

# TDCOSMO III: Dark matter substructure meets dark energy - the effects of (sub)halos on strong-lensing measurements of $H_0$

D. Gilman<sup>1</sup>\*, S. Birrer<sup>2</sup>, and T. Treu<sup>1</sup>

<sup>1</sup> Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA

<sup>2</sup> Kavli Institute for Particle Astrophysics and Cosmology and Department of Physics, Stanford University, Stanford, CA 94305, USA

July 2nd, 2020

## ABSTRACT

Time delay cosmography uses the arrival time delays between images in strong gravitational lenses to measure cosmological parameters, in particular the Hubble constant  $H_0$ . The lens models used in time delay cosmography omit dark matter subhalos and line-of-sight halos because their effects are assumed to be negligible. We explicitly quantify this assumption by analyzing realistic mock lens systems that include full populations of dark matter subhalos and line-of-sight halos, applying the same modeling assumptions used in the literature to infer  $H_0$ . We base the mock lenses on six quadruply-imaged quasars that have delivered measurements of the Hubble constant, and quantify the additional uncertainties and/or bias on a lens-by-lens basis. We show that omitting dark substructure does not bias inferences of  $H_0$ . However, perturbations from substructure contribute an additional source of random uncertainty in the inferred value of  $H_0$  that scales as the square root of the lensing volume divided by the longest time delay. This additional source of uncertainty, for which we provide a fitting function, ranges from 0.6 – 2.4%. It may need to be incorporated in the error budget as the precision of cosmographic inferences from single lenses improves, and sets a precision limit on inferences from single lenses.

**Key words.** gravitational lensing: strong - cosmology: dark matter - cosmology: cosmological parameters - methods: statistical

## 1. Introduction

Time delay cosmography with multiply-imaged quasars delivers percent-level measurements of the Hubble constant  $H_0$  (Chen et al. 2019; Birrer et al. 2019; Wong et al. 2020; Shajib et al. 2019), and substructure lensing studies use strong gravitational lenses to measure the halo mass function below  $10^9 M_\odot$ , thereby constraining the nature of dark matter (Dalal & Kochanek 2002; Nierenberg et al. 2014; Hezaveh et al. 2016b; Birrer et al. 2017; Vegetti et al. 2018; Hsueh et al. 2019; Gilman et al. 2020a,b). To date, these two lines of research are pursued independently. Dark substructure is omitted from the lens models used in time delay cosmography, while substructure lensing analyses do not explicitly marginalize over the uncertainties of cosmological parameters. In this paper, we move to bridge the gap between these fields by quantifying how dark substructure affects the precision and accuracy of measurements of  $H_0$  obtained from strong lenses.

The impetus to carry out this analysis now stems from a recent challenge to the standard  $\Lambda$ CDM cosmology. Strong-lensing measurements of  $H_0$  in the local Universe presented by the HOLICOW collaboration (Suyu et al. 2017), and from Type Ia supernova, carried out by the SH0ES collaboration (Riess et al. 2019), differ from the value of  $H_0$  derived from the Cosmic Microwave Background (Planck Collaboration et al. 2018). The tension between the CMB data and six strongly-lensed quasars alone is  $3.1\sigma$ , and exceeds  $5\sigma$  when lensing is combined with supernova data (Wong et al. 2020; Riess 2019). A resolution to this tension may require new physics. Before reaching this con-

clusion it is necessary to quantify potential sources of bias and measurement uncertainties relevant for time delay cosmography.

Oguri (2007) pointed out that dark matter subhalos surrounding the main deflector in galaxy-scale strong lenses alter the arrival times between lensed images. The arrival times provide a direct, one-step method to measure  $H_0$ , so it is conceivable that perturbations from substructure to these data could bias, or contribute to the error budget, of cosmographic analyses. Keeton & Moustakas (2009) (see also Congdon et al. 2010; Cyr-Racine et al. 2016) carried out the first comprehensive analysis of time delay perturbations from substructure, and showed that populations of dark subhalos impart arrival time perturbations on the scale of a few hours. This level of perturbation lies safely below current measurement precision, which measures image arrival times down to a few-days precision (e.g. Bonvin et al. 2018a; Courbin et al. 2018; Millon et al. 2020). On these grounds, substructure is assumed to be negligible and it is omitted from the strong lens models used for cosmography.

However, as the precision of time delay cosmography improves it is timely to revisit this assumption and quantify the potential impact of substructure both in terms of accuracy and precision. First, Keeton & Moustakas (2009) did not account for field halos along the line of sight to strong lenses, which can significantly outnumber subhalos. Because the number of line-of-sight halos depends on the lens and source redshifts, the amplitude of time delay perturbations should evolve with redshift, as should any resulting uncertainties and/or bias in the  $H_0$  value inferred from the time delays. Second, it is possible that perturbations from dark matter halos on lensing observables propagate to  $H_0$  inferences through indirect channels, not directly through the time delays. The collective lensing effects of a population of

\* gilmanda@ucla.edu

halos could be absorbed by the global mass model of the lens system, or the *macromodel*, in which case substructure would impart stochastic perturbations to the macromodel.

In this work, we quantify these effects by applying the same modeling assumptions used in cosmographic analyses to mock lenses that include full populations of subhalos and line-of-sight halos. Using the image positions and the surface brightness of lensed arcs to constrain the macromodel, we infer  $H_0$  from the mock lenses with lens models that do not include substructure, as is currently done in the literature. Comparing the inferred  $H_0$  with the ‘true’ value of  $H_0$  used to create the mocks, we quantify to what degree dark substructure biases or increases the uncertainties reported in cosmographic inferences, had it been included in the lens model. The mock lenses are based on six real quadruply-imaged quasars systems modeled by the HOLiCOW, STRIDES, and SHARP teams (from now on, we refer to this collaboration as TDCOSMO) that have delivered measurements of  $H_0$ , allowing us to estimate the effects of substructure for each individual system. By construction, our simulations isolate the role of substructure from other sources of uncertainty, such as the mass sheet transformation (Schneider & Sluse 2013) and systematic uncertainties associated with the lens macromodel (e.g. Kochanek 2019).

A necessary component of our analysis is the creation of hundreds of realistic mock lenses with full line-of-sight and subhalo populations. The dark matter halos in the mock systems simultaneously perturb the arrival times and the flux ratios. In addition, halos occasionally produce gravitational imaging residuals in extended images that lead to the discovery of individual dark matter halos in strong lens systems, a technique known as gravitational imaging (Koopmans 2005; Vegetti et al. 2010; Hezaveh et al. 2016b; Vegetti et al. 2018). Based on our simulations, we comment on results relevant for gravitational imaging, and in particular recent suggestions to measure the convergence power spectrum of dark matter halos (e.g. Hezaveh et al. 2016b; Cyr-Racine et al. 2016). As the focus of this work is time delay cosmography, we relegate discussion of these topics to Appendices B and D.

This paper is organized as follows: Section 2 describes our approach for estimating  $H_0$  uncertainties from simulations that include substructure in mock lenses. Section 3 describes the specific modeling choices implemented in the simulations. Section 4 describes the results of the analysis, which we summarize concisely in Section 5. We make concluding remarks in Section 6.

We perform gravitational lensing computations with *lenstronomy*<sup>1</sup> (Birrer & Amara 2018). Computations that involve the halo mass function and the matter power spectrum are executed with *colossus* (Diemer 2018). We assume a flat  $\Lambda$ CDM cosmology using the parameters from WMAP9 (Hinshaw et al. 2013) ( $\Omega_m = 0.28$ ,  $\sigma_8 = 0.82$ ), and set  $h = 0.733$  (see Section 2). The conclusions we present are independent of the choice of cosmological parameters.

## 2. Cosmographic inference on mock lenses that include substructure

This section describes our method to estimate the effect of dark substructure on cosmographic inferences. A brief outline is as follows: First, we create realistic mock datasets with substructure included in the lens models. Next, we analyze the data as would be done in a modern cosmographic inference, ignoring substructure in the model. Finally, we investigate how the

omission of dark substructure affects cosmographic inferences through comparison with a reference dataset that contains no substructure.

We begin by introducing the time delay distance  $D_{\Delta t}$ , and describe the observables from strong lens systems used to constrain it, in Section 2.1. In Section 2.2, we discuss the task of inferring  $H_0$  in the presence of full populations of dark matter halos in a Bayesian context. Section 2.3 describes how we create mock datasets with substructure included the lens model. Section 2.4 details our procedure for using these mock datasets to quantify any bias or additional uncertainties associated with the presence of substructure in strong lenses.

### 2.1. The time delay distance and strong lensing observables

Strong lensing constrains the Hubble constant through a measurement of an absolute distance scale  $D_{\Delta t}$ , referred to as the *time delay distance*  $D_{\Delta t} = (1 + z_d) \frac{D_d D_s}{D_{ds}}$ . The factors  $D_d$ ,  $D_s$ ,  $D_{ds}$  are angular diameter distances to the deflector, the source, and from the deflector to the source<sup>2</sup>, and  $z_d$  is the deflector redshift. The combination of cosmological distances entering  $D_{\Delta t}$  depends on cosmology, with a particularly strong dependence on  $H_0$ , with  $D_{\Delta t} \propto \frac{1}{H_0}$ .

The data vector  $\mathbf{d} = (\theta_{\text{img}}, \mathbf{t}, s)$  that constrains quantities such as  $D_{\Delta t}$  includes the four lensed quasar image positions  $\theta_{\text{img}} = (\theta_A, \theta_B, \theta_C, \theta_D)$ , the time delays between the images  $\mathbf{t} = (t_{12}, t_{13}, t_{14})$ , and the observed surface brightness  $s(\theta)$  at angular position  $\theta$  in the image plane.

All matter in the lens system, dark or baryonic, affects the time delays, image positions, and the arc surface brightness. Two equations encode the dependence of these observables on the deflection angles  $\alpha$ , and on angular diameter distances and cosmology. First, the ray tracing equation describes the propagation of light rays through lines of sight populated by dark matter halos (Blandford & Narayan 1986)

$$\theta_K(\theta) = \theta - \frac{1}{D_s} \sum_{k=1}^{K-1} D_{k,s} \alpha_k (D_k \theta_k). \quad (1)$$

This is a recursive relation for angular coordinates on the  $k$ th lens plane, given a coordinate on the sky  $\theta$ , the deflection angle  $\alpha$ , and the angular diameter distance from the  $k$ th lens plane to the source plane,  $D_{k,s}$ . Equation 1 gives the surface brightness in the image plane  $s(\theta)$  by ray-tracing through the lens model to the source plane, and the multiple images of the background quasar appear at the coordinates  $\theta$  that each map to the same position on the source plane.

The time delay between images located at coordinates  $\theta_A$  and  $\theta_B$  is

$$t_{AB} = \frac{D_{\Delta t}^{\text{eff}}}{c} (\phi^{\text{eff}}(\theta_A) - \phi^{\text{eff}}(\theta_B)). \quad (2)$$

where the effective Fermat potential is given by

$$\phi^{\text{eff}}(\theta_X) \equiv \sum_{k=1}^{K-1} \frac{1 + z_k}{1 + z_d} \frac{D_k D_{k+1} D_{ds}}{D_d D_s D_{k,k+1}} \times \left[ \frac{(\theta_k^X - \theta_{k+1}^X)^2}{2} - \xi_{k,k+1} \Psi(\theta_k^X) \right], \quad (3)$$

<sup>2</sup> Through the text  $D_{i,j}$  stands for an angular diameter distance from the  $i$ th lens plane to the  $j$ th. A subscript  $s$  stands for the source plane.

<sup>1</sup> <https://github.com/sibirrer/lenstronomy>

where  $\Psi(\theta)$  is the projected gravitational potential at position  $\theta$ ,  $\xi_{i,j} \equiv \frac{D_{i,j}D_s}{D_iD_{i,s}}$ , and where we use shorthand notation  $\theta_K(\theta_X) \equiv \theta_K^X$  for Equation 1 evaluated at a specific angular position  $\theta_X$ .

Assuming a cosmology, there is a unique distance scale  $D_{\Delta t}^{\text{eff}}$ , given a value of  $H_0$ . Following the standard approach (e.g. Wong et al. 2020; Shajib et al. 2020), we can define an effective time delay distance  $D_{\Delta t}^{\text{eff}}$  that normalizes the measured time delays and the model-predicted Fermat potential, and compute  $H_0$  from this effective distance scale.

A typical cosmographic analysis includes nuisance parameters  $\mathbf{n}$ , such as the lens mass profile, the lens light profile, the source light profile, and other factors that can affect the inference such as the microlensing time delay (Tie & Kochanek 2018), and the external convergence associated with large-scale structure near the lens and along the line of sight (Schneider & Sluse 2013; Birrer et al. 2016). Obtaining a  $D_{\Delta t}$  posterior from a strong lens system requires marginalizing the likelihood function over these nuisance parameters

$$p(\mathbf{d}|D_{\Delta t}) = \int \mathcal{L}(\mathbf{d}|D_{\Delta t}, \mathbf{n}) p(\mathbf{n}) d\mathbf{n}. \quad (4)$$

The evaluation of Equation 4 involves simultaneously reconstructing the lensed background source and the lensed image.

## 2.2. Directly computing $H_0$ posteriors with substructure

To incorporate substructure, we describe dark field halos and subhalos with population-level statistical properties specified by a set of hyper-parameters  $\mathbf{q}_{\text{cdm}}$ . These hyper-parameters describe, for example, the form of the field halo and subhalo mass functions, the average concentration-mass relation of halos, the probability distribution for their spatial distribution, etc., assuming Cold Dark Matter (CDM). Typically, cosmographic analyses ignore  $\mathbf{q}_{\text{cdm}}$  altogether when measuring  $D_{\Delta t}$ , assuming  $p(D_{\Delta t}|\mathbf{d}, \mathbf{q}_{\text{cdm}}) \approx p(D_{\Delta t}|\mathbf{d})$ . In other words, the assumption in the literature is that inferences on  $H_0$  do not depend on the nature of small-scale dark matter structure in the lens system, or at least that other factors in the analyses, such as the nuisance parameters  $\mathbf{n}$ , dominate the  $D_{\Delta t}$  error budget such that dark substructure can be safely ignored.

The direct route towards testing this assumption involves a daunting calculation of a likelihood function. We can account for dark substructure in Equation 4 by generating realizations of dark matter halos specified by parameters  $\mathbf{m}_{\text{sub}}$  from the model specified by hyper-parameters  $\mathbf{q}_{\text{cdm}}$ . The likelihood that accounts for  $\mathbf{q}_{\text{cdm}}$  includes an additional integral over all possible configurations  $\mathbf{m}_{\text{sub}}$

$$\mathcal{L}(\mathbf{d}|D_{\Delta t}, \mathbf{q}_{\text{cdm}}) = \int p(\mathbf{d}|D_{\Delta t}, \mathbf{m}_{\text{sub}}, \mathbf{n}) \times p(\mathbf{m}_{\text{sub}}|\mathbf{q}_{\text{cdm}}) p(\mathbf{n}) d\mathbf{m}_{\text{sub}} d\mathbf{n} \quad (5)$$

Incorporating substructure by directly evaluating the integral over  $\mathbf{m}_{\text{sub}}$  is not feasible because Equation 5 is computationally intractable. The problem stems mainly from integrating  $p(\mathbf{d}|D_{\Delta t}, \mathbf{m}_{\text{sub}}, \mathbf{n})$  over  $\mathbf{n}$ , because  $\mathbf{m}_{\text{sub}}$  includes potentially tens of thousands of halos whose deflection angles need to be continuously re-evaluated while performing the integral over  $\mathbf{n}$ .

## 2.3. Forward modeling substructure perturbations on lensing data

To circumvent the direct evaluation of the intractable integral in Equation 5, we instead create realistic mock datasets computed

with substructure in the lens model using a known ‘true’ value of  $H_0$ . We then apply same assumptions used in the literature (ignore substructure) to analyze these systems with the goal of recovering the ‘true’  $H_0$  value. The degree to which we fail to recover the ‘true’  $H_0$ , and the degree to which the uncertainties increase relative to a control sample with no substructure in the lens system, allowing us to quantify any bias or increased uncertainties associated with substructure.

The mock lenses we generate are realistic analogs of the lens systems RXJ1131-1231, PG1115+080, HE0435-1223, B1608+656, WFI2033-4723, and DESJ0408+5354, each of which has been used to measure  $H_0$  (see the discussion in Section 3.5 for details). Through the procedure outlined in the remainder of this section, we determine on a case-by-case basis to what degree omitting substructure from the lens model results in under-estimated uncertainties and/or bias in the inferred value of  $H_0$ .

To begin, we describe two sets of mock data created for each lens system: the *control data* and the *substructure-perturbed data*:

- *The control data*: The control data includes no substructure, and the same parameterization for the mass model is used to create the data as is used to fit the resulting mock lens. The uncertainty in  $H_0$  inferred from this dataset is determined entirely by the measurement precision of the time delays, and the degree to which the lensed images and extended arc constrain the mass profile. We tune the signal to noise in the mock data until the  $H_0$  inference we obtain from this dataset has the same error budget from time delays and lens modeling as those quoted in Table 3 of Wong et al. (2020), or the uncertainties reported by Shajib et al. (2020) in the case of DESJ0408.
- *The substructure-perturbed data*: We first generate a realization of dark matter field halos and subhalos. Next, starting from the lens model used to create the control dataset, we adjust the mass model of the main deflector to recover the same image positions as in the control data, with the subhalos and line-of-sight halos included in the system. We place the same extended background source used to create the control data at the new source position obtained from adjusting the macro-model, and proceed to compute the time delays and the surface brightness of the extended images. The mock lens obtained through this process is indistinguishable by eye from the control data, but the time delays and lensed arc include the subtle lensing signatures of dark matter substructure.

Both the control and substructure-perturbed models are created using the same value of  $H_0 = 73.3 \text{ km s}^{-1} \text{ Mpc}$ , the value quoted by Wong et al. (2020), and we apply the same Gaussian PSF with a FWHM of 50 milli-arcseconds to the simulated lenses, comparable to the imaging resolution from the Hubble Space Telescope. We stress that our results are independent of the specific choice of  $H_0$ .

## 2.4. Estimating $H_0$ uncertainties from un-modeled dark substructure

We fit the mock datasets using smoothly-parameterized mass profiles for the main deflector mass profile while ignoring substructure in the lens system. The model applied to the control data is the same as the model used to create the data, so we should fit the mock data down to the noise; this is a perfect model scenario. We then apply the same modeling assumptions

to the substructure-perturbed datasets, fitting them with the same smoothly-parameterized mass profile. To isolate the role of substructure, we assume perfectly-measured image positions and do not add astrometric uncertainties to the lensed point sources. The flux ratios are not used to constrain the lens model, although we can compute them jointly with the time delays (see Appendix D). We perform the lens modeling on 200 mock datasets, each of which corresponds to one realization of dark matter halos. This gives 200 corresponding posterior distributions of  $H_0$ , which we combine by adding them together.

In our procedure, mimicking the actual measurements, we include additional model complexity in the mock datasets, rather than in the model applied to the data. The requirement that must be met in order to safely make this approximation is that each of the substructure-perturbed datasets wield the same constraining power over  $H_0$ . We take two measures to ensure this is the case: first, because the mock ‘measured’ time delays change slightly from realization to realization in the substructure-perturbed datasets, a fixed uncertainty in units of days translates to different percent precision in  $H_0$ . To keep the same constraining power of the time delays the same for each individual realization, we operate with *relative* uncertainties on the time delays, using the same uncertainties as those measured from the real lens systems studied in the literature.

Second, by re-weighting posterior samples we ensure that we operate at fixed uncertainty in the logarithmic slope of the main deflector mass profile  $\gamma$ , and the Einstein radius  $R_{\text{Ein}}$ . Operating at fixed uncertainty in  $\gamma$  and  $R_{\text{Ein}}$  serves two purposes: first, power-law mass models fit to the lens system HE0435-1223 (to use a specific example) have logarithmic slopes  $\gamma = 1.9 \pm 0.05$  (Chen et al. 2019). Substructure realizations that lead to global mass profiles with profile slopes substantially different than this are therefore not realistic analogs of the lens in question. Quantitatively, the realizations that lead to this mismatch between the model and the data do not contribute to the integral in Equation 5, because the likelihood associated with these realizations would be extremely low. This argument applies to both  $\gamma$  and  $R_{\text{Ein}}$ . The second reason for operating at fixed  $\gamma$  and  $R_{\text{Ein}}$  is that both of these parameters are, in principle, constrained by the stellar kinematics of the main deflector. Enforcing equality between the  $\gamma$  and  $R_{\text{Ein}}$  posteriors ensures that the modeled-predicted velocity dispersion matches the velocity dispersion that would have been ‘measured’ from the mock datasets. By incorporating information that could, in principle, be constrained by kinematic data, we isolate the role of substructure in the lens system.

After this series of steps, we are ready to make comparisons between the  $H_0$  inference from the control data and the inferences from the substructure-perturbed data. If substructure is irrelevant for time delay cosmography, then we expect the  $H_0$  posteriors obtained from the control data and the substructure-perturbed dataset to be identical. The degree to which the  $H_0$  posteriors differ allows us to quantify how the reported values of  $H_0$  from the TDCOSMO analyses would have been affected, had substructure been included in the lens model. Compared to the control data, the  $H_0$  inferences from the substructure-perturbed data will be inflated by some amount  $\sigma_{\text{sub}}$ , independent of the other sources of uncertainty  $\sigma_{\text{reported}}$ . The total uncertainty is therefore

$$\sigma_{\text{total}}^2 = \sigma_{\text{reported}}^2 + \sigma_{\text{sub}}^2. \quad (6)$$

A bias in  $H_0$  will manifest as different medians between the control and substructure-perturbed posteriors.

### 3. Simulation Setup

The procedure outlined in the previous section requires a prescription for rendering populations of dark matter halos, and models for the lens mass distribution, lens light profile, and source light profile. We describe the parameterizations and modeling choices for the subhalo and line-of-sight halo populations in Sections 3.1. Section 3.2 describes how subhalos and line-of-sight halos are generated symmetrically around the light rays throughout the lensing volume, and the addition of negative convergence profiles to correct for the additional mass added in substructure. Sections 3.3 and 3.4 describe the models we implement for the main deflector mass profile and the background source light profile, respectively. We reiterate that the lens models we use omit complexity in the form of external convergence or more complicated mass profiles in order to isolate the perturbative effects of substructure.

#### 3.1. The CDM subhalo and field halo populations

We use the same parameterizations for the CDM mass function as those implemented by (Gilman et al. 2020a,b); we briefly review the essential features in the following paragraphs.

Both subhalos and field halos along the line of sight, are modeled as truncated NFW profiles (Navarro et al. 1996; Baltz et al. 2009) defined with a density contrast of 200 with respect to the critical density at  $z = 0$ . We use the mass-concentration relation presented by Diemer & Joyce (2019). The truncation radius for subhalos depends on the mass of the subhalo, and the position inside the host halo, while field halos are truncated at  $r_{50}$ . The truncation of field halos at radii  $> r_{200}$  keeps the total mass in line-of-sight halos finite with a negligible impact on their lensing properties.

We draw field halo masses from a Sheth-Tormen mass function (Sheth et al. 2001) with the addition of the two-halo term  $\xi_{2\text{halo}}(M_{\text{halo}}, z)$

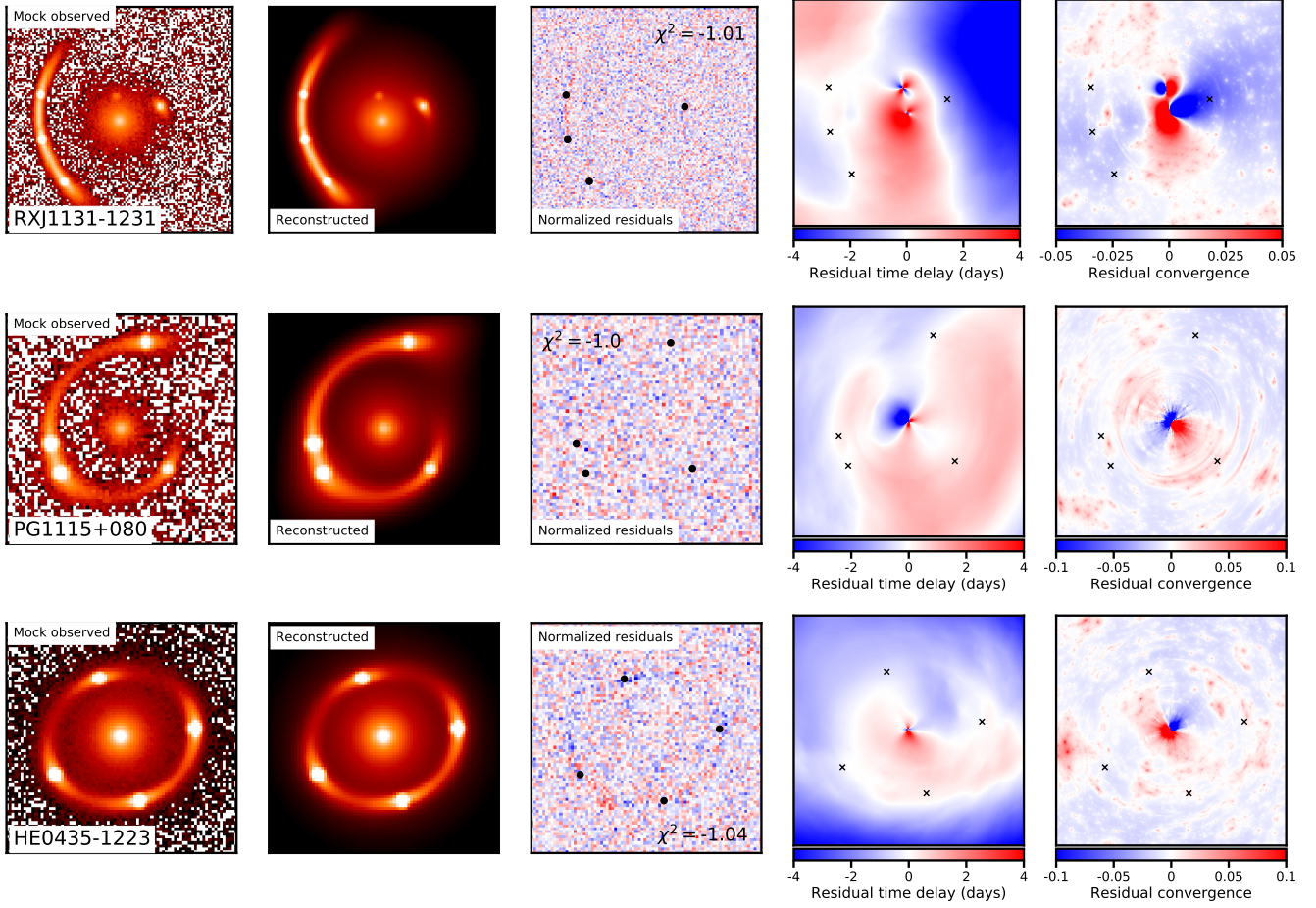
$$\frac{d^2 N_{\text{los}}}{dmdV} = (1 + \xi_{2\text{halo}}(M_{\text{host}}, z)) \frac{d^2 N}{dmdV} \Big|_{\text{ShethTormen}}. \quad (7)$$

The two-halo term boosts the amount of structure near the host halo, which has a mass  $M_{\text{host}}$ . We render subhalos from a power-law mass function defined in projection

$$\frac{d^2 N_{\text{sub}}}{dmdA} = \frac{\Sigma_{\text{sub}}}{m_0} \left( \frac{m}{m_0} \right)^\alpha \mathcal{F}(M_{\text{host}}, z) \quad (8)$$

with logarithmic slope  $\alpha = -1.9$  (Springel et al. 2008) and a pivot mass  $m_0 = 10^8 M_\odot$ . The function  $\mathcal{F}(M_{\text{host}}, z)$  encodes the evolution of the projected number density of subhalos with host halo mass and redshift (see Gilman et al. 2020a), and is calibrated using the semi-analytic modeling code *galacticus* (Benson 2012). It is given by  $\log_{10}(\mathcal{F}) = k_1 \log_{10}\left(\frac{M_{\text{host}}}{10^{13.5} M_\odot}\right) + k_2 \log_{10}(1+z)$ , with  $k_1 = 0.88$  and  $k_2 = 1.7$ . We assume a host halo mass of  $10^{13.3} M_\odot$  for each mock lens, based on the population mean halo mass for strong lensing galaxies inferred by Lagattuta et al. (2010). We use a normalization  $\Sigma_{\text{sub}} = 0.025 \text{ kpc}^{-2}$  based on the measurements from (Gilman et al. 2020a). These modeling choices are intended to be representative of the subhalo populations around massive elliptical galaxies.

We render both halos and subhalos in the mass range  $10^6 - 10^9 M_\odot$ . We have carried out tests with a lower minimum mass threshold, and verify that our results do not depend on this choice. We assume halos more massive than  $10^9$  would host a visible galaxy, and would therefore be modeled explicitly in



**Fig. 1.** The panels show the mock observed lens that includes substructure in the lens model (far left), the reconstructed lensed image (second from left), the normalized imaging residuals (center), the residual time delay surface (second from right), and the residual convergence (far right). The residual time delay surface is the ‘true’ time delay surface from the mock lens minus the model-predicted time delay surface from the lens model fit to the data. Similarly, the residual convergence is the full multi-plane convergence (Equation 10) minus the convergence from the model fit to the mock data. The lensed quasar image positions are marked as points (crosses) in the center (right) panels. Each row corresponds to one substructure-perturbed dataset for the lens systems based on RXJ1131-1231, PG1115+080, and HE0435-1223.

cosmographic analyses. The choice of  $10^9$  solar masses is conservative in the sense that a larger mass would increase the uncertainties we attribute to un-detected substructure. The rendering volume is a double-cone geometry with an opening angle of 15 times the Einstein radius of each lens. We have tested larger opening angles, and verify that our results are unchanged.

### 3.2. Lens cone symmetry and convergence corrections

At each lens plane along the line of sight<sup>3</sup>, we generate halos uniformly distributed in two dimensions around a coordinate center that tracks the path of a lensed light ray backwards from the observer towards the lens mass centroid. Behind the main deflector, the rendering area tracks the light as it is deflected by background galaxies, as is the case in HE0435, WFI2033, and DESJ0408. Without shifting the rendering area to track lensed light rays, substructure is not distributed symmetrically around the path of the light, and the asymmetric mass distribution biases the inferred value of  $H_0$ . The shifting of substructure along a lensed light path is particularly important for lens systems with large satellites and/or line-of-sight galaxies, as the source coordinates in these systems can be offset from the lens center by as much as 1 arcsec, comparable to a typical lens Einstein radius.

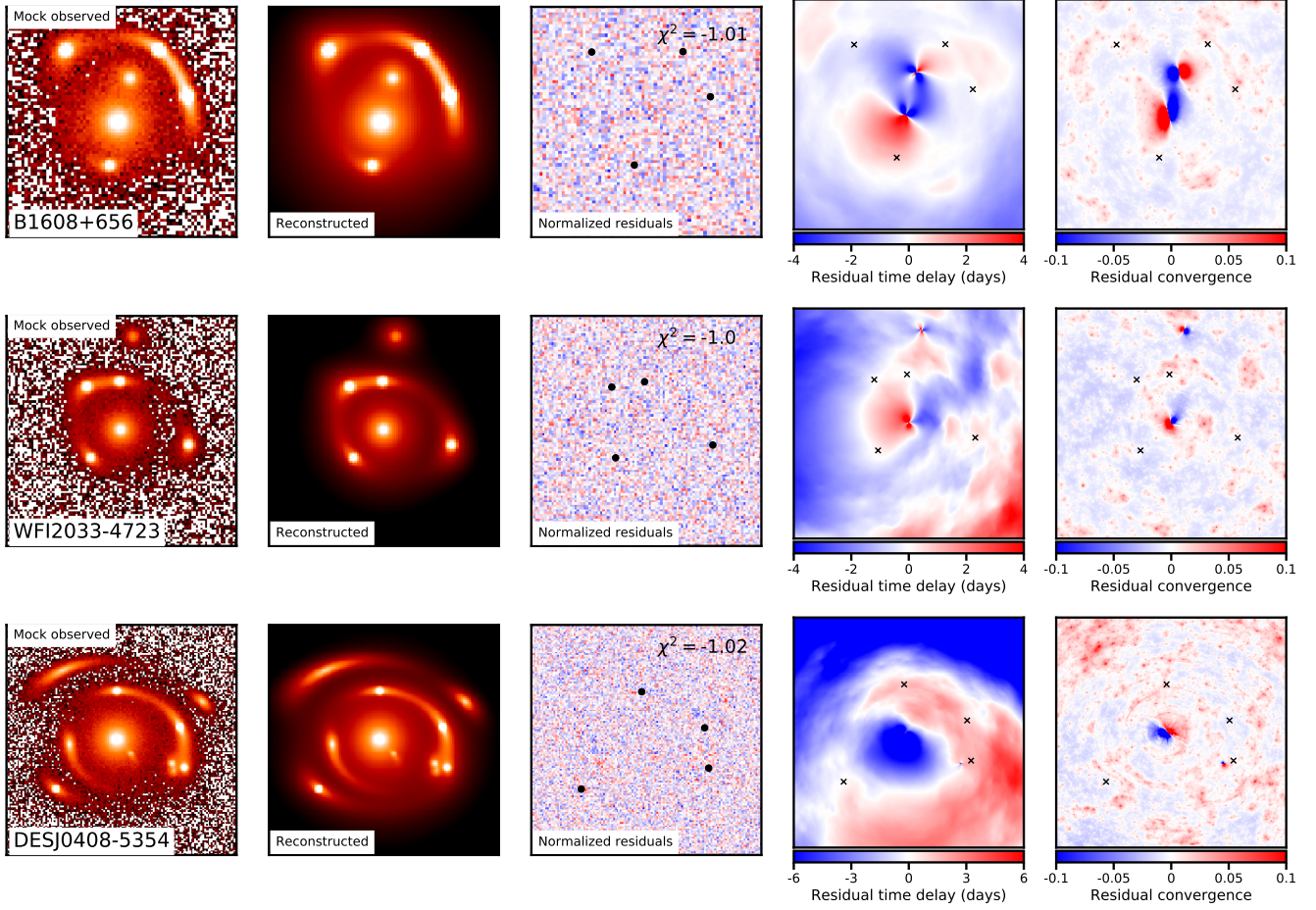
ordinate in these systems can be offset from the lens center by as much as 1 arcsec, comparable to a typical lens Einstein radius.

We add negative convergence at each lens plane to subtract the mean mass rendered in (sub)halos, ensuring that the average dark matter density in the rendering volume is equal to the mean dark matter density of the Universe. For line-of-sight halos, the correction is in the form of a uniform (negative) convergence sheet. The correction is more complicated for subhalos, as they are spatially distributed following a cored NFW profile. The cored spatial distribution, with a central core radius set to half the scale radius of the host, is intended to mimic the tidal disruption of subhalos that pass close to the central galaxy. The net convergence profile that needs to be subtracted for the subhalos is formally given by a convolution of the subhalo density profile with their spatial distribution function, but we are able to approximate this profile with a simple cored NFW profile normalized such that the residual convergence from substructure near the Einstein radius is zero on average.

Omitting the mass sheet corrections trivially biases  $H_0$  by re-scaling each time delay through the addition of a net positive convergence in the form of dark matter halos. This effect is similar to the external convergence that is accounted for in cosmographic inferences, which affects the inferred time delay

<sup>3</sup> Lens planes are placed uniformly in redshift with a spacing of 0.02.





**Fig. 2.** The figure is the same as Figure 1, but shows examples of the mock observed and reconstructed lens (left), the imaging residuals (center), the residual time delay surface (second from right), and the residual convergence (far right) maps for the lens systems B1608+656, WFI2033-4723, and DESJ0408-5354.

distance as

$$D_{\Delta t} = \frac{D_{\Delta t, \text{model}}}{1 - \kappa_{\text{ext}}}. \quad (9)$$

While the ensemble of substructure-perturbed datasets have a residual convergence of zero on average, on a realization-by-realization basis the net residual convergence from substructure fluctuates around the mean dark matter density in the Universe. This introduces a mass-sheet transformation from random populations of dark matter halos that is distinct from the external convergence  $\kappa_{\text{ext}}$ , which comes from large-scale structure in the lensing volume. The fluctuations of the substructure convergence cause fluctuations in the model-predicted time delays, and therefore contribute to the  $H_0$  error budget.

### 3.3. The main deflector mass and light profiles

Since our goal is to isolate the effects of substructure, we use the same mass and light profiles to create the control and substructure-perturbed datasets as we use in the lens models fit to these data. The main deflector is modeled as a power-law ellipsoid with logarithmic mass profile slope  $\gamma$ , plus external shear  $\gamma_{\text{ext}}$ . To keep the analogs as similar to the real lenses as possible, we choose a logarithmic profile slope  $\gamma$  consistent with the mean value inferred in the lens modeling by TDCOSMO.

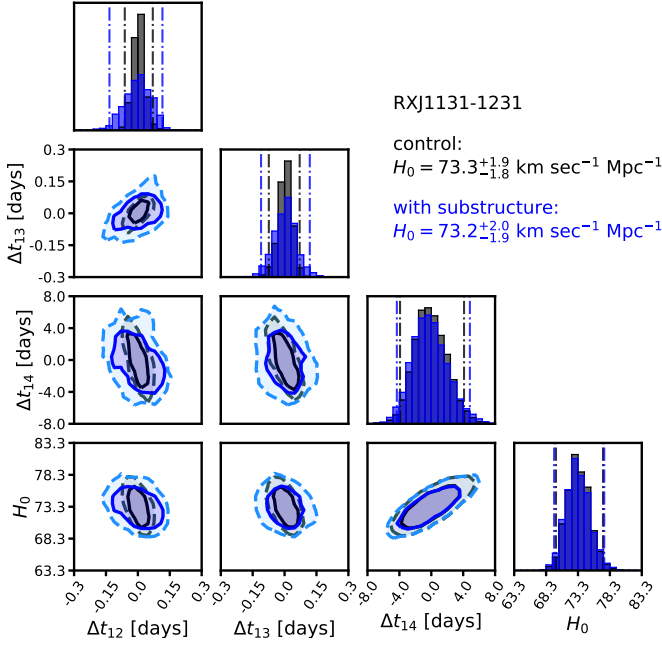
All of the lens systems we consider have either a nearby satellite galaxy in the main lens plane, a massive galaxy along

the line of sight near the main deflector (in projection), or both. We model these objects as singular isothermal spheres, using Einstein radii and positions inferred by the lens modeling from the TDCOSMO teams. We add an elliptical Sérsic light profile for the main deflector light, and circular Sérsic light profiles for the satellite or nearby galaxies. The parameters describing the macromodels used to create the mock lenses are listed in Table A.1.

### 3.4. Background source light profiles

The mock datasets are all created using an elliptical Sérsic model for the background source. The exact parameters describing these profiles are unique to each lens, and are chosen such that the lensed arc in each mock resembles the extended features in the real lens it is based on.

Modeling lenses for cosmography involves a simultaneous reconstruction of the background source light and the lensed image. Using an elliptical Sérsic profile to model the background source (the same model as is used to create the mock lens), we detect significant residuals in the reconstructed lensed images that result from substructure in the lens system. Following the approach taken in the literature, when we detect these residuals we add additional complexity to the background source. We in-



**Fig. 3.** The joint posterior distribution of the residual time delays  $\Delta t$ , defined as the model-predicted time delay minus the ‘true’ time delay, and the Hubble constant, computed for the lens system RXJ1131-1231. The contours show the 68% and 95% confidence intervals for the control posterior that includes no substructure (black), and the the combined posteriors from the inferences on lens models that include substructure (blue). Dashed lines in the histograms show 95% confidence intervals. The median and 68% confidence intervals of the  $H_0$  inference are printed above the figure.

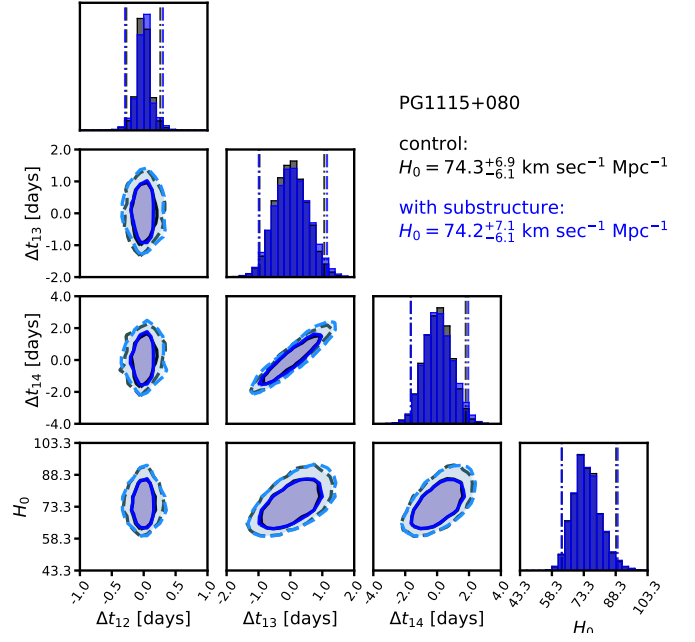
corporate this complexity by adding shapelets<sup>4</sup> (Refregier 2003; Birrer et al. 2015). The amount small-scale detail in the background source is controlled by the parameter  $n_{\max}$ , which we set to  $n_{\max} = 8$ . In the case of DESJ0408+5354, we follow Shajib et al. (2020) and add additional set of shapelets with  $n_{\max} = 3$  to one of the additional lensed sources.

### 3.5. Discussion of individual lens systems

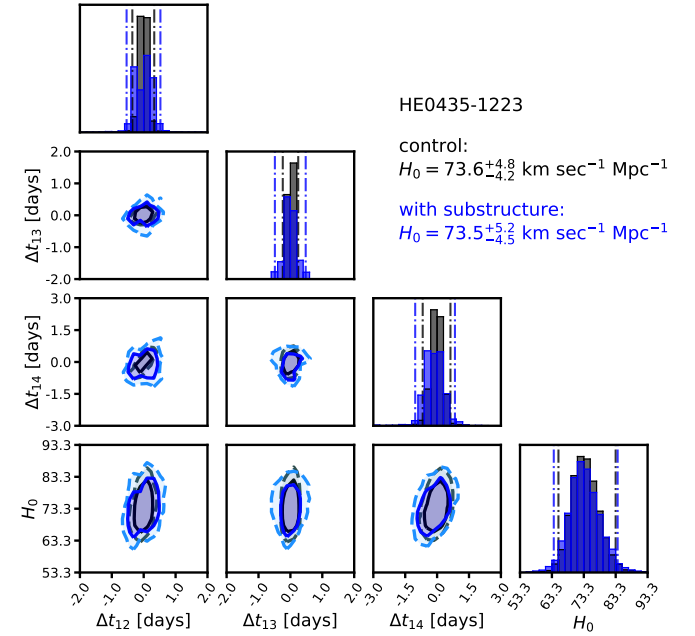
We construct realistic analogs of six strong lens systems analyzed by the TDCOSMO collaboration. We briefly comment on each system in following paragraphs, highlighting both cosmographic inferences and substructure lensing analyses performed with each system.

- RXJ1131-1231 This quad was discovered by Sluse et al. (2003), and consists of source at  $z = 0.658$  lensed by a galaxy at  $z = 0.295$  (Sluse et al. 2003, 2007). The time delays are measured at  $\sim 1.6\%$  precision by Tewes et al. (2013), and exhibit a relatively long maximum delay of  $\sim 90$  days. This system has been modeled extensively with the aim of constraining cosmology (Suyu et al. 2013, 2014; Birrer et al. 2016; Chen et al. 2016, 2018). The lens mass model includes a satellite galaxy inside the Einstein radius. Perturbations from substructure in the lensed arc of this system were analyzed by Birrer et al. (2017), and were used to place a lower limit on the mass of a thermal relic dark matter particle of  $2.2\text{keV}$ .

<sup>4</sup> See Appendix B for a discussion of how additional structure in the background source model can absorb the perturbations from substructure in extended images.

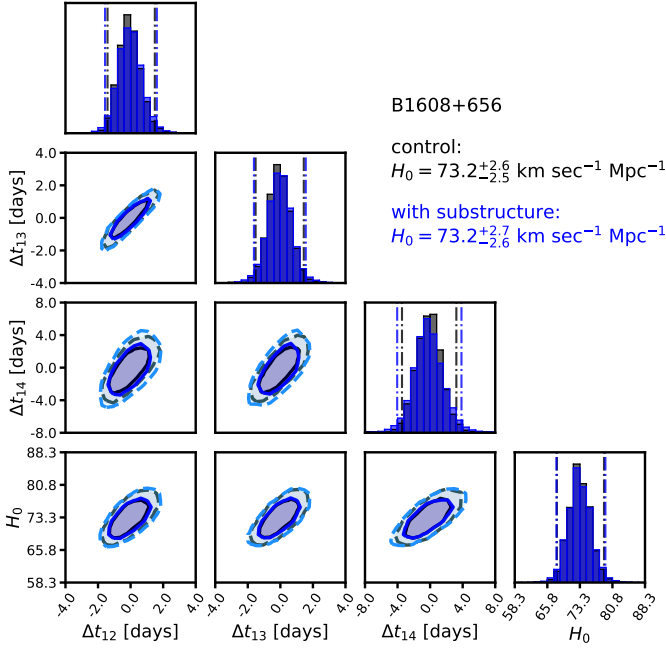


**Fig. 4.** The figure is the same as Figure 3, but shows the posteriors of the residual time delays and  $H_0$  computed for the lens system PG1115+080.

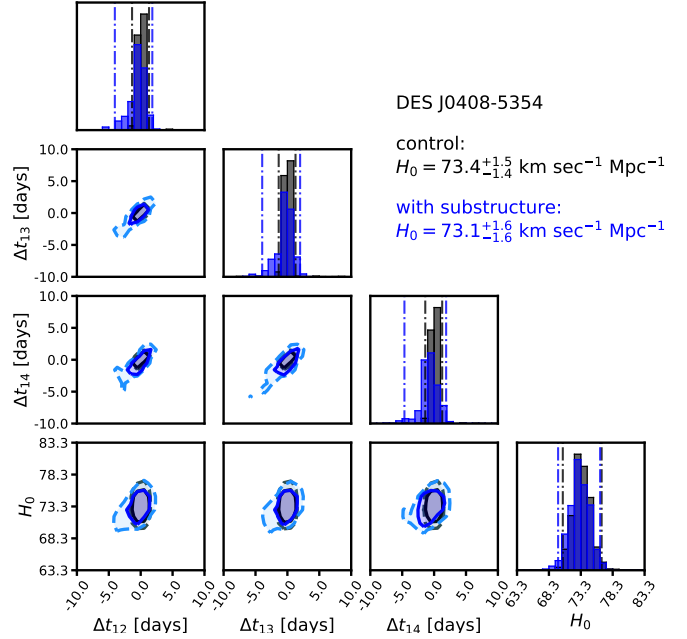


**Fig. 5.** The figure is the same as Figure 3, but shows the posteriors of the residual time delays and  $H_0$  computed for the lens system HE0435-1223.

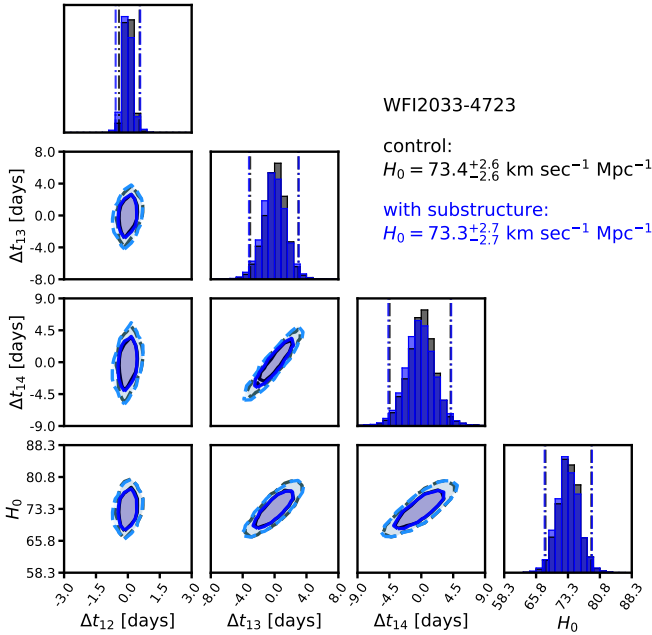
- PG1115+080 This quad was discovered by Weymann et al. (1980), and consists of a source at redshift  $z = 1.722$  lensed by a galaxy at  $z_d = 0.310$  (Tonry 1998). The time delays were measured at  $\sim 6.4\%$  precision by Bonvin et al. (2018b), and the longest delay is  $\sim 18$  days. This system was recently modeled with both adaptive optics and HST imaging by (Chen et al. 2019). We explicitly model the nearby cluster of galaxies as a single SIS profile projected  $\sim 10$  arcsec away from the lens center. This system was first used to measure the abundance of dark subhalos by Dalal & Kochanek (2002), and more recently was used to constrain the free-streaming length of dark matter Hsueh et al. (2019); Gilman



**Fig. 6.** The figure is the same as 3, but shows the posteriors of the residual time delays and  $H_0$  computed for the lens system B1608+656.



**Fig. 8.** The figure is the same as 3, but shows the posteriors of the residual time delays and  $H_0$  computed for the lens system DESJ0408-5354.



**Fig. 7.** The figure is the same as 3, but shows the posteriors of the residual time delays and  $H_0$  computed for the lens system WFI2033-4723.

et al. (2020a), and the mass-concentration relation of CDM halos (Gilman et al. 2020b).

- HE0435+1223 This quad was discovered by Wisotzki et al. (2002), with lens/source redshifts of  $z = 0.45$  and  $z_s = 1.69$ , respectively (Morgan et al. 2005). The time delays were measured at  $\sim 5.3\%$  precision (Bonvin et al. 2017), with a maximum delay of  $\sim 14$  days. This system was analyzed for cosmography by Bonvin et al. (2017); Wong et al. (2017); Chen et al. (2019). We explicitly model the galaxy located behind the main deflector at  $z = 0.78$  as an SIS profile, placing it at an angular position corrected by  $\sim 0.65$  arcsec due to fore-

ground lensing effects from the main deflector<sup>5</sup>. Nierenberg et al. (2017) measured narrow-line flux ratios for this system (see also Nierenberg et al. 2020), and used them to look for substructure near the images. The narrow-line flux ratio measurements for this system were also used to measure the halo mass function and mass-concentration relation of CDM halos (Gilman et al. 2020a,b).

- B1608+656 This lens system was discovered by Myers et al. (1995), with time delay measurements and early cosmographic inferences carried out by Fassnacht et al. (2002) and Koopmans et al. (2003). The lens/source redshifts for this system are  $z_d = 0.45$ , and  $z_s = 1.69$  (Fassnacht et al. 1996; Myers et al. 1995). The longest delay in this system is  $\sim 76$  days, measured at  $\sim 2\%$  precision. This system is unique in that main deflector consists of two merging galaxies, although (Suyu et al. 2009) show it is still well-described by a single power law profile. We follow Suyu et al. (2009) and model the larger object as a power law ellipsoid, and included smaller object as an SIS profile.
- WFI2033-4723 This quad was discovered by Morgan et al. (2004). The lens and source lie at  $z = 0.66$  and  $z = 1.66$ , respectively (Sluse et al. 2019). The time delays were measured by Bonvin et al. (2019) with a precision of  $\sim 2\%$ , and a maximum delay of  $\sim 60$  days. The main lensing galaxy has a small satellite located  $\sim 2$  arcsec away from the lens centroid. In addition, there is a galaxy located behind the main deflector at  $z = 0.745$  situated  $\sim 3.5$  arcseconds away from the lens center<sup>6</sup>. This system was modeled for cosmography by Rusu et al. (2019), and for substructure lensing inferences by (Gilman et al. 2020a,b).
- DESJ0408-5354 This lens system stands out from the rest due to the presence of two strongly-lensed sources, in addition to the quadruply-imaged quasar at  $z = 2.355$ , discovered

<sup>5</sup> For additional discussion, see Section 5.8 in Gilman et al. (2020a)

<sup>6</sup> As in the case of HE0435+1223, since this galaxy lies behind the main lens plane the coordinate listed in Table A.1 is corrected for foreground lensing effects.



by Lin et al. (2017). The time delays for this system were measured by Courbin et al. (2018) at a precision of 1% with the longest delay  $\sim 150$  days. Shajib et al. (2020) modeled this system for cosmography, and exploited the additional information provided by the multiple lensed sources to measure  $H_0$  at 3.8% precision (not including substructure). The main deflector is accompanied by a satellite galaxy that splits one of the lensed images, and several other galaxies along the line of sight. Shajib et al. (2020) modeled six of these line-of-sight galaxies, but in our mock lens models we only include the nearest two (G3 and G4, using the naming convention from Shajib et al. (2020)). This simplification does not affect our conclusions, since we are matching the precision of the original inference by construction.

We do not explicitly model doubly-imaged quasar SDSS 1206+4332, which was analyzed for cosmography by Birrer et al. (2019), because the analysis pipeline we develop is designed for quadruple-image lenses. However, based on the results obtained for the six systems we do model, we can estimate the fractional uncertainty in  $H_0$  from substructure in SDSS 1206+4332 (see Section 4.2).

The control lens models used for each system are summarized in Table A.1.

## 4. Results

In this section we apply the procedure outlined in Section 2 to the six strong lens systems described in the previous section. In Section 4.1, we show examples of the mock and reconstructed lensed images for each system, as well as maps of the residual time delay surface and convergence. Section 4.2 presents the main results of this analysis, in which we quantify how unmodeled dark substructure contributes to uncertainties and/or bias in  $H_0$ .

### 4.1. Reconstructed lenses: imaging residuals, time delay anomalies, and convergence

Figures 1 and 2 show examples of the mock observed and reconstructed lenses for each of the six systems we analyze. The panels on the far left show the mock observed lens system, which includes substructure. The reconstructed lens shown in the second from left panels includes no substructure. The central panels show the imaging residuals between the mock observed lens and the reconstruction, and displays the reduced  $\chi^2$  obtained from the fit to the imaging data. We discard realizations with reduced  $\chi^2 < -1.1$ , as a structure massive enough to produced large imaging residuals would likely be detected and explicitly modeled. In Appendix B we comment on the implications of these imaging residuals for gravitational imaging of dark substructure, and in particular on the implications for recent proposals to measure the convergence power spectrum of halos through gravitational imaging residuals.

The second from right panels show the residual time delay surface, which is the difference between the ‘true’ time delay surface of the control dataset, and the model-predicted time delay surface from the reconstructed lens system. Dark matter halos imprint small perturbations on the time delay surface, which, if they were somehow detected, could be interpreted as ‘time delay anomalies’, similar to the ‘flux ratio anomalies’ attributed to dark substructure.

The far right panels of Figures 1 and 2 show the residual convergence<sup>7</sup> between the full lens system that includes substructure, and the smooth model fit to the mock data. We use the definition of an effective convergence in terms of the deflection angles  $\alpha$  (Gilman et al. 2019)

$$\kappa_{\text{eff}} = \frac{1}{2} \left( \frac{\partial \alpha}{\partial x} + \frac{\partial \alpha}{\partial y} \right), \quad (10)$$

which generalizes lensing convergence to systems with line-of-sight structure<sup>8</sup>. The residual convergence shown in the right-most panels of Figures 1 and 2 is simply  $\kappa_{\text{eff}} - \kappa_{\text{reconstructed}}$ , where  $\kappa_{\text{reconstructed}}$  is the convergence from the smooth model fit to the mock data and  $\kappa_{\text{eff}}$  is computed from the mock lens system that includes substructure. Structures in the convergence maps that are tangentially sheared around the Einstein radius come from halos located between the main deflector and the source that produce highly nonlinear lensing effects.

### 4.2. $H_0$ inferences with dark substructure

Figures 3, 4, 5, 6, 7 and 8 show the joint posterior distribution for the inferred value of  $H_0$  and the residual time delays  $\Delta t$ , defined as the true time delay from the mock data minus the model-predicted time delay. The posterior from the control data is shown in black, and the posterior obtained from the substructure-perturbed datasets are shown in blue. The joint posterior distributions of  $H_0$  and the lens macromodel parameters are shown in Appendix C.

The median inferred value of  $H_0$  can differ from the ‘true’ value of  $H_0 = 73.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$  if the  $H_0$  posterior is asymmetric, as is the case for PG1115+080. This does not suggest the presence of a bias in the inference, because a bias would manifest as different medians between the control and substructure-perturbed posteriors. We plot the difference between the medians of the two posteriors as a function of the lensing volume divided by the longest time delay (the significance of these units will be discussed in the following paragraph) in Figure 9. The median values of the two posteriors are statistically consistent to within the error bars, which are given the uncertainty of the control dataset divided by  $\sqrt{N}$ , where  $N = 200$  is the total number of realizations. We conclude that the omission of substructure from strong lens models used for cosmography does not introduce a bias greater than 0.3%.

While we find no evidence for a bias in  $H_0$ , substructure alters the lens model and changes the Fermat potential at the image positions, which leads to an increased scatter in the time delays that factors directly into the  $H_0$  error budget. The additional uncertainty caused by substructure evolves with redshift and geometry in a predictable way. From Equation 2, the fractional uncertainty in  $H_0$  scales as

$$\frac{\delta H_0}{H_0} \propto \frac{\delta t}{t_{\text{max}}} \quad (11)$$

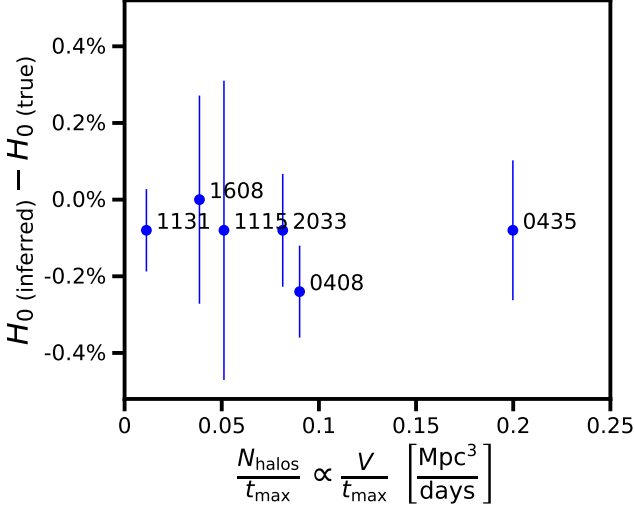
where  $t_{\text{max}}$  is the longest time delay, which is usually the most precisely measured and therefore the most constraining observable, and  $\delta t$  is a perturbation to the time delays from dark matter halos. Using analytic arguments, Keeton & Moustakas (2009) and Cyr-Racine et al. (2016) showed that  $\delta t$  is proportional to

<sup>7</sup> Convergence is typically defined as the projected mass in units of the critical density for lensing.

<sup>8</sup> The deflection angle  $\alpha$  is just the summation term to the right of the minus sign in Equation 1

**Table 1.** Summary of the sources of uncertainty affecting cosmographic inferences. This table is the same as Table 3 in the paper by Wong et al. (2020), plus the additional uncertainty term from substructure.

source of uncertainty	RXJ1131-1231	PG1115+080	HE0435-1223	B1608+656	WFI2033-4723	DESJ0408-5354
time delays	1.6%	6.4%	5.3%	1.7%	2.9%	1.8%
external convergence	3.3%	2.7%	2.8%	6.4%	5.7%	3.3%
lens modeling	2.2%	5.7%	2.5%	3.0%	2.2%	1.0%
substructure	0.6%	1.2%	2.4%	0.9%	1.2%	1.1%
$\sigma_{\text{reported}}$	4.3%	9.0%	6.5%	7.3%	6.8%	3.9%
$\sigma_{\text{total}}$	4.3%	9.1%	6.9%	7.4%	6.9%	4.1%

**Fig. 9.** The y-axis shows the bias in  $H_0$ , defined as the median of the  $H_0$  posterior obtained from the substructure-perturbed datasets minus the median of the posterior from the control datasets. The mock lenses are sorted on the x-axis by the lensing volume divided by the longest time delay, a dimensional quantity that determines the additional uncertainty caused by substructure in the lens system. The error bars show the statistical uncertainty on the median, which is given by  $\frac{\sigma_{\text{control}}}{\sqrt{N}}$ , where  $\sigma_{\text{control}}$  is the uncertainty in  $H_0$  from the control dataset and  $N = 200$  is the number of realizations. We find no evidence for bias in the inferred value of  $H_0$  stemming from substructure that is not modeled in cosmographic inferences.

$\sqrt{f_{\text{sub}}}$ , where  $f_{\text{sub}}$  is proportional to the total number of subhalos. We can generalize these results to realizations with line-of-sight halos by replacing  $f_{\text{sub}}$  by  $N_{\text{halo}}$ , where  $N_{\text{halo}}$  is just the total number of halos. The total number of halos is proportional to the lensing volume  $V$ , so the scaling in the fractional uncertainty in  $H_0$  is

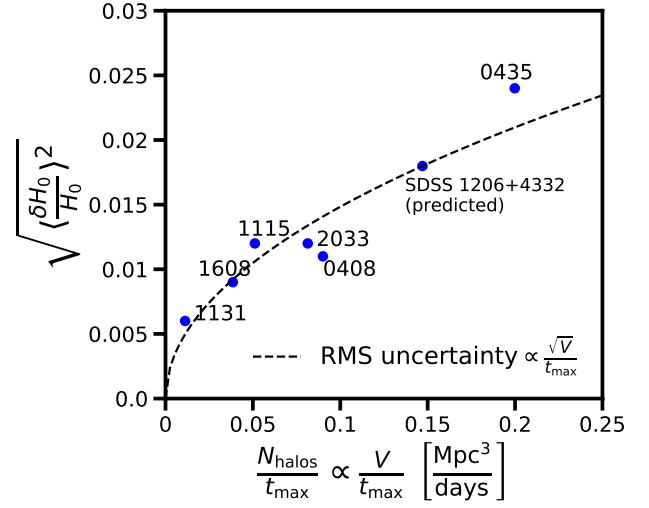
$$\frac{\delta H_0}{H_0} \propto \frac{\sqrt{V}}{t_{\text{max}}}. \quad (12)$$

The volume  $V$  is defined in terms of the comoving distances to the lens  $d_{\text{lens}}$ , to the source  $d_{\text{src}}$ , and the Einstein radius (converted to radians)  $R_{\text{Ein}}$

$$V = \pi (d_{\text{lens}} R_{\text{Ein}})^2 d_{\text{src}}. \quad (13)$$

Figure 10 shows the trend predicted by Equation 12 alongside the fractional uncertainties obtained for each lens. The dashed line fit in Figure 10 is given by

$$\frac{\delta H_0}{H_0} = 0.026 \left( \frac{V}{1 \text{ Mpc}^3} \right)^{0.5} \left( \frac{10 \text{ days}}{t_{\text{max}}} \right). \quad (14)$$

**Fig. 10.** The r.m.s. fractional uncertainty in  $H_0$  (y-axis) is plotted as a function of the number of halos divided by the longest time delay, which is proportional to the lensing volume divided by the longest time delay (x-axis). The labeled blue points show the fractional uncertainty for each of the six systems analyzed in this work. The dashed line shows the fitting function given by Equation 14. We use the fitting function in Equation 14 to estimate the contribution of substructure to the uncertainty in  $H_0$  for the lens system SDSS 1206+4332.

This simple relation captures the leading-order effects of substructure, but there are additional factors that could affect the uncertainties and produce scatter around the  $\frac{\sqrt{V}}{t_{\text{max}}}$  curve. For example, Oguri (2007) noted that lenses with symmetric image configurations are more sensitive to gravitational potential perturbations than lenses with asymmetric image configurations. Also, the Sheth-Tormen halo mass function in the range  $10^6 - 10^9 M_{\odot}$  has a mild redshift dependence that we do not account for in the fitting function.

We can use Equation 14 to estimate the fractional uncertainty in  $H_0$  given the lens redshift, the source redshift, the Einstein radius, and the longest time delay. For the doubly-imaged system SDSS 1206+4332, Equation 14 gives an uncertainty of 1.8%. This system was not analyzed explicitly in this work since the analysis pipeline is designed for quadruple-image lenses, but the physical arguments that motivate the scaling  $\frac{\sqrt{V}}{t_{\text{max}}}$  should still apply to SDSS 1206+4332.

We summarize the total error budget for each lens system, including the additional source of uncertainty from dark substructure, in Table 1. For a joint inference from a sample of lenses, we compute the contribution to the error budget from substructure

$\sigma_{\text{net}}$ , given by  $\sigma_{\text{net}}^2 = (\sum \sigma_i^2)^{-1}$ , where  $\sigma_i$  are the uncertainties from substructure of each lens system included in the joint inference. Based on the results for the lenses we explicitly model summarized in Table 1, plus the 1.8% uncertainty estimated using Equation 14 for SDSS 1206+4332, we find  $\sigma_{\text{net}} = 0.5\%$ . This is an independent source of uncertainty that should be summed in quadrature with other contributions.

## 5. Summary

We have carried out an analysis of how dark matter subhalos and field halos along the line of sight affect inferences of the Hubble constant  $H_0$  obtained through time delay cosmography. Using mock lenses based on six quadruply-imaged quasars modeled by the TDCOSMO collaboration, we quantify on a case-by-case basis how dark substructure would have affected the precision and accuracy of cosmographic inferences from each system, had substructure been included in the model. Our main results are summarized as follows:

- We show that excluding dark substructure from lens models used to infer cosmological parameters does not bias inferences on  $H_0$  at the level of the 0.3%, the numerical and statistical precision of our work. This rules out the possibility that explicitly modeling substructure in cosmographic inferences would alleviate the  $H_0$  tension between early and late-Universe probes. It also suggests that substructure cannot explain the anticorrelation between the lens redshift and the inferred time delay distance noted by Wong et al. (2020).
- Substructure perturbs the image arrival time delays and the global mass model, which translates into uncertainties on the value of  $H_0$  derived from the time delays and the Fermat potential predicted by the lens model. The additional uncertainties from substructure listed in Table 1 range from 0.6% in the case of RXJ1131-1231 to 2.4% in the case of HE0435-1223. This source of uncertainty is an independent contribution to the error budget that should be added in quadrature with other contributions. This additional source of uncertainty does not significantly alter the current error budget per lens, but it will not be completely negligible as the precision per system improves, e.g. with the addition of spatially resolved kinematics (Shajib et al. 2018). Combining our results for the uncertainties in single lenses, the overall contribution of substructure to the error budget of the TDCOSMO sample, which should be added in quadrature with the other sources of uncertainty, is 0.5%.
- We show that the perturbation to  $H_0$  scales as the square root of the lensing volume divided by the longest time delay. We provide a simple fitting function that uses this scaling to predict the level of perturbation for a strong lens system given the lens and source redshifts, the Einstein radius, and the longest delay.

## 6. Discussion and conclusions

Our results are consistent with previous analysis that show  $H_0$  inferences are internally consistent (Millon et al. 2019; Wong et al. 2020), and are consistent with cosmological distances inferred from Type Ia supernova (Pandey et al. 2019). Since the level of perturbation from line-of-sight halos evolves with redshift, if un-modeled halos biased  $H_0$  inferences significantly it

would lead to inconsistencies between lenses at different redshifts, and between strong-lensing based inferences on  $H_0$  with independent probes of the distance ladder.

While we find no evidence for bias, we conclude that the omission of substructure from strong lens models results in a slight underestimate of the uncertainties in the measured values of  $H_0$ . The additional uncertainty term, which should be summed in quadrature with the other contributions to the error budget, ranges from less than one percent in the case of RXJ1131-1231, to 2.4% in the case of HE0435-1223, increasing the total uncertainties in these systems from 4.3% to 4.4% and 6.5% to 6.9%, respectively. When inferences from the seven lenses analyzed by TDCOSMO are combined, we estimate that substructure contributes an additional 0.5% uncertainty. This is a subdominant contribution to the overall error budget that we reiterate should be summed in quadrature with other contributions.

To make use of the results presented in this paper for future cosmographic inferences, we suggest two possibilities. First, one could obtain a rough estimate of the level of perturbation using the fitting function plotted in Figure 10, and possibly incorporate this additional uncertainty when quoting results on  $H_0$ . Second, the analysis performed in this work could be repeated for new lens systems. The first approach sacrifices precision for speed, while the second approach is more time consuming, but gives a more precise estimate for the additional uncertainty from substructure.

Finally, we note that dark matter halos should also alter the arrival times between images of lensed supernova, such as supernova ‘‘Refsdal’’, which is multiply imaged by an early-type galaxy in the cluster MACS J1149.5+2223 (Kelly et al. 2015). We can extrapolate the results of this work to the Refsdal system to approximate the impact of substructure on future cosmographic inferences. The Refsdal system consists of four images arranged in an Einstein cross configuration separated by  $\sim 1''$  with a maximum delay of 24 days (Rodney et al. 2016), and a fifth image  $\sim 30''$  away with a time delay of  $\sim 1$ yr. Coincidentally, both pairs of Einstein radii and time delays give an uncertainty of 1.3% in  $H_0$  from substructure, per Equation 14, because the long time delay between  $S_X$  and the Einstein cross counteracts the larger image separation.

## Acknowledgments

We thank Anwar Shajib for answering many questions, and for sharing the MCMC chains computed for DESJ0408-5354. We also thank Adriano Agnello, Frederic Courbin, and Dominique Sluse for comments on the manuscript.

DG, TT, and SB acknowledge support by the US National Science Foundation through grant AST-1714953 and AST-1906976. This work used computational and storage services associated with the Hoffman2 Shared Cluster provided by the UCLA Institute for Digital Research and Education’s Research Technology Group.

## References

- Baltz, E. A., Marshall, P., & Oguri, M. 2009, JCAP, 1, 015
- Benson, A. J. 2012, NewA, 17, 175
- Birrer, S. & Amara, A. 2018, Physics of the Dark Universe, 22, 189
- Birrer, S., Amara, A., & Refregier, A. 2015, ApJ, 813, 102
- Birrer, S., Amara, A., & Refregier, A. 2016, JCAP, 2016, 020
- Birrer, S., Amara, A., & Refregier, A. 2017, JCAP, 2017, 037
- Birrer, S., Treu, T., Rusu, C. E., et al. 2019, MNRAS, 484, 4726
- Blandford, R. & Narayan, R. 1986, ApJ, 310, 568
- Bonvin, V., Chan, J. H. H., Millon, M., et al. 2018a, AA, 616, A183

- Bonvin, V., Chan, J. H. H., Millon, M., et al. 2018b, AA, 616, A183
- Bonvin, V., Courbin, F., Suyu, S. H., et al. 2017, MNRAS, 465, 4914
- Bonvin, V., Millon, M., Chan, J. H. H., et al. 2019, AA, 629, A97
- Çağan Şengül, A., Tsang, A., Diaz Rivero, A., et al. 2020, arXiv e-prints, arXiv:2006.07383
- Chen, G. C. F., Chan, J. H. H., Bonvin, V., et al. 2018, MNRAS, 481, 1115
- Chen, G. C. F., Fassnacht, C. D., Suyu, S. H., et al. 2019, MNRAS, 490, 1743
- Chen, G. C. F., Suyu, S. H., Wong, K. C., et al. 2016, MNRAS, 462, 3457
- Congdon, A. B., Keeton, C. R., & Nordgren, C. E. 2010, ApJ, 709, 552
- Courbin, F., Bonvin, V., Buckley-Geer, E., et al. 2018, AA, 609, A71
- Cyr-Racine, F.-Y., Keeton, C. R., & Moustakas, L. A. 2019, PhysRevD, 100, 023013
- Cyr-Racine, F.-Y., Moustakas, L. A., Keeton, C. R., Sigurdson, K., & Gilman, D. A. 2016, PhysRevD, 94, 043505
- Dalal, N. & Kochanek, C. S. 2002, ApJ, 572, 25
- Diaz Rivero, A., Cyr-Racine, F.-Y., & Dvorkin, C. 2018, PhysRevD, 97, 023001
- Diemer, B. 2018, ApJS, 239, 35
- Diemer, B. & Joyce, M. 2019, ApJ, 871, 168
- Fassnacht, C. D., Womble, D. S., Neugebauer, G., et al. 1996, ApJL, 460, L103
- Fassnacht, C. D., Xanthopoulos, E., Koopmans, L. V. E., & Rusin, D. 2002, ApJ, 581, 823
- Gilman, D., Birrer, S., Nierenberg, A., et al. 2020a, MNRAS, 491, 6077
- Gilman, D., Birrer, S., Treu, T., Nierenberg, A., & Benson, A. 2019, MNRAS, 487, 5721
- Gilman, D., Du, X., Benson, A., et al. 2020b, MNRAS, 492, L12
- Hezaveh, Y., Dalal, N., Holder, G., et al. 2016a, JCAP, 2016, 048
- Hezaveh, Y. D., Dalal, N., Marrone, D. P., et al. 2016b, ApJ, 823, 37
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, The Astrophysical Journal Supplement Series, 208, 19
- Hsueh, J. W., Enzi, W., Vegetti, S., et al. 2019, MNRAS, 2780
- Keeton, C. R. & Moustakas, L. A. 2009, ApJ, 699, 1720
- Kelly, P. L., Rodney, S. A., Treu, T., et al. 2015, Science, 347, 1123
- Kochanek, C. S. 2019, arXiv e-prints, arXiv:1911.05083
- Koopmans, L. V. E. 2005, MNRAS, 363, 1136
- Koopmans, L. V. E., Treu, T., Fassnacht, C. D., Blandford, R. D., & Surpi, G. 2003, ApJ, 599, 70
- Lagattuta, D. J., Fassnacht, C. D., Auger, M. W., et al. 2010, ApJ, 716, 1579
- Lin, H., Buckley-Geer, E., Agnello, A., et al. 2017, ApJL, 838, L15
- Millon, M., Courbin, F., Bonvin, V., et al. 2020, arXiv e-prints, arXiv:2006.10066
- Millon, M., Galan, A., Courbin, F., et al. 2019, arXiv e-prints, arXiv:1912.08027
- Morgan, N. D., Caldwell, J. A. R., Schechter, P. L., et al. 2004, AJ, 127, 2617
- Morgan, N. D., Kochanek, C. S., Pevunova, O., & Schechter, P. L. 2005, AJ, 129, 2531
- Myers, S. T., Fassnacht, C. D., Djorgovski, S. G., et al. 1995, ApJL, 447, L5
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
- Nierenberg, A. M., Gilman, D., Treu, T., et al. 2020, MNRAS, 492, 5314
- Nierenberg, A. M., Treu, T., Brammer, G., et al. 2017, MNRAS, 471, 2224
- Nierenberg, A. M., Treu, T., Wright, S. A., Fassnacht, C. D., & Auger, M. W. 2014, MNRAS, 442, 2434
- Oguri, M. 2007, ApJ, 660, 1
- Pandey, S., Raveri, M., & Jain, B. 2019, arXiv e-prints, arXiv:1912.04325
- Planck Collaboration, Akrami, Y., Arroja, F., et al. 2018, arXiv e-prints, arXiv:1807.06205
- Refregier, A. 2003, MNRAS, 338, 35
- Riess, A. G. 2019, Nature Reviews Physics, 2, 10
- Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, ApJ, 876, 85
- Rodney, S. A., Strolger, L. G., Kelly, P. L., et al. 2016, ApJ, 820, 50
- Rusu, C. E., Wong, K. C., Bonvin, V., et al. 2019, arXiv e-prints, arXiv:1905.09338
- Schneider, P. & Sluse, D. 2013, AA, 559, A37
- Shajib, A. J., Birrer, S., Treu, T., et al. 2020, MNRAS, 494, 6072
- Shajib, A. J., Birrer, S., Treu, T., et al. 2019, MNRAS, 483, 5649
- Shajib, A. J., Treu, T., & Agnello, A. 2018, MNRAS, 473, 210
- Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1
- Sluse, D., Claeskens, J. F., Hutsemekers, D., & Surdej, J. 2007, AA, 468, 885
- Sluse, D., Rusu, C. E., Fassnacht, C. D., et al. 2019, MNRAS, 490, 613
- Sluse, D., Surdej, J., Claeskens, J. F., et al. 2003, AA, 406, L43
- Springel, V., Wang, J., Vogelsberger, M., et al. 2008, MNRAS, 391, 1685
- Suyu, S. H., Auger, M. W., Hilbert, S., et al. 2013, ApJ, 766, 70
- Suyu, S. H., Bonvin, V., Courbin, F., et al. 2017, MNRAS, 468, 2590
- Suyu, S. H., Marshall, P. J., Blandford, R. D., et al. 2009, ApJ, 691, 277
- Suyu, S. H., Treu, T., Hilbert, S., et al. 2014, ApJL, 788, L35
- Tewes, M., Courbin, F., Meylan, G., et al. 2013, AA, 556, A22
- Tie, S. S. & Kochanek, C. S. 2018, MNRAS, 473, 80
- Tonry, J. L. 1998, AJ, 115, 1
- Vegetti, S., Despali, G., Lovell, M. R., & Enzi, W. 2018, MNRAS, 481, 3661
- Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., & Gavazzi, R. 2010, MNRAS, 408, 1969
- Weymann, R. J., Latham, D., Angel, J. R. P., et al. 1980, Nat, 285, 641
- Wisotzki, L., Schechter, P. L., Bradt, H. V., Heinmüller, J., & Reimers, D. 2002, AA, 395, 17
- Wong, K. C., Suyu, S. H., Auger, M. W., et al. 2017, MNRAS, 465, 4895
- Wong, K. C., Suyu, S. H., Chen, G. C. F., et al. 2020, MNRAS[arXiv:1907.04869]

## Appendix A: Baseline macro lensmodels

Table A.1 lists the parameters describing the lens macromodels for the mock lens analogs of the six strong lens systems considered in this work.

## Appendix B: Hiding dark matter halos in extended images with shapelets

As a byproduct of this analysis, we obtain hundreds of mock lenses with realistic line-of-sight and substructure populations that we model with state of the art image reconstruction techniques. These simulations are relevant to the field of gravitational imaging of dark substructure in lensed arcs (Vegetti et al. 2010; Hezaveh et al. 2016b; Vegetti et al. 2018). Dark substructure occasionally produces imaging residuals in lensed arcs that betray the presence of individual halos, or provide a means to infer quantities such as the convergence power spectrum of halos in the lens system (Hezaveh et al. 2016a; Diaz Rivero et al. 2018; Cyr-Racine et al. 2019; Çağan Şengül et al. 2020).

As the convergence itself is not observable, it must be derived from the surface brightness residuals from the lens modeling (Cyr-Racine et al. 2019). Keeping all else fixed, adding additional complexity to the background source light partially removes gravitational imaging residuals in extended images. The gravitational imaging residuals therefore depend explicitly on how the unknown background source is modeled.

To illustrate, the two rows in Figure show examples of imaging residuals computed for the mock analogs of B1608+656 and PG1115+080. The far left panels in Figure B.1 show the imaging residuals without shapelets added to the background source, and the second from left panels show the same mock lens system, with the same population of dark matter halos, modeled with shapelets included in the background source. The population of line-of-sight halos and subhalos that perturbs the mock images are shown in the third panels, using the definition of the multi-plane convergence described in Section 4.1. The far right panel shows the difference between the source surface brightness between the models that include and exclude shapelets. In these two cases, adding shapelets to the background source removes imaging residuals from the extended images.

In Figure B.2, we show the stacked power spectra of imaging residuals from 100 different realizations. The shaded regions represent show 68% bootstrap confidence intervals around the median in each distance bin plotted on the x-axis. Without shapelets included in the background source, substructure leads to enhanced power on small scales, as expected. The addition of shapelets, which in this example act as a proxy for any additional source complexity that is typically required to model lensed images, removes the signal.

While the effects of very massive halos that produce direct hits on extended images evidently cannot be completely absorbed by the background source (e.g. Vegetti et al. 2010; Hezaveh et al. 2016b), the collective effects of the many smaller halos that affect the convergence power spectrum can be partially absorbed by a reconfiguration of the lens and source model. The degree to which this occurs depends on the specifics of the source model, and on the quality of the imaging data. In principle, the fact that the source is multiply imaged breaks degeneracies between the source structure and imaging residuals (Cyr-Racine et al. 2019), but with finite resolution and signal to noise the effect persists. Higher-resolution images, or an identification of what scales are dominated by substructure pertur-

bations rather than source morphology, may help deal with this systematic challenge.

## Appendix C: Joint posterior distributions of $H_0$ with lens macromodel parameters

Figures C.1 through C.6 show the joint posterior distributions of macromodel parameters with the Hubble constant for the six lens systems analyzed in this work.

## Appendix D: Joint posterior distributions of time delays and flux ratios

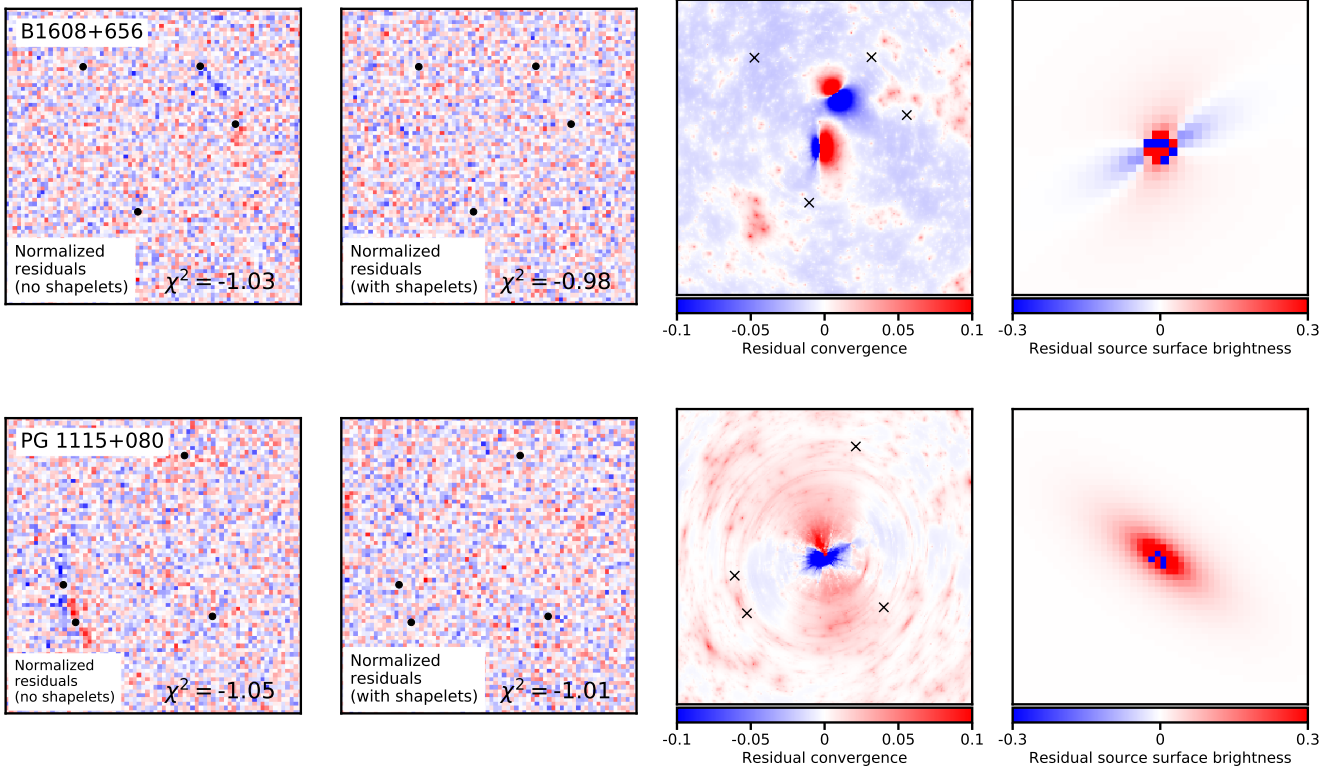
The initial motivation for considering the effects of substructure on time delays was to probe the nature of dark matter through these data. The thinking was that time delay ‘anomalies’, or measured time delays that differ from the model-predicted time delays, encode the abundance and spatial distribution of dark substructure (e.g. Keeton & Moustakas 2009; Cyr-Racine et al. 2016).

The Shapiro delay associated with dark (sub)halos depends on the projected gravitational potential of halos, and is a long-range interaction. This could, in principle, complement the information encoded by image flux ratios, which are highly localized probes of dark matter structure and provide an avenue to probe small-scale structure in strong lenses (Dalal & Kochanek 2002; Nierenberg et al. 2014; Hsueh et al. 2019; Gilman et al. 2020a,b). Figures D.1 through D.6 show the joint distributions of the residual time delays and residual flux ratios for the six lens systems analyzed. The flux ratios are computed with a background source parameterized as a Gaussian with a FWHM of 25 pc.

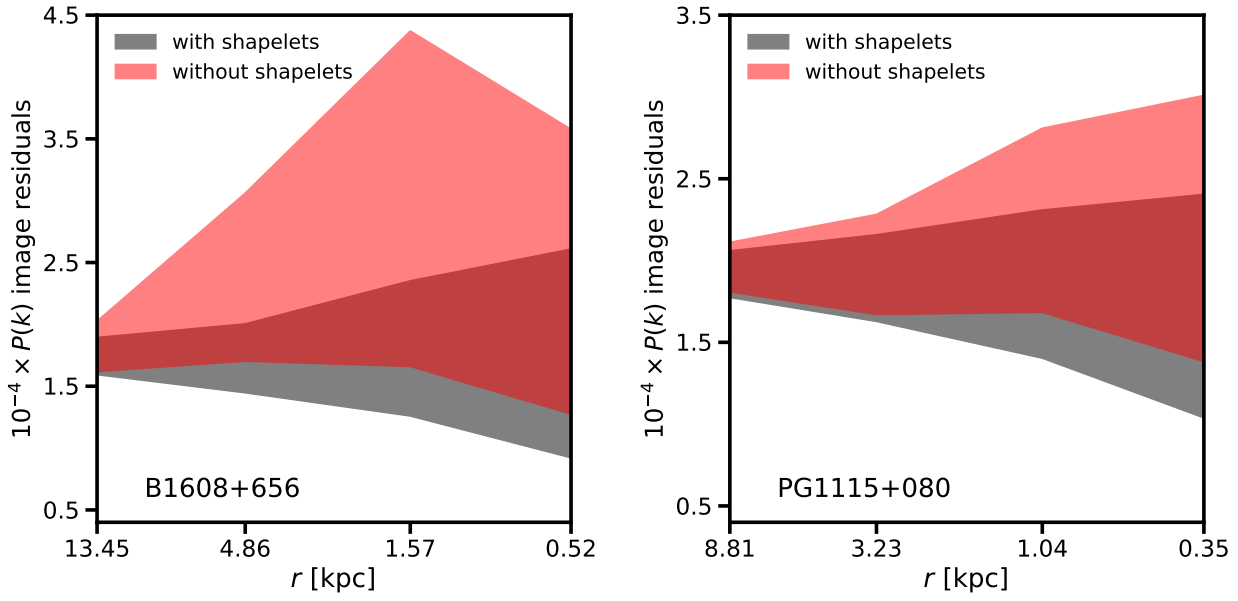


**Table A.1.** Lens model parameters used to create the baseline datasets. The main deflector is modeled as an elliptical power law plus external shear, and other satellites or field galaxies are modeled with spherical power law profiles. Parameters from left to right: Einstein radius of the main deflector, x and y mass centroid of main deflector, main deflector ellipticity and position angle, logarithmic slope of the main mass profile, external shear magnitude and position angle. The columns labeled  $R_i$ ,  $x_i$ ,  $y_i$ ,  $z_i$  contain the Einstein radii, coordinates, and redshifts of satellite galaxies, or galaxies along the line of sight, that are included in the mock lens models. Angular positions marked with an asterisk indicate that this coordinate is corrected for foreground lensing effects by the main deflector.

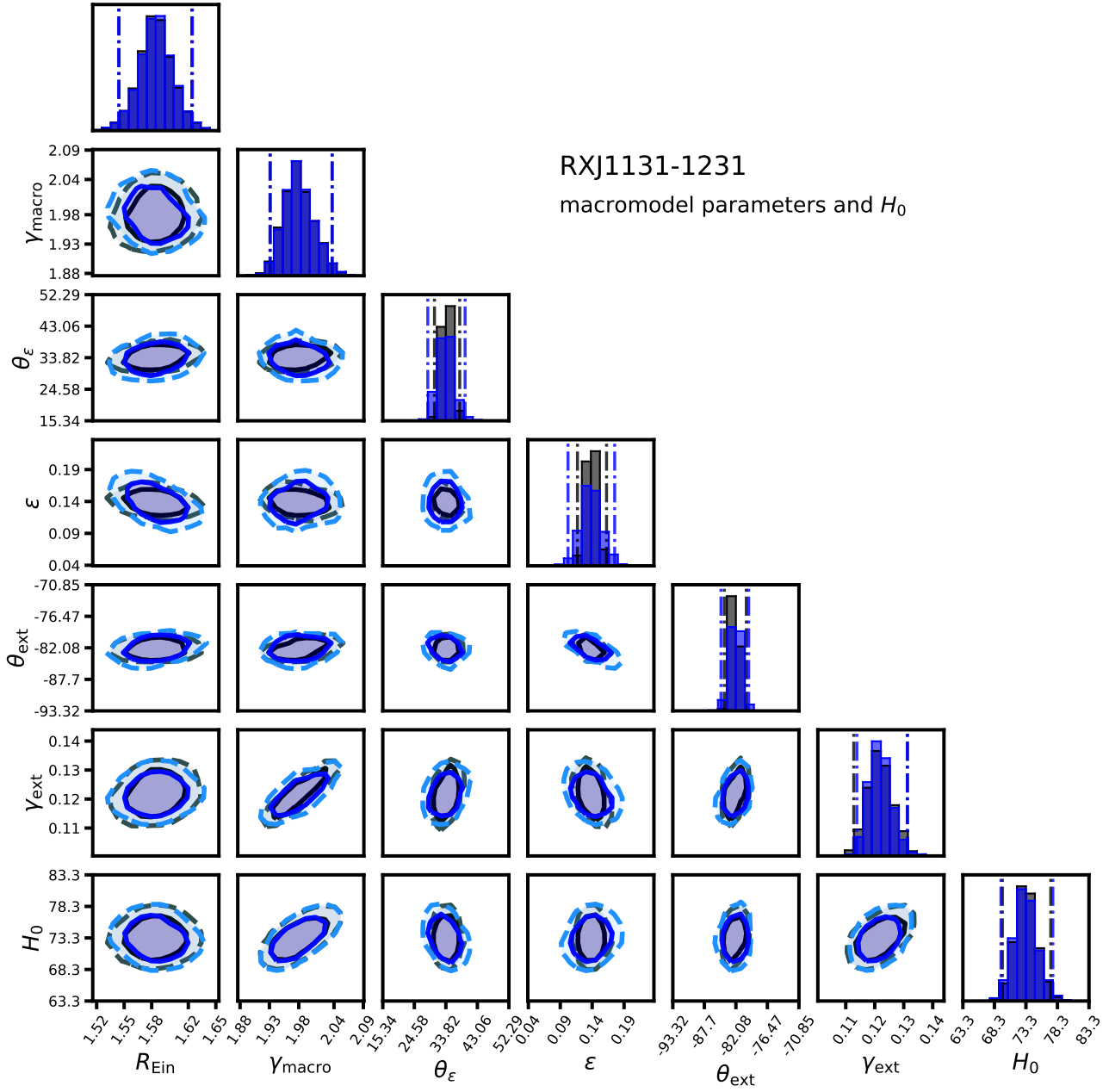
Lens system	$R_{\text{Ein}}$	$g_x$	$g_y$	$\epsilon$	$\theta_\epsilon$	$\gamma$	$\gamma_{\text{ext}}$	$\theta_{\text{ext}}$	$R_1$	$x_1$	$y_1$	$z_1$	$R_2$	$x_2$	$y_2$	$z_2$
RXJ1131-1231	1.58	-0.01	-0.03	0.14	34	1.98	0.12	-82	0.28	-0.1	0.61	$z_d$	-	-	-	-
PG1115-080	1.05	0.0	0.0	0.01	-71	2.20	0.03	-56	2.0	-9.21	-3.91	$z_d$	-	-	-	-
HE0435-1223	1.17	-0.02	0.02	0.07	-81	1.98	0.05	10	0.35	-2.27*	1.98*	$z_d + 0.33$	-	-	-	-
B1608+656	0.90	0.01	0.06	0.51	-4	2.08	0.10	26	0.26	-0.71	0.13	$z_d$	-	-	-	-
WFI2033-4723	1.0	-0.02	0.02	0.22	57	1.95	0.17	-8	0.03	0.24	2.04	$z_d$	0.93	-3.63*	-0.68*	$z_d + 0.09$
DESJ0408-5354	1.72	0.14	0.00	0.39	23	1.92	0.11	5	0.22	-1.58	-0.95	$z_d$	0.77	1.13*	-7.45*	$z_d + 0.17$



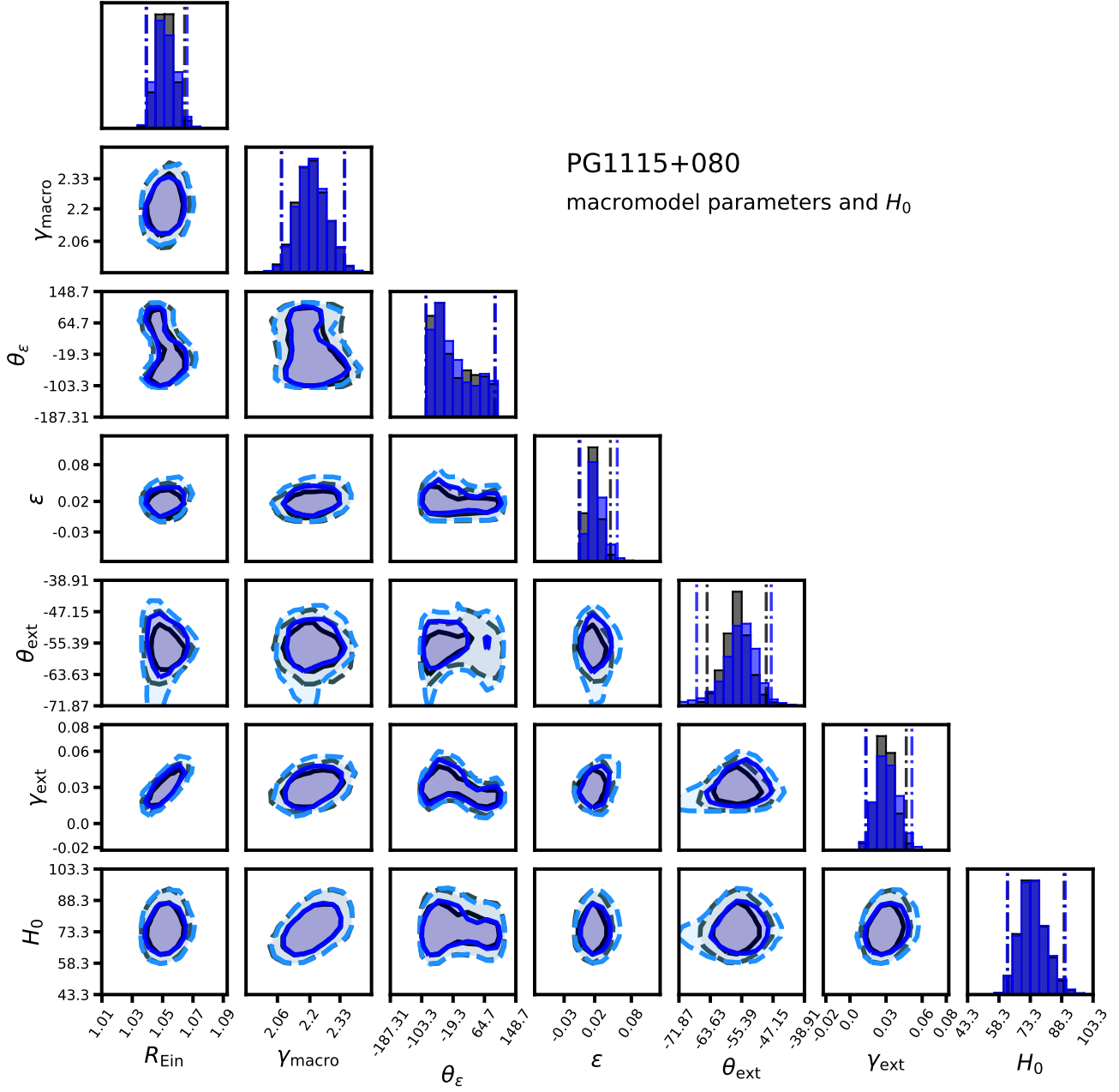
**Fig. B.1.** The panels show examples of imaging residuals, convergence, and residual source surface brightness for one mock lens system based on B1608+656 (top row) and PG1115+080 (bottom row). The far left panel shows the gravitational imaging residuals when only an elliptical Sérsic profile, the same profile used to create the mock data, is used to model the background source. The second from left panel shows the imaging residuals when shapelets are added to the source light profile. The dark matter halos that perturb the lensed arc in the mock data are shown in the third from left panel, and the residual source surface brightness is shown in the far right panel.



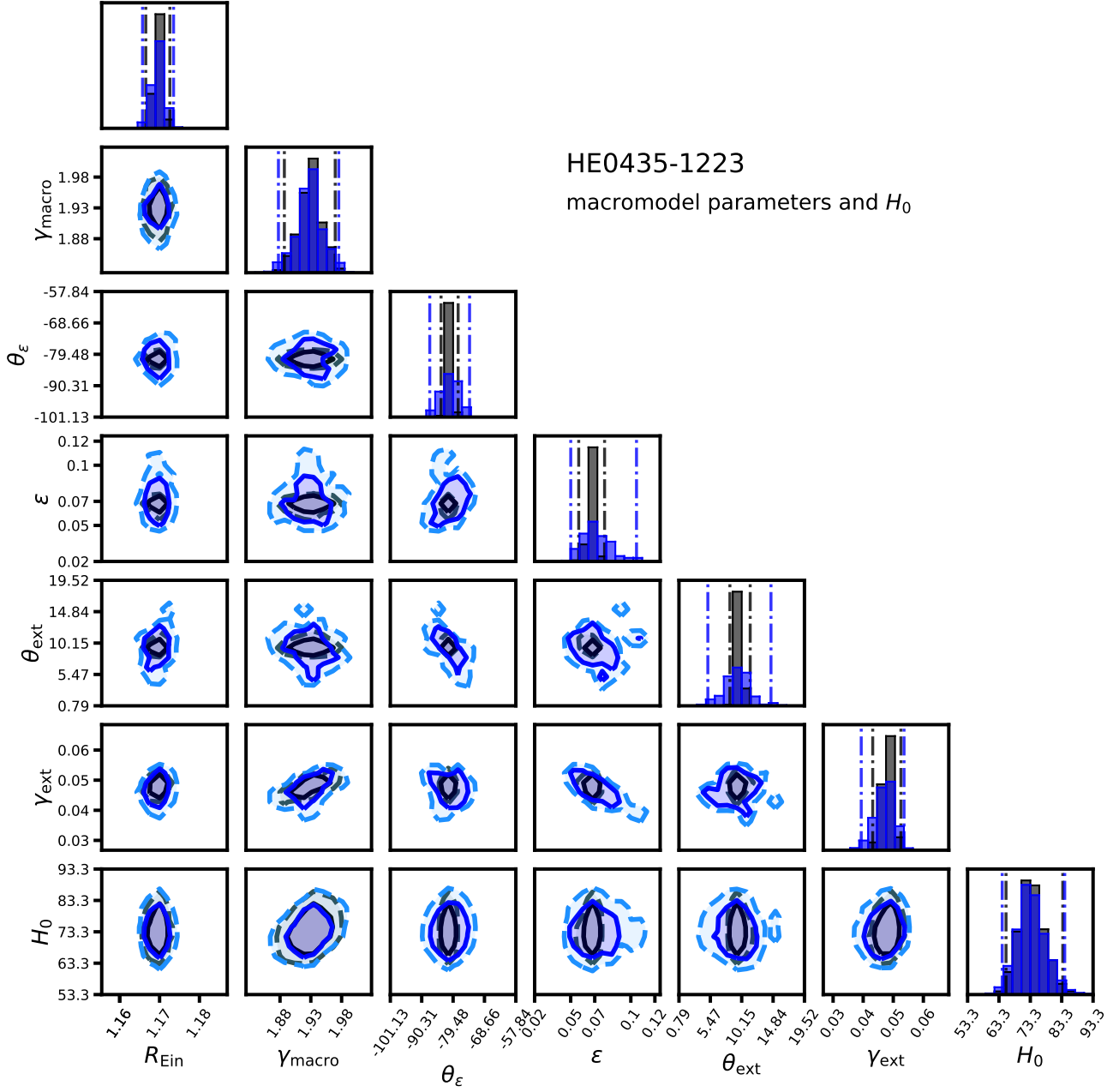
**Fig. B.2.** The stacked power spectrum of gravitational imaging residuals computed from 100 residual maps like those in Figure B.1. The manner in which substructure affects the power spectrum of imaging residuals on small scales depends on how the background source is modeled.



**Fig. C.1.** Joint distribution of selected macromodel parameters and the inferred Hubble constant for the control data set (black) and the substructure-perturbed dataset (blue). The parameters are (reading the x-axis labels from the left) the Einstein radius, the logarithmic slope of the main deflector, the ellipticity position angle, the ellipticity, the external shear position angle, the external shear strength, and the Hubble constant. Vertical lines show 95% confidence intervals and the contours show 68% and 95% confidence intervals. The distributions show are computed for the lens system RXJ1131-1231.

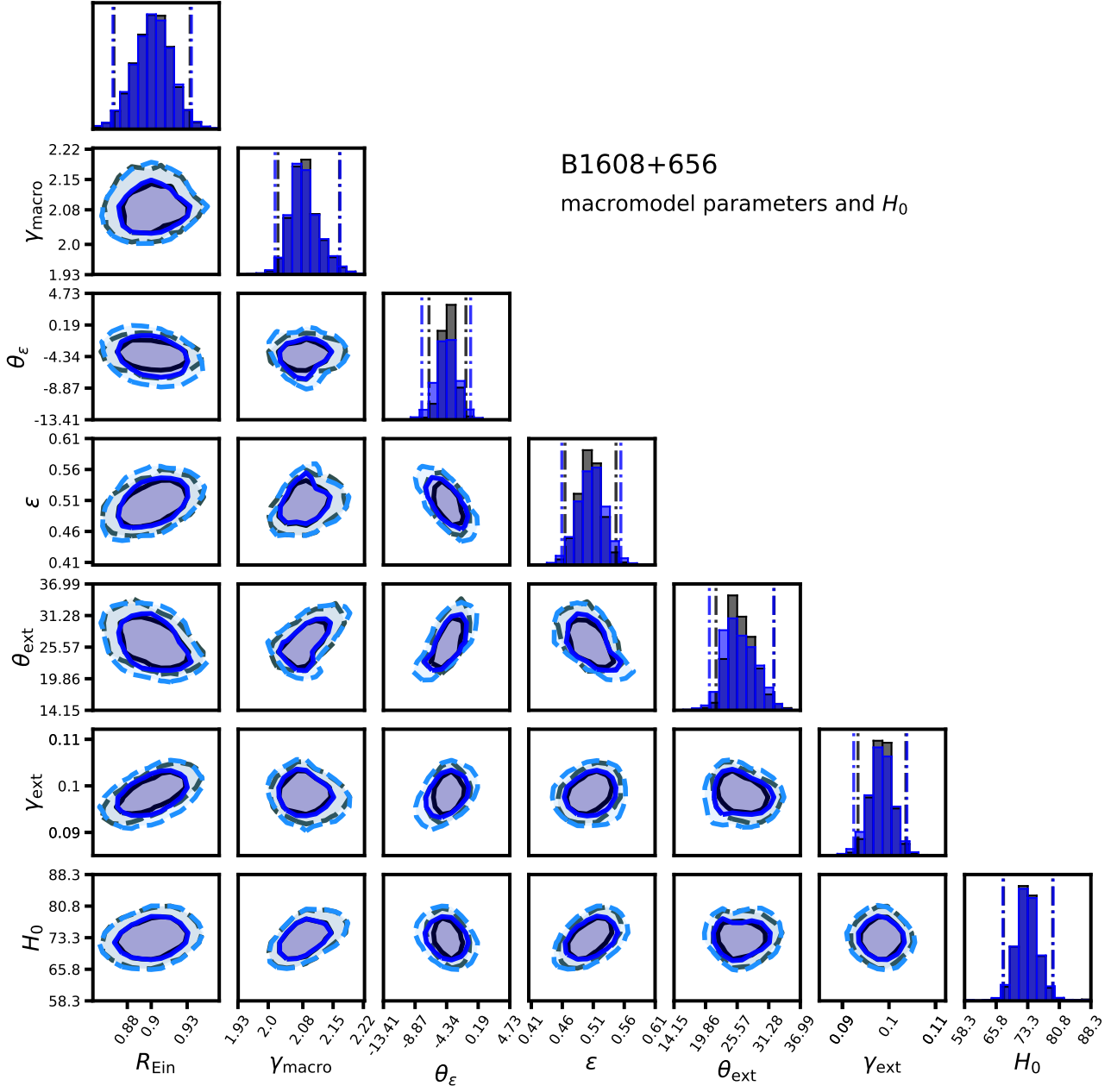


**Fig. C.2.** Joint distribution of selected macromodel parameters and the inferred Hubble constant for the lens system PG1115+080. See Figure C.1 for a description of the parameters.

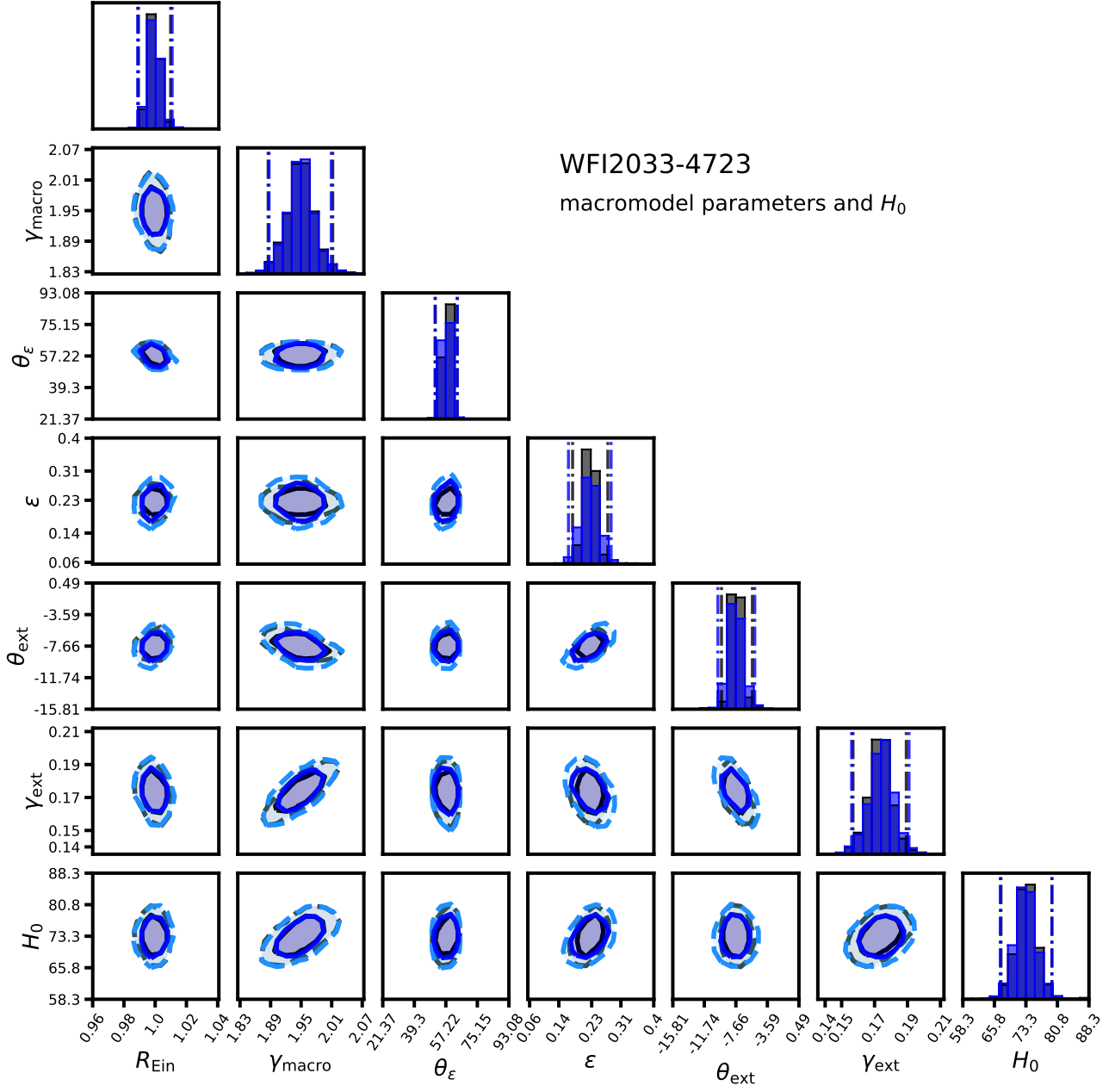


**Fig. C.3.** Joint distribution of selected macromodel parameters and the inferred Hubble constant for the lens system HE0435-1223. See Figure C.1 for a description of the parameters.

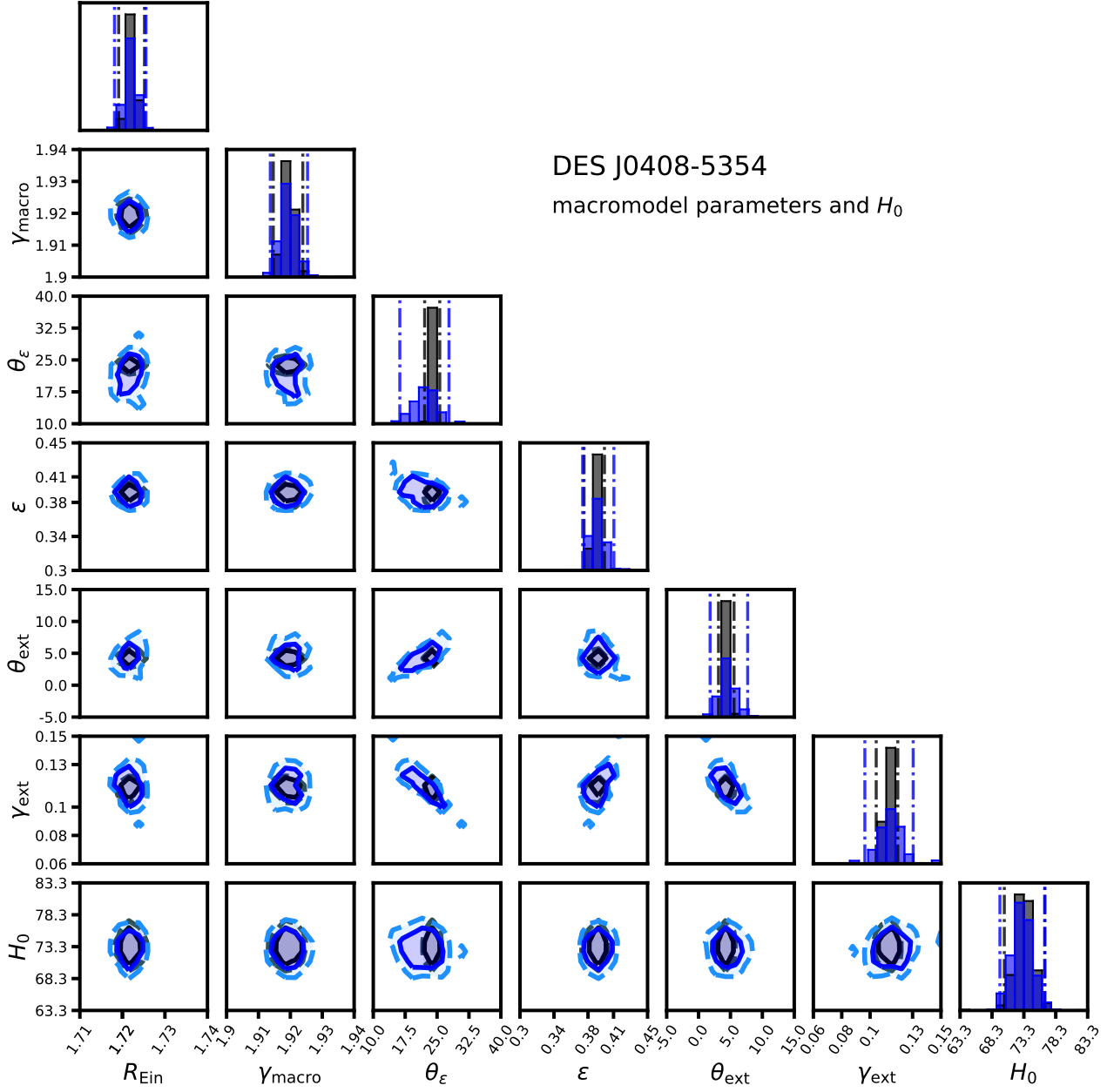




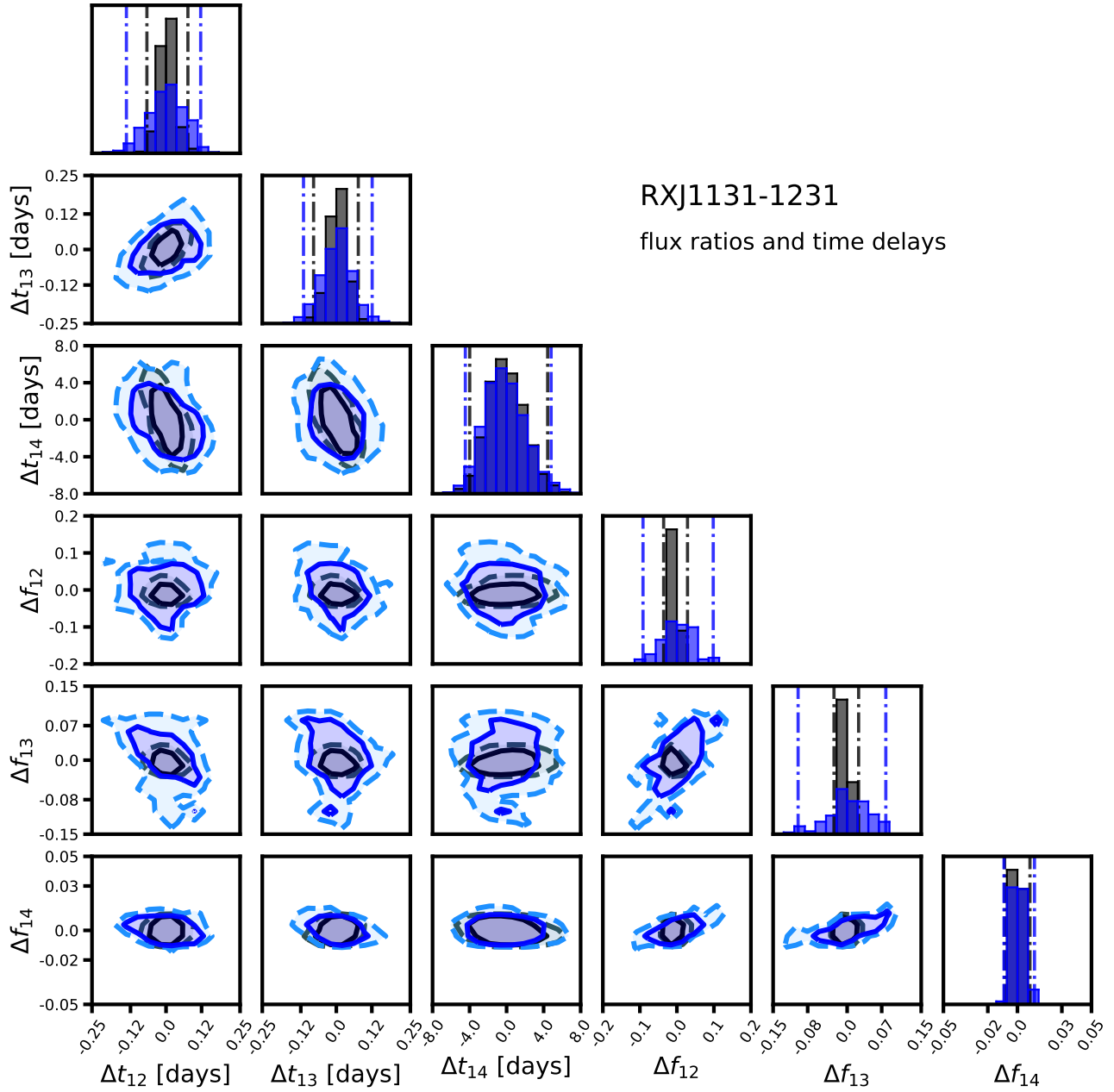
**Fig. C.4.** Joint distribution of selected macromodel parameters and the inferred Hubble constant for the lens system B1608+656. See Figure C.1 for a description of the parameters.



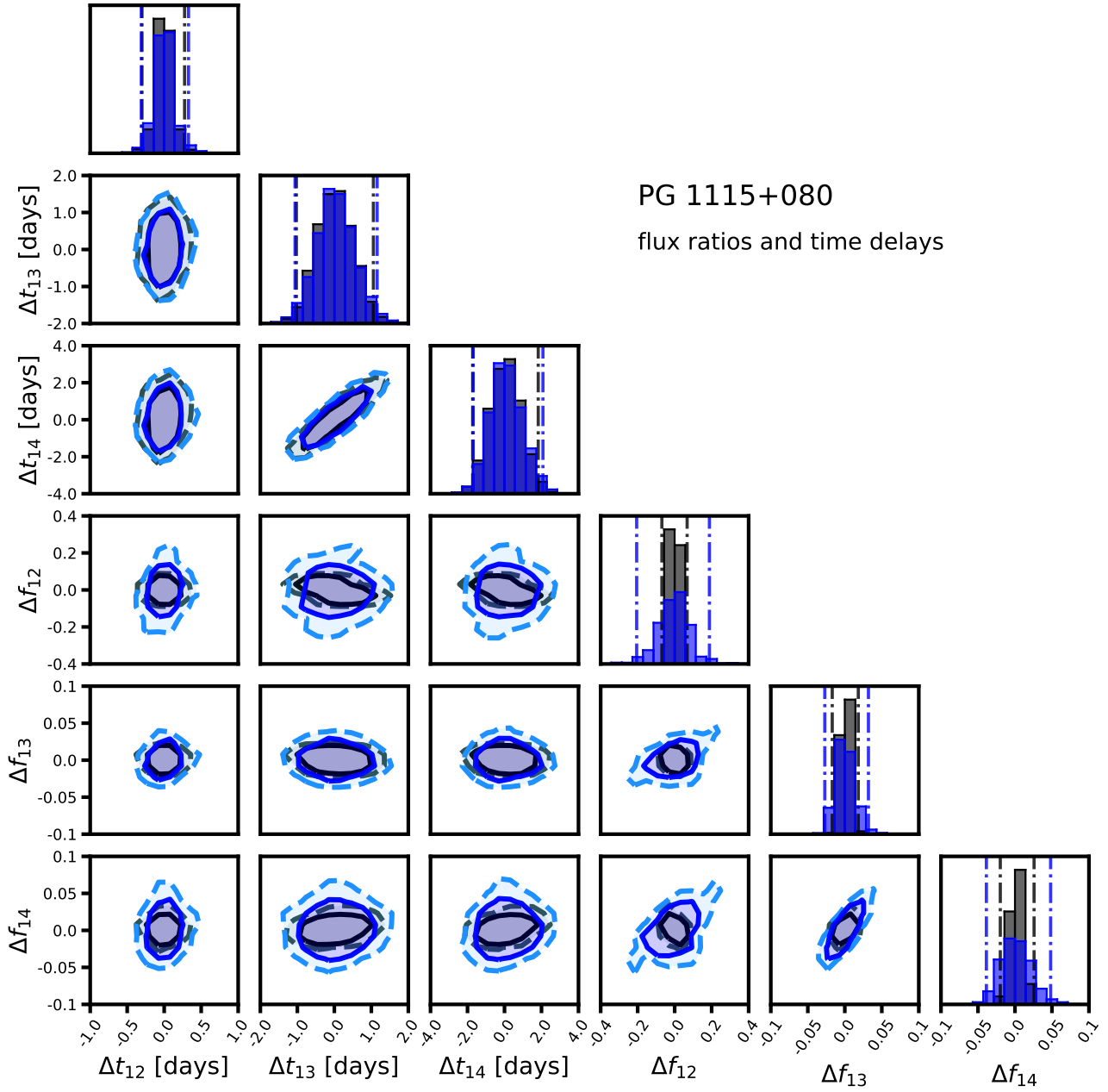
**Fig. C.5.** Joint distribution of selected macromodel parameters and the inferred Hubble constant for the lens system WFI2033-4723. See Figure C.1 for a description of the parameters.



**Fig. C.6.** Joint distribution of selected macromodel parameters and the inferred Hubble constant for the lens system DESJ0408-5354. See Figure C.1 for a description of the parameters.

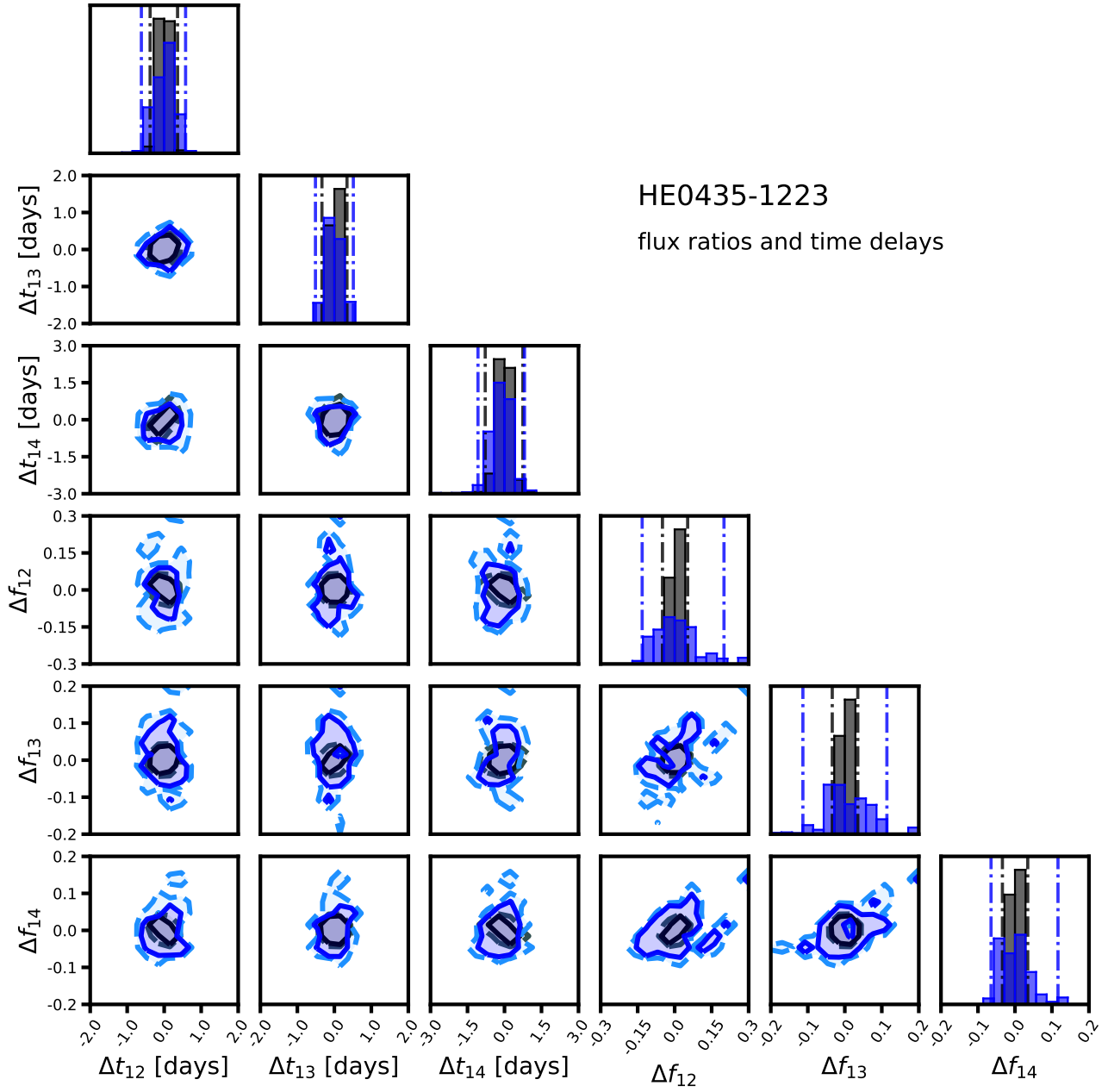


**Fig. D.1.** Joint distribution of flux ratio and time delay residuals (model minus data) computed for the lens system RXJ1131-1231.

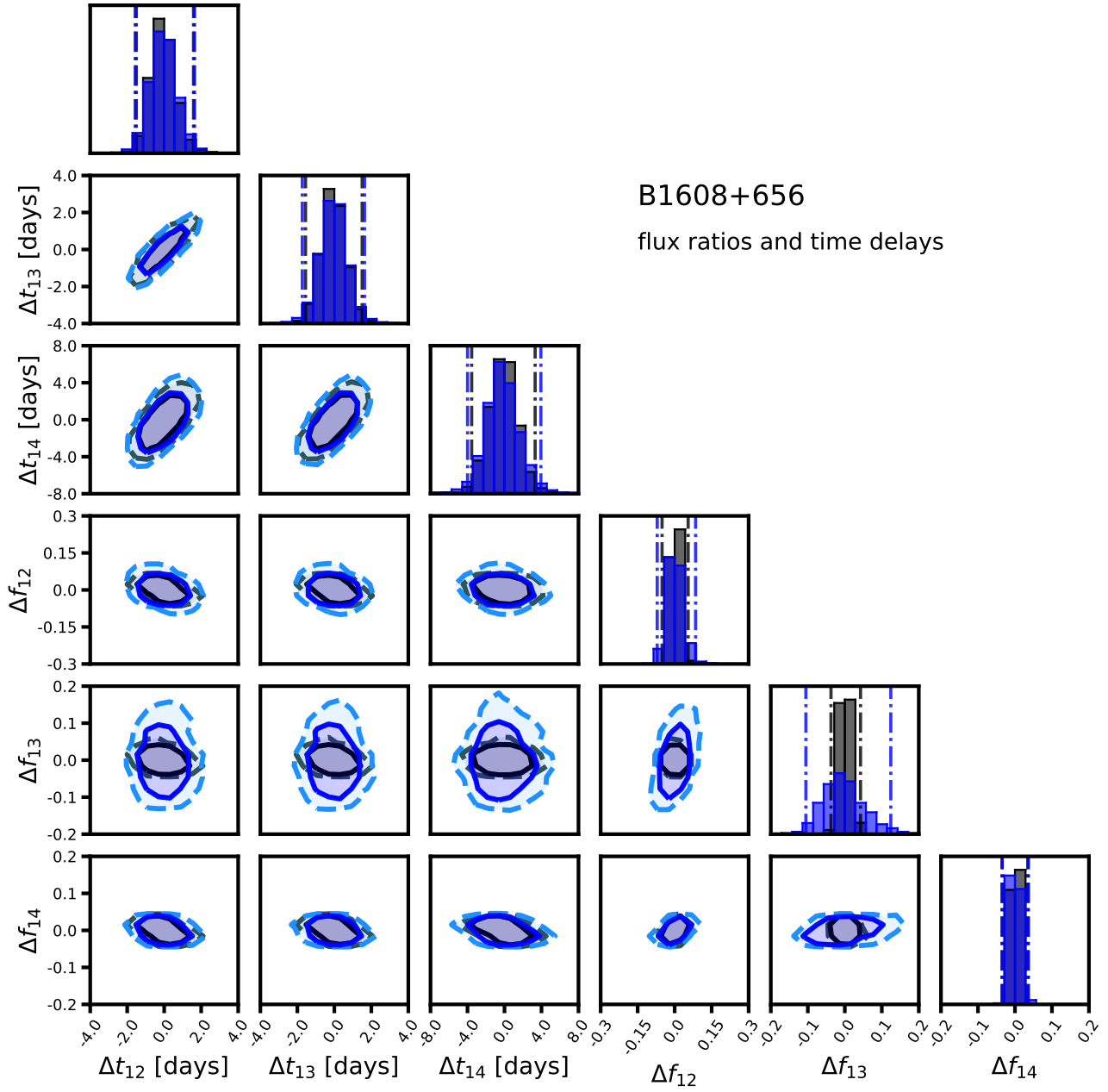


**Fig. D.2.** Joint distribution of flux ratio and time delay residuals (model minus data) computed for the lens system PG1115+080.

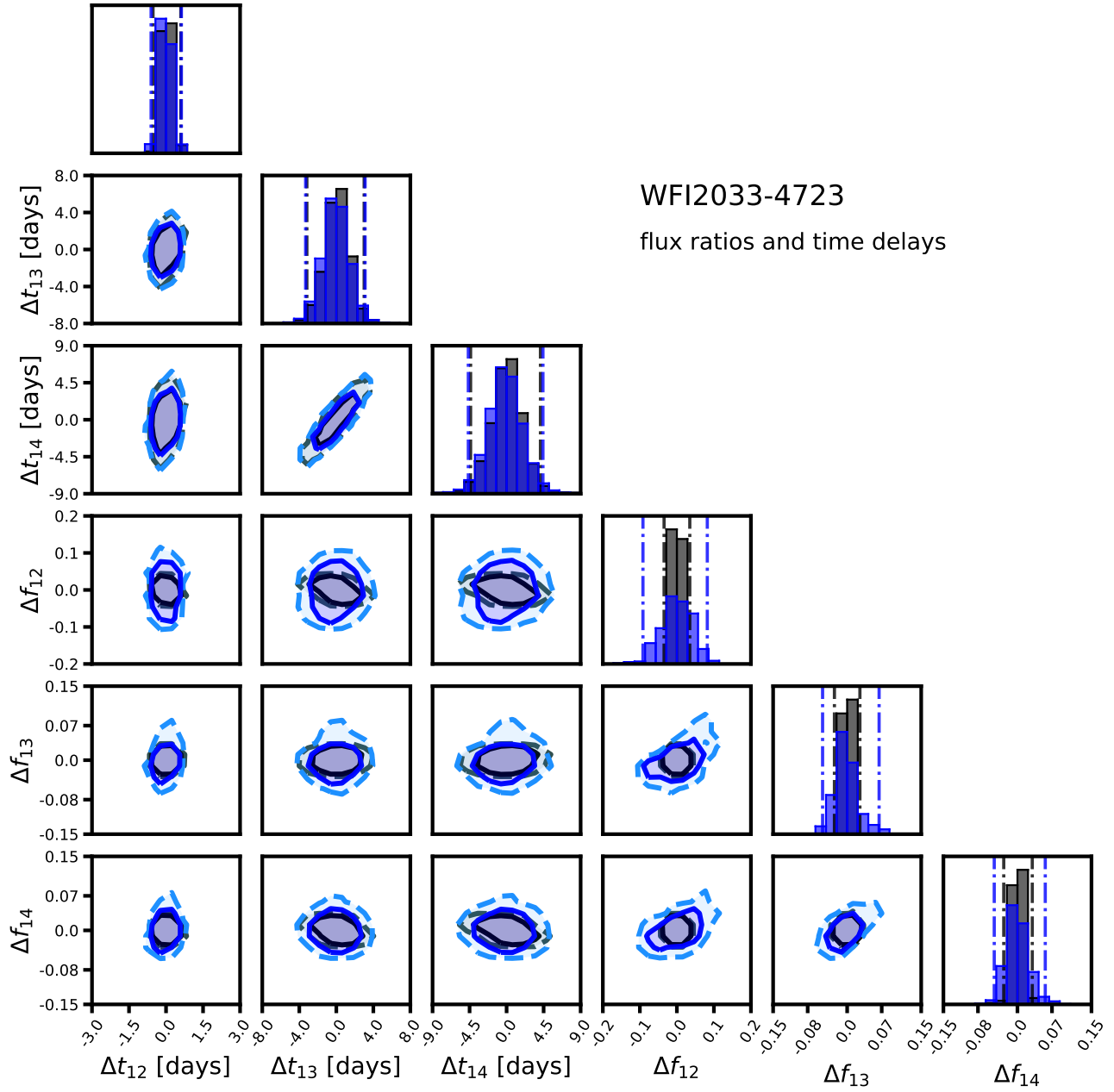




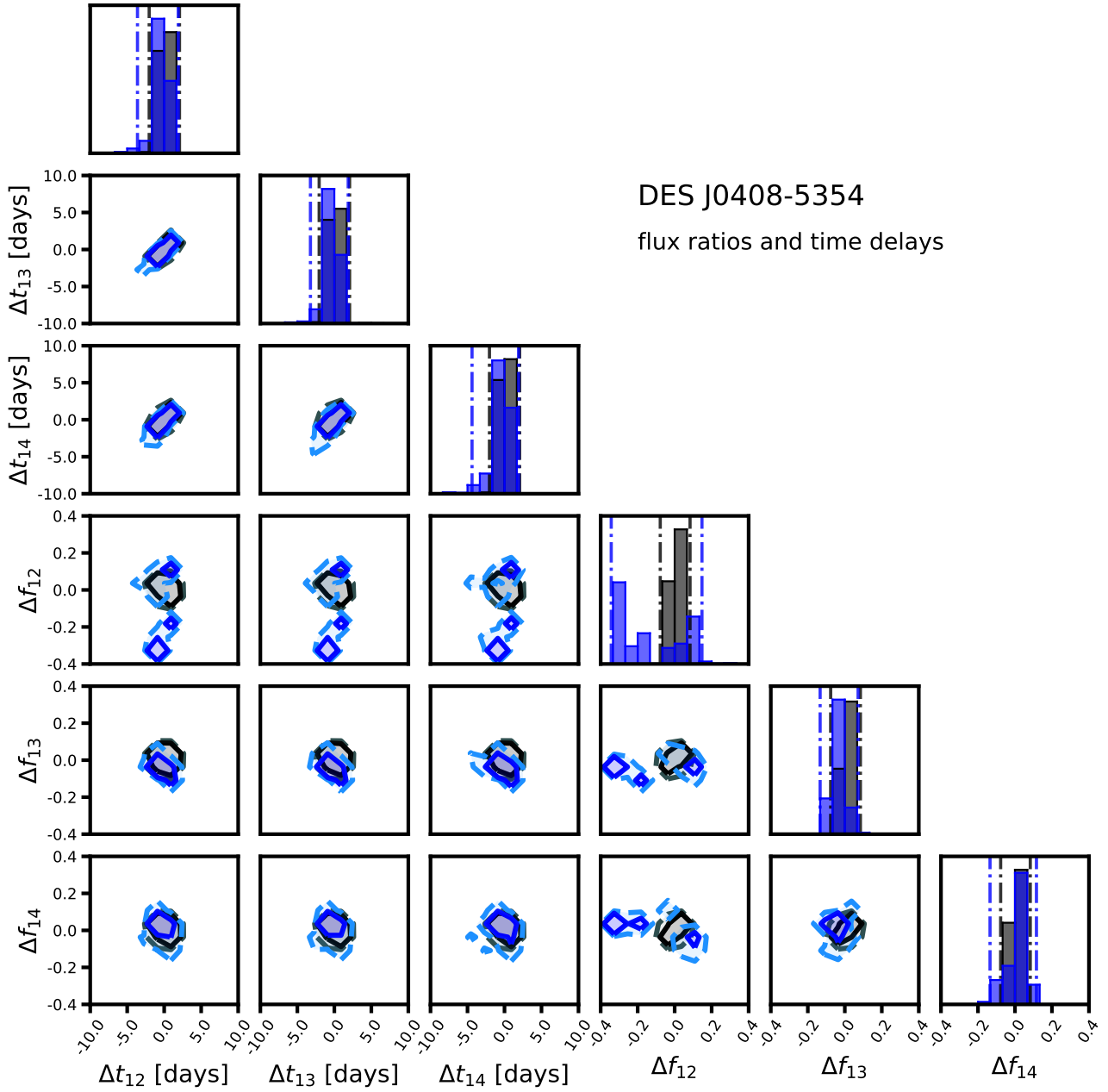
**Fig. D.3.** Joint distribution of flux ratio and time delay residuals (model minus data) computed for the lens system HE0435-1223.



**Fig. D.4.** Joint distribution of flux ratio and time delay residuals (model minus data) computed for the lens system B1608-656.



**Fig. D.5.** Joint distribution of flux ratio and time delay residuals (model minus data) computed for the lens system WFI2033-4723.



**Fig. D.6.** Joint distribution of flux ratio and time delay residuals (model minus data) computed for the lens system DESJ0408-5354.