred quantities are also considered in the selection of this small subsample. One B-grade model, G62734, is not included because its central dark matter content is poorly constrained. This and the other 14 lower-grade models are considered to be worth revisiting but were not successful enough to present alongside the cleaner examples we present here.

To discuss inferred quantities, we estimate bivariate PDFs for the quantities using a Gaussian kernel-density estimate from the final 10 000 iterations. The best estimate for each inferred quantity listed in Table 3 is determined at the maximum of one of these bivariate PDFs,

Table 3. Results of four highest scoring models. z_{lens} and z_{source} are the redshifts of the foreground deflector and background source. Type refers to the configuration of foreground + background template types. θ_E is the Einstein radius calculated from the model mass distribution and lensing distances. Remaining quantities are integrated within θ_E. M_E is the total enclosed Einstein mass. f_{DM} is the enclosed dark matter fraction. L is the enclosed luminosity in the r-band enclosed. M_*/L is the enclosed stellar mass-to-light ratio. Grade is an evaluation of the quality of the fit to the image according to the scheme outlined in Table 1.

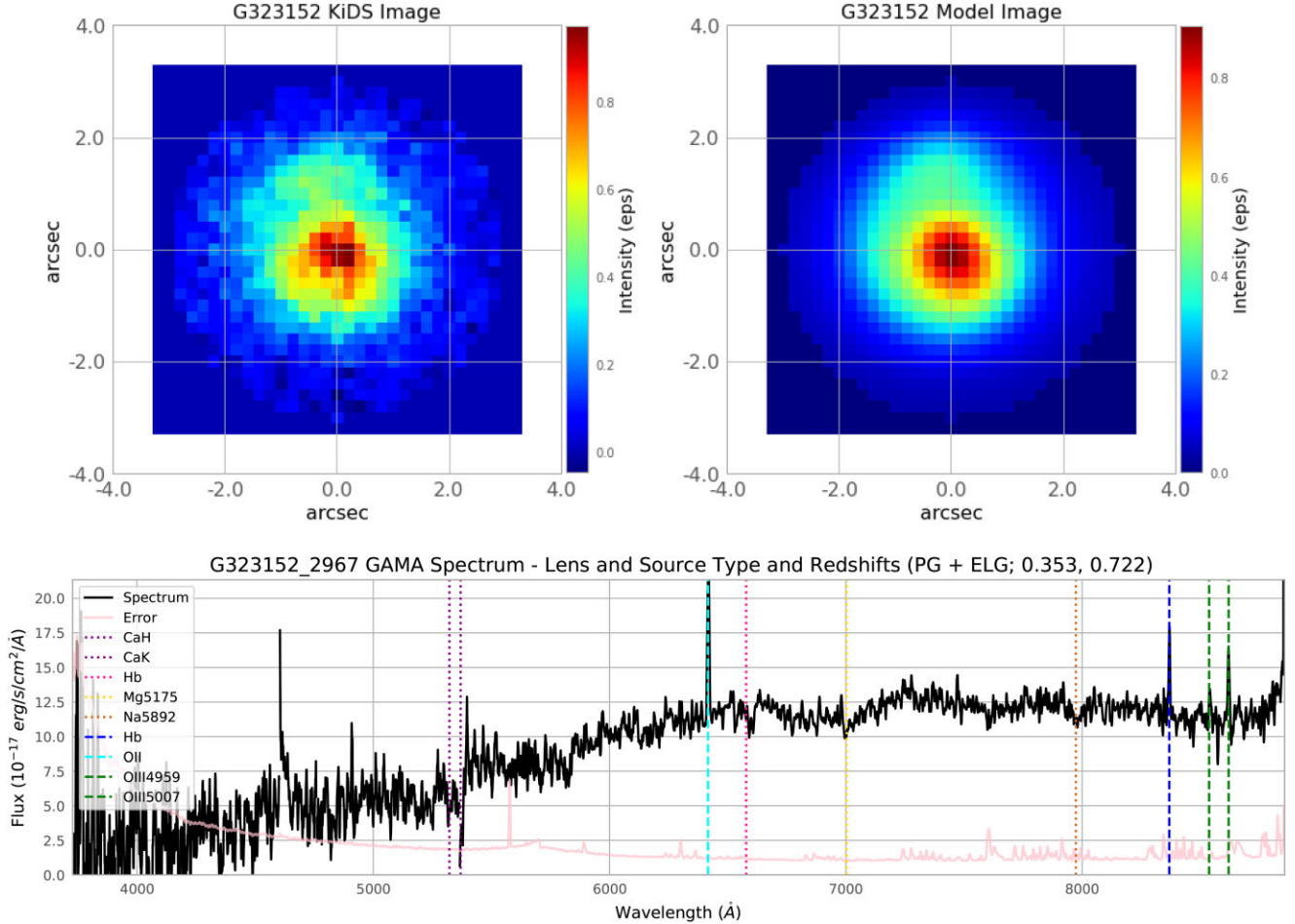| GAMA ID	z_{lens}	z_{source}	Type	$M_*/M_\odot$	$M_E/M_\odot$	f_{DM}	θ_E	θ_{eff}	M_*/L_g	$L_g/L_\odot$	Grade
138582	0.325	0.433	ELG + ELG	9.91e + 10	8.69e + 11	0.888	1.20	1.99	7.70	1.98e + 10	A
323152	0.353	0.722	PG + ELG	1.31e + 11	4.65e + 11	0.717	1.27	2.78	4.98	2.69e + 10	A
513159	0.289	0.701	PG + ELG	7.31e + 10	7.29e + 11	0.900	1.72	2.44	4.85	1.52 + 10	B
250289	0.401	0.720	PG + ELG	5.47e + 11	8.82e + 11	0.375	1.55	2.44	8.92 (6.86 r)	6.11e + 10 (7.95e + 10 r)	B

Figure 4. G323152. A-Grade. *Upper left*: The observed image shows an apparent arc feature above the central lens galaxy light-profile. *Upper right*: The model image captures the extra light reasonably, though without the exact shape. This could be a result of internal substructure or the impact of shear along the line of sight, both of which are unaccounted for in the model. *Lower*: The GAMA spectrum shows strong line features at the redshifts at 0.353 and 0.722 identified by AUTOZ. Dotted lines identify foreground lens galaxy absorption features (H, K, Hβ, Mg, and Na) at $z = 0.353$, and dashed lines show background source emission features (Hβ, [O II], [O III]) at $z = 0.722$.

where we use uncorrelated values as much as possible. We show the observed image, maximum-likelihood model fit, and spectrum for the highest scoring model in Fig. 4. The other three models listed in Table 3, as well as G62734, are shown and discussed in more detail in Figs C1–C4 of Appendix C.

We are primarily interested in studying the stellar and dark matter content in the central regions of the lens galaxies. Although the mass and light profiles in the models are inferred to a larger radius, mass measurements via strong lensing are the most precise when considering only mass within the Einstein radius of the galaxy. It is

important to note that the Einstein radius is a feature individual to each system, so the quantities are not calculated within a uniform radius from the centre of each galaxy. Values for each Einstein radius can be found in Table 3.

6.2 Comparing highest-quality model results

We show the four highest-quality models for comparison. Figs 5–7 show the four models in parameter space of interest to our study. All of the plotted quantities are taken within the Einstein radius (see
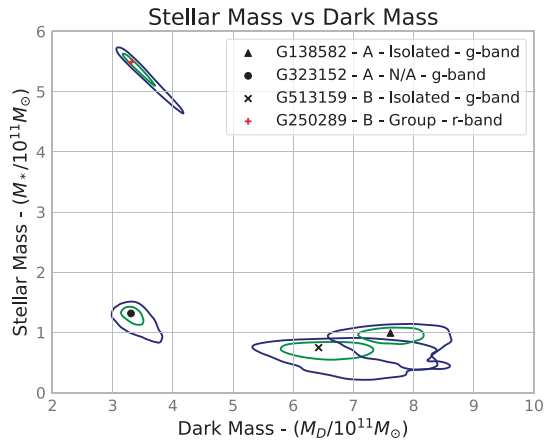
Stellar Mass vs Dark Mass

Figure 5. Mass components integrated within the Einstein radius of each of the four best-fitting models described in Section 6 and Table 3. The legend shows the GAMA identifier, quality grade, environment classification, and the SDSS bandpass used for the final model fitting. Green and blue contours about each point enclose 1σ (39 per cent) and 2σ (86 per cent) of the two-dimensional PDF, respectively.

Table 3). Each model is identified with a different marker, and B-grade group-member galaxy G250289 is indicated with a red marker to remind the reader that the final model fit utilized the r-band. Green and blue contours enclose 1σ (39 per cent) and 2σ (86 per cent) of the two-dimensional PDF, respectively. With the small sample size shown here, we do not intend to address questions of assembly bias and galaxy formation mechanisms. These plots are intended to discuss the cleanest subset of our sample in the context of what can be considered more thoroughly in future work.

Fig. 5 shows the integrated stellar mass and dark mass enclosed within the Einstein radius of the lens models. These are obtained by integrating over the Sérsic stellar mass profile and elliptical NFW profile. The galaxies have total enclosed Einstein mass values of order $M_E \sim 3-8 \times 10^{11}$ M$_\odot$, which is expected since lensing galaxies are typically quite massive. Assembly bias would show itself here as a trend where group central galaxies tend to have higher stellar mass than isolated galaxies at the same dark matter halo mass. The only group-member galaxy, G250289 (marked with a red cross), lies in one of the smaller dark matter haloes and has the highest stellar mass. G323152 (marked with a black circle, not listed in GAMA GROUPFINDING catalog) has a similar dark mass and about one-fifth of the stellar mass compared to G250289. Conversely, the two isolated galaxies have the highest dark masses and the lowest stellar masses. The higher stellar mass in G250289 could have more to do with effects from the difference in r-band and g-band S/N than a physical interpretation. With so few data, it is difficult to determine how much of an effect this difference has on the results of the models.

The total lensing (Einstein) mass enclosed within the Einstein radius is generally well constrained. We want our models to further constrain the fraction of this lensing mass that is dark matter. The fraction of dark matter is not directly constrained by a model prior but is very sensitive to the assumed forms of mass and light profiles and the constraints placed upon those. Fig. 6 shows the g-band integrated luminosity and dark matter fraction (both calculated within the Einstein radius) for each model. The uncertainties in the fraction of dark matter along the x-axis are relatively small, which is surprising given the inherent degeneracy of stellar and dark mass in lens modelling. The small parameter space explored could indicate a lack of flexibility of the models' stellar and dark
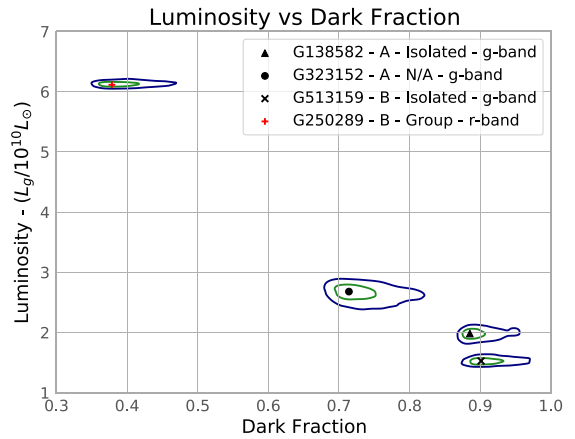
Luminosity vs Dark Fraction

Figure 6. SDSS g-band luminosity and dark matter fraction integrated within the Einstein radius of each of the four best-fitting models described in Section 6 and Table 3. Legend and marker information are the same as in Fig. 5.
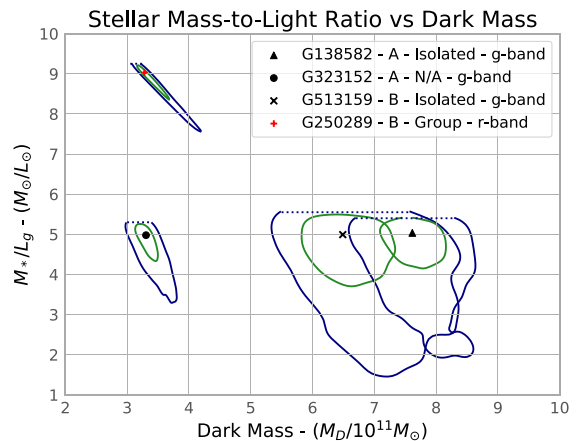
Stellar Mass-to-Light Ratio vs Dark Mass

Figure 7. Stellar mass-to-light ratio (M_*/L_g) in g-band and dark mass integrated within the Einstein radius of each of the four best-fitting models described in Section 6 and Table 3. Dashed lines at the 2σ contours indicate boundaries corresponding to upper constraints placed on the mass-to-light ratio of the models (see Section 4.2). Legend and marker information are the same as in Fig. 5.

profile priors, perhaps an excessive constraint or weighting towards one mass component over the other. As discussed in Section 4.2, the careful selection of priors can be challenging. Compare again G250289 (red cross) and G323152 (black circle), which have similar enclosed dark masses. Even corrected (scaled down) from r-band to g-band, the enclosed luminosity for G250289 is 2–4 times greater than the other three, which could be why the model attributes a higher fraction of the Einstein mass to the stellar component. These models have relatively similar total Einstein masses enclosed within similar Einstein radii around 1.2–1.7 arcsec. G250289 has an Einstein radius of ~1.5 arcsec that is typical and within the range of the other model values, so additional luminosity and stellar mass is not a result of an overextended radius of integration.

Fig. 7 shows the g-band stellar mass-to-light ratio compared to the enclosed dark mass. Dotted lines at the 2σ contours indicate upper constraints placed on the mass-to-light ratio. This figure completes our discussion of the degeneracy. In summary, the model has two options for attributing the lensing mass: (i) To favour the stellar
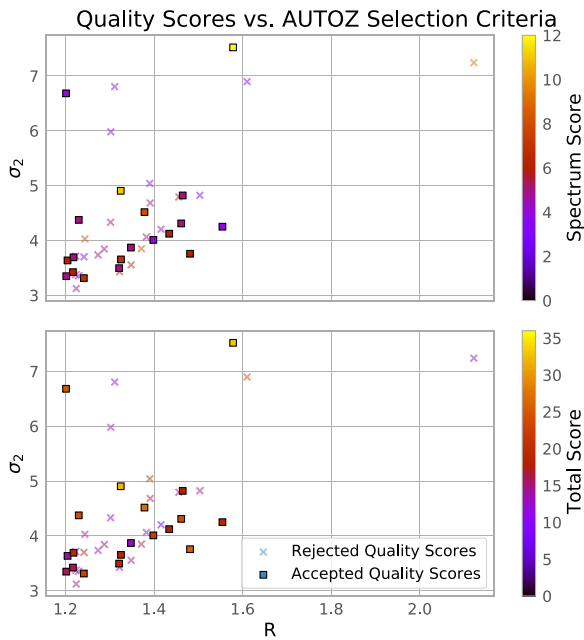
Figure 8. *Upper*: Spectrum quality scores. *Lower*: Total score shown as scatter plot colour variations scaled with the colourbar on the right of each plot. The axes of the scatter plot are the selection criteria from Fig. 1. Vertical axes are the second-highest cross-correlation peak σ_2, and horizontal axes are the R parameter. Dark purple colours indicate poor scores, with higher scores at orange and yellow. Little correlation can be seen between these parameters and the results of the models or their subjective spectrum scores. In the upper plot, four low-scoring spectra with high σ_2 correspond to overlapping emission lines.

component, a higher stellar mass can be the result of a heavier stellar population, and (ii) conversely, a very large, centrally concentrated dark matter halo can make up for a lower stellar mass and luminosity. This is all expected. Bounds of integration for luminosity, stellar mass, and dark mass are all dependent on the measure of the Einstein radius, which is in turn dependent on the total mass which the model is attempting to parse into stellar and dark components. It is a complex problem with degenerate variables that is only constrained by meaningful priors. In our cleanest subsample, the most identifiable differences occur for the candidate that was modelled from a different photometric bandpass. This is another experimental design decision based on limitations of the data that significantly affected our ability to analyse the resulting models. Thus, even our best models suffer.

7 AUTOZ CONSIDERATIONS POST-MODELLING

Here, we summarize the results of our quality control procedure described in Section 5 and give several recommendations for removing contaminants. Appendix D gives a more detailed breakdown of the candidates and models in the context of the cases discussed in Section 3.2.

We discovered some failure modes for our method of utilizing the AUTOZ algorithm for background source redshift determination. Revisions to our initial procedure removed about half of the candidates. Few single cases should be excluded without consideration; for example, some redshift determinations were kept even though there appeared to potentially be a falsely attributed line. However, absorption and emission lines must both be checked for overlap regardless of the template configuration type, and template type

should be questioned with this information. Several of the lens ELG galaxies turned out to have prominent absorption features at the same redshift. The binary identification of PG and ELG, as we have done here, is an oversimplification that hinders both our classification and understanding of the spectra and expected galaxy properties.

The two A-grade candidate models were acquired with AUTOZ redshift matches that we consider to be cases that deserve extra care (where the foreground galaxy is best matched with an ELG template). This is interesting because one may be tempted to cut these cases entirely in order to obtain a clean and fairly homogeneous sample (i.e. large elliptical galaxy lensing a bright ELG, PG + ELG). We find the more careful consideration of other possible cases to be fruitful. If we had adopted a selection that focused only on configurations where the primary template matched a passive lensing galaxy, then ~ 20 per cent of our final selection, including the two highest-scoring models, would have never been considered. Care should be taken in order to reap the benefits of this expanded population while minimizing contamination.

In the big picture, the improvement of spectral quality in wide-field surveys is essential for making this work in an automated way over large sample sizes, but an automated redshift algorithm like AUTOZ could be optimized for background source redshift determination. Our subjective quality scores show little correlation to the AUTOZ output parameters that we used for the initial selection, as shown in Fig. 8. The two axes show the AUTOZ selection parameter space, composed of the second cross-correlation peak σ_2 and the R parameter. Three of the candidates with the highest σ_2 show very poor spectrum scores because they are instances of overlapping emission lines. This suggests that a higher threshold for σ_2 may not actually yield higher-quality background source redshift matches. We now introduce other options for maintaining a clean sample selection with AUTOZ while expanding the sample size in future works.

7.1 Recommendations to remove contaminants from AUTOZ selection

One could quite easily remove contaminants during the automated selection. Each of these recommendations pays particular attention to the redshifts of ELGs in both the foreground lens and background source positions. The two most convincing (A-grade) spectra and models were cases of ELG+, so we want to remove as many contaminants as possible without the blanket removal of either of these cases. In order to maintain the applicability of this procedure to an even larger set of data than is considered here, we recommend the following selection criteria be implemented when adopting automated redshift determinations:

(i) **Remove ELG+ELG and PG + PG matches where emission or absorption lines redshift to overlapping wavelengths between foreground lens and background source.** One can calculate lens and source redshift combinations that result in overlapping observed emission lines similarly to the procedure described in Holwerda et al. (2015). The following equation defines a region of parameter space between a lower and upper linear function of $(1 + z_{source})$ to $(1 + z_{lens})$. Within this region, an overlap will occur for a given pair of rest-frame emission line wavelengths:

$$\left| \frac{1 + z_{source}}{1 + z_{lens}} - \frac{\lambda_{r,lens}}{\lambda_{r,source}} \right| < \frac{A}{R} \tag{4}$$

where z_{source} and z_{lens} are the source and lens redshifts, $\lambda_{r,\,lens}$ and $\lambda_{r,\,source}$ are the rest-frame wavelengths of emission lines from lens and source, R is the spectral resolution of the instrument, and A is a
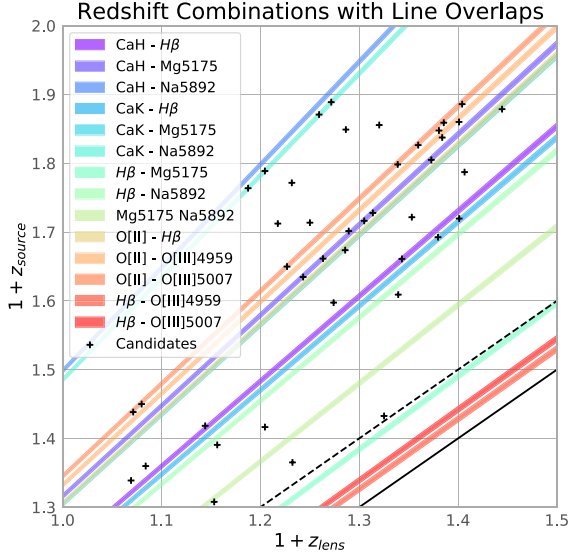
Figure 9. Combinations of foregrounds lens and background source redshift that result in overlapping important emission or absorption line features. Coloured line-regions are calculated with equation (4) and indicate parameter space where the labelled line features will overlap. Each is labelled in the legend as 'source feature - lens feature'. Using these functions identifies 19 of the 20 overlaps in the 42 candidates and all 6 that made the final selection of 19. The region in the lower right between the solid and dashed black lines shows the selection criterion utilized in the initial AUTOZ selection, which removed candidates where the source redshift was within 0.1 of the lens redshift.

coefficient that widens the range of exclusion for potential overlapping features. The equation implicitly accounts for the dependence of resolution on observed wavelength. Fig. 9 shows the regions where a redshift combination of foreground lens and background source results in overlapping lines given GAMA's spectral resolution of $R \sim 1300$ and $A = 4$ (i.e. overlapping emission lines are closer than four times the smallest resolvable wavelength difference). This prescription identifies most of the cases of overlap that were flagged by direct visual inspection of the spectra. We retained one of them with the lowest possible D-grade because its SDSS-BOSS spectrum showed fairly reasonable source Hβ and [O III] couplet lines at higher-wavelength that were not in the range of the GAMA spectrum.

(ii) **Remove + ELG configurations where Hβ and [O III] couplet emission lines are redshifted beyond the wavelength range of the observation.** Table 4 and Fig. 10 show the redshift limits beyond which Hβ and [O III] $\lambda\lambda$4959,5007 lines will be above the survey upper wavelength limit for several optical spectroscopic surveys. Spectroscopy in the 1-μm range is necessary in order to detect the emission lines from sources at $z \sim 1$ and will become more important for modelling lenses with foreground lens redshifts higher than $z \sim 0.5$. DESI and the upcoming 4MOST have higher resolution

Table 5. Two models with AUTOZ primary redshift template match to the background source and secondary match to the foreground lens ($\sigma_{lens} < \sigma_{source}$). Type refers to the foreground + background configuration of galaxy templates. Grade is an evaluation of the quality of the fit to the image according to the scheme outlined in Table 1. G323152 is one of the highest scoring models in this study.

GAMA ID	z_{lens}	z_{source}	σ_{lens}	σ_{source}	Type	Grade
323152	0.353	0.722	**7.52**	11.32	PG + ELG	A
262836	0.418	0.144	**3.87**	10.23	ELG + PG	D

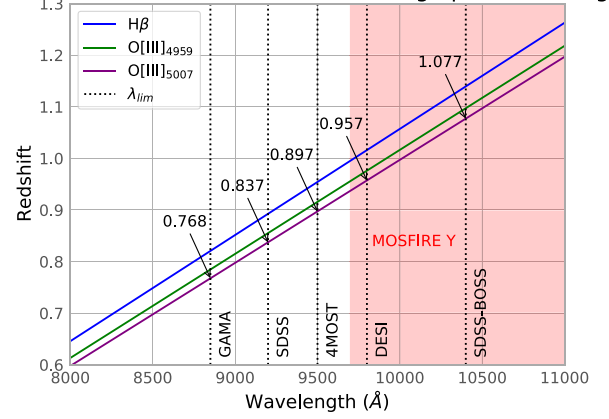**Figure 10.** Blue, green, and purple lines show the redshift of rest-frame Hβ and [O III] $\lambda\lambda$4959,5007. Dotted vertical lines show the upper wavelength limit of various spectroscopic surveys, and the overlaid arrows point to the redshift at which the first of these three lines will disappear in the survey observations. The red shaded region indicates the very near-IR coverage of MOSFIRE Y-band (0.97–1.12 μm) that could also potentially reveal background source emission lines at around $z \sim 1$ and greater.

and cover extended optical wavelength ranges that correspond to these redshifts, which will reduce the significance of this problem. Background source redshifts could also be assessed by looking for emission lines in very near-infrared (e.g. MOSFIRE Y-band, 0.97–1.12 μm) spectra, where possible. In principle, template-based automated redshift identification could be run for each separately, which would negate some of the difficulties inherent to identifying the two distinct signals within the single observation. The obvious negative to this option is the need for additional observations.

(iii) **Remove ELG+ matches with any of the following characteristics: (a) low foreground lens redshift (less than $z \sim 0.2$ for our sample), (b) ELG + PG configuration, and (c) primary redshift match to the background source.** ELG+ configurations with low foreground lens redshifts suffered from several failure conditions in addition to being a less-likely configuration than PG+. For GAMA,
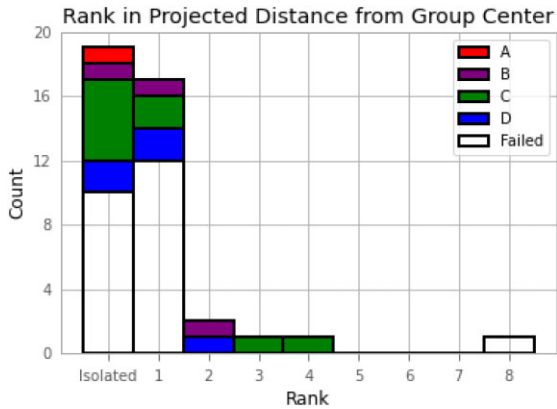
Table 4. Five spectroscopic surveys and their upper wavelength limits limits. The right three columns show the source redshift at which the given emission line will be redshifted beyond the upper wavelength limit of the observation.

Survey	λ_{lim} (Å)	z_{source} Hβ	[O III] λ4959	[O III] λ5007
AAT (GAMA/DEVILS)	8850	0.821	0.785	0.768
SDSS (original)	9200	0.893	0.855	0.837
4MOST (lo-res)	9500	0.954	0.916	0.897
DESI	9800	1.016	0.976	0.957
SDSS-BOSS	10400	1.139	1.097	1.077

Figure 11. Stacked histogram. Quality scores for 41 of 42 candidates in reference to data from GAMA GROUPFINDING catalogs (one candidate does not have environment data). Location on the *x*-axis distinguishes 'Isolated' from group member galaxies, which are further separated by the rank in projected distance from the centre of mass of the group. A rank of 1 indicates that the lens candidate is the central galaxy of the associated group. Colours indicate the quality grade of models, with no colour indicating models that were not accepted.

the noisy short-wavelength end of the observed spectral range corresponds to emission lines with z_{lens} less than ∼0.2, leading to mistaken classification of noise peaks as emission lines. While this is specific in part to GAMA's wavelength-specific spectral performance, several of these low-redshift foreground lens matches are also ELG+PG configurations, of which almost all were removed in quality scoring. The one remaining candidate had the lowest score of all those accepted. Perhaps even more condemning is the fact that many of these low-redshift ELG+ foreground lens matches were also cases where the background source was assigned the primary redshift. This trio of doubtful cases coincided for several of the candidates that were removed from our sample during quality scoring.

8 DISCUSSION

8.1 GAMA environment

The sample of 42 KiDS lens candidates and the subsample of 19 with accepted grade A–D models are both close to evenly split between group-member and isolated galaxies according to the GAMA GROUPFINDING metrics. 22 (8 graded) are associated with groups and 19 (9 graded) are isolated. (G323152) is not represented in GAMA group catalogs.

We note that most strong lensing galaxies should be the most massive galaxies in their halo, either in a group or in isolation. Fig. 11 shows the rank of the proximity of the object to the centre of mass of the group relative to other group members, with one indicating the closest or centre-most galaxy. Most of the candidates shown here are group central galaxies, but our models failed for 11 of those. On the other hand, four of the five candidates that are not the central galaxy of their group were accepted and given grades, comprising 40 per cent of the graded group-galaxy members. Low scores for group member galaxies could mean that their identification as lenses is a false-positive given their proximity to other significantly large galaxies. Alternatively, if these are lenses, the modelled mass structure of the lensing galaxy as a single light-mass component and assumption of its presence in the centre of the dark matter halo may not be accurate enough to reproduce the lensing observables. If this were the case,
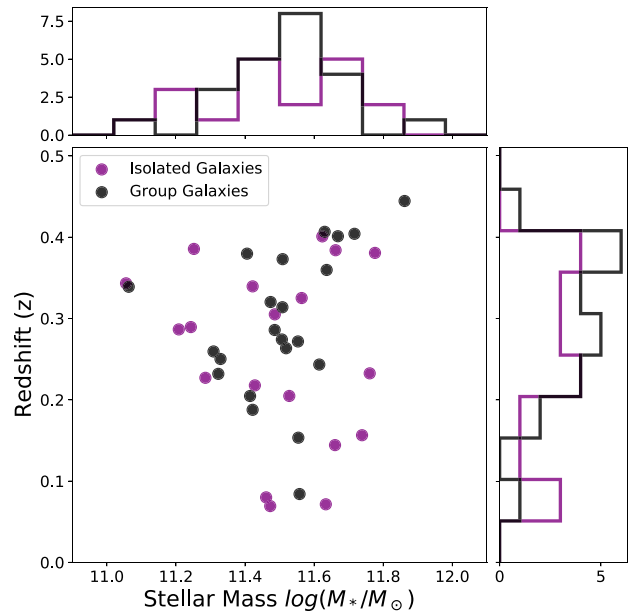

Figure 12. Redshifts and stellar masses of AUTOZ sample from GAMA-DR3 STELLARMASSESLAMBDAR V20 (Taylor et al. 2011) determined by stellar population and separated by group-member and isolated galaxies according to GAMA team internal GROUPFINDING catalogs (Robotham et al. 2011). There is no clear distinction between the subsamples in either observable.

one might expect the group centrals to be more easily modelled than the subdominant or 'satellite' galaxies with ranks of 2 or greater, which is not apparent in Fig. 11. The two B-grade group galaxies are ranked 1 and 2, indicating that one of them is a central while the other has at least one companion that is competing for dominance in the group. The rank 2 group member is G62734, which was removed from the final selection because its dark matter content was poorly constrained. This could be a result of this galaxy's distance from the centre of the group mass.

Compared with the SLACS study in Treu et al. (2009), in which 12 of 70 (17 per cent) were associated with groups, our KiDS/GAMA strong lens candidate sample and selected subsample of 19 models is more highly represented by group-member galaxies. Definitions of group membership based on environmental parameters are not the same between these studies. The nearly 50/50 split between group-member and isolated galaxies in our sample does not necessarily support or dispute a preference for overdense environments by lensing (and all massive) elliptical galaxies. However, the high completeness of GAMA compared with SDSS may instead suggest that our sample minimizes the apparent environmental preference. The distinction of group association here could be affected by selection bias, as those designated as isolated could in fact be groups with satellite members beyond the GAMA flux limit. If this were the case, there should be a systematic bias in isolated galaxies towards higher redshift. Fig. 12 shows that neither subsample of group member or isolated candidates is significantly distinguished in redshift or stellar mass.

With more data and better measurements than we have accomplished here, one may be able to compare observations to the scatter in the upper plot of fig. 9 of Zehavi et al. (2018), where for fixed dark halo mass, higher stellar-mass galaxies tend to exist in denser environments. Note that the modelled mass components here are calculated within the Einstein radius and not the full extent of the galaxy. The majority of the dark halo component should extend well beyond
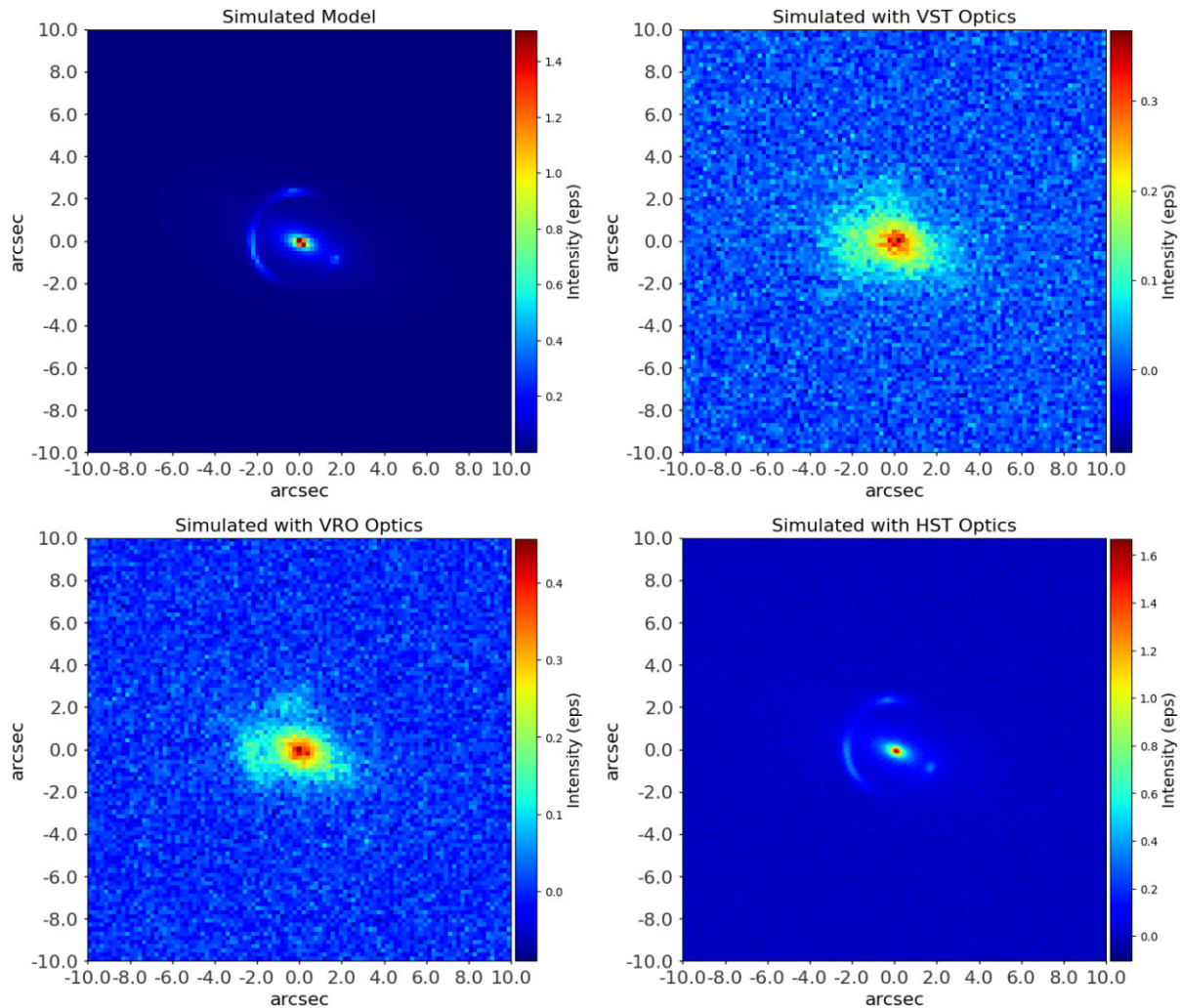
Figure 13. *Upper left*: Maximum log-likelihood model for G3629152 shown with pixel-scale 0.2 arcsec pixel^{-1}. Other images are simulated with identical background sky and convolved with optics of *upper right*: VST (r-band PSF 0.65 arcsec, pixel scale 0.2), *lower left*: LSST at Rubin (PSF 0.5, pixel scale 0.2), and *lower right*: ACS on HST (PSF 0.1, pixel scale 0.05). Imaging from space-based observatories or AO would allow for better model-fitting and tighter uncertainties for future efforts.

the stellar halo, and these high-mass lensing galaxies are more likely to exist in more massive dark matter haloes ($\log(M_h/h^{-1}\,\mathrm{M_\odot}) \sim$ 12–14) where the suggested environmental trend is less supported. The precision and numbers required to test assembly bias will require more refinement of the methods discussed in this study as well as the power of more sophisticated surveys to come.

8.2 Future work: a place for ground-based observations

Realistic lens modelling by fitting mass and light profile parameters is a complex problem with a large number of parameters. With even the highest-quality ground-based imaging offered by the likes of the KiDS, the angular resolution is insufficient to constrain the *individual* model solutions to levels where one can make strong inferences about the *individual* lens galaxies. There are simply too many solutions that fit the image to a high probability, which inflates the uncertainty to levels that make it difficult for one to draw conclusions from the inferred quantities. These uncertainties on a single lens can be significantly constrained with the level of imaging afforded by adaptive optics (AO) or space-based instruments. Fig. 13 shows a model solution for one of the lens models after being

simulated with the optics for three observatories: (i) VST used for KiDS, which was the instrument that collected the original image, (ii) LSST at the Vera Rubin Observatory (VRO) representing the next generation of ground-based observatories, and (iii) Advanced Camera for Surveys on the Hubble Space Telescope. The same model-fitting procedure applied to HST images or observations with AO of the same lensing galaxies would result in error estimates an order of magnitude better than the results we achieve here. Alternatively, future systematic modelling of orders of magnitude more ground-based, lower-resolution observations (as we expect to achieve with observatories like the VRO) can result in similar precision. Constraints at the population level (made possible with these larger sample sizes) can enhance higher-resolution individual measurements through Bayesian hierarchical frameworks. Our work demonstrates the value of wide-field, lower-resolution surveys as a complementary tool to the expensive and hyper-competitive observing campaigns that are the default for strong lens studies.

The next generation of spectroscopic surveys is already underway, e.g. the DEVILS deep survey on the AAT (Davies et al. 2018; Holwerda et al. 2021) and the DESI redshift survey. One can expect increased numbers of spectroscopic lensing candidates as

well as opportunities to identify the redshift of the potential background source. More comprehensive spectroscopic surveys are being planned with the 4MOST instrument (de Jong et al. 2012; Depagne & 4MOST consortium 2015). These planned surveys include extragalactic ones such as the two-tiered Wide Area Vista Extragalactic Survey (Driver et al. 2019), the Optical, Radio Continuum and HI Deep Spectroscopic Survey (Duncan et al. in preparation), and a cosmological low-S/N wide-area survey (Richard et al. 2019). These 4MOST surveys are expected to achieve high completeness in their target fields and yield a boon of spectroscopically confirmed strong lensing systems with the same advantages exploited by the procedure outlined here at better spectral resolution and wider fields.

Identifications of strong gravitational lenses through imaging are also expected to increase in the near future with observations by the likes of the VRO, Euclid, and the Roman Space Telescope, in addition to improved machine learning techniques. Following the discussion in Knabel et al. (2020), one expects the selection functions of the spectroscopic surveys and these optical and near-infrared imaging surveys to show a limited improvement in overlap. The analysis presented here, however, shows that in the overlap a useful subset of strong lenses can be utilized for modelling by imaging and spectroscopy combined.

The lenses discussed here, and in future similar ground-based efforts, are also ideal candidates for deeper follow-up observation with higher-resolution imaging and spectroscopy. These follow-up observations could include Integral Field Unit observations to measure the stellar population characteristics across the elliptical, chart the foreground lens galaxy kinematics, as well as study the background source light and stellar population characteristics. These considerations are more aligned with and have been sufficiently described in the existing literature and will not be discussed further here.

9 CONCLUSIONS

We arrive at the following conclusions from our analysis of strong lens candidates in GAMA and KiDS using AUTOZ and PYAUTOLENS:

(i) Meaningful strong-lens studies can be conducted by applying lens-modelling methods such as those we have outlined here to large imaging and spectroscopic surveys.

(ii) Automated template-matching redshift algorithms like AUTOZ can be utilized to determine reliable background source redshifts required for lens modelling. Careful consideration should be taken in cleaning the algorithm's selection, following the recommendations outlined in Section 7.1.

(iii) Limits of optical resolution in large ground-based surveys present significant challenges to the uniqueness of solutions in our Bayesian modelling of individual strong lenses.

(iv) As sample sizes grow, refinements to these techniques can produce lensing measurements in *quantities* that will offer considerable statistical power. This approach is complementary to the more detailed modelling of individual lenses that is possible with deeper and higher resolution observations.

DATA AVAILABILITY

KiDS images and data used in this paper are available from the Astro-WISE Data base Viewer Web Service (dbview.astro-wise.org). LinKS-specific data can be found at the LinKS website (https://www.astro.rug.nl/lensesinkids/). The GAMA AUTOZ catalog is available from GAMA-DR3 website (http://www.gama-survey.org/dr3/, AATSPECAUTOZALL, SPECALL, LAMDARSTELLARMASSES, SERSICCATSDSS, and KCORR_AUTO_Z00 catalogs) and the team internal GROUPFINDING catalog for the full GAMA fields will be made available in GAMA-DR4 (Driver et al. in preparation).

REFERENCES

Alpaslan M. et al., 2014, MNRAS, 440, L106
Alpaslan M. et al., 2015, MNRAS, 451, 3249
Artale M. C., Zehavi I., Contreras S., Norberg P., 2018, MNRAS, 480, 3978
Astropy Collaboration, 2013, A&A, 558, A33
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., Bolton A. S., Gavazzi R., Koopmans L. V. E., Marshall P. J., Moustakas L. A., Burles S., 2010a, ApJ, 724, 511
Auger M. W., Treu T., Gavazzi R., Bolton A. S., Koopmans L. V. E., Marshall P. J., 2010b, ApJL, 721, L163
Baldry I. K. et al., 2014, MNRAS, 441, 2440
Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379

Behroozi P. et al., 2020, MNRAS, 499, 5702

Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Moustakas L. A., 2006, ApJ, 638, 703

Brough S., Tran K.-V., Sharp R. G., von der Linden A., Couch W. J., 2011, MNRASL, 414, L80

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Chaves-Montero J., Angulo R. E., Schaye J., Schaller M., Crain R. A., Furlong M., Theuns T., 2016, MNRAS, 460, 3100

Contreras S., Zehavi I., Padilla N., Baugh C. M., Jiménez E., Lacerna I., 2019, MNRAS, 484, 1133

Contreras S., Chaves-Montero J., Zennaro M., Angulo R. E., 2021, MNRAS, 507, 3412

Cui W., Davé R., Peacock J. A., Anglés-Alcázar D., Yang X., 2021, Nature Astron., 5, 1069

Davies L. J. M. et al., 2018, MNRAS, 480, 768

de Jong J. T. A., Verdoes Kleijn G. A., Kuijken K. H., Valentijn E. A., 2013, Exp. Astron., 35, 25

de Jong J. T. A. et al., 2015, A&A, 582, A62

de Jong J. T. A. et al., 2017, A&A, 604, A134

de Jong R. S. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460T

Depagne E., 2015, Astrophysics and Space Science Proceedings, Vol. 39, Asteroseismology of Stellar Populations in the Milky Way. Springer International Publishing Switzerland, Sesto, p. 147

Diemer B., 2018, ApJS, 239, 35

Driver S. P. et al., 2009, Astron. Geophys., 50, 5.12

Driver S. P. et al., 2011, MNRAS, 413, 971

Driver S. P. et al., 2019, The Messenger, 175, 46

Eisenstein D. J. et al., 2001, AJ, 122, 2267

Etherington A. et al., 2022, MNRAS, 517, 3275

Falcón-Barroso J., Sánchez-Blázquez P., Vazdekis A., Ricciardelli E., Cardiel N., Cenarro A. J., Gorgas J., Peletier R. F., 2011, A&A, 532, A95

Foreman-Mackey D., 2016, J. Open Sour. Softw., 1, 24

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

Hearin A. P., Watson D. F., van den Bosch F. C., 2015, MNRAS, 452, 1958

Hearin A. P., Zentner A. R., van den Bosch F. C., Campbell D., Tollerud E., 2016, MNRAS, 460, 2552

Holwerda B. W. et al., 2015, MNRAS, 449, 4277

Holwerda B. W. et al., 2019, AJ, 158, 103

Holwerda B. W., Knabel S., Steele R. C., Strolger L., Kielkopf J., Jacques A., Roemer W., 2021, MNRAS, 505, 1316

Hopkins A. M., 2018, Publ. Astron. Soc. Aust., 35, 39

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

Hunter J. D., 2007, Comput. Sci. Eng., 9, 90

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

Kelvin L. S. et al., 2012, MNRAS, 421, 1007

Kinney A. L., Calzetti D., Bohlin R. C., McQuade K., Storchi-Bergmann T., Schmitt H. R., 1996, ApJ, 467, 38

Knabel S. et al., 2020, AJ, 160, 223

Kuijken K. et al., 2019, A&A, 625, A2

Lam S. K., Pitrou A., Seibert S., 2015, Proc. 2nd Workshop on the LLVM Compiler Infrastructure in HPC - LLVM '15. ACM, New York, p. 1

Lastennet E. et al., 2002, Astrophys. Space Sci., 280, 83

Le Borgne J.-F. et al., 2003, A&A, 402, 433

Li R. et al., 2020, ApJ, 899, 30

Lin Y.-T., Mandelbaum R., Huang Y.-H., Huang H.-J., Dalal N., Diemer B., Jian H.-Y., Kravtsov A., 2016, ApJ, 819, 119

Liske J. et al., 2015, MNRAS, 452, 2087

Loveday J. et al., 2012, MNRAS, 420, 1239

Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, MNRAS, 432, 1544

Matthee J., Schaye J., Crain R. A., Schaller M., Bower R., Theuns T., 2017, MNRAS, 465, 2381

McCarthy K. S., Zheng Z., Guo H., Luo W., Lin Y.-T., 2022, MNRAS, 509, 380

Nightingale J. W., Dye S., 2015, MNRAS, 452, 2940

Nightingale J. W., Dye S., Massey R. J., 2018, MNRAS, 478, 4738

Nightingale J. W., Hayes R. G., Griffiths M., 2021a, J. Open Sour. Softw., 6, 2550

Nightingale J. W. et al., 2021b, J. Open Sour. Softw., 6, 2825

Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825

Petrillo C. E. et al., 2017, MNRAS, 472, 1129

Petrillo C. E. et al., 2019a, MNRAS, 482, 807

Petrillo C. E. et al., 2019b, MNRAS, 484, 3879

Planck Collaboration XIII, 2016, A&A, 594, A13

Posti L., Fall S. M., 2021, A&A, 649, A119

Price-Whelan A. M. et al., 2018, AJ, 156, 123

Richard J. et al., 2019, The Messenger, 175, 50

Robotham A. S. G. et al., 2011, MNRAS, 416, 2640

Sanchez-Blazquez P. et al., 2006, MNRAS, 371, 703

Sérsic J. L., 1968, Atlas de Galaxias Australes. Observatorio Astronomico, Cordoba, Argentina

Somerville R. S., Popping G., Trager S. C., 2015, MNRAS, 453, 4337

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

Speagle J. S., 2020, MNRAS, 493, 3132

Speagle J. S. et al., 2019, MNRAS, 490, 5658

Suyu S. H., Marshall P. J., Hobson M. P., Blandford R. D., 2006, MNRAS, 371, 983

Taylor R. et al., 2011, MNRAS, 418, 1587

Treu T., Stiavelli M., Casertano S., Møller P., Bertin G., 2002, ApJ, 564, L13

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

van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22

van der Walt S., Schönberger J. L., Nunez-Iglesias J., Boulogne F., Warner J. D., Yager N., Gouillart E., Yu T., 2014, PeerJ, 2, e453

Van Rossum G., Drake F. L., 2009, Python 3 Reference Manual. CreateSpace, Scotts Valley, CA

Velander M. et al., 2014, MNRAS, 437, 2111

Virtanen P. et al., 2020, Nature Methods, 17, 261

Warren S. J., Dye S., 2003, ApJ, 590, 673

Zehavi I., Contreras S., Padilla N., Smith N. J., Baugh C. M., Norberg P., 2018, ApJ, 853, 84

Zehavi I., Kerby S. E., Contreras S., Jiménez E., Padilla N., Baugh C. M., 2019, ApJ, 887, 17

Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, MNRAS, 443, 3044

APPENDIX A: PREPARING DATA FOR MODELLING

Images and weight maps are 101×101 pixel ($\sim 20 \times 20$ arcsec^2) cutouts from coadded images of KiDS tile observations acquired from the publicly available Astro-WISE Data base Viewer Web Service.[4] g- and r-band images are cut out centred on the object's RA and Dec., recentred to the brightest pixel in the central (lens) galaxy light profile, and converted to eps (electrons per second) for modelling. KiDS image pixel values in the Astro-WISE Data base are given in calibrated flux units relative to the flux corresponding to magnitude 0 and are converted to 'brightness' units of electron counts by multiplying by the tile's average gain, which includes additional factors necessary for this conversion. PYAUTOLENS is by default set to be optimally utilized with units of electrons per second (eps), which is acquired by dividing by the exposure time (1800 s for r-band, 900 for g-band).

PYAUTOLENS requires input of the PSF and noise map for each image. The inverse square root of the weight map corresponding to

---

[4] dbview.astro-wise.org

the cutout image gives the rms noise, which is converted to electron counts and squared to recreate the background sky. We then add this image to the corresponding cutout image and take the square root to give the noise map, after which we convert to eps. We generate a Gaussian PSF for each image from the average full width at half maximum (FWHM) of the PSF for each image.

We next mark pixel-positions of the distorted images of the lensed background source in each image, when visible, using a GUI distributed with PYAUTOLENS. During a lens model fit, PYAUTOLENS casts aside all mass models where these image-pixel positions do not trace within a designated threshold of one another in the source plane. This narrows the parameter space that is searched and ensures that the model fits the observed image features of interest.

We generate three masks for the three searches with each candidate: (i) *lens mask* – a circular aperture tailored to show only the lens galaxy (on the order of but usually slightly less the effective radius, typically around 1–1.3 arcsec); (ii) *source mask* – a circular annular aperture showing only the light we determine to be the lensed background source features (with inner radius about the size of the circular lens mask and outer radius around 3 arcsec); and (iii) *full mask* – a circular aperture of typically around 3 arcsec that includes most of the light from the lens and source features and masks as many peripheral contaminants as possible.

APPENDIX B: LENS MODELLING PIPELINE

This section details continue the description of our lens modelling methods summarized in Section 4.1. Through experimentation, we have designed a pipeline composed of a chain of three DYNESTY searches that we use as a template for fitting each lens. The variety of lensing configurations, image quality, etc. force us to tailor aspects of each model-fit individually, in particular alternating the masks that segment the foreground lens light and distorted background source light. In order to institute the least bias possible, we allow the models to probe a wide range of possible solutions for each parameter. The shape of the prior distribution has significant effects on the performance of the search. We use uniform, log uniform, and Gaussian functions depending on the parameter and informative auxiliary observations. In the following sections, we describe this three-step automated pipeline, where from here on we refer to a 'search' as a model-fit performed by the non-linear search DYNESTY. Each subsequent search in the chain has more complexity in the form of additional parameters, which we balance in computational time by passing priors from previous search outputs. For each non-linear search in the chain, the priors are described in Table B1, and the DYNESTY settings are given in Table B2.

B1 Search 1 – lens light

Search 1 is the simplest and quickest of the three searches and focuses on returning an accurate lens light profile. The subtraction of this modelled light from the observed image should then show the lensed features of the background source. This search fits an elliptical Sérsic profile (Sérsic 1968),

$$I(R) = I_e e^{-b_n[(\frac{R}{R_e})^{1/n}-1]} \qquad (B1)$$

where R is angular radius from the centre of the profile, I_e is the intensity at the effective radius R_e, $b_n \approx 2n - 0.327$, and n is the Sérsic index. PYAUTOLENS generates an image from these parameters in the image-plane and fits to the observed r-band image. The purpose of this search is to infer a high likelihood Sérsic lens light model,

which serves two purposes for search 2: (i) it provides a lens-light subtracted image and; (ii) it provides lens light priors that are passed to subsequent searches. Because the image is centred during pre-processing, the distribution can be initialized with fairly tight constraints. Elliptical components, ϵ_1 and ϵ_2, are defined as

$$\epsilon_1 = \epsilon_y = \frac{1+b/a}{1-b/a}\sin 2\alpha \qquad (B2)$$

$$\epsilon_2 = \epsilon_x = \frac{1+b/a}{1-b/a}\cos 2\alpha \qquad (B3)$$

where b and a are the semimajor and semiminor axes of the ellipse, and α is the position angle. The intensity is parametrized according to electrons per second and therefore takes a wide log-uniform distribution. The Sérsic index prior covers a wide range of reasonable values with a uniform distribution.

For the r-band images, many of the lensed background source's features are positioned within the lens galaxy's effective radius. Search 1 therefore struggles to deblend the lens and source light, and the distorted arcs of background source light are attributed to the foreground lens galaxy. In some cases, this leads to a model solution that describes a lens light profile that is very large and very elliptical. To mitigate this systematic effect, we use the aforementioned *lens* mask for search 1 and constraints on the effective radius to assist the search to focus on fitting the lens light and not the source light. The residuals and uncertainties of this search therefore tend to be quite high. In fact, the residuals often outline the lensed images of the source itself, and the resulting maximum log-likelihood is lower than the value inferred in the second and third searches.

B2 Search 2 – Lens mass and source light

Search 2 focuses on the light from the background source. This component is modelled as a spherical exponential light profile in the source plane defined at the source redshift. The exponential light profile corresponds to the simple $n = 1$ case of equation (B1) and is parametrized using its (source-plane) centre, effective radius, and intensity. With higher-resolution imaging data, the background source light profile could be fit with a more detailed model. The background source centre coordinates are initialized to values within 2 arcsec of the line of sight of the foreground lens centre. The intensity of the background source light is again set to a wide log-uniform prior distribution, as for the lens light in Search 1. The source effective radius is initialized with a Gaussian distribution around a typical disc galaxy size, as discussed in Section 4.2. To map coordinates to the source plane, the lens galaxy's total mass is modelled as a singular isothermal elliptical (SIE) profile. The mass profile's Einstein radius prior is a wide Gaussian distribution centred at 1.0 arcsec with a hard upper limit of 2.5 arcsec. The centre and elliptical components of the SIE are paired with the light profile with the assumption that the ellipses will be aligned. The lens light profile takes prior distributions passed from the results of search 1. By fixing the centre of the lens profiles to the results of search 1, the lens model in this search is reduced to 10 free parameters. This, in addition to taking informed Gaussian priors from search 1 for the lens light, helps the model to focus on solutions that fit the source-light instead of systematic solutions that fit artefacts in the data. We use the annular *source* mask that removes the lens light from the observed image and therefore further focuses the model on fitting the source light. We also utilize PYAUTOLENS's position resampling functionality, whereby the brightest pixels in the lensed source are marked (via a GUI). Again, the results of this search are passed as priors to Search 3.

Table B1. Details about model searches and priors for three-step lens model-fitting with PYAUTOLENS. Each phase fits a number of free parameters that model light and mass profiles of the lens and source galaxies by exploring the parameter space according to the prior's PDF. Parameters fit in searches 1 and 2 are input as Gaussian or fixed priors for subsequent searches. 'Elliptical components' are related to the axis ratio and position angle as in equations (B2) and (B3). See Sections B and 4.2 for more details about searches, profiles, and priors.

Search	# Free Parameters	Fit	Profile	Prior	PDF
1	7	Lens light	Elliptical Sérsic	Centre (y, x)	Uniform (-0.3–0.3 arcsec)
				Elliptical comps (ϵ_1, ϵ_2)	Gaussian (mean = 0.0, σ = 0.3)
				Intensity	Log Uniform (10^{-6}–10^6 eps)
				Effective radius	Gaussian (GAMA-DR3 r mean and σ arcsec *or* SLACS 7 kpc $\pm$ 3.3 at lens distance, upper limit = mean + 3σ)
				Sérsic Index	Uniform (0.5 - 8.0)
2	10	Lens light	Elliptical Sérsic	Centre (y, x)	Prior passed from search 1 (fixed)
				Elliptical comps (ϵ_1, ϵ_2)	Prior passed from search 1 (Gaussian)
				Intensity	Prior passed from search 1 (Gaussian)
				Effective radius	Prior passed from search 1 (Gaussian)
				Sérsic index	Prior passed from search 1 (Gaussian)
		Lens stellar mass	Elliptical isothermal	Centre (y, x)	Paired to lens light prior (fixed)
				Elliptical comps (ϵ_1, ϵ_2)	Paired to lens light prior (Gaussian)
				Einstein radius	Gaussian (mean = 1.0, σ = 0.5, limit = 0–2.5 arcsec)
		Source light	Spherical exponential	Centre (y, x)	Uniform (-2.0–2.0 arcsec)
				Intensity	Log Uniform (10^{-6}–10^6 eps)
				Effective radius	Gaussian (7.5 kpc ± 2.5 at source distance, upper limit = mean + 3σ)
3	14	Lens stellar light and mass	Elliptical Sérsic	Centre (y, x)	Prior passed from search 2 (fixed)
				Elliptical comps (ϵ_1, ϵ_2)	Prior passed from search 1 (Gaussian)
				Intensity	Prior passed from search 2 (Gaussian)
				Effective radius	Prior passed from search 2 (Gaussian)
				Sérsic index	Prior passed from search 2 (Gaussian)
				Mass-to-light ratio	Log uniform (limits calculated)
		Lens dark mass	Elliptical NFW	Centre (y, x)	Paired to stellar mass prior (fixed)
				Elliptical comps (ϵ_1, ϵ_2)	Gaussian (mean = 0.0, σ = 0.3)
				κ_s	Uniform (0.0–1.0)
				Scale radius	Gaussian (calculated from SLACS-IV mean and σ)
		Source light	Spherical exponential	Centre (y, x)	Prior passed from Search 2 (Gaussian)
				Intensity	Prior passed from Search 2 (Gaussian)
				Effective radius	Prior passed from Search 2 (Gaussian)

Table B2. DYNESTY non-linear search settings for each of the three searches of model-fitting. These settings balance computational cost with a thorough exploration of parameter space. Relaxed settings (e.g. low *n* live points and high evidence tolerance) are useful for expediting initial fits that inform later fits. The trade-off is less well-defined uncertainty and a chance that the global maximum likelihood fit has been missed in favour of a local one. See Section 4.1 for a thorough description of PYAUTOLENS and DYNESTY search settings.

Search	*n* live points	Evidence tolerance	Steps per walk	Acceptance fraction	Positions threshold	Subgrid size
1	200	0.5	10	0.3	N/A	2 × 2 subpixels
2	300	0.25	10	0.3	1.5 arcsec	2 × 2 subpixels
3	500	0.25	10	0.3	1.5 arcsec	2 × 2 subpixels

B3 Search 3 – Combined lens and source models

Search 3 fits every component of the system. To model the foreground lens, we use a combined elliptical Sérsic mass-light profile for the stellar component and an elliptical NFW profile for the dark matter halo. Background source light is modelled again as a spherical exponential profile. Priors are passed from search 2 (see Table B1), except for the dark matter halo profile and stellar mass-to-light ratio. The prior distributions for these crucial parameters are determined by calculating central and limiting values according to a more careful process, as described in Section 4.2. The *full* mask including all the lens and source features is used to remove background features

and other contaminants that exist in the periphery, which saves computational time. Lensed image positions are again used to discard unphysical mass models. This final search produces a reasonable fit to the complexities introduced by each component and gives uncertainties on each of the inferred quantities. Additional disc and bulge components, cores, and multiple galaxies at different planes along the line of sight can be fitted and would allow more precise and realistic models. These improvements to realism are unhelpful here given the quality of imaging available for the objects in question but would be simply applied in future studies following the same principles outlined in our strategy.

APPENDIX C: HIGHEST-QUALITY MODEL RESULTS

Figs C1-C4 show the observed image, model image, and spectra for some of the most successful models.
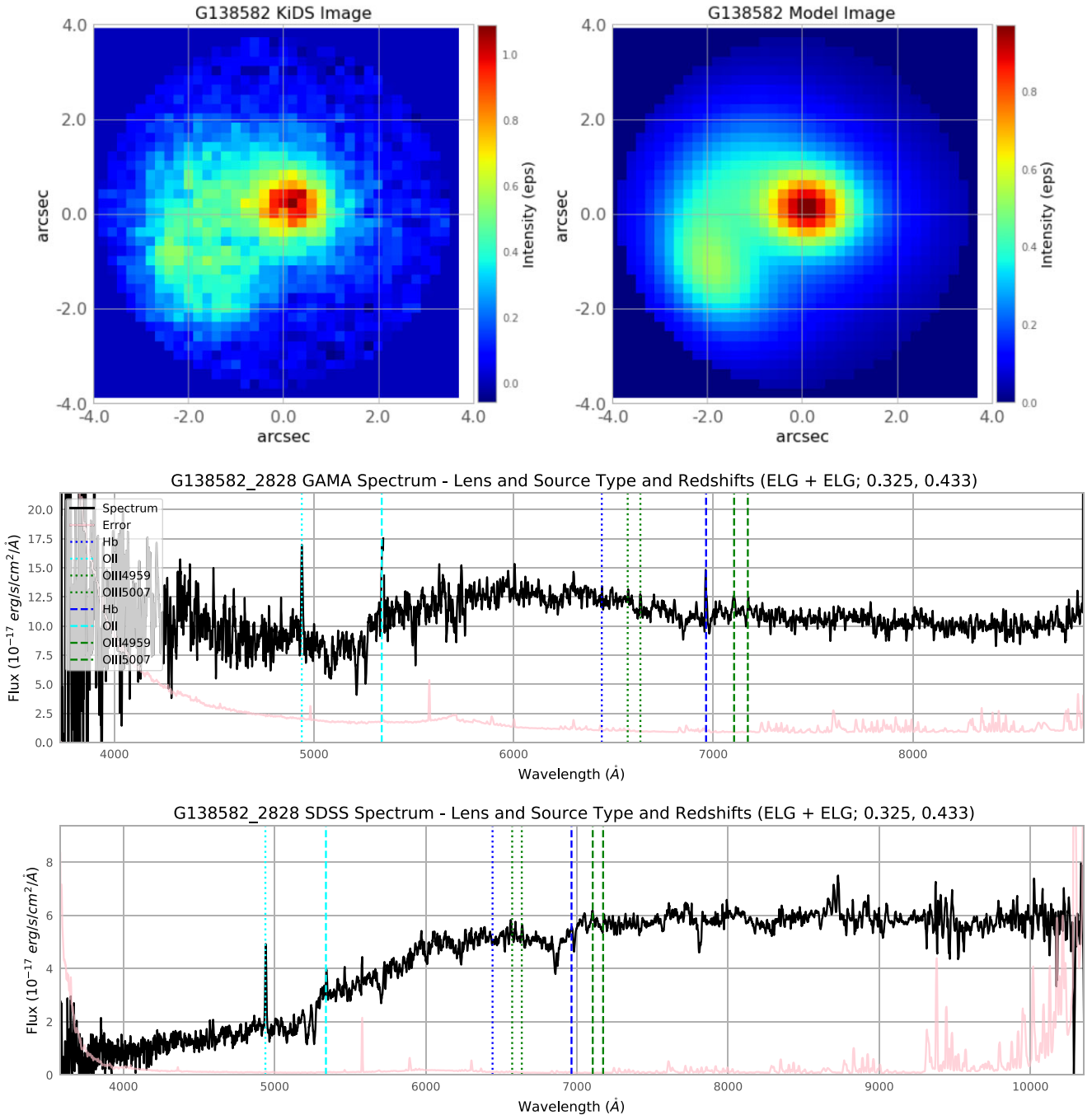
Figure C1. G138582. A-Grade. *Upper left*: The observed image shows a single bright elongated feature in the lower left of the foreground lens galaxy profile with a tail in the upper part of the feature. *Upper right*: The model image correctly captures the shape of the image with lensing characteristics. *Lower*: The GAMA and SDSS spectra both show reasonably strong emission lines (Hβ, [O II], [O III]) for the foreground lens galaxy at $z = 0.325$ (dotted) and the background source galaxy at $z = 0.433$ (dashed). Foreground lens CaH&K absorption lines are also easily identified but are left unmarked here to show the presence of emission lines for this spectrum's ELG + ELG AUTOZ template match.
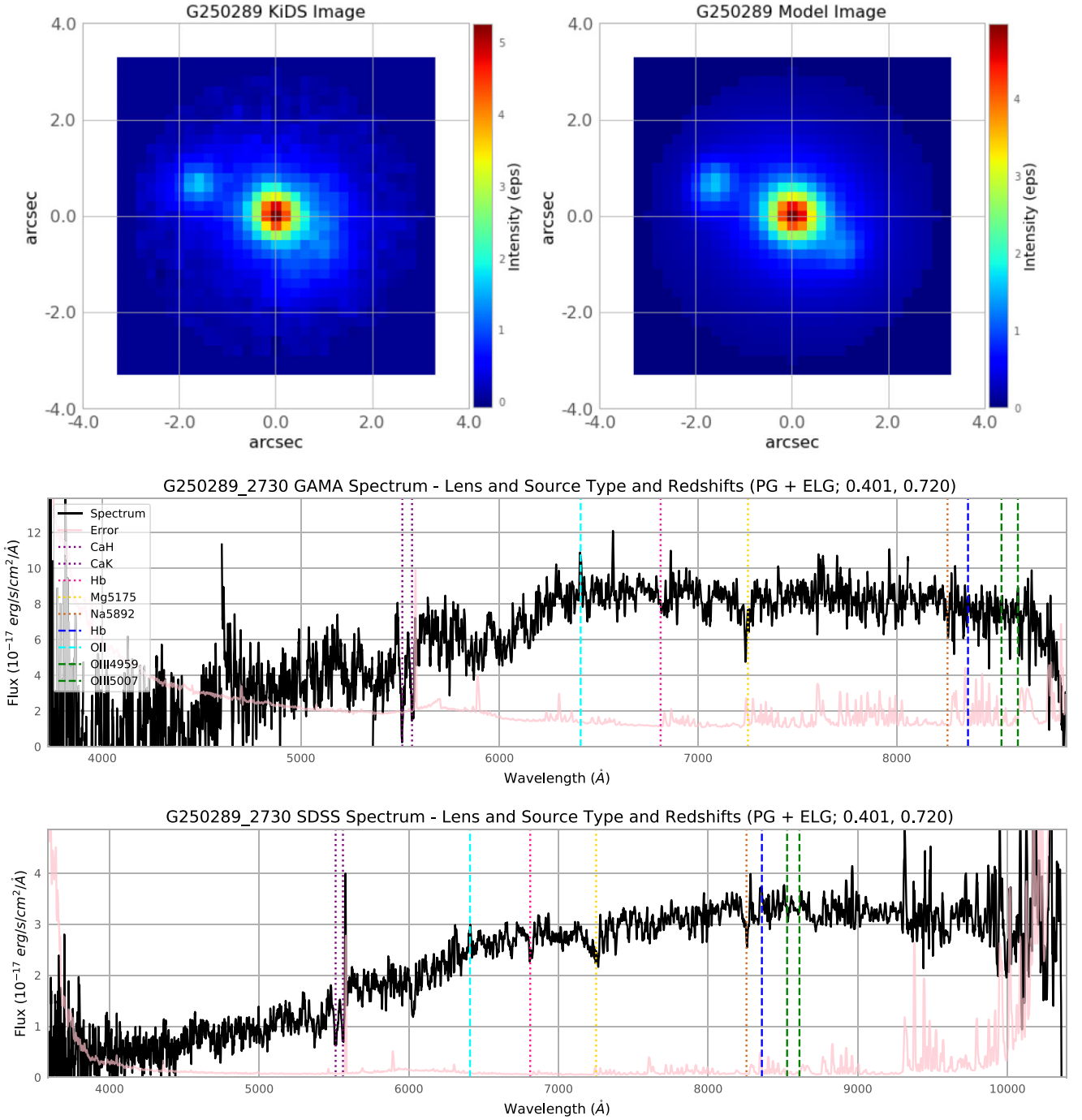
Figure C2. G250289. B-Grade. *Upper left*: The observed image shows a doubly imaged source with a near-elliptical shape in the upper left with respect to the foreground lens and an arc mirrored across the lens to the lower right that blends somewhat with the foreground lens light. *Upper right*: The model reconstructs the locations of both images, but the mirrored image in the lower right lacks the stretched elongated shape, which may be an effect of internal structure that is unaccounted for in the model. *Lower*: The absorption features shown with dotted lines (CaH, CaK, Hβ, Mg, and Na) of the foreground lens galaxy at $z = 0.401$ are particularly strong. Emission of [O II] from the source galaxy at $z = 0.720$ appears in both spectra with dashed lines, as well as weaker features from Hβ and [O III]. However, the SDSS spectrum appears to show a stronger [O III] $\lambda4959$ than [O III] $\lambda5007$, which should not be the case. The weak background source-flux could be because much of the the upper left source feature in the observed image is outside the 1- and 1.5-arcsecond GAMA and SDSS apertures.
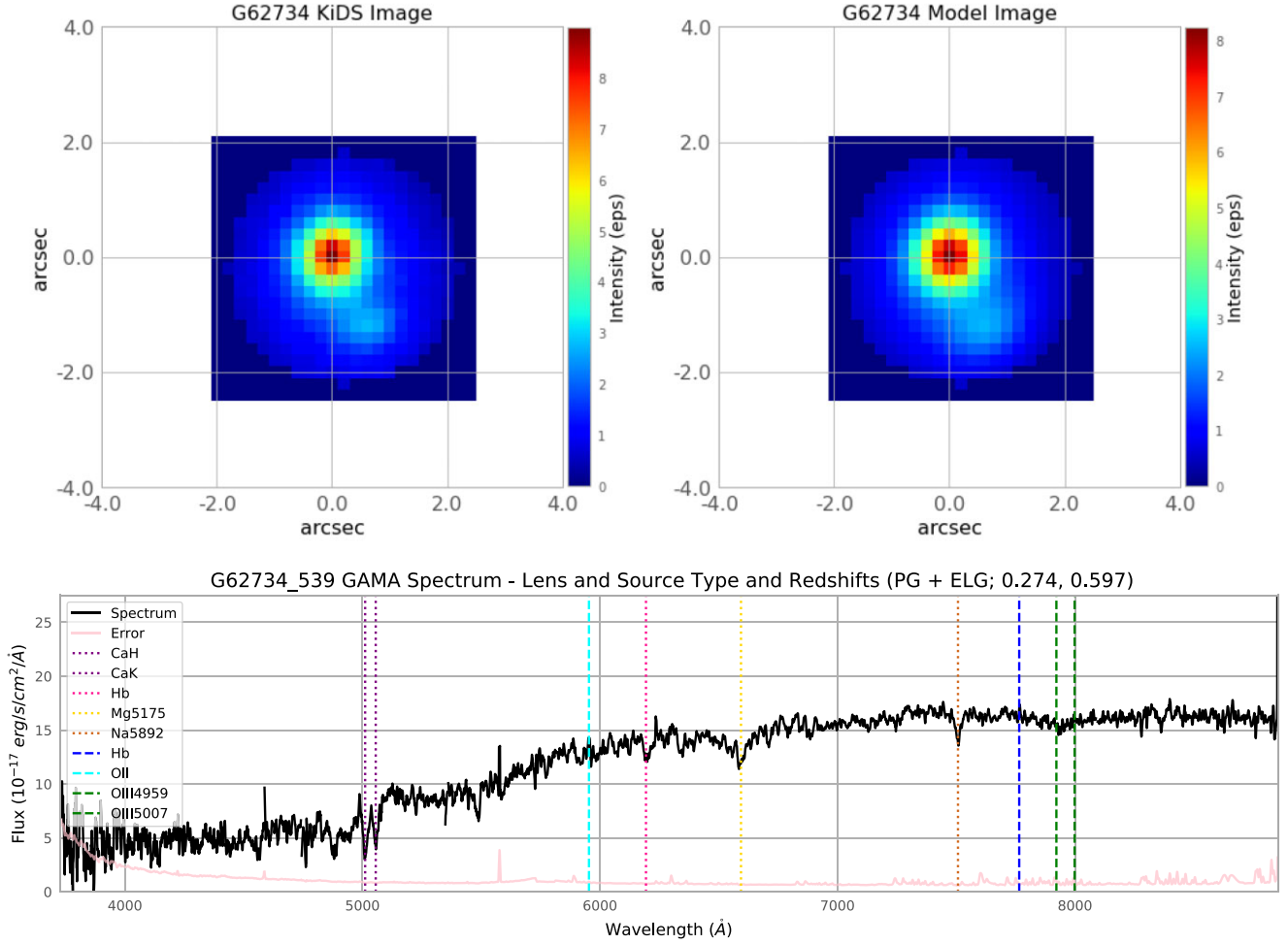
Figure C3. G62734. B-Grade. **Dark mass poorly constrained, so not included in further analysis alongside the other A- and B-grade models.** *Upper left*: The observed image shows an image in the lower right with respect to the foreground lens profile. *Upper right*: The shape and location of the lensed source feature in the lower right are well-fit in the model image, but there is some extra light surrounding the lensed source feature that may be due to the less-sophisticated spherical exponential light profile that we use to model the source light. The exact reconstruction of the source light profile is not the main goal of this exercise, though higher resolution imaging would make it worth further constraining with more flexible priors. *Lower*: Foreground lens absorption features (CaH, CaK, Hβ, Mg, and Na) are clearly shown with dotted lines at $z = 0.274$, and weak emission features can be identified with dashed lines at $z = 0.597$.
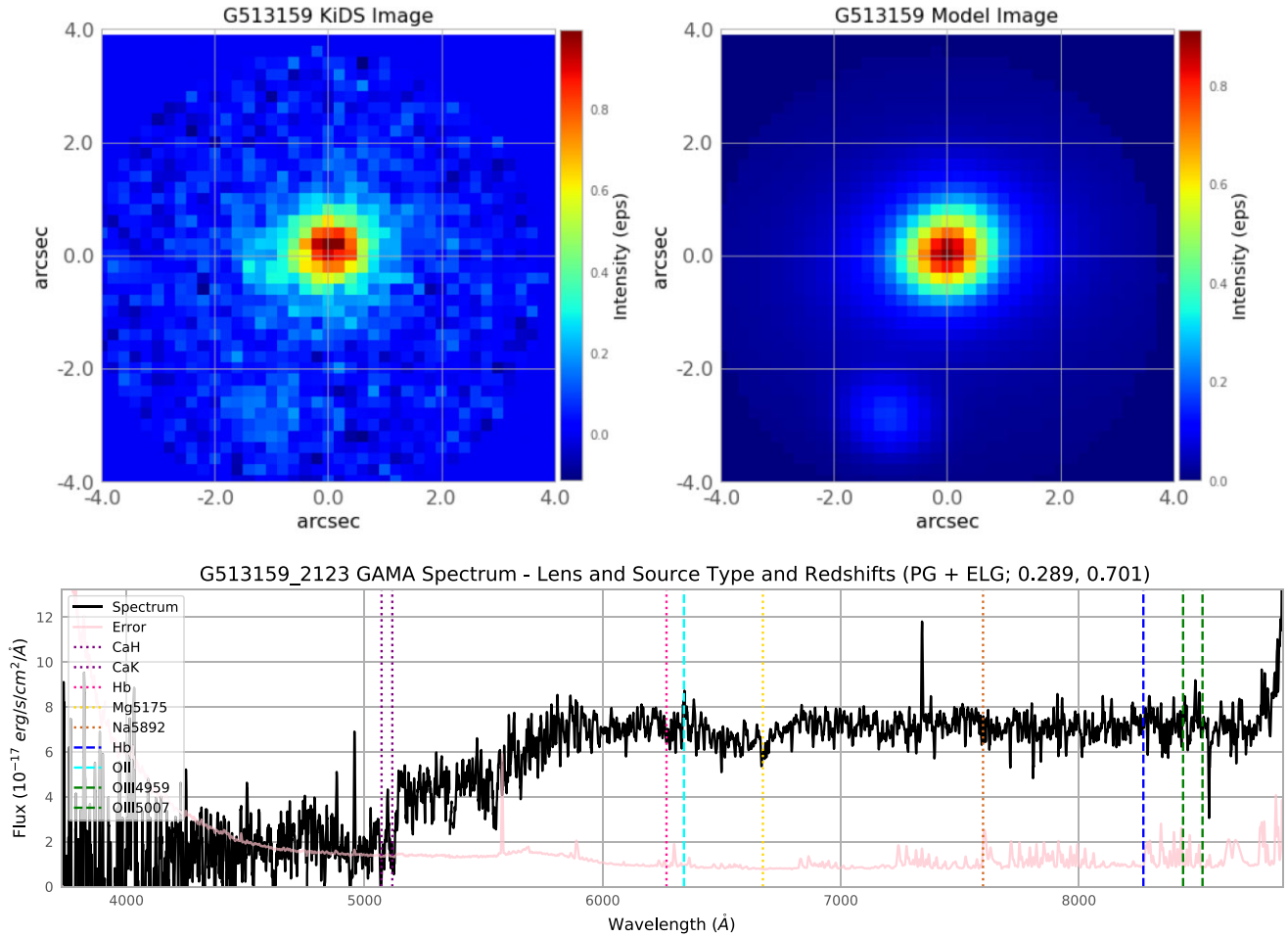
Figure C4. G513159. B-Grade. *Upper left*: The observed image shows a feature around 3 arcsec away from the central foreground lens profile that may be a lensed source feature. *Upper right*: The model successfully accounts for the position and flux of the extra light through lensing. *Lower*: The GAMA spectrum shows strong CaH and CaK features with dotted lines for the foreground lens galaxy at $z = 0.289$ and possible emission line features ([O II] and [O III]) with dashed lines at $z = 0.701$. Some expected features are plotted but not well defined in the spectrum.

APPENDIX D: SPECTRUM QUALITY CONTROL

We return to the specific cases described in Section 3.2 to see how they affected our final quality scoring and subsample selection in the interest of retaining the most true positives while minimizing the inclusion of false positives.

D1 When there are overlapping emission or absorption lines...

Almost half of the candidates (20 of 42) we selected initially by AUTOZ output had overlapping line features of some kind in their spectrum. 12 of these were overlapping absorption features; eight were emission features. Six of the 19 candidates that were accepted following critical quality control had overlaps. The presence of an overlap affected the scoring of the individual spectrum, which was accepted only if the other background source line features were well-shown. We classify the cases of overlap as 'on-template' or 'off-template' in reference to the background source template. 'Off-template' overlaps are cases when the overlapping line is an emission or absorption feature for a background source PG or ELG, respectively, as opposed to 'on-template' overlaps, where the overlap is emission or absorption for ELG or PG respectively. Nine of the 20 cases of overlap were 'on-template'. The other 11 were 'off-

template'. The six overlapping cases that were retained in the final subsample of 19 graded models consisted of one on-template and five off-template overlaps.

Seven of the eight candidates with overlapping *emission* features include background source ELGs (three PG + ELG and four ELG + ELG), and one is a background source PG (ELG + PG). Two of these are retained in the 19 graded models. Recall from Section 3.2 that all emission line overlaps are between the lens [O III] $\lambda\lambda$4959,5007 couplet and the source [O II] λ3727. It appears that these emission lines can have a significant effect even when one of the templates is a PG. Of the 12 *absorption* feature overlaps, eight were PG + ELG, two were ELG + ELG, and two were PG + PG. Five of these off-template PG+ELG absorption line overlaps make our final selection of 19 candidates, with a grade B, two C's, and two D's. One of the highest-scoring candidates (G250289, PG + ELG) had an overlap of foreground-lens Hβ and background-source CaK absorption features, but the emission lines from the background source were well defined and gave confidence to the redshift determination. This case is a bit odd considering the background galaxy was fit to an ELG template and would presumably be most heavily weighted by emission features. None of the ELG+ELG or PG + PG configurations with overlaps were accepted. One of the two candidates noted in Section 3 that were removed from the
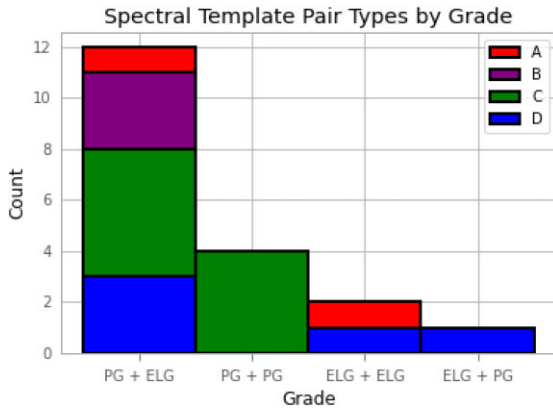
Figure D1. Stacked histogram of the four possible configurations of PG and ELG, written as foreground + background, separated by their quality grade (A, B, C, and D) as described in Section 5. The large majority of successful models were composed of a passive foreground lens galaxy and emission line background source galaxy, which is expected. Other configurations are less likely, but one of the two A-grade models came from an ELG + ELG configuration.

Holwerda et al. (2015) sample had overlapping features. The other had a reasonable spectrum and failed for other reasons.

D2 When the lens is described as an emission line galaxy...

As discussed in Section 3.2, the case of an ELG acting as the foreground lens is less likely than a case where a PG acts as the lens. Only 3 ELG+ configurations were retained in the graded subsample of 19 candidates, two of which were low-scoring D-grades. Only one of the ELG + PG configurations was accepted and given a D-grade. Fig. D1 shows the foreground + background configurations for the 21 candidates in the final selection in the same manner shown in Fig. 2, now with quality grades. A, B, C, and D grades are blue, green, purple, and red, respectively. Interestingly, one of the highest scoring (A-grade candidate G138582) candidates was one of the ELG + ELG matches. As shown in Fig. C1, the emission lines from lens and source are clearly determined, and the resulting model was one of the most successful of this study. This example highlights the potential value of including (though with critical evaluation) the ELG+ foreground lens template configurations in the selection. Still, the template configurations shown in Fig. D1 mostly reaffirm the validity of the assumption that passive large elliptical galaxies provide the clearest and most usable foreground lenses. Further, since more background source +ELG template configurations have higher scores relative to + PG configurations, this again shows that the flux from strong emission lines in the background source is more detectable than the continuum and absorption features of a PG.

D3 When source emission lines are redshifted beyond observed wavelength range...

27 of the initial 42 Autoz spectra had + ELG configurations (i.e. background source is an ELG), eight of which had Hβ and [O III] $\lambda\lambda$4959,5007 emission lines redshifted beyond the GAMA upper wavelength limit of 8850 Å. These features would be present in the longer-wavelength upper range of the SDSS-BOSS spectrum for all 27 + ELG candidates, but not all were measured in SDSS-BOSS. Three of the 19 candidates had all three above line features redshifted beyond the survey upper wavelength limit. Because these

three objects were also measured with SDSS-BOSS spectroscopy and included in GAMA-DR3 SpecAll, their emission lines redshifted beyond 8850 Å were detectable, but the Autoz match did not have access to those wavelengths. These were two C-grades and a D-grade.

D4 When primary redshift is background source...

10 of the initial 42 aAutoz spectra featured higher cross-correlation peaks to the background source than to the foreground lens (i.e. σ_1 is the match to the background source). Two of those are included in the final graded 19 models. These two cases are shown in Table 5. One of these is one of the two highest-scoring candidates (A-grade, candidate G323152, PG + ELG). G323152 represents the case described in Section 3.2 where very strong emission lines from the background source are interpreted as the primary redshift match instead of the lower redshift passive continuum. The other candidate with σ_1 assigned to the background source flux is a D-grade with ELG+PG configuration. As mentioned before, we expect this configuration with the primary match to the background source redshift to be far less likely. Still, as with the ELG + ELG matches discussed in the previous section, the A-grade example of this case reinforces the value of including the Autoz configurations where σ_1 is at higher redshift than σ_2.

D5 Additional curiosities, overlaps, failures of our utilization of AUTOZ

Two of the ELG + PG configurations show the lens [O III] λ4959 line straddled by the H and K lines of the background source PG. This is a case where a 'peak' between the two absorption valleys can be mistakenly considered an emission line feature. These are two of four foreground lens redshift matches below z ∼ 0.1. The other two have overlaps between lens [O III] λ5007 and source [O II] λ3727. A revision of the initial selection strategy could have extended the redshift cutoff to z ∼ 0.1 with no change to the sample. Two others appear to have emission lines fairly close to absorption lines, which might also give the impression of a peak or valley where it actually does not exist. One of these was accepted in the 19 and was given a grade of D.

APPENDIX E: SUBSAMPLE SELECTION AND CONTEXT

We find that the majority of machine learning candidates did not pass our selection criteria covering Autoz output parameters. This is predictable in light of the results of Knabel et al. (2020).

From the 421 LinKS candidates in the GAMA equatorial regions, there are 348 matching Autoz entries (including duplicates) for 300 unique LinKS candidates. 59 of these entries pass the $R \geq$

Table E1. Duplicate Autoz entries for LinKS lens candidates. Boldface text indicates the selected entry. Type refers to foreground + background template matches. z_1 and z_2 refer to redshift matches corresponding to Autoz cross-correlation peaks σ_1 and σ_2. R is a parameter that weights σ_2 to third and fourth matches.

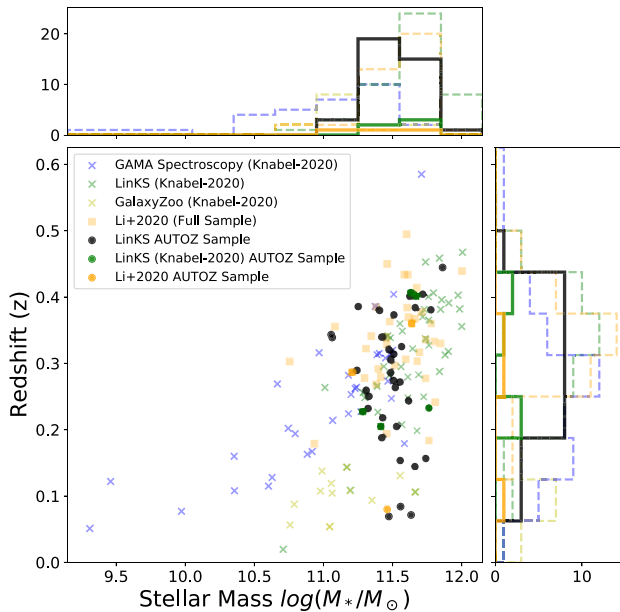| GAMA ID	Type	z_1	z_2	σ_1	σ_2	R
G544226	**PG + ELG**	**0.227**	**0.650**	**9.393**	**7.240**	**2.122**
	PG + ELG	0.650	0.227	6.294	6.410	0.650
G262874	**ELG + ELG**	**0.386**	**0.859**	**6.222**	**3.422**	**1.217**
	ELG + PG	0.386	0.195	9.339	4.817	1.416

Figure E1. Stellar masses and redshifts of the AUTOZ sample with deeper coloured circular markers shown against the candidates discussed in Knabel et al. (2020) with faded X's for context. LinKS candidates (shown in green for the LinKS subsample selected in Knabel et al. (2020) and black for those that were not) and 'BG' candidates from (Li et al. 2020; orange) have high stellar masses at intermediate redshift $log(M_*/M_\odot) \sim 11$–11.75 at $z \sim 0.2$–0.5. Blue and yellow X's are spectroscopy and citizen-science candidates selected in Knabel et al. (2020).

1.2 criterion, and 56 of these have galaxy–galaxy template matches. Four of those entries are duplicates, leaving 52 candidates (including six from Knabel et al. 2020). We remove 12 of these through our redshift criteria, which leaves 42 (40 unique) LinKS AUTOZ foreground + background redshift matches. The two duplicates are shown in Table E1, with the accepted matches in bold text. For G544226, both entries show a PG template match at redshift $z = 0.227$ with an ELG at redshift $z = 0.650$. The accepted entry shows higher σ_1, σ_2, and R, and it attributes the primary redshift match to the foreground lens galaxy. The other entry is an example where σ_1 can refer to the background source galaxy and σ_2 to the foreground lens galaxy, effectively reversing which shows 'better' match while still identifying the redshifts and type correctly. Note that σ_1 and σ_2 for the rejected entry are quite close (6.294 and 6.410, respectively). Both entries for the other duplicate candidate show the same primary match. The entry that is rejected has a secondary match to an ELG template at much closer redshift, which is most likely a false match. We remove the one LinKS candidate with a low redshift success probability and are left with 39 LinKS AUTOZ-selected candidates,

six of which were included in the final LinKS candidate selection of Knabel et al. (2020).

32 of 48 Li-BG candidates in the GAMA equatorial fields have a match in AUTOZ, with 53 entries including duplicates. Eight candidates (with no duplicates) are selected by the R criterion. Five of those eight candidates are removed by our redshift criteria, leaving three unique candidates for analysis. One GalaxyZoo candidate has a match in the AUTOZ catalog, but it does not pass selection criteria for followup.

In order to briefly contextualize this selection in reference to some of the results and conclusions drawn in Knabel et al. (2020), we show the AUTOZ sample of 42 candidates selected in this work in Fig. E1, with circular markers in comparison with the candidates discussed in Knabel et al. (2020) shown in the background with X's. Stellar mass estimates and lens redshifts shown here are from GAMA-DR3 STELLARMASSESLAMBDAR catalog. The AUTOZ sample is slightly lower in stellar mass on average than the LinKS subsample as selected in Knabel et al. (2020), with a mean and median $\log M_*$ of (11.50, 11.51) compared to (11.61, 11.67). A Kolmogorov–Smirnov test of the stellar masses between the LinKS AUTOZ subsample and the LinKS subsample as selected in Knabel et al. (2020) results in a KS-metric of 0.352 with a p-value of 0.007, indicating a statistically significant disparity between the masses of the two selections. In fact, when compared to the GAMA spectroscopy subsample, as selected in Knabel et al. (2020), the KS-test results are almost identical (metric 0.353, p-value 0.007). The bulk of AUTOZ candidates hovers in the parameter space overlapping the upper mass end of the GAMA spectroscopic candidates and the lower mass end of the LinKS from Knabel et al. (2020) candidates, which is reasonable if they are to be large enough to have distinguishable features for identification by machine learning while being small enough to have a higher chance of flux from the lensing features being collected in the 1 arcsec GAMA spectroscopic fiber aperture.

Two candidates in the AUTOZ sample have σ_2 and R values that would place them in the selection space defined for the Holwerda et al. (2015) blended spectra candidates. One of them (G184530) was not selected in that study because it is an ELG + PG configuration (i.e. the emission line match is at closer redshift). The other (G544226) was removed because Holwerda et al. (2015) removed candidates near the alias of $(1 + z_1)/(1 + z_2) = 1.343 \pm 0.002$, corresponding to an overlap between redshifted [O II] $\lambda 3727$ and [O III] $\lambda 5007$ emission lines. G544226 then would have been the one overlap between the GAMA spectroscopic and LinKS machine learning catalogs in Knabel et al. (2020) if it had not been removed. G544226 made the selection for high-quality candidates in Knabel et al. (2020). With a redshift of $z = 0.227$ and $\log M_* = 11.29$, it existed squarely in the overlap of parameters space between the GAMA spectroscopy and LinKS machine learning candidates.

This paper has been typeset from a TEX/LATEX file prepared by the author.