

# H0LiCOW – IV. Lens mass model of HE 0435–1223 and blind measurement of its time-delay distance for cosmology

Kenneth C. Wong,<sup>1,2★†</sup> Sherry H. Suyu,<sup>3,2,4</sup> Matthew W. Auger,<sup>5</sup> Vivien Bonvin,<sup>6</sup> Frederic Courbin,<sup>6</sup> Christopher D. Fassnacht,<sup>7</sup> Alekski Halkola, Cristian E. Rusu,<sup>7</sup> Dominique Sluse,<sup>8</sup> Alessandro Sonnenfeld,<sup>9</sup> Tommaso Treu,<sup>10</sup> Thomas E. Collett,<sup>11</sup> Stefan Hilbert,<sup>12,13</sup> Leon V. E. Koopmans,<sup>14</sup> Philip J. Marshall<sup>15</sup> and Nicholas Rumbaugh<sup>7</sup>

<sup>1</sup>National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

<sup>2</sup>Institute of Astronomy and Astrophysics, Academia Sinica (ASIAA), PO Box 23-141, Taipei 10617, Taiwan

<sup>3</sup>Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748 Garching, Germany

<sup>4</sup>Physik-Department, Technische Universität München, James-Frank-Straße 1, D-85748 Garching, Germany

<sup>5</sup>Institute of Astronomy, University of Cambridge, Madingley Rd, Cambridge CB3 0HA, UK

<sup>6</sup>Laboratoire d'Astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauvigny, CH-1290 Versoix, Switzerland

<sup>7</sup>Department of Physics, University of California Davis, 1 Shields Avenue, Davis, CA 95616, USA

<sup>8</sup>STAR Institute, Quartier Agora – Allée du six Août, 19c B-4000 Liège, Belgium

<sup>9</sup>Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), University of Tokyo, Chiba 277-8583, Japan

<sup>10</sup>Department of Physics and Astronomy, University of California, Los Angeles, CA 90095-1547, USA

<sup>11</sup>Institute of Cosmology and Gravitation, University of Portsmouth, Burnaby Rd, Portsmouth PO1 3FX, UK

<sup>12</sup>Exzellenzcluster Universe, Boltzmannstr. 2, D-85748 Garching, Germany

<sup>13</sup>Ludwig-Maximilians-Universität, Universitäts-Sternwarte, Scheinerstr. 1, D-81679 München, Germany

<sup>14</sup>Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, the Netherlands

<sup>15</sup>Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94035, USA

Accepted 2016 November 24. Received 2016 November 21; in original form 2016 June 30

## ABSTRACT

Strong gravitational lenses with measured time delays between the multiple images allow a direct measurement of the time-delay distance to the lens, and thus a measure of cosmological parameters, particularly the Hubble constant,  $H_0$ . We present a blind lens model analysis of the quadruply imaged quasar lens HE 0435–1223 using deep *Hubble Space Telescope* imaging, updated time-delay measurements from the COSmological MONitoring of GRAvitational Lenses (COSMOGRAIL), a measurement of the velocity dispersion of the lens galaxy based on Keck data, and a characterization of the mass distribution along the line of sight. HE 0435–1223 is the third lens analysed as a part of the  $H_0$  Lenses in COSMOGRAIL's Wellspring (H0LiCOW) project. We account for various sources of systematic uncertainty, including the detailed treatment of nearby perturbers, the parametrization of the galaxy light and mass profile, and the regions used for lens modelling. We constrain the effective time-delay distance to be  $D_{\Delta t} = 2612_{-191}^{+208}$  Mpc, a precision of 7.6 per cent. From HE 0435–1223 alone, we infer a Hubble constant of  $H_0 = 73.1_{-6.0}^{+5.7}$  km s<sup>-1</sup> Mpc<sup>-1</sup> assuming a flat  $\Lambda$ CDM cosmology. The cosmographic inference based on the three lenses analysed by H0LiCOW to date is presented in a companion paper (H0LiCOW Paper V).

**Key words:** gravitational lensing: strong – cosmological parameters – distance scale.

## 1 INTRODUCTION

The flat  $\Lambda$  cold dark matter ( $\Lambda$ CDM) cosmological model is the concordance model of our Universe today. It is consistent with a variety of independent experiments, including an analysis of the cosmic microwave background (CMB) by the *Planck* mission (Planck

\*E-mail: [kcwong19@gmail.com](mailto:kcwong19@gmail.com)

†EACOA Fellow.

Collaboration XIII 2016). The *Planck* results provide the most precise cosmological parameter constraints to date, under the assumption of spatial flatness. However, there is no physical reason to assume flatness, and if the flatness assumption is relaxed, there are strong degeneracies among the cosmological parameters inferred from CMB data, particularly with the Hubble constant,  $H_0$  (e.g. Freedman et al. 2012; Riess et al. 2016). Therefore, an independent determination of  $H_0$  is crucial for understanding the nature of the Universe (e.g. Hu 2005; Suyu et al. 2012a; Weinberg et al. 2013).

The idea of using gravitational lens time delays to measure the Hubble constant dates back to Refsdal (1964). In practice, gravitational lens time delays provide a one-step method to determine the distance and hence the Hubble constant (e.g. Vanderriest et al. 1989; Keeton & Kochanek 1997; Schechter et al. 1997; Kochanek 2003; Koopmans et al. 2003; Saha et al. 2006; Oguri 2007; Fadely et al. 2010; Suyu et al. 2010, 2013; Sereno & Paraficz 2014; Rathna Kumar, Stalin & Prabhu 2015; Birrer, Amara & Refregier 2016; Chen et al. 2016). This method is independent of the cosmic distance ladder (e.g. Riess et al. 2011; Freedman et al. 2012) and serves as a key test of possible systematic effects in individual  $H_0$  probes. This method rests on the fact that light rays emitted from the source at the same instant will take different paths through space–time at each of the image positions. These paths have different lengths and traverse different gravitational potentials before reaching the observer, leading to an offset in arrival times. If the source exhibits variations in its flux, the delays can be measured by monitoring the lensed images. The measured time delays can be used to calculate the time-delay distance, a combination of angular diameter distances among the observer, lens and source. The time-delay distance is primarily sensitive to  $H_0$ , with weaker dependence on other cosmological parameters (e.g. Coe & Moustakas 2009; Treu & Marshall 2016).

However, a precise and accurate determination of  $H_0$  through this method requires a variety of observational data. A dedicated long-term monitoring campaign is necessary to obtain accurate time delays, as the uncertainty in  $H_0$  is directly related to the relative uncertainty in the measured time delays. Deep, high-resolution imaging is required to accurately model the lens using the extended source images, which is needed to break degeneracies between the mass profile and the underlying cosmology (e.g. Kochanek 2002; Warren & Dye 2003). In order to reduce the effects of the mass sheet degeneracy (e.g. Falco, Gorenstein & Shapiro 1985; Gorenstein, Shapiro & Falco 1988; Saha 2000; Schneider & Sluse 2013; Xu et al. 2016), a measurement of the lens galaxy’s velocity dispersion (e.g. Koopmans et al. 2003; Koopmans 2004) and an estimate of the external convergence,  $\kappa_{\text{ext}}$ , along the line of sight (LOS) is needed.  $\kappa_{\text{ext}}$  can also bias the lens model parameters if unaccounted for (e.g. Collett et al. 2013; Greene et al. 2013; McCully et al. 2014, 2016).

In an effort to provide an accurate independent estimate of  $H_0$  using time-delay lenses, we use a number of new data sets as part of our project,  $H_0$  Lenses in COSMOGRAIL’s Wellspring (HOLiCOW), to model five lensed quasars. These data sets include high-resolution imaging with the *Hubble Space Telescope* (*HST*), precise time-delay measurements from the COSmological MONitoring of GRAVItational Lenses (COSMOGRAIL; Courbin et al. 2005; Eigenbrod et al. 2005; Bonvin et al. 2016b) project and from Very Large Array monitoring (Fassnacht et al. 2002), a photometric and spectroscopic survey to characterize the LOS mass distribution to estimate  $\kappa_{\text{ext}}$  in these systems and stellar velocity dispersion measurements of the strong lens galaxies. With five separate lenses, we plan to account for systematic uncertainties and obtain a robust constraint on  $H_0$  to <3.5 per cent precision.

In this paper, we present the results of a detailed lens modelling analysis of the gravitational lens HE 0435–1223 using new high-resolution imaging data from *HST*. HE 0435–1223 is the third HOLiCOW system analysed in this manner, following B1608+656 (Suyu et al. 2010) and RXJ1131–1231 (Suyu et al. 2013, 2014). This paper is the fourth in a series of papers detailing our analysis of HE 0435–1223. The other papers include an overview of the HOLiCOW project (Suyu et al. 2016, hereafter **HOLiCOW Paper I**), a spectroscopic survey of the HE 0435–1223 field and a characterization of the groups along the LOS (Sluse et al. 2016, hereafter **HOLiCOW Paper II**), a photometric survey of the HE 0435–1223 field and an estimate of  $\kappa_{\text{ext}}$  due to the external LOS structure (Rusu et al. 2016, hereafter **HOLiCOW Paper III**), and a presentation of our latest time-delay measurements for HE 0435–1223 and the cosmological inference from our combined analysis of HE 0435–1223, B1608+656 and RXJ1131–1231 (Bonvin et al. 2016a, hereafter **HOLiCOW Paper V**).

This paper is organized as follows. We provide a brief overview of using time-delay lenses for cosmography in Section 2. In Section 3, we describe the observational data used in our analysis. We describe our lens modelling procedure in Section 4. The time-delay distance results and their implications for cosmology are presented in Section 5. We summarize our main conclusions in Section 6. Throughout this paper, all magnitudes given are on the AB system.

## 2 TIME-DELAY COSMOGRAPHY

### 2.1 Time-delay distance

When a source is gravitationally lensed, the light travel time from the source to the observer depends on both the path length of the light rays and the gravitational potential of the lens through which the rays pass. For a single lens plane, the excess time delay of an image at an angular position  $\theta = (\theta_1, \theta_2)$  with corresponding source position  $\beta = (\beta_1, \beta_2)$  relative to the case of no lensing is

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

where  $D_{\Delta t}$  is the time-delay distance and  $\psi(\theta)$  is the lens potential. The time-delay distance (Refsdal 1964; Schneider, Ehlers & Falco 1992; Suyu et al. 2010) is defined<sup>1</sup> as

$$D_{\Delta t} \equiv (1 + z_d) \frac{D_d D_s}{D_{\text{ds}}}, \quad (2)$$

where  $z_d$  is the lens redshift,  $D_d$  is the angular diameter distance to the lens,  $D_s$  is the angular diameter distance to the source, and  $D_{\text{ds}}$  is the angular diameter distance between the lens and the source. Since  $D_{\Delta t}$  has units of distance, it is inversely proportional to  $H_0$ .

For lens systems with multiple deflectors at distinct redshifts, the observed time delays depend on various combinations of the angular diameter distances measured between us, the multiple deflectors and the source, and the observed time delays are no longer proportional to a single time-delay distance. The observed image positions depend on the multiplane lens equation (e.g. Blandford & Narayan 1986; Kovner 1987; Schneider et al. 1992; Petters, Levine & Wambsgans 2001; Collett & Auger 2014; McCully et al. 2014). However, for a system where the lensing is dominated by a single

<sup>1</sup> For historical reasons, the time-delay distance is written in terms of angular diameter distances. A more natural definition is  $D_{\Delta t} \equiv \hat{D}_d \hat{D}_s / \hat{D}_{\text{ds}}$  where  $\hat{D}$  are the proper distances that the photons have travelled.

plane, the observed time delays are primarily sensitive to the time-delay distance defined in equation (2), with the deflector redshift as that of the primary strong lens plane. We show in Section 4.9 that this approximation is valid for HE 0435–1223 and our results can thus be interpreted as a constraint on  $D_{\Delta t}(z_d, z_s)$ , which we refer to as the effective time-delay distance measured by this system. Hereafter,  $D_{\Delta t}$  refers to this effective time-delay distance unless otherwise indicated.

For variable sources such as active galactic nuclei (AGN), it is possible to monitor the fluxes of the lensed images at positions  $\theta_i$  and  $\theta_j$  and measure the time delay,  $\Delta t_{ij} \equiv t(\theta_i, \beta) - t(\theta_j, \beta)$ , between them (e.g. Vanderriest et al. 1989; Schechter et al. 1997; Fassnacht et al. 1999, 2002; Kochanek et al. 2006; Courbin et al. 2011). The lens potentials at the two image positions,  $\psi(\theta_i)$  and  $\psi(\theta_j)$ , as well as the source position,  $\beta$ , can be determined from a mass model of the system. Therefore, lenses with measured time delays and accurate lens models can be used to constrain  $D_{\Delta t}$ .

A complicating factor in using time-delay lenses for cosmography is the fact that all mass along the LOS contributes to the lens potential that the light rays pass through. These external perturbers not only affect the lens model of the system, but also lead to additional focusing and defocusing of the light rays, which in turn affects the measured time delays (e.g. Seljak 1994). If unaccounted for, these external perturbers can lead to biased inferences of  $D_{\Delta t}$ . If effects of LOS perturbers are small, they can be approximated by an external convergence term in the lens plane,  $\kappa_{\text{ext}}$  (neglecting the  $1 - \beta$  terms that enter into a more accurate prescription; Keeton 2003; McCully et al. 2014). The true  $D_{\Delta t}$  is related to the  $D_{\Delta t}^{\text{model}}$  inferred from a mass model that does not account for  $\kappa_{\text{ext}}$  by

$$D_{\Delta t} = \frac{D_{\Delta t}^{\text{model}}}{1 - \kappa_{\text{ext}}}. \quad (3)$$

$\kappa_{\text{ext}}$  cannot be constrained from the lens model due to the mass sheet degeneracy (e.g. Falco et al. 1985; Gorenstein et al. 1988; Saha 2000), in which the addition of a uniform mass sheet and a rescaling of the source plane coordinates can affect the product of the time delays and  $H_0$  but leave other observables unchanged.

The above degeneracy caused by  $\kappa_{\text{ext}}$  can be broken or substantially mitigated by estimating the mass distribution along the LOS (e.g. Fassnacht et al. 2006; Momcheva et al. 2006, 2015; Williams et al. 2006; Wong et al. 2011). However, for perturbers that are very massive or projected very close to the lens, they may need to be included explicitly in the mass model, as their higher order effects need to be properly accounted for (McCully et al. 2016). On the other hand, the lens profile is also degenerate with the time-delay distance in that the radial profile slope is tightly correlated with the time-delay distance (e.g. Kochanek 2002; Wucknitz 2002; Suyu 2012). The profile degeneracy affects models that share the same form of mass density profile (e.g. a power-law density profile), as well as models with different density profiles (described analytically or not). Furthermore, the profile degeneracy can mimic the effects of the mass sheet degeneracy since different profiles can exactly or approximately be mass sheet transformations of one form or another (e.g. Schneider & Sluse 2013, 2014; Unruh, Schneider & Sluse 2016). With reasonable assumptions about the lens galaxy's mass profile, these degeneracies can be reduced by augmenting the lensing data with stellar kinematics measurements of the lens galaxy (e.g. Treu & Koopmans 2002; Koopmans et al. 2003; Auger et al. 2010; Suyu et al. 2014). Including the velocity dispersion in the modelling helps to constrain any internal uniform mass component from a local galaxy group that the dynamics is sensitive to (Koopmans 2004).

## 2.2 Joint inference

Our inference of  $D_{\Delta t}$  follows that of Suyu et al. (2013), but with some important modifications. Our observational data sets are denoted by  $\mathbf{d}_{\text{HST}}$  for the *HST* imaging data,  $\Delta \mathbf{t}$  for the time delays,  $\sigma$  for the velocity dispersion of the lens galaxy and  $\mathbf{d}_{\text{LOS}}$  for the properties of the LOS mass distribution determined from our photometric and spectroscopic data. We want to obtain the posterior probability distribution function (PDF) of the model parameters  $\nu$  given the data,  $P(\xi | \mathbf{d}_{\text{HST}}, \Delta \mathbf{t}, \sigma, \mathbf{d}_{\text{LOS}}, \mathbf{A})$ . The vector  $\xi$  includes the lens model parameters  $\nu$ , the cosmological parameters  $\pi$  (Section 4.9) and nuisance parameters representing the external convergence ( $\kappa_{\text{ext}}$ ; Section 4.4) and anisotropy radius ( $r_{\text{ani}}$ ; Section 4.3), each of which we introduce and discuss in the sections indicated.  $\mathbf{A}$  denotes a discrete set of assumptions we make about the form of the model, including the choices we have to make about the data modelling region, the setup of the source reconstruction grid, the treatment of the various deflector mass distributions, etc. In general,  $\mathbf{A}$  cannot be fully captured by continuous parameters. By Bayes' theorem, we have that

$$P(\xi | \mathbf{d}_{\text{HST}}, \Delta \mathbf{t}, \sigma, \mathbf{d}_{\text{LOS}}, \mathbf{A}) \propto P(\mathbf{d}_{\text{HST}}, \Delta \mathbf{t}, \sigma, \mathbf{d}_{\text{LOS}} | \xi, \mathbf{A}) P(\xi | \mathbf{A}), \quad (4)$$

where  $P(\mathbf{d}_{\text{HST}}, \Delta \mathbf{t}, \sigma, \mathbf{d}_{\text{LOS}} | \xi, \mathbf{A})$  is the joint likelihood function and  $P(\xi | \mathbf{A})$  is the prior PDF for the parameters given our assumptions. Since the data sets are independent, the likelihood can be separated,

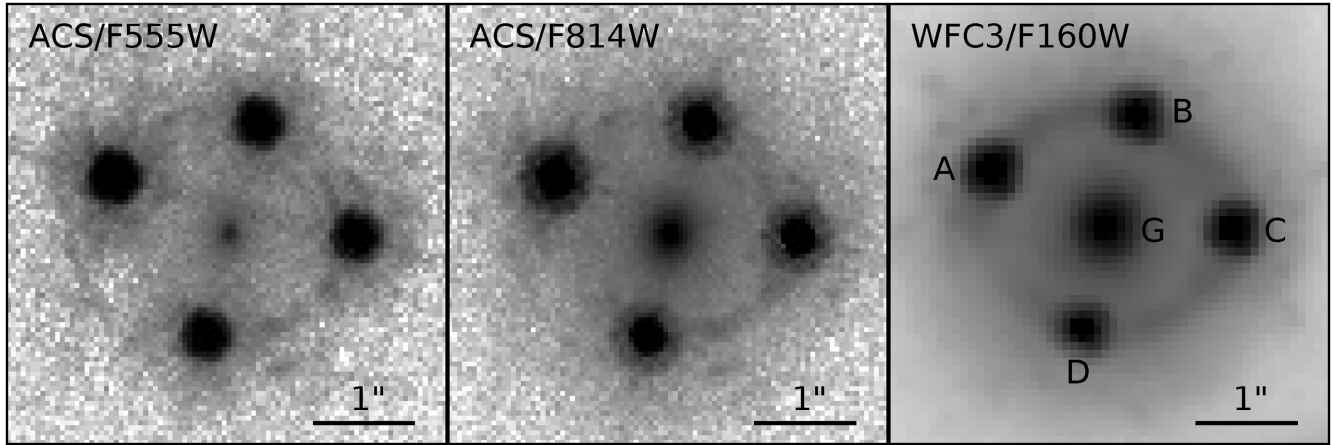
$$\begin{aligned} P(\mathbf{d}_{\text{HST}}, \Delta \mathbf{t}, \sigma, \mathbf{d}_{\text{LOS}} | \xi, \mathbf{A}) &= P(\mathbf{d}_{\text{HST}} | \xi, \mathbf{A}) \\ &\times P(\Delta \mathbf{t} | \xi, \mathbf{A}) \\ &\times P(\sigma | \xi, \mathbf{A}) \\ &\times P(\mathbf{d}_{\text{LOS}} | \xi, \mathbf{A}). \end{aligned} \quad (5)$$

We note that equation (5) assumes the approximation that the LOS can be decoupled from the lens model. We can calculate the individual likelihoods separately and combine them as in equation (5) to get the final posterior PDF for a given set of assumptions.

In Section 4.7, we lay out a range of systematics tests where we vary the content of  $\mathbf{A}$  and repeat the inference of  $\xi$ . Such a sensitivity analysis is important for checking the magnitude of various known but unmodelled systematic effects, but it leaves us with the question of how to combine the results. We note that the marginalization integral over these assumptions can be approximated as a sum as follows (denoting all four data sets by  $\mathbf{d}$ ),

$$\begin{aligned} P(\xi | \mathbf{d}) &= \int P(\xi | \mathbf{d}, \mathbf{A}) P(\mathbf{A} | \mathbf{d}) d\mathbf{A} \\ &\propto \sum_k P(\xi | \mathbf{d}, \mathbf{A}_k) P(\mathbf{d} | \mathbf{A}_k) \\ &\propto \sum_k P(\xi | \mathbf{d}, \mathbf{A}_k), \end{aligned} \quad (6)$$

provided the following two statements are true: first, that the prior PDF over possible assumptions is uniform, and that our sampling of possible assumptions is fair. We choose reasonable variations in the systematic effects to try to achieve this. The second is that the evidence  $P(\mathbf{d} | \mathbf{A}_k)$  does not change appreciably between inferences; this is likely to be true if the goodness of fit does not change, and the parameter priors and volumes are not very different. Under these assumptions, equation (6) shows that a *sum* of the posterior PDFs is



**Figure 1.** *HST* images of HE 0435–1223. Shown are cutouts of the lens system used for lens modelling in the ACS/*F555W* (left), ACS/*F814W* (middle) and WFC3/*F160W* (right) bands. The images are 4.5 arcsec on a side. The scale is indicated in the bottom right of each panel. The main lens galaxy (G) and lensed quasar images (A, B, C, D) are marked.

an approximation to the posterior PDF marginalized over the tested systematic effects.

### 3 DATA

HE 0435–1223 (J2000:  $4^{\text{h}}38^{\text{m}}14^{\text{s}}.9$ ,  $-12^{\circ}17'14''.4$ ) is a quadruply lensed quasar discovered by Wisotzki et al. (2002) as part of the Hamburg/ESO survey for bright QSOs (Wisotzki et al. 2000). The main deflector is a massive elliptical galaxy at a redshift of  $z_{\text{d}} = 0.4546 \pm 0.0002$  (Morgan et al. 2005), and the source redshift is  $z_{\text{s}} = 1.689$ .<sup>2</sup> Our spectroscopic observations reveal that the lens is a part of a galaxy group with a velocity dispersion of  $\sigma = 471 \pm 100 \text{ km s}^{-1}$  measured from 12 member galaxies (HOLiCOW Paper II), which is independently confirmed by Wong et al. (2011) and Wilson et al. (in preparation) based on a spectroscopic study by Momcheva et al. (2006, 2015). We present the *HST* imaging used for lens modelling in Section 3.1, the time delays measured by COSMOGRAIL in Section 3.2, the spectroscopy of the lens galaxy for measuring the lens stellar velocity dispersion in Section 3.3 and ground-based imaging and spectroscopy to characterize the lens environment in Section 3.4.

#### 3.1 *HST* imaging

We obtain deep *HST* observations of HE 0435–1223 using the Wide Field Camera 3 (WFC3) IR channel in the *F160W* band (Program #12889; PI: Suyu). The details of these observations are presented in HOLiCOW Paper I, which we summarize here. Using a combination of short (44 s) and long (599 s) exposures, we reconstruct the brightness distribution of both the lensed AGN and host galaxy. We reduce the images using DRIZZLEPAC.<sup>3</sup> The images are drizzled to a final pixel scale of 0.08 arcsec without masking the bright AGN pixels, as they are well characterized.

We also use archival observations from the Advanced Camera for Surveys (ACS) on *HST* in the *F555W* and *F814W* filters

(Program #9744; PI: Kochanek). The images are reduced using MULTIDRIZZLE<sup>4</sup> with charge transfer inefficiency taken into account (e.g. Anderson & Bedin 2010; Massey et al. 2010). The final pixel scale of the reduced images is 0.05 arcsec.

We create cutouts of the *HST* images around the lens and define an arcmask in each band in which we perform the modelling. For the ACS bands, we use a  $90 \times 90$  pixel cutout (4.5 arcsec on a side), and for the WFC3/*F160W* band, we use a  $60 \times 60$  pixel cutout (4.8 arcsec on a side). These cutouts are shown in Fig. 1.

To generate the initial point spread function (PSF) of the exposures, we first select three stars in the field that are close to the lens galaxy in angular separation to minimize CCD distortion effects, and which have approximately the same brightness as the lensed AGN images to avoid any PSF broadening effects. We then simultaneously fit these stars with a Moffat profile plus a regularized fine-pixel array. The exposures are sky-subtracted prior to the PSF fitting. The details of this fitting procedure are described by Cantale et al. (2016a) and are based on ideas presented in Magain, Courbin & Sohy (1998). A successful application of the procedure is presented by Cantale et al. (2016b). We then use this initial PSF as the starting point for our iterative PSF correction procedure (see Section 4.1).

The weight images are constructed as follows. We take a large, relatively sparse area of the image and approximate the background noise as the normalized median absolute deviation (NMAD), defined as  $\text{NMAD} \equiv 1.48 \times \text{median}(|p_i - \text{median}(p)|)$ , where  $p_i$  is the value of pixel  $i$  and  $\text{median}(p)$  is the median of all pixels in the selected area. We use the NMAD, which is a good approximation to the standard deviation, as it is less sensitive to outliers. We create a ‘noise image’ that has the same dimensions as the lens galaxy cutout with all pixels initialized to the value of the background noise. We then add Poisson noise to this noise image by taking all pixels in the lens galaxy cutout where the flux is greater than the background noise level and adding in quadrature the square root of each pixel value (normalized by its effective exposure time) to the corresponding pixel in the noise image (this is because the units of the science image are counts per second). The noise image is then squared and inverted to obtain the weight image. We note that

<sup>2</sup> We note that Sluse et al. (2012) measure an updated source redshift of  $z_{\text{s}} = 1.693$ . We use the original value of  $z_{\text{s}} = 1.689$  in our analysis but verify that using this updated measurement does not impact our results.

<sup>3</sup> DRIZZLEPAC is a product of the Space Telescope Science Institute, which is operated by AURA for NASA.

<sup>4</sup> MULTIDRIZZLE is a product of the Space Telescope Science Institute, which is operated by AURA for NASA.

while the background noise for the WFC3 IR camera depends on the number of non-destructive reads, we verify that the number of reads in the region of the lensed arc is the same as for the blank sky patch used for estimating the background noise, so this procedure is valid.

When modelling with these weights, there are large residuals near the AGN image centres due to our inability to model the PSF on a grid of pixels with sufficient accuracy. This can lead to biased results as the model will be influenced by these relatively small areas rather than the large-scale features of the source, so we compensate for this by reducing the weight in these regions (e.g. Suyu 2012). We scale the weight in these regions by a power law such that a pixel originally given a noise value of  $p_i$  is rescaled to a noise value of  $A \times p_i^b$ . The  $A$  and  $b$  are constants that are different for each band and are chosen such that the normalized residuals in the AGN image regions are approximately consistent with the normalized residuals in the rest of the arc region.

We note that in determining the effective exposure time on a pixel-by-pixel basis, we turn off the bad pixel masking in a  $3 \times 3$  pixel region around the brightest pixel of each of the lensed AGN images. This is done because allowing bad pixel masking results in interpolations of the image pixels that cause the four AGN images to exhibit different PSF profiles, which complicates our iterative PSF correction scheme (Section 4.1). Turning off the bad pixel mask produces more faithfully and consistently the four AGN images. Since the majority of the lens mass model constraints come from the lensing arcs away from the centres of the AGN images, we have checked that these arcs do not have bad pixels that would affect our lens mass model.

### 3.2 Time-delay measurements

Time-delay measurements for HE 0435–1223 were initially given in Courbin et al. (2011). Further monitoring of the system by COSMOGRAIL has since improved the time-delay accuracy and precision, completing the data from Courbin et al. (2011) with  $\sim 1300$  exposures of 6 min each for a total of 301 new observing nights ranging from 2010 September to 2016 April. The details of the data acquisition and time-delay measurements used in our analysis are presented in H0LiCOW Paper V, but we summarize the main results here.

The data treatment follows the procedure described by Tewes et al. (2013b). Each observing epoch is corrected following the standard reduction steps (bias subtraction, flat-fielding and sky correction). The PSF is estimated following the procedure described in Section 3.1. The exposures are then normalized using bright, non-saturated stars in the field of view. The photometry of the four images of HE 0435–1223 is obtained on each exposure using the Magain et al. (1998) deconvolution photometry presented in Cantale et al. (2016b). The light curves obtained with this method are presented in fig. 2 of H0LiCOW Paper V.

The measurement of the time delays between each pair of images follows the formalism introduced by Tewes, Courbin & Meylan (2013a). The common intrinsic variability of the quasar and the four independent extrinsic variability curves are fitted using free-knot splines. The curves are then shifted in time to optimize the fit. The uncertainties on the time-delay measurements are estimated using a Monte Carlo approach. A set of 1000 synthetic light curves are drawn, mimicking the light curves and the time-delay constraining power of the observed data (Tewes et al. 2013a). It is important that the synthetic data sets span a range of plausible true time delays, as this allows us to verify that the estimator accurately responds

to these input delays (i.e. does not suffer from lethargy, described in Rathna Kumar et al. 2015) and has not been involuntarily fine-tuned to recover a particular value of the time delay. Various tests on the data reduction process and curve-shifting technique have been performed successfully to ensure the reliability of the time-delay measurements. We use the time delays relative to image A:  $\Delta t_{AB} = -8.8 \pm 0.8$  d,  $\Delta t_{AC} = -1.1 \pm 0.7$  d and  $\Delta t_{AD} = -13.8 \pm 0.9$  d, where the uncertainties represent  $1\sigma$  confidence intervals.

### 3.3 Stellar velocity dispersion of lens galaxy

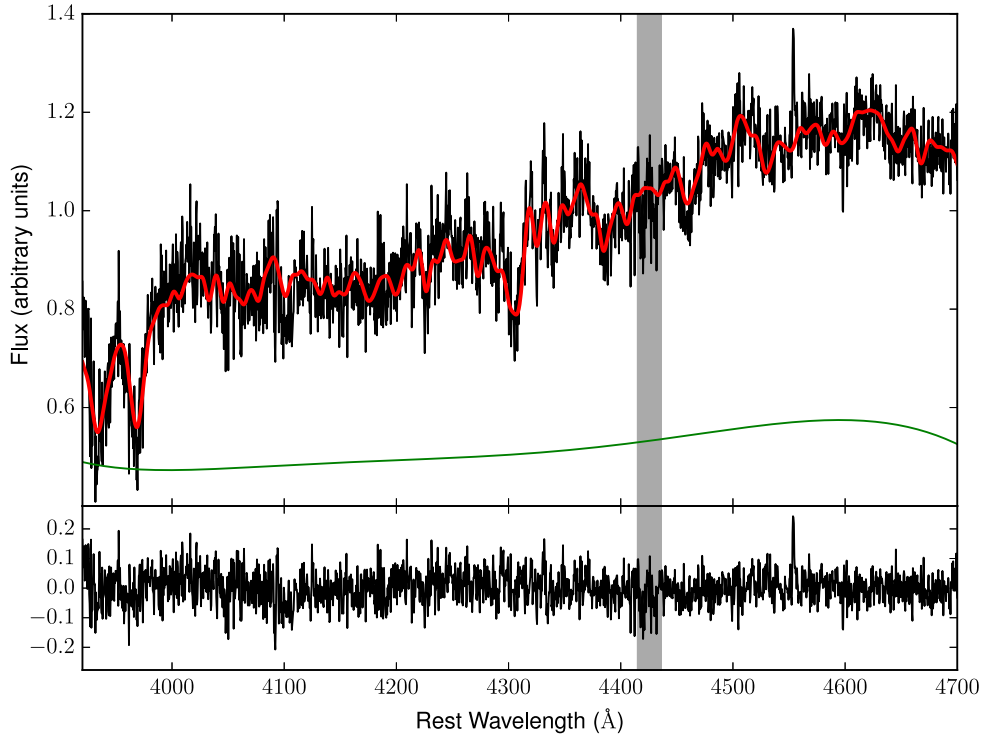
HE 0435–1223 was observed with the Low-Resolution Imaging Spectrometer (LRIS; Oke et al. 1995) on the Keck I telescope on 2011 January 4. Six exposures of 1200 s were obtained in 0.8 arcsec seeing with the red arm of the spectrograph using the 831/8200 grating, which has a dispersion of  $0.58 \text{ \AA pixel}^{-1}$  and yields an effective resolution  $\sigma_{\text{res}} \sim 37 \text{ km s}^{-1}$ . The 0.75 arcsec slit was oriented to intersect the eastern- and western-most lensed QSO images (i.e. at a position angle of  $76^\circ$ ) and a 4-pixel (0.54 arcsec) aperture was used to extract 1D spectra from each exposure. These six spectra were then resampled to a single spectrum using spline interpolation and rejecting pixels affected by cosmic rays or other artefacts; the resulting spectrum is shown in Fig. 2. The velocity dispersion was obtained following the same procedure as in Suyu et al. (2010, 2013), resulting in an inference of  $\sigma = 222 \text{ km s}^{-1}$  with a statistical uncertainty of  $11 \text{ km s}^{-1}$  and a systematic uncertainty of  $\sim 10 \text{ km s}^{-1}$  due to the templates used, the region of the spectrum that was fitted, and the order of the polynomial continuum. We therefore adopt an overall uncertainty of  $\sigma_\sigma = 15 \text{ km s}^{-1}$ . This measurement is in agreement with a previous determination of  $\sigma = 222 \pm 34 \text{ km s}^{-1}$  by Courbin et al. (2011) within a 1 arcsec aperture.

### 3.4 Lens environment: photometry and spectroscopy

To account for the effects of LOS structure, we have obtained deep multiband photometry and multi-object spectroscopy in the HE 0435–1223 field to characterize the external mass distribution. Details of the photometric observations and inference on  $\kappa_{\text{ext}}$  are presented in H0LiCOW Paper III, and the details of the spectroscopic data are presented in H0LiCOW Paper II, but we summarize the data here.

Our wide-field photometric data consist of ground-based *ugriJHK<sub>s</sub>* observations, as well as 3.6, 4.5, 5.8 and 8.0  $\mu\text{m}$  observations with the *Spitzer* Infrared Array Camera (IRAC; Fazio et al. 2004). We infer photometric redshifts and stellar masses using PSF-matched photometry measured with SExtractor. We use LEPHARE (Ilbert et al. 2006) to measure stellar masses for the best-fitting redshift using the spectral energy distribution (SED) templates employed by CFHTLenS (Velander et al. 2014), which assume a Chabrier (2003) initial mass function.

The wide-field spectroscopic data are taken with a combination of Keck/LRIS, the Focal Reducer/low-dispersion Spectrograph 2 (FOR2; Appenzeller et al. 1998) on the Very Large Telescope (VLT), and the Gemini Multi-Object Spectrograph (GMOS; Hook et al. 2004), and are combined with existing spectroscopic observations of this field (Momcheva et al. 2006, 2015). It is particularly important to model the most significant perturbers, as their effects may not be adequately accounted for by external shear alone (McCully et al. 2014). McCully et al. (2016) find that the most significant perturbers are those that are massive, projected close to the lens, and that are in the foreground of the lens redshift. H0LiCOW Paper II presents an estimate of the relative significance of nearby perturbers to HE 0435–1223 as quantified by their flexion shift,



**Figure 2.** Top: Keck/LRIS spectrum of HE 0435–1223 with the best-fitting model overplotted in red and a polynomial continuum, which accounts for contamination from the lensed QSO images and template mismatch, shown in green. We find that  $\sigma = 222 \pm 15 \text{ km s}^{-1}$ , including systematic uncertainties due to the templates used, the region of the spectrum that was fitted, and the order of the polynomial continuum. The grey vertical band represents a wavelength range that is excluded from the fit due to the presence of a strong Mg II absorption system. Bottom: residuals from the best fit.

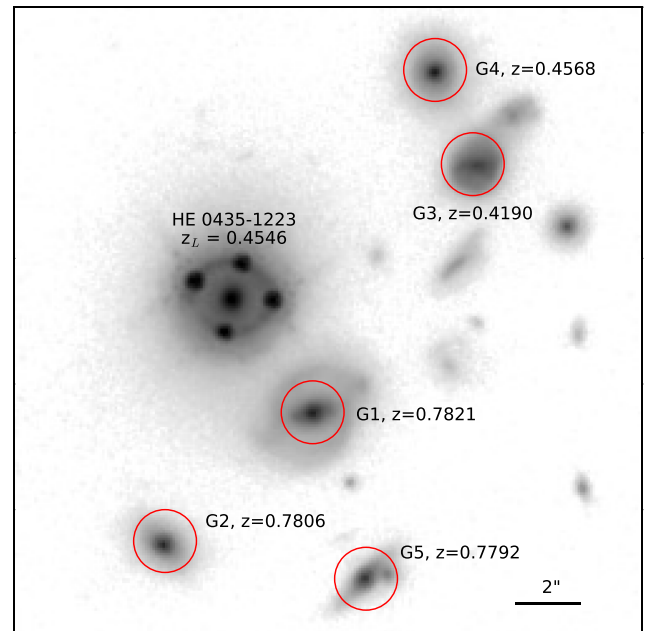
$\Delta_3x$  (McCully et al. 2016), and finds that at most, the five nearest perturbers should be accounted for explicitly, with all other perturbers having a negligible influence. Fig. 3 shows the lens and the relative positions and redshifts of these five perturbers, all brighter than  $i = 22.5$  mag and projected within 12 arcsec of the lens.

## 4 LENS MODELLING

In this section, we describe our procedure to simultaneously model the images in the three *HST* bands and the time delays to infer the lens model parameters.

### 4.1 Overview

We perform our lens modelling using *GLEE*, a software package developed by S. H. Suyu and A. Halkola (Suyu & Halkola 2010; Suyu et al. 2012b). The lensing mass distribution is described by a parametrized profile. The extended host galaxy of the source is modelled separately on a  $40 \times 40$  pixel grid with curvature regularization (Suyu et al. 2006). The lensed quasar images are modelled as point sources convolved with the PSF. By modelling the quasar images on the image plane independently from the extended host galaxy light distribution, we allow for variations in quasar fluxes due to microlensing, time delays and substructure. The lens galaxy light distribution is modelled using Chameleon profiles (defined as the difference of two non-singular  $r^{-2}$  elliptical profiles; Kassiola & Kovner 1993; Dutton et al. 2011), which are a good approximation to Sérsic profiles. We represent the galaxy light distribution as the sum of two Chameleon profiles with a common centroid. We use Chameleon profiles rather than Sérsic profiles because they provide a similarly good fit to the data (see Sections 4.7 and 5) and it is



**Figure 3.** *HST*/WFC3 *F160W* image of a  $20 \text{ arcsec} \times 20 \text{ arcsec}$  field around HE 0435–1223. The angular scale is indicated in the bottom right corner. The five most significant nearby perturbers are marked with red circles, and the redshifts of the perturbers are indicated. G1 is included explicitly in our model, as it is the most massive and nearest in projection to HE 0435–1223. We also test the effects of including the other perturbers as one of our systematics tests.

more straightforward to link their parameters to the mass parameters in our tests of alternative mass models. Model parameters of the lens and source are constrained through Markov Chain Monte Carlo (MCMC) sampling.

Since we account for G1 at a different redshift from the main lens galaxy, we make use of the full multiplane lens equation (e.g. Blandford & Narayan 1986; Kovner 1987; Schneider et al. 1992; Petters et al. 2001; Collett & Auger 2014; McCully et al. 2014) in our modelling. We vary  $H_0$  directly in our models and use this distribution to calculate the effective model time-delay distance  $D_{\Delta t}^{\text{model}}$ . In calculating  $D_{\Delta t}^{\text{model}}$ , we assume  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $w = -1$ , although we show that relaxing these assumptions shifts the resulting  $D_{\Delta t}^{\text{model}}$  distributions by  $<1$  per cent (Section 4.9).

## 4.2 Mass model

Our primary mass model for the lens galaxy is a singular power-law elliptical mass distribution (hereafter ‘SPEMD’; Barkana 1998), although we also test a model consisting of a baryonic component that traces the light distribution and a separate dark matter component (hereafter the ‘composite’ model; see Section 4.7). We also include an external shear in the strong lens plane. Past studies have shown that a power-law model provides a good general description of typical lens galaxies at the length-scales we are interested in (e.g. Koopmans et al. 2006, 2009; Suyu et al. 2009; Auger et al. 2010; Barnabè et al. 2011; Sonnenfeld et al. 2013).

We also explicitly include the most nearby massive perturbing galaxy [G1 in Fig. 3;  $z = 0.7821$ ,  $\log(M_*/M_\odot) = 10.9$ ] that is projected  $\sim 4.5$  arcsec away from the lens, which is close enough that its influence may not be adequately described by external shear (HOLiCOW Paper II; see also McCully et al. 2016). G1 is modelled as a singular isothermal sphere, which is a reasonable assumption as higher order moments of its potential will have a small influence at the position of the main lens galaxy. G1 is treated using the full multiplane lens equation, as detailed by Suyu et al. (in preparation).

Our SPEMD model has the following free parameters:

- (i) position  $(\theta_1, \theta_2)$  of the centroid (allowed to vary independently from the centroid of the light distribution);
- (ii) *Einstein* radius  $\theta_E$ ;
- (iii) minor-to-major axis ratio  $q$  and associated position angle  $\theta_q$ ;
- (iv) 3D slope of the power-law mass distribution  $\gamma'$ ;
- (v) external shear  $\gamma_{\text{ext}}$  and associated position angle  $\theta_\gamma$ ;<sup>5</sup>
- (vi) *Einstein* radius of G1;
- (vii) the cosmological parameter  $H_0$ .

In principle, our lens is drawn from a selection function and the choice of model priors may introduce a bias on the inferred time-delay distance. However, since the selection function is not well known and these biases are negligibly small for an analysis like ours (Collett & Cunningham 2016), we conservatively assume uniform priors on the model parameters.

To get a starting point for our model, we run a preliminary model where only the positions and time delays of the lensed quasar images are used as constraints and G1 is not included. This preliminary model is fast and easy to optimize, and we use the output parameters as the initial parameters of our primary model.

Our constraints on the primary lens model include the positions of the lensed quasar images, the measured time delays and the sur-

face brightness of the pixels in the ACS/F555W, ACS/F814W and WFC3/F160W images that are fit simultaneously. We first model the lens system individually in each band to iteratively update the PSFs using the lensed AGN images themselves in a manner similar to Chen et al. (2016), but with the PSF corrections and source intensity reconstructed simultaneously in our case rather than separately (Suyu et al., in preparation). We then fix these ‘corrected’ PSFs and use them in our final models that simultaneously use the surface brightness distribution in all three bands as constraints. We do not enforce any similarity of pixel values at the same spatial position across different bands. In our MCMC sampling, we vary the light parameters of the lens galaxy and quasar images, the mass parameters of the lens galaxy, the external shear, the *Einstein* radius of G1 and  $H_0$ . The quasar positions are linked across all three bands, but the other light parameters are allowed to vary independently.

Fig. 4 shows the data and the lens model results in each of the three bands, as well as the source reconstruction. Our model reproduces the surface brightness structure of the lensed AGN and host galaxy in all three bands. There are some small residuals in the region of the lensed arc away from the AGN images. We attribute these to compact star formation regions in the host galaxy, as our model maps these features to similar locations in the source plane. We test a model where the region near these residuals are masked out and find that our  $D_{\Delta t}$  inference is consistent to within our systematic uncertainties (Section 4.7).

## 4.3 Kinematics

We follow Suyu et al. (2010) and Sonnenfeld et al. (2012) to compute the LOS stellar velocity dispersion of the strong lens galaxy through the spherical Jeans equation (see also Treu & Koopmans 2002; Koopmans et al. 2003). For a given lens model, we obtain the 3D density profile of the lens galaxy by taking the spherical deprojection of the circularized surface mass density profile. The resulting 3D density profile assumes an analytical form for both the power-law and the composite model. The 3D distribution of tracers is obtained by applying the same procedure to the surface brightness distribution of the lens galaxy, which we model as a Hernquist (1990) profile. We also tested a Jaffe (1983) profile that has been shown to produce similar results (Suyu et al. 2010), and find that the results are affected by less than 1 per cent level. We parametrize the orbital anisotropy profile with an Osipkov–Merritt model (Osipkov 1979; Merritt 1985)

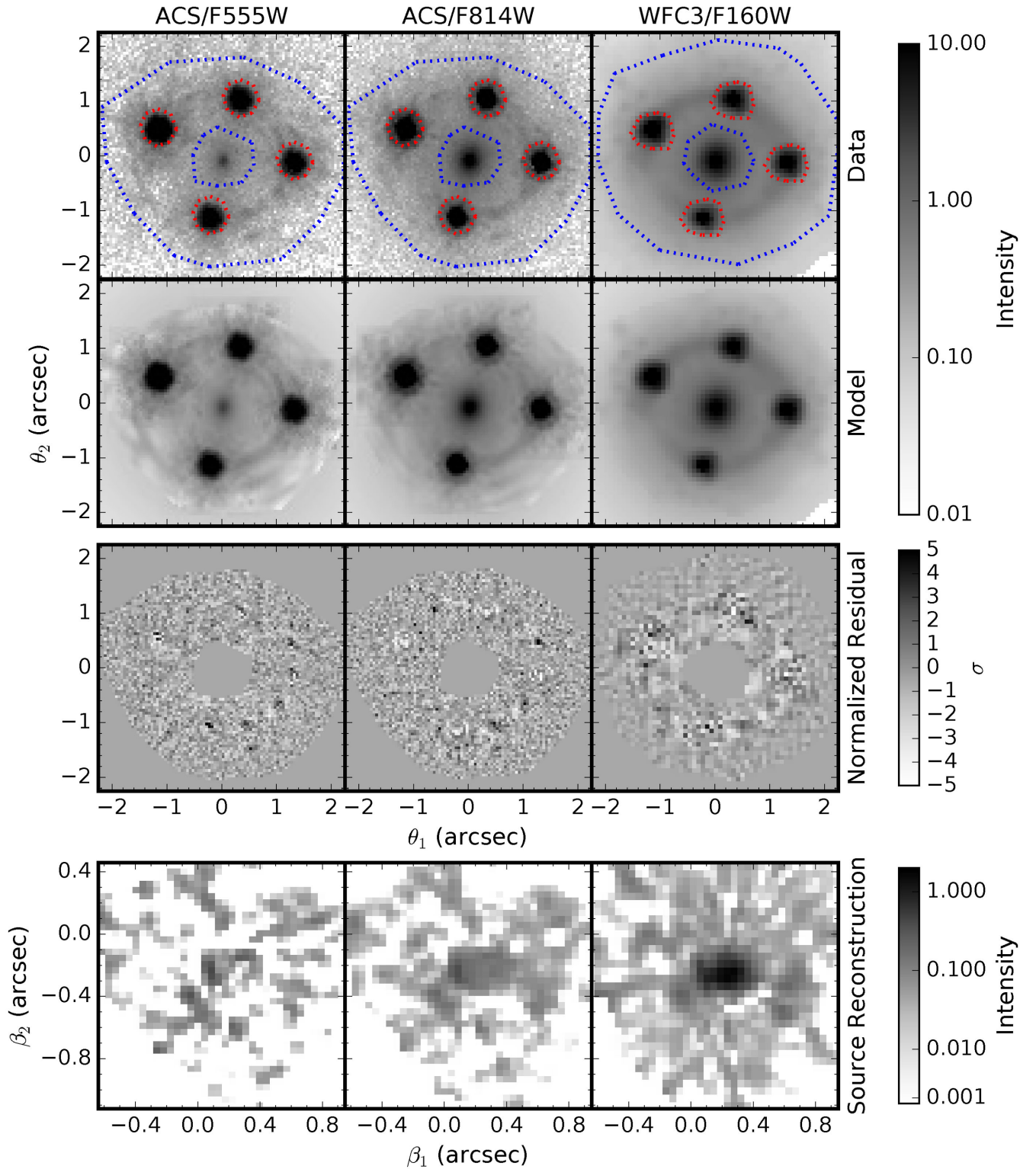
$$\frac{\sigma_\theta^2}{\sigma_r^2} = 1 - \frac{r^2}{r_{\text{ani}}^2 + r^2}. \quad (7)$$

Given values of the lens mass parameters in Section 4.2, the external convergence  $\kappa_{\text{ext}}$  in Section 4.4 and the anisotropy radius  $r_{\text{ani}}$ , we then calculate the LOS velocity dispersion profile by numerically integrating the solutions of the spherical Jeans equation as given by Mamon & Łokas (2005). Finally, we calculate the integral over the spectroscopic slit of the seeing-convolved brightness-weighted LOS velocity dispersion  $\sigma^P$  (equation 20 of Suyu et al. 2010), which we then compare to the measurements to compute the likelihood of the kinematics data,

$$P(\sigma | \mathbf{v}, \boldsymbol{\pi}, \kappa_{\text{ext}}, r_{\text{ani}}) = \frac{1}{\sqrt{2\pi}\sigma_\sigma} \exp \left[ -\frac{(\sigma^P(\mathbf{v}, \boldsymbol{\pi}, \kappa_{\text{ext}}, r_{\text{ani}}) - \sigma)^2}{2\sigma_\sigma^2} \right], \quad (8)$$

where  $\sigma = 222 \text{ km s}^{-1}$  and  $\sigma_\sigma = 15 \text{ km s}^{-1}$  (Section 3.3). We adopt a uniform prior on  $r_{\text{ani}}$  between 0.5 and 5 times the effective radius,

<sup>5</sup>  $\theta_\gamma$  is defined to be the direction of the shear itself, i.e. orthogonal to the direction of the mass producing the shear.



**Figure 4.** Lens model results for ACS/F555W (left), ACS/F814W (middle) and WFC3/F160W (right). Shown are the observed image (top row), the reconstructed image predicted by the model (second row), the normalized residual within the arc-mask region (defined as the difference between the data and model, normalized by the estimated uncertainty of each pixel; third row) and the reconstructed source (bottom row). The blue dotted lines indicate the arc-mask region used for fitting the extended source, and the red dotted lines indicate the AGN mask region where the power-law weighting is applied. The colour bars show the scale in the respective panels. The results shown here are for the fiducial model, but the results for the other systematics tests (Section 4.7) are qualitatively similar.



$r_{\text{eff}}$ , which we determine to be  $r_{\text{eff}} = 1.33$  arcsec from our lens light fitting<sup>6</sup> in the *F814W* filter.

#### 4.4 External convergence

In *HOLiCOW Paper III*, we estimate the external convergence using weighted number counts in a manner similar to Greene et al. (2013, see also Fassnacht et al. 2011). We use the weighted counts to select corresponding lines of sight from the  $\kappa_{\text{ext}}$  catalogues produced from the Millennium Simulation by Hilbert et al. (2009) and, thus, to get a  $\kappa_{\text{ext}}$  distribution. We use the  $\kappa_{\text{ext}}$  distribution from *HOLiCOW Paper III* that was derived by combining three constraints: the unweighted galaxy number counts, the counts weighted by  $1/r$ , and external shear matching that from our lens modelling, which gives a median external convergence at the position of HE 0435–1223 of  $\kappa_{\text{ext}} = 0.003$ , with 16 per cent and 84 per cent percentiles of  $\kappa_{\text{ext}} = -0.016$  and  $\kappa_{\text{ext}} = 0.034$ , respectively. Although the external shear can change slightly among different models, these changes generally affect the  $\kappa_{\text{ext}}$  distribution at the  $\sim 0.005$  level or smaller, which we can safely neglect. Since we are explicitly including the nearest LOS perturber in our mass model, this galaxy does not contribute to the inferred external shear, nor do we want it to be double counted in the external convergence. We therefore exclude galaxies projected within 5 arcsec of the main lens galaxy when calculating the relative galaxy number counts<sup>7</sup> for both the simulated and real lines of sight.

The host galaxy group likely has a small effect as the external shear is small, and an estimate of its flexion shift (*HOLiCOW Paper II*) indicates that it is a less significant perturber than G1. In addition, a weak lensing analysis of the field (Tihhonova et al., in preparation) finds a conservative  $3\sigma$  upper limit of  $\kappa_{\text{ext}} = 0.04$  at the lens position, further suggesting that the group does not significantly affect our analysis. The external convergence contribution of the host galaxy group is implicitly included in our model through the procedure of *HOLiCOW Paper III*.

#### 4.5 Blind analysis

Throughout our analysis, we blind the  $H_0$  values in our lens model and the inferred time-delay distance values to avoid confirmation bias using a similar procedure as Suyu et al. (2013). This is done by subtracting the median of the parameter PDFs from the distribution when displaying plots. This allows us to measure the precision and relative offsets of these parameter distributions and their correlation with other parameters without being able to see the absolute value. This eliminates the tendency for experimenters to stop investigating systematic errors when they obtain an answer consistent with the ‘expected’ result. After finalizing our analysis, writing our paper draft with blinded  $D_{\Delta t}$  distributions, and coming to a consensus among the coauthors during a collaboration telecon on 2016 June 16, we unblind the results and do not make any further changes to the models. There is also no iteration between the lens modelling and time delay measurements (i.e. the delays are measured once and

used as they are; see *HOLiCOW Paper V*). Throughout this paper, we show blinded  $D_{\Delta t}$  distributions until Section 5, where we reveal the absolute  $D_{\Delta t}$  values from our inference.

#### 4.6 Inferring the time-delay distance

Our inference on  $D_{\Delta t}$  using all of the available data is calculated as in equations (4) and (5). We use importance sampling (e.g. Lewis & Bridle 2002) to combine the velocity dispersion and external convergence distributions with the  $D_{\Delta t}^{\text{model}}$  inferred from our lens model. Specifically, for each set of lens parameters  $\nu$  from our lens mass model MCMC chain, we draw a sample of  $\kappa_{\text{ext}}$  from the distribution in Section 4.4 and a sample of  $r_{\text{ani}}$  from the uniform distribution  $[0.5, 5]r_{\text{eff}}$ . With these, we can then compute the kinematics likelihood in equation (8) for the joint sample  $\{\nu, \kappa_{\text{ext}}, r_{\text{ani}}\}$  and use this to weight the joint sample. From the effective *model* time-delay distance computed from our multiplane lensing ( $D_{\Delta t}^{\text{model}}$ ) and the external convergence ( $\kappa_{\text{ext}}$ ), we can then compute the effective time-delay distance ( $D_{\Delta t}$ ) via equation (3), keeping its absolute value blinded until we finalize our analysis. The resulting distribution of  $D_{\Delta t}$  encapsulates the cosmological information from HE 0435–1223.

#### 4.7 Systematics tests

In this section, we describe a range of tests of the effects of various systematics in our modelling. In addition to a basic ‘fiducial’ model, we perform inferences given the following sets of assumptions:

- (i) a model with the image plane cutout region in all bands increased by 10 pixels in both the  $\theta_1$  and  $\theta_2$  directions.
- (ii) A model with the arcmask region increased by 1 pixel on both the inner and outer edges. To compensate for the larger arcmask region, we increase the source plane resolution to  $50 \times 50$  pixels in all bands.
- (iii) A model with the arcmask region increased by 2 pixels on both the inner and outer edges. To accommodate the larger arcmask, we also increase the image plane cutout region by 10 pixels in all bands. To compensate for the larger arcmask region, we increase the source plane resolution to  $50 \times 50$  pixels in all bands.
- (iv) A model where the regions near the AGN images are given zero weight rather than being scaled by a power-law weighting.
- (v) A model where the regions near the AGN images scaled by the power-law weighting is increased by 1 pixel around the outer edge.
- (vi) A model where the regions near the AGN images scaled by the power-law weighting is increased by 2 pixels around the outer edge.
- (vii) A model where the light profile of the lens galaxy is represented by the sum of two Sérsic profiles rather than the sum of two Chameleon profiles.
- (viii) A model including the five most significant nearby perturbers (shown in Fig. 3) rather than just G1. The relative *Einstein* radii of the perturbers, assumed to be singular isothermal spheres, are calculated from their stellar masses (*HOLiCOW Paper III*), assuming a relationship between velocity dispersion and stellar mass from Bernardi et al. (2011). The ratio of *Einstein* radii is fixed, but with a global scaling allowed to vary freely. This is done to prevent the model from optimizing the perturbers’ *Einstein* radii in a way that would be inconsistent with their measured redshifts and stellar masses. The galaxies’ stellar masses are computed assuming the cosmology of the Millennium Simulation ( $H_0 = 73 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,

<sup>6</sup> We use the double Sérsic model of the lens galaxy light to determine  $r_{\text{eff}}$  because the Chameleon profile does not provide an accurate description at large radii.

<sup>7</sup> For our model that includes the five nearest perturbers, we run a test where we calculate  $\kappa_{\text{ext}}$  excluding a larger region. The corresponding shift in  $\kappa_{\text{ext}}$  affects our final  $D_{\Delta t}$  distribution by  $\sim 0.2$  per cent at most, so we neglect this effect.

$\Omega_m = 0.25$ ,  $\Omega_\Lambda = 0.75$ ; Springel et al. 2005; Hilbert et al. 2009), but we verify that for alternative cosmologies, their stellar masses change by  $<0.02$  dex, and the ratios of their *Einstein* radii therefore are affected by a negligible amount.

(ix) A ‘composite’ model with separate stellar and dark matter components. The details of this model are discussed in Section 4.8.

(x) The composite model with the regions near the AGN images scaled by the power-law weighting increased by 1 pixel around the outer edge.

(xi) The composite model with the arc-mask region increased by 1 pixel on both the inner and outer edges and a  $50 \times 50$  pixel source plane resolution.

As described in Section 2.2, we combine the MCMC chains from all of these tests. In doing so, we effectively assume that (1) these various tests sample a reasonable distribution of assumptions that we could have made when modelling the system, and that these assumptions have equal prior probability, and (2) neither the goodness of fit nor the parameter space prior volume are appreciably different between the tests. We verify that the goodness of fit does not change appreciably during this procedure (see Section 5). We weight the different MCMC chains equally and concatenate them, resulting in a set of samples that characterizes our final posterior PDF for  $D_{\Delta t}$ . This procedure folds the systematic uncertainty due to our modelling assumptions into our final uncertainty on the inferred parameters.

#### 4.8 Comparison of power-law and composite models

We follow Suyu et al. (2014) to construct the composite model of baryons and dark matter as one of our systematic tests. The composite model consists of mass components associated with each of the four non-singular isothermal elliptical profiles (making up the two Chameleon profiles) in the lens galaxy light model in the WFC3/*F160W* band scaled by an overall mass-to-light (*M/L*) ratio. We use the *F160W* band because it probes the rest-frame near-infrared and thus should be the best tracer of stellar mass. The dark matter component is modelled as an elliptical NFW (Navarro, Frenk & White 1996) potential with its centroid linked to that of the light centroid in *F160W*. This is motivated by Dutton & Treu (2014), who find that non-contracted NFW profiles are a good representation for the dark matter haloes of massive elliptical galaxies.

The composite model has the following free parameters:

- (i) *M/L* ratio for the baryonic component;
- (ii) NFW halo normalization  $\kappa_{0,h}$  (defined as  $\kappa_{0,h} \equiv 4\kappa_s$ ; Golse & Kneib 2002);
- (iii) NFW halo scale radius  $r_s$ ;
- (iv) NFW halo minor-to-major axis ratio  $q$  and associated position angle  $\theta_q$ ;
- (v) external shear  $\gamma_{\text{ext}}$  and associated position angle  $\theta_\gamma$ ;
- (vi) *Einstein* radius of G1;
- (vii) the cosmological parameter  $H_0$ .

We set a Gaussian prior of  $r_s = 14.3 \text{ arcsec} \pm 2.0 \text{ arcsec}$  based on the results of Gavazzi et al. (2007) for lenses in the Sloan Lens ACS Survey (SLACS; Bolton et al. 2006) sample, which encompasses the redshift and stellar mass of HE 0435–1223. All other parameters are given uniform priors. We note that the relative amplitudes of the two Chameleon profiles representing the stellar light distribution of the lens galaxy can vary during the modelling, whilst the relative amplitudes are fixed in the mass profiles. To account for this, we adopt an iterative approach where we run

a series of MCMC chains and update the (fixed) relative amplitudes of the associated mass components to match that of the light components after each chain. We iterate until the relative change in the light profile amplitudes reach a point where the inferred  $D_{\Delta t}$  stabilizes, then combine the MCMC chains after this point into a single distribution to represent the composite model. The remaining two composite models (with a larger arc-mask or AGN mask) use fixed relative amplitudes of the mass components from the latest iteration of the original composite model.

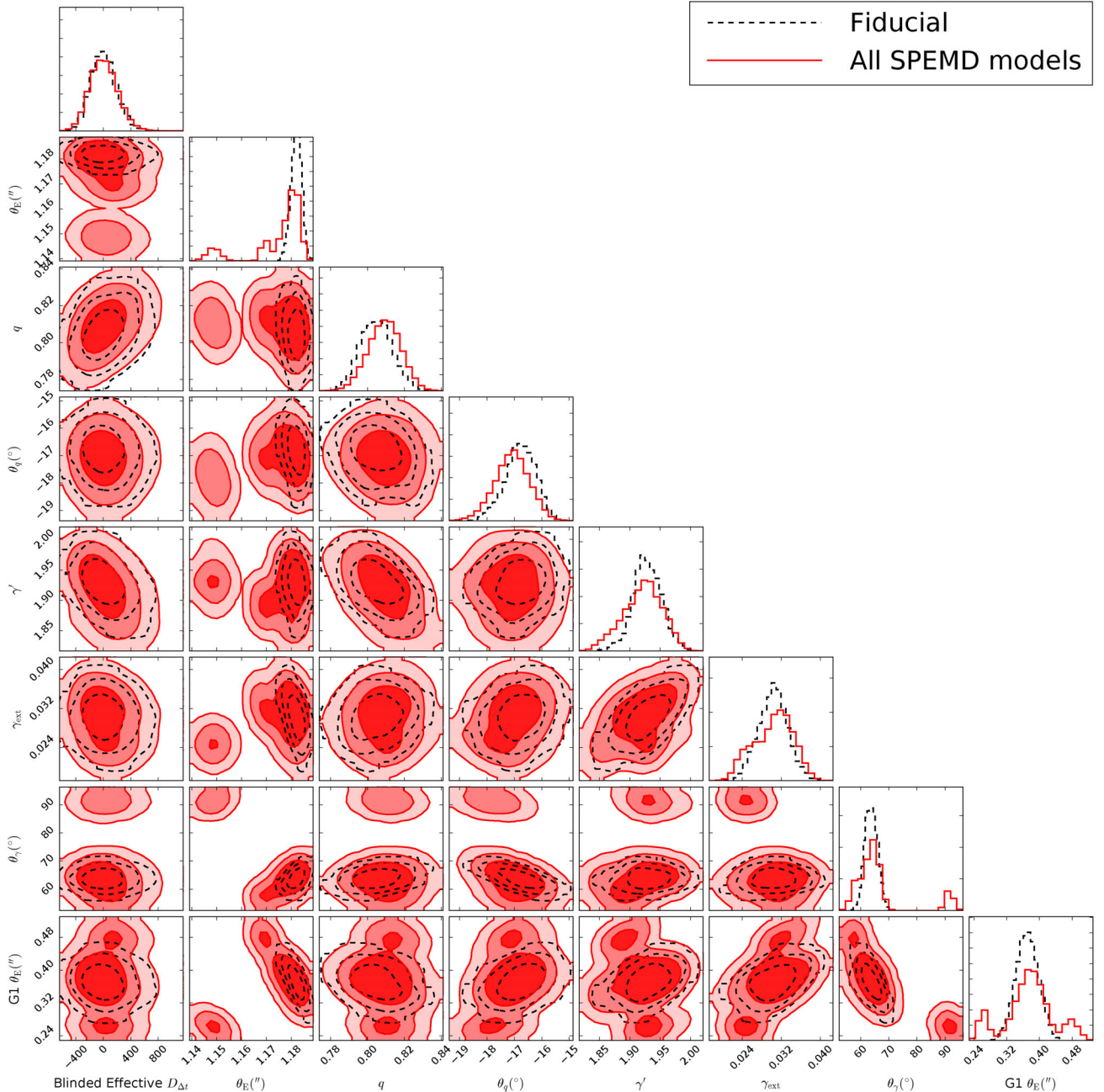
The marginalized parameter distributions of the SPEMD model are shown in Fig. 5. We show the combined distributions of all SPEMD models as well as the fiducial model separately. The parameter statistics for each model are given in Appendix A. We note that two particular models stand out. The model with the arc-mask expanded by 1 pixel and a  $50 \times 50$  source grid prefers a smaller *Einstein* radius for the main lens galaxy and a larger *Einstein* radius for G1. This degeneracy is likely due to systematics associated with the source pixel size (Suyu et al. 2013), as this model has a smaller source pixel size than the others. The 5-perturber model prefers a smaller *Einstein* radius for both the main lens galaxy and G1, as well as a very different  $\theta_\gamma$ . This is not surprising, as the addition of the extra perturbers in the lens model contributes to the integrated LOS lensing effect, reducing the contribution needed from the main lens and G1, as well as changing the external shear needed to fit the data. The offset between the mass centroid and the light centroid in the *F160W* band is typically  $\sim 0.002$  arcsec.

We show the marginalized parameter distributions of the composite model in Fig. 6. Again, we show the combined distributions as well as the main composite model separately, and the parameter statistics for each model are given in Appendix A. The main composite model appears to have some degenerate or bimodal features, but this is because this model itself is the combination of several separate models with slightly different relative amplitudes between the two Chameleon components, as mentioned above. The model with a larger arc-mask and source grid prefers a larger G1 *Einstein* radius, similar to the analogous SPEMD model. The dark matter fraction within the *Einstein* radius for the composite models is  $f_{\text{DM}} \sim 45$  per cent.

We compare the physical parameters of our ‘fiducial’ power-law model to the composite model. The results are shown in Table 1, with the parameter statistics for all composite models given in Appendix A. We note that the external shear strength of the composite model is smaller than that of the power-law model, which we attribute to a degeneracy between  $\gamma_{\text{ext}}$  and the internal ellipticity of the mass model. When external shear is removed, the composite model’s critical curves appear slightly more elliptical than those of the power-law model, supporting this interpretation. As mentioned in Section 4.4, the difference in  $\gamma_{\text{ext}}$  between these models has a negligible effect on the  $\kappa_{\text{ext}}$  distribution.

#### 4.9 Impact of different cosmologies

In the multilens-plane modelling, we need to sample the cosmological parameters in order to carry out the ray tracing. Throughout our analysis, we only vary  $H_0$ , keeping other cosmological parameters fixed ( $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $w = -1$ ). This is done for computational reasons, as the MCMC sampling becomes inefficient when they are all allowed to vary simultaneously. In principle,  $D_{\Delta t}$  has a weak dependence on these other cosmological parameters. We test their impact by rerunning the fiducial model while allowing combinations of them to vary with uniform priors. The resulting effective  $D_{\Delta t}$  distributions, shown in Fig. 7, have peaks that are consistent



**Figure 5.** Marginalized parameter distributions from our SPEMD lens model results. We show the fiducial model (dashed black contours) and the combined results from our systematics tests (shaded red contours). The contours represent the 68.3 per cent, 95.4 per cent and 99.7 per cent quantiles.

to within 1 per cent of the absolute value, demonstrating that the results are insensitive to these extra cosmological parameters at the level of accuracy that we are currently working at. In future, when errors shrink further, this sampling will be included. We conclude that with the current level of precision, we are justified in deriving the posterior distribution function of the time-delay distance by varying  $H_0$  only for computational efficiency. We emphasize that this does not affect in any way the generality of our results and that the resulting posterior distribution function is robust and can be interpreted in any cosmological model.

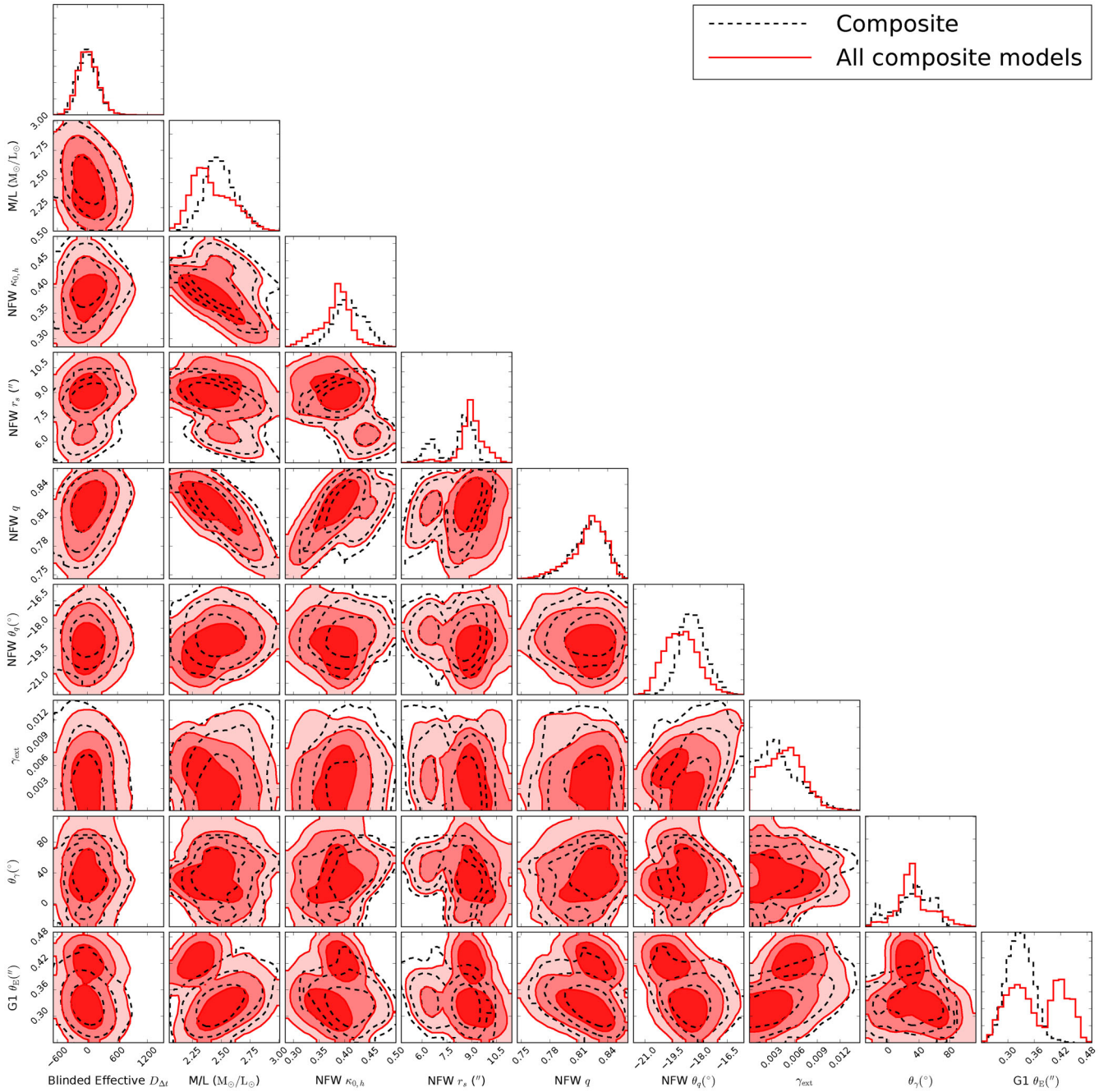
To expand on this point, it is instructive to consider multiplane lensing (e.g. Blandford & Narayan 1986; Kovner 1987; Kochanek & Apostolakis 1988; Schneider et al. 1992; Petters et al. 2001; Collett

& Auger 2014; McCully et al. 2014; Schneider 2014) for the case of two lens planes, as we have in most of our models. Defining  $\theta_1$ ,  $\theta_2$  and  $\theta_3$  as the angular coordinates on the main lens plane, the G1 lens plane and the source plane, respectively, the multiplane lens equations in this case are

$$\theta_2 = \theta_1 - \frac{D_{12}}{D_2} \hat{\alpha}_1(D_1\theta_1), \quad (9)$$

$$\theta_3 = \theta_1 - \frac{D_{13}}{D_3} \hat{\alpha}_1(D_1\theta_1) - \frac{D_{23}}{D_3} \hat{\alpha}_2(D_2\theta_2), \quad (10)$$

where  $D_i$  is the angular diameter distance from the observer to plane  $i$ ,  $D_{ij}$  is the angular diameter distance between planes  $i$  and  $j$  and



**Figure 6.** Marginalized parameter distributions from our composite lens model results. We show the main composite model (dashed black contours) and the combined results from our systematics tests (shaded red contours). The contours represent the 68.3 per cent, 95.4 per cent and 99.7 per cent quantiles.

$\hat{\alpha}_i$  is the deflection angle at plane  $i$ . Scaling the deflection angle relative to the source (third) plane, we have the scaled deflection angle as

$$\alpha_i(\theta_i) = \frac{D_{i3}}{D_3} \hat{\alpha}_i(D_i \theta_i). \quad (11)$$

By further defining

$$\beta_{ij} = \frac{D_{ij}}{D_j} \frac{D_3}{D_{i3}}, \quad (12)$$

we can rewrite equations (9) and (10) as

$$\theta_2 = \theta_1 - \beta_{12} \alpha_1(\theta_1), \quad (13)$$

$$\theta_3 = \theta_1 - \alpha_1(\theta_1) - \alpha_2(\theta_2). \quad (14)$$

The multiplane time delay has contributions from the geometric delays between planes and the gravitational delay at each mass plane:

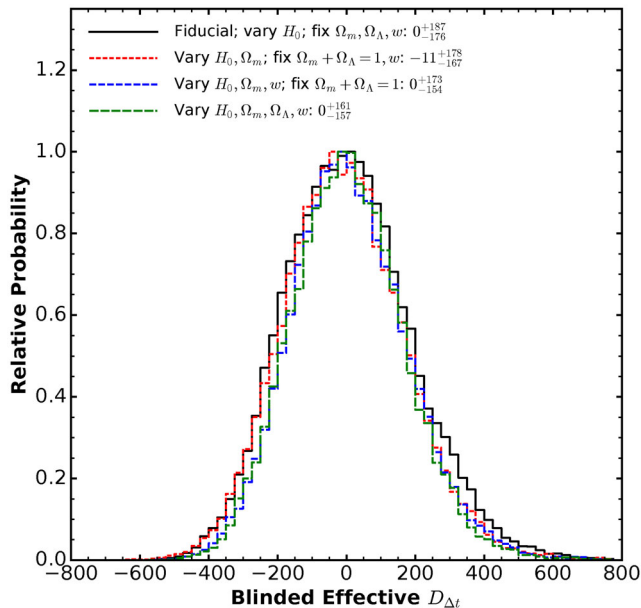
$$t = \frac{D_{\Delta t}^2}{c} \left[ \frac{1}{2} |\theta_2 - \theta_1|^2 - \beta_{12} \psi_1(\theta_1) \right] + \frac{D_{\Delta t}^{23}}{c} \left[ \frac{1}{2} |\theta_3 - \theta_2|^2 - \psi_2(\theta_2) \right], \quad (15)$$

**Table 1.** Lens model parameters. Reported values are medians, with errors corresponding to the 16th and 84th percentiles. Angles are measured east of north.

Parameter	Marginalized constraints
Fiducial singular power-law ellipsoid model	
$\theta_E$ (arcsec) <sup>a</sup>	$1.182^{+0.002}_{-0.002}$
$q$	$0.80^{+0.01}_{-0.01}$
$\theta_q$ (°)	$-16.8^{+0.5}_{-0.6}$
$\gamma'$	$1.93^{+0.02}_{-0.02}$
$\gamma_{\text{ext}}$	$0.030^{+0.003}_{-0.003}$
$\theta_\gamma$ (°)	$63.7^{+2.4}_{-2.2}$
G1 $\theta_E$ (arcsec)	$0.37^{+0.03}_{-0.03}$
Composite Model	
Stellar $M/L$ ( $M_\odot/L_\odot$ ) <sup>b</sup>	$2.5^{+0.1}_{-0.1}$
NFW $\kappa_{0,h}$	$0.41^{+0.03}_{-0.03}$
NFW $r_s$ (arcsec)	$8.43^{+0.58}_{-1.94}$
NFW $q$	$0.82^{+0.01}_{-0.02}$
NFW $\theta_q$ (°)	$-18.4^{+0.7}_{-0.7}$
$\gamma_{\text{ext}}$	$0.004^{+0.003}_{-0.002}$
$\theta_\gamma$ (°)	$34.4^{+22.9}_{-32.5}$
G1 $\theta_E$ (arcsec)	$0.33^{+0.03}_{-0.03}$

Notes. <sup>a</sup>Spherical-equivalent *Einstein* radius.

<sup>b</sup> $M/L$  within  $\theta_E$  for rest-frame *V* band. The given uncertainties are only statistical and do not include systematic effects. The stellar mass is calculated assuming  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.3$  and  $\Omega_\Lambda = 0.7$ , but changes in the cosmology affect the  $M/L$  by a negligible amount.



**Figure 7.** PDF of  $D_{\Delta t}$  for the various cosmologies. We compare the fiducial model to one in which  $\Omega_m$  is allowed to vary (with  $\Omega_m + \Omega_\Lambda = 1$ ), one in which  $w$  is also allowed to vary, and one in which  $\Omega_m$ ,  $\Omega_\Lambda$  and  $w$  are all allowed to vary independently. The distributions are blinded by subtracting the median of the fiducial model PDF. The different cosmology tests are indicated by the legend, and the median and 68 percent quantiles of the  $D_{\Delta t}$  distributions are given. The median of the blinded effective time-delay distance PSF is insensitive to the extra cosmological parameters to within 1 percent.

**Table 2.** Effective time-delay distance. Reported values are medians, with errors corresponding to the 16th and 84th percentiles.  $\chi^2$  values are computed within the fiducial arcmask and outside the AGNmask+2pix region for a fair comparison among models. The models are grouped such that those that use the same data set are together. For models with a larger arcmask, we calculate  $\chi^2$  for a source grid resolution that approximately matches that of the other models so that we can fairly compare them.

Model	$D_{\Delta t}$ (Mpc)	$\chi^2$
Fiducial	$2532^{+187}_{-176}$	11 024.9
Sérsic profiles	$2722^{+209}_{-185}$	11 001.5
5 perturbers	$2642^{+187}_{-173}$	11 002.0
Composite	$2646^{+202}_{-188}$	11 014.1
AGN mask+1pix	$2507^{+189}_{-171}$	11 029.6
Composite,AGN mask+1pix	$2656^{+211}_{-194}$	11 032.2
Arcmask+1pix,50x50src	$2741^{+170}_{-150}$	11 097.7
Composite,Arcmask+1pix,50x50src	$2665^{+195}_{-171}$	11 121.5
img+10pix	$2532^{+184}_{-146}$	11 090.2
Arcmask+2pix,img+10pix,50x50src	$2636^{+178}_{-152}$	11 074.3
AGN mask weight=0	$2518^{+195}_{-181}$	10 921.6
AGN mask+2pix	$2528^{+187}_{-166}$	11 065.4
<b>Total</b>	<b><math>2612^{+208}_{-191}</math></b>	–

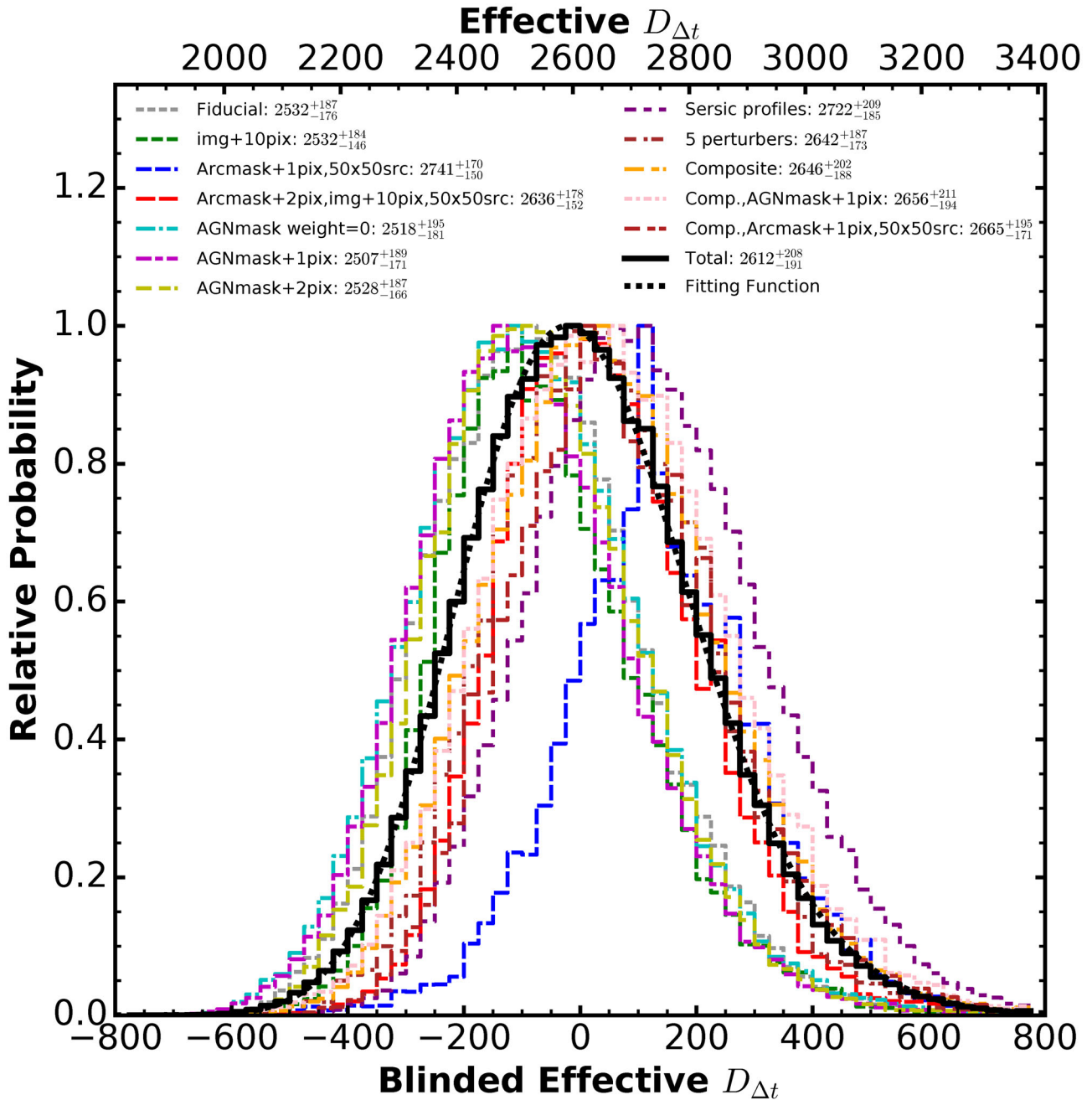
where  $\psi_i$  is the lens potential related to the scaled deflection angle via  $\nabla\psi_i = \alpha_i$ , and the time-delay distances between planes are

$$D_{\Delta t}^{ij} \equiv (1 + z_i) \frac{D_i D_j}{D_{ij}}, \quad (16)$$

with  $z_i$  being the redshift of plane  $i$ . From equation (15), we see that the time delay depends on the two time-delay distances and  $\beta_{12}$ . In general, it is difficult to constrain all these distance quantities independently. In fact, in multiplane modelling, we adopt specific cosmological models to compute the distances ( $D_{ij}$  and  $D_i$ ) for the ray tracing, and compare the time-delay distance measurements from these different background cosmologies. For the case of HE 0435–1223, where G1 is not strongly lensing the background source but merely perturbs it, the effect on the time delays from G1 is weak. The lack of sensitivity to  $\Omega_m$  and  $w$  suggests that HE 0435–1223 is not sensitive to  $\beta_{12}$  at an interesting level to probe it directly in the same way as a double source plane lens (e.g. Gavazzi et al. 2008; Collett & Auger 2014). In HE 0435–1223, we find instead that the time delays are mostly set by the strong lens, and we can measure the effective  $D_{\Delta t}$ , which is  $D_{\Delta t}^{13}$ , that is independent of assumptions on the background cosmology, as demonstrated in Fig. 7. This robust distance determination then permits us to constrain any reasonable cosmological model via the distance–redshift relation.

## 5 RESULTS

The marginalized posterior  $D_{\Delta t}$  distributions for our lens model are given in Table 2. We report the median and 68 percent quantiles for each of the models described in Section 4.7, as well as a final distribution that combines all of the chains. These distributions are shown in Fig. 8. The blinded distributions, shown on the bottom  $x$ -axis of Fig. 8, were the only values seen until the unblinding. The velocity dispersion and external convergence have



**Figure 8.** PDF of  $D_{\Delta t}$  for the various models, as indicated by the legend. The median and 68 per cent quantile of each distribution is given. The thick black line represents the sum of all the distributions, which accounts for the various systematic uncertainties. The dotted black line is the skewed lognormal distribution (equation 17) fit to the final distribution. The bottom  $x$ -axis shows the blinded result, which is obtained by subtracting the median of the combined PDF from the absolute  $D_{\Delta t}$  values. The top  $x$ -axis shows the true  $D_{\Delta t}$  values. Throughout our blind analysis, the top  $x$ -axis was hidden until our analysis was finalized.

been included in these distributions. Each of the chains representing a different systematics test is given equal weight because the goodness-of-fit is comparable. Our final constraint on the effective time-delay distance in HE 0435–1223 is  $D_{\Delta t} = 2612^{+208}_{-191}$  Mpc. We note that our fiducial model parameters are consistent with an identical model run only using the  $F160W$  band as constraints.

Table 2 also shows the  $\chi^2$  for each model. The  $\chi^2$  values are calculated within the fiducial arcmask and outside of the AGN mask+2 pixel region to ensure a fair comparison among the different models. The  $\chi^2$  is calculated by summing the square of the normalized residual pixels (third row of Fig. 4) within this region.

The number of degrees of freedom,  $N_{\text{dof}}$ , is the number of pixels in this region across all three bands ( $N_d = 9577$ ) minus the number of lens mass/light model parameters minus a  $\Gamma$  term that represents the effective number of source pixels accounting for source regularization (see Suyu et al. 2006).  $\Gamma$  is calculated separately for each of the models’ arcmask and AGN mask regions. The typical  $N_{\text{dof}}$  for our models is  $\sim 8400$ – $8600$ . Most of the residual  $\chi^2$  is associated with a few compact star-forming regions in the host galaxy that cannot be modelled at the resolution of our source pixel grid (Fig. 4). Our tests show that masking out these regions affects the  $D_{\Delta t}$  distribution by less than our systematic uncertainties (see Section 4.2). We

**Table 3.** Cosmological parameter constraints from HE 0435–1223. Reported values are medians, with errors corresponding to the 16th and 84th percentiles. *Planck* priors are the Planck Collaboration XIII (2016) chains from baseline high-L *Planck* power spectra and low-L temperature and LFI polarization (*plikHM\_TT\_lowTEB*).

Model name	Description	Parameter priors	Marginalized cosmological parameters
UH <sub>0</sub>	Flat $\Lambda$ CDM cosmology, fixed $\Omega_\Lambda$ $\Omega_m = 1 - \Omega_\Lambda = 0.32$	$H_0$ uniform in [0,150]	$H_0 = 74.3^{+6.0}_{-5.4}$
U $\Lambda$ CDM	Flat $\Lambda$ CDM cosmology $\Omega_m = 1 - \Omega_\Lambda$	$H_0$ uniform in [0,150] $\Omega_\Lambda$ uniform in [0,1]	$H_0 = 73.1^{+5.7}_{-6.0}$ $\Omega_\Lambda = 0.51^{+0.34}_{-0.34}$
$\text{o}\Lambda$ CDM + <i>Planck</i>	Non-flat $\Lambda$ CDM cosmology $\Omega_k = 1 - \Omega_\Lambda - \Omega_m$	<i>Planck</i> prior for $\{H_0, \Omega_\Lambda, \Omega_m\}$	$H_0 = 63.5^{+3.7}_{-3.7}$ $\Omega_m = 0.35^{+0.04}_{-0.04}$ $\Omega_\Lambda = 0.66^{+0.03}_{-0.03}$ $\Omega_k = -0.011^{+0.010}_{-0.011}$
wCDM + <i>Planck</i>	Flat wCDM cosmology $\Omega_m = 1 - \Omega_\Lambda$	<i>Planck</i> prior for $\{H_0, \Omega_\Lambda, w\}$	$H_0 = 83.7^{+9.2}_{-9.0}$ $\Omega_\Lambda = 0.80^{+0.04}_{-0.05}$ $w = -1.52^{+0.27}_{-0.27}$

note that for a fair comparison, the  $\chi^2$  for models with larger arc-masks are calculated on a source plane pixel scale that gives them approximately the same source resolution as the other models ( $41 \times 41$  pixels for the arcmask+1 pixel models,  $45 \times 45$  pixels for the arcmask+2 pixel model). The typical absolute change in  $\chi^2$  for 1 pixel changes<sup>8</sup> in the source grid resolution is  $\sim 60$ – $70$ . We take this as the uncertainty in  $\chi^2$ , and the  $\chi^2$  values are all very close among models that use the same data set. Therefore, we are justified in weighting each of the models equally.

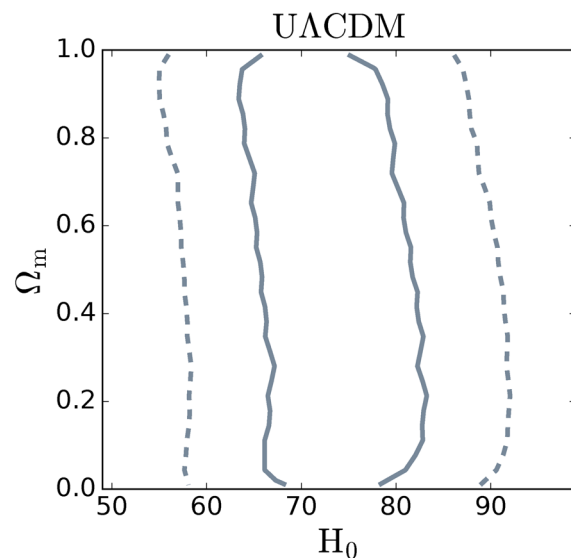
We fit a skewed lognormal function to the  $D_{\Delta t}$  distribution, as this function provides an accurate parametrized representation of our result (Suyu et al. 2010). The distribution has the form

$$P(D_{\Delta t}) = \frac{1}{\sqrt{2\pi}(x - \lambda_D)\sigma_D} \exp\left[-\frac{(\ln(x - \lambda_D) - \mu_D)^2}{2\sigma_D^2}\right], \quad (17)$$

where  $x = D_{\Delta t}/(1 \text{ Mpc})$ ,  $\lambda_D = 653.9$ ,  $\mu_D = 7.5793$  and  $\sigma_D = 0.10312$ . We plot this best-fitting function along with the final  $D_{\Delta t}$  distribution in Fig. 8. The median, 68 per cent, and 95 per cent quantiles of the  $D_{\Delta t}$  distribution and the best-fitting function agree to within  $\sim 0.1$  per cent, indicating that this function is an accurate representation.

Based on our inferred effective time-delay distance, we can calculate cosmological parameters for a variety of cosmological models, which are described in Table 3. For the U $\Lambda$ CDM cosmology, we constrain the Hubble constant to be  $H_0 = 73.1^{+5.7}_{-6.0} \text{ km s}^{-1} \text{ Mpc}^{-1}$ , giving a precision of  $\sim 8$  per cent from just this single lens system. This value is in good agreement with the latest distance ladder results ( $H_0 = 73.24 \pm 1.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ; Riess et al. 2016) and higher than the latest *Planck* measurement for a similar cosmology ( $H_0 = 67.8 \pm 0.9 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ; Planck Collaboration XIII 2016). Fig. 9 shows the posterior distribution of  $H_0$  and  $\Omega_m$  in U $\Lambda$ CDM. Fixing  $\Omega_\Lambda$  in the UH<sub>0</sub> model does not change the inferred  $H_0$  significantly ( $H_0 = 74.3^{+6.0}_{-5.4} \text{ km s}^{-1} \text{ Mpc}^{-1}$ ). Our results for the  $\text{o}\Lambda$ CDM + *Planck* model suggest a Universe consistent with spatial flatness. Interestingly, the wCDM + *Planck* model prefers a dark energy equation of state parameter that is in mild tension with  $w = -1$  at the  $\sim 2\sigma$  level. The results for each of our models are summarized in Table 3.

<sup>8</sup> A 1 pixel change in source grid resolution roughly corresponds to the changes in source pixel size across our different models.



**Figure 9.** Posterior distribution of  $H_0$  and  $\Omega_m$  for the U $\Lambda$ CDM cosmology determined from the time-delay distance inference of HE 0435–1223. The contours represent the 68 per cent and 95 per cent quantiles of the distribution.  $\Omega_m$  has a weak influence on  $D_{\Delta t}$ , so it is not well constrained. The marginalized value of  $H_0$  for this cosmology is  $73.1^{+5.7}_{-6.0} \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

The results for HE 0435–1223 presented here can be combined with previous analyses of B1608+656 (Suyu et al. 2010) and RXJ1131–1231 (Suyu et al. 2013, 2014) to produce stronger constraints on cosmology. A full analysis of the implications of our  $D_{\Delta t}$  inference for a variety of cosmologies using constraints from all three H0LiCOW lenses analysed to date is presented in H0LiCOW Paper V.

## 6 CONCLUSIONS

We have performed a blind analysis of the gravitational lens HE 0435–1223 using new deep *HST* imaging, high-precision time delays from COSMOGRAIL, a measurement of the lens galaxy velocity dispersion, and spectroscopic and photometric data to constrain the mass distribution along the LOS. Our model is able to reproduce the surface brightness structure of the lensed AGN and

host galaxy in all three *HST* bands, as well as the measured time delays. Combining these data sets and accounting for various sources of systematic uncertainty in the lens modelling, we constrain the effective time-delay distance to be  $D_{\Delta t} = 2612_{-191}^{+208}$  Mpc, giving a precision of 7.6 per cent. For a flat  $\Lambda$ CDM cosmology with uniform priors on  $H_0$  and  $\Omega_{\Lambda}$ , we constrain the Hubble constant to be  $H_0 = 73.1_{-6.0}^{+5.7}$  km s $^{-1}$  Mpc $^{-1}$  (a precision of  $\sim 8$  per cent), in good agreement with the latest distance ladder results. A detailed analysis of the implications of our  $D_{\Delta t}$  constraint on a variety of cosmologies is presented in [HOLiCOW Paper V](#).

Upcoming analyses of the remaining two HOLiCOW systems will complete the sample of five time-delay lenses and constrain  $H_0$  to  $<3.5$  per cent precision. Our extensive blind analysis of HE 0435–1223 demonstrates the utility of gravitational lens time delays as a precise and independent cosmological probe. With hundreds of new lensed AGN expected to be discovered in current and future wide-field imaging surveys (Oguri & Marshall 2010), we expect time-delay cosmography to provide competitive cosmological constraints throughout the next decade.

## ACKNOWLEDGEMENTS

We thank the referee for reviewing this paper and providing helpful commentary. We thank Adriano Agnello, Roger Blandford, Geoff Chih-Fan Chen, Xuheng Ding, Yashar Hezaveh, Kai Liao, John McKean, Georges Meylan, Danka Paraficz, Chiara Spinello, Malte Tewes, Olga Tihonova and Simona Vegetti for their contributions to the HOLiCOW project. We thank Simon Birrer for useful discussions and feedback. We thank Bau-Ching Hsieh for computing support on the SuMIRE computing cluster. We thank Michelle Wilson for sharing the lens galaxy group properties from an ongoing independent study. HOLiCOW and COSMOGRAIL are made possible thanks to the continuous work of all observers and technical staff obtaining the monitoring observations, in particular at the Swiss Euler telescope at La Silla Observatory. Euler is supported by the Swiss National Science Foundation. KCW is supported 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 is supported by the Max Planck Society through the Max Planck Research Group. This work is supported in part by the Ministry of Science and Technology in Taiwan via grant MOST-103-2112-M-001-003-MY3. VB and FC are supported by the Swiss National Science Foundation (SNSF). CDF and CER are funded through the NSF grant AST-1312329, ‘Collaborative Research: Accurate cosmology with strong gravitational lens time delays’ and the *HST* grant GO-12889. DS acknowledges funding support from a *Back to Belgium* grant from the Belgian Federal Science Policy (BELSPO). TT thanks the Packard Foundation for generous support through a Packard Research Fellowship, the NSF for funding through NSF grant AST-1450141, ‘Collaborative Research: Accurate cosmology with strong gravitational lens time delays’. SH acknowledges support by the DFG cluster of excellence ‘Origin and Structure of the Universe’ ([www.universe-cluster.de](http://www.universe-cluster.de)). LVEK is supported in part through an NWO-VICI career grant (project number 639.043.308). PJM acknowledges support from the U.S. Department of Energy under contract number DE-AC02-76SF00515. This paper is based on observations made with the NASA/ESA *HST*, obtained at the Space Telescope Science Institute, which is operated by the Association of

Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with Programs #12889 and #9744. Support for program #12889 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. Some of 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 Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

## REFERENCES

- Anderson J., Bedin L. R., 2010, *PASP*, 122, 1035  
 Appenzeller I. et al., 1998, *The Messenger*, 94, 1  
 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  
 Barkana R., 1998, *ApJ*, 502, 531  
 Barnabè M., Czoske O., Koopmans L. V. E., Treu T., Bolton A. S., 2011, *MNRAS*, 415, 2215  
 Bernardi M., Roche N., Shankar F., Sheth R. K., 2011, *MNRAS*, 412, 684  
 Birrer S., Amara A., Refregier A., 2016, *J. Cosmology Astropart. Phys.*, 8, 020  
 Blandford R., Narayan R., 1986, *ApJ*, 310, 568  
 Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Moustakas L. A., 2006, *ApJ*, 638, 703  
 Bonvin V. et al., 2016a, *MNRAS*, preprint ([arXiv:1607.01790](https://arxiv.org/abs/1607.01790)) (HOLiCOW Paper V)  
 Bonvin V., Tewes M., Courbin F., Kuntzer T., Sluse D., Meylan G., 2016b, *A&A*, 585, A88  
 Cantale N., Courbin F., Tewes M., Jablonka P., Meylan G., 2016a, *A&A*, 589, A81  
 Cantale N. et al., 2016b, *A&A*, 589, A82  
 Chabrier G., 2003, *PASP*, 115, 763  
 Chen G. C.-F. et al., 2016, *MNRAS*, 462, 3457  
 Coe D., Moustakas L. A., 2009, *ApJ*, 706, 45  
 Collett T. E., Auger M. W., 2014, *MNRAS*, 443, 969  
 Collett T. E., Cunnington S. D., 2016, *MNRAS*, 462, 3255  
 Collett T. E. et al., 2013, *MNRAS*, 432, 679  
 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  
 Dutton A. A., Treu T., 2014, *MNRAS*, 438, 3594  
 Dutton A. A. et al., 2011, *MNRAS*, 417, 1621  
 Eigenbrod A., Courbin F., Vuissoz C., Meylan G., Saha P., Dye S., 2005, *A&A*, 436, 25  
 Fadely R., Keeton C. R., Nakajima R., Bernstein G. M., 2010, *ApJ*, 711, 246  
 Falco E. E., Gorenstein M. V., Shapiro I. I., 1985, *ApJ*, 289, L1  
 Fassnacht C. D., Pearson T. J., Readhead A. C. S., Browne I. W. A., Koopmans L. V. E., Myers S. T., Wilkinson P. N., 1999, *ApJ*, 527, 498  
 Fassnacht C. D., Xanthopoulos E., Koopmans L. V. E., Rusin D., 2002, *ApJ*, 581, 823  
 Fassnacht C. D., Gal R. R., Lubin L. M., McKean J. P., Squires G. K., Readhead A. C. S., 2006, *ApJ*, 642, 30  
 Fassnacht C. D., Koopmans L. V. E., Wong K. C., 2011, *MNRAS*, 410, 2167  
 Fazio G. G. et al., 2004, *ApJS*, 154, 10  
 Freedman W. L., Madore B. F., Scowcroft V., Burns C., Monson A., Persson S. E., Seibert M., Rigby J., 2012, *ApJ*, 758, 24



- Gavazzi R., Treu T., Rhodes J. D., Koopmans L. V. E., Bolton A. S., Burles S., Massey R. J., Moustakas L. A., 2007, *ApJ*, 667, 176
- Gavazzi R., Treu T., Koopmans L. V. E., Bolton A. S., Moustakas L. A., Burles S., Marshall P. J., 2008, *ApJ*, 677, 1046
- Golse G., Kneib J.-P., 2002, *A&A*, 390, 821
- Gorenstein M. V., Shapiro I. I., Falco E. E., 1988, *ApJ*, 327, 693
- Greene Z. S. et al., 2013, *ApJ*, 768, 39
- Hernquist L., 1990, *ApJ*, 356, 359
- Hilbert S., Hartlap J., White S. D. M., Schneider P., 2009, *A&A*, 499, 31
- Hook I. M., Jørgensen I., Allington-Smith J. R., Davies R. L., Metcalfe N., Murowinski R. G., Crampton D., 2004, *PASP*, 116, 425
- Hu W., 2005, in Wolff S. C., Lauer T. R., eds, *ASP Conf. Ser. Vol. 339, Observing Dark Energy*. Astron. Soc. Pac., San Francisco, p. 215
- Ilbert O. et al., 2006, *A&A*, 457, 841
- Jaffe W., 1983, *MNRAS*, 202, 995
- Kassiola A., Kovner I., 1993, *ApJ*, 417, 450
- Keeton C. R., 2003, *ApJ*, 584, 664
- Keeton C. R., Kochanek C. S., 1997, *ApJ*, 487, 42
- Kochanek C. S., 2002, *ApJ*, 578, 25
- Kochanek C. S., 2003, *ApJ*, 583, 49
- Kochanek C. S., Apostolakis J., 1988, *MNRAS*, 235, 1073
- Kochanek C. S., Morgan N. D., Falco E. E., McLeod B. A., Winn J. N., Dembicky J., Ketzbeck B., 2006, *ApJ*, 640, 47
- Koopmans L. V. E., 2004, preprint ([astro-ph/0412596](https://arxiv.org/abs/astro-ph/0412596))
- Koopmans L. V. E., Treu T., Fassnacht C. D., Blandford R. D., Surpi G., 2003, *ApJ*, 599, 70
- 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
- Kovner I., 1987, *ApJ*, 316, 52
- Lewis A., Bridle S., 2002, *Phys. Rev. D*, 66, 103511
- 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., 2016, *ApJ*, preprint ([arXiv:1601.05417](https://arxiv.org/abs/1601.05417))
- Magain P., Courbin F., Sohy S., 1998, *ApJ*, 494, 472
- Mamon G. A., Łokas E. L., 2005, *MNRAS*, 363, 705
- Massey R., Stoughton C., Leauthaud A., Rhodes J., Koekemoer A., Ellis R., Shaghoulouian E., 2010, *MNRAS*, 401, 371
- 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
- Morgan N. D., Kochanek C. S., Pevunova O., Schechter P. L., 2005, *AJ*, 129, 2531
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Oguri M., 2007, *ApJ*, 660, 1
- Oguri M., Marshall P. J., 2010, *MNRAS*, 405, 2579
- Oke J. B. et al., 1995, *PASP*, 107, 375
- Osipkov L. P., 1979, *Pis'ma Astron. Zh.*, 5, 77
- Petters A. O., Levine H., Wambsganss J., 2001, *Singularity Theory and Gravitational Lensing*. Birkhäuser, Boston
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- Rathna Kumar S., Stalin C. S., Prabhu T. P., 2015, *A&A*, 580, A38
- Refsdal S., 1964, *MNRAS*, 128, 307
- Riess A. G. et al., 2011, *ApJ*, 730, 119
- Riess A. G. et al., 2016, *ApJ*, 826, 56
- Rusu C. E. et al., 2016, *MNRAS*, preprint ([arXiv:1607.01047](https://arxiv.org/abs/1607.01047)) (H0LiCOW Paper III)
- Saha P., 2000, *AJ*, 120, 1654
- Saha P., Coles J., Macciò A. V., Williams L. L. R., 2006, *ApJ*, 650, L17
- Schechter P. L. et al., 1997, *ApJ*, 475, L85
- Schneider P., 2014, *A&A*, 568, L2
- 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*. Springer-Verlag, New York
- Seljak U., 1994, *ApJ*, 436, 509
- Sereno M., Paraficz D., 2014, *MNRAS*, 437, 600
- Sluse D., Hutsemékers D., Courbin F., Meylan G., Wambsganss J., 2012, *A&A*, 544, A62
- Sluse D. et al., 2016, *MNRAS*, preprint ([arXiv:1607.00382](https://arxiv.org/abs/1607.00382)) (H0LiCOW Paper II)
- 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
- Springel V. et al., 2005, *Nature*, 435, 629
- Suyu S. H., 2012, *MNRAS*, 426, 868
- 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., 2012a, preprint ([arXiv:1202.4459](https://arxiv.org/abs/1202.4459))
- Suyu S. H. et al., 2012b, *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., 2016, *MNRAS*, preprint ([arXiv:1607.00017](https://arxiv.org/abs/1607.00017)) (H0LiCOW Paper I)
- Tewes M., Courbin F., Meylan G., 2013a, *A&A*, 553, A120
- Tewes M. et al., 2013b, *A&A*, 556, A22
- Treu T., Koopmans L. V. E., 2002, *ApJ*, 575, 87
- Treu T., Marshall P. J., 2016, *A&AR*, preprint ([arXiv:1605.05333](https://arxiv.org/abs/1605.05333))
- Unruh S., Schneider P., Sluse D., 2016, *A&A*, preprint ([arXiv:1606.04321](https://arxiv.org/abs/1606.04321))
- Vanderricht C., Schneider J., Herpe G., Chevreton M., Moles M., Wlerick G., 1989, *A&A*, 215, 1
- Velander M. et al., 2014, *MNRAS*, 437, 2111
- Warren S. J., Dye S., 2003, *ApJ*, 590, 673
- Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, *Phys. Rep.*, 530, 87
- Williams K. A., Momcheva I., Keeton C. R., Zabludoff A. I., Lehár J., 2006, *ApJ*, 646, 85
- Wisotzki L., Christlieb N., Bade N., Beckmann V., Köhler T., Vanelle C., Reimers D., 2000, *A&A*, 358, 77
- 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
- Wucknitz O., 2002, *MNRAS*, 332, 951
- Xu D., Sluse D., Schneider P., Springel V., Vogelsberger M., Nelson D., Hernquist L., 2016, *MNRAS*, 456, 739

## APPENDIX A: MODEL PARAMETERS

We show the marginalized parameter constraints for each of the SPEND models in Table A1 and for each of the composite models in Table A2.

**Table A1.** SPEMD model parameters. Reported values are medians, with errors corresponding to the 16th and 84th percentiles. Angles are measured east of north.

Parameter	Marginalized constraints								
	Fiducial	Img+10	Arc+1,50src	Arc+2,Im+10,50src	AGNwht=0	AGNmask+1	AGNmask+2	5 pert.	Sérsic
$\theta_E$ (arcsec) <sup>a</sup>	1.182 <sup>+0.002</sup> <sub>-0.002</sub>	1.182 <sup>+0.002</sup> <sub>-0.002</sub>	1.169 <sup>+0.002</sup> <sub>-0.001</sub>	1.176 <sup>+0.002</sup> <sub>-0.002</sub>	1.180 <sup>+0.002</sup> <sub>-0.002</sub>	1.182 <sup>+0.002</sup> <sub>-0.002</sub>	1.181 <sup>+0.002</sup> <sub>-0.002</sub>	1.149 <sup>+0.003</sup> <sub>-0.003</sub>	1.181 <sup>+0.002</sup> <sub>-0.002</sub>
$q$	0.80 <sup>+0.01</sup> <sub>-0.01</sub>	0.81 <sup>+0.01</sup> <sub>-0.01</sub>	0.81 <sup>+0.01</sup> <sub>-0.01</sub>	0.81 <sup>+0.01</sup> <sub>-0.01</sub>	0.81 <sup>+0.01</sup> <sub>-0.01</sub>	0.80 <sup>+0.01</sup> <sub>-0.01</sub>	0.80 <sup>+0.01</sup> <sub>-0.01</sub>	0.81 <sup>+0.01</sup> <sub>-0.01</sub>	0.81 <sup>+0.01</sup> <sub>-0.01</sub>
$\theta_q$ (°)	-16.8 <sup>+0.5</sup> <sub>-0.6</sub>	-17.0 <sup>+0.5</sup> <sub>-0.5</sub>	-17.1 <sup>+0.5</sup> <sub>-0.4</sub>	-16.6 <sup>+0.5</sup> <sub>-0.6</sub>	-16.8 <sup>+0.5</sup> <sub>-0.5</sub>	-17.1 <sup>+0.6</sup> <sub>-0.6</sub>	-17.3 <sup>+0.5</sup> <sub>-0.6</sub>	-17.8 <sup>+0.5</sup> <sub>-0.5</sub>	-17.0 <sup>+0.6</sup> <sub>-0.6</sub>
$\gamma'$	1.93 <sup>+0.02</sup> <sub>-0.02</sub>	1.95 <sup>+0.02</sup> <sub>-0.03</sub>	1.89 <sup>+0.02</sup> <sub>-0.02</sub>	1.91 <sup>+0.02</sup> <sub>-0.01</sub>	1.94 <sup>+0.02</sup> <sub>-0.02</sub>	1.94 <sup>+0.02</sup> <sub>-0.02</sub>	1.94 <sup>+0.02</sup> <sub>-0.02</sub>	1.93 <sup>+0.02</sup> <sub>-0.01</sub>	1.87 <sup>+0.03</sup> <sub>-0.02</sub>
$\gamma_{\text{ext}}$	0.030 <sup>+0.003</sup> <sub>-0.003</sub>	0.033 <sup>+0.003</sup> <sub>-0.003</sub>	0.032 <sup>+0.002</sup> <sub>-0.002</sub>	0.030 <sup>+0.003</sup> <sub>-0.004</sub>	0.033 <sup>+0.003</sup> <sub>-0.003</sub>	0.032 <sup>+0.003</sup> <sub>-0.003</sub>	0.031 <sup>+0.003</sup> <sub>-0.003</sub>	0.025 <sup>+0.002</sup> <sub>-0.002</sub>	0.026 <sup>+0.003</sup> <sub>-0.003</sub>
$\theta_\gamma$ (°)	63.7 <sup>+2.4</sup> <sub>-2.2</sub>	65.0 <sup>+1.9</sup> <sub>-1.8</sub>	57.7 <sup>+1.2</sup> <sub>-1.6</sub>	60.6 <sup>+2.1</sup> <sub>-1.7</sub>	63.6 <sup>+1.9</sup> <sub>-1.9</sub>	65.3 <sup>+1.9</sup> <sub>-2.0</sub>	65.4 <sup>+2.0</sup> <sub>-2.0</sub>	-88.5 <sup>+1.4</sup> <sub>-1.3</sub>	63.1 <sup>+2.7</sup> <sub>-2.7</sub>
G1 $\theta_E$ (arcsec)	0.37 <sup>+0.03</sup> <sub>-0.03</sub>	0.38 <sup>+0.02</sup> <sub>-0.03</sub>	0.48 <sup>+0.02</sup> <sub>-0.02</sub>	0.40 <sup>+0.02</sup> <sub>-0.02</sub>	0.39 <sup>+0.02</sup> <sub>-0.03</sub>	0.37 <sup>+0.03</sup> <sub>-0.03</sub>	0.37 <sup>+0.03</sup> <sub>-0.03</sub>	0.26 <sup>+0.01</sup> <sub>-0.01</sub>	0.35 <sup>+0.03</sup> <sub>-0.03</sub>

Note. <sup>a</sup>Spherical-equivalent *Einstein* radius.

**Table A2.** Composite model parameters. Reported values are medians, with errors corresponding to the 16th and 84th percentiles. Angles are measured east of north.

Parameter	Marginalized constraints		
	Composite	Composite, AGNmask+1	Composite, Arcmask+1,50src
Stellar $M/L$ ( $M_\odot/L_\odot$ ) <sup>a</sup>	2.5 <sup>+0.1</sup> <sub>-0.1</sub>	2.6 <sup>+0.2</sup> <sub>-0.2</sub>	2.3 <sup>+0.1</sup> <sub>-0.1</sub>
NFW $\kappa_{0,h}$	0.41 <sup>+0.03</sup> <sub>-0.03</sub>	0.36 <sup>+0.03</sup> <sub>-0.03</sub>	0.39 <sup>+0.01</sup> <sub>-0.01</sub>
NFW $r_s$ (arcsec)	8.43 <sup>+0.58</sup> <sub>-1.94</sub>	9.43 <sup>+0.69</sup> <sub>-0.94</sub>	8.96 <sup>+0.28</sup> <sub>-0.26</sub>
NFW $q$	0.82 <sup>+0.01</sup> <sub>-0.02</sub>	0.81 <sup>+0.02</sup> <sub>-0.02</sub>	0.83 <sup>+0.01</sup> <sub>-0.01</sub>
NFW $\theta_q$ (°)	-18.4 <sup>+0.7</sup> <sub>-0.7</sub>	-18.6 <sup>+0.7</sup> <sub>-0.7</sub>	-19.7 <sup>+0.6</sup> <sub>-0.6</sub>
$\gamma_{\text{ext}}$	0.004 <sup>+0.003</sup> <sub>-0.002</sub>	0.003 <sup>+0.003</sup> <sub>-0.002</sub>	0.006 <sup>+0.002</sup> <sub>-0.002</sub>
$\theta_\gamma$ (°)	34.4 <sup>+22.9</sup> <sub>-32.5</sub>	44.6 <sup>+26.8</sup> <sub>-36.2</sub>	28.3 <sup>+6.0</sup> <sub>-7.6</sub>
G1 $\theta_E$ (arcsec)	0.33 <sup>+0.03</sup> <sub>-0.03</sub>	0.32 <sup>+0.03</sup> <sub>-0.03</sub>	0.42 <sup>+0.03</sup> <sub>-0.02</sub>

Note. <sup>a</sup> $M/L$  within  $\theta_E$  for rest-frame  $V$  band. The given uncertainties are only statistical and do not include systematic effects. The stellar mass is calculated assuming  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.3$  and  $\Omega_\Lambda = 0.7$ , but changes in the cosmology affect the  $M/L$  by a negligible amount.

## APPENDIX B: INVERSE MAGNIFICATION TENSORS

The components of the inverse magnification tensor are

$$\mathcal{A}_{ij}(\boldsymbol{\theta}) = \frac{\partial \beta_i}{\partial \theta_j}, \quad (\text{B1})$$

where  $i = 1, 2$ ,  $j = 1, 2$ ,  $\boldsymbol{\beta} = (\beta_1, \beta_2)$  is the source plane coordinates, and  $\boldsymbol{\theta} = (\theta_1, \theta_2)$  is the coordinates of the image plane [which is also the first lens plane,  $\boldsymbol{\theta} = (\theta_{11}, \theta_{12})$ ].

The general multiplane lens equation is

$$\boldsymbol{\theta}_j = \boldsymbol{\theta}_1 - \sum_{i=1}^{j-1} \beta_{ij} \boldsymbol{\alpha}_i(\boldsymbol{\theta}_i), \quad (\text{B2})$$

where  $\beta_{ij}$  is given by equation (12) [note the difference between  $\beta_{ij}$  with two indices and the source coordinates  $\beta_i$  with one index]. This is the general form of equations (13) and (14). For  $N$  lens planes, the source coordinates are  $\boldsymbol{\beta} = \boldsymbol{\theta}_{N+1}$ . For the case of two lens planes, as we have in our model,  $\boldsymbol{\beta} = \boldsymbol{\theta}_3$ . We present the inverse magnification tensors at the positions of the lensed quasar images in Table B1. While the inverse magnification tensor is symmetric for single-plane lensing, this is not true for multiplane lensing.

**Table B1.** Inverse magnification tensor.

Model	A	B	C	D
Fiducial	$\begin{bmatrix} 0.732^{+0.018}_{-0.016} \\ -0.302^{+0.007}_{-0.007} \\ 0.328^{+0.008}_{-0.008} \end{bmatrix}$	$\begin{bmatrix} -0.105^{+0.004}_{-0.004} \\ 0.281^{+0.008}_{-0.007} \\ -0.105^{+0.005}_{-0.005} \end{bmatrix}$	$\begin{bmatrix} 0.832^{+0.020}_{-0.017} \\ -0.024^{+0.003}_{-0.003} \\ 0.846^{+0.015}_{-0.022} \end{bmatrix}$	$\begin{bmatrix} -0.220^{+0.008}_{-0.007} \\ 0.249^{+0.008}_{-0.008} \\ -0.220^{+0.009}_{-0.006} \end{bmatrix}$
img+10pix	$\begin{bmatrix} 0.743^{+0.014}_{-0.020} \\ -0.308^{+0.009}_{-0.007} \\ 0.333^{+0.010}_{-0.010} \end{bmatrix}$	$\begin{bmatrix} -0.105^{+0.005}_{-0.005} \\ 0.286^{+0.007}_{-0.009} \\ 0.333^{+0.010}_{-0.010} \end{bmatrix}$	$\begin{bmatrix} 0.846^{+0.015}_{-0.022} \\ -0.024^{+0.003}_{-0.003} \\ 0.776^{+0.014}_{-0.014} \end{bmatrix}$	$\begin{bmatrix} 0.216^{+0.007}_{-0.007} \\ 0.254^{+0.007}_{-0.009} \\ 0.870^{+0.021}_{-0.027} \end{bmatrix}$
Arcmask+1pix,50x50src	$\begin{bmatrix} 0.692^{+0.013}_{-0.013} \\ -0.292^{+0.005}_{-0.006} \\ 0.313^{+0.006}_{-0.006} \end{bmatrix}$	$\begin{bmatrix} -0.099^{+0.003}_{-0.003} \\ 0.271^{+0.007}_{-0.006} \\ 0.809^{+0.018}_{-0.016} \end{bmatrix}$	$\begin{bmatrix} -0.029^{+0.002}_{-0.002} \\ 0.178^{+0.004}_{-0.004} \\ 0.805^{+0.014}_{-0.013} \end{bmatrix}$	$\begin{bmatrix} -0.207^{+0.005}_{-0.006} \\ 0.230^{+0.006}_{-0.005} \\ -0.211^{+0.004}_{-0.005} \end{bmatrix}$
Arcmask+2pix,img+10pix,50x50src	$\begin{bmatrix} 0.711^{+0.011}_{-0.011} \\ -0.296^{+0.005}_{-0.005} \\ 0.319^{+0.005}_{-0.005} \end{bmatrix}$	$\begin{bmatrix} -0.102^{+0.003}_{-0.003} \\ 0.274^{+0.005}_{-0.005} \\ 0.830^{+0.019}_{-0.014} \end{bmatrix}$	$\begin{bmatrix} 0.805^{+0.014}_{-0.013} \\ 0.180^{+0.003}_{-0.002} \\ 0.837^{+0.019}_{-0.018} \end{bmatrix}$	$\begin{bmatrix} 0.201^{+0.006}_{-0.005} \\ 0.241^{+0.006}_{-0.005} \\ -0.219^{+0.007}_{-0.008} \end{bmatrix}$
AGN mask weight=0	$\begin{bmatrix} 0.736^{+0.017}_{-0.016} \\ -0.307^{+0.007}_{-0.007} \\ 0.332^{+0.008}_{-0.008} \end{bmatrix}$	$\begin{bmatrix} -0.103^{+0.004}_{-0.004} \\ 0.286^{+0.008}_{-0.008} \\ 0.332^{+0.008}_{-0.008} \end{bmatrix}$	$\begin{bmatrix} -0.007^{+0.002}_{-0.002} \\ 0.185^{+0.005}_{-0.005} \\ -0.023^{+0.003}_{-0.003} \end{bmatrix}$	$\begin{bmatrix} 0.211^{+0.007}_{-0.007} \\ 0.863^{+0.024}_{-0.023} \\ 0.250^{+0.008}_{-0.007} \end{bmatrix}$
AGN mask+1pix	$\begin{bmatrix} 0.737^{+0.016}_{-0.015} \\ -0.306^{+0.007}_{-0.007} \\ 0.334^{+0.007}_{-0.007} \end{bmatrix}$	$\begin{bmatrix} -0.106^{+0.004}_{-0.004} \\ 0.286^{+0.008}_{-0.008} \\ 0.334^{+0.007}_{-0.007} \end{bmatrix}$	$\begin{bmatrix} 0.840^{+0.017}_{-0.017} \\ 0.187^{+0.004}_{-0.004} \\ -0.023^{+0.003}_{-0.003} \end{bmatrix}$	$\begin{bmatrix} 0.216^{+0.006}_{-0.006} \\ 0.863^{+0.023}_{-0.022} \\ -0.223^{+0.007}_{-0.007} \end{bmatrix}$
AGN mask+2pix	$\begin{bmatrix} 0.734^{+0.016}_{-0.015} \\ -0.306^{+0.006}_{-0.007} \\ 0.334^{+0.007}_{-0.007} \end{bmatrix}$	$\begin{bmatrix} -0.105^{+0.004}_{-0.004} \\ 0.285^{+0.007}_{-0.008} \\ -0.090^{+0.003}_{-0.003} \end{bmatrix}$	$\begin{bmatrix} 0.837^{+0.018}_{-0.016} \\ -0.023^{+0.003}_{-0.003} \\ 0.802^{+0.014}_{-0.012} \end{bmatrix}$	$\begin{bmatrix} 0.216^{+0.006}_{-0.006} \\ 0.858^{+0.023}_{-0.020} \\ -0.216^{+0.005}_{-0.006} \end{bmatrix}$
5 perturbors	$\begin{bmatrix} 0.707^{+0.012}_{-0.011} \\ -0.301^{+0.005}_{-0.005} \\ 0.317^{+0.006}_{-0.005} \end{bmatrix}$	$\begin{bmatrix} -0.090^{+0.003}_{-0.003} \\ 0.277^{+0.006}_{-0.005} \\ 0.826^{+0.015}_{-0.012} \end{bmatrix}$	$\begin{bmatrix} -0.019^{+0.002}_{-0.002} \\ 0.181^{+0.004}_{-0.004} \\ 0.784^{+0.022}_{-0.018} \end{bmatrix}$	$\begin{bmatrix} 0.202^{+0.005}_{-0.005} \\ 0.237^{+0.006}_{-0.005} \\ -0.205^{+0.007}_{-0.008} \end{bmatrix}$
Sérsic profiles	$\begin{bmatrix} 0.688^{+0.019}_{-0.015} \\ -0.285^{+0.007}_{-0.008} \\ 0.309^{+0.009}_{-0.008} \end{bmatrix}$	$\begin{bmatrix} -0.098^{+0.004}_{-0.004} \\ 0.262^{+0.008}_{-0.007} \\ 0.801^{+0.025}_{-0.019} \end{bmatrix}$	$\begin{bmatrix} -0.008^{+0.002}_{-0.002} \\ 0.173^{+0.005}_{-0.004} \\ -0.022^{+0.002}_{-0.002} \end{bmatrix}$	$\begin{bmatrix} 0.198^{+0.008}_{-0.007} \\ 0.790^{+0.027}_{-0.021} \\ 0.231^{+0.008}_{-0.007} \end{bmatrix}$
Composite	$\begin{bmatrix} 0.737^{+0.034}_{-0.026} \\ -0.295^{+0.010}_{-0.011} \\ 0.315^{+0.012}_{-0.010} \end{bmatrix}$	$\begin{bmatrix} -0.104^{+0.006}_{-0.007} \\ 0.276^{+0.011}_{-0.010} \\ 0.837^{+0.028}_{-0.025} \end{bmatrix}$	$\begin{bmatrix} 0.833^{+0.028}_{-0.027} \\ -0.016^{+0.007}_{-0.015} \\ -0.029^{+0.006}_{-0.016} \end{bmatrix}$	$\begin{bmatrix} 0.212^{+0.008}_{-0.008} \\ 0.834^{+0.029}_{-0.026} \\ 0.244^{+0.010}_{-0.009} \end{bmatrix}$
Composite,AGN mask+1pix	$\begin{bmatrix} 0.738^{+0.025}_{-0.026} \\ -0.304^{+0.012}_{-0.011} \\ 0.329^{+0.011}_{-0.012} \end{bmatrix}$	$\begin{bmatrix} -0.103^{+0.006}_{-0.006} \\ 0.282^{+0.011}_{-0.013} \\ 0.855^{+0.028}_{-0.030} \end{bmatrix}$	$\begin{bmatrix} 0.843^{+0.027}_{-0.028} \\ -0.022^{+0.003}_{-0.003} \\ 0.795^{+0.013}_{-0.014} \end{bmatrix}$	$\begin{bmatrix} 0.221^{+0.009}_{-0.008} \\ 0.855^{+0.030}_{-0.032} \\ -0.208^{+0.005}_{-0.005} \end{bmatrix}$
Composite,Arcmask+1pix,50x50src	$\begin{bmatrix} 0.705^{+0.012}_{-0.013} \\ -0.298^{+0.006}_{-0.006} \\ 0.319^{+0.006}_{-0.006} \end{bmatrix}$	$\begin{bmatrix} -0.097^{+0.003}_{-0.003} \\ 0.275^{+0.006}_{-0.006} \\ 0.814^{+0.013}_{-0.014} \end{bmatrix}$	$\begin{bmatrix} -0.008^{+0.001}_{-0.001} \\ 0.178^{+0.003}_{-0.004} \\ -0.025^{+0.002}_{-0.002} \end{bmatrix}$	$\begin{bmatrix} 0.196^{+0.004}_{-0.005} \\ 0.811^{+0.014}_{-0.015} \\ 0.236^{+0.005}_{-0.006} \end{bmatrix}$

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