Statistical strong lensing: IV. Inferences with no individual source redshifts

Sonnenfeld, A.

Statistical strong lensing. IV. Inferences with no individual source redshifts

Alessandro Sonnenfeld[1]

Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, the Netherlands
e-mail: sonnenfeld@strw.leidenuniv.nl

ABSTRACT

Context. Strong lensing mass measurements require the knowledge of the redshift of both the lens and the source galaxy. Traditionally, spectroscopic redshifts are used for this purpose. Upcoming surveys, however, will lead to the discovery of $\sim 10^5$ strong lenses, and it will be very difficult to obtain spectroscopic redshifts for most of them. Photometric redshift measurements will also be very challenging due to the blending between lens and source light.
Aims. The goal of this work is to demonstrate how to carry out an inference of the structural properties of the galaxy population from the analysis of a set of strong lenses with no individual source redshift measurements, and to assess the loss in precision compared to the case in which spectroscopic redshifts are available.
Methods. Building on the formalism introduced in Paper III, I developed a method that allows a statistical strong lensing inference to be carried out while marginalising over the source redshifts. This method, which relies on the knowledge of the properties of the unlensed background source population and of the selection function of the survey, generalises an approach known as photogeometric redshift, originally introduced by the Strong Lensing Legacy Survey collaboration. I tested the method on simulated data consisting of a subset of 137 strong lenses that is complete above a cut in observational space.
Results. The method recovers the properties of the galaxy population with a precision that is comparable to that attainable in the case in which individual source redshifts are known.
Conclusions. The photogeometric redshift method is a viable approach for the analysis of large sets of strong lenses provided that the background source population properties and lens selection function are well known.

Key words. Gravitational lensing: strong – Galaxies: fundamental parameters

1. Introduction

Strong gravitational lensing allows for some of the most precise mass measurements in astrophysics. The angular separation between strongly lensed images depends on the mass distribution of the lens, as well as on the distances between the observer, the lens, and the background source. When the redshifts of the lens and source galaxies are known, by modelling the angular structure of the strongly lensed images it is possible to obtain measurements of the total projected mass enclosed within the images with a precision of a few percent, as that is the typical uncertainty associated with the lens modelling step (Bolton et al. 2008).

Traditionally, lens and source redshifts are obtained via spectroscopy. Spectroscopic follow-up observations, however, can be expensive in terms of exposure time. In the case of the Strong Lensing Legacy Survey (SL2S; Ruff et al. 2011; Gavazzi et al. 2012), for example, 1-hour-long observations on 8-metre-class telescopes were needed to obtain the redshift of the background sources. Moreover, most strongly lensed sources span the redshift range $1 - 3$ (Sonnenfeld et al. 2013). A substantial fraction of them do not have any strong emission lines in the optical region of the spectrum, and as a result they go undetected when observed with optical spectroscopy (Ruff et al. 2011; Sonnenfeld et al. 2013). Near-infrared observations usually guarantee higher chances of success (Sonnenfeld et al. 2015, 2019), but very few of the existing spectrographs have both the broad wavelength coverage needed to maximise the chances of detecting an emis-

sion line and the high spectral resolution required to minimise the contamination from atmospheric emission.

Upcoming surveys such as Euclid[1], the Legacy Survey of Space and Time (LSST[2]) and the Chinese Space Station Telescope (CSST) will lead to the discovery of $\sim 10^5$ new strong lenses (Collett 2015). These lensing data have the potential to enable unique measurements of the mass structure of galaxies. Sonnenfeld & Cautun (2021, hereafter Paper I) showed how with a sample of 10^3 strong lenses it is possible to calibrate stellar mass-to-light ratio measurements of massive galaxies with 0.03 dex precision. The analysis of Paper I, however, was based on the assumptions that the redshifts of all lens galaxies and background sources were known from spectroscopy. It is unclear whether such spectroscopic data will become available in the near future.

The standard alternative to spectroscopy is photometric redshift (photo-z) measurements. The accuracy of photo-zs depends critically on that of the photometric measurements on which they are based. Most lenses are massive elliptical galaxies, which are bright and with an easily characterised spectral energy distribution. This makes it possible to obtain very precise photometric redshift measurements for them (see for example Vakili et al. 2019). Lensed sources, however, are fainter. Their images have typical $i-$band magnitudes around $24 - 25$ and are usually found $1 - 2''$ away from the lens centre (Sonnenfeld et al. 2019). Con-

---

[1] https://www.euclid-ec.org/
[2] https://www.lsst.org/

sequently, they are often blended with the light from the lens, which is usually brighter at all but the bluest wavelengths. This makes it difficult to determine the colours of lensed sources. As a result, photometric redshifts have so far found very little use in galaxy-scale strong lensing, with the exception of cases in which high-resolution imaging data in many bands are available (Sonnenfeld et al. 2012).

When it is not possible to obtain individual redshift measurements, the only remaining option is to rely on statistics: if the redshift distribution of the population of background galaxies is known, then it is possible to guess the source redshift of a lens, in a probabilistic sense. Since lensed sources typically span a broad redshift range, this approach can no longer provide percent mass measurements on individual systems. Nevertheless, high precision can be gained with the statistical combination of a large number of lenses.

A similar approach is used in weak lensing analyses, where, for the determination of the lensing critical surface mass density, the population distribution of the source redshifts, $P(z_s)$, is used instead of individual photo-z measurements (see for example Hildebrandt et al. 2017). In the case of strong lensing, however, there is an important complication due to the selection function of the strong lens population. The probability of a galaxy being strongly lensed by a foreground object varies with its redshift, and so does the probability of it being detected by a strong lens survey. As a result, the redshift distribution of strongly lensed sources is very different from that of unlensed galaxies: for example, the abundance of sources at a redshift close to that of the lenses is highly suppressed. Therefore, it is not possible to directly use an estimate of $P(z_s)$ relative to the general background galaxy population as a prior on the source redshift of a strong lens, as that would lead to a biased measurement.

Ruff et al. (2011) tackled this problem in the following way in their study of SL2S lenses. They assumed that galaxies have an isothermal density profile. Then, they estimated a prior probability distribution for the lens deflection strength, $P(\sigma_{SIE})$, of the lenses in the SL2S sample by taking the observed stellar velocity dispersion function of galaxies and re-weighting it to take into account the strong lensing cross-section and the luminosity selection of the parent galaxy sample. The prior probability $P(\sigma_{SIE})$ was then used, together with the observed Einstein radius of each lens, to obtain a posterior probability distribution of the source redshift. This method of obtaining the redshift of a strongly lensed source is known as photogeometric redshift.

There are two problems with the Ruff et al. (2011) method. The first is that it rests on the assumption of an isothermal density profile for the foreground galaxy population. The density profile of galaxies, while on average close to isothermal, varies with properties such as stellar mass, half-light radius, and redshift (Sonnenfeld et al. 2013). Therefore, the isothermal assumption introduces a bias. While this bias is probably not significant for the Ruff et al. (2011) analysis, it is set to become an important source of systematic errors once precision improves. More generally, if the purpose of the lensing observations is to determine the structural properties of the galaxy population, then the correct approach is to infer these properties jointly with the redshifts of the background sources, instead of keeping them fixed. The second problem is that Ruff et al. (2011) did not account for the selection associated with the lens detection phase, which, in general, varies with source redshift. Both of these problems can be overcome with the formalism introduced in Sonnenfeld (2022, Paper III from here on).

The purpose of this paper is to show how to carry out a population study of a sample of strong lenses with no individual

source redshift measurements, and to quantify the loss in precision compared to the case in which spectroscopic redshifts are available. The method that I propose relies on a good knowledge of the properties of the unlensed background galaxy population, as well as the selection function of the strong lensing survey. As with Paper III, this is more easily done when the sample is complete above a well-defined cut in observational space. I tested the method on a simulated population of foreground galaxies and background sources, from which a complete subset of strong lenses was drawn.

The structure of this work is the following. In Sect. 2 I introduce the theoretical formalism on which this work is based. In Sect. 3 I describe the simulations used to test the analysis method. In Sect. 4 I describe the model that I used to fit the simulated data. In Sect. 5 I show the results of the experiment. I discuss the results and draw conclusions in Sect. 6.

Throughout the paper I assume a flat Λ cold dark matter cosmology with $\Omega_M = 0.3$ and $H_0 = 70 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$. Masses and luminosities are in solar units. The Python code used for the simulation and analysis of the lens sample can be found on GitHub[3].

2. Photogeometric redshifts

This section explains (1) how information from strong lensing data, together with prior knowledge of the background source population, can be used to constrain the redshift of a lensed source, and (2) how the properties of a population of galaxies can be inferred by statistically combining data from a subset of lenses, even when their source redshifts are not known. The following derivation builds on the formalism developed in Paper III.

The population of strong lenses from a survey with selection criterion S can be described with the following probability distribution:

$$P_{SL}(\boldsymbol{\psi}_g, \boldsymbol{\psi}_s | S) \propto P_g(\boldsymbol{\psi}_g) P_s(\boldsymbol{\psi}_s) P_{det}(\boldsymbol{\psi}_g, \boldsymbol{\psi}_s | S), \tag{1}$$

where $\boldsymbol{\psi}_g$ and $\boldsymbol{\psi}_s$ are parameters describing the foreground galaxy and background source, P_g and P_s are the galaxy and source population distribution before lensing, and P_{det} is the probability of a lens-source pair being detected as a strong lens by the survey. The source redshift, z_s, is one of the parameters describing the source. The source parameter array, $\boldsymbol{\psi}_s$, can then be written as

$$\boldsymbol{\psi}_s = (\boldsymbol{\psi}_s^{(-z_s)}, z_s), \tag{2}$$

where $\boldsymbol{\psi}_s^{(-z_s)}$ includes all of the source parameters (for instance, the magnitude, the position, etc.) except the redshift.

Given a strong lens with observed data $\mathbf{d}$, the marginal posterior probability of the source redshift is

$$P(z_s | \mathbf{d}) \propto \int d\boldsymbol{\psi}_g d\boldsymbol{\psi}_s^{(-z_s)} P_{SL}(\boldsymbol{\psi}_g, \boldsymbol{\psi}_s^{(-z_s)}, z_s | S) P(\mathbf{d} | \boldsymbol{\psi}_g, \boldsymbol{\psi}_s^{(-z_s)}, z_s). \tag{3}$$

If the strong lens population distribution is known exactly from prior information, then one can use that knowledge to constrain $P(z_s | \mathbf{d})$ via the above equation. This is the basic principle of photogeometric redshift.

As the term photogeometric suggests, the constraints on the source redshift consist of a photometric part, which is given by the knowledge, informed by the lens model, of the unlensed surface brightness distribution of the source, and a geometric part,

---

[3] `https://github.com/astrosonnen/strong_lensing_tools`

which is directly related to the lens configuration. To clarify this concept, it is useful to consider the following example.

We take a strong lens consisting of a foreground galaxy, described by parameters ψ_g, and a background point source. The data, $\mathbf{d}$, include the positions and apparent magnitudes of the multiple images of the source. These data constrain the source redshift via the factor $P(\mathbf{d}|\psi_g, \psi_s^{(-z_s)}, z_s)$ in Eq. 3, which is the geometric term: modifying the source redshift varies the predicted image positions and magnitudes as a result of the dependence of the lensing deflection angle on the observer-lens-source geometry.

The photometric constraint on z_s is instead given by the factor $P_{SL}(\psi_g, \psi_s^{(-z_s)}, z_s|S)$ in Eq. 3. This can be written as

$$P_{SL}(\psi_g, \psi_s^{(-z_s)}, z_s|S) = P(\psi_g, \psi_s^{(-z_s)}|S)P(z_s|\psi_g, \psi_s^{(-z_s)}, S), \qquad (4)$$

where I used the rule of conditional probability. The first factor in the right-hand side of the above equation is the prior probability of the lens and source parameters (except the source redshift) being ψ_g and $\psi_s^{(-z_s)}$ given the fact that the lens is included in the survey with selection criterion S. The second factor, $P(z_s|\psi_g, \psi_s^{(-z_s)}, S)$, is the prior probability of the source redshift given the remaining model parameters and the selection criterion. One of the parameters entering $\psi_s^{(-z_s)}$ is the intrinsic (unlensed) source apparent magnitude, which can be used to further constrain the source redshift given the knowledge of the distribution of background sources in redshift-magnitude space. For example, if the source is intrinsically very bright, large values of the redshift will be disfavoured.

Source redshift information is rarely needed for its own sake: in most cases, z_s is required to constrain lens models. In other words, the factor P_g, which enters Eq. 3 via the expansion of Eq. 1, is uncertain, and the source redshift is one of the ingredients needed to constrain it. Ruff et al. (2011) solved the problem of determining the structural parameters of lenses with no source redshifts as follows. They used information on P_g that was independent of their strong lens survey, the SL2S, together with the assumption of an isothermal density profile for the galaxies, to infer the z_s of sources with no spectroscopic redshifts. Then, they combined these estimates of z_s with the observed Einstein radii and stellar velocity dispersion to infer the evolution of the density profile of the lenses, effectively obtaining an updated estimate of P_g. Their procedure ensured that no data were used more than once, but it is internally inconsistent because their inferred density profiles of the lenses were different from isothermal.

The correct approach, when the main goal is to constrain the galaxy population properties, is to infer P_g while marginalising over the source redshifts. If we summarise the galaxy distribution with a set of parameters, η, then we can obtain the posterior probability distribution of η given the data as

$$P(\eta|\mathbf{d}) \propto P(\eta) \prod_i^{N_{lens}} \int d\psi_{g,i} d\psi_{s,i}^{(-z_s)} dz_{s,i} P(\mathbf{d}_i|\psi_g, \psi_{s,i}^{(-z_s)}, z_s) \times$$
$$\times P_{SL}(\psi_{g,i}, \psi_{s,i}^{(-z_s)}, z_{s,i}|\eta, S), \quad (5)$$

where the product is taken over a sample of N_{lens} strong lenses. This is the approach adopted in this paper.

The integrands in the right-hand side of Eq. 5 contain the expression for the photogeometric redshift given by Eq. 3. In this case, however, the probability distribution of the lens parameters, P_{SL}, is not fixed but is a function of the galaxy population parameters, η. If a new lens is observed from a survey with the

same characteristics, then its photogeometric redshift can be estimated by marginalising over all possible values of η given our knowledge gathered from the analysis of the first N_{lens} lenses:

$$P(z_{s,N_{lens}+1}|\mathbf{d}_{N_{lens}+1}) \propto \int d\eta P(\eta|\mathbf{d}) \int d\psi_{g,N_{lens}+1} d\psi_{s,N_{lens}+1}^{(-z_s)}$$
$$P(\mathbf{d}_{N_{lens}+1}|\psi_{g,N_{lens}+1}, \psi_{s,N_{lens}+1}^{(-z_s)}) P_{SL}(\psi_{g,N_{lens}+1}, \psi_{s,N_{lens}+1}^{(-z_s)}|\eta, S). \qquad (6)$$

3. Simulations

In order to demonstrate how the formalism introduced in the previous section can be applied to a practical case, I generated a simulated sample of strong lenses with the following general characteristics. As in the previous papers of this series, I assumed foreground galaxies to have an axisymmetric mass distribution and background sources to be point-like. Each galaxy consists of the sum of two concentric components: a dark matter halo and a stellar bulge. I defined the galaxy population among which strong lenses are searched for by applying a luminosity cut. The galaxy sample is assumed to be complete above this cut. Lens-source systems that produce at least two images with observed magnitude brighter than a given threshold are listed as strong lenses detected by the survey. The survey is assumed to be complete above this limit. In the next few sections, I provide the details of how this sample was created.

3.1. Foreground galaxy population

I generated 10^4 galaxies, with luminosity, L, drawn from the following probability distribution

$$P(L) \sim \left(\frac{L}{L^*}\right)^{\alpha_L} \exp\left\{-\frac{L}{L^*}\right\}, \qquad (7)$$

with $\log L^* = 10.8$, $\alpha_L = -1.0$, and $\log L > 10.5$. Given the luminosity of each galaxy, I obtained its stellar mass by drawing the base-10 logarithm of the stellar mass-to-light ratio, Υ_*, from the following distribution:

$$\log \Upsilon_* \sim \mathcal{N}(0.5, 0.1^2), \qquad (8)$$

where the notation $\mathcal{N}(\mu, \sigma^2)$ indicates a Gaussian distribution with mean μ and variance σ^2. The combination of the luminosity function, Eq. 7, and the mass-to-light ratio distribution of Eq. 8 produces a stellar mass function that is approximately that of quiescent galaxies at $0.2 < z < 0.5$, as measured by Muzzin et al. (2013) under the assumption of a Chabrier (2003) stellar initial mass function but shifted towards higher masses by ~ 0.1 dex.

I described the mass distribution of the stellar component as an axisymmetric de Vaucouleurs profile, with half-light radius drawn from the following stellar mass-size relation:

$$\log R_e \sim \mathcal{N}(1.0 + 0.8(\log M_* - 11.5), 0.15^2). \qquad (9)$$

Subsequently, I assigned a dark matter halo mass, M_h, to each galaxy. I defined M_h as the mass enclosed within a sphere with average density equal to 200 times the critical density of the Universe. I drew the value of $\log M_h$ from the following Gaussian distribution:

$$\log M_h \sim \mathcal{N}(13.0 + 1.5(\log M_* - 11.5), 0.2^2). \qquad (10)$$

The mean of this Gaussian scales with a power 1.5 of the stellar mass, M_*. I assumed a spherical Navarro, Frenk, and White

(NFW; Navarro et al. 1997) profile for the dark matter distribution, with a fixed ratio between the virial radius and the scale radius, corresponding to a concentration of 5:

$$\frac{r_{200}}{r_s} = 5, \tag{11}$$

where r_{200} is the radius of the sphere enclosing a dark matter mass equal to M_h. For simplicity, the foreground galaxies are assumed to be all at the same redshift, $z_g = 0.4$. The distribution in $\log L$, $\log M_h$, $\log \Upsilon$, and $\log R_e$ of the sample of foreground galaxies is shown as solid black contours in Fig. 1.

3.2. Background source population

I described each background source by means of its $i-$band apparent magnitude, m_s, and its redshift, z_s. I first assumed that their UV absolute magnitude, M, is drawn from the following luminosity function:

$$N(M) = 0.4 \ln 10 \phi^* \left(10^{-0.4(M-M^*)}\right)^{\alpha_s+1} \exp\left\{-10^{-0.4(M-M^*)}\right\}, \tag{12}$$

with $\phi^* = 7.02 \times 10^{-3}$ Mpc^{-3} mag^{-1}, $M^* = -19.68$, and $\alpha_s = -1.32$. These values correspond to the rest-frame 1500Å luminosity function of $z \approx 2$ galaxies, as measured by Parsa et al. (2016).

Then, I assumed that all of the sources have the same spectral energy distribution, given by the spectral template 'starb1' from the Kinney-Calzetti Spectral Atlas of Galaxies (Calzetti et al. 1994; Kinney et al. 1996). For ease of computation, I only considered galaxies with $m_s < 27$ and $z_g+0.1 < z_s < 4$. The resulting distribution in z_s and m_s is shown as solid black lines in Fig. 1. Once integrated over z_s, the background source population has an average number density of $n_{bkg} = 162$ arcmin^{-2}.

3.3. Strong lens definition

Given each foreground galaxy, I randomly placed background sources behind it, assuming a uniform spatial distribution with average projected number density n_{bkg}. Then, I solved the lens equation to determine their observed image positions. For simplicity, I assumed the background sources to be massless so that the only source of lensing deflection was that of the foreground galaxy. Axisymmetric lenses of the kind considered in this experiment can produce up to three images of the same source. I focused on the two brightest images, at positions θ_1 and θ_2. With the label 1 I indicate the main image, that is, the one outside of the tangential critical curve, which is usually the brightest. Image 2, instead, is located between the radial and the tangential critical curve, opposite to image 1 with respect to the lens centre (which is the same notation used in all of the papers of this series).

Given the intrinsic apparent magnitude, m_s, of a strongly lensed source and the lensing magnification at the two image positions, I computed the magnitudes of the two lensed images, m_1 and m_2. Then, I added a Gaussian observational noise with dispersion $\epsilon_m = 0.1$ mag to obtain their corresponding observed values:

$$m_1^{obs} \sim \mathcal{N}(m_1, \epsilon_m^2) \tag{13}$$
$$m_2^{obs} \sim \mathcal{N}(m_2, \epsilon_m^2). \tag{14}$$

I defined as a strong lens any galaxy with a strongly lensed source for which the observed magnitude of image 2 is brighter than 26. The lens selection criterion is then

$$S = \left\{m_2^{obs} < 26\right\}. \tag{15}$$

This condition mimics a selection based on the detection of a counter-image to the main arc, which usually allows the lensing nature of a lens candidate to be determined with high confidence. I assumed that all such lenses are included in the survey, which is therefore complete above the observational cut given by Eq. 15. All of the other galaxies are treated as non-detections.

With this definition, the simulation produced 137 strong lenses. Two of the lenses have more than one strongly lensed source: in those cases, I only considered one lensed source each, chosen at random. The distribution in foreground galaxy properties, background source properties, and Einstein radius of the strong lenses is shown in Fig. 1. Compared to the general population, strong lens galaxies tend to be more luminous and, consequently, to have higher halo masses. Additionally, they are more compact at fixed luminosity (that is, their half-light radius is on average smaller) and have a higher stellar mass-to-light ratio. These last two trends are quantified in Table 1.

The distribution of strongly lensed sources differs from the general population of background galaxies in two ways. First, the low-redshift part of the distribution is suppressed. This is because the strong lensing cross-section decreases with source redshift at fixed lens redshift. Second, the abundance of faint galaxies ($m_s > 26$) is greatly reduced. This is a consequence of the detection limit of Eq. 15.

4. Model

I fitted the simulated strong lensing data with a model that describes the joint distribution of the galaxy population, the background source population, and the strong lens detection efficiency of the survey, using the same method introduced in Paper III. Each aspect of the model is described in the following sections.

4.1. Foreground galaxies

I described each galaxy with the sum of a stellar component, consisting of a de Vaucouleurs profile with known luminosity and half-light radius and unknown mass-to-light ratio, and a dark matter halo, consisting of an NFW profile. Moreover, I assumed that the concentration parameter of the NFW profile, Eq. 11, is known exactly from prior information. This is not a realistic assumption, but it is a useful one: it simplifies calculations while still allowing the main goal of this paper to be pursued, which is to demonstrate the accuracy of the inference method itself.

Each galaxy is therefore described by four parameters: the luminosity, L, the stellar mass-to-light ratio, Υ_*, the half-light radius, R_e, and the halo mass, M_h. The galaxy parameter set is therefore
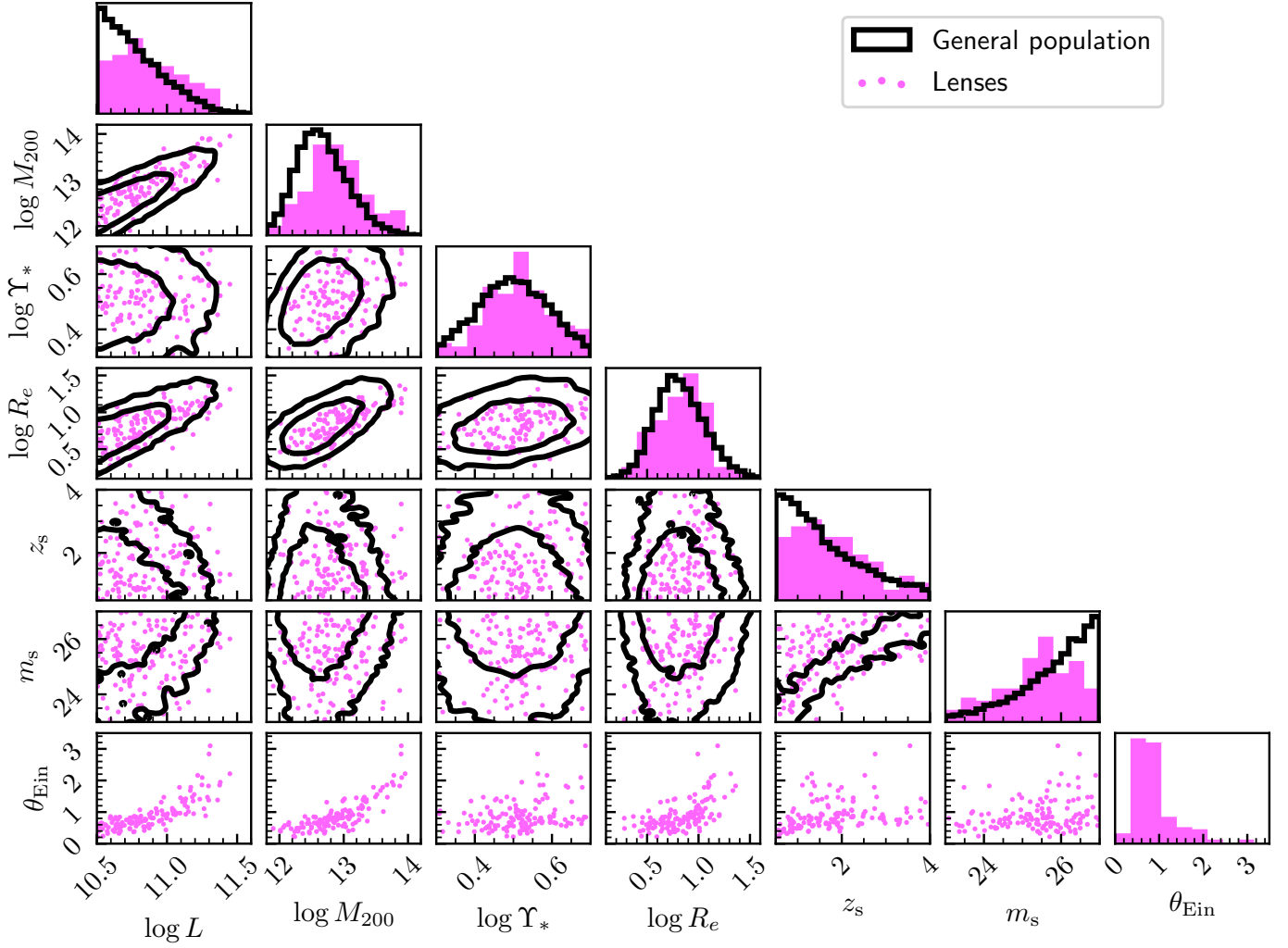
$$\psi_g \equiv \{L, \Upsilon_*, R_e, M_h\}. \tag{16}$$

I assumed the following probability distribution for ψ_g:

$$P_g(\psi_g|\eta) = \mathcal{L}(L)\mathcal{U}(\log \Upsilon_*)\mathcal{R}(\log R_e)\mathcal{H}(\log M_h). \tag{17}$$

The factor $\mathcal{L}$ is the luminosity distribution of the sample, which I fixed to the true distribution used to generate the mock. The factor $\mathcal{U}$ is the distribution in the stellar mass-to-light ratio, which I assumed to be a Gaussian in $\log \Upsilon_*$,

$$\mathcal{U}(\log \Upsilon_*) \sim \mathcal{N}(\mu_\Upsilon, \sigma_\Upsilon^2), \tag{18}$$

with mean μ_Υ and dispersion σ_Υ.

Fig. 1. Simulated sample. Solid lines show the distribution in the luminosity, halo mass, stellar mass-to-light ratio, and half-light radius of the general population of foreground galaxies as well as the distribution in redshift and $i-$band magnitude of the general population of background sources. Contour levels correspond to 68% and 95% enclosed probability. Pink points show the values of the above quantities, as well as of the Einstein radius, of the strong lenses. Histograms are normalised to one.

The factor $\mathcal{R}$ is the distribution in half-light radius, which I modelled as the following Gaussian in $\log R_e$:

$$\mathcal{R}(\log R_e) \sim \mathcal{R}(\mu_{R,0} + \beta_R(\log M_* - \log M_*^p), \sigma_R^2), \tag{19}$$

with

$$\log M_*^p = 11.0 + \mu_\Upsilon. \tag{20}$$

In words, the mean $\log R_e$ scales with a power β_R of the stellar mass, and the pivot point of the stellar mass-size relation, M_*^p, is the average stellar mass of a galaxy with luminosity $\log L = 11$. The parameter $\mu_{R,0}$, then, is the average $\log R_e$ of a galaxy with luminosity $\log L = 11$. This quantity can be measured with high precision on the sample of galaxies from which lenses are drawn; therefore, I assume that it is known exactly.

Finally, the factor $\mathcal{H}$ is the distribution in halo mass. Similarly to the $\mathcal{R}$ factor, this is modelled as

$$\mathcal{H}(\log M_h) \sim \mathcal{N}(\mu_{h,0} + \beta_h(\log M_* - \log M_*^p), \sigma_h^2). \tag{21}$$

The distribution $P_g(\psi_g|\eta)$ is a function of the following set of free parameters:

$$\eta \equiv \{\mu_\Upsilon, \sigma_\Upsilon, \beta_R, \sigma_R, \mu_{h,0}, \beta_h, \sigma_h\}. \tag{22}$$

The goal is to infer η given the data. The distribution $P_g(\psi_g|\eta)$ reduces to the true distribution used to generate the mock for

$$\eta^{(\text{true})} = \{0.5, 0.1, 0.8, 0.15, 13.0, 1.5, 0.2\}. \tag{23}$$

4.2. Background sources

Sources are point-like, and therefore they can be fully described with their apparent magnitude, m_s, their redshift, z_s, and their position relative to the lens centre, β:

$$\psi_s \equiv \{m_s, z_s, \beta\}. \tag{24}$$

I assumed that the distribution in ψ_s of the background galaxy population is known exactly from prior observations. I discuss possible strategies to drop this assumption in Sect. 6.

The probability distribution $P_s(\psi_s)$ is therefore fixed to the one used to generate the mock (shown in black in Fig. 1). This can be written as

$$P_s(\psi_s) = U_{\mathbb{R}^2}(\beta)P(z_s)P(m_s|z_s). \tag{25}$$

In the above equation, $U_{\mathbb{R}^2}$ is a uniform distribution on the sky, $P(m_s|z_s)$ is

$$P(m_s|z_s) \propto \left(10^{-0.4(m_s - m_s^*(z_s))}\right)^{\alpha_s + 1} \exp\left\{-10^{-0.4(m_s - m_s^*(z_s))}\right\}, \tag{26}$$

where $m_s^*(z_s)$ is obtained from M^* via the same spectral template used to generate the mock, and $P(z_s)$ is

$$P(z_s) \propto \frac{dV_C}{dz} \frac{dn_{bkg}}{dz}, \tag{27}$$

which is the product of the derivatives with respect to the redshift of the comoving volume and the surface density of background sources.

4.3. Lens detection efficiency

As Sect. 3.3 explains, the lens sample is complete above the cut of Eq. 15. This implies that the detection efficiency of a strong lens, expressed in terms of model quantities, is given by

$$P_{det}(\boldsymbol{\psi}_g, \boldsymbol{\psi}_s | S) = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{m_2(\boldsymbol{\psi}_g, \boldsymbol{\psi}_s) - 26}{\sqrt{2}\epsilon_m}\right)\right]. \tag{28}$$

4.4. Inferring the galaxy population parameters

I fitted the model described in this section to the data relative to the 137 strong lenses. For each lens, I assumed that the following data are available: (1) the positions of the two images of the lensed source, θ_1^{obs} and θ_2^{obs}, measured exactly; (2) the half-light radius of the galaxy, $R_e^{(obs)}$, measured exactly; (3) the luminosity of the lens galaxy, $L^{(obs)}$, measured exactly; and (4) the magnitudes of the two images, m_1^{obs} and m_2^{obs}, measured with uncertainty ϵ_m.

In principle, the remaining galaxies in the sample (the non-detections) could be used as additional constraints, following the method introduced in Paper III. However, Paper III also showed that, when relative magnification information is available (that is, when the magnitudes of the lensed images are used as constraints), the benefit of adding the non-detections is negligible.

Following Paper III, the posterior probability distribution of the galaxy population parameters, given the data of the strong lenses, is given by

$$P(\boldsymbol{\eta}|\mathbf{d}) \propto P(\boldsymbol{\eta}) \prod_i P(\mathbf{d}_i|\boldsymbol{\eta}), \tag{29}$$

where the product is taken over the strong lenses, and $\mathbf{d}_i$ is the data relative to the i−th lens. Each factor $P(\mathbf{d}_i|\boldsymbol{\eta})$ can be written as the following integral:

$$P(\mathbf{d}_i|\boldsymbol{\eta}) = \int d\boldsymbol{\psi}_g d\boldsymbol{\psi}_s P(\mathbf{d}_i|\boldsymbol{\psi}_g, \boldsymbol{\psi}_s) P_{SL}(\boldsymbol{\psi}_g, \boldsymbol{\psi}_s|\boldsymbol{\eta}, S), \tag{30}$$

where

$$P_{SL}(\boldsymbol{\psi}_g, \boldsymbol{\psi}_s|\boldsymbol{\eta}, S) = A(\boldsymbol{\eta}) P_g(\boldsymbol{\psi}_g|\boldsymbol{\eta}) P_s(\boldsymbol{\psi}_s) P_{det}(\boldsymbol{\psi}_g, \boldsymbol{\psi}_s|S), \tag{31}$$

and $A(\boldsymbol{\eta})$ is a constant that ensures that P_{SL} is normalised to unity. The likelihood of observing the data given the lens and source parameters, $P(\mathbf{d}_i|\boldsymbol{\psi}_g, \boldsymbol{\psi}_s)$, is

$$P(\mathbf{d}_i|\boldsymbol{\psi}_g, \boldsymbol{\psi}_s) = \delta\left(R_e - R_e^{(obs)}\right) \delta\left(L - L^{(obs)}\right) \delta\left(\theta_1(\boldsymbol{\psi}_g, \beta) - \theta_1^{obs}\right)$$
$$\delta\left(\theta_2(\boldsymbol{\psi}_g, \beta) - \theta_2^{obs}\right) P(m_1^{obs}, m_2^{obs}|\boldsymbol{\psi}_g, \boldsymbol{\psi}_s), \tag{32}$$

where $P(m_1^{obs}, m_2^{obs}|\boldsymbol{\psi}_g, \boldsymbol{\psi}_s)$ is the product of two Gaussians, with means m_1^{obs} and m_2^{obs} and dispersion ϵ_m.

Equation 30 is a seven-dimensional integral. The integrals over the galaxy half-light radius and luminosity are trivial, in virtue of the Dirac delta functions of Eq. 32. Similarly, the fact

that the two image positions are known exactly allows the integration to be performed over two lens parameters analytically. After integrating over the source position β and $\log \Upsilon_*$, Eq. 32 becomes

$$P(\mathbf{d}|\boldsymbol{\eta}) = \int dz_s dm_s d\log M_h P(m_1^{obs}, m_2^{obs}|M_h, L\Upsilon_*^*, R_e, \beta^*)$$
$$\times A(\boldsymbol{\eta}) P(L, R_e, M_h, \Upsilon_*^*|\boldsymbol{\eta}) P_s(z_s, m_s, \beta^*)$$
$$\times P_{det}(\boldsymbol{\psi}_g, z_s, m_s, \beta^*|S) |\det J|_{(\Upsilon_*, \beta)=(\Upsilon_*^*, \beta^*)}, \tag{33}$$

where I dropped the subscript i and set $L = L^{(obs)}$ and $R_e = R_e^{(obs)}$ for ease of notation. In the above equation, Υ_*^* and β^* are the values of the stellar mass-to-light ratio and source position needed to reproduce the two image positions, while the last factor in the integrand is the Jacobian determinant of the following variable change:

$$(\log \Upsilon_*, \beta) \rightarrow (\theta_1, \theta_2). \tag{34}$$

I sampled the posterior probability distribution of $\boldsymbol{\eta}$ via a Markov chain Monte Carlo. I assumed a uniform prior on all of the parameters, over the range listed in Table 1. At each draw of $\boldsymbol{\eta}$, I computed both the normalisation constant, $A(\boldsymbol{\eta})$, and the integrals of Eq. 33 via Monte Carlo integration with importance sampling.

4.5. The redshift-lens mass degeneracy

Before showing the results of the inference over the entire lens population, it is useful to consider the problem of the degeneracy between the source redshift and the lens mass parameters in the case of a single lens. The lens model used for the fit has five degrees of freedom: the stellar and halo mass of the lens galaxy and the magnitude, redshift, and position of the background source. If the magnitudes of the two images and the redshift of the source were known exactly (in addition to the image positions), then the model would be fully constrained, and the result of the fit would be equal to the truth. In the case of unknown source redshift, however, the uncertainty on z_s propagates to the other lens model parameters.

In order to understand the amplitude of this effect, I fitted a lens from the simulated sample, assuming perfect measurements, while assuming a range of values of the source redshift. Figure 2 shows the inferred values of Υ_*, M_h as a function of z_s. Both the inferred stellar mass-to-light ratio and the halo mass increase with decreasing source redshift: this is because, as the source gets closer to the lens, the value of the surface critical density for strong lensing becomes larger, and therefore a higher mass is required to produce the deflection angles needed to fit the image positions. More precisely, when varying the source redshift from $z_s = 1$ to $z_s = 3$, the inferred stellar mass-to-light ratio changes from $\log \Upsilon_* = 0.55$ to $\log \Upsilon_* = 0.42$. While this is a relatively large uncertainty, the combination of a large set of lenses allows this source of noise to be be beaten down statistically, as the next section shows.

5. Results

5.1. Galaxy population inference

Figure 3 shows the posterior probability distribution of the parameters that describe the halo mass and stellar mass-to-light ratio distribution of the galaxy population. The median and 68%
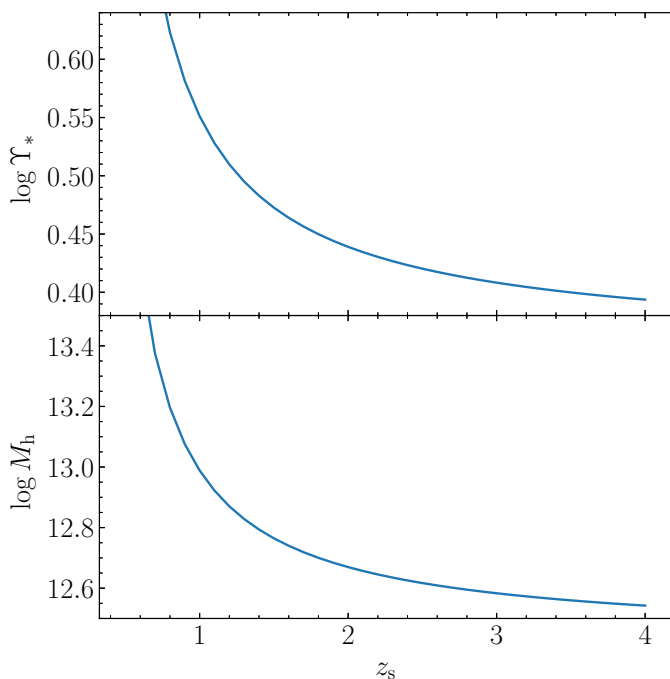
Fig. 2. Stellar mass-to-light ratio (upper panel) and halo mass (lower panel) obtained by fitting the lens model introduced in Sect. 3.1 to a mock strong lens from the simulated sample as a function of the assumed source redshift. The fit was done assuming perfect measurements of the two image positions and magnitudes.

credible region of the marginal posterior of all of the model parameters (including those describing the size distribution) is reported in Table 1. The true values of all of the parameters are accurately recovered. The inference on the average $\log \Upsilon_*$ parameter, in particular, is $\mu_\Upsilon = 0.498 \pm 0.018$. Moreover, as in Paper III, the model is able to correct for the lensing selection bias: it favours values of the parameters closer to those of the general galaxy population and disfavours values measured on the lenses alone (these values are shown in Fig. 3 and reported in Table 1), though with low significance.

In order to compare this result with a case in which source redshifts are available, I repeated the analysis by fixing the z_s of each lens to the true value. The resulting posterior probability distribution is also shown in Fig. 3 and reported in Table 1. The uncertainties obtained when z_s is known are comparable to the case with no individual z_s information. Only the inference on the parameters that describe the stellar mass-to-light ratio is slightly more precise. This means that the primary factor that determines the uncertainty in this experiment is the limited amount of information that the lensing data provide, and not the lack of source redshift information. The uncertainties associated with the individual source redshifts are effectively reduced by the statistical combination of the lenses in the sample.

This result depends on the relative importance of the source redshift uncertainty in the error budget. In turn, this can vary with the quality of the lensing data, the flexibility of the model used to fit the data, the intrinsic scatter in the lens and source population, and the sample size. Exploring the sensitivity of the results to these various aspects is beyond the purpose of this paper. We can consider, however, the case in which the two-component galaxy model of Sect. 3.1 is replaced with a singular isothermal sphere (SIS). An SIS is fully described by a single parameter that, in the case in which the source redshift is known, can be constrained exactly with the two image positions of the lensed

source. In other words, uncertainties associated with the individual lens models vanish when assuming an SIS. As a result, the uncertainty associated with the source redshifts acquires a bigger weight in the error budget, and therefore the difference in precision between the known and unknown individual redshift cases becomes larger. Nevertheless, massive galaxies are known to have a more complex mass structure than a simple SIS (see e.g. Sonnenfeld et al. 2013): in all practical cases, when analysing real samples of lenses, models with at least two degrees of freedom are to be preferred.

5.2. Photogeometric redshifts on new lenses

Once the properties of the population of galaxies have been determined, it is possible to use this knowledge to predict the source redshifts of individual strong lenses. We can consider a new lens, discovered by a survey with identical properties to the one on which the mock sample was based. The marginal posterior probability of the source redshift of this lens is given by Eq. 6. Figure 4 shows this posterior probability for an example lens, fitted with the same model of Sect. 4 and using the posterior probability of the galaxy population parameters, $P(\eta|\mathbf{d})$, obtained earlier.

In order to illustrate the way in which the knowledge of the galaxy population informs this quantity, Fig. 4 also shows the marginal posterior probability on z_s obtained (1) assuming flat priors on the lens parameters $\log M_h$ and $\log \Upsilon_*$ and (2) neglecting the lens detection efficiency term, P_{det}. The two distributions are very different: the information on the galaxy population properties and on the selection function correctly disfavours solutions with source redshifts very close to the lens. On the contrary, when this information is not taken into account, the posterior probability in the source redshift stays similar to the prior distribution of the background source population, which is highest at the lowest redshifts.

The width of the $P(z_s|\mathbf{d}_{N_{lens}+1})$ distribution is the result of (1) uncertainties associated with the lens model, (2) uncertainties on the galaxy population distribution parameters, η, and (3) the width of the $P_{SL}(\psi_g, \psi_s|\eta)$ distribution. In order to isolate these different sources of uncertainty, I repeated the analysis assuming perfect lensing measurements (that is, assuming no errors on the magnitudes of the two images) and fixing η to the true value. The resulting posterior probability distribution, also shown in Fig. 4, is very similar to the one obtained without these additional assumptions. This means that the main sources of uncertainty on the photogeometric redshift are the width of the galaxy and source population distributions.

6. Discussion and summary

The analysis described in the previous sections relied on the assumption of a perfect knowledge of the distribution of the background source population, $P_s(\psi_s)$. As discussed in Paper III, it can be challenging to acquire this information in practice, especially if the selection criteria of a survey involve more aspects than just the magnitudes and positions of the lensed images. For example, if the detection probability is also a function of surface brightness, then the model ψ_s must also include parameters such as galaxy size.

In principle, $P_s(\psi_s)$ can be determined by studying the population of unlensed galaxies. However, because of lensing magnification, the appearance of strongly lensed sources, and therefore their detectability, can vary with changes in ψ_s that occur below
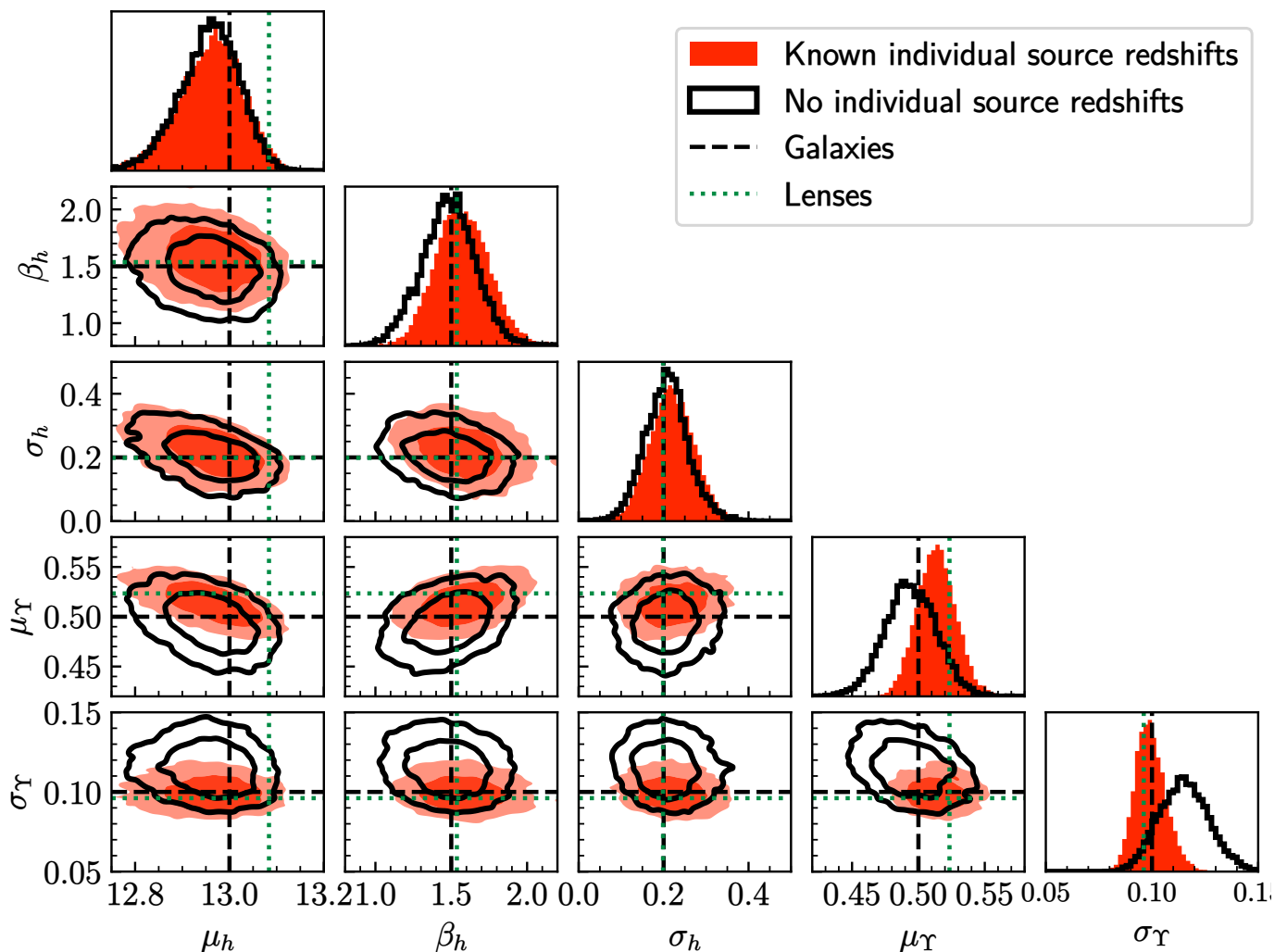
Fig. 3. Posterior probability distribution of the parameters describing the distribution in M_h and Υ_* of the galaxy population, obtained by fitting the simulated sample of 137 strong lenses with the model of Sect. 4. Solid black contours show the inference obtained with no individual source redshift information. Red filled contours show the inference obtained with known individual source redshifts. Dashed black lines mark the true values of the parameters. Dotted green lines indicate the values of the parameters obtained by fitting the model to the subsample of strong lenses.

Table 1. Inference on the galaxy population model parameters.

Parameter	Truth	Lenses	Prior	No individual redshifts	Known individual redshifts	Description
$\mu_{h,0}$	13.00	13.08	$U(12.5, 13.5)$	12.96 ± 0.06	12.96 ± 0.07	Mean $\log M_h$ at luminosity $\log L = 11$ and average Υ_*
β_h	1.50	1.54	$U(0.0, 3.0)$	1.48 ± 0.17	1.57 ± 0.18	Scaling of the mean $\log M_h$ with $\log M_*$
σ_h	0.20	0.20	$U(0.0, 0.5)$	0.21 ± 0.05	0.22 ± 0.05	Scatter in $\log M_h$ around the mean
μ_Υ	0.50	0.52	$U(0.3, 0.7)$	0.494 ± 0.019	0.514 ± 0.015	Mean $\log \Upsilon_*$
σ_Υ	0.10	0.10	$U(0.0, 0.3)$	0.115 ± 0.012	0.099 ± 0.007	Scatter in $\log \Upsilon_*$ around the mean
β_R	0.80	0.70	$U(0.0, 2.0)$	0.76 ± 0.05	0.79 ± 0.05	Scaling of $\log R_e$ with $\log M_*$
σ_R	0.15	0.14	$U(0.0, 0.5)$	0.164 ± 0.011	0.161 ± 0.010	Scatter in $\log R_e$ around the mean

Notes. Column (2): True values of the population parameters. Column (3): Values of the lens population parameters, obtained by fitting the model of Eq. 17 to the sample of detected strong lenses. Column (4): Priors on the parameters. Column (5): Median and 68% credible interval of the marginal posterior probability distribution of each parameter. Column (6): Inference with individual source redshift information.

the resolution limit of the observations with which the unlensed galaxy population is studied. A possible solution, then, is to determine $P_s(\boldsymbol{\psi}_s)$ along with the galaxy population, adding flexibility to the model. This will introduce an additional source of uncertainty. In that case, the redshift dependence of $P_s(\boldsymbol{\psi}_s)$ could

be constrained by obtaining spectroscopic redshifts for a small subset of the background sources in the strong lens sample. I leave the exploration of the case in which the source population distribution term is no longer fixed for future study.
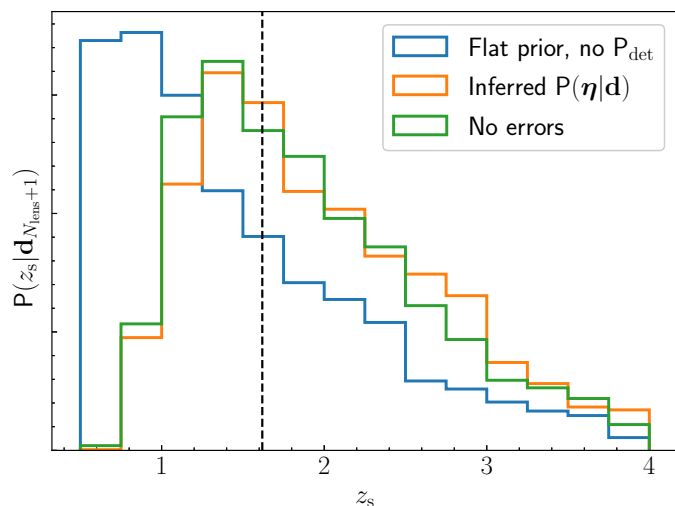
Fig. 4. Photogeometric redshift of a strongly lensed source, assuming a lens model from the family of Sect. 3.1. The blue histogram was obtained by fitting the lensing data assuming a flat prior on $\log M_h$ and $\log \Upsilon_*$ and the source population distribution of Sect. 3.2, with no correction for the strong lensing selection. The orange histogram was obtained assuming the full model, including the selection function correction, and using the values of the galaxy population parameters, $\boldsymbol{\eta}$, inferred in the experiment as a prior. The green histogram was obtained neglecting observational errors and fixing the values of $\boldsymbol{\eta}$ to the truth. The dashed vertical line marks the true value of the source redshift.

This paper revisited the concept of photogeometric redshift, addressing the theoretical shortcomings present in the original formulation of Ruff et al. (2011). The experiment carried out demonstrates that it is possible to use strong lenses with no individual source redshift information to make accurate inferences of the properties of the galaxy population. When dealing with a large sample of lenses ($N_{\text{lens}} > 100$), the statistical noise from the uncertainty on the individual source redshifts is greatly reduced. In the case in which the properties of the background (unlensed) source population are known exactly, its contribution to the error budget on the galaxy population parameters is only secondary.

The key elements that are necessary for a successful application of this method are (1) the knowledge of the lens selection function and (2) the knowledge of the distribution in surface brightness profile-redshift space of the background galaxy population. The problem of determining the former can be greatly simplified by identifying a subset of the strong lens population that is complete above an observational cut, as shown in Paper III. Measuring the latter is potentially made difficult by finite resolution effects, the importance of which depends on the specific method used to discover lenses. Nevertheless, the distribution of the source properties can be inferred along with the properties of the foreground galaxy population by extending the model. In summary, the method introduced in this paper provides a viable solution to the problem of analysing large samples of lenses when spectroscopic follow-up is not possible.

References

Bolton, A. S., Burles, S., Koopmans, L. V. E., et al. 2008, ApJ, 682, 964
Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582
Chabrier, G. 2003, PASP, 115, 763
Collett, T. E. 2015, ApJ, 811, 20
Gavazzi, R., Treu, T., Marshall, P. J., Brault, F., & Ruff, A. 2012, ApJ, 761, 170
Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, MNRAS, 465, 1454
Kinney, A. L., Calzetti, D., Bohlin, R. C., et al. 1996, ApJ, 467, 38
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Parsa, S., Dunlop, J. S., McLure, R. J., & Mortlock, A. 2016, MNRAS, 456, 3194
Ruff, A. J., Gavazzi, R., Marshall, P. J., et al. 2011, ApJ, 727, 96
Sonnenfeld, A. 2022, A&A, 659, A132, (Paper III)
Sonnenfeld, A. & Cautun, M. 2021, A&A, 651, A18, (Paper I)
Sonnenfeld, A., Jaelani, A. T., Chan, J., et al. 2019, A&A, 630, A71
Sonnenfeld, A., Treu, T., Gavazzi, R., et al. 2012, ApJ, 752, 163
Sonnenfeld, A., Treu, T., Gavazzi, R., et al. 2013, ApJ, 777, 98
Sonnenfeld, A., Treu, T., Marshall, P. J., et al. 2015, ApJ, 800, 94
Vakili, M., Bilicki, M., Hoekstra, H., et al. 2019, MNRAS, 487, 3715