Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **490,** 1743–1773 (2019)
Advance Access publication 2019 September 12

doi:10.1093/mnras/stz2547

# A SHARP view of H0LiCOW: $H_0$ from three time-delay gravitational lens systems with adaptive optics imaging

Geoff C.-F. Chen [1]★ Christopher D. Fassnacht,[1] Sherry H. Suyu,[2,3,4] Cristian E. Rusu [5,6]† James H. H. Chan [7] Kenneth C. Wong,[5,8] Matthew W. Auger,[9,10] Stefan Hilbert,[11,12] Vivien Bonvin,[7] Simon Birrer,[13] Martin Millon,[7] Léon V. E. Koopmans,[14] David J. Lagattuta [15] John P. McKean,[14,16] Simona Vegetti,[2] Frederic Courbin,[7] Xuheng Ding,[13] Aleksi Halkola,[17] Inh Jee,[2] Anowar J. Shajib [13] Dominique Sluse,[18] Alessandro Sonnenfeld [8,19] and Tommaso Treu[13]

*Affiliations are listed at the end of the paper*

Accepted 2019 September 5. Received 2019 August 29; in original form 2019 July 9

## ABSTRACT

We present the measurement of the Hubble constant, $H_0$, with three strong gravitational lens systems. We describe a blind analysis of both PG 1115+080 and HE 0435−1223 as well as an extension of our previous analysis of RXJ 1131−1231. For each lens, we combine new adaptive optics (AO) imaging from the Keck Telescope, obtained as part of the SHARP (Strong-lensing High Angular Resolution Programme) AO effort, with *Hubble Space Telescope* (*HST*) imaging, velocity dispersion measurements, and a description of the line-of-sight mass distribution to build an accurate and precise lens mass model. This mass model is then combined with the COSMOGRAIL-measured time delays in these systems to determine $H_0$. We do both an AO-only and an AO + *HST* analysis of the systems and find that AO and *HST* results are consistent. After unblinding, the AO-only analysis gives $H_0 = 82.8^{+9.4}_{-8.3}$ km s$^{-1}$ Mpc$^{-1}$ for PG 1115+080, $H_0 = 70.1^{+5.3}_{-4.5}$ km s$^{-1}$ Mpc$^{-1}$ for HE 0435−1223, and $H_0 = 77.0^{+4.0}_{-4.6}$ km s$^{-1}$ Mpc$^{-1}$ for RXJ 1131−1231. The joint AO-only result for the three lenses is $H_0 = 75.6^{+3.2}_{-3.3}$ km s$^{-1}$ Mpc$^{-1}$. The joint result of the AO + *HST* analysis for the three lenses is $H_0 = 76.8^{+2.6}_{-2.6}$ km s$^{-1}$ Mpc$^{-1}$. All of these results assume a flat Λ cold dark matter cosmology with a uniform prior on $\Omega_m$ in [0.05, 0.5] and $H_0$ in [0, 150] km s$^{-1}$ Mpc$^{-1}$. This work is a collaboration of the SHARP and H0LiCOW teams, and shows that AO data can be used as the high-resolution imaging component in lens-based measurements of $H_0$. The full time-delay cosmography results from a total of six strongly lensed systems are presented in a companion paper.

**Key words:** gravitational lensing: strong – instrumentation: adaptive optics – distance scale.

## 1 INTRODUCTION

### 1.1 Distance measurement discrepancy: a 4.4$\sigma$ tension on the value of $H_0$

The temperature anisotropies of the cosmic microwave background (CMB) and density correlations of baryon acoustic oscillations (BAOs) obtained with the Wilkinson Microwave Anisotropy Probe, *Planck* satellite, and BAO surveys provide strong support to the standard flat ΛCDM cosmological model (e.g. Komatsu et al. 2011; Hinshaw et al. 2013; Planck Collaboration VI 2018). Under a few strong assumptions, such as flatness and constant dark energy density, these data give sub-per cent precision on the parameters of the standard cosmological model (e.g. Anderson et al. 2014; Kazin et al. 2014; Ross et al. 2015).

Intriguingly, distance measurements from Type Ia supernova (SN) that have been calibrated by the local distance ladder are smaller than the predictions from the CMB data given the flat ΛCDM model (see the illustration in fig. 4 in Cuesta et al. 2015), leading to a ∼4.4$\sigma$ tension in $H_0$ between the value predicted by

★ E-mail: chfchen@ucdavis.edu
† Subaru Fellow.

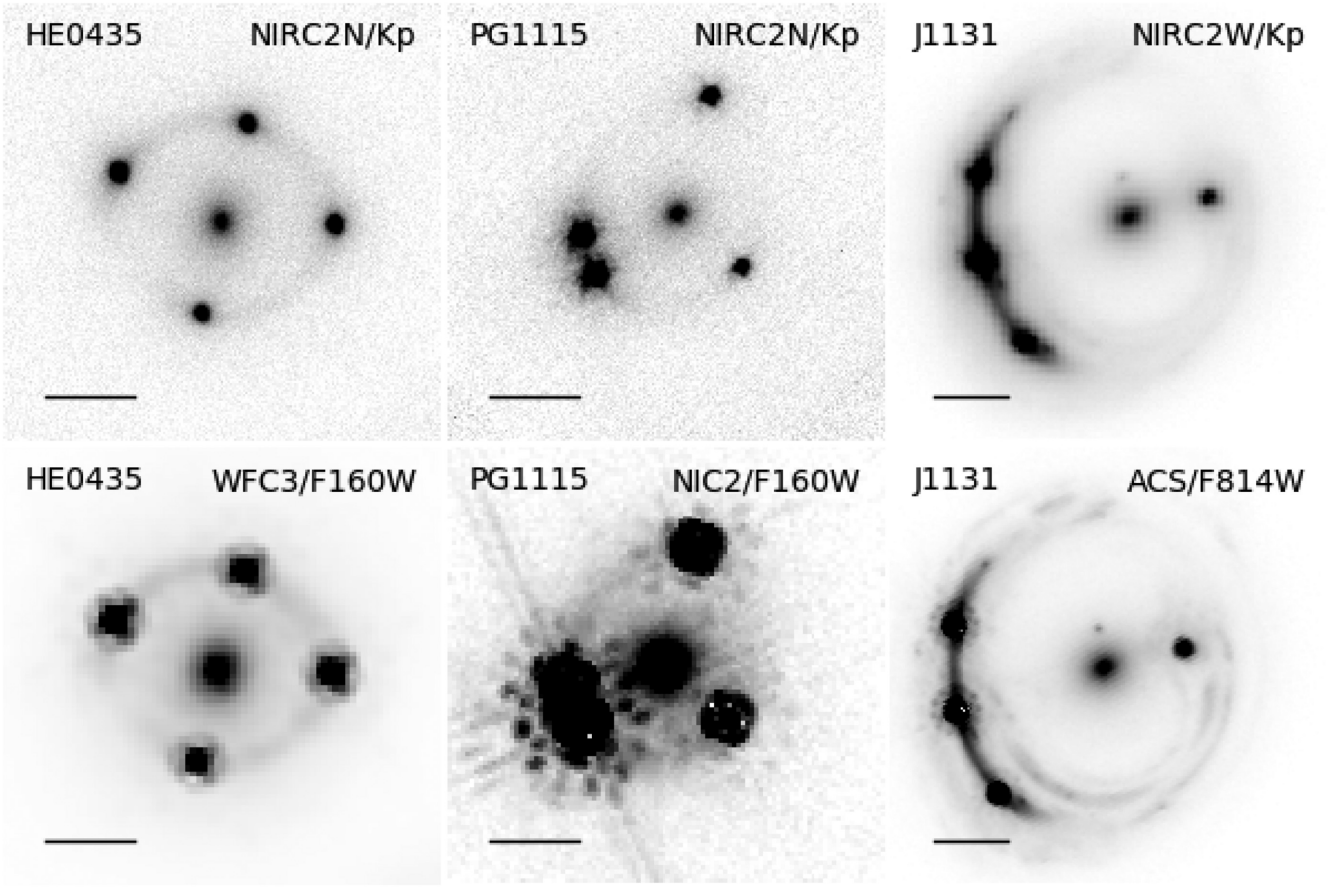


**Figure 1.** AO (top row) and *HST* (bottom row) images of the three gravitational lens systems. The solid horizontal line represents 1 arcsec scale. The foreground main lenses are located in the centre of the lens systems. The multiple lensed images and the extended arc around the lensing galaxy are from the background AGN and its host galaxy.

the CMB and the local value (=74.03 $\pm$ 1.42 km s$^{-1}$ Mpc$^{-1}$; Riess et al. 2019). The SN data can also be calibrated by the inverse distance ladder method to yield a model-dependent value of $H_0$. Under the assumption of the standard pre-recombination physics, combining BAO and SN with the CMB-calibrated physical scale of the sound horizon gives $H_0 = 67.3 \pm 1.1$ km s$^{-1}$ Mpc$^{-1}$ (Aubourg et al. 2015). A recent blind analysis with additional SN data from Dark Energy Survey gives $H_0 = 67.77 \pm 1.30$ km s$^{-1}$ Mpc$^{-1}$ (Macaulay et al. 2019). Both results are in excellent agreement with the *Planck* value ($H_0 = 67.27 \pm 0.6$ km s$^{-1}$ Mpc$^{-1}$) under the assumption of flat $\Lambda$CDM model (Planck Collaboration VI 2018). Furthermore, even without using the CMB anisotropy, the combination of BAO data with light element abundances produces *Planck*-like $H_0$ values (Addison et al. 2018). This indicates that systematic errors, especially in the *Planck* data analysis, most likely are not the main driver of the $H_0$ discrepancies. Similarly, the local distance ladder analyses have also passed a range of systematic checks (e.g. Efstathiou 2014; Cardona, Kunz & Pettorino 2017; Zhang et al. 2017; Dhawan, Jha & Leibundgut 2018; Feeney, Mortlock & Dalmasso 2018; Follin & Knox 2018; Riess et al. 2019), and rule out the local void scenario (Keenan, Barger & Cowie 2013; Fleury, Clarkson & Maartens 2017; Kenworthy, Scolnic & Riess 2019; Shanks, Hogarth & Metcalfe 2019)

There have been several attempts to address this $\sim$4.4$\sigma$ tension by extending the standard cosmological model, either by changing the size of the sound horizon in the early Universe (e.g. Heavens, Jimenez & Verde 2014; Wyman et al. 2014; Cuesta et al. 2015; Alam et al. 2017; Agrawal et al. 2019; Kreisch, Cyr-Racine & Doré 2019; Poulin et al. 2019) or by altering the expansion history (Efstathiou 2003; Linder 2004; Moresco et al. 2016; Alam et al. 2017).

Recent studies (e.g. Bernal, Verde & Riess 2016; Joudaki et al. 2018; Aylor et al. 2019; Lemos et al. 2019) have also tried to directly reconstruct $H(z)$ in order to investigate the $H_0$ tension in the context of a possibly poor understanding of the evolution of dark energy density. From an empirical point of view, the current SN and BAO data sets only support $w(z) = -1$ within the redshift range where data are available (Cuesta et al. 2016). A very recent and dramatic decrease in $w$ or the presence of strong dark energy at $3 < z < 1000$ may escape detection and still generate a high value of $H_0$ (Riess et al. 2016). Standard sirens could possibly explore the $z > 3$ range in the future and provide a high-precision $H_0$ measurement (Chen, Fishbach & Holz 2018b). Nevertheless, it is also important to note that some $H_0$-value tension remains even if we do not consider the distance ladder constraints. For example, the high-$\ell$ CMB power spectrum prefers an even lower $H_0$ value than that from the low-$\ell$ CMB power spectrum (Addison et al. 2016; Planck Collaboration VI 2018).

Given the various tensions across different data sets, any convincing resolution to the $H_0$ tensions, either due to unknown systematics or due to new physics, needs to simultaneously resolve multiple disagreements. Therefore, comparing the distance measurements among independent and robust methodologies to cross-examine

**Table 1.** Details of the AO observations.

| Lens | Instrument | Date | $t_{exp}$ |
|---|---|---|---|
| HE 0435−1223 | Keck/NIRC2-N | 2010-12-02 | 8100 |
| | Keck/NIRC2-N | 2012-01-01 | 2400 |
| PG 1115+080 | Keck/NIRC2-N | 2012-05-15 | 900 |
| | Keck/NIRC2-N | 2017-04-11 | 900 |
| RXJ 1131−1231 | Keck/NIRC2-W | 2012-05-16 | 1800 |
| | Keck/NIRC2-W | 2012-05-18 | 1800 |

the $H_0$ tension is probably the only way to shed light on the true answer.

### 1.2 Distance measurement from time-delay cosmography

Time-delay cosmography not only is a completely independent technique of the distance ladder methods, but also has the advantage of being a one-step measurement of combined cosmological distances. In addition, time-delay cosmography is a complementary and cost-effective alternative compared to Type Ia SN or BAO (Suyu et al. 2013; Tewes et al. 2013b). In a time-delay gravitational lens, the combined cosmological distance we can measure is called the time-delay distance (Suyu et al. 2010), which is a ratio of the angular diameter distances in the system

$$D_{\Delta t} \equiv (1+z_\ell)\frac{D_\ell D_s}{D_{\ell s}} \propto H_0^{-1}, \quad (1)$$

where $D_\ell$ is the distance to the lensing galaxy, $D_s$ is the distance to the background source, and $D_{\ell s}$ is the distance between the lens and the source. Furthermore, we can make a separate determination of $D_\ell$ by measuring the velocity dispersion of the lensing galaxy (Jee, Komatsu & Suyu 2015; Birrer, Amara & Refregier 2016; Jee et al. 2016; Birrer et al. 2019). First proposed by Refsdal (1964), the $H_0$ measurement requires modelling the mass in the lensing galaxy and along the line of sight, and measuring the time delays between multiple images via a monitoring program. The advantage of this method is that $D_{\Delta t}$ is primarily sensitive to $H_0$ and insensitive to the neutrino physics and spatial curvature, but still sensitive to the properties of dark energy (Bonvin et al. 2017; Birrer et al. 2019).

The H0LiCOW collaboration[1] is using strong gravitational lens systems to measure cosmological parameters (Suyu et al. 2017). The most recent measurement of $H_0$ from the collaboration used the doubly lensed quasar system, SDSS J1206+4332, to derive $H_0 = 68.8^{+5.4}_{-5.1}$ km s$^{-1}$ Mpc$^{-1}$ for that lens system alone, as well as combining the new system with previous H0LiCOW lenses (Suyu et al. 2009, 2010, 2013, 2014; Bonvin et al. 2017; Rusu et al. 2017; Sluse et al. 2017; Wong et al. 2017) to obtain a joint inference on $H_0$ with 3 per cent precision: $H_0 = 72.5^{+2.1}_{-2.3}$ km s$^{-1}$ Mpc$^{-1}$ (Birrer et al. 2019). This result agrees with the Riess et al. (2019) value within the $1\sigma$ uncertainties. More recently, the collaboration completed its analysis of WFI2033−4723 using the time delays from COSMOGRAIL[2] (Bonvin et al. 2019; Rusu et al. 2019a; Sluse et al. 2019).

Achieving the goal of obtaining a 1 per cent or better measurement of $H_0$ with time-delay cosmography requires a significantly larger sample of lensed quasars with high-quality data than has been analysed to date. Many new lensed quasars have been discovered (e.g. Lin et al. 2017; Schechter et al. 2017; Agnello et al. 2018; Ostrovski et al. 2018; Williams et al. 2018; Lemon, Auger & McMahon 2019; Rusu et al. 2019b) using state-of-the-art lens-finding techniques applied to current large sky surveys (e.g. Joseph et al. 2014; Agnello 2017; Ostrovski et al. 2017; Petrillo et al. 2017; Lanusse et al. 2018; Spiniello et al. 2018; Treu et al. 2018; Avestruz et al. 2019), and more are expected to be found with the Large Synoptic Survey Telescope (Oguri & Marshall 2010). Hence, a 1 per cent $H_0$ measurement from time-delay cosmography is a realistic expectation in the near future (e.g. Jee et al. 2015, 2016, 2019; de Grijs et al. 2017; Shajib, Treu & Agnello 2018; Suyu et al. 2018) if we can control the systematic effects to a sub-per cent level (Dobler et al. 2015; Liao et al. 2015; Ding et al. 2018).

[1] $H_0$ Lenses in COSMOGRAIL's Wellspring, www.h0licow.org.
[2] COSmological MOnitoring of GRAvItational Lenses.

**Table 2.** Lensing parameters for creating the magnification maps of MTDE. The values of $\kappa$, $\gamma$, and $\kappa_*/\kappa$ of HE 0435−1223 and PG 1115+080 are from our lens models.

| Name | img | $\kappa$ | $\gamma$ | $\kappa_*/\kappa$ |
|---|---|---|---|---|
| HE 0435−1223 | A | 0.473 | 0.358 | 0.347 |
| (this work) | B | 0.630 | 0.540 | 0.361 |
| | C | 0.494 | 0.327 | 0.334 |
| | D | 0.686 | 0.575 | 0.380 |
| PG 1115+080 | A1 | 0.424 | 0.491 | 0.259 |
| (this work) | A2 | 0.451 | 0.626 | 0.263 |
| | B | 0.502 | 0.811 | 0.331 |
| | C | 0.356 | 0.315 | 0.203 |
| RXJ 1131−1231 | A | 0.526 | 0.410 | 0.429 |
| (this work) | B | 0.459 | 0.412 | 0.434 |
| | C | 0.487 | 0.306 | 0.414 |
| | D | 0.894 | 0.807 | 0.581 |

### 1.3 Lens modelling with adaptive optics data

A critical input to achieving precise modelling of the mass distribution in the lensing galaxy is sensitive high-resolution imaging in which extended emission from the background source is detected. For this reason, many models of the lensing potential are based on imaging from *Hubble Space Telescope* (*HST*) (e.g. Suyu et al. 2009, 2010; Birrer, Amara & Refregier 2015; Birrer et al. 2016, 2019; Wong et al. 2017). Adaptive optics (AO) observations with ground-based telescopes can provide imaging with angular resolution that is comparable to or better than *HST* data, especially for systems that are faint in the optical and bright at near-infrared wavelengths, thus providing an attractive alternative for modelling the lens mass distribution (e.g. Lagattuta et al. 2012; Chen et al. 2016).

However, the challenge of using AO data is the unstable point spread function (PSF). Chen et al. (2016) showed that with a new iterative PSF reconstruction method applied to RXJ 1131−1231, the reconstructed PSF allows one to model the AO imaging down to the noise level, as well as providing tighter constraints on the lens model than were obtained from *HST* imaging of the system. In this new analysis, we apply the PSF reconstruction method to three lens systems, HE 0435−1223, PG 1115+080, and RXJ 1131−1231, which have not only high-resolution AO imaging and *HST* imaging, but also measured time delays, stellar velocity dispersions, and studies of their environments. We then infer their time-delay distances via detailed lens mass modelling. The AO data for RXJ 1131−1231 have already been analysed by Chen et al. (2016), but the analyses based on the AO imaging of HE 0435−1223 and PG 1115+080 are presented here for the first time.

**Table 3.** Quasar source parameters: the black hole mass ($M_{\rm bh}$), the Eddington ratio ($L/L_{\rm E}$), the accretion efficiency ($\eta$), the inclination angle (incl.), and the references.

| Name | $M_{\rm bh}$ ($M_\odot$) | $L/L_{\rm E}$ | $\eta$ | incl. (deg) | Reference |
|---|---|---|---|---|---|
| HE 0435−1223 | $5.75 \times 10^8$ | 0.1 | 0.1 | 0 | Sluse et al. (2012) |
| PG 1115+080 | $1.2 \times 10^9$ | 0.1 | 0.1 | 0 | Morgan et al. (2010) |
| RXJ 1131−1231 | $1.3 \times 10^8$ | 0.1 | 0.1 | 0 | Dai et al. (2010) |

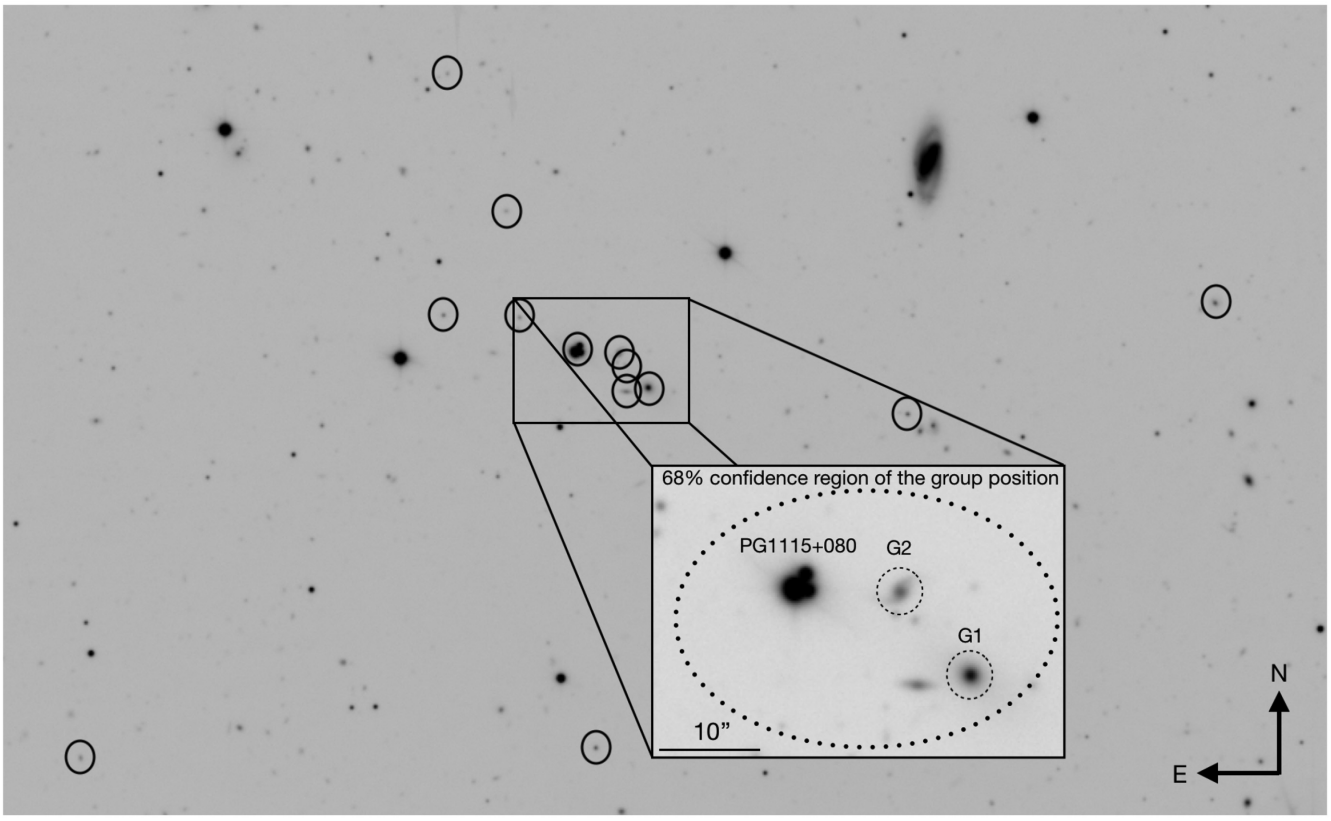


**Figure 2.** 2.2 m MPIA telescope image of the FOV around PG 1115+080. We show the galaxies/group that we explicitly model here. As PG 1115+080 is embedded in a nearby group that consists of 13 galaxies labelled with solid circles, we model not only the main lens but also the group explicitly. The dotted circle represents $1\sigma$ uncertainty of the priors of the group position ($\Delta$RA $= 23\farcs4$ and $\Delta$Dec. $= 15\farcs84$) measured in Wilson et al. (2016). Since G1 and G2 have the first two largest values of $\Delta_3 x$, we model either G1 or both G1 and G2 explicitly in addition to the main lens. We label G1 and G2 with the dashed circles.

The Keck AO imaging data are part of the Strong-lensing High Angular Resolution Programme (SHARP; Fassnacht et al., in preparation), which aims to study the nature of dark matter using high-resolution AO imaging (e.g. Lagattuta, Auger & Fassnacht 2010; Lagattuta et al. 2012; Vegetti et al. 2012; Hsueh et al. 2016, 2017, 2018; Spingola et al. 2018). The time-delay measurements are provided by the COSMOGRAIL group (e.g. Courbin et al. 2005; Vuissoz et al. 2007, 2008; Courbin et al. 2011; Rathna Kumar et al. 2013; Tewes, Courbin & Meylan 2013a; Tewes et al. 2013b; Bonvin et al. 2017, 2018), which aims to provide the highest-precision measurements of time delays. The lens environment of RXJ 1131−1231 and HE 0435−1223 studies is provided by the H0LiCOW team (Suyu et al. 2014; Rusu et al. 2017; Sluse et al. 2017)

The outline of the paper is as follows. In Section 2, we briefly recap the basics of obtaining inferences on cosmography with time-delay lenses and the statistical tools we use for this process. We describe the observations of HE 0435−1223, PG 1115+080, and RXJ 1131−1231 with the AO imaging system at the Keck Observatory in Section 3. In Section 4, we describe the models that we use to analyse the data. In Section 5, we elaborate the detailed lens modelling and the properties of the reconstructed PSF for each system. In Section 6, we present the joint cosmological inference from the AO imaging only as well as from combined AO plus *HST*. We summarize in Section 7.

## 2 BASIC THEORY

We briefly introduce the relation between cosmology and gravitational lensing in Section 2.1 and the joint inference of all information in Section 2.2.

### 2.1 Time-delay cosmography

When a compact variable background source, such as an active galactic nucleus (AGN) or SN, sitting inside its host galaxy is

strongly lensed by a foreground object, the distorted host galaxy shape combined with the time delay between the multiple images allows one to precisely determine a particular size of the system. One can express the excess time delays as

$$\Delta t = (1-\lambda)\frac{D_{\Delta t}}{c}\left[\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta})\right], \quad (2)$$

where $\boldsymbol{\theta}$, $\boldsymbol{\beta}$, and $\psi(\boldsymbol{\theta})$ are the image location, the source location, and the projected 2D lensing gravitational potential, respectively (Refsdal 1964; Shapiro 1964). The $\lambda$ parameter represents the lack of perfect knowledge of the full mass distribution, as discussed later.

The advantage of this formulation is the separability of the cosmographic information, contained in the $D_{\Delta t}$ parameter, and the lens modelling. This allows one to infer cosmographic information without the need for cosmological priors in the lens modelling. Although for the lens systems that contain multiple lensing galaxies at different redshifts a particular cosmological model needs to be applied to the analysis, the resulting effective $D_{\Delta t}$ is robust against cosmological model assumptions, in that the posterior distribution of the effective $D_{\Delta t}$ shows nearly identical distribution (within 1 per cent) for different background cosmological models (Wong et al. 2017; Rusu et al. 2019a).

Because of the mass-sheet transformation (MST; Falco, Gorenstein & Shapiro 1985; Schneider & Sluse 2013, 2014), the determination of $D_{\Delta t}$ is subject to an understanding of $\lambda$ (see also the discussion in Birrer et al. 2019). Thus, additional priors and information from simulations, environmental data, and the stellar velocity dispersion of the lensing galaxy are required to constrain the degeneracy between different mass profiles and the degeneracy between the mass profile and the mass sheet contributed by the environment (Suyu et al. 2010; Fassnacht, Koopmans & Wong 2011; Rusu et al. 2017; Tihhonova et al. 2018). In addition, a prior on the source size (Birrer et al. 2016) or having a background source of known brightness (e.g. a lensed SN; Grillo et al. 2018) can also put constraints on $\lambda$. We refer interested readers to Treu & Marshall (2016) and Suyu et al. (2018) for more details.

### 2.2 Joint inference

In this work, we will present our joint inference on $D_{\Delta t}$. We use $\boldsymbol{d}_i$ to denote the imaging data, where $i$ = HE, PG, and RXJ represent HE 0435−1223, PG 1115+080, and RXJ 1131−1231, respectively; we use $\boldsymbol{\Delta t}_i$ for time delays, $\boldsymbol{d}_{\mathrm{ENV}_i}$ to characterize the lens environments, and $\sigma_i$ for the stellar velocity dispersions of the lensing galaxies. Here, $\boldsymbol{\eta}_i$ are the parameters we want to infer from the data, and $\boldsymbol{A}$ denotes the discrete assumptions that we made in the models (e.g. whether the lensing galaxy is modelled using a power-law or NFW + stellar mass distribution). The posterior of $\boldsymbol{\eta}_i$ can be expressed as

$$P(\boldsymbol{\eta}_i|\boldsymbol{d}_i, \boldsymbol{\Delta t}_i, \sigma_i, \boldsymbol{d}_{\mathrm{ENV}_i}, \boldsymbol{A}) \propto P(\boldsymbol{d}_i, \boldsymbol{\Delta t}_i, \sigma_i, \boldsymbol{d}_{\mathrm{ENV}_i}|\boldsymbol{\eta}_i, \boldsymbol{A})P(\boldsymbol{\eta}_i|\boldsymbol{A}), \quad (3)$$

where $P(\boldsymbol{d}_i, \boldsymbol{\Delta t}_i, \sigma_i, \boldsymbol{d}_{\mathrm{ENV}_i}|\boldsymbol{\eta}_i, \boldsymbol{A}_i)$ is the joint likelihood for each lens. Since we assume that the environment can be decoupled from the lens, and that the data sets are independent,

$$P(\boldsymbol{d}_i, \boldsymbol{\Delta t}_i, \sigma_i, \boldsymbol{d}_{\mathrm{ENV}_i}|\boldsymbol{\eta}_i, \boldsymbol{A}_i) = P(\boldsymbol{d}_i|\boldsymbol{\eta}_i, \boldsymbol{A}_i)P(\boldsymbol{\Delta t}_i|\boldsymbol{\eta}_i, \boldsymbol{A}_i)P(\sigma_i|\boldsymbol{\eta}_i, \boldsymbol{A})P(\boldsymbol{d}_{\mathrm{ENV}_i}|\boldsymbol{\eta}_i, \boldsymbol{A}_i). \quad (4)$$

Note that we do not use the flux ratios of the lensed quasars as constraints on the model.

In order to explore the unmodelled systematic uncertainties that may arise from modelling choices, we vary the content of $\boldsymbol{A}$ for each lens. The marginalized integral can be expressed as

$$P(\boldsymbol{\eta}_i|\boldsymbol{d}_{i,\mathrm{tot}}) = \int P(\boldsymbol{\eta}_i|\boldsymbol{d}_{i,\mathrm{tot}}, \boldsymbol{A}_i)P(\boldsymbol{A}_i)\mathrm{d}\boldsymbol{A}_i \approx \sum_k P(\boldsymbol{\eta}_i|\boldsymbol{d}_{i,\mathrm{tot}}, \boldsymbol{A}_{i,k})P(\boldsymbol{A}_{i,k}), \quad (5)$$

where $\boldsymbol{A}_{i,k}$ and $\boldsymbol{d}_{i,\mathrm{tot}}$ represent the different model choices and all data sets for the lens system $i$, respectively. For ranking the models, we follow Birrer et al. (2019) to estimate the evidence, $P(\boldsymbol{A}_{i,k})$, by using the Bayesian information criterion (BIC), which is defined as

$$\mathrm{BIC} = \ln(n)k - 2\ln(\hat{L}), \quad (6)$$

where $n$ is the number of data points including the lens imaging, eight AGN positions, three time delays, and one velocity dispersion, $k$ is the number of free parameters in the lens model that are given uniform priors, plus two source-position parameters, plus one anisotropy radius to predict the velocity dispersion, and $\hat{L}$ is the maximum likelihood of the model, which is the product of the AGN position likelihood, the time-delay likelihood, the pixelated image plane likelihood, and the kinematic likelihood. The image plane likelihood is the Bayesian evidence of the pixelated-source intensity reconstruction using the arcmask imaging data (see Suyu & Halkola 2010) times the likelihood of the lens model parameters within the image plane region that excludes the arcmask. We follow Birrer et al. (2019) and calculate the relative BIC and weighting for the SPEMD and composite models separately to avoid biases due to our choice of lens model parameterization.

## 3 DATA

The analysis in this paper is based on new Keck AO and archival *HST* observations of three gravitational lens systems. In this section, we describe the lens sample and the data acquisition and analysis.

### 3.1 The sample

The sample consists of three well-known lensed quasar systems. Images of these lens systems are shown in Fig. 1.

(i) *HE 0435−1223:* The HE 0435−1223 system (J2000: $4^{\mathrm{h}}38^{\mathrm{m}}14\overset{\mathrm{s}}{.}9$, $12^{\circ}17'14\overset{''}{.}4$) is a quadruply lensed quasar discovered by Wisotzki et al. (2002). The main lensing galaxy is at a redshift of $z_\ell = 0.4546$ (Morgan et al. 2005), and the source redshift is $z_\mathrm{s} = 1.693$ (Sluse et al. 2012). The lens resides inside a galaxy group that contains at least 12 galaxies, with a velocity dispersion $\sigma = 471 \pm 100\,\mathrm{km\,s^{-1}}$ (e.g. Momcheva et al. 2006; Wong et al. 2011; Wilson et al. 2016; Sluse et al. 2017). Wong et al. (2017) measured the stellar velocity dispersion of the lensing galaxy to be $\sigma = 222 \pm 15\ \mathrm{km\,s^{-1}}$. The time delays of this system were measured by Bonvin et al. (2017) with ∼6.5 per cent uncertainties.

(ii) *PG 1115+080:* This four-image system was the second strong gravitational lens system to be discovered (Weymann et al. 1980). The system is located at $11^{\mathrm{h}}18^{\mathrm{m}}16\overset{\mathrm{s}}{.}899$, $+7^{\circ}45'58\overset{''}{.}502$ (J2000). The background quasar with a redshift of $z_\mathrm{s} = 1.722$ is lensed by a galaxy with $z_\ell = 0.3098$ (Henry & Heasley 1986; Christian, Crabtree & Waddell 1987; Tonry 1998). The lensed images are in a classic 'fold' configuration, with an image pair A1 and A2 near the critical curve. The lens resides inside a galaxy group with 13 known members that has a velocity dispersion of

$\sigma = 390 \pm 60\,\mathrm{km\,s^{-1}}$ (Wilson et al. 2016). Tonry (1998) measured the stellar velocity dispersion of the lensing galaxy to be $\sigma = 281 \pm 25\,\mathrm{km\,s^{-1}}$. The time delays of this system were measured by Bonvin et al. (2018) with ∼6.4 per cent uncertainties.

(iii) *RXJ 1131−1231:* The RXJ 1131−1231 system (J2000: $11^{\mathrm{h}}31^{\mathrm{m}}52^{\mathrm{s}}$, $-12^{\circ}31'59''$) is a quadruply lensed quasar discovered by Sluse et al. (2003). The spectroscopic redshifts of the lensing galaxy and the background source are at $z_\ell = 0.295$ (Suyu et al. 2013) and $z_\mathrm{s} = 0.657$ (Sluse et al. 2007), respectively. Suyu et al. (2013) measured the stellar velocity dispersion of the lensing galaxy to be $\sigma = 323 \pm 20\,\mathrm{km\,s^{-1}}$. The time delays were measured by Tewes et al. (2013b) with ∼1.5 per cent uncertainties.

### 3.2 Keck adaptive optics imaging

All three lens systems were observed at $K'$-band with the Near-Infrared Camera 2 (NIRC2), sitting behind the AO bench on the Keck II Telescope, as part of the SHARP AO effort (Fassnacht et al., in preparation). The targets were observed with either the narrow camera set-up, which provides a roughly 10 arcsec × 10 arcsec field of view and a pixel scale of 9.942 mas, or the wide camera that gives a roughly 40 arcsec × 40 arcsec field of view and a pixel scale of 39.686 mas. Details of the observations are provided in Table 1.

The NIRC2 data were reduced using the SHARP PYTHON-based pipeline, which performs a flat-field correction, sky subtraction, correction of the optical distortion in the images, and a co-addition of the exposures. During the distortion correction step, the images are resampled to produce final pixel scales of 10 mas pixel$^{-1}$ for the narrow camera and 40 mas pixel$^{-1}$ for the wide camera. The narrow camera pixels oversample the PSF, which has typical full width at half-maximum (FWHM) values of 60–90 mas. Therefore, to improve the modelling efficiency for the narrow camera data, we perform a 2 × 2 binning of the images produced by the pipeline to obtain images that have a 20 mas pixel$^{-1}$ scale. Further details on RXJ 1131−1231, which was observed with the NIRC2 wide camera, can be found in Chen et al. (2016).

### 3.3 *Hubble Space Telescope* imaging

All three lens systems have been observed by *HST* (GO-9375, PI: Kochanek; GO-9744, PI: Kochanek; GO-12889, PI: Suyu). The *HST* imaging of both RXJ 1131−1231 (Suyu et al. 2013) and HE 0435−1223 (Wong et al. 2017) was analysed in the previous work. Therefore, the inferences from these previous models are combined with those from the new AO models in Section 5. In contrast, the *HST* data for PG 1115+080 have not been modelled using the latest pixelated techniques, although Treu & Koopmans (2002) combined lensing geometry and velocity dispersion to study the content of the luminous matter and dark matter profiles. Therefore, we perform a joint modelling procedure on the AO and *HST* data for PG 1115+080.

### 3.4 MPIA 2.2 m imaging

The contribution of the line-of-sight (LOS) mass distribution to the lensing requires deep wide-field imaging of the region surrounding the lens system. For HE 0435−1223 and RXJ 1131−1231, we have used *HST*/ACS or Subaru SuprimeCam imaging, which have been analysed as part of our previous work on these systems (Suyu et al. 2014; Rusu et al. 2017). To achieve the requisite combination of depth and area for PG 1115+080, we co-added 95 images of the field taken with the Wide Field Imager (Baade et al. 1999) mounted at the Cassegrain focus of the MPIA 2.2 m telescope. The camera provides a pixel scale of 0.238 arcsec. The data were obtained through ESO $BB\#R_\mathrm{c}/162$ filter. These are the same data used by COSMOGRAIL to measure the time delay for this system (see section 2.1 of Bonvin et al. 2018, for details); each image has been exposed for 330 s and covers a field of view of ∼8 arcmin × 16 arcmin. The data reduction process follows the standard procedure, including master bias subtraction, master flat fielding, sky subtraction, fringe pattern removal, and finally exposure-to-exposure normalization using SEXTRACTOR on field stars prior to co-adding the exposures. The final co-added image has an effective seeing of 0.86 arcsec.

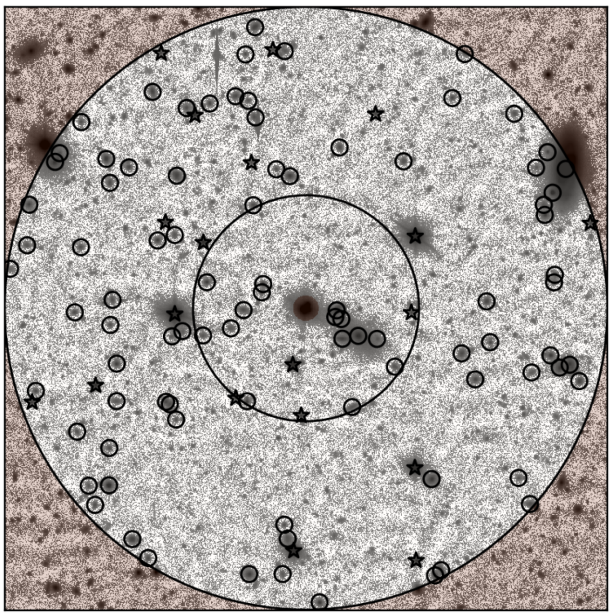

**Figure 3.** *Rc*-band 240 arcsec × 240 arcsec image around PG 1115+080. These are the same data shown in Fig. 2, but with contrast chosen to better show the detected objects. The large circles mark the 45 and 120 arcsec radii apertures centred on the lens. The inner 5 arcsec and outer >120 arcsec masked regions are shown in colour. For all objects $R \leq 23$, small circles mark galaxies and star symbols mark stars.

## 4 LENS MODELS

In this section, we describe the models that we use for fitting the high-resolution imaging data, including the lens mass models in Section 4.1, lens light models in Section 4.2, the models for constraining the MST in Section 4.3, and the time-delay prediction models in Section 4.4. We use GLEE, a strong lens modelling code developed by S. H. Suyu and A. Halkola to model the lens arc, lens light, and lens AGNs simultaneously (Suyu & Halkola 2010; Suyu et al. 2012). The source galaxy is reconstructed on a regular grid, and we choose curvature as our regularization form (Suyu et al. 2006). The AO PSF is reconstructed by the method developed in Chen et al. (2016), which uses the four lensed quasars as prior information to iteratively reconstruct the AO PSF from the imaging data. The tests in Chen et al. (2016) showed that the PSF reconstruction method can recover the structure of the host galaxy and PSF in mock data sets.

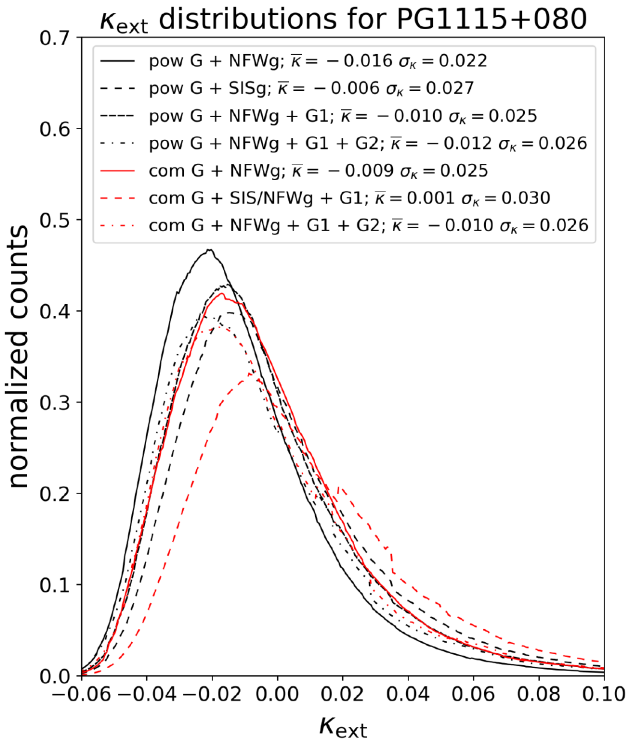


**Figure 4.** Distributions of $\kappa_{ext}$ for PG 1115+080, for various lens models and their associated shear values. The constraints used to produce the distributions from the Millennium Simulation are the external shear $\gamma$, plus the combination of weighted counts corresponding to galaxy number counts inside the 45 and 120 arcsec apertures, as well as number counts weighted by the inverse of the distance of each galaxy to the lens. The numerical constraints are reported in Table C1. The distributions with G1 or G2 marked are calculated by removing those galaxies from the weighted count constraints, since these galaxies are explicitly included in the lens models. The size of the histogram bin is $\Delta\kappa_{ext} = 0.00\,055$. As the original distributions are noisy, we plot their convolution with a large smoothing window of size $30 \times \Delta\kappa_{ext}$. In the legend, 'pow' refers to the power-law model, 'com' refers to the composite model, $\bar{\kappa}$ refers to the median of the distribution, and $\sigma_\kappa$ refers to the semi-difference of the 84 and 16 percentiles of the distribution.

### 4.1 Mass models

The following three analytical functions are used for modelling the main lens, nearby groups, and nearby galaxies:

(i) *SPEMD:* Many studies have shown that a power-law model provides a good first-order description of the lensing galaxies for galaxy–galaxy lensing (e.g. Koopmans et al. 2006, 2009; Suyu et al. 2009; Auger et al. 2010; Barnabè et al. 2011; Sonnenfeld et al. 2013). Thus, for every lens, we model the mass distribution of the lensing galaxy with a singular power-law elliptical mass distribution (Barkana 1998). The main parameters include radial slope ($\gamma'$), Einstein radius ($\theta_E$), and the axial ratio of the elliptical isodensity contour ($q$).

(ii) *Composite:* We follow Suyu et al. (2014) and test a composite (baryonic + dark matter) model. The baryonic component is modelled by multiplying the lens surface brightness distribution by a constant *M*/*L* ratio parameter (see Section 4.2). For the dark matter component, we adopt the standard NFW profile (Navarro, Frenk & White 1996; Golse & Kneib 2002) with the following parameters: halo normalization ($\kappa_s$), halo scale radius ($r_s$), and halo minor-to-major axial ratio ($q$), as well as associated position angle ($\theta_q$). Note that the ellipticity is implemented in the potential for the dark matter. The ellipticities of our three lenses are all inside the range where elliptical potential models are a good description of elliptical mass distributions and thus can be reliably applied to observational data (Golse & Kneib 2002).

(iii) *SIS:* Singular isothermal sphere (SIS) models are used to describe the nearby group and the individual galaxies inside the group of PG 1115+080, the nearby galaxies of HE 0435−1223, and the satellite of RXJ 1131−1231.

### 4.2 Lens light models

The following two analytical functions are used to model the lens light distribution:

(i) *2Sérsic:* We model the light distribution of the lens galaxy with two concentric elliptical Sérsic profiles. For all three lens systems, we found that a single Sérsic profile was insufficient for modelling the light distributions.

(ii) *2Chameleon:* The chameleon profile is the difference of two isothermal profiles. It mimics a Sérsic profile and enables computationally efficient lens modelling (Dutton et al. 2011). The parametrized Chameleon profile can be found in Suyu et al. (2014). We convert the Chameleon light profile to mass with an additional constant *M*/*L* ratio parameter when modelling the composite model described in Section 4.1.

### 4.3 Strategies for mitigating the mass-sheet transformation

The MST is a known degeneracy in lens modelling, in which one can transform a projected mass distribution, $\kappa(\theta)$, into infinite sets of $\kappa_\lambda(\theta)$ via

$$\kappa_\lambda(\theta) = (1-\lambda)\kappa(\theta) + \lambda, \tag{7}$$

without degrading the fit to the imaging. The corresponding time-delay distance changes via

$$D_{\Delta t} = D^{\rm true}_{\Delta t,\lambda} = \frac{D^{\rm model}_{\Delta t}}{1-\lambda}. \tag{8}$$

This degeneracy can be produced by both the lens environment and an incorrect description of the mass distribution in the lensing galaxy. We discuss the approach for estimating the contribution from the environment in Section 4.3.1, and for ranking the mass models with kinematic information in Section 4.3.2.

#### *4.3.1 Mass along the line of sight*

If we perfectly know the true $\kappa(\theta)$, then the role of $\lambda$ in equation (7) can be understood by looking at the behaviour far from the lensing galaxy, i.e. as $\theta \to \infty$. In this regime, the mass distribution of the lensing galaxy, $\kappa(\theta)$, approaches 0 and, thus, $\lambda$ can be interpreted as a constant-density mass sheet contributed by the lens environment. It is exactly the first-order form produced by LOS structure when its effect is small. Therefore, in our models we identify $\lambda$ with $\kappa_{ext}$, the physical convergence associated with LOS structures that do not affect the kinematics of the strong lens galaxy (e.g. Suyu et al. 2010; Wong et al. 2017; Birrer et al. 2019)

$$D_{\Delta t} = \frac{D^{\rm model}_{\Delta t}}{1-\kappa_{ext}}. \tag{9}$$

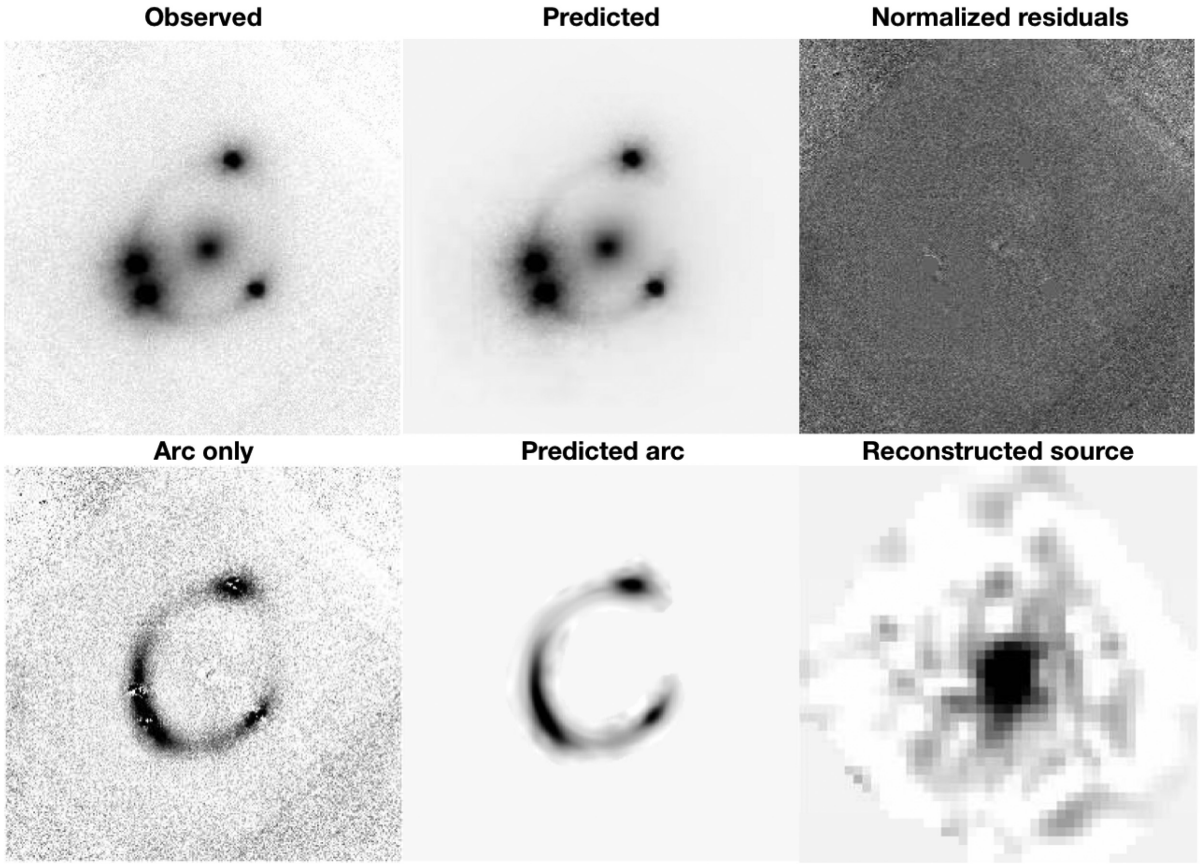


**Figure 5.** PG 1115+080 AO image reconstruction of the most probable model with a source grid of 43 × 43 pixels and 59 × 59 pixels PSF for convolution of spatially extended images. *Top left*: The PG 1115+080 AO image. *Top middle*: The predicted image of all components, including lens light, arc light, and AGN light. *Top right*: Image residuals, normalized by the estimated $1\sigma$ uncertainty of each pixel. *Bottom left*: The arc-only image that removes the lens light and AGN light from the observed image. *Bottom middle*: Predicted lensed image of the background AGN host galaxy. *Bottom right*: The reconstructed host galaxy of the AGN in the source plane.

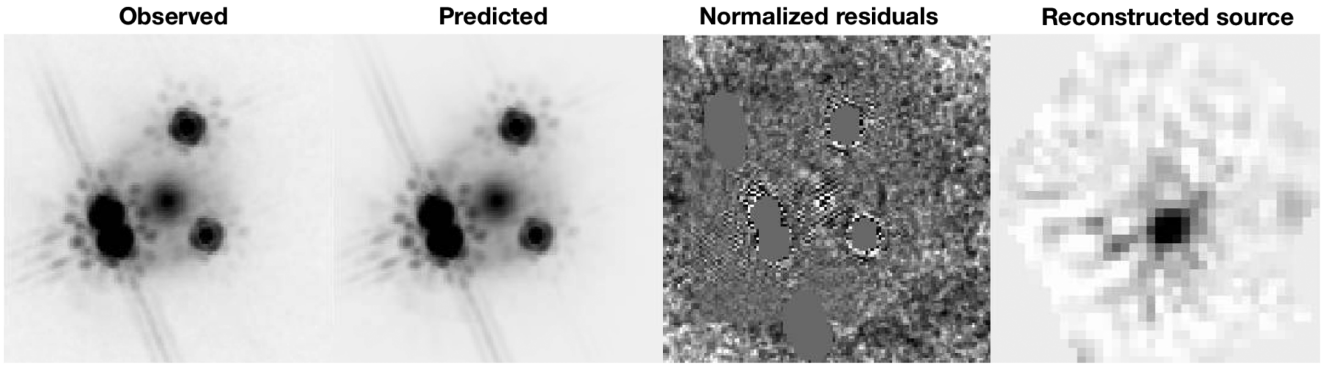


**Figure 6.** PG 1115+080 *HST* image reconstruction of the most probable model with a source grid of 38 × 38 pixels. From the left to the right: the PG 1115+080 *HST* image, the predicted image of all components, including lens light, arc light, and AGN light, image residuals normalized by the estimated $1\sigma$ uncertainty of each pixel, and the reconstructed host galaxy of the AGN in the source plane.

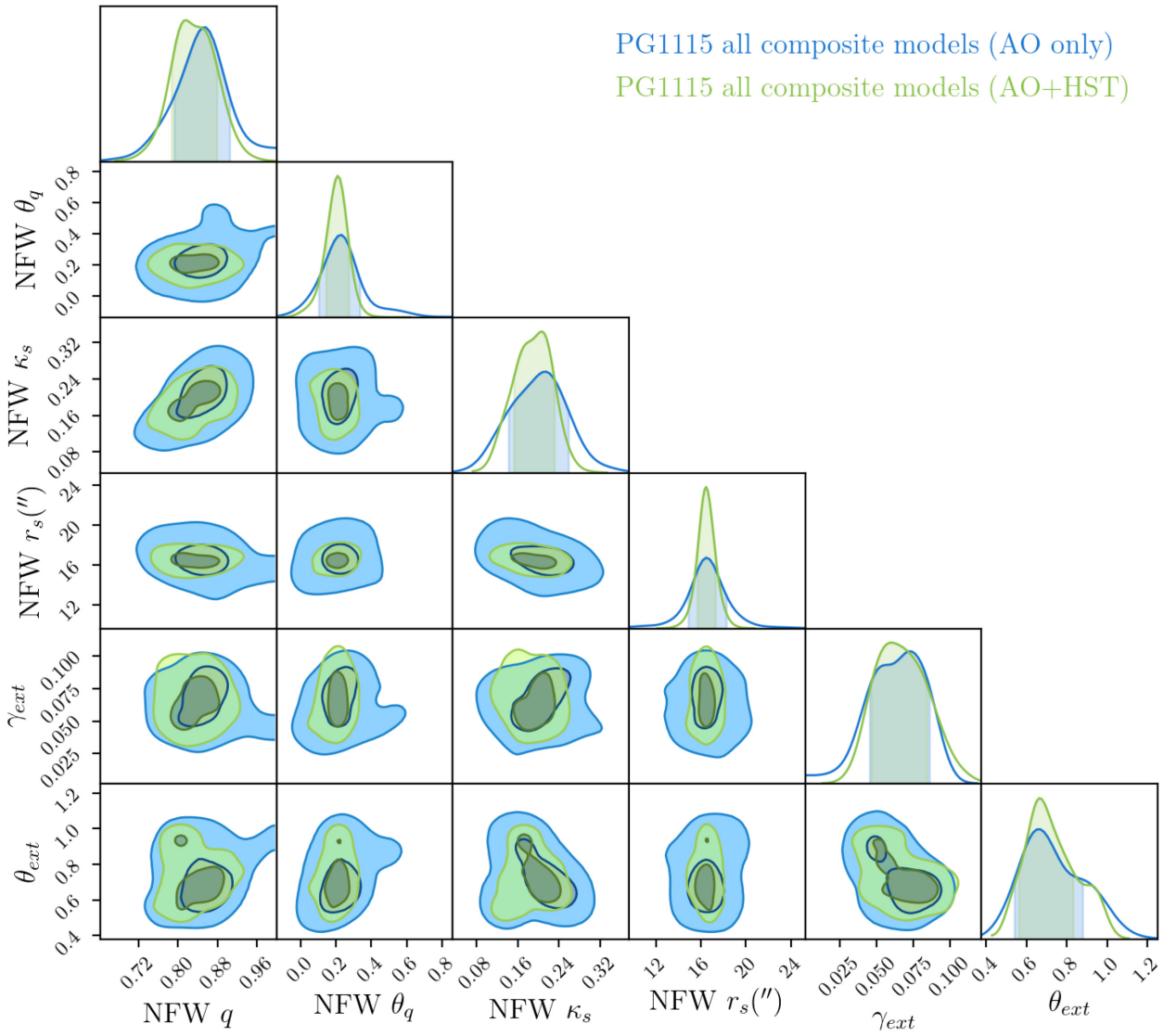


**Figure 7.** Marginalized parameter distributions from the PG 1115+080 composite lens model results. We show the comparison between using only AO imaging data and both AO and *HST* imaging data. The contours represent the 68.3 and 95.4 per cent quantiles.

In addition, the second-order distortion from the LOS also produces a tidal stretching on the lens images.

We use the following models to capture these two effects.

(i) *Shear:* Shear distorts the lensed image shapes and thus it can be detected in the modelling process. We express the lens potential in polar coordinates $(\theta, \varphi)$ to model the external shear on the imaging plane

$$\psi_{\rm ext}(\theta, \varphi) = \frac{1}{2}\gamma_{\rm ext}\theta^2 \cos 2(\varphi - \varphi_{\rm ext}), \quad (10)$$

where $\gamma_{\rm ext}$ is the shear strength and $\varphi_{\rm ext}$ is the shear angle. The shear position angle of $\varphi_{\rm ext} = 0^\circ$ corresponds to a shearing along $\theta_1$, whereas $\varphi_{\rm ext} = 90^\circ$ corresponds to shearing along $\theta_2$.[3]

(ii) *Millennium Simulation:* Due to the MST, the lens images do not provide direct information on $\kappa_{\rm ext}$. We thus use the results of ray tracing by Hilbert et al. (2009) through the Millennium Simulation (Springel et al. 2005) to statistically estimate the mass contribution along the line of sight to our lenses. This technique was first employed by Suyu et al. (2010), who took the ratio of observed galaxy number counts in an aperture around a lens to those in a control survey (from Fassnacht et al. 2011), in order to measure the local over/underdensity of galaxies in the lens fields. They then selected lines of sight of similar over/underdensity from the Millennium Simulation, with their corresponding values of the convergence, thus producing a probability distribution for $\kappa_{\rm ext}$: $P(\kappa_{\rm ext}|\boldsymbol{d}_{\rm ENV})$. Suyu et al. (2013) later used, in addition to the number counts, the shear value inferred from lens modelling as an additional constraint on $\kappa_{\rm ext}$. Greene et al. (2013) showed that further constraints, in the form of weighted number counts, can be derived by incorporating physical quantities relevant to lensing such as the distance of each galaxy to the lens, redshifts, luminosities, and stellar masses. Birrer et al. (2019) expressed the technique as an application of approximate Bayesian computing and further combined weighted number count constraints from multiple

[3]Our (right-handed) coordinate system $(\theta_1, \theta_2)$ has $\theta_1$ along the East–West direction and $\theta_2$ along the North–South direction.

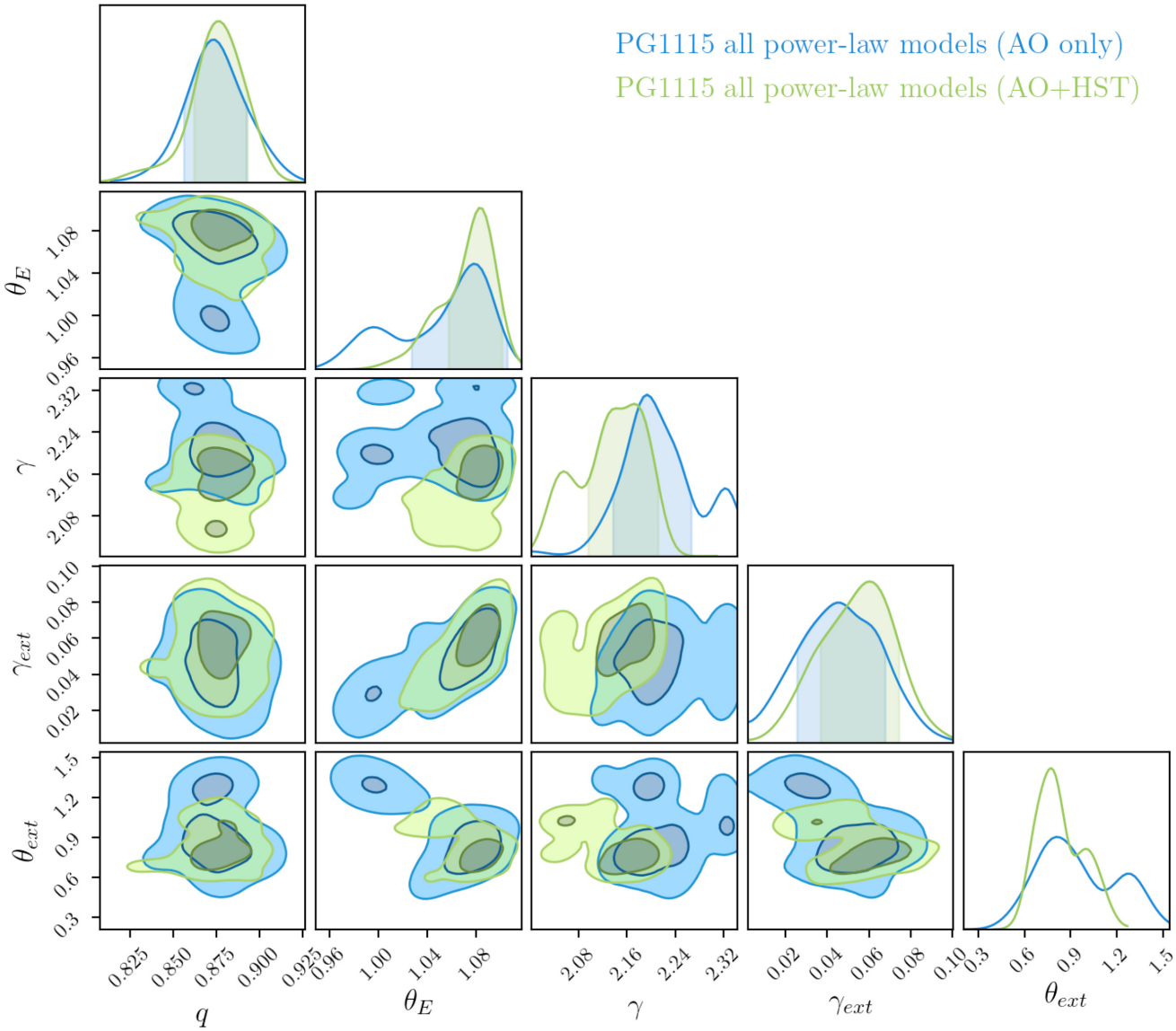


**Figure 8.** Marginalized parameter distributions from the PG 1115+080 power-law lens model results. We show the comparison between using only AO imaging data and both AO and *HST* imaging data. The contours represent the 68.3 and 95.4 per cent quantiles.

aperture radii. In Sections 5.1.5, 5.2.3, and 5.3.2, we provide an implementation of this technique for each of our three lenses, customized to the nature of the available environment data. For the main numerical and mathematical details of the implementation, we refer the reader to Rusu et al. (2017).

If the mass along the line of sight is large enough such that we cannot ignore the higher-order terms (flexion and beyond), we need to model that mass explicitly. McCully et al. (2014, 2017) give a quantitative term, flexion shift ($\Delta_3 x$), which estimates the deviations in lensed image positions due to third-order (flexion) terms, and suggest that if $\Delta_3 x$ is higher than $10^{-4}$ arcsec (for the typical galaxy-scale lenses that we are studying here), one should model the perturbers explicitly to avoid biasing $H_0$ at the (sub-) per cent level. We follow this convention to choose which galaxies are included in our models. Note that the $\Delta_3 x$ threshold is based on mock data that do not include extended arcs. Therefore, this criterion should already be conservative since real lens imaging with extended arcs provides more constraining power than the point-source imaging that was used to set the threshold.

#### 4.3.2 *Lens kinematics*

Conversely, even if we perfectly know the mass along the line of sight, the value of $\lambda$ remains uncertain because we do not know the true $\kappa(\theta)$ distribution. As pointed out by Schneider & Sluse (2013), when assuming a specific mass profile, one artificially breaks the internal MST (Koopmans 2004), a special case of the source-position transformation (Schneider & Sluse 2014; Wertz & Orthen 2018; Wertz, Orthen & Schneider 2018). However, recent work based on the Illustris simulation has indicated that this effect may be of less concern for massive galaxies such as those in the H0LiCOW sample (Xu et al. 2016). The MST allows many different mass distributions within the Einstein radius of the lens, as long as the integrated $\kappa$ within the Einstein ring is preserved. However, the stellar velocity dispersion is sensitive to the integrated value of

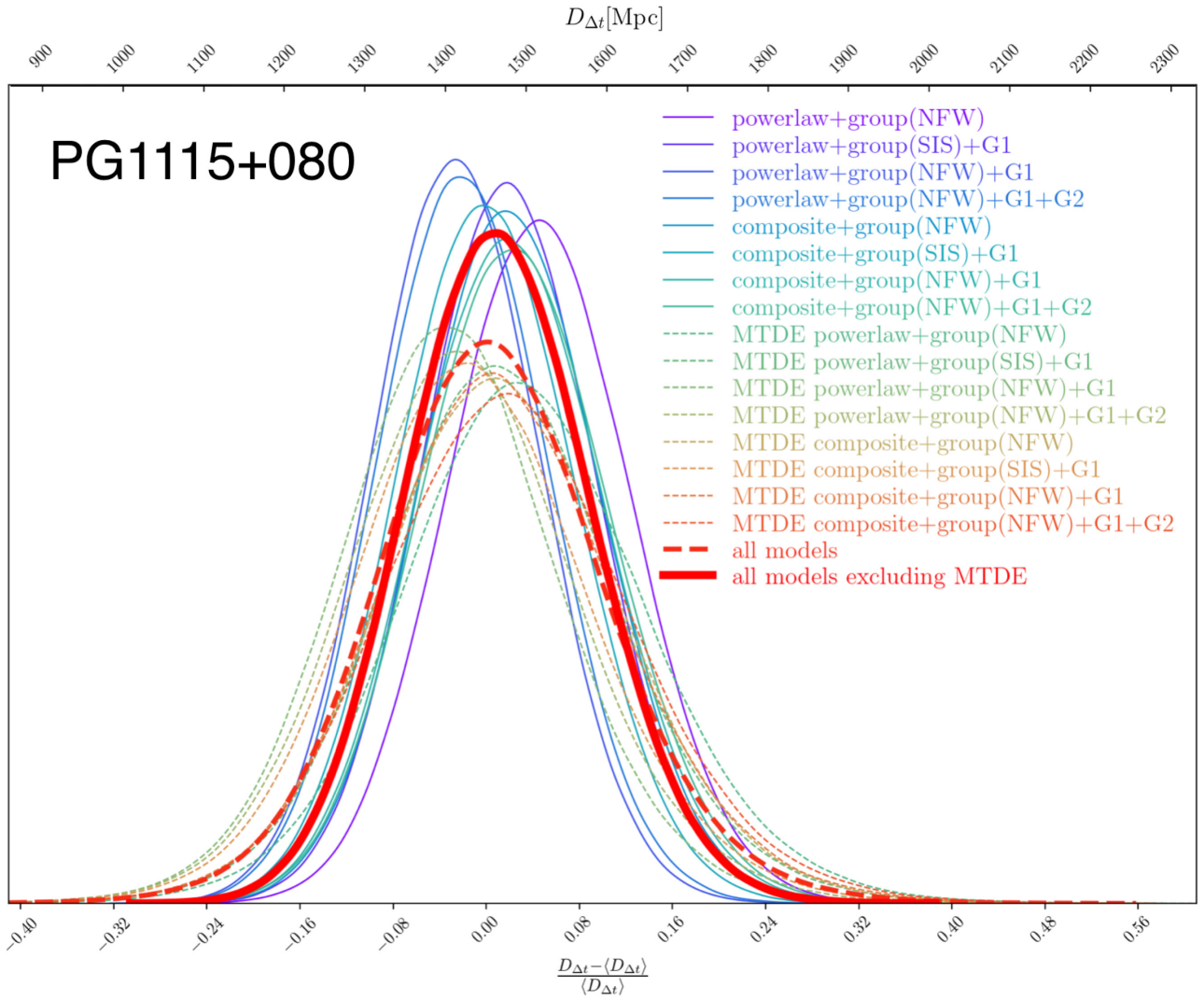


**Figure 9.** The $D_{\Delta t}$ of different model choices from PG 1115+080. We explore 160 model choices in total with all different combinations of choices among two kinds of main lens models, various mass models for the group, five different resolutions of the reconstructed source, and three different priors of the accretion disc sizes (or no MTDE). The solid lines are the cases without including the MTDE, while the dashed lines are the cases including MTDE. Each dashed line has marginalized three different kinds of accretion disc sizes. The last chain, which excludes the MTDE cases, is used to infer the value of $H_0$.

$\kappa$ within the *effective* radius of the lensing galaxy, which is often different from the Einstein radius. Thus, we can use the observed stellar velocity dispersion of the lensing galaxy to rank different mass models (Treu & Koopmans 2002, 2004).

In the lens modelling of RXJ 1131−1231, Suyu et al. (2014) have shown that by including the velocity dispersion, one can obtain a robust $D_{\Delta t}$ when considering both power-law and composite models. Sonnenfeld (2018) also shows that velocity dispersion is the key to obtaining an unbiased $H_0$ measurement. Hence, we follow Suyu et al. (2014) in adopting the composite model and also incorporating the velocity dispersion into the modelling to mitigate this internal mass-profile degeneracy.

There are three different components needed to predict the velocity dispersion.

(i) *A 3D mass distribution:* Following Suyu et al. (2010), one can obtain the 3D lens mass from the lens modelling by assuming spherical symmetry.[4] In general, the spherically symmetric 3D mass density of the lens can be expressed as

$$\rho_{\rm local}(r) = \rho_0 r_0^n F_n(r), \tag{11}$$

where $\rho_0 r_0^n$ and $F_n(r)$ are the normalization and the mass density distribution. By integrating $\rho_{\rm local}$ within a cylinder with radius given by the Einstein radius, $R_{\rm Ein}$, one obtains

$$\begin{aligned} M_{\rm local} &= 4\pi \int_0^\infty {\rm d}z \int_0^{R_{\rm Ein}} \rho_0 r_0^n F_n(\sqrt{s^2+z^2}) s \, {\rm d}s \\ &= \rho_0 r_0^n M_{\rm 2D}(R_{\rm Ein}), \end{aligned} \tag{12}$$

[4]Yıldırım, Suyu & Halkola (2019) show that the spherical symmetry assumption does not produce obvious signs of bias on inferring the $D_\ell$ and $D_{\Delta t}$.

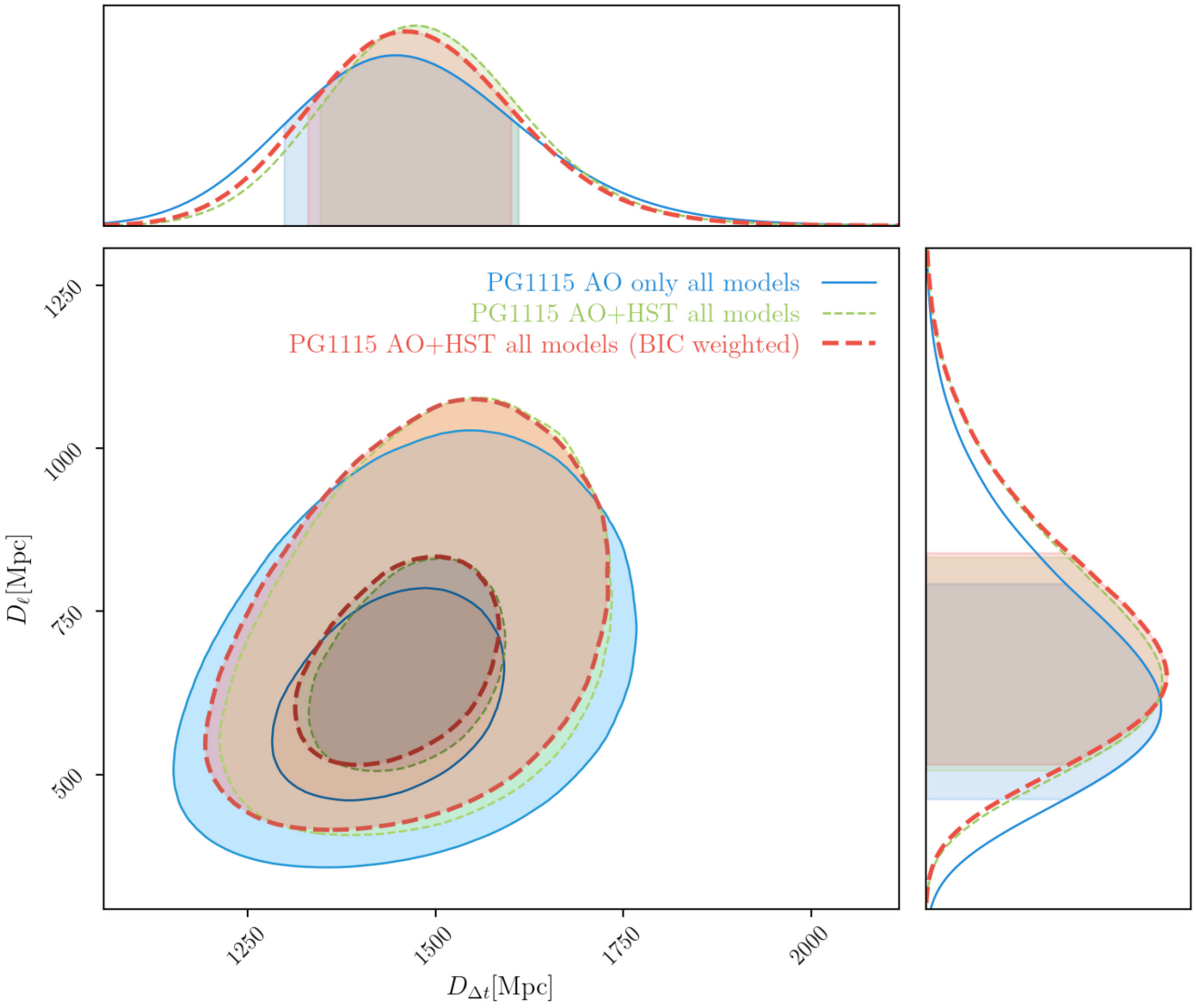


**Figure 10.** The unblinded $D_{\Delta t}$ and $D_\ell$ of PG 1115+080. We plot the results with only using AO imaging, using both AO and *HST* imaging, and the BIC weighted results.

where $M_{\rm 2D}(R_{\rm Ein})$ is the projected mass within $R_{\rm Ein}$. The mass contained in $M_{\rm local}$ is

$$M_{\rm local} = M_{\rm Ein} - M_{\rm ext} = \pi R_{\rm Ein}^2 \Sigma_{\rm cr}(1-\kappa_{\rm ext}), \tag{13}$$

where $M_{\rm ext}$ represents the mass contribution from $\kappa_{\rm ext}$ and

$$\Sigma_{\rm cr} = \frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_\ell D_{\ell {\rm s}}} \tag{14}$$

is the critical surface mass density. Combining equation (12) with equation (13), the normalization in equation (11) can be expressed as

$$\rho_0 r_0^n = \frac{\pi R_{\rm Ein}^2 \Sigma_{\rm cr}(1-\kappa_{\rm ext})}{M_{\rm 2D}(R_{\rm Ein})}. \tag{15}$$

Substituting this in equation (11), we obtain

$$\rho_{\rm local} = \frac{\pi R_{\rm Ein}^2 \Sigma_{\rm cr}(1-\kappa_{\rm ext})}{M_{\rm 2D}(R_{\rm Ein})} F_n(r). \tag{16}$$

Although there is $(1-\kappa_{\rm ext})$ in equation (16), the normalization of the local mass density distribution remains invariant (Yıldırım et al. 2019) as $\Sigma_{\rm cr}$ can be re-expressed as

$$\Sigma_{\rm cr} = \frac{c^2}{4\pi G}\frac{D_{\Delta t}}{1+z_\ell}\frac{1}{D_\ell^2} = \frac{c^2}{4\pi G}\frac{D_{\Delta t}^{\rm model}}{(1+z_\ell)(1-\kappa_{\rm ext})}\frac{1}{D_\ell^2}, \tag{17}$$

where $(1-\kappa_{\rm ext})$ term cancels out in equation (16).

(ii) *An anisotropy component:* We assume the anisotropy component in the form of an anisotropy radius, $r_{\rm ani}$, in the Osipkov–Merritt

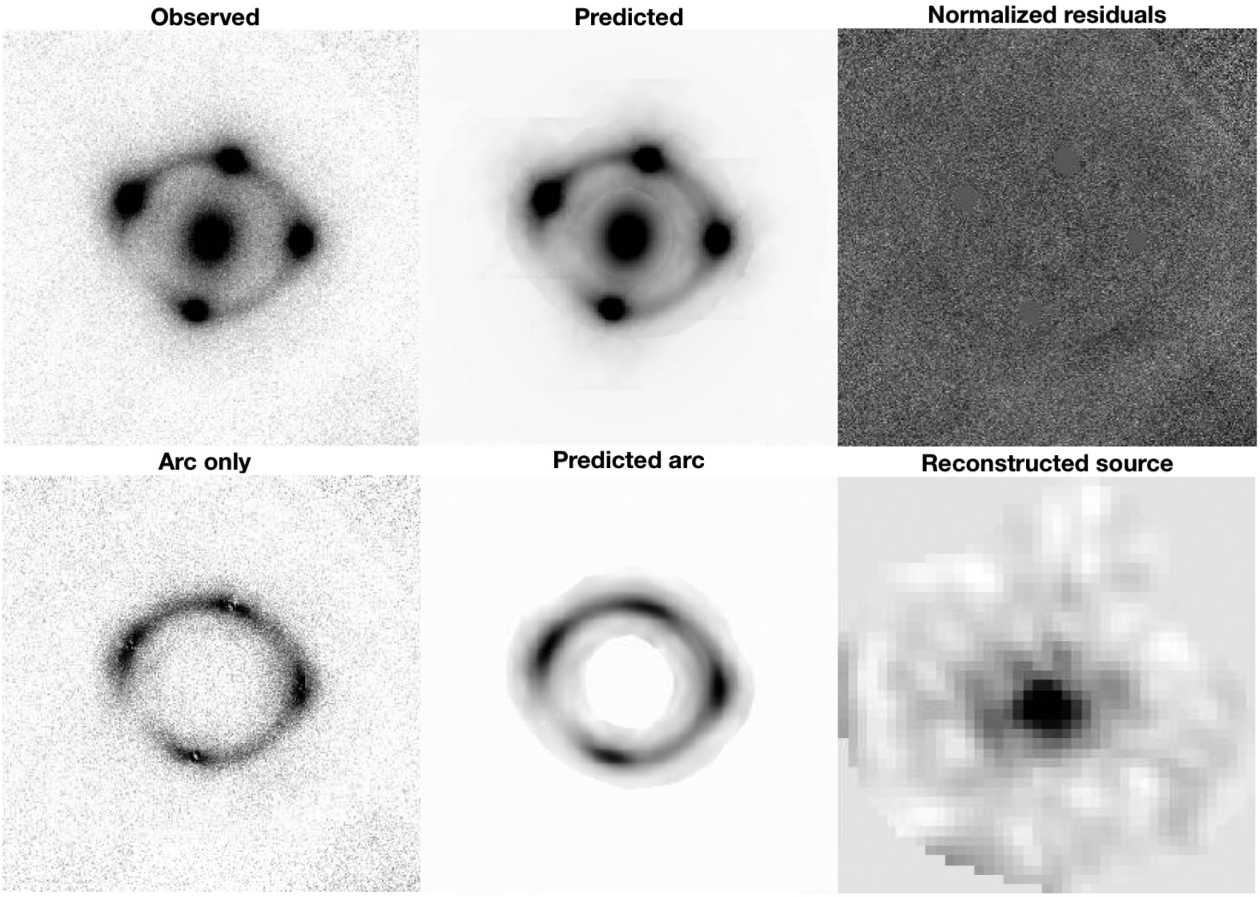


**Figure 11.** HE 0435−1223 AO image reconstruction of the most probable model with a source grid of 50 × 50 pixels and 69 × 69 pixels PSF for convolution of spatially extended images. *Top left*: HE 0435−1223 AO image. *Top middle*: Predicted lensed image of the background AGN host galaxy. *Top right*: Predicted light of the lensed AGNs and the lens galaxies. *Bottom left*: The arc-only image that removes the lens light and AGN light from the observed image. *Bottom middle*: Predicted lensed image of the background AGN host galaxy. *Bottom right*: The reconstructed host galaxy of the AGN in the source plane.

(OM) formulation[5] (Osipkov 1979; Merritt 1985)

$$\beta_{\rm ani} = \frac{r^2}{r_{\rm ani}^2 + r^2}, \tag{18}$$

where $r_{\rm ani} = 0$ is pure radial orbits and $r_{\rm ani} \to \infty$ is isotropic with equal radial and tangential velocity dispersions.

(iii) *A stellar component:* We assume a Hernquist profile[6] (Hernquist 1990)

$$\rho_* = \frac{I_0 a}{2\pi r(r+a)^3}, \tag{19}$$

for the power-law model, where $I_0$ is the normalization term and the scale radius can be related to the effective radius by $a = 0.551 r_{\rm eff}$.

[5]We further tested an additional anisotropy model, namely, a two-parameter extension of OM, and found that the uncertainty on the inferred $D_\ell$ due to anisotropy models is comparable to that from the choice of different mass models (i.e. power-law versus composite models). Thus, we should not be underestimating the uncertainty on $D_\ell$ (Jee et al. 2019). In addition, different anisotropy models have negligible impact on the $D_{\Delta t}$ measurement.

[6]Suyu et al. (2010) showed that different types of stellar distribution functions, namely, Hernquist versus Jaffe, produced nearly identical PDFs for the cosmological parameters.

For the composite model, the stellar component is represented by the light profile multiplied by a constant mass-to-light ratio.

With the aforementioned three components, we follow Sonnenfeld et al. (2012) and calculate the 3D radial velocity dispersion by numerically integrating the solutions of the spherical Jeans equation (Binney & Tremaine 1987)

$$\frac{1}{\rho_*}\frac{{\rm d}\left(\rho_* \sigma_{\rm r}^2\right)}{{\rm d}r} + 2\frac{\beta_{\rm ani}\sigma_{\rm r}^2}{r} = -\frac{GM(r)}{r^2}, \tag{20}$$

given the $\kappa_{\rm ext}$ from Section 4.3.1. Note that since the LOS velocity dispersion has a degeneracy between its anisotropy and the mass profile (Dejonghe 1987), we marginalize the sample of $r_{\rm ani}$ over a uniform distribution $[0.5, 5]r_{\rm eff}$. To compare with the data, we can get the seeing-convolved luminosity-weighted LOS velocity dispersion

$$(\sigma^{\rm P})^2 = \frac{\int_{\mathcal{A}}[I(R)\sigma_{\rm s}^2 * \mathcal{P}]{\rm d}\mathcal{A}}{\int_{\mathcal{A}}[I(R) * \mathcal{P}]{\rm d}\mathcal{A}}, \tag{21}$$

where $R$ is the projected radius, $I(R)$ is the light distribution, $\mathcal{P}$ is the PSF convolution kernel (Mamon & Łokas 2005), and $\mathcal{A}$ is the aperture. The luminosity-weighted LOS velocity dispersion is given

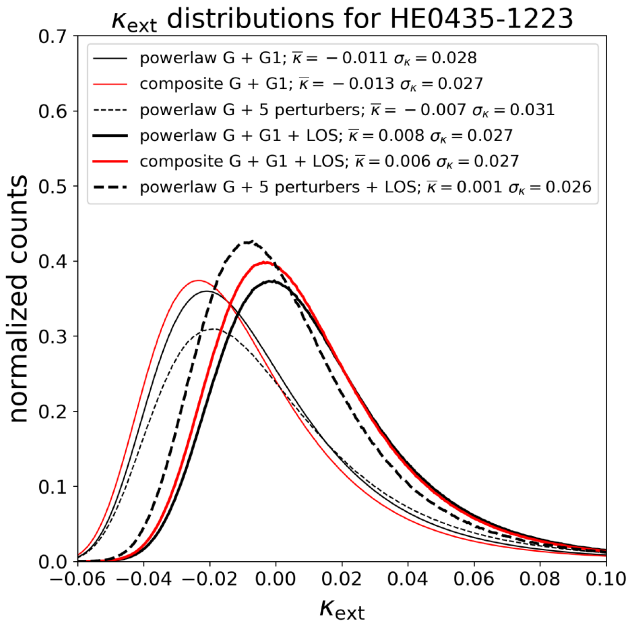


**Figure 12.** Distributions of $\kappa_{\rm ext}$ for HE 0435−1223, for three different lens models and their associated shear values. The constraints used to produce the distributions from the Millennium Simulation are either only the external shear $\gamma$ or the shear plus the combination of weighted counts corresponding to galaxy number counts inside the 45 arcsec aperture, as well as number counts weighted by the inverse of the distance of each galaxy to the lens (the '+LOS' models). The numerical constraints are reported in Table C1. For the `powerlaw + G1` and `composite + G1` models, we used an inner mask of 5 arcsec around the lens, and for the `powerlaw + 5 perturbers` model, we used a mask of 12 arcsec radius. The five perturbers are indicated in fig. 3 from Wong et al. (2017). Three additional galaxies enter the 12 arcsec radius inner mask, and we slightly boosted their distance from the lens, in order to avoid masking them. See caption of Fig. 4 for additional details.

by

$$I(R)\sigma_{\rm s}^2 = 2\int_R^\infty \left(1-\beta_{\rm ani}\frac{R^2}{r^2}\right)\frac{\rho_*\sigma_{\rm r}^2 r\,{\rm d}r}{\sqrt{r^2-R^2}}. \quad (22)$$

For a system with significant perturbers at a different redshift from the main lens (e.g. HE 0435−1223), we assume a flat ΛCDM cosmology with $H_0$ uniform in [0, 150] km s$^{-1}$ Mpc$^{-1}$, $\Omega_{\rm m} = 0.3$, and $\Omega_{\rm m} = 1 - \Omega_\Lambda$ to calculate the critical density and rank the models by the predicted velocity dispersion. Note that this assumption does not affect the generality of the conclusion. For our single lens plane systems (i.e. RXJ 1131−1231 and PG 1115+080), we use the measured velocity dispersion to constrain $D_{\rm s}/D_{\ell{\rm s}}$ and then combine with the measurement of $D_{\Delta t}$ to infer the value of $D_\ell$ without assuming any cosmological model (Birrer et al. 2016, 2019). The further advantage of this method is that $D_\ell$ is not affected by $\kappa_{\rm ext}$ (Jee et al. 2015).

### 4.4 Microlensing time-delay prediction models

In Section 2, we showed that the time delays between multiple images are due to the geometry and the gravitational potential that the light passes through. Tie & Kochanek (2018) introduce a possible new microlensing effect on the time delays that can shift the light curves depending on the structure of the accretion disc in the lensed quasar and the density of the stars in the lensing galaxy. They estimated this effect under the assumption of a lamp-post model for the accretion disc, where a large part of the disc lights up concurrently on light-travel scales that are a significant fraction of the time delays between the lensed images. The observed time delay could thus be affected by differential magnification of the individual images, resulting from microlensing by stars in the lensing galaxy. However, the lamp-post model is only one choice for how to represent the accretion disc; other accretion disc models for which variability is different from the lamp-post model are possible (e.g. Dexter & Agol 2011). We follow Chen et al. (2018a) and present the $D_{\Delta t}$ measurements both with and without the lamp-post assumption. However, we only consider the case without the microlensing effect in our final $H_0$ determination since it is not clear at this point which is the proper disc model to use. Note that it also was not applied in previous H0LiCOW work to infer the final $H_0$ measurement.

A more detailed description of this effect and how to estimate the probability distribution of the microlensing time-delay effect (MTDE) can be found in Bonvin et al. (2018). We briefly summarize the technique here. We generate magnification maps using GPU-D (Vernardos & Fluke 2014), which incorporates a graphics processing unit implementation of the inverse ray-shooting technique (Kayser, Refsdal & Stabell 1986). All magnification maps have dimension of 8192 × 8192 pixels over a scale of $20\langle R_{\rm E}\rangle$, where

$$\langle R_{\rm E}\rangle = \sqrt{\frac{D_{\rm s}D_{\ell{\rm s}}}{D_\ell}\frac{4G\langle M_\star\rangle}{c^2}}. \quad (23)$$

We choose the Salpeter initial mass function with mean mass $\langle M_\star\rangle = 0.3{\rm M}_\odot$ and the ratio between the upper and lower masses $M_{\rm upper}/M_{\rm lower} = 100$ (Kochanek 2004). We consider a standard thin disc model (Shakura & Sunyaev 1973). Given the disc size of lensed quasar, the average microlensing time delay at each position on a magnification map can be derived using equation (10) of Tie & Kochanek (2018). The parameters that are used to estimate the probability distribution of the microlensing time delay for each system are listed in Tables 2 and 3.

To fold this effect into time-delay modelling, we use

$$\Delta t_{ij} = (D_{\Delta t}/c)\Delta\tau_{ij} + (t_i - t_j), \quad (24)$$

where the first term on the right-hand side is the same as in equation (2) and $t_i - t_j$ is the extra delay caused by the MTDE between images $i$ and $j$ (see details in Chen et al. 2018a).

## 5 LENS MODELLING

Both HE 0435−1223 and RXJ 1131−1231 have been extensively modelled using the extended lensed emission seen in high-resolution *HST* imaging of the systems (Suyu et al. 2014; Wong et al. 2017), but our modelling techniques have not yet been applied to PG 1115+080. Therefore, in this section we begin with a description of the modelling of PG 1115+080 in Section 5.1, and then describe HE 0435−1223 in Section 5.2 and RXJ 1131−1231 in Section 5.3.

For PG 1115+080, we model the *HST* and AO imaging simultaneously. However, for HE 0435−1223 and RXJ 1131−1231, we only model the AO imaging since the *HST* imaging has already been modelled (Suyu et al. 2014; Wong et al. 2017), and then combine the two modelling outputs to obtain a joint inference on $H_0$ (see Section 6.2).

Our analyses of PG 1115+080 and HE 0435−1223 are blind, as in Suyu et al. (2013) and Rusu et al. (2019a), in order to avoid

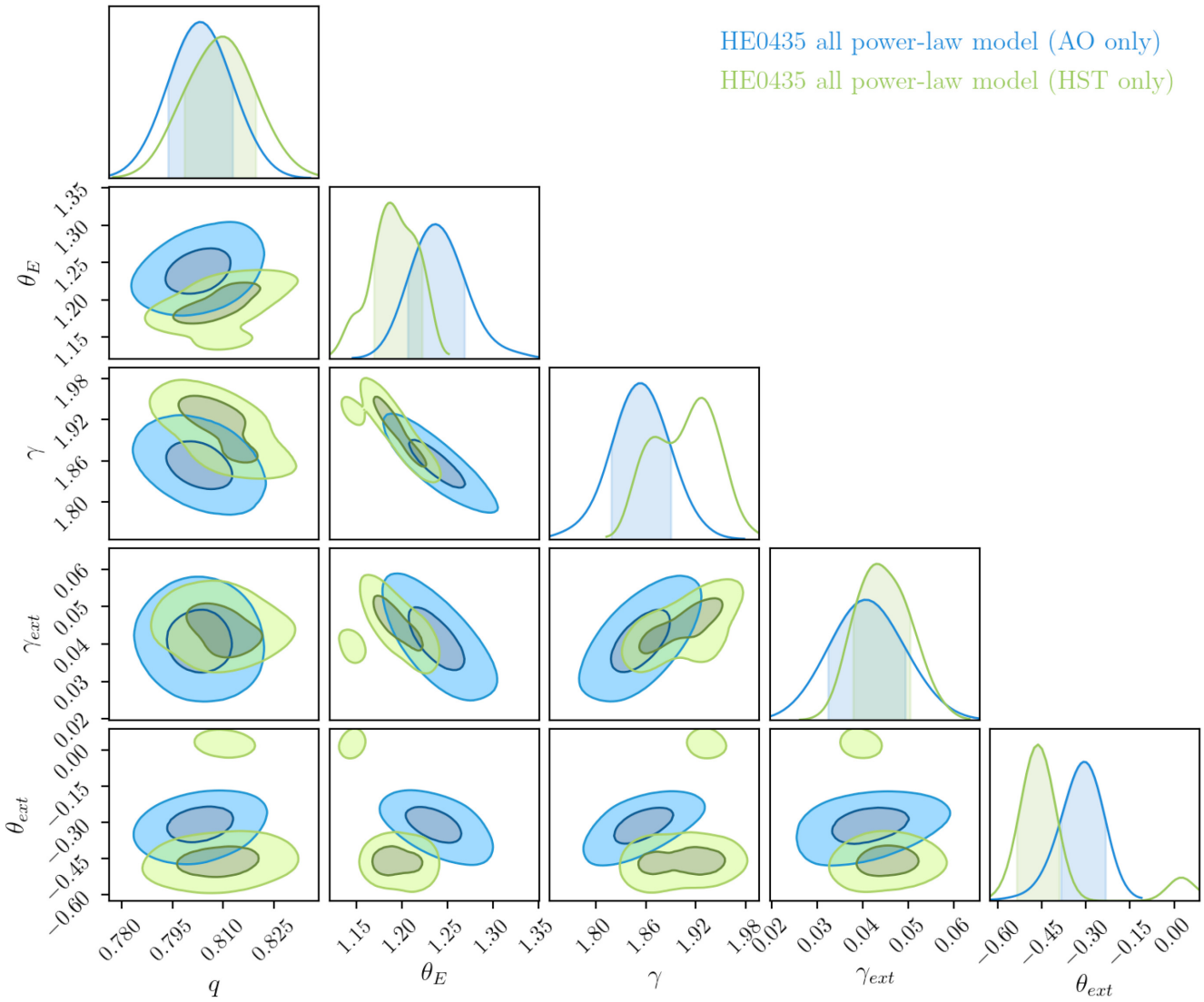


**Figure 13.** Marginalized parameter distributions from the power-law lens model results for HE 0435−1223. We show the comparison between using only AO imaging data and using only *HST* imaging data. The contours represent the 68.3 and 95.4 per cent quantiles.

confirmation bias. That is, the values of $D_{\Delta t}$, $D_\ell$ (if computed), and $H_0$ were kept blind until all co-authors came to a consensus to reveal the values during a collaboration telecon on June 5. The analysis was frozen after we unblinded the results and no changes were made to any of the numerical results. The time between unblinding and submission was used to polish the text and figures of the manuscript, and carrying out the detailed comparison of the AO- and *HST*-based analysis.

In contrast, the RXJ 1131−1231 analysis was not done blindly as the AO data for this system were used to develop the PSF reconstruction technique. On top of the power-law model we have done in Chen et al. (2016), we further test the composite model and use both models to infer $D_{\Delta t}$ and $D_\ell$.

To better control the systematics due to the choice of lens modelling technique, we run Markov chain Monte Carlo (MCMC) sampling with different source resolutions in each model. This approach was used because Suyu et al. (2013) have shown that the effects of the pixelated-source grid resolution dominate the uncertainty on the lens modelling when using a modelling code, such as GLEE, that implements the pixelated-source reconstruction technique.

### 5.1 PG 1115+080 modelling

PG 1115+080 is a single-plane lens system embedded in a nearby group that consists of 13 known galaxies (see the solid circles in Fig. 2). If we model the lens without including the group, the mass profile shows a very steep slope ($\gamma \sim 2.35$; note that $\gamma = 2$ corresponds to the isothermal profile), which has also been found in previous studies of this system (e.g. Keeton & Kochanek 1997; Treu & Koopmans 2002), and a strong shear ($\gamma_{\rm ext} \sim 0.15$), which comes from the nearby group. Wilson et al. (2016) showed that, compared with the other 11 groups along the light of sight, the nearby group contributes the largest convergence at the lens position. Furthermore, McCully et al. (2017) indicate that the group produces a significant flexion shift. Thus, it is crucial to model not only the main lens but also the group explicitly if we want to obtain an unbiased $H_0$ measurement.

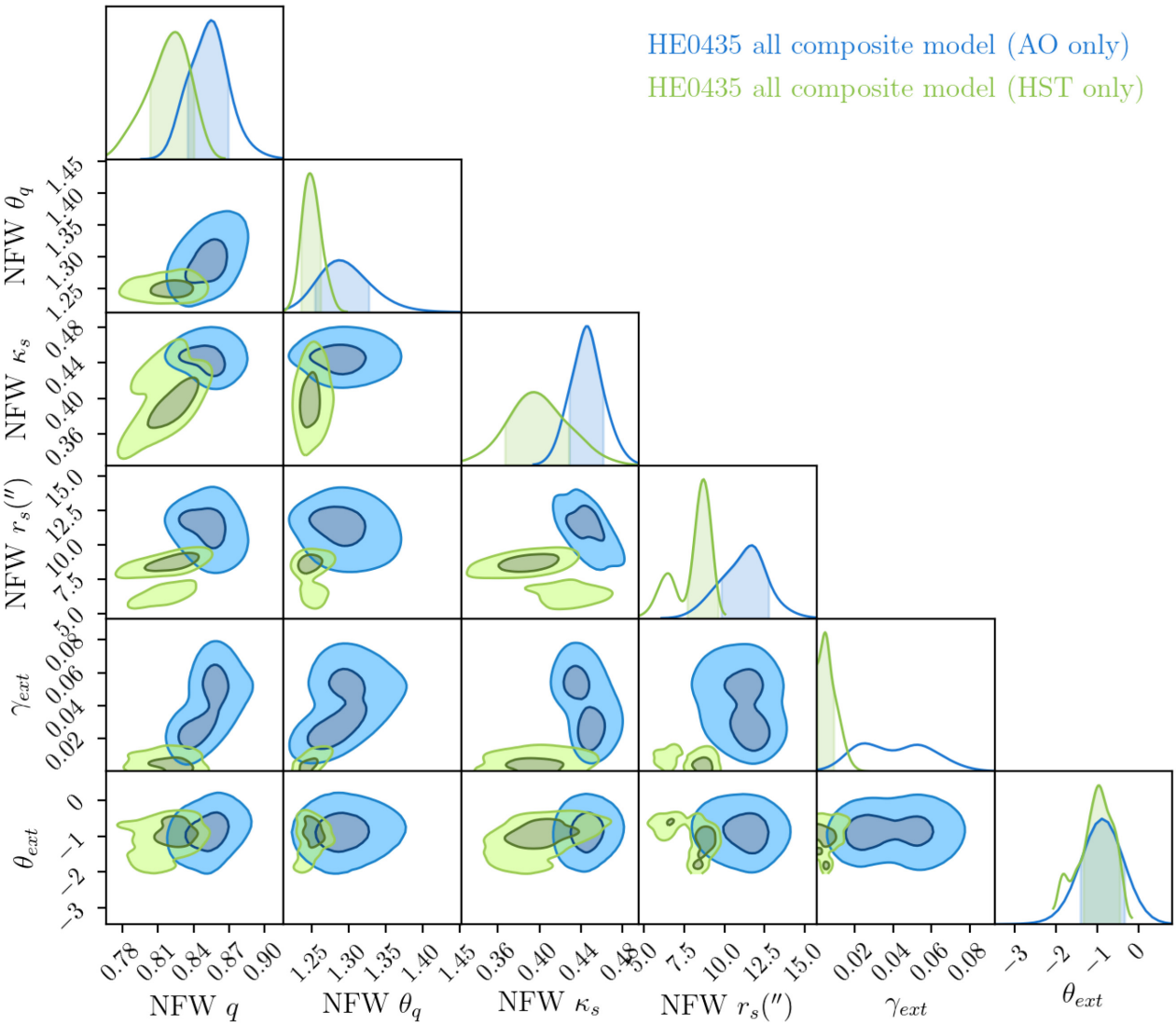


**Figure 14.** Marginalized parameter distributions from the composite lens model results for HE 0435−1223. We show the comparison between using only AO imaging data and using only *HST* imaging data. The contours represent the 68.3 and 95.4 per cent quantiles.

#### 5.1.1 *The PSF of PG 1115*

For the *HST* imaging, we use TINYTIM (Krist & Hook 1997) to generate the PSFs with different spectral index, $\alpha$, of a power law from −0.4 to −2.5 and different focuses[7] from 0 to 10. We find that the best fit is the PSF with focus equal to 0 and spectral index equal to −1.6. We use this TINYTIM PSF as the initial guess and then apply the PSF correction method while modelling the *HST* imaging. For the AO imaging, we follow the criteria described in section 4.4.3 in Chen et al. (2016) and perform eight iterative steps to create the final PSF and make sure the size of the PSF for convolution is large enough so that the results are stable. The FWHM of the AO PSF is 0.07 arcsec, while the FWHM of the *HST* PSF is 0.15 arcsec. We show the reconstructed AO PSF in Fig. A1.

[7]The flux per unit frequency interval is $F\nu = C\nu^{\alpha}$, where $\alpha$ is the power-law index and $C$ is a constant; focus is related to the breathing of the second mirror, which is between 0 and 10.

#### 5.1.2 *Main lens*

We follow two approaches to modelling the mass distribution in the main lensing galaxy.

(i) *SPEMD + 2Sérsic + shear:* We first choose the SPEMD density profile to model the extended arc and reconstruct the source structure on a pixelated grid (Suyu et al. 2006). We found that a single Sérsic profile is not sufficient to describe the light distribution, so we model it with two concentric elliptical Sérsic profiles with free relative position angles and ellipticities. By comparing the mass and light components, we found a similar result to Yoo et al. (2005), namely that the position of the centre of mass is very close to the centre of light, with $|\Delta r| \approx 0.015$ arcsec. This implies that the offset between the projected centre of dark matter and baryonic matter is small.

(ii) *Composite + 2Chameleon + shear:* We also model the main lens with composite model. Because the SPEMD + 2Sérsic + shear model indicated that the dark matter and baryonic centroids were consistent, we link the centroid of NFW profile to the centroid of two

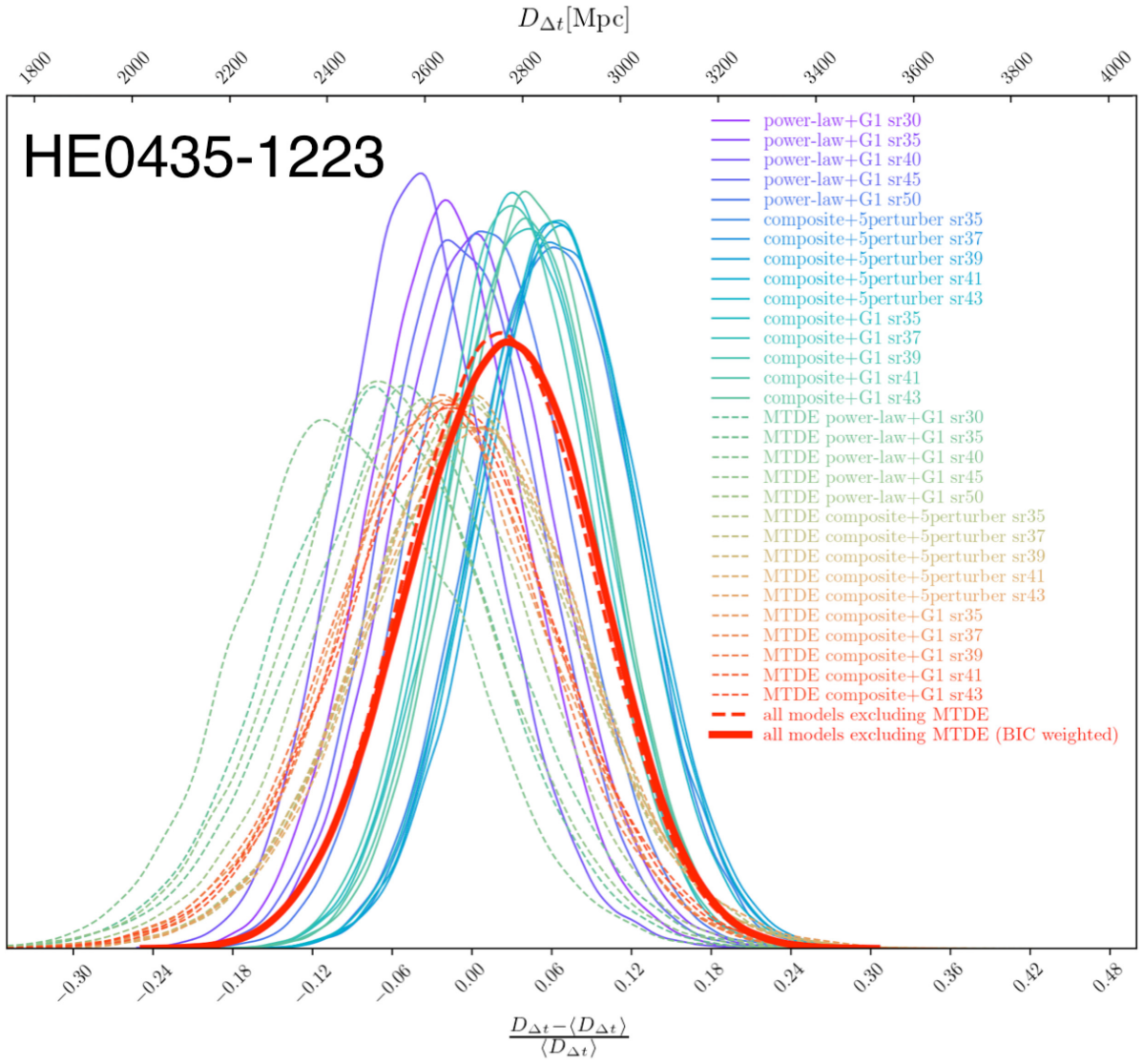


**Figure 15.** The various model choices of HE 0435−1223 with and without considering MTDE.

concentric chameleon profiles. We also follow Wong et al. (2017) and Rusu et al. (2019a) to iteratively update the relative amplitudes of the associated mass components to match those of the light components, as the relative amplitudes of light components can vary when we run the MCMC chains, while the relative amplitudes are fixed in the mass profiles.

### *5.1.3 Nearby group*

Based on the velocity dispersion of the nearby group, $\sigma_{\rm group} = 390 \pm 60$ km s$^{-1}$, the inferred group mass is around $10^{13}$–$10^{14}$ h$^{-1}$ M$_\odot$ (Wilson et al. 2016). Oguri (2006) has shown that in this mass regime the mass profile is too complicated to be described by either a simple NFW profile or SIS profile, as it is a transition between the two. Thus, we use an NFW profile as the fiducial model, but also model the group with an SIS profile as a systematic check. In the following, we show how we determine reasonable priors on the NFW and SIS profiles.

(i) *Group (NFW):* We follow Wong et al. (2011) and use $M_{\rm vir}$–$c_{\rm vir}$ relationships based on WMAP5 results in Macciò, Dutton & van den Bosch (2008) to translate the observed velocity dispersion to scale radius and the normalization of the NFW profile (we compared the priors by assuming WMAP1 and WMAP3 and found that the difference is negligible, so our results are robust to variations in assumed $M_{\rm vir}$–$c_{\rm vir}$ relation). Here, we briefly recap the process. We use the measured velocity dispersion and its uncertainties, assuming it is a Gaussian distribution, to get a probability distribution for the group virial mass, $M_{\rm vir}$. Then, we can obtain the concentration, $c_{\rm vir}$, from the $M_{\rm vir}$–$c_{\rm vir}$ relationship assuming a reasonable scatter of 0.14 in $\log c_{\rm vir}$ (Bullock et al. 2001; Wong et al. 2011). With the critical density and the characteristic overdensity at the lens redshift (Eke, Navarro & Frenk 1998; Eke, Navarro & Steinmetz 2001), we can obtain $r_{\rm vir}$ and a prior probability on the scale radius, $r_{\rm s}$, via $c_{\rm vir} = r_{\rm vir}/r_{\rm s}$. The prior probability distribution of the normalization can be calculated by combining $r_{\rm s}$, the central density of the halo ($\rho_0$), and the critical surface density for lensing ($\Sigma_{\rm cr}$).

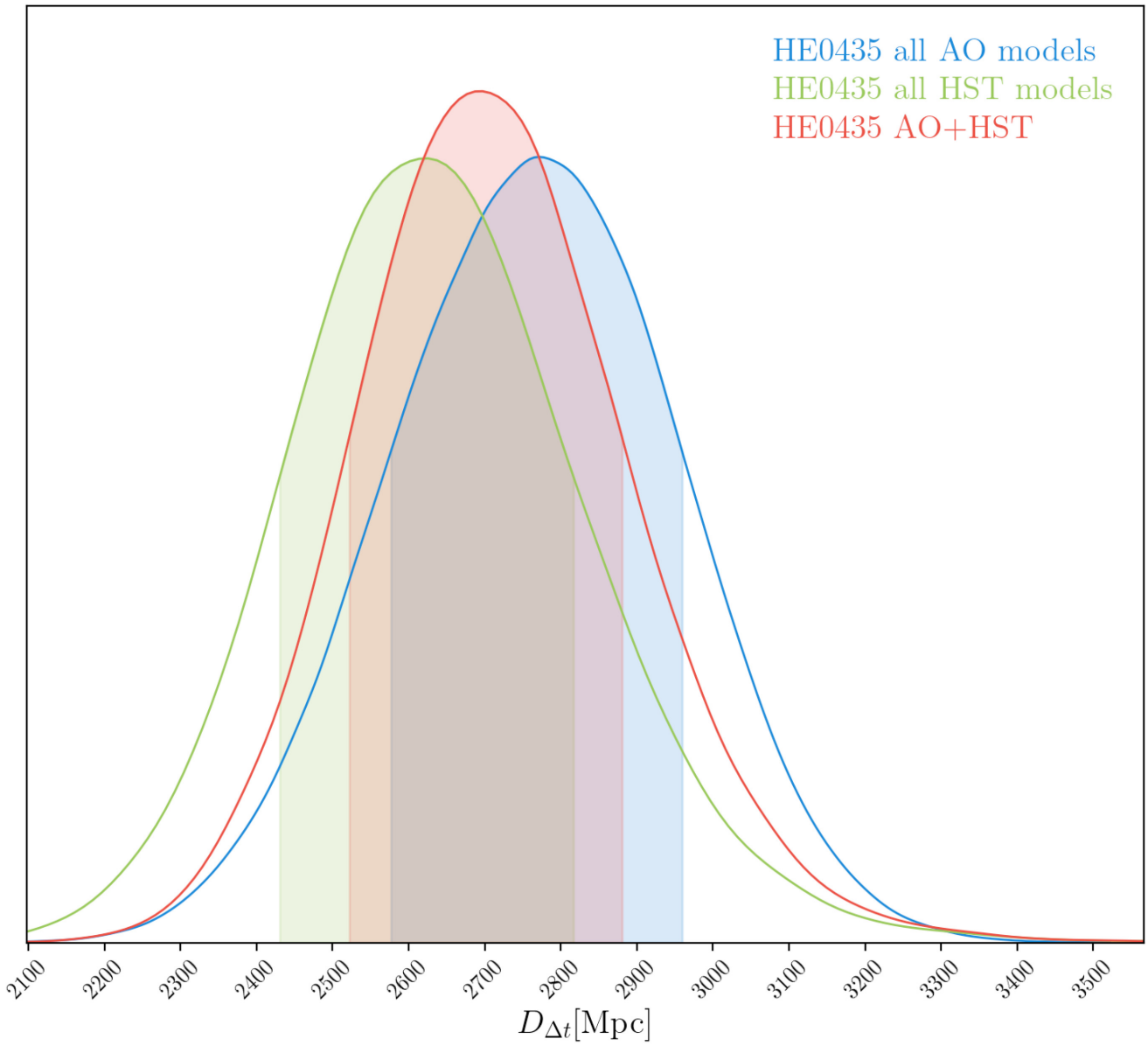


**Figure 16.** Constraints on $D_{\Delta t}$ from HE 0435−1223 when using AO imaging data only, *HST* imaging data only, or AO + *HST*. Because of the significant multiplane lensing needed for this system, we cannot compute a $D_\ell$.

(ii) *Group (SIS):* We convert the velocity dispersion to an Einstein radius via

$$\sigma^2 = \theta_{\rm E} \frac{c^2}{4\pi} \frac{D_{\rm s}}{D_{\ell {\rm s}}}, \tag{25}$$

to get a prior on $\theta_{\rm E}$ of $1.4 \pm 0.2$ arcsec for G1 and $0.4 \pm 0.4$ arcsec for G2.

For these two models, we also put a prior on the position of the group (see Fig. 2) based on Wilson et al. (2016).

### 5.1.4 Nearby perturbing galaxies: G1 and G2

As some of the galaxies inside the group are close to the main lens, these perturbers could individually affect the main lens beyond the second-order distortion terms. We calculate $\Delta_3 x$ of the nearby galaxies using the notation and definition in McCully et al. (2017). As $\Delta_3 x$ is expressed in terms of the Einstein radius of these perturbers, we convert the measured velocity dispersions of G1 and G2 ($250 \pm 20$ and $130 \pm 60$ km s$^{-1}$, respectively, from Tonry 1998) into corresponding Einstein radii using equation (25). For the other galaxies lacking a measurement of the velocity dispersion, we assume that they are located at the group redshift (this assumption maximizes the value of $\Delta_3 x$), and use their relative luminosities compared to either G1 or G2 (depending on the morphology, since G1 is a spiral), to infer a velocity dispersion from the Faber–Jackson relation (Faber & Jackson 1976). We find $\log \Delta_3 x({\rm G1}) = -3.68^{+0.13}_{-0.14}$ (in units of log(arcsec)) and $\log \Delta_3 x({\rm G2}) = -4.01^{+0.75}_{-1.07}$, whereas the remaining galaxies have $\log \Delta_3 x < -4$, and we therefore neglect them. Thus, to test for systematic effects, we model the group either as a single group profile or as a group halo plus

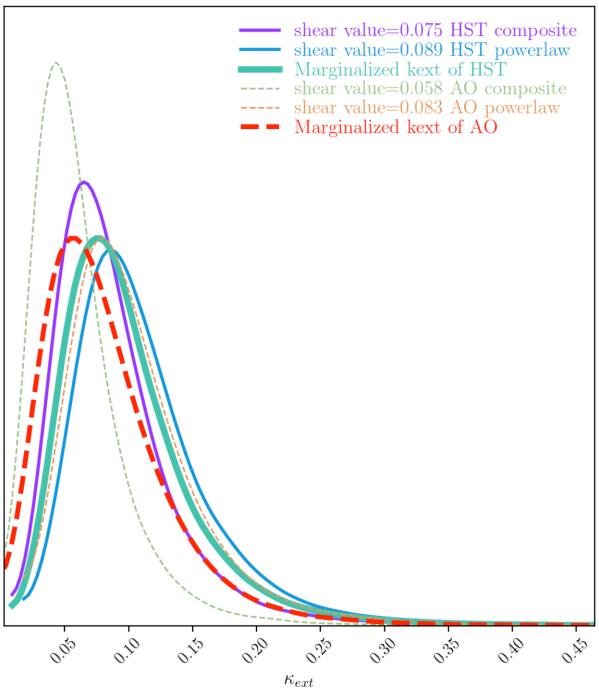


**Figure 17.** The comparison of the $\kappa_{\rm ext}$ for RXJ 1131−1231, based on the number counts and the shear values inferred from the AO imaging and *HST* imaging. Note that the high $\kappa_{\rm ext}$ value is because RXJ 1131−1231 is dominated by the very large shear constraint (see more details in Suyu et al. 2013).

either G1, or both G1 and G2, where the galaxies are modelled as SIS mass distributions.

### 5.1.5 *LOS analysis and the external convergence*

The technique of inferring $P(\kappa_{\rm ext}|\boldsymbol{d}_{\rm ENV})$, based on the Millennium Simulation and observed weighted galaxy number counts, was briefly described in Section 4.3.1. As implemented by Rusu et al. (2017), it requires wide-field, broad-band images to compute photometric redshifts and other physical properties of the galaxies surrounding the lens. The deepest multiband images currently available are provided by the Sloan Digital Sky Survey (SDSS; Adelman-McCarthy et al. 2008) and the Panoramic Survey Telescope and Rapid Response System (Chambers et al. 2016), but these are still relatively shallow, which may result in a biased $\kappa_{\rm ext}$ (Collett et al. 2013). For our $\kappa_{\rm ext}$ analysis for PG 1115+080, we therefore use the deep co-added data set from the MPIA 2.2 m telescope described in Section 3.4.

The co-added image has a limiting magnitude of 25.36 ± 0.08, deeper than the control survey, CFHTLenS ($r = 24.88 \pm 0.16$). We perform source detection in this image using SExtractor (Bertin & Arnouts 1996). For a fair comparison with the control survey, we need to convert our $R_{\rm c}$ magnitudes to $r$-band magnitudes. However, as we only have a single band, we cannot compute colour terms. Fortunately, a cross-match of the detections in our field with those in SDSS[8] shows that, after correcting for the zero-point offset, the scatter is small, with an rms of ∼0.10 mag, which we add to the photometric error budget. However, we choose a brighter magnitude limit $r \leq 23$ mag, in order to be able to use the purely morphological galaxy–star classification of CFHTLenS (Hildebrandt et al. 2012), where objects down to $i < 23$ mag are classified based solely on the FLUX_RADIUS parameter measured by SExtractor, and because, roughly, $r - i \sim 0.5$ for our cross-matches with SDSS. Due to the good seeing of our data, we find a clear stellar locus that allows us to determine a good classification threshold for FLUX_RADIUS, using the methodology in Coupon et al. (2009). We show the 240 arcsec × 240 arcsec cut-out of the field of view (FOV) in Fig. 3, marking the sources detected down to our magnitude limit.

We compute relative weighted galaxy number counts in terms of simple counts (a weight of unity) as well as using as weight the inverse of the distance between each galaxy in the field and the lens (weighting by $1/r$). We do this inside both the 45 and 120 arcsec radius apertures, using the technique from Rusu et al. (2017) and the galaxy catalogue produced earlier, with the exception that we use $r$-band magnitudes for both the lens field and the CFHTLenS fields, whereas Rusu et al. (2017) used $i$-band magnitudes. When doing this, we ignore the galaxies confirmed as part of the galaxy group, as the group is explicitly incorporated in our lensing models, and we need to compute $\kappa_{\rm ext}$ without its contribution. In addition, we account for the galaxies that are expected to be part of the group, but are missed due to the spectroscopic incompleteness, as described in Appendix D. We report our results in Table C1. These numbers are mostly consistent with the unit value, indicating that, after removing the contribution of the galaxy group, the field around the lens is of average density. Finally, we compute $P(\kappa_{\rm ext}|\boldsymbol{d}_{\rm ENV}, \gamma)$ following the technique presented in Birrer et al. (2019), which combines the constraints from both apertures. The combination of apertures results in a tighter distribution, as shown by Rusu et al. (2019a). We show the resulting distributions, corresponding to the various tests of systematics from Section 5.1.6, in Fig. 4 and the summary table in Appendix C.

### 5.1.6 *Systematic tests and unblinding results*

We summarize the choices that we explore for the mass modelling, including the nearby group/galaxies. For each of the models, we set the weights for the regions containing the AGN images to zero and fix the mass centroid for G1 and G2 at the centre of its light distribution. When modelling the galaxy group as an NFW profile, we use the $M_{\rm vir}$–$c_{\rm vir}$ relation from Macciò et al. (2008), based on a WMAP5 cosmology (we found that the impact of using different cosmology is negligible).

(i) SPEMD + 2Sérsic + NFW group.
(ii) SPEMD + 2Sérsic + NFW group + G1.
(iii) SPEMD + 2Sérsic + SIS group + G1.
(iv) SPEMD + 2Sérsic + NFW group + G1 + G2.
(v) Composite lens + NFW group.
(vi) Composite lens + NFW group + G1.
(vii) Composite lens + SIS group + G1.
(viii) Composite lens + NFW group + G1 + G2.
(ix) For all of these models, we test five different source resolutions. See details in Appendix F.

To address the MST, we use the $P(\kappa_{\rm ext}|\boldsymbol{d}_{\rm ENV}, \gamma)$ from the previous section and importance sampling the measured velocity dispersion, $\sigma = 281 \pm 25$ km s$^{-1}$ (Tonry 1998), which was obtained inside a 1.0 arcsec$^2$ aperture with a seeing of 0.8 arcsec. When sampling $D_{\Delta t}$ with the time-delay measurements from Bonvin et al. (2018), we

[8]We ignore the negligible differences between the SDSS and CFHTLenS $r$-band filters.

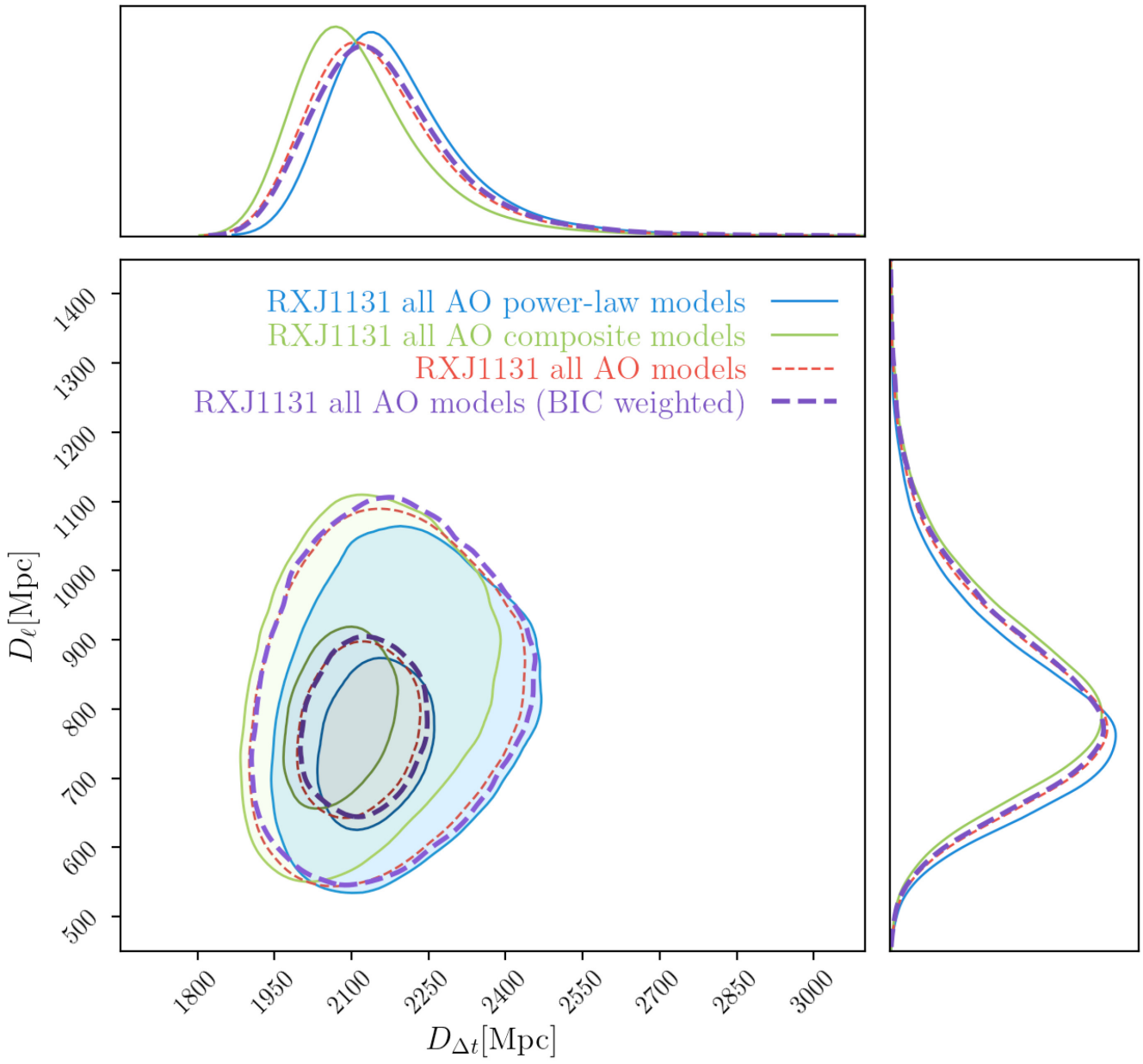


**Figure 18.** The $D_{\Delta t}$ and $D_\ell$ of RXJ 1131−1231 based on the analysis of the AO data. The power-law model and composite model show consistent results. We show the equally weighted marginalized results or BIC weighted marginalized results. The latter are used to do the cosmological inference.

follow Chen et al. (2018a) to sample the $D_{\Delta t}$ both with and without considering the MTDE. The parameters for estimating the star density and accretion disc model can be found in Chen et al. (2018a). In sum, we explore 160 modelling choices in total, with all different combination of choices among two kinds of main lens models, various mass models for the group, five different resolutions of the reconstructed source, and three different priors of the accretion disc sizes (or we ignore the MTDE). We show the AO imaging reconstruction in Fig. 5 and *HST* imaging reconstruction in Fig. 6. In Figs 7 and 8, we present the posteriors of the important parameters of AO-only results and *HST* + AO results for the composite model and the power-law model, respectively. The $D_{\Delta t}$ distributions from the 160 different model choices are shown in Fig. 9. The final $D_{\Delta t}$–$D_\ell$ without considering MTDE is shown in Fig. 10. We show the ΔBIC value of each model without considering MTDE in Appendix F. We found that the $D_{\Delta t}$ values inferred from various model choices are statistically consistent. The uncertainties of the final marginalized $D_{\Delta t}$ and $D_\ell$ are ∼9.6 and ∼29 per cent, respectively.

### 5.2 HE 0435−1223 modelling

For the HE 0435−1223 system, we model only the AO data, and then combine the results with the *HST*-based modelling of Wong et al. (2017). This lens system presents a bit of complexity because there are significant contributions to the lensing signal from galaxies at multiple redshifts. In particular, there are five important perturbers (G1–G5) that are close in projection to HE 0435−1223 (see fig. 3 in Wong et al. 2017). Based on the $\Delta_3 x$ criterion of McCully

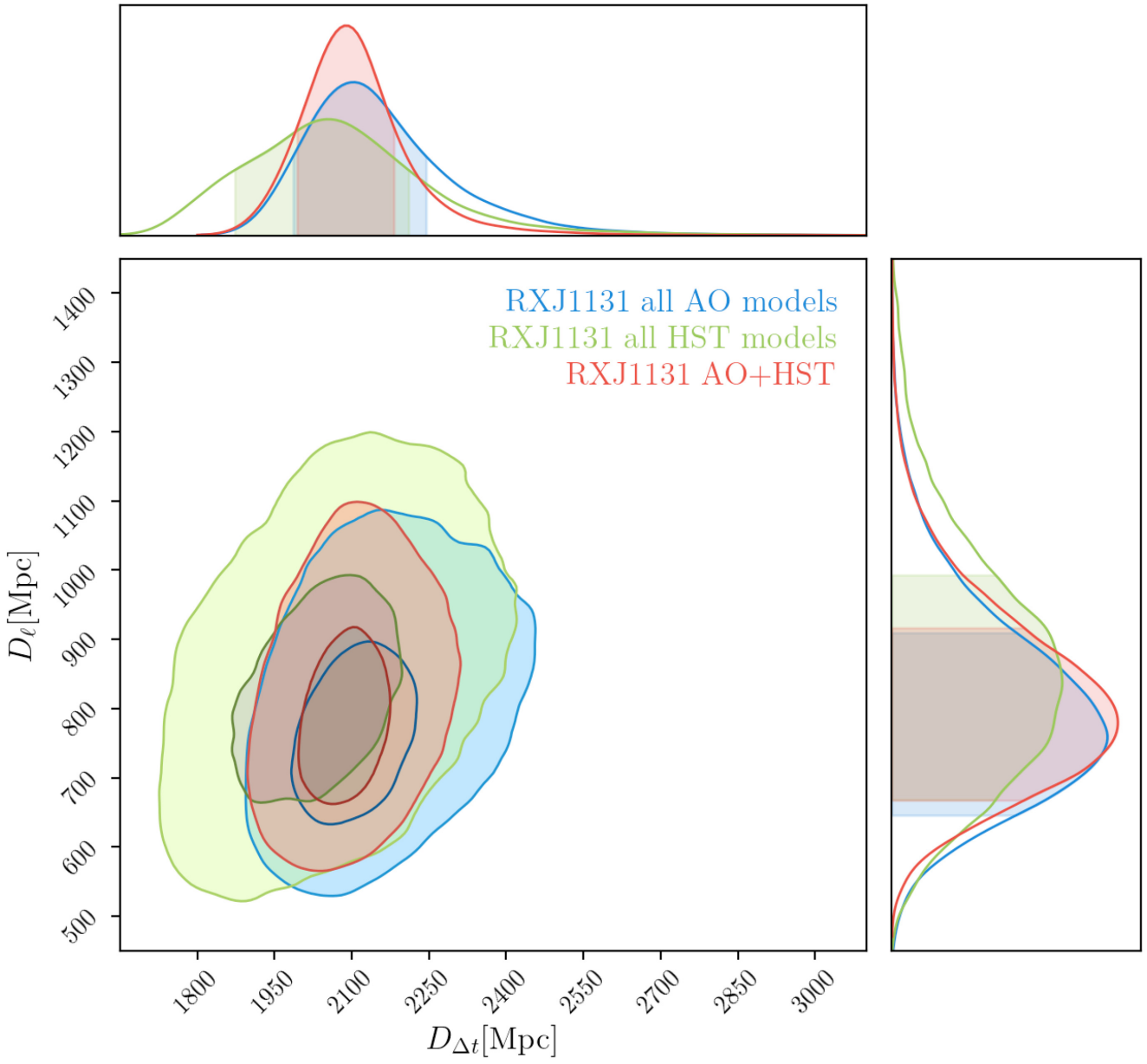


**Figure 19.** The $D_{\Delta t}$ and $D_\ell$ of RXJ 1131−1231 based on the combination of AO and *HST*.

et al. (2014, 2017), we should include the most massive nearby perturber, G1, explicitly in the model. However, Sluse et al. (2017) show that although the $\Delta_3 x$ values of the other four galaxies are not above the threshold when considered individually, when considered together they do show a significant effect. Since G1–G5 are located at different redshifts, we follow Wong et al. (2017) and model this system through the multiplane lens equation (e.g. Blandford & Narayan 1986; Schneider, Ehlers & Falco 1992; Collett & Auger 2014; McCully et al. 2014; Wong et al. 2017). In this case, there is no single time-delay distance, and therefore a particular cosmological model needs to be applied to the analysis. However, if the lens system is dominated by a single primary lens, as is the case for HE 0435−1223, then we can define an effective time-delay distance, $D_{\Delta t}^{\rm eff}(z_\ell, z_s)$, which is fairly robust to changes in the assumed cosmology.

### 5.2.1 *The AO PSF of HE 0435−1223*

We follow the same criteria described in Section 5.1.1 and perform 13 iterative correction steps to obtain the final AO PSF of HE 0435−1223. The FWHM of the reconstructed HE 0435−1223 AO PSF is 0.07 arcsec (see Fig. A1).

### 5.2.2 *Lens model choices*

As we did for PG 1115+080, we model the main lens with either an SPEMD or a composite model. For the composite model, we follow Wong et al. (2017) and set the Gaussian prior for the scale radius to 4.3 ± 2.0 arcsec based on scaling relations derived from the SLACS sample (Gavazzi et al. 2007). The most massive perturber, G1, is modelled as an SIS profile. When modelling G1–

G5 simultaneously as SIS distributions, we fix the ratios of their Einstein radii by estimating their stellar masses (Rusu et al. 2017) and then using Bernardi et al. (2011) to convert these to velocity dispersions and then to Einstein radii. We follow Wong et al. (2017) and fix the ratio of Einstein radii, but the global scaling is allowed to vary.

#### 5.2.3 LOS analysis and the external convergence

For this system, we have gathered wide-field imaging in a variety of filters, as well as conducting targeted spectroscopy (Rusu et al. 2017; Sluse et al. 2017). Our results on $P(\kappa_{\rm ext}|\boldsymbol{d}_{\rm ENV}, \gamma)$ are presented in Rusu et al. (2017). In this work, we use the shear values determined from our lens modelling of the AO data to update the weighted number counts for the system. Otherwise, we follow the analysis of Rusu et al. (2017) in order to conduct a direct comparison of $H_0$ from the *HST* and AO data sets.

#### 5.2.4 Systematic tests and unblinding results

We list the systematic tests we have done here. For each of the models, we set the weights for the regions containing the AGN images to zero and fix the mass centroid for G1 at the centre of its light distribution.

(i) A power-law model plus G1 as an SIS.
(ii) A composite model plus G1 as an SIS.
(iii) A composite model plus the five perturbers (G1–G5).
(iv) For all of these models, we test five different source resolutions. See details in Appendix G.

To assess the MST, for each model we perform the importance sampling given the measured velocity dispersion, $\sigma = 222 \pm 15\,{\rm km\,s^{-1}}$ inside a 0.54 arcsec × 0.7 arcsec aperture with a seeing of 0.8 arcsec (Wong et al. 2017). We show the AO imaging reconstruction in Fig. 11. The external convergence in Fig. 12. The comparisons between AO results and HST results in two mass models are shown in Fig. 13. and Fig. 14. We show the posteriors of $D_{\Delta t}$ in different model choices in Fig. 15 and the posteriors of joint $D_{\Delta t}$ in Fig. 16.

Note that for the baryonic component in the composite model, the mass distribution is based on the light distribution in the *HST* imaging. This is so because an insufficient knowledge about the structure in the wings of the AO PSF introduces a degeneracy between the reconstructed PSF structure and lens galaxy light (see the discussion in Appendix B).

### 5.3 RXJ 1131−1231 modelling

A detailed discussion of the RXJ 1131−1231 lens modelling of the AO imaging can be found in the paper of Chen et al. (2016). To summarize, we used the power-law mass distribution to model the lens potential and used two concentric Sérsic profiles to model the lens light. The satellite galaxy of the main deflector was modelled as an SIS profile. We modelled only the lensing galaxy plus satellite, and did not consider $\kappa_{\rm ext}$. The modelling marginalized over five different source resolutions in order to better control the systematics. In this paper, we further explore a different mass model and turn the previous work and the new results in this paper into cosmology.

#### 5.3.1 Main lens and satellite

To add to the previous power-law model, we test a composite model with different source resolutions in this paper. We follow Suyu et al. (2014) to set a Gaussian prior on the NFW scale radius of $18.6 \pm 2.6$ arcsec, based on the weak lensing analysis of SLACS lenses (Gavazzi et al. 2007) that have similar velocity dispersions to RXJ 1131−1231. For the other parameters, we set uniform priors. We model the satellite light distribution with a circular Sérsic profile, and the satellite mass as an SIS distribution whose centroid is linked to the light centroid.

Note that due to the degeneracy between the reconstructed PSF structure and lens galaxy light, the baryonic mass distribution in the composite model is also based on the light distribution in the *HST* imaging.

#### 5.3.2 LOS analysis and the external convergence

As in Suyu et al. (2013), we use a combination of observations and simulations to estimate the contribution of the LOS mass distribution for RXJ 1131−1231, i.e. $P(\kappa_{\rm ext})$, $P(\kappa_{\rm ext}|\gamma)$, and $P(\kappa_{\rm ext}|\boldsymbol{d}_{\rm ENV}, \gamma)$. Here, $\gamma$ is the external shear required by the mass models of the main lensing galaxy, while $\boldsymbol{d}_{\rm ENV}$ is the relative overdensity of galaxies within a 45 arcsec aperture that is centred on the lens. This overdensity, $\zeta_1^{45\,{\rm arcsec}} = 1.4 \pm 0.05$ (following the notation in Birrer et al. 2019), is calculated from galaxies with apparent *HST*/ACS F814W magnitudes $18.5 \leq m \leq 24.5$ in both the lens and control samples (Fassnacht et al. 2011). The overdensity and shear values are combined with the simulated lensing data based on the Millennium Simulation (Springel et al. 2005; Hilbert et al. 2009) together with the semi-analytic galaxy model of Henriques et al. (2015), to get the probability distributions for $\kappa_{\rm ext}$. We show the results in Fig. 17.

#### 5.3.3 Systematic tests

We list the systematic tests we have done including those done in our previous work. In all of the models, the regions near the AGN images are given zero weight.

(i) SPEMD + 2Sérsic lens model. We re-run the model since we did not link the satellite mass position to its light position in Chen et al. (2016).
(ii) A composite model.
(iii) For these models, we test five different source resolutions. See details in Appendix H.

We use the observed velocity dispersion, $323 \pm 20\,{\rm km\,s^{-1}}$ (Suyu et al. 2013), given the $\kappa_{\rm ext}$ in Section 5.3.2 to sample $D_{\Delta t}$ and $D_\ell$ without assuming cosmology. We plot the posteriors of $D_{\Delta t}$ and $D_\ell$ in Fig. 18 and the joint results in Fig. 19.

## 6 COSMOLOGICAL INFERENCE

We present the cosmological inferences based on the distance measurements (see Table 4) of the three gravitational lenses that have AO imaging data, *HST* imaging data, velocity dispersion measurements, LOS studies, and time-delay measurements. In particular, we present the cosmological inferences based on only the AO imaging data in Section 6.1, while we present the cosmological inferences based on a combination of both the AO and *HST* imaging in Section 6.2.

**Table 4.** The summary of the distance measurements from the three lenses.

| Lens names | AO-only $D_{\Delta t}$ (Mpc) | AO-only $D_\ell$ (Mpc) | AO + *HST* $D_{\Delta t}$ (Mpc) | AO + *HST* $D_\ell$ (Mpc) |
|---|---|---|---|---|
| RXJ 1131−1231 | $2120^{+140}_{-120}$ | $770^{+150}_{-120}$ | $2086^{+97}_{-91}$ | $780^{+140}_{-110}$ |
| HE 0435−1223 | $2770^{+190}_{-190}$ | – | $2700^{+180}_{-170}$ | – |
| PG 1115+080 | $1440^{+170}_{-150}$ | $600^{+190}_{-140}$ | $1460^{+140}_{-130}$ | $650^{+190}_{-130}$ |

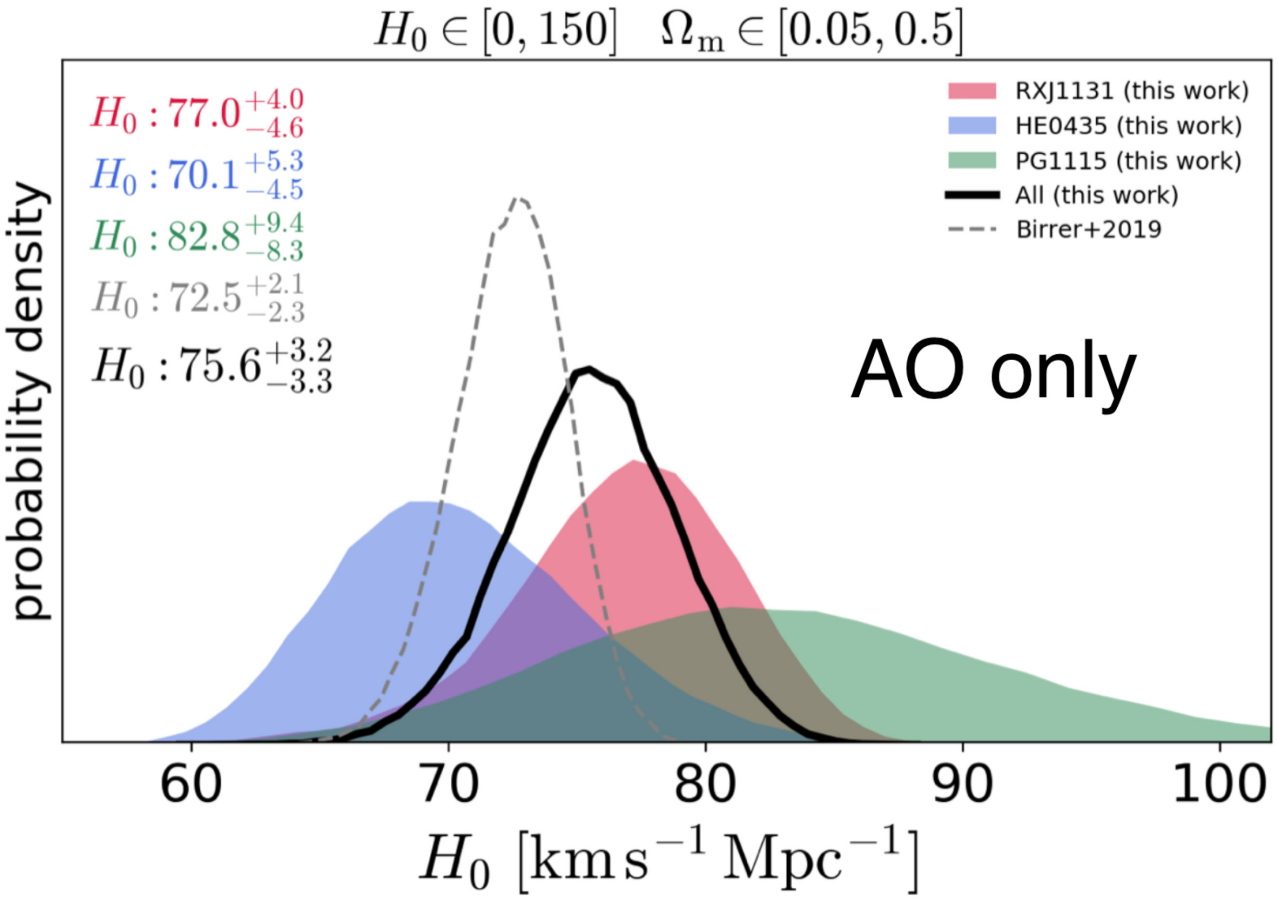


**Figure 20.** Marginalized posterior probability distributions for $H_0$ in the UΛCDM cosmology using the constraints from the three AO-only strong lenses (RXJ 1131−1231, HE 0435−1223, PG 1115+080). The overlaid histograms present the distributions for each individual strong lens, and the solid black line corresponds to the distribution resulting from the joint inference from all three data sets. The dashed line shows the latest joint $H_0$ from the H0LiCOW collaboration (Birrer et al. 2019). The quoted values of $H_0$ in the top-left corner of each panel are the median, 16th, and 84th percentiles.

### 6.1 Cosmological inference from AO strong lensing

Fig. 20 presents the marginalized posterior probability distribution function (PDF) for $H_0$ assuming a flat ΛCDM model that has a uniform prior on $H_0$ in the range [0, 150] km s$^{-1}$ Mpc$^{-1}$ and a uniform prior on $\Omega_m$ in the range of [0.05, 0.5]. After unblinding, we find that PG 1115+080 AO imaging yields $H_0 = 82.8^{+9.4}_{-8.3}$ km s$^{-1}$ Mpc$^{-1}$, HE 0435−1223 AO imaging yields $H_0 = 70.1^{+5.3}_{-4.5}$ km s$^{-1}$ Mpc$^{-1}$, and RXJ 1131−1231 AO imaging yields $H_0 = 77.0^{+4.0}_{-4.6}$ km s$^{-1}$ Mpc$^{-1}$. The joint analysis of three AO lenses yields $H_0 = 75.6^{+3.2}_{-3.3}$ km s$^{-1}$ Mpc$^{-1}$.

### 6.2 Cosmological inference from AO and *HST* strong lensing imaging

As the AO imaging of HE 0435−1223 and RXJ 1131−1231 is modelled separately from the *HST* imaging, we have developed a Bayesian approach to properly combine the *HST* and AO results for these two lenses (see details in Appendix E). In short, since the *HST* and AO images are independent data sets, we can get the joint probability distribution by multiplying their probability distributions, as long as the prior on the joint parameters is used

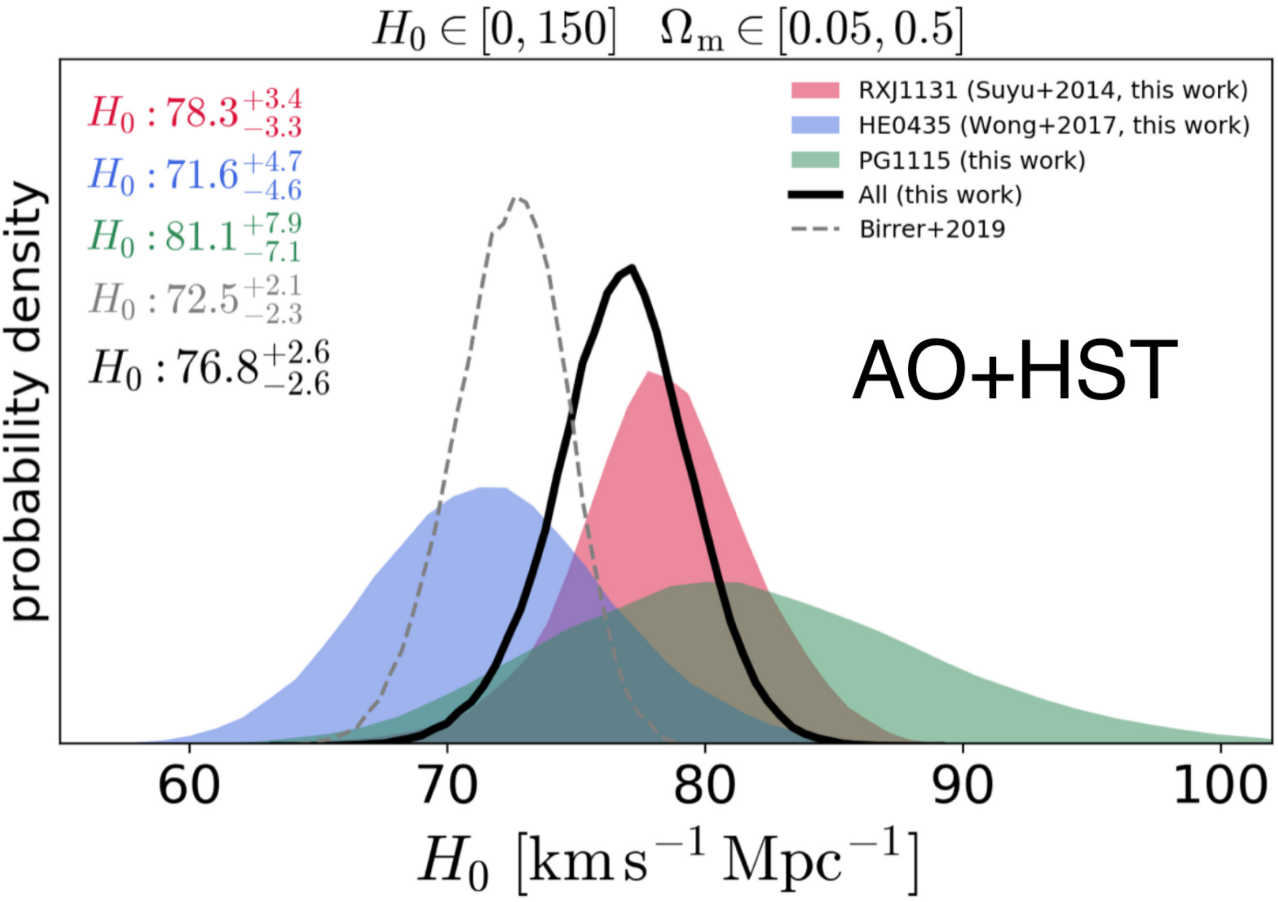


**Figure 21.** Marginalized posterior probability distributions for $H_0$ in the UΛCDM cosmology using the constraints from the three AO + *HST* strong lenses (RXJ 1131−1231, HE 0435−1223, PG 1115+080). The overlaid histograms present the distributions for each individual strong lens, and the solid black line corresponds to the distribution resulting from the joint inference from all three data sets. The dashed line shows the latest joint $H_0$ from the H0LiCOW collaboration (Birrer et al. 2019). The quoted values of $H_0$ in the top-left corner of each panel are the median, 16th, and 84th percentiles.

only once. We can express the joint posterior as

$$P(\boldsymbol{\eta}|\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H}, {\rm MS}) \propto \frac{P(\boldsymbol{\eta}|\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H}, {\rm MS}) P(\boldsymbol{\eta}|\boldsymbol{d}_{\rm AO}, {\rm H}, {\rm MS})}{P(\boldsymbol{\eta}|{\rm H}, {\rm MS})}, \quad (26)$$

where $P(\boldsymbol{\eta}|\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H}, {\rm MS})$ reflects the previous modelling results, $P(\boldsymbol{\eta}|\boldsymbol{d}_{\rm AO}, {\rm H}, {\rm MS})$ is obtained by this work, $P(\boldsymbol{\eta}|{\rm H}, {\rm MS})$ is the prior used in both analyses, H is the lens model, and MS is the Millennium Simulation.

Fig. 21 presents the marginalized posterior PDF for $H_0$ assuming flat ΛCDM model. We found that the joint AO + *HST* of HE 0435−1223 implies a value of the Hubble constant of $H_0 = 71.6^{+4.7}_{-4.6}$ km s$^{-1}$ Mpc$^{-1}$, the joint AO + *HST* of PG 1115+080 implies $H_0 = 81.1^{+7.9}_{-7.1}$ km s$^{-1}$ Mpc$^{-1}$, and the joint AO + *HST* of RXJ 1131−1231 implies $H_0 = 78.3^{+3.4}_{-3.3}$ km s$^{-1}$ Mpc$^{-1}$. The combination of three AO + *HST* lenses yields $H_0 = 76.8^{+2.6}_{-2.6}$ km s$^{-1}$ Mpc$^{-1}$.

We found that after combining AO and *HST* imaging, the dominant sources of uncertainty of both PG 1115+080 and HE 0435−1223 are time-delay measurements, while the dominant source of uncertainty of RXJ 1131−1231 is the LOS mass distribution.

## 7 CONCLUSIONS

We did the blind analysis on both PG 1115+080 and HE 0435−1223 as well as an extension of our previous analysis of RXJ 1131−1231. For each system, we combined the AO imaging, *HST* imaging, the measurements of the lens galaxy's velocity dispersion, the LOS studies from deep wide-area spectroscopic as well as photometric data, and the time-delay measurements from state-of-the-art light-curve fitting algorithm to infer the value of $H_0$. We find that the high S/N AO and *HST* imaging data yield consistent results, providing an important validation of the AO PSF reconstruction techniques for high-precision lensing work.

This paper demonstrates the ability of using AO imaging to constrain the mass model as well as the value of $H_0$. Furthermore, we show that combining AO imaging with *HST* imaging can further tighten the uncertainties from the lens mass model and thus improve the precision of the determination of $H_0$.

In this paper, we infer the value of $H_0$ under the assumption of a flat ΛCDM model that has a uniform prior on $H_0$ in the range [0, 150] km s$^{-1}$ Mpc$^{-1}$ and a uniform prior on $\Omega_m$ in the range of [0.05, 0.5]. After unblinding, PG 1115+080 AO imaging yields $H_0 = 82.8^{+9.4}_{-8.3}$ km s$^{-1}$ Mpc$^{-1}$, HE 0435−1223 AO imaging yields $H_0 = 70.1^{+5.3}_{-4.5}$ km s$^{-1}$ Mpc$^{-1}$, and RXJ 1131−1231 AO imaging

yields $H_0 = 77.0^{+4.0}_{-4.6}$ km s$^{-1}$ Mpc$^{-1}$. The joint analysis of three AO lenses yields $H_0 = 75.6^{+3.2}_{-3.3}$ km s$^{-1}$ Mpc$^{-1}$.

The joint AO + *HST* of PG 1115+080 yields $H_0 = 81.1^{+7.9}_{-7.1}$ km s$^{-1}$ Mpc$^{-1}$, the joint AO + *HST* of HE 0435−1223 yields $H_0 = 71.6^{+4.7}_{-4.6}$ km s$^{-1}$ Mpc$^{-1}$, and the joint AO + *HST* of RXJ 1131−1231 yields $H_0 = 78.3^{+3.4}_{-3.3}$ km s$^{-1}$ Mpc$^{-1}$. The combination of three AO + *HST* lenses yields $H_0 = 76.8^{+2.6}_{-2.6}$ km s$^{-1}$ Mpc$^{-1}$.

We refer the reader to the paper by Wong et al., where the results presented here will be combined with a self-consistent analysis of three previously published systems to carry out a full cosmological investigation.

## ACKNOWLEDGEMENTS

GC-FC acknowledges support from the Ministry of Education in Taiwan via Government Scholarship to Study Abroad (GSSA). CDF and GC-FC acknowledge support for this work from the National Science Foundation under grant no. AST-1715611. SH acknowledges support by the DFG cluster of excellence 'Origin and Structure of the Universe' (www.universe-cluster.de). LVEK acknowledges the support of an NWO-VIDI grant (no. 639.043.308).

This work was supported by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan. KCW was supported in part by an EACOA Fellowship awarded by the East Asia Core Observatories Association, which consists of the Academia Sinica Institute of Astronomy and Astrophysics, the National Astronomical Observatory of Japan, the National Astronomical Observatories of the Chinese Academy of Sciences, and the Korea Astronomy and Space Science Institute. SHS thanks the Max Planck Society for support through the Max Planck Research Group. JHHC, VB, and FC acknowledge support from the Swiss National Science Foundation (SNSF). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (COSMICLENS: grant agreement no. 787866). XD and TT acknowledge support from the Packard Foundation through a Packard Research Fellowship and from the NSF through grant AST-1714953. GC-FC thanks Jen-Wei Hsueh for the useful discussions and feedback. The data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

This work is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/IRFU, at the Canada–France–Hawaii Telescope (CFHT), which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. CFHTLenS data processing was made possible thanks to significant computing support from the NSERC Research Tools and Instruments grant program.

This work used computational and storage services associated with the Hoffman2 Shared Cluster provided by UCLA Institute for Digital Research and Education's Research Technology Group. Data analysis was in part carried out on common use data analysis computer systems at the Astronomy Data Center (ADC) of the National Astronomical Observatory of Japan.

## REFERENCES

Addison G. E., Huang Y., Watts D. J., Bennett C. L., Halpern M., Hinshaw G., Weiland J. L., 2016, ApJ, 818, 132
Addison G. E., Watts D. J., Bennett C. L., Halpern M., Hinshaw G., Weiland J. L., 2018, ApJ, 853, 119
Adelman-McCarthy J. K. et al., 2008, ApJS, 175, 297
Agnello A., 2017, MNRAS, 471, 2013
Agnello A. et al., 2018, MNRAS, 479, 4345
Agrawal P., Cyr-Racine F.-Y., Pinner D., Randall L., 2019, preprint (arXiv: 1904.01016)
Alam S. et al., 2017, MNRAS, 470, 2617
Anderson L. et al., 2014, MNRAS, 441, 24
Andreon S., Hurn M. A., 2010, MNRAS, 404, 1922
Aubourg É. et al., 2015, Phys. Rev. D, 92, 123516
Auger M. W., Treu T., Bolton A. S., Gavazzi R., Koopmans L. V. E., Marshall P. J., Moustakas L. A., Burles S., 2010, ApJ, 724, 511
Avestruz C., Li N., Zhu H., Lightman M., Collett T. E., Luo W., 2019, ApJ, 877, 58
Aylor K., Joy M., Knox L., Millea M., Raghunathan S., Kimmy Wu W. L., 2019, ApJ, 874, 4
Baade D. et al., 1999, Messenger, 95, 15
Barkana R., 1998, ApJ, 502, 531
Barnabè M., Czoske O., Koopmans L. V. E., Treu T., Bolton A. S., 2011, MNRAS, 415, 2215
Bernal J. L., Verde L., Riess A. G., 2016, J. Cosmol. Astropart. Phys., 10, 019
Bernardi M., Roche N., Shankar F., Sheth R. K., 2011, MNRAS, 412, 684
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Binney J., Tremaine S., 1987, Galactic Dynamics. Princeton Univ. Press, Princeton, NJ, p. 747
Birrer S., Amara A., Refregier A., 2015, ApJ, 813, 102
Birrer S., Amara A., Refregier A., 2016, J. Cosmol. Astropart. Phys., 8, 020
Birrer S. et al., 2019, MNRAS, 484, 4726
Blandford R., Narayan R., 1986, ApJ, 310, 568
Bonvin V. et al., 2017, MNRAS, 465, 4914
Bonvin V. et al., 2018, A&A, 616, A183
Bonvin V. et al., 2019, A&A, 629, A97
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Cappellari M. et al., 2013, MNRAS, 432, 1709
Cardona W., Kunz M., Pettorino V., 2017, J. Cosmol. Astropart. Phys., 2017, 056
Chambers K. C. et al., 2016, preprint (arXiv:1612.05560)
Chen G. C.-F. et al., 2016, MNRAS, 462, 3457
Chen G. C.-F. et al., 2018a, MNRAS, 481, 1115
Chen H.-Y., Fishbach M., Holz D. E., 2018b, Nature, 562, 545
Christian C. A., Crabtree D., Waddell P., 1987, ApJ, 312, 45
Collett T. E., Auger M. W., 2014, MNRAS, 443, 969
Collett T. E. et al., 2013, MNRAS, 432, 679
Coupon J. et al., 2009, A&A, 500, 981
Courbin F., Eigenbrod A., Vuissoz C., Meylan G., Magain P., 2005, in Mellier Y., Meylan G., eds, Proc. IAU Symp. 225, Gravitational Lensing Impact on Cosmology. Cambridge Univ. Press, Cambridge, p. 297
Courbin F. et al., 2011, A&A, 536, A53
Cuesta A. J., Verde L., Riess A., Jimenez R., 2015, MNRAS, 448, 3463
Cuesta A. J. et al., 2016, MNRAS, 457, 1770

Dai X., Kochanek C. S., Chartas G., Kozłowski S., Morgan C. W., Garmire G., Agol E., 2010, ApJ, 709, 278
de Grijs R., Courbin F., Martínez-Vázquez C. E., Monelli M., Oguri M., Suyu S. H., 2017, Space Sci. Rev., 212, 1743
Dejonghe H., 1987, MNRAS, 224, 13
Dexter J., Agol E., 2011, ApJ, 727, L24
Dhawan S., Jha S. W., Leibundgut B., 2018, A&A, 609, A72
Ding X. et al., 2018, preprint (arXiv:1801.01506)
Dobler G., Fassnacht C., Treu T., Marshall P. J., Liao K., Hojjati A., Linder E., Rumbaugh N., 2015, ApJ, 799, 168
Dutton A. A. et al., 2011, MNRAS, 417, 1621
Efstathiou G., 2003, MNRAS, 343, L95
Efstathiou G., 2014, MNRAS, 440, 1138
Eke V. R., Navarro J. F., Frenk C. S., 1998, ApJ, 503, 569
Eke V. R., Navarro J. F., Steinmetz M., 2001, ApJ, 554, 114
Faber S. M., Jackson R. E., 1976, ApJ, 204, 668
Falco E. E., Gorenstein M. V., Shapiro I. I., 1985, ApJ, 289, L1
Fassnacht C. D., Koopmans L. V. E., Wong K. C., 2011, MNRAS, 410, 2167
Feeney S. M., Mortlock D. J., Dalmasso N., 2018, MNRAS, 476, 3861
Fleury P., Clarkson C., Maartens R., 2017, J. Cosmol. Astropart. Phys., 2017, 062
Follin B., Knox L., 2018, MNRAS, 477, 4534
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
Golse G., Kneib J.-P., 2002, A&A, 390, 821
Greene Z. S. et al., 2013, ApJ, 768, 39
Grillo C. et al., 2018, ApJ, 860, 94
Heavens A., Jimenez R., Verde L., 2014, Phys. Rev. Lett., 113, 241302
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2663
Henry J. P., Heasley J. N., 1986, Nature, 321, 139
Hernquist L., 1990, ApJ, 356, 359
Hilbert S., Hartlap J., White S. D. M., Schneider P., 2009, A&A, 499, 31
Hildebrandt H. et al., 2012, MNRAS, 421, 2355
Hinshaw G. et al., 2013, ApJS, 208, 19
Hsueh J. W., Fassnacht C. D., Vegetti S., McKean J. P., Spingola C., Auger M. W., Koopmans L. V. E., Lagattuta D. J., 2016, MNRAS, 463, L51
Hsueh J.-W. et al., 2017, MNRAS, 469, 3713
Hsueh J.-W., Despali G., Vegetti S., Xu D. A., Fassnacht C. D., Metcalf R. B., 2018, MNRAS, 475, 2438
Jee I., Komatsu E., Suyu S. H., 2015, J. Cosmol. Astropart. Phys., 11, 033
Jee I., Komatsu E., Suyu S. H., Huterer D., 2016, J. Cosmol. Astropart. Phys., 4, 031
Jee I., Suyu S. H., Komatsu E., Fassnacht C. D., Hilbert S., Koopmans L. V. E., 2019, in press
Joseph R. et al., 2014, A&A, 566, A63
Joudaki S., Kaplinghat M., Keeley R., Kirkby D., 2018, Phys. Rev. D, 97, 123501
Kayser R., Refsdal S., Stabell R., 1986, A&A, 166, 36
Kazin E. A. et al., 2014, MNRAS, 441, 3524
Keenan R. C., Barger A. J., Cowie L. L., 2013, ApJ, 775, 62
Keeton C. R., Kochanek C. S., 1997, ApJ, 487, 42
Kenworthy W. D., Scolnic D., Riess A., 2019, ApJ, 875, 145
Kochanek C. S., 2004, ApJ, 605, 58
Komatsu E. et al., 2011, ApJS, 192, 18
Koopmans L. V. E., 2004, preprint (astro-ph/0412596)
Koopmans L. V. E., Treu T., Bolton A. S., Burles S., Moustakas L. A., 2006, ApJ, 649, 599
Koopmans L. V. E. et al., 2009, ApJ, 703, L51
Kreisch C. D., Cyr-Racine F.-Y., Doré O., 2019, preprint (arXiv:1902.00534)
Krist J. E., Hook R. N., 1997, in Casertano S., Jedrzejewski R., Keyes T., Stevens M., eds, The 1997 HST Calibration Workshop with a New Generation of Instruments. Space Telescope Science Institute, Baltimore, MD, p. 192
Lagattuta D. J., Auger M. W., Fassnacht C. D., 2010, ApJ, 716, L185
Lagattuta D. J., Vegetti S., Fassnacht C. D., Auger M. W., Koopmans L. V. E., McKean J. P., 2012, MNRAS, 424, 2800
Lanusse F., Ma Q., Li N., Collett T. E., Li C.-L., Ravanbakhsh S., Mandelbaum R., Póczos B., 2018, MNRAS, 473, 3895
Lemon C. A., Auger M. W., McMahon R. G., 2019, MNRAS, 483, 4242
Lemos P., Lee E., Efstathiou G., Gratton S., 2019, MNRAS, 483, 4803
Liao K. et al., 2015, ApJ, 800, 11
Lin H. et al., 2017, ApJ, 838, L15
Linder E. V., 2004, Phys. Rev. D, 70, 043534
Macaulay E. et al., 2019, MNRAS, 486, 2184
Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, MNRAS, 391, 1940
McCully C., Keeton C. R., Wong K. C., Zabludoff A. I., 2014, MNRAS, 443, 3631
McCully C., Keeton C. R., Wong K. C., Zabludoff A. I., 2017, ApJ, 836, 141
Mamon G. A., Łokas E. L., 2005, MNRAS, 363, 705
Merritt D., 1985, AJ, 90, 1027
Momcheva I., Williams K., Keeton C., Zabludoff A., 2006, ApJ, 641, 169
Momcheva I. G., Williams K. A., Cool R. J., Keeton C. R., Zabludoff A. I., 2015, ApJS, 219, 29
Moresco M., Jimenez R., Verde L., Cimatti A., Pozzetti L., Maraston C., Thomas D., 2016, J. Cosmol. Astropart. Phys., 2016, 039
Morgan N. D., Kochanek C. S., Pevunova O., Schechter P. L., 2005, AJ, 129, 2531
Morgan C. W., Kochanek C. S., Morgan N. D., Falco E. E., 2010, ApJ, 712, 1129
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Oguri M., 2006, MNRAS, 367, 1241
Oguri M., Marshall P. J., 2010, MNRAS, 405, 2579
Osipkov L. P., 1979, Pis'ma Astron. Zh., 5, 77
Ostrovski F. et al., 2017, MNRAS, 465, 4325
Ostrovski F. et al., 2018, MNRAS, 473, L116
Petrillo C. E. et al., 2017, MNRAS, 472, 1129
Planck Collaboration VI, 2018, preprint (arXiv:1807.06209)
Poulin V., Smith T. L., Karwal T., Kamionkowski M., 2019, Phys. Rev. Lett., 122, 221301
Rathna Kumar S. et al., 2013, A&A, 557, A44
Refsdal S., 1964, MNRAS, 128, 307
Riess A. G. et al., 2016, ApJ, 826, 56
Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, ApJ, 876, 85
Ross A. J., Samushia L., Howlett C., Percival W. J., Burden A., Manera M., 2015, MNRAS, 449, 835
Rusu C. E. et al., 2017, MNRAS, 467, 4220
Rusu C. E. et al., 2019a, preprint (arXiv:1905.09338)
Rusu C. E., Berghea C. T., Fassnacht C. D., More A., Seman E., Nelson G. J., Chen G. C. F., 2019b, MNRAS, 486, 4987
Rusu C. E. et al., 2019c, in press
Schechter P. L., Morgan N. D., Chehade B., Metcalfe N., Shanks T., McDonald M., 2017, AJ, 153, 219
Schneider P., Sluse D., 2013, A&A, 559, A37
Schneider P., Sluse D., 2014, A&A, 564, A103
Schneider P., Ehlers J., Falco E. E., 1992, Gravitational Lenses, XIV, Vol. 560. Springer-Verlag, Berlin, p. 112
Shajib A. J., Treu T., Agnello A., 2018, MNRAS, 473, 210
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shanks T., Hogarth L. M., Metcalfe N., 2019, MNRAS, 484, L64
Shapiro I. I., 1964, Phys. Rev. Lett., 13, 789
Sluse D. et al., 2003, A&A, 406, L43
Sluse D., Claeskens J. F., Hutsemékers D., Surdej J., 2007, A&A, 468, 885
Sluse D., Hutsemékers D., Courbin F., Meylan G., Wambsganss J., 2012, A&A, 544, A62
Sluse D. et al., 2017, MNRAS, 470, 4838
Sluse D. et al., 2019, preprint (arXiv:1905.08800)
Sonnenfeld A., 2018, MNRAS, 474, 4648
Sonnenfeld A., Treu T., Gavazzi R., Marshall P. J., Auger M. W., Suyu S. H., Koopmans L. V. E., Bolton A. S., 2012, ApJ, 752, 163
Sonnenfeld A., Treu T., Gavazzi R., Suyu S. H., Marshall P. J., Auger M. W., Nipoti C., 2013, ApJ, 777, 98

Spingola C., McKean J. P., Auger M. W., Fassnacht C. D., Koopmans L. V. E., Lagattuta D. J., Vegetti S., 2018, MNRAS, 478, 4816
Spiniello C. et al., 2018, MNRAS, 480, 1163
Springel V. et al., 2005, Nature, 435, 629
Suyu S. H., Halkola A., 2010, A&A, 524, A94
Suyu S. H., Marshall P. J., Hobson M. P., Blandford R. D., 2006, MNRAS, 371, 983
Suyu S. H., Marshall P. J., Blandford R. D., Fassnacht C. D., Koopmans L. V. E., McKean J. P., Treu T., 2009, ApJ, 691, 277
Suyu S. H., Marshall P. J., Auger M. W., Hilbert S., Blandford R. D., Koopmans L. V. E., Fassnacht C. D., Treu T., 2010, ApJ, 711, 201
Suyu S. H. et al., 2012, ApJ, 750, 10
Suyu S. H. et al., 2013, ApJ, 766, 70
Suyu S. H. et al., 2014, ApJ, 788, L35
Suyu S. H. et al., 2017, MNRAS, 468, 2590
Suyu S. H., Chang T.-C., Courbin F., Okumura T., 2018, Space Sci. Rev., 214, 91
Tewes M., Courbin F., Meylan G., 2013a, A&A, 553, A120
Tewes M. et al., 2013b, A&A, 556, A22
Tie S. S., Kochanek C. S., 2018, MNRAS, 473, 80
Tihhonova O. et al., 2018, MNRAS, 477, 5657
Tonry J. L., 1998, AJ, 115, 1
Treu T., Koopmans L. V. E., 2002, MNRAS, 337, L6
Treu T., Koopmans L. V. E., 2004, ApJ, 611, 739
Treu T., Marshall P. J., 2016, A&AR, 24, 11
Treu T. et al., 2018, MNRAS, 481, 1041
Vegetti S., Lagattuta D. J., McKean J. P., Auger M. W., Fassnacht C. D., Koopmans L. V. E., 2012, Nature, 481, 341
Vernardos G., Fluke C. J., 2014, Astron. Comput., 6, 1
Vuissoz C. et al., 2007, A&A, 464, 845
Vuissoz C. et al., 2008, A&A, 488, 481
Wertz O., Orthen B., 2018, A&A, 619, A117
Wertz O., Orthen B., Schneider P., 2018, A&A, 617, A140
Weymann R. J. et al., 1980, Nature, 285, 641
Williams P. R. et al., 2018, MNRAS, 477, L70
Wilson M. L., Zabludoff A. I., Ammons S. M., Momcheva I. G., Williams K. A., Keeton C. R., 2016, ApJ, 833, 194
Wisotzki L., Schechter P. L., Bradt H. V., Heinmüller J., Reimers D., 2002, A&A, 395, 17
Wong K. C., Keeton C. R., Williams K. A., Momcheva I. G., Zabludoff A. I., 2011, ApJ, 726, 84
Wong K. C. et al., 2017, MNRAS, 465, 4895
Wyman M., Rudd D. H., Vanderveld R. A., Hu W., 2014, Phys. Rev. Lett., 112, 051302
Xu D., Sluse D., Schneider P., Springel V., Vogelsberger M., Nelson D., Hernquist L., 2016, MNRAS, 456, 739
Yıldırım A., Suyu S. H., Halkola A., 2019, preprint (arXiv:1904.07237)
Yoo J., Kochanek C. S., Falco E. E., McLeod B. A., 2005, ApJ, 626, 51
Zhang B. R., Childress M. J., Davis T. M., Karpenka N. V., Lidman C., Schmidt B. P., Smith M., 2017, MNRAS, 471, 2254

## APPENDIX A: COMPARISON OF THE RECONSTRUCTED AO PSF FROM THREE LENSES

We show the comparison of the reconstructed AO PSF in Fig. A1. The 2D plots of the three PSF clearly show that the atmosphere disturbance produces various structures of PSF with a core plus a wing, while the radial average intensity indicates that the intensity gradient of the cores is very similar inside 0.1 arcsec.

## APPENDIX B: DEGENERACY BETWEEN AO PSF WING AND LENS LIGHT

The PSF reconstructed method developed in Chen et al. (2016) allows us to model the AO imaging down to the noise level and recover $D_{\Delta t}$ when we adopt the power-law model. However, we found that there exists a degeneracy between AO PSF wing and the lens light if we reconstruct the AO PSF from the AO imaging only. While it does not cause a problem for the power-law model since the power-law mass is constrained by the arc, the composite model instead could yield a biased $H_0$ because we assume that the baryonic matter distribution follows the lens light. Thus, in the case of HE 0435−1223 and RXJ 1131−1231 where we do not reconstruct the PSF simultaneously from *HST* and AO imaging, we fix the mass distribution of the baryonic matter to the baryonic mass inferred from *HST* imaging.

In the case of PG 1115+080, we found that the degeneracy can be broken by simultaneously modelling the AO imaging and *HST* imaging as they share the same mass model. The additional constraint allows us to better characterize the AO PSF wing. Thus, given the same PSF, both power-law model and composite model can be modelled down to the noise level. In the future, 2D kinematic data could also potentially further break the degeneracy between the PSF wing and lens light, as it provides the information to characterize the dark and baryonic matter content (Cappellari et al. 2013).

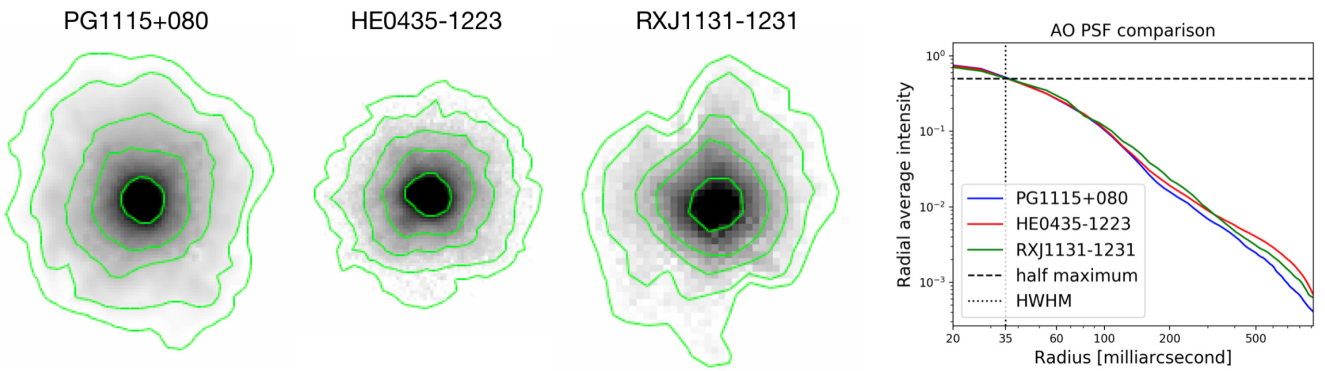


**Figure A1.** The left three figures are the reconstructed AO PSF of PG 1115+080, HE 0435−1223, and RXJ 1131−1231, respectively. The right-hand panel is the comparison of the radial average intensity of the PSF, which shows the core plus its wings. The isointensity contours represent 0.07, 0.01, 0.005, 0.0025, and 0.0015.

**Table C1.** Joined constraints used to estimate $\kappa_{\rm ext}$.

| Lens | Model name | $\gamma_{\rm ext}$ | $\zeta_1^{45\,\rm arcsec}$ | $\zeta_{1/r}^{45\,\rm arcsec}$ | $\zeta_1^{120\,\rm arcsec}$ | $\zeta_{1/r}^{120\,\rm arcsec}$ |
|---|---|---|---|---|---|---|
| RXJ 1131−1231 | Power law G | $0.083 \pm 0.003$ | $1.4^{+0.05}_{-0.05}$ | | | |
| RXJ 1131−1231 | Composite G | $0.058 \pm 0.005$ | $1.4^{+0.05}_{-0.05}$ | | | |
| HE 0435−1223 | Power law G + G1 | $0.041 \pm 0.018$ | | | | |
| HE 0435−1223 | Composite G + G1 | $0.026 \pm 0.026$ | | | | |
| HE 0435−1223 | Composite G + 5 perturbers | $0.056 \pm 0.026$ | | | | |
| HE 0435−1223 | Power law G + G1 + LOS | $0.041 \pm 0.018$ | $1.27^{+0.05}_{-0.05}$ | $1.31^{+0.05}_{-0.05}$ | | |
| HE 0435−1223 | Composite G + G1 + LOS | $0.026 \pm 0.026$ | $1.27^{+0.05}_{-0.05}$ | $1.31^{+0.05}_{-0.05}$ | | |
| HE 0435−1223 | Composite G + 5 perturbers + LOS | $0.056 \pm 0.026$ | $1.21^{+0.05}_{-0.05}$ | $1.17^{+0.05}_{-0.05}$ | | |
| PG 1115+080 | Power law G + NFW group | $0.027 \pm 0.009$ | $1.11^{+0.01}_{-0.11}$ | $1.27^{+0.01}_{-0.23}$ | $1.10^{+0.05}_{-0.09}$ | $0.98^{+0.03}_{-0.10}$ |
| PG 1115+080 | Power law G + SIS group | $0.061 \pm 0.009$ | $1.11^{+0.01}_{-0.11}$ | $1.27^{+0.01}_{-0.23}$ | $1.10^{+0.05}_{-0.09}$ | $0.98^{+0.03}_{-0.10}$ |
| PG 1115+080 | Power law G + NFW group + G1 | $0.054 \pm 0.009$ | $1.00^{+0.01}_{-0.11}$ | $1.02^{+0.03}_{-0.12}$ | $1.09^{+0.04}_{-0.09}$ | $0.96^{+0.04}_{-0.10}$ |
| PG 1115+080 | Power law G + NFW group + G1 + G2 | $0.058 \pm 0.012$ | $0.89^{+0.01}_{-0.11}$ | $0.84^{+0.06}_{-0.24}$ | $1.07^{+0.05}_{-0.09}$ | $0.94^{+0.04}_{-0.10}$ |
| PG 1115+080 | Composite G + NFW group | $0.048 \pm 0.009$ | $1.11^{+0.01}_{-0.11}$ | $1.27^{+0.01}_{-0.23}$ | $1.10^{+0.05}_{-0.09}$ | $0.98^{+0.03}_{-0.10}$ |
| PG 1115+080 | Composite G + SIS group + G1 | $0.072 \pm 0.008$ | $1.00^{+0.01}_{-0.11}$ | $1.02^{+0.03}_{-0.12}$ | $1.09^{+0.04}_{-0.09}$ | $0.96^{+0.04}_{-0.10}$ |
| PG 1115+080 | Composite G + NFW group + G1 | $0.072 \pm 0.008$ | $1.00^{+0.01}_{-0.11}$ | $1.02^{+0.03}_{-0.12}$ | $1.09^{+0.04}_{-0.09}$ | $0.96^{+0.04}_{-0.10}$ |
| PG 1115+080 | Composite G + NFW group + G1 + G2 | $0.060 \pm 0.008$ | $0.89^{+0.01}_{-0.11}$ | $0.84^{+0.06}_{-0.24}$ | $1.07^{+0.05}_{-0.09}$ | $0.94^{+0.04}_{-0.10}$ |

## APPENDIX C: CONSTRAINTS USED TO ESTIMATE $\kappa_{\rm EXT}$

Here, we show the $\kappa_{\rm ext}$ distribution of HE 0435−1223 in Fig. 12 and present the summary table of the constraints used to estimate $\kappa_{\rm ext}$ for all three systems in Table C1.

## APPENDIX D: ACCOUNTING FOR THE MISSING GALAXY GROUP MEMBERS IN PG 1115+080 DUE TO SPECTROSCOPIC INCOMPLETENESS

The spectroscopic coverage of the FOV around PG 1115+080 is incomplete. Down to $R_{\rm c} \leq 22.5$ and within 120 arcsec radius around the lens, there are 63 galaxies, out of which 33 have spectroscopy (Wilson et al. 2016), 11 of which are part of the galaxy group at $z = 0.31$, including the lensing galaxy. This means that there may be other galaxies within this magnitude range and radius from the lens that are also part of the galaxy group associated with the lensing galaxy, but that are missed due to spectroscopic incompleteness.[9] In Section 5.1.3, we have specifically computed the lensing properties of this group, based on its physical properties derived by Momcheva et al. (2015). As a result, when we compute $\kappa_{\rm ext}$ at the location of the lens, using the weighted number counts approach, we must remove the galaxies that are part of this group, as the convergence from the group has already been included in the lensing models, and must not be double counted. While the galaxies that are known to be part of the group can easily be removed, we must also account for the galaxies expected to be missed due to our spectroscopic incompleteness.

Following the technique presented in Rusu et al. (2019a), we use two different approaches to estimate the number of missing galaxies that are expected to be part of the group, but that are not identified as group members because of our spectroscopic incompleteness. In the first approach, we use the knowledge provided by the number of known group members, the number of galaxies with spectroscopy, and the total number of detected galaxies (within the given magnitude and aperture radius), and we apply Poisson statistics to estimate a number of $10 \pm 5$ missing galaxies (median, 16th, and 84th percentiles). In the second approach, we use the group velocity dispersion and virial radius from Wilson et al. (2016), and we estimate the expected number of galaxies inside the virial radius using the empirical relation from Andreon & Hurn (2010). Using the measured offset from the group centroid to the lens, and propagating all uncertainties, we measure the expected number of missing galaxies at the intersection of the sphere of virial radius and the 120 arcsec radius cylinder centred on the lens to be $1^{+3}_{-1}$. We plot the distributions of these numbers in Fig. D1. The first approach predicts a significantly larger number of missing galaxies than the second. In fact, due to the small value of the velocity dispersion, the second method would only predict a total number of $8^{+6}_{-4}$ galaxies, therefore less than the number of confirmed group members, unless we enforce this constraint. This discrepancy may be due to the shallow absolute magnitude limit of $M_{\rm V} = -20$ used by Andreon & Hurn (2010), corresponding to $r \sim 21$ at $z \sim 0.3$, therefore significantly brighter than our limiting magnitude. We note, however, that the two techniques produced results that were in agreement for a different lens, described in Rusu et al. (2019c).

In view of the above, and also due to the fact that it avoids any physical assumption, we consider the first method to be more reliable. Finally, when computing weighted galaxy counts, we do this by randomly sampling 10 times from the distribution of missing galaxy numbers, and then randomly excluding that number

[9]For RXJ 1131−1231 and HE 0435−1223, we have not incorporated any galaxy groups in the lens models; therefore, we do not need to remove the contribution of such groups from our estimate of the external convergence based on number counts. Hence, the spectroscopic incompleteness is not relevant to our analysis.

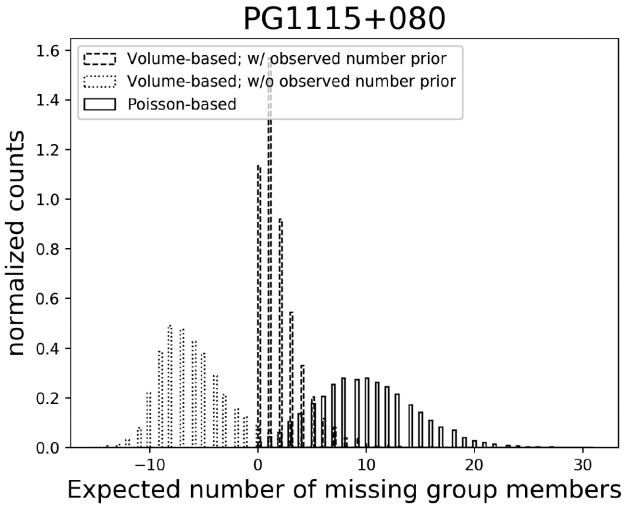


**Figure D1.** Estimated number of missing galaxy group members inside the ≤120 arcsec radius from the lens system for PG 1115+080, computed with two methods, with or without imposing the prior that the group consists of at least the number of galaxies spectroscopically confirmed to be members.

of galaxies from our catalogue of galaxies inside the 120 arcsec apertures.

## APPENDIX E: COMBINING THE INDEPENDENT MEASUREMENTS

Consider that we have following sets of information, AO imaging ($\boldsymbol{d}_{\rm AO}$), *HST* imaging ($\boldsymbol{d}_{HST}$), velocity dispersion ($\sigma$), and environment data ($\boldsymbol{d}_{\rm ENV}$), which can be used to constrain the difference of the Fermat potential. We already have two independent measurements, $P(\boldsymbol{\eta}|\boldsymbol{d}_{\rm AO}, {\rm H}, {\rm MS})$ and $P(\boldsymbol{\eta}|\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H}, {\rm MS})$. The goal is to combine them and obtain the joint constraint, $P(\boldsymbol{\eta}|\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H}, {\rm MS})$, where $\boldsymbol{\eta} = (\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta})$, $\phi^{\rm H}_{ij}$ is the difference of the model Fermat potentials (not the true Fermat potential, $\phi^{\rm True}_{ij}$) at imaging $i$ and $j$ that we are interested in, $\kappa_{\rm ext}$ is the convergence from the line of sight, $\gamma_{\rm ext}$ is the external shear inferred from the imaging data, and $\boldsymbol{\zeta}$ are the other parameters that we want to marginalize upon. H is one kind of the lens model (i.e. the power-law model or the composite model) and MS is Millennium Simulation.

We start with the joint constraint and step by step link to the independent measurements. With Bayes' theorem, the joint constraint can be expressed as

$$P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H}, {\rm MS}) = P(\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}|\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}, {\rm MS}) \cdot \frac{P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|{\rm H}, {\rm MS})}{P(\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}|{\rm H}, {\rm MS})}, \quad \text{(E1)}$$

where

$$P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|{\rm H}, {\rm MS}) = P(\phi^{\rm H}_{ij}|{\rm H}) P(\kappa_{\rm ext}|{\rm H,MS}) P(\gamma_{\rm ext}|{\rm H}) P(\boldsymbol{\zeta}|{\rm H}) \quad \text{(E2)}$$

are the priors on the parameters of the mass model and MS. Note that Millennium Simulation naturally provides more $\kappa_{\rm ext}$, which close to mean density and lens model implicitly assumes $\kappa_{\rm ext} < 1$, so $P(\kappa_{\rm ext}|{\rm H}, {\rm MS})$ is a non-flat prior with an upper bound (<1).

The next step is to separate the data sets into $[\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}]$ and $[\boldsymbol{d}_{\rm AO}]$. Since the data are all independent, we can write down

$$P(\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}|\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}, {\rm MS}) = P(\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}|\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}, {\rm MS}) \cdot P(\boldsymbol{d}_{\rm AO}|\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}, {\rm MS}). \quad \text{(E3)}$$

Although $\boldsymbol{d}_{\rm AO}$ do not have the direct constraint power on $\kappa_{\rm ext}$, the shear value inferred from $\boldsymbol{d}_{\rm AO}$ implicitly helps constrain $\kappa_{\rm ext}$. Furthermore, we leave $\kappa_{\rm ext}$ in the last term in equation (E3) because we want to link to $P(\boldsymbol{\eta}|\boldsymbol{d}_{\rm AO}, {\rm H}, {\rm MS})$ in the future steps. Equation (E1) becomes

$$P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H}, {\rm MS}) = P(\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}|\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}, {\rm MS}) \cdot \frac{P(\boldsymbol{d}_{\rm AO}|\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}, {\rm MS}) P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|{\rm H}, {\rm MS})}{P(\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}|{\rm H}, {\rm MS})}, \quad \text{(E4)}$$

where

$$P(\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}|\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}, {\rm MS}) = P(\boldsymbol{d}_{HST}|\phi^{\rm H}_{ij}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}) P(\sigma|\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}, {\rm MS}) \cdot P(\boldsymbol{d}_{\rm ENV}|\gamma_{\rm ext}, \kappa_{\rm ext}, {\rm MS}). \quad \text{(E5)}$$

The last term in equation (E5) tells us that the shear value inferred from the lens imaging can also help us to further constrain the convergence because of the correlation between the $\gamma_{\rm ext}$ and $\kappa_{\rm ext}$ in MS.

Bayes' theorem tells us that

$$P(\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}|\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}, {\rm MS}) = P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H}, {\rm MS}) \cdot \frac{P(\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}|{\rm H}, {\rm MS})}{P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|{\rm H}, {\rm MS})} \quad \text{(E6)}$$

and

$$P(\boldsymbol{d}_{\rm AO}|\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}, {\rm H}, {\rm MS}) = \frac{P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|\boldsymbol{d}_{\rm AO}, {\rm H}, {\rm MS}) P(\boldsymbol{d}_{\rm AO}|{\rm H}, {\rm MS})}{P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|{\rm H}, {\rm MS})}. \quad \text{(E7)}$$

By substituting equation (E6) and equation (E7) into equation (E4), finally we obtain

$$P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H}, {\rm MS}) \propto P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|\boldsymbol{d}_{HST}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H}, {\rm MS}) \cdot \frac{P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|\boldsymbol{d}_{\rm AO}, {\rm H}, {\rm MS})}{P(\phi^{\rm H}_{ij}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|{\rm H}, {\rm MS})}. \quad \text{(E8)}$$

The first term on the right-hand side in equation (E8) is done by Wong et al. (2017), while the numerator in the second term is from AO data alone. Thus, based on equation (E8), in order to get the joint constraint, we need to multiply these two posteriors and divide by the non-uniform priors (e.g. $P(\kappa_{\rm ext}|{\rm H}, {\rm MS})$) used in both data sets. This is so because we need to get rid of the extra constraining power from doubly using the same non-uniform priors. Therefore,

the denominator in equation (E8) becomes $P(\kappa_{\rm ext}|{\rm H, MS})$. Note that $\phi_{ij}^{\rm H}$ is the model Fermat potential, but what we want to obtain is $P(\phi_{ij}^{\rm True})$, which can be expressed as

$$
\begin{aligned}
&P(\phi_{ij}^{\rm True}|\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H, MS}) \\
&= \int {\rm d}\kappa_{\rm ext} \int {\rm d}\phi_{ij}^{\rm H} P(\phi_{ij}^{\rm H}, \kappa_{\rm ext}) \delta(\phi_{ij}^{\rm True} - \phi_{ij}^{H}(1-\kappa_{\rm ext})) \\
&= \int {\rm d}\kappa_{\rm ext} \int {\rm d}\phi_{ij}^{\rm H} P(\phi_{ij}^{\rm H}, \kappa_{\rm ext}) \frac{\delta(\phi_{ij}^{\rm H} - \phi_{ij}^{\rm True}/(1-\kappa_{\rm ext}))}{|-(1-\kappa_{\rm ext})|} \\
&= \int {\rm d}\kappa_{\rm ext} \frac{P(\phi_{ij}^{\rm True}/(1-\kappa_{\rm ext}), \kappa_{\rm ext})}{1-\kappa_{\rm ext}}, \qquad \text{(E9)}
\end{aligned}
$$

where

$$
\begin{aligned}
P(\phi_{ij}^{\rm H}, \kappa_{\rm ext}) = \int & P(\phi_{ij}^{\rm H}, \kappa_{\rm ext}, \gamma_{\rm ext}, \boldsymbol{\zeta}|\boldsymbol{d}_{HST}, \boldsymbol{d}_{\rm AO}, \sigma, \boldsymbol{d}_{\rm ENV}, {\rm H, MS}) \\
& \times {\rm d}\boldsymbol{\zeta}\, {\rm d}\gamma_{\rm ext}. \qquad \text{(E10)}
\end{aligned}
$$

## APPENDIX F: SUMMARY OF PG 1115+080 LENS MODELS WITH RESPECT TO THE BIC VALUE

We present the BIC of the power-law models in Table F1 and composite models in Table F2.

**Table F1.** Total 20power-law models of PG 1115+080 ordered in increased ΔBIC value.

| Main lens model | Perturbers | $S_{\rm r}$ | ΔBIC | Posterior weight |
|---|---|---|---|---|
| SPEMD | Group (NFW) | 37 × 37 | 0 | 1 |
| SPEMD | Group (NFW) + G1 + G2 | 37 × 37 | 11 | 0.9820 |
| SPEMD | Group (NFW) + G1 + G2 | 39 × 39 | 18 | 0.9469 |
| SPEMD | Group (NFW) + G1 | 37 × 37 | 31 | 0.8967 |
| SPEMD | Group (NFW) | 38 × 38 | 35 | 0.8338 |
| SPEMD | Group (NFW) + G1 + G2 | 33 × 33 | 41 | 0.7614 |
| SPEMD | Group (NFW) + G1 | 41 × 41 | 46 | 0.6827 |
| SPEMD | Group (NFW) + G1 + G2 | 35 × 35 | 46 | 0.6827 |
| SPEMD | Group (NFW) + G1 | 39 × 39 | 54 | 0.5198 |
| SPEMD | Group (SIS) + G1 | 36 × 36 | 72 | 0.4415 |
| SPEMD | Group (SIS) + G1 | 34 × 34 | 74 | 0.3684 |
| SPEMD | Group (NFW) + G1 | 35 × 35 | 88 | 0.3019 |
| SPEMD | Group (NFW) + G1 + G2 | 41 × 41 | 90 | 0.2433 |
| SPEMD | Group (SIS) + G1 | 32 × 32 | 91 | 0.1930 |
| SPEMD | Group (NFW) | 32 × 32 | 92 | 0.1510 |
| SPEMD | Group (NFW) | 36 × 36 | 95 | 0.1172 |
| SPEMD | Group (NFW) | 34 × 34 | 114 | 0.0911 |
| SPEMD | Group (NFW) + G1 | 43 × 43 | 126 | 0.0721 |
| SPEMD | Group (SIS) + G1 | 40 × 40 | 127 | 0.0597 |
| SPEMD | Group (SIS) + G1 | 38 × 38 | 148 | 0.0536 |

**Table F2.** Total 20 composite models of PG 1115+080 ordered in increased ΔBIC value.

| Main lens model | Perturbers | $S_{\rm r}$ | ΔBIC | Posterior weight |
|---|---|---|---|---|
| Composite | Group (SIS) + G1 | 37 × 37 | 0 | 1 |
| Composite | Group (NFW) | 30 × 30 | 34 | 0.9799 |
| Composite | Group (NFW) | 32 × 32 | 55 | 0.9409 |
| Composite | Group (NFW) | 36 × 36 | 67 | 0.8853 |
| Composite | Group (NFW) + G1 + G2 | 39 × 39 | 69 | 0.8162 |
| Composite | Group (NFW) + G1 + G2 | 37 × 37 | 80 | 0.7374 |
| Composite | Group (NFW) + G1 | 33 × 33 | 88 | 0.6529 |
| Composite | Group (NFW) + G1 | 35 × 35 | 90 | 0.5664 |
| Composite | Group (NFW) + G1 | 32 × 32 | 96 | 0.4815 |
| Composite | Group (NFW) + G1 + G2 | 33 × 33 | 118 | 0.4010 |
| Composite | Group (NFW) + G1 + G2 | 35 × 35 | 122 | 0.3273 |
| Composite | Group (NFW) | 34 × 34 | 132 | 0.2620 |
| Composite | Group (SIS) + G1 | 31 × 31 | 138 | 0.2057 |
| Composite | Group (SIS) + G1 | 35 × 35 | 145 | 0.1584 |
| Composite | Group (NFW) + G1 | 39 × 39 | 146 | 0.1200 |
| Composite | Group (SIS) + G1 | 33 × 33 | 149 | 0.0897 |
| Composite | Group (SIS) + G1 | 39 × 39 | 156 | 0.0669 |
| Composite | Group (NFW) + G1 + G2 | 31 × 31 | 163 | 0.0506 |
| Composite | Group (NFW) + G1 | 37 × 37 | 193 | 0.0402 |
| Composite | Group (NFW) | 38 × 38 | 232 | 0.0352 |

## APPENDIX G: SUMMARY OF HE 0435−1223 LENS MODELS WITH RESPECT TO THE BIC VALUE

We present the BIC of the power-law models in Table G1 and composite models in Table G2.

**Table G1.** Total five power-law models of HE 0435−1223 ordered in increased ΔBIC value.

| Main lens model | Perturbers | $S_r$ | ΔBIC | Posterior weight |
|---|---|---|---|---|
| SPEMD | G1 | 50 × 50 | 0 | 1 |
| SPEMD | G1 | 40 × 40 | 5 | 0.9743 |
| SPEMD | G1 | 45 × 45 | 7 | 0.9328 |
| SPEMD | G1 | 35 × 35 | 14 | 0.8914 |
| SPEMD | G1 | 30 × 30 | 16 | 0.8658 |

**Table G2.** Total 10 composite models of HE 0435−1223 ordered in increased ΔBIC value.

| Main lens model | Perturbers | $S_r$ | ΔBIC | Posterior weight |
|---|---|---|---|---|
| Composite | G1–G5 | 43 × 43 | 0 | 1 |
| Composite | G1 | 43 × 43 | 10 | 0.9855 |
| Composite | G1 | 41 × 41 | 11 | 0.9580 |
| Composite | G1–G5 | 41 × 41 | 12 | 0.9201 |
| Composite | G1 | 39 × 39 | 16 | 0.8756 |
| Composite | G1–G5 | 39 × 39 | 19 | 0.8288 |
| Composite | G1–G5 | 37 × 37 | 20 | 0.7844 |
| Composite | G1 | 35 × 35 | 21 | 0.7466 |
| Composite | G1 | 37 × 37 | 21 | 0.7466 |
| Composite | G1–G5 | 35 × 35 | 23 | 0.7046 |

## APPENDIX H: SUMMARY OF RXJ 1131−1231 LENS MODELS WITH RESPECT TO THE BIC VALUE

We present the BIC of the power-law models in Table H1 and composite models in Table H2.

**Table H1.** Total five power-law models of RXJ 1131−1231 ordered in increased ΔBIC value.

| Main lens model | Perturbers | $S_r$ | ΔBIC | Posterior weight |
|---|---|---|---|---|
| SPEMD | Satellite (SIS) | 79 × 79 | 0 | 1 |
| SPEMD | Satellite (SIS) | 77 × 77 | 85 | 0.4618 |
| SPEMD | Satellite (SIS) | 75 × 75 | 176 | 0.0912 |
| SPEMD | Satellite (SIS) | 73 × 73 | 499 | 0.0071 |
| SPEMD | Satellite (SIS) | 71 × 71 | 795 | 0.0002 |

**Table H2.** Total five composite models of RXJ 1131−1231 ordered in increased ΔBIC value.

| Main lens model | Perturbers | $S_r$ | ΔBIC | Posterior weight |
|---|---|---|---|---|
| Composite | Satellite (SIS) | 71 × 71 | 0 | 1 |
| Composite | Satellite (SIS) | 70 × 70 | 398 | 0.4050 |
| Composite | Satellite (SIS) | 73 × 73 | 915 | 0.0575 |
| Composite | Satellite (SIS) | 74 × 74 | 1641 | 0.0025 |
| Composite | Satellite (SIS) | 72 × 72 | 1798 | 0.0000 |

[1]*Department of Physics, University of California, Davis, CA 95616, USA*
[2]*Max Planck Institute for Astrophysics, Karl-Schwarzschild-Strasse 1, D-85740 Garching, Germany*
[3]*Physik-Department, Technische Universität Müchen, James-Franck-Str 1, D-85748 Garching, Germany*
[4]*Academia Sinica Institute of Astronomy and Astrophysics (ASIAA), 11F of ASMAB, No. 1, Section 4, Roosevelt Road, Taipei 10617, Taiwan*
[5]*National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[6]*Subaru Telescope, National Astronomical Observatory of Japan, 650 N Aohoku Pl, Hilo, HI 96720, USA*
[7]*Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland*
[8]*Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), University of Tokyo, Chiba 277-8583, Japan*
[9]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[10]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[11]*Exzellenzcluster Universe, Boltzmannstr 2, D-85748 Garching, Germany*
[12]*Ludwig-Maximilians-Universitĺat, Universitäts-Sternwarte, Scheinerstr 1, D-81679 Munchen, Germany*
[13]*Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA*
[14]*Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, the Netherlands*
[15]*Univ Lyon, Univ Lyon 1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France*
[16]*ASTRON, Netherlands Institute for Radio Astronomy, PO Box 2, NL-7990 AA Dwingeloo, the Netherlands*
[17]*Py'orrekuja 5 A, FI-04300 Tuusula, Finland*
[18]*STAR Institute, Quartier Agora – Allée du six Aôut, 19c, B-4000 Liège, Belgium*
[19]*Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands*

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