

QUANTIFYING ENVIRONMENTAL AND LINE-OF-SIGHT EFFECTS IN MODELS OF STRONG GRAVITATIONAL LENS SYSTEMS

CURTIS MCCULLY^{1,2,*} CHARLES R. KEETON³ KENNETH C. WONG^{4,5,†} ANN I. ZABLUDOFF⁶

¹Las Cumbres Observatory Global Telescope Network, 6740 Cortona Dr Suite 102, Goleta, CA 93117-5575, USA

²Department of Physics, University of California, Santa Barbara, CA 93106-9530, USA

³Department of Physics and Astronomy, Rutgers, the State University of New Jersey, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA.

⁴Institute of Astronomy and Astrophysics, Academia Sinica (ASIAA), P.O. Box 23-141, Taipei 10617, Taiwan

⁵National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

⁶Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721

Draft version September 28, 2016

ABSTRACT

Matter near a gravitational lens galaxy or projected along the line of sight (LOS) can affect strong lensing observables by more than contemporary measurement errors. We simulate lens fields with realistic three-dimensional mass configurations (self-consistently including voids), and then use lens models to quantify biases and uncertainties associated with different ways of treating the lens environment (ENV) and LOS. We identify the combination of mass, projected offset, and redshift that determines the importance of a perturbing galaxy for lensing. Foreground structures have a stronger effect on the lens potential than background structures, due to non-linear effects in the foreground and downweighting in the background. There is dramatic variation in the net strength of ENV/LOS effects across different lens fields; modeling fields individually yields stronger priors for H_0 than ray tracing through N-body simulations. Lens systems in groups tend to have stronger ENV/LOS contributions than non-group lenses. In models, ignoring mass outside the lens yields poor fits and biased results. Adding external shear can account for tidal stretching from galaxies at redshifts $z \geq z_{\text{lens}}$, but it requires corrections for external convergence and cannot reproduce non-linear effects from foreground galaxies. Using the tidal approximation is reasonable for most perturbers as long as non-linear redshift effects are included. Yet even then, the scatter in H_0 is limited by the lens profile degeneracy. Asymmetric image configurations produced by highly elliptical lens galaxies are less sensitive to the lens profile degeneracy, so they offer appealing targets for precision lensing analyses in future surveys like LSST and Euclid.

1. INTRODUCTION

Strong gravitational lensing is an important probe for many facets of cosmology. Analysis of strong lenses has led to constraints on the masses and properties of dark matter halos of galaxies (e.g., Keeton et al. 1998; Treu et al. 2006; Koopmans et al. 2006; Barnabè et al. 2009; Auger et al. 2010; Treu et al. 2010; Lagattuta et al. 2012; Wong et al. 2014), substructure in galaxy halos (e.g., Mao & Schneider 1998; Metcalf & Madau 2001; Dalal & Kochanek 2002; Vegetti & Vogelsberger 2014; Hezaveh et al. 2014), and the Hubble constant, independent of the cosmic distance ladder (e.g., Refsdal 1964; Keeton & Kochanek 1997; Kochanek 2003; Saha et al. 2006; Oguri 2007; Suyu et al. 2010, 2013; Birrer et al. 2015; Chen et al. 2016; Bonvin et al. 2016; Suyu et al. 2016; Wong et al. 2016). Strong lensing may also be employed to constrain the properties of dark energy (e.g., Turner 1990; Linder 2004, 2011; Cao et al. 2012; Treu et al. 2013).

In recent years, both the quantity and quality of strong lens data have improved. There are now roughly a hundred known lensed QSOs (e.g., CASTLeS¹; Oguri et al. 2006; Myers et al. 2003) and a comparable number of strongly lensed galaxies (e.g., Bolton et al. 2008; Belokurov et al. 2009). These samples will increase dra-

matically in the near future with LSST and Euclid (e.g., LSST Science Collaboration et al. 2009; Coe & Moustakas 2009; Refregier et al. 2010; Oguri & Marshall 2010; Collett 2015). Here, we focus on lensed QSOs because the compact source can vary rapidly enough to enable measurements of lens time delays (though much of this discussion also applies to strongly lensed supernovae; Kelly et al. 2015, 2016; Rodney et al. 2016). The relative positions and fluxes of lensed QSO images are routinely measured to high precision using the *Hubble Space Telescope* (e.g., Lehar et al. 2000; Sluse et al. 2012, and references therein; CASTLeS Collaboration). Our understanding of gravitational lenses and the constraints they place on cosmology is no longer limited by observations, but by systematic uncertainties.

One of the key issues is that lens galaxies are not isolated systems. Matter in the environment of the lens galaxy and projected along the line of sight (hereafter ENV/LOS) can produce perturbations in the lensing potential that cannot be ignored as we enter the era of “Precision Lensing” (which we define by the goal of constraining the Hubble constant to $< 1\%$). Mass that is not physically associated with lens galaxy but lies close in projection can also produce perturbations to the lens potential (e.g., Bar-Kana 1996; Momcheva et al. 2006; Wong et al. 2011; Collett & Cunningham 2016). Jaroszynski & Kostrzewa-Rutkowska (2014) show that omitting ENV/LOS effects can often lead to unsuccessful fits, es-

* cmccully@lco.net

† EACOA Fellow

¹ <http://www.cfa.harvard.edu/castles/>

pecially when time delay constraints are included. Collett & Cunningham (2016) argue that lensed QSOs are more likely to have an overdense LOS than the global population which can lead to a bias in the inferred Hubble constant from a sample of lensed QSOs.

To lowest order, using the “tidal approximation,” perturbing galaxies contribute “external convergence,” which produces extra (de-)magnification, and “external shear,” which stretches the images of the source, transforming circles to ellipses. External convergence due to the ENV/LOS is one of the dominant components of the uncertainty budget for measurements of the Hubble constant (Suyu 2012).

Our goal for this work is to quantify the effects of mass outside the main lens galaxy on the measured cosmology,² addressing the following questions: Which individual perturbing galaxies are the most important? What is the range of ENV/LOS contributions? What drives the bias and scatter in reproducing lensing observables from lens systems in realistic ENV/LOS mass distributions? Which lens systems produce the strongest constraints on the Hubble constant?

To answer these questions, we build mass models for the ENV/LOS using extensive photometric and spectroscopic data for ~ 25 strong lensing systems (Momcheva et al. 2006; Williams et al. 2006; Momcheva et al. 2015). We embed a lens galaxy in these models and ray trace through the full three-dimensional matter distributions to generate mock lensing observables. We then fit these lensing observables with three different lensing models, which have different methodologies for treating the ENV/LOS, to understand the precision and accuracy of the recovered lens galaxy properties and H_0 .

The first method (which we term “Lens-Only”) neglects mass outside the main lens altogether. This approach almost surely produces biased constraints on the Hubble constant, and so it is rarely used in studies that derive constraints on precision cosmology, but it still appears in some studies that focus on lens and source properties (e.g., Calanog et al. 2014; Hezaveh et al. 2013).

The second type of model (“Lens+Shear”) accounts for tidal effects in the lens plane by treating the magnitude and direction of the external shear as free parameters to be optimized for individual lens systems. This widely-used approach has several limitations. First, the tidal approximation neglects higher-order effects beyond shear, which may be significant for objects sufficiently close to the optical axis. This approach also neglects non-linear effects that arise from having mass in multiple redshift planes (McCully et al. 2014; Jaroszynski & Kostrzewa-Rutkowska 2012). Also, lens models cannot directly constrain the external convergence because of the mass sheet degeneracy (Falco et al. 1985). To avoid biases in the derived cosmological parameters, corrections for external convergence must be applied using additional constraints such as weak lensing (Nakajima et al. 2009; Fadely et al. 2010) or the number density of galaxies near the lens (Suyu et al. 2010, 2013; Collett et al. 2013; Greene et al. 2013; Rusu et al. 2016; Wong et al. 2016) calibrated by ray tracing through cosmo-

logical simulations (e.g., Hilbert et al. 2009). However, these corrections are derived from statistical studies that may have limited applicability to individual lens environments. Also, the shear derived from lens models does not always match the shear calculated from modeling the ENV/LOS directly (Wong et al. 2011).

The third type of model (“3-D Lens”) is to build full, three-dimensional lens models that explicitly account for mass in the environment and along the LOS. While our modeling approach is more complicated and observationally expensive, it yields a direct connection between the convergence and shear in lens models and the physical mass in the beam. The full multi-plane lens equation is not computationally tractable for use with these complex models, so we use the framework we developed in McCully et al. (2014, hereafter M14). In our framework, most (or all) of the ENV/LOS galaxies are treated using the tidal approximation, but the full three dimensional structure of the mass distribution is preserved. This has two major benefits. The first is that we can balance efficiency (by treating galaxies in the tidal approximation) and accuracy (by including non-linear redshift effects not captured by a simple, external shear). The second benefit is that by varying which galaxies are treated exactly, we can disentangle non-linear effects produced by galaxies along the LOS, and higher order terms produced by ENV/LOS galaxies that are close in projection to the lens.

Section 2 describes our methodology for handling 3-D mass models, including a new treatment of voids in our multi-plane lensing formalism (Section 2.3). Once we have the mock lensing data, we fit different types of lens models to test how well we can recover parameters related to the input galaxy and cosmology. In Section 3.1 we use the models to understand what makes a perturbing galaxy important for lensing, and we derive a quantity that shows the dependence on the perturber’s mass, projected offset, and redshift. In Section 3.2 we use this quantity to characterize ENV/LOS effects and compare different lens fields. In Section 3.3 we then examine the parameters recovered from the different types of fitted models to understand statistical and systematic uncertainties that are associated with ENV/LOS effects. Finally, in Section 3.4 we vary the properties of the main lens galaxy to ascertain which lens systems will provide stronger constraints on the Hubble constant.

Throughout this paper we assume a cosmology with $\Omega_M = 0.274$, $\Omega_\Lambda = 0.726$, and $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

2. GRAVITATIONAL LENSING IN 3-D

Our goal is to quantify ENV/LOS effects using realistic lensing beams. Since the analysis is inherently three-dimensional, we begin by reviewing multi-plane lensing as the analytical infrastructure that is necessary any time we consider mass at a different redshift than the main lens galaxy (Section 2.1). To examine ENV/LOS effects, we build mock lenses using realistic matter distributions based on our photometric and spectroscopic observations. Our mass models consist of two components: discrete mass structures such as group and galaxy halos, which are discussed in Section 2.2; and a smooth background mass density, which can account for under-dense regions (voids) and can be treated within our multi-plane lensing framework using the new approach described in

² Here we focus on the Hubble constant, but similar arguments apply for time-delay cosmography measurements that have been proposed to be used to measure dark energy (Treu et al. 2013).

Section 2.3. We ray trace from many source positions through these mass models to produce simulated lensing observables including image positions, flux ratios, and time delays, as discussed in Section 2.4. Finally, we fit these lensing observables with different models (described in Section 2.5) to understand whether different assumptions about the ENV/LOS lead to bias and/or scatter in the recovered values of lens galaxy parameters and the Hubble constant.

2.1. Multi-plane Lensing

As we are interested in how mass in the ENV/LOS affects lensing measurements, we must go beyond the standard lens equation that only includes a single lens. Instead we use the multi-plane lens equation (which makes the thin-lens approximation at each redshift of interest; Petters et al. 2001, see Petkova et al. 2014 for an implementation of multi-plane lensing in the GLAMER code, Metcalf & Petkova 2014) that traces a light ray from the observer to the source, bending the ray at each redshift plane of interest. The full multi-plane lens equation is given by

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

where

$$\beta_{ij} = \frac{D_{ij} D_s}{D_j D_{is}}. \quad (2)$$

Our mass models typically contain hundreds of components, but many of the corresponding terms in the lens equation can be treated with the tidal approximation characterized by convergence and shear.³ In our previous paper (M14; see Schneider 2014 for an alternative derivation), we presented a hybrid framework for multi-plane lensing that incorporates all of the tidal planes (weak lenses) into a set of matrices so the sum includes only “main” lens planes (strong lenses, i.e., those not treated with the tidal approximation). Our framework recovers the full multi-plane lens equation when all of the redshift planes are treated as “main” planes. Even when it uses the tidal approximation, our framework still includes non-linear effects associated with having mass at different redshifts. The generalized expressions for the lens equation, magnification tensor, and time delay are as follows:

$$\mathbf{x}_j = \mathbf{B}_j \mathbf{x}_1 - \sum_{\ell \in \{\ell_\mu < j\}} \mathbf{C}_{\ell j} \boldsymbol{\alpha}_\ell(\mathbf{x}_\ell), \quad (3)$$

$$\mathbf{A}_j = \mathbf{B}_j - \sum_{\ell \in \{\ell_\mu < j\}} \mathbf{C}_{\ell j} \boldsymbol{\Gamma}_\ell \mathbf{A}_\ell, \quad (4)$$

$$T = \frac{1}{2} \mathbf{x}_s \cdot \mathbf{F}_s \mathbf{x} + \sum_{\ell \in \{\ell_\mu\}} \left[\frac{1}{2} \tau_{\ell s} \mathbf{x}_\ell \cdot \boldsymbol{\alpha}_\ell - \frac{1}{2} \mathbf{x}_s \cdot \mathbf{G}_{\ell s} \boldsymbol{\alpha}_\ell - \tau_{\ell s} \phi_\ell \right], \quad (5)$$

where the sums run over “main” planes, $\boldsymbol{\alpha}_\ell$ is the deflec-

tion in main plane ℓ , and

$$\tau_{ij} = \frac{1 + z_i}{c} \frac{D_i D_j}{D_{ij}}. \quad (6)$$

All of the tidal effects are characterized by the following matrices:

$$\mathbf{B}_j = \mathbf{I} - \sum_{i=1, i \notin \{\ell_\mu\}}^{j-1} \beta_{ij} \boldsymbol{\Gamma}_i \mathbf{B}_i, \quad (7)$$

$$\mathbf{C}_{\ell j} = \beta_{\ell j} \mathbf{I} - \sum_{i=\ell+1, i \notin \{\ell_\mu\}}^{j-1} \beta_{\ell i} \boldsymbol{\Gamma}_i \mathbf{C}_{\ell i}, \quad (8)$$

$$\mathbf{F}_j \equiv - \sum_{i=1, i \notin \{\ell_\mu\}}^{j-1} \tau_{is} \boldsymbol{\Gamma}_i \mathbf{B}_i, \quad (9)$$

$$\mathbf{G}_{\ell j} \equiv \tau_{\ell s} \mathbf{I} - \sum_{i=\ell+1, i \notin \{\ell_\mu\}}^{j-1} \tau_{is} \boldsymbol{\Gamma}_i \mathbf{C}_{\ell i}. \quad (10)$$

These matrices can be computed once at the start of any lens modeling analysis and stored for repeated use yielding a large increase in computational efficiency. We note that the \mathbf{C} and \mathbf{G} matrices depend on ratios of cosmic distances, so these matrices are only constant if the cosmological parameters besides H_0 are fixed. When other cosmological parameters are allowed to vary, e.g., constraining Ω_Λ , it is necessary to recompute these matrices. One possibility to overcome this computational challenge is to compute a grid of \mathbf{C} and \mathbf{G} matrices and interpolate to a given set of cosmological parameters.

Note in eq. 3 that each main plane needs to be evaluated using the positions \mathbf{x}_ℓ , which are not generally the same as the positions \mathbf{x}_1 on the sky and must be computed with the lens equation. This distinction gives rise to non-linearities that cannot be mimicked by a simple shear. Our framework is therefore more accurate than fitting a simple shear, but it is also more efficient than the full multi-plane lens equation because the recursive sums include only main planes. It fills the gap between using the full multi-plane lens equation (which can be computationally expensive) and treating everything as a simple external shear (which omits higher-order effects that can be significant for objects projected near the lens and does not include non-linear redshift effects).

If we take the observer to be in plane $i = 0$, we can write $\beta_{0j} = 1$ and $\tau_{0j} = 0$. We then realize that the \mathbf{B} and \mathbf{C} matrices are related, as are the \mathbf{F} and \mathbf{G} matrices:

$$\mathbf{B}_j = \mathbf{C}_{0j} \quad \text{and} \quad \mathbf{F}_j = \mathbf{G}_{0j}. \quad (11)$$

If there is a single main plane, the multi-plane lens equation reduces to

$$\mathbf{x}_s = \mathbf{B}_s \mathbf{x}_1 - \mathbf{C}_{\ell s} \boldsymbol{\alpha}_\ell(\mathbf{B}_\ell \mathbf{x}_1). \quad (12)$$

This resembles the standard single-plane lens equation except that \mathbf{B}_ℓ appears in the argument of the deflection of the main lens galaxy, leading to non-linear effects that we quantify and explore using our numerical results in the following sections.

2.2. Building Mass Models from Observations

³ Sections 3.1 and 3.3 discuss the objective criterion we use to determine which perturbors can be treated with the tidal approximation.

Next we consider our methodology for constructing mass models from photometric and spectroscopic observations. Our mass models are made of two components: a smooth background density and discrete objects, namely galaxies and group dark matter halos. We first consider the discrete structures in our models.

We include discrete mass structures using an updated version of the methodology in Wong et al. (2011). We start by introducing galaxies detected in our photometric observations into our mass models. The photometric catalogs are based on 12 nights of imaging data from the 4m Mayall telescope on Kitt Peak and the 4m Blanco telescope on Cerro Tololo (Williams et al. 2006). The photometric data is highly complete down to $I \leq 21.5$. For many of these objects, we have spectroscopic redshifts, which allows us to include them at their measured 3-D position. The spectroscopic sample is presented in Momcheva et al. (2015), and consists of 40 nights of 6.5m telescope observations using the LDSS-2, LDSS-3, IMACS, and Hectospec instruments (Sluse et al. 2016 present a similar dataset). For LOS galaxies that lack spectroscopy, we must assign a redshift to get its 3-D position. If the field is in the SDSS footprint, we assign photometric redshifts from SDSS DR9 (Ahn et al. 2012; Csabai et al. 2003). If not, we draw from a smoothed version of the spectroscopic redshift distribution in that particular field.

Once we have the 3-D position of the discrete mass structures (RA, Dec, and redshift), we must choose a mass distribution for the galaxy. We model each galaxy as a truncated singular isothermal sphere. This is a surprisingly good approximation for observed galaxy-scale lenses (e.g., Rusin et al. 2003; Gavazzi et al. 2007; Koopmans et al. 2009, but see Xu et al. 2016). Throughout our analysis, most of the galaxies included in our mass models are significantly offset in projection from the main lens galaxy limiting our sensitivity to exact mass distribution of the galaxies along the LOS.

The truncation radii are set to be equivalent to r_{200} . The velocity dispersions and r_{200} are drawn from scaling relations (including scatter) derived from a galaxy-galaxy weak lensing analysis in the CFHTLS (Brimioulle et al. 2013). We use the luminosity scaling relationships that evolve with redshift to find r_{200} and velocity dispersion. We adopt the relations found in Brimioulle et al. (2013). The relationship for r_{200} is given by $r_{200}/r_0 = (L/L_0)^{\eta_r}$ where $L_0 = 1.6 \times 10^{10} h^{-2} L_\odot (1+z)$. For blue galaxies ($B-V \leq 0.7$), $r_0 = 133 \pm 3 h^{-1}$ kpc and $\eta_r = 0.40 \pm 0.08$, while for red galaxies ($B-V > 0.7$), $r_0 = 198 \pm 3 h^{-1}$ kpc and $\eta_r = 0.38 \pm 0.04$. For the velocity dispersion scaling relations, Brimioulle et al. (2013) give $\sigma/\sigma_0 = (L/L_0)^{\eta_\sigma}$. L_0 is the same as for the radius scaling relation. For blue galaxies, $\sigma_0 = 123 \pm 3$ km s $^{-1}$ and $\eta_\sigma = 0.24 \pm 0.03$, while for red galaxies, $\sigma_0 = 173 \pm 2$ km s $^{-1}$ and $\eta_\sigma = 0.25 \pm 0.03$.

For lenses that are in a group or cluster, we treat the group member galaxies in a slightly different way. We compute a total dynamical mass for the group (Girardi et al. 1998; Momcheva et al. 2006; Ammons et al. 2014; Momcheva et al. 2015; Wilson et al. submitted) and apportion the mass between the group dark matter halo and the group galaxies. The dark matter halo is assumed to be a spherical NFW profile. The halo concentration is taken from the mass-concentration relation from simula-

tions by Zhao et al. (2009), with an assumed scatter of 0.14 dex (Bullock et al. 2001). The fraction of the total group mass assigned to the halo, f_{halo} , is randomly drawn from a uniform distribution such that the group galaxies are not truncated below twice their effective radii, nor beyond their R_{200} (the radius at which the mean density enclosed is 200 times the matter density at that redshift). The remainder of the mass, $(1 - f_{\text{halo}})M_{\text{group}}$, is assigned to the group galaxies, which are scaled such that they have the same density at their truncation radii. All group galaxies and the group halo are assumed to be in the same redshift plane as the lens, even though there may be slight redshift differences due to peculiar velocities.

Some lenses have groups along the LOS, (e.g., B0712; Fassnacht & Lubin 2002). Groups along the LOS of sight are treated in less detail. The galaxies are included as described above, but mass is not apportioned to a group NFW halo. For this current work, we are less concerned about reproducing exact lens fields and are more concerned about producing *realistic* 3-D mass models so that we can study how the LOS affects lens models. In future models we will account for groups along the LOS using an updated version of the methodology in Ammons et al. (2014) (see also Wilson et al. submitted).

This methodology allows us to explicitly include matter that we see (e.g., galaxies). However, there may be matter that we do not observe like faint galaxies or dark matter filaments. We include the rest of that mass that we do not explicitly include in our discrete mass structures described above in a smooth, uniform mass component which is described in more detail in the next section.

2.3. Setting the Smooth Background Density

We now turn our attention to the underlying, smooth density component of our mass models. In traditional single-plane lensing, it is reasonable to assume that the lens galaxy does not significantly contribute to the overall geometry of the universe, and thus to place it on top of a smooth background given by the mean density of the universe. As we begin to include more and more galaxies along the line sight, however, the cosmological effects become non-negligible. Simply adding the galaxies on top of the mean density would lead to biased results. If our mass models were 100% complete, there should be no smooth background component at all. In practice, we need to find an intermediate approach that scales the background density to account for incompleteness in our mass models (which can arise, for example, from having flux-limited photometric and spectroscopic data).

There are several ways we could do this. One possibility is to start with an *empty* beam and add both galaxies and a smooth density field. However, we would still need to account for the mean density of the universe outside the beam when computing cosmic distances. A different approach, which we prefer, is to start with a *mean density* beam. We can then subtract a smooth density field that corresponds to the amount of mass that we have included in our mass model by discrete structures (e.g., galaxies). The benefit is that we can then compute cosmic distances using the standard mean density universe expressions while still accounting for the lensing effects produced by having a smooth background density that is less than the mean density. In practice, we can esti-

mate our mass completeness (which is explicitly included as discrete structures) relative to the smooth mean density of the universe (which we note that this is different than the completeness of the number of observed galaxies). We can then subtract off the fraction of mass in our discrete structures from the mean density of the universe.

Our goal is to do these calculations using our multi-plane framework. By doing this, we can still take advantage of the computational efficiency of the framework, while maintaining the ability to include non-linear redshift effects in lens models. In our framework, we can subtract from the mean density by inserting redshift planes with negative mass density (i.e., negative convergence) in between the planes that contain galaxies. Our framework, by construction, is built of discrete mass structures. However, in the limit of a large number of thin planes, the sums over “void” planes becomes integrals that we now derive.

Consider a plane that represents the projection of a thin slab extending from proper distance r_p to $r_p + dr_p$. If the density contrast within this slab is $\delta\rho_p = \rho_p - \bar{\rho}_p$, the lensing effects are characterized by a convergence

$$d\kappa = \frac{\delta\rho_p dr_p}{\Sigma_{\text{cr}}} \quad (13)$$

where the critical surface density for lensing is

$$\Sigma_{\text{cr}} = \frac{c^2}{4\pi G} \frac{D_s}{D_\ell D_{\ell s}} \quad (14)$$

Converting to comoving distances $X(z)$, we can write

$$d\kappa = \frac{4\pi G}{c^2} \frac{X(X_s - X)}{X_s} \frac{\delta\rho_p}{(1+z)^2} dX. \quad (15)$$

In order to achieve the desired subtraction, we set the density contrast to be

$$\delta\rho_p = -f(z)\bar{\rho}(z) \quad (16)$$

where $f(z)$ is the mass completeness in our models (which can vary with redshift), and the mean density of the universe at redshift z is $\bar{\rho}(z) = \rho_c \Omega_m (1+z)^3$, where $\rho_c = 3H_0^2/8\pi G$ is the critical density of the universe today.

We now need to incorporate the “void” planes into our LOS framework. The general expression for the \mathbf{C} matrix⁴ between redshifts z_i and z_j is given in equation 8. The sum in that equation now includes both tidal and void planes, but we can handle the sums over void planes by redefining the β factors as follows. If there is only mean density between redshifts z_i and z_j , then there are no planes in the sum and hence $\mathbf{C}_{ij} = \beta_{ij}\mathbf{I}$: the connection between the planes is an isotropic tensor that depends on distance ratios. Now suppose there are only void planes between z_i and z_j , with $\mathbf{\Gamma}_k \rightarrow \Delta\kappa_k\mathbf{I}$. Since void planes are isotropic, \mathbf{C}_{ij} will still be isotropic. Therefore, let us define a new $\hat{\beta}_{ij}$ such that $\mathbf{C}_{ij} = \hat{\beta}_{ij}\mathbf{I}$. We point out that $\mathbf{C}_{ik} = \hat{\beta}_{ik}\mathbf{I}$ for all $i < k < j$. Using this definition along with the definition of \mathbf{C} in equation

8, we see that $\hat{\beta}_{ij}$ must follow the recursion relation

$$\hat{\beta}_{ij} = \beta_{ij} - \sum_{k=i+1}^{j-1} \hat{\beta}_{ik}\beta_{kj}\Delta\kappa_k. \quad (17)$$

Now if plane j is a tidal plane that includes a galaxy, we write the shear tensor as $\mathbf{\Gamma}_j + \Delta\kappa_j\mathbf{I}$. Then,

$$\mathbf{C}_{ij+1} = \hat{\beta}_{ij} - \hat{\beta}_{jj+1}\mathbf{\Gamma}_j\mathbf{C}_{ij}, \quad (18)$$

where we have used $\beta_{kk+1} = \hat{\beta}_{kk+1}$.

Next we consider $j+1$ as a void plane and $j+2$ as a tidal plane and find that

$$\mathbf{C}_{ij+2} = \hat{\beta}_{ij+2} - \hat{\beta}_{jj+2}\mathbf{\Gamma}_j\mathbf{C}_{ij} \quad (19)$$

$$\mathbf{C}_{ij+3} = \hat{\beta}_{ij+3} - \hat{\beta}_{jj+3}\mathbf{\Gamma}_j\mathbf{C}_{ij} - \hat{\beta}_{j+2j+3}\mathbf{\Gamma}_{j+2}\mathbf{C}_{ij+2}. \quad (20)$$

This generalizes to

$$\mathbf{C}_{ij} = \hat{\beta}_{ij} - \sum_{k=i+1}^{j-1} \hat{\beta}_{ik}\mathbf{\Gamma}_k\mathbf{C}_{ik}, \quad (21)$$

where the sum is now only over tidal planes. All of the effects of void planes are now contained in the $\hat{\beta}$ factors.

In the limit of a large number of thin void planes, the sum in equation 17 becomes a continuous integral:

$$\hat{\beta}_{ij} = \beta_{ij} - \int_{z_i}^{z_j} \hat{\beta}(z_i, z)\beta(z, z_j)d\kappa(z) \quad (22)$$

This expression is recursive and thus unwieldy in practice, but it suggests that $\hat{\beta}$ can be described by a differential equation. We are going to take derivatives of 22 with respect to comoving distances, so it is first useful to examine derivatives of β_{ij} . In terms of comoving distances, X_i ,

$$\beta_{ij} = \frac{D_{ij}D_s}{D_{is}D_j} = \frac{(X_j - X_i)X_s}{X_j(X_s - X_i)} \quad (23)$$

$$\frac{d\beta_{ij}}{dX_j} = \frac{X_iX_s}{X_j^2(X_s - X_i)} \quad (24)$$

$$\frac{d^2\beta_{ij}}{dX_j^2} = -2\frac{X_iX_s}{X_j^3(X_s - X_i)} = -\frac{2}{X_j} \frac{d\beta_{ij}}{dX_j} \quad (25)$$

Now, differentiating equation 22 with respect to $X_j = X(z_j)$ yields

$$\frac{d\hat{\beta}_{ij}}{dX_j} = \frac{d\beta_{ij}}{dX_j} - \int_{X_i}^{X_j} \hat{\beta}(z_i, z) \frac{d\beta(z, z_j)}{dX_j} \frac{d\kappa}{dX} dX. \quad (26)$$

where we use $\beta(z_j, z_j) = 0$. The second derivative is

$$\begin{aligned} \frac{d^2\hat{\beta}_{ij}}{dX_j^2} &= \frac{d^2\beta_{ij}}{dX_j^2} - \hat{\beta}_{ij} \frac{d\kappa}{dX_j} \frac{d\beta(z, z_j)}{dX_j} \Big|_{z \rightarrow z_j} \\ &\quad - \int_{X_i}^{X_j} \hat{\beta}(z_i, z) \frac{d^2\beta(z, z_j)}{dX_j^2} \frac{d\kappa}{dX} dX. \end{aligned} \quad (27)$$

Now we combine equations 26 and 27 in the following

⁴ The following argument also applies to the \mathbf{B} matrices since they are a subset of the \mathbf{C} matrices (see eq. 11).

way:

$$\begin{aligned} \frac{d^2 \hat{\beta}_{ij}}{dX_j^2} + \frac{2}{X_j} \frac{d\hat{\beta}_{ij}}{dX_j} = & \left[\frac{d^2 \beta_{ij}}{dX_j^2} + \frac{2}{X_j} \frac{d\beta_{ij}}{dX_j} \right] - \hat{\beta}_{ij} \frac{d\kappa}{dX_j} \frac{d\beta(z, z_j)}{dX_j} \Big|_{z \rightarrow z_j} \\ & - \int_{X_i}^{X_j} \hat{\beta}(z_i, z) \left[\frac{d^2 \beta(z, z_j)}{dX_j^2} + \frac{2}{X_j} \frac{d\beta_{ij}}{dX_j} \right] \frac{d\kappa}{dX} dX. \end{aligned} \quad (28)$$

The terms in square brackets vanish by equation 25. Then using equations 15 and 26 in the remaining term on the right-hand side yields

$$\frac{d^2 \hat{\beta}_{ij}}{dX_j^2} + \frac{2}{X_j} \frac{d\hat{\beta}_{ij}}{dX_j} = -\frac{4\pi G}{c^2} \frac{\delta\rho_p}{(1+z_j)^2} \hat{\beta}_{ij}. \quad (29)$$

This is our desired differential equation, which can be solved using standard numerical techniques. Since it is a second-order equation, we need two boundary conditions. The first is $\hat{\beta}_{ii} = 0$. The second condition is based on the first derivative with

$$\left. \frac{d\hat{\beta}_{ij}}{dX_j} \right|_{j \rightarrow i} = \frac{D_s}{D_i D_{is}}. \quad (30)$$

We can follow a similar set of arguments for time delays and the \mathbf{G}_{ij} matrices. As before, if there are only void planes between z_i and z_j , we can define $\mathbf{G}_{ij} = \hat{\tau}_{ij} \mathbf{I}$. This implies that $\hat{\tau}_{ij}$ should follow the following relation:

$$\hat{\tau}_{ij} = \tau_{is} - \sum_{k=i+1}^{j-1} \hat{\beta}_{ik} \tau_{ks} \Delta\kappa_k. \quad (31)$$

The continuous limit is

$$\hat{\tau}_{ij} = \tau_{is} - \int_{X_i}^{X_j} \hat{\beta}(z_i, z) \tau(z, z_s) \frac{d\kappa}{dX} dX \quad (32)$$

We now take a derivative of $\hat{\tau}_{ij}$ with respect to X_j and obtain

$$\frac{d\hat{\tau}_{ij}}{dX_j} = -\hat{\beta}_{ij} \tau_{js} \frac{d\kappa}{dX_j}. \quad (33)$$

This is a first order differential equation, so we only need one boundary condition. The simplest choice is $\hat{\tau}_{ij} \rightarrow \tau_{is}$ as $j \rightarrow i$ because in that case there is no void correction. Following a similar argument to what we used for \mathbf{C}_{ij} then yields

$$\mathbf{G}_{ij} = \hat{\tau}_{ij} - \sum_{k=i+1}^{j-1} \hat{\tau}_{ik} \Gamma_k \mathbf{C}_{ik}. \quad (34)$$

These updates to the \mathbf{C}_{ij} and \mathbf{G}_{ij} matrices provide a straightforward way to account for the lensing effects of a smooth density field. In principle the density can be arbitrary, but in practice we use it to compensate for the matter that is explicitly included in our mass models. This correction avoids biases that would be introduced if we simply added all galaxies on top of the mean mass density of the universe.

In practice, we parameterize the mass completeness, $f(z)$, with a simple analytic function. Using all 23 fields

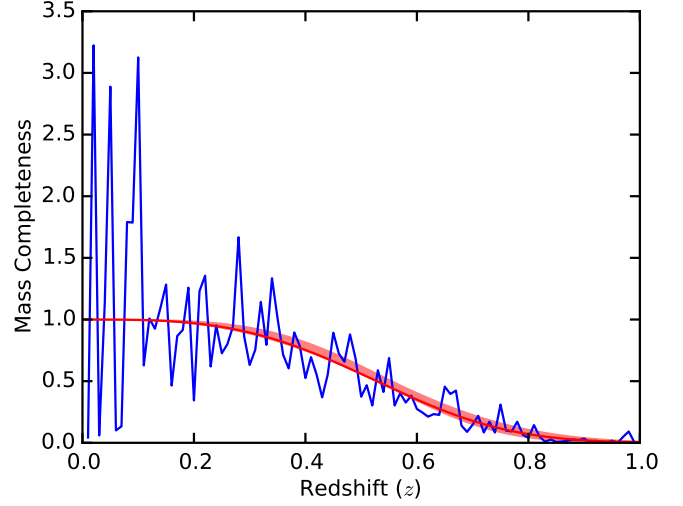


FIG. 1.— Mass completeness as a function of redshift. The blue line shows the average mass per redshift bin, relative to a smooth universe, across the 23 fields for which we have constructed 3-D mass models. The red line shows the best fit to the data using a functional form of $f(z) = e^{-z^a/b}$ with best fit parameters $a = 3.23$ and $b = 0.183$. The red band shows the 16-84% confidence interval derived from MCMC. We then use this mass completeness function to set the lensing contribution from the smooth background mass density. This allows us to self consistently include voids. Because this is a fit to the average across mass models, individual beams can have higher or lower than the mean density of the universe. The only assumption is that the average of our beams should be approximately the mean density of the universe.

for which we have constructed mass models, we fit the average mass per redshift bin, relative to a smooth universe, using a function of the form $f(z) = e^{-z^a/b}$ as shown in Figure 1. This function is a generalization of a Gaussian that captures the near-total completeness at low redshift (with $f(z=0) = 1$) and smoothly decreases in a similar manner to our observed mass completeness. We find best fit values of $a = 3.23$ and $b = 0.183$. We then run MCMC (Foreman-Mackey et al. 2013) to estimate the uncertainties on the fitted parameters to the mass completeness. The confidence 1- σ confidence interval is shown in the red band in Figure 1.

We adopt this mass completeness function for all subsequent analysis. Because we have used the average across all of our beams instead of fitting each beam individually, this allows for individual beams to be over or under dense compared to the mean density of the universe. The only assumption is that the mean of our total sample is approximately the mean density of the universe in each redshift bin.

2.4. Generating Mock Lensing Observables

Next, we place a lens galaxy into our ENV/LOS mass models, put a variety of sources behind it, and solve the lens equation to generate “mock” lenses.⁵ We treat the main lens galaxy using a ellipsoidal power law model given by

$$\kappa = \frac{b^{2-\eta}}{2\xi^{2-\eta}} \quad \text{where} \quad \xi^2 = x^2 + \frac{y^2}{(1-e)^2} \quad (35)$$

⁵ All of our simulations and modeling use the *lensmodel* package (Keeton 2001).

Here ξ is the elliptical radius and e is the ellipticity. The power law index η is chosen so the enclosed mass scales as $M(r) \propto r^\eta$. The Einstein radius for this model is $R_E = b\eta^{\eta-2}$. Our initial models use $\eta = 1$ which corresponds to a singular isothermal ellipsoid.

We treat the azimuthal angle of the main lens galaxy and the position of the source as nuisance parameters to marginalize. Specifically, we draw random values for the lens galaxy orientation from a uniform distribution. We choose source positions from a magnification weighted distribution to approximate the magnification bias that is present in real observations (see Keeton & Zabludoff 2004). Overall, we generate a sample of 300 quad and 300 double mock image configurations for each mass model.

We use the full recursive lens equation (eq. 1 with no approximations) to ray trace through our mass models and generate mock lensing observables: image configuration, image positions, time delays, and flux ratios (similar to Petkova et al. 2014). We imagine observing these lenses with fiducial measurement uncertainties of 3 milliarcseconds for the positions of the lensed images and the main galaxy, 5% for the fluxes, and 1 day for the time delays. These values are typical of observations with instruments such as *HST* and monitoring campaigns such as COSMOGRAIL (Eigenbrod et al. 2005). We do not explicitly add observational scatter to the mock data; preliminary tests indicate that such measurement noise introduces a “floor” in the χ^2 values and the scatter in recovered lens parameters but does not qualitatively change our results.

Two-image lens systems have the following observables:

- 2×2 image positions (x, y)
- 2 fluxes
- 1 time delay
- 1×2 lens galaxy position (x, y)

giving a total of 9 constraints.

By contrast, four-image systems have:

- 4×2 image positions (x, y)
- 4 fluxes
- 3 time delays
- 1×2 lens galaxy position (x, y)

giving a total of 17 constraints.

2.5. Fitting Lens Models

We can then fit the simulated observables as we would for real lens systems. We consider three types of models that treat the ENV/LOS in different ways to see if we can recover the true, input lens galaxy parameters and the Hubble constant. We consider the following models: (1) The “Lens-Only” model ignores ENV/LOS effects altogether. (2) The “Lens+Shear” model attempts to account for ENV/LOS effects by fitting an external shear in the main lens plane (e.g., Suyu et al. 2013). (3) The “3-D Lens” model uses our multi-plane framework developed in M14. However, most (or all) of the ENV/LOS galaxies

are treated using the tidal approximation. Their effects are included in the models using the C and G matrix sets. The C and G matrices are not allowed to vary in the fits for this analysis, but in principle we could be marginalize over them. This allows us to disentangle non-linear effects, which are accounted for in our 3-D Lens models, and higher order terms which are not. If we removed the 3-D structure from our models, we would recover the Lens-Only model. The models all have the following 10 free parameters:

- Source position (x, y)
- Source flux
- Hubble constant (h)
- Einstein radius of the main lens galaxy
- Lens galaxy position (x, y) .
- Lens galaxy ellipticity and orientation (e, θ_e)
- Lens galaxy power law index (η)

The Lens+Shear model has two additional free parameters that characterize the external shear; we use the pseudo-Cartesian components of the shear, $\gamma_c = \gamma \cos 2\theta_\gamma$ and $\gamma_s = \gamma \sin 2\theta_\gamma$.

Given the numbers of parameters and constraints, quad lenses are well constrained with 7 degrees of freedom (dof) for the Lens-Only and 3-D Lens models and 5 dof for the Lens+Shear model. Double lenses are, however, under constrained with -1 dof for the Lens-Only and the 3-D Lens models and -3 dof for the Lens+Shear model.

Because the Lens+Shear model has a different number of dof, care must be taken when comparing the results to the other models. We scale the χ^2 value by the 95% confidence limit for the number of dof to ensure a fair comparison across model types.

In some cases it may be reasonable to put priors on the lens galaxy ellipticity and orientation based on the light profile (e.g., Bolton et al. 2008), but in lensing there is not always a strong relation between the light and mass (Bruderer et al. 2015, and references therein). Therefore we do not place any priors on the lens galaxy parameters. We do assume a weak Gaussian prior on the Hubble constant, $h = 0.71 \pm 0.3$, which ensures that models are reasonable but otherwise has little effect on the results for quad lenses. The prior plays a stronger role in double lenses since they are under constrained. Overall, the results from double lenses are qualitatively similar to those from quad lenses but considerably broader. For the remainder of the discussion we focus on quad lenses because they are better constrained.

In all cases, the model parameters are varied to find the best fit and allowed range. We seek to understand how the assumptions and approximations in the various models cause parameters to shift away from their true values. The Lens-Only models ignore the ENV/LOS entirely, so they serve as control test: the best fit values for the lens galaxy parameters and the Hubble constant move away from the truth in an attempt to account for the LOS/ENV effects allowing us to characterize the overall importance of ENV/LOS effects. The Lens+Shear

models have two extra free parameters, and we aim to test whether the shear can account for ENV/LOS effects well enough to provide robust constraints on the Hubble constant. Finally, the 3-D Lens models are based on the same mass models that were used to generate the mock lensing observables. The analysis is not circular (or only a test of the code), however, because the fits treat most or all of the ENV/LOS galaxies using the tidal approximation (whereas the mock data are generated with the full multi-plane lens equation). This approach allows us to isolate the effects of non-linear effects from higher order terms beyond the shear and convergence. Also, even with perfect knowledge of the ENV/LOS, there may be degeneracies between model parameters in the 3-D Lens models that we will explore below.

For the recovered lens quantities, we use Markov Chain Monte Carlo (MCMC) methods to calculate the full posterior probability distribution. When presenting marginalized results for a given lens parameter, we show the median value of the fitted parameters and estimate the scatter by measuring half of the difference between the 16th and 84th percentiles. When plotting results, we show fractional changes in the fitted parameters to emphasize the variation.

We now have all of the machinery in place to study the ENV/LOS contribution to lensing. We now turn to the results of our simulations.

3. RESULTS

To understand the contribution of the ENV/LOS to lensing, we break the problem into four pieces. We first examine the perturbations due to individual ENV/LOS galaxies to understand what combination of mass, projected offset, and redshift determine a perturber's importance (Section 3.1). We then apply the lessons to realistic mass models that contain hundreds of ENV/LOS galaxies (Section 3.2). After that, we fit the mock lensing observables generated with our realistic mass models to determine how well different types of models can recover the input galaxy parameters and cosmology (Section 3.3). Finally, we place different lens systems into the ENV/LOS mass models to understand how the mass and ellipticity of the lens galaxy and the position of the source affect a lens's sensitivity to ENV/LOS effects when inferring the Hubble constant (Section 3.4).

3.1. Which Individual Environment/LOS Galaxies are the Most Important?

While there are hundreds of galaxies in our beams, they are not equally important for lensing. Conceptually, we want to understand what properties control an individual perturber's influence on the lensed images. When characterizing an individual perturber, we test the effects of mass, projected offset from the main lens galaxy, and redshift. We initially control for redshift, placing the perturbing galaxy at the same redshift as the main lens but varying the mass and projected offset (Section 3.1.1). We then move the perturbing galaxy in redshift (Section 3.1.2). Based on the results from M14, we expect there to be a difference between perturbers in the foreground and background of the main lens due to non-linear effects that we aim to quantify. We use both equations and simulations to identify a quantity, based on the flexion produced by a perturbing galaxy, that characterizes

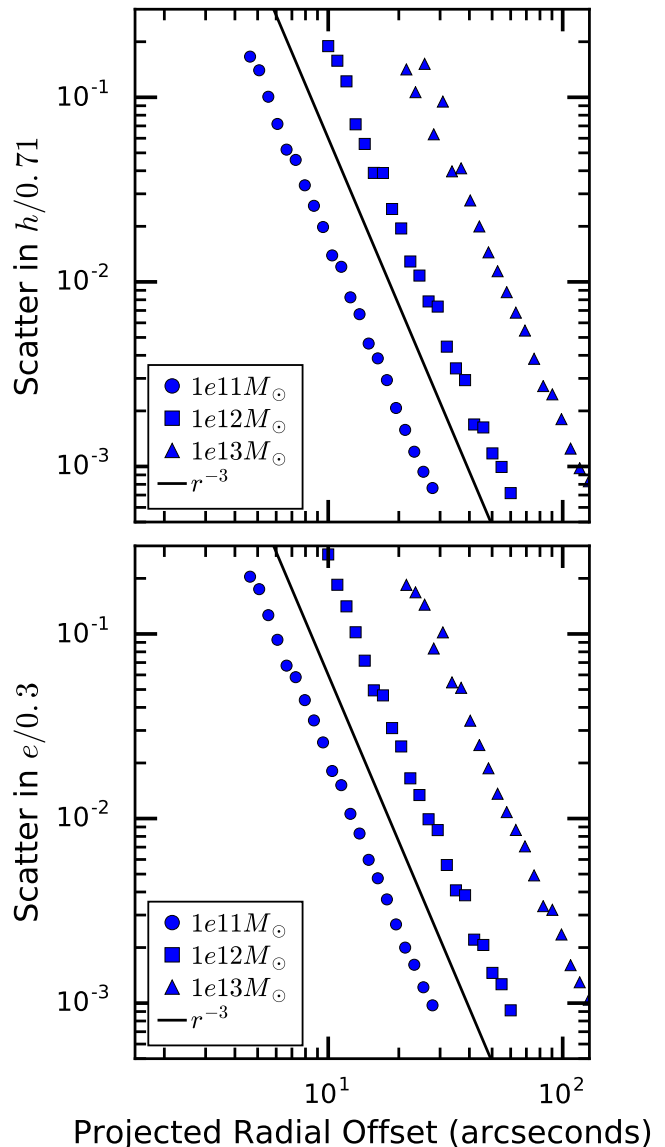


FIG. 2.— Deviations in lens model parameters for simulations with a single perturbing galaxy at the same redshift as the main lens galaxy. The horizontal axis is the radial offset of the perturber, and the vertical axis is the scatter in recovered values of the Hubble constant (h , top) and ellipticity (e , bottom). The points show results from simulations for three decades in the perturber's mass. The parameter deviations follow a r^{-3} power law (indicated with a line) because our models omit third-order terms and higher in the lens potential. The deviations also scale with perturber mass as $M \propto R_p^2$. Thus, understanding lens potential perturbations provides a robust way to characterize a perturbing galaxy's effect on the Hubble constant.

how much a given galaxy affects the recovered value of the Hubble constant.

3.1.1. Massive Galaxies Projected Near the Lens are More Important

We begin our characterization of individual perturbing galaxies with the simplest case: we examine one main lens galaxy and one perturbing galaxy in the same plane (removing any redshift effects). We first derive analytic expectations for how an individual perturbing galaxy affects the lens potential, and then show that the devia-

tions in the recovered Hubble constant scale in the same way as the lens potential perturbations. To derive our expectations for the behavior of our models, we make some simplifying assumptions (e.g., a point mass perturber), but we note that we only make these assumptions when deriving our criterion to characterize the individual contribution of a perturbing galaxy along the LOS. When fitting our realistic lens beams, we use truncated SIS potentials to model the galaxies in the ENV and along the LOS (see Section 2.2).

Consider a single main lens galaxy with deflection, α_g , and a single perturbing galaxy with deflection, α_p . Both are at the same redshift in this case. It is useful to define the lensing potential as

$$\alpha \equiv \nabla \phi. \quad (36)$$

If the perturbing galaxy does not overlap the lensed images, its gravitational effect is the same as a point mass so we can write the lens potential as

$$\phi_p = R_p^2 \ln |\mathbf{x} - \mathbf{r}| \quad (37)$$

where R_p is the Einstein radius of the perturber, \mathbf{x} is the image position in the redshift plane of the perturber, and \mathbf{r} is the position of the perturber. If we let $|\mathbf{x}| = x$, $|\mathbf{r}| = r$, and θ = the angle between the perturber and the image position as measured from the origin, we can rewrite the potential using the law of cosines as

$$\phi_p = \frac{1}{2} R_p^2 \ln(r^2 + x^2 - rx \cos \theta). \quad (38)$$

If the perturbing galaxy is far from the lensed images, then $x \ll r$ and we can expand the potential in a Taylor series:

$$\phi_p = R_p^2 \left[\ln(r) - \cos(\theta) \frac{x}{r} - \frac{1}{2} \cos(2\theta) \frac{x^2}{r^2} - \frac{1}{3} \cos(3\theta) \frac{x^3}{r^3} + \dots \right]. \quad (39)$$

This expression can be extended beyond a point mass model as

$$\phi_p = \phi(0) + \alpha^i(0)x^i + \frac{1}{2}\Gamma^{ij}x^ix^j + \frac{1}{6}\mathcal{F}^{ijk}x^ix^jx^k + \dots \quad (40)$$

where we have adopted the Einstein convention of summing over repeated indices and defined the tidal and flexion⁶ tensors:

$$\Gamma^{ij} \equiv \frac{\partial^2 \phi_p}{\partial x^i \partial x^j} \bigg|_{x=0}, \quad (41)$$

$$\mathcal{F}^{ijk} \equiv \frac{\partial^3 \phi_p}{\partial x^i \partial x^j \partial x^k} \bigg|_{x=0}. \quad (42)$$

For a point mass, the tidal and flexion amplitudes scale as $\Gamma \propto R_p^2/r^2$ and $\mathcal{F} \propto R_p^2/r^3$.

In the expansions, the zeroth order terms correspond to shifts in the zeropoint of the potential, which are unobservable. The first order terms correspond to a uniform deflection, which can be absorbed by translating

⁶ Throughout this work, we refer to third-order terms collectively as “flexion” terms, but see Bacon et al. (2006) for more discussion.

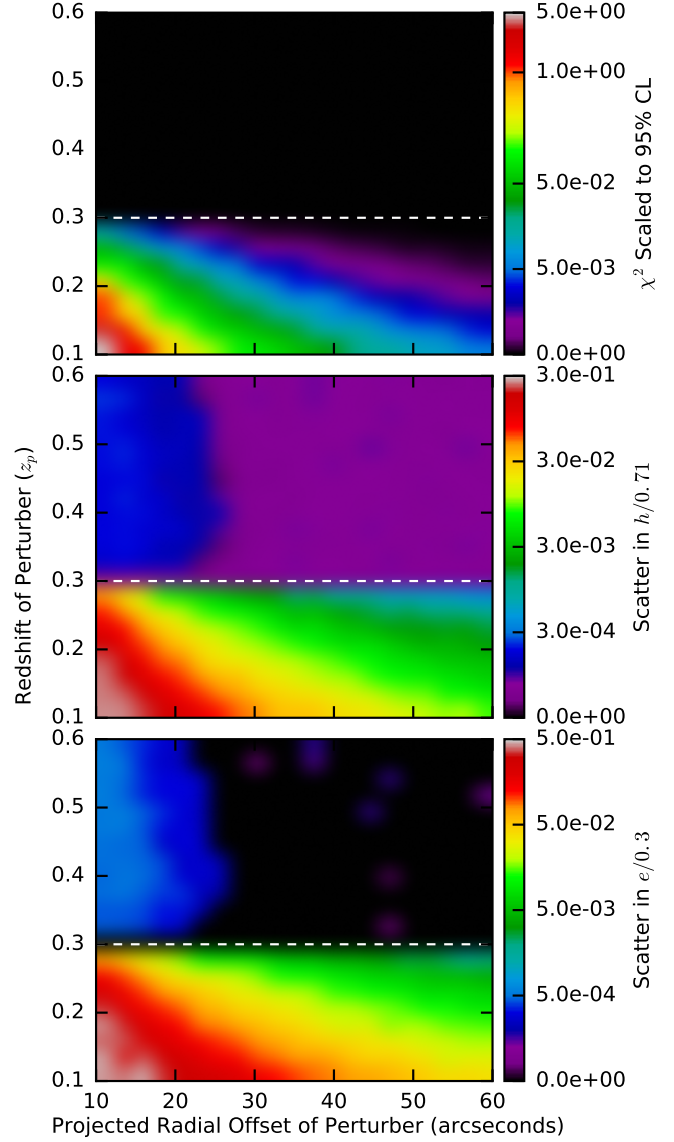


FIG. 3.— Recovered lens model parameters for simulations with only tidal effects from an LOS perturber. The simulations shown here do not include higher order terms, to emphasize non-linear effects. The main lens has a redshift of $z_{\text{lens}} = 0.3$ (marked by the white dashed line) and the perturber has a mass of $10^{12} M_{\odot}$. We use an asinh color scaling for dynamic range; (at large values, asinh acts like a logarithm, but at small values, it becomes linear). When the perturber is in the background, an effective shear can account for the LOS effects (see eq. 49). When the perturber is in the foreground, however, there are non-linear effects that cannot be mimicked by a shear in the lens plane.

the source plane. Therefore only the terms at second order and higher have observable consequences.

To quantify the environmental effects, we run simulations with an SIE main lens galaxy and a point mass perturber. We choose the Einstein radius of the main lens to be $R_E = 1.0''$, and we place the lens and perturber at redshift $z_{\text{lens}} = z_p = 0.3$ in front of a source at redshift $z_{\text{src}} = 2.0$. We consider three decades of perturber mass. Figure 2 shows the results in terms of scatter in the recovered values of the Hubble constant h and the ellipticity e .

The parameter scatter increases as the perturber gets

closer to the main galaxy and higher order terms become more important. The scalings follow power laws that are consistent with the deviations in the lens potential. Our models omit third-order terms and higher, so we expect the scatter to scale as r^{-3} . The scatter in the ellipticity and the Hubble constant are both proportional to the mass of the perturbing galaxy.

Therefore, the scatter in the recovered Hubble constant created by an individual perturbing galaxy is proportional to M/r^3 (note that $M \propto R_p^2$ for a point mass), which is the same scaling as the deviations in the lens potential. These results quantify our intuition that massive galaxies near the main lens are important for lensing.

Below, we will use these mass and offset scalings to characterize the strength of a perturbing ENV/LOS galaxy, but this current analysis only applies if the perturbing galaxy is at the same redshift as the main lens.

3.1.2. Foreground Perturbors are More Important than Background

Now that we understand the mass and offset scalings, we examine how the redshift of the perturbing galaxy affects its contribution to the lens potential.

For a single perturbing galaxy, the lens equation 1 has the form

$$\mathbf{x}_2 = \mathbf{x}_1 - \beta_{12}\alpha_1(\mathbf{x}_1) \quad (43)$$

and

$$\mathbf{x}_s = \mathbf{x}_1 - \beta_{1s}\alpha_1(\mathbf{x}_1) - \beta_{2s}\alpha_2(\mathbf{x}_2). \quad (44)$$

Combining these equations, recalling that $\beta_{is} \equiv 1$, and setting $\beta \equiv \beta_{12}$ lets us write the two-plane lens equation as

$$\mathbf{x}_s = \mathbf{x}_1 - \alpha_1(\mathbf{x}_1) - \alpha_2(\mathbf{x}_1 - \beta\alpha_1(\mathbf{x}_1)). \quad (45)$$

In the following discussion it is useful to recall that β varies monotonically with redshift, ranging from $\beta = 0$ when $z_p = z_{\text{lens}}$ to $\beta = 1$ when the perturber redshift reaches either $z_p = 0$ (for a foreground perturber) or $z_p = z_{\text{src}}$ (for a background perturber).

To make further progress, we need to distinguish situations when the perturber is in front of or behind the main lens galaxy. In M14, we showed that background perturbors only produce an additional shear while foreground perturbors produce non-linear effects.

Lens Potential Perturbations from a Background Galaxy

Suppose the perturbing galaxy is in the background. The first part of this discussion parallels Keeton (2003). We have $\alpha_1 = \alpha_g$ and $\alpha_2 = \alpha_p$. Substituting this into equation 45 yields

$$\mathbf{x}_s = \mathbf{x}_1 - \alpha_g(\mathbf{x}_1) - \alpha_p(\mathbf{x}_1 - \beta\alpha_g(\mathbf{x}_1)). \quad (46)$$

We Taylor expand α_p and omit the unobservable terms corresponding to the constant potential and deflection:

$$\begin{aligned} x_s^i &= x_1^i - \alpha_g^i(\mathbf{x}_1) - \Gamma^{ij}(x_1^j - \beta\alpha_g^j(\mathbf{x}_1)) \\ &\quad - \frac{1}{2}\mathcal{F}^{ijk}(x_1^j - \beta\alpha_g^j(\mathbf{x}_1))(x_1^k - \beta\alpha_g^k(\mathbf{x}_1)) + \mathcal{O}(\mathbf{x}_1^4). \end{aligned} \quad (47)$$

Here we use index notation for clarity.

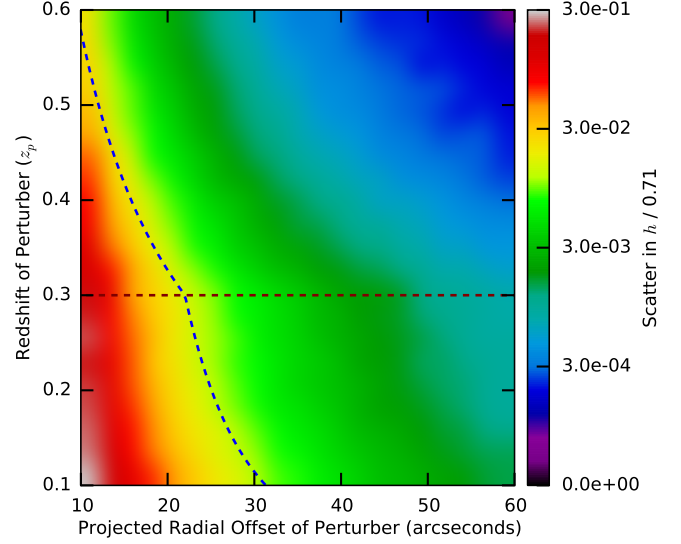


FIG. 4.— Scatter in the Hubble constant, h , for perturbors at different projected offsets and redshifts. The main lens has a redshift of $z_{\text{lens}} = 0.3$ which is marked with a dotted line, and the perturbing galaxy has a mass of $10^{12} M_{\odot}$. The dashed line shows a contour of constant flexion shift, $\Delta_3 x$, from equations 55 and 62 (picking a different value of $\Delta_3 x$ would simply rescale the contour). The effects due to a perturber in the background are downweighted (see equation 55), and therefore the perturber needs to be closer in projection to have the same effect than as if it were at the same redshift as the lens. In the foreground, the curve flares out, implying that foreground perturbors can be farther away in projection to have the same effect if they were at the lens redshift. This occurs because the Einstein radius of a point mass increases as redshift decreases, yielding stronger lensing effects. The curve matches the shape of the color contours, indicating that our theoretical expression from the lens potential captures the redshift dependence of LOS effects in the recovered parameters, like the Hubble constant. Thus, we can use the flexion shift to compare effects of LOS galaxies even if they are at different redshifts.

If we truncate the expansion at tidal terms and return to vector/tensor notation, we can write

$$\mathbf{x}_s = (\mathbf{I} - \Gamma)\mathbf{x}_1 - (\mathbf{I} - \beta\Gamma)\alpha_g(\mathbf{x}_1) + \mathcal{O}(\mathbf{x}_1^3). \quad (48)$$

This equation looks very similar to the single plane lens equation, but with some multiplicative factors. In fact, if we multiply both sides by $(\mathbf{I} - \beta\Gamma)^{-1}$ from the left, introduce a scaled source coordinate $\mathbf{u}_{\text{eff}} \equiv (\mathbf{I} - \beta\Gamma)^{-1}\mathbf{x}_s$, and define an effective shear by

$$\mathbf{I} - \Gamma_{\text{eff}} \equiv (\mathbf{I} - \beta\Gamma)^{-1}(\mathbf{I} - \Gamma), \quad (49)$$

then we can rewrite equation 48 as

$$\mathbf{u}_{\text{eff}} = (\mathbf{I} - \Gamma_{\text{eff}})\mathbf{x}_1 - \alpha_g(\mathbf{x}_1), \quad (50)$$

which is equivalent to the single plane lens equation with an external shear in the lens plane (see also Schneider 1997). We are allowed to rescale the source plane because the position of the source is unobservable and we typically only measure the ratio of the fluxes of different images, which is insensitive to the absolute intrinsic luminosity of the source.⁷

⁷ The freedom to rescale the source plane vanishes if the intrinsic luminosity of the source is known. Thus, lensed standard candles, such as type Ia supernovae (Kelly et al. 2016; Rodney et al. 2016; Kelly et al. 2015; Patel et al. 2014; Kolatt & Bartelmann 1998), offer a novel way to break the mass sheet degeneracy.

In the tidal approximation, we again expect errors in the recovered lens parameters to be associated with the largest terms in the lens potential that have been neglected: the third order flexion terms. We therefore use the terms involving \mathcal{F} to quantify the lens potential perturbations.⁸ We approximate the main lens galaxy as a singular isothermal sphere (SIS) with deflection

$$\alpha_g = R_E \hat{\mathbf{r}}. \quad (51)$$

Then we can write the lens equation 48 in the tidal approximation as

$$\mathbf{x}_s = \mathbf{x}_1 - R_E \hat{\mathbf{r}} - \mathbf{\Gamma}(\mathbf{x}_1 - \beta R_E \hat{\mathbf{r}}). \quad (52)$$

For comparison, the lens equation with flexion terms is

$$\begin{aligned} \mathbf{x}_s = & \mathbf{x}'_1 - R_E \hat{\mathbf{r}} - \mathbf{\Gamma}(\mathbf{x}'_1 - \beta R_E \hat{\mathbf{r}}) \\ & - \frac{1}{2}(\mathbf{x}'_1 - \beta R_E \hat{\mathbf{r}})\mathcal{F}(\mathbf{x}'_1 - \beta R_E \hat{\mathbf{r}}). \end{aligned} \quad (53)$$

where we recognize that the solutions \mathbf{x}'_1 of this equation may differ from the solutions \mathbf{x}_1 of equation 52 because of the flexion effects. If we assume that $\mathbf{\Gamma} \ll 1$ (which applies if this is truly a perturbing galaxy and not a main lens galaxy), and we subtract equations 53 and 52, we can define the image shift caused by third order terms to be $\Delta_3 \mathbf{x} \equiv \mathbf{x}'_1 - \mathbf{x}_1$, and thus find

$$\Delta_3 \mathbf{x} = \frac{1}{2}(\mathbf{x}'_1 - \beta R_E \hat{\mathbf{r}})\mathcal{F}(\mathbf{x}'_1 - \beta R_E \hat{\mathbf{r}}). \quad (54)$$

Now, if we define the perturber to be colinear with the image position ($\theta = 0$)⁹ and assume that the positions of the multiple images are $|\mathbf{x}'_1| \approx |\mathbf{x}_1| \approx R_E$, we can write the magnitude of the flexion shift as

$$\Delta_3 x = \frac{R_E^2 R_p^2}{r^3} (1 - \beta)^2. \quad (55)$$

We call this quantity, $\Delta_3 x$, the “flexion shift” because it characterizes the perturbations to the image positions caused by the third-order (“flexion”) terms from a perturber. This quantity has units of arcseconds and it gives us a way to quantify and compare the ENV/LOS contributions from different perturbing galaxies.

We note that in this derivation, we have used the expression for \mathcal{F} for a point mass, which is reasonable when the perturber is projected far from the images. Strictly speaking, r here is the unlensed distance in the redshift plane of the perturber. To convert to offset as observed on the sky r' , we use the lens equation to yield $r' = r - \beta R_E$. As long as the perturbing galaxy is many Einstein radii away from the lens, we can put $r' \approx r$. (If that were not the case, the perturber would need to be treated as a main plane anyway.)

Lens Potential Perturbations from a Foreground Galaxy

We now turn our attention to a perturber in the foreground of the lens. In this case $\alpha_1 = \alpha_p$ and $\alpha_2 = \alpha_g$,

⁸ Recall from Section 3.1.1 that the flexion terms scale with mass and projected offset as M/r^3 ; now we generalize now to include redshift.

⁹ If we consider all possible angles, the root mean square value has an extra factor of $\sqrt{2}$ that does not affect the scalings.

so the lens equation has the form

$$\mathbf{x}_s = \mathbf{x}_1 - \alpha_p(\mathbf{x}_1) - \alpha_g(\mathbf{x}_1 - \beta \alpha_p(\mathbf{x}_1)). \quad (56)$$

Taylor expanding α_p yields

$$\begin{aligned} x_s^i = & x_1^i - \alpha_p^i(0) - \Gamma^{ij} x_1^j - \frac{1}{2} \mathcal{F}^{ijk} x_1^j x_1^k + \dots \\ & - \alpha_g^i \left(x_1^i - \beta \alpha_p^i(0) - \beta \Gamma^{ij} x_1^j - \frac{1}{2} \mathcal{F}^{ijk} x_1^j x_1^k + \dots \right). \end{aligned} \quad (57)$$

Now truncating at tidal terms yields

$$\mathbf{x}_s = (\mathbf{I} - \mathbf{\Gamma})\mathbf{x}_1 - \alpha_g((\mathbf{I} - \beta \mathbf{\Gamma})\mathbf{x}_1 + \mathcal{O}(\mathbf{x}_1^3)) + \mathcal{O}(\mathbf{x}_1^3). \quad (58)$$

This looks similar to equation 48, but there is a key difference: instead of just having a multiplicative effect on the source position like the background perturber, the deflection from the foreground perturber enters the lens equation *inside the argument* of the deflection of the main galaxy. A foreground perturber has complicated effects because it creates a difference between the coordinates we see on the sky and the coordinates in the plane of the main lens galaxy. There is no way to define an effective shear that fully captures the non-linear effects of a foreground perturber (see also M14).

In principle, one can define a scaled coordinate based on the argument of the deflection to transform this equation to look like the standard lens equation (e.g., Schneider 1997; Keeton 2003). This requires care, however, because the new quantities do not correspond to the *observed* image positions that are typically used in lens modeling. An alternative is to rescale the mass of the main lens galaxy to account for the change in the argument of the deflection (Schneider 1997), but this also requires care because the mass predicted by the lens models is no longer the true physical mass of the lens.

To examine the differences between the foreground and background perturbers, we simulated a $10^{12} M_\odot$ perturbing galaxy at a variety of projected offsets and redshifts. To emphasize the contribution from non-linear effects, in this set of simulations we generate the mock data using the tidal approximation (i.e., there are no higher order terms). This allows us to focus entirely on non-linear effects. Figure 3 shows that an external shear can account for a background perturber (to within numerical precision), as discussed above. By contrast, shear *cannot* mimic the non-linear effects of a foreground LOS galaxy.

We now seek an expression for the flexion shift $\Delta_3 x$ caused by a foreground perturber to compare to the background case. Following an analysis similar to that above, and again using an SIS main lens and a point mass perturber, we can write the lens equation in the tidal approximation as

$$\mathbf{x}_s = \mathbf{x}_1 - R_E \hat{\mathbf{r}} - \mathbf{\Gamma}(\mathbf{x}_1 - \beta R_E \hat{\mathbf{r}}) \quad (59)$$

while the lens equation with flexion terms is

$$\mathbf{x}_s = \mathbf{x}'_1 - R_E \hat{\mathbf{r}} - \mathbf{\Gamma}(\mathbf{x}_1 - \beta R_E \hat{\mathbf{r}}) - \frac{1}{2} \mathbf{x}_1 \mathcal{F} \mathbf{x}_1. \quad (60)$$

Subtracting these equations and rearranging yields

$$\Delta_3 \mathbf{x} = \frac{1}{2} \mathbf{x}_1 \mathcal{F} \mathbf{x}_1. \quad (61)$$

In this case, the multiple images form at $\mathbf{x}_2 \approx R_E \hat{\mathbf{r}}$, implying that $R_E \approx \mathbf{x}_1 - \beta \mathbf{\Gamma} \mathbf{x}_1$. If we assume that $\mathbf{\Gamma} \ll 1$

(as above), we can write $\mathbf{x}_1 \approx R_E \hat{\mathbf{r}}$, giving

$$\Delta_3 x = \frac{R_E^2 R_p^2}{r^3}. \quad (62)$$

This equation is similar to equation 55, but without any β factors. Therefore, not only can foreground perturbers not be mimicked by an external shear because of non-linear effects, but also the effects of background perturbers are downweighted compared to foreground perturbers.

Scatter in Lens Model Parameters Tracks Lens Potential Perturbations

We now use simulations to test whether our expression for the deviation in the lens potential, the flexion shift, $\Delta_3 x$ (eqs. 55 and 62), provides a useful way to characterize the importance of ENV/LOS perturbers and their contribution to the recovered lens parameters and Hubble constant. For these simulations, we fix the perturber mass to be $10^{12} M_\odot$ but vary its projected offset and redshift. In these simulations, we treat the perturber as a point mass, which is not a valid assumption when the projected offset or redshift is small, but is reasonable in the regime where we anticipate using the tidal approximation. Figure 4 shows the scatter in the recovered value of the Hubble constant (the median value is not shown because it matches the input value; there is no bias). Generally speaking, the scatter increases as the offset decreases and higher order terms become more significant. In the background, the perturber’s effects become weaker as the redshift offset increases because of the β factors in equation 55; while in the foreground, the perturber’s effects become stronger as $z \rightarrow 0$ because its angular Einstein radius increases. Both of these scalings are well described by our functional form for $\Delta_3 x$, which indicates that we have successfully defined a quantity to characterize the contribution of perturbing galaxies.

Many lens systems have significant contributions from in their local environment, (e.g., HE0435–1223 has a neighbor, Kochanek 2006; while MG0414+0534 (Tonry & Kochanek 1999), RXJ1131–1231 (hereafter RXJ1131 Sluse et al. 2003), and B2114+022 (King et al. 1999) all have satellites that are presumably close enough to matter), and while some models do treat nearby perturbers exactly (e.g., Kochanek 2006; Fadely & Keeton 2012), typically the decision whether to include a neighbor galaxy in a model is *ad hoc*. The flexion shift, $\Delta_3 x$, provides a quantitative criterion to compare the importance of LOS galaxies even if they are at different redshifts. Any given perturber can be treated with the tidal approximation if its $\Delta_3 x$ value is small, but it must be treated explicitly if its $\Delta_3 x$ value is large (see Sluse et al. 2016 for an application of the flexion shift to observed LOS galaxy catalogs).

3.2. Characterizing Realistic Beams

Now that we have a quantity to assess the importance of individual galaxies, we can apply this criterion to realistic mass models that have hundreds of ENV/LOS galaxies. This allows us to characterize which ENV/LOS galaxies produce the strong effects on the inferred Hubble constant. The perturbing ENV/LOS galaxies that produce strongest ENV/LOS effects can then be treated

with more care (i.e., as “main” planes in our hybrid framework). The flexion shift criterion also gives us the ability to compare the mass models for different lens fields, in principle allowing us to select fields that do not have strong contributions as a “gold sample” of lenses for future surveys like LSST and Euclid.

To gain intuition about how the flexion shift applies to realistic beams, we first calculate the flexion shift for all of the perturbing ENV/LOS galaxies in the lens system RXJ1131, which has a source at redshift $z_{\text{src}} = 0.658$ and a lens in a group at redshift $z_{\text{lens}} = 0.2936$. We adopt an Einstein radius of $R_E = 1.5''$, close to the measured value (Suyu et al. 2013) and also close to the peak of the observed distribution (Sonnenfeld et al. 2013), and an ellipticity of $e = 0.3$. Figure 5 illustrates $\Delta_3 x$ values for galaxies in this field. While $\Delta_3 x$ generally increases toward the lens, there is no single radial cut that can be used to determine the importance of a perturber because perturber mass and redshift must also be taken into account. By following contours of constant $\Delta_3 x$, as illustrated in right-hand panel, we can determine the effective radial offset for the perturber if it were in the main lens plane. Since background perturbers are downweighted by $(1 - \beta)^2$, they must have a smaller projected offset to have the same effect as if they were in the main lens plane. Perturbing galaxies in the foreground are slightly upweighted because the Einstein radius of a point mass perturber increases as we move it to lower redshift, which means that foreground galaxies can have a larger projected offset to have the same effect as if they were at the main lens redshift.

Figure 6 shows a histogram of all the $\Delta_3 x$ values for this field. The flexion shifts range from $\sim 5 \times 10^{-3}$ arcseconds down to $< 10^{-10}$ arcseconds, with background perturbers generally having smaller values than foreground perturbers or group members. In the next section we consider the quantitative threshold for $\Delta_3 x$ needed to achieve desired accuracy and precision in lens model results.

While external convergence does not capture all LOS effects (it omits non-linearities in eq. 1), it is a useful quantity for comparing to previous work. Suyu et al. (2013) (see also Rusu et al. 2016) used the Millennium simulation along with galaxy number counts around RXJ1131 to estimate a prior probability distribution for the external convergence used to constrain the Hubble constant; their distribution is shown in red in Figure 7. We compute the effective convergence (as defined in M14 and references therein) directly from our mass models for RXJ1131. We consider many realizations that account for observational uncertainties and scatter in the relations used to assign masses to galaxies described in Section 2.2. This yields a distribution of “direct” convergence calculations that is shown by the blue histogram in Figure 7. We find that the peaks of the distributions are roughly consistent. However, our “direct” convergence calculations produce a tighter distribution by a factor of ~ 4 . This is likely because we build mass models based on the observed beam for each individual lens rather than using statistical results from N-body simulations. The narrower distribution of external convergences translates into a stronger prior on measuring the Hubble constant.

We can also characterize values of the flexion shift for a wider range of fields. Figure 8 shows $\Delta_3 x$ values for

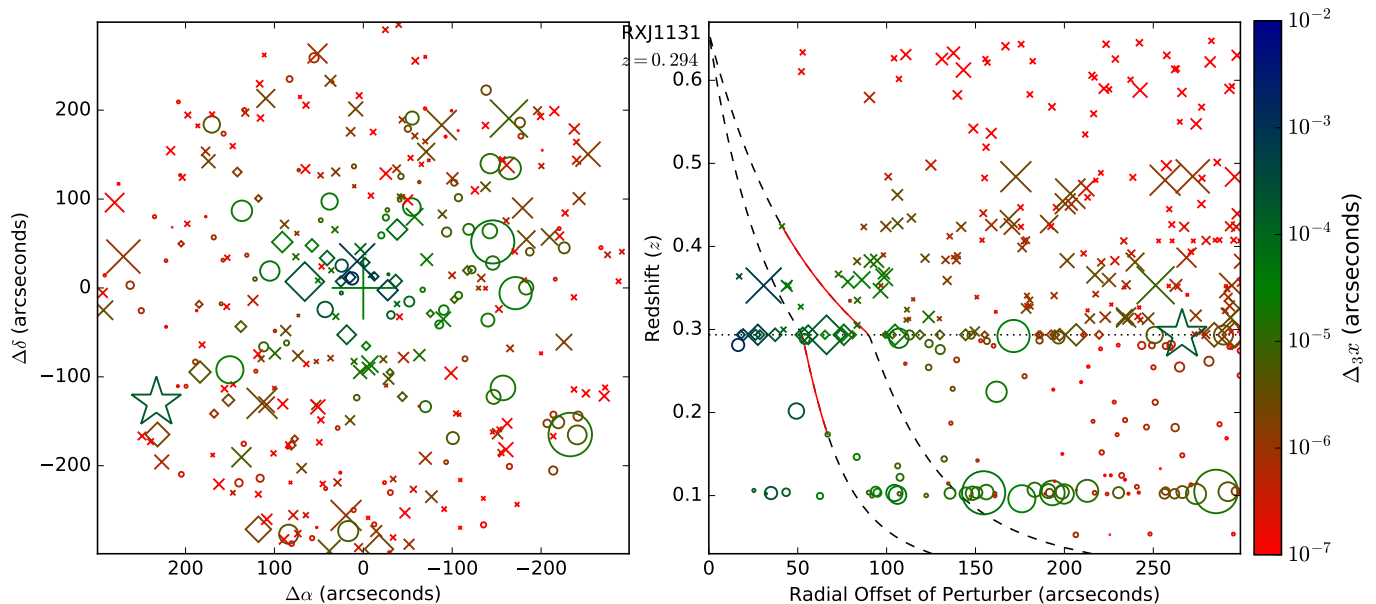


FIG. 5.— Illustration of the flexion shift for galaxies in the field of RXJ1131. In the left panel, each galaxy is shown at its position on the sky. The area of each point is proportional to the mass of the galaxy. X's and O's mark galaxies behind and in front of the main lens galaxy, respectively, while diamonds mark members in the group around the main lens, and the star indicates the location of the common group halo. The color of the points represents the strength of the lensing effects measured by the flexion shift Δ_{3x} . The right panel shows the same galaxies plotted in the r - z projection, similar to Figures 3 and 4. The main lens redshift is indicated by the dotted line. Two Δ_{3x} contours have been shown to guide the eye. The red section of the contour illustrates how to map a perturbing galaxy to its effective offset as if it were in the main lens plane. There is no simple radial cut that selects the most important LOS galaxies; a more complicated quantity like the flexion shift, Δ_{3x} , is necessary.

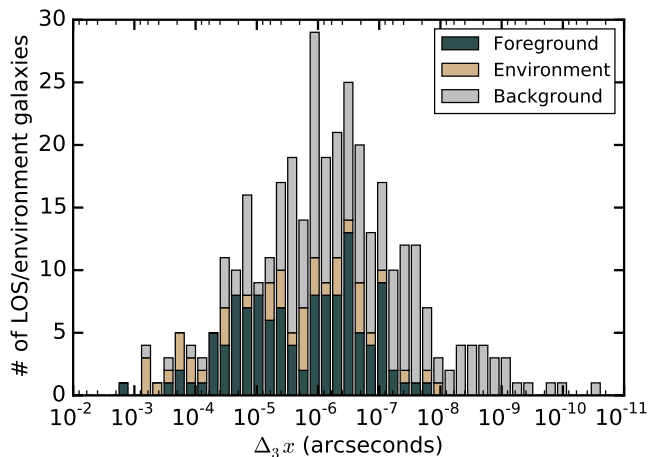


FIG. 6.— Histogram of flexion shifts for galaxies within $5'$ of RXJ1131. The dark (light) gray bins represent perturbing galaxies in the foreground (background) of the main lens galaxy, while the tan bins in between correspond to group members in the immediate environment of the main lens. The range of flexion shifts for this field spans ~ 8 orders of magnitude. We have reversed the x-axis for comparison to subsequent figures. Perturbers with large a flexion shift, and therefore a large contribution to the lens potential, are on the left. For reference, when generating our χ^2 values, we have assumed a positional uncertainty of 3×10^{-3} arcseconds. Few individual perturbing galaxies have that large of flexion shift, but the cumulative flexion is easily larger than our assumed measurement uncertainties. As anticipated, foreground perturbers and group members tend to have larger Δ_{3x} values than background perturbers.

nine other multiply-imaged QSO fields from Wong et al. (2011). The fields are all complex; there is not a simple radial cut that divides the sample into exact and tidal perturbers. There is striking diversity in the strength of

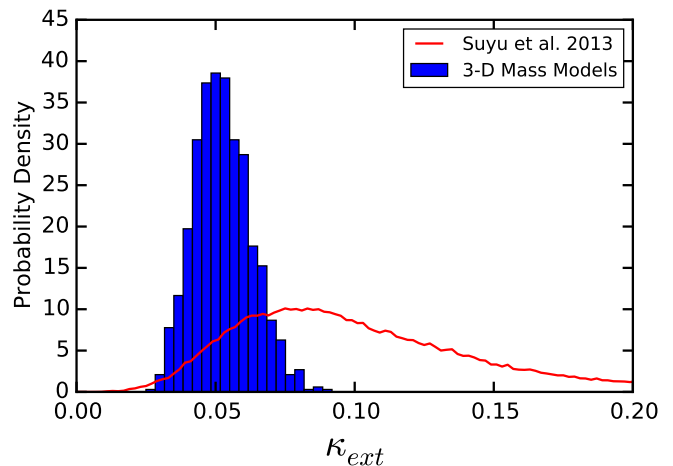


FIG. 7.— Probability distribution for the external convergence in the RXJ1131 field. The blue histogram shows the effective convergence computed directly from our 3-D mass models, accounting for measurement uncertainties and scatter in the relations used to assign masses to galaxies. The red curve shows the external convergence distribution derived using galaxy number counts from the Millennium simulation by Suyu et al. (2013). The peaks of the distributions are roughly consistent, but our models favor slightly lower values of external convergence. Because we build mass models for the specific, observed field, we obtain a ~ 4 times narrower distribution that effectively translates into a stronger prior on the Hubble constant.

external effects. There is no single number of galaxies that is guaranteed to be the “right” number to include in lens models. Each field must be considered on an individual basis.

While each field is unique, we do expect that some fields will have stronger ENV/LOS contributions than

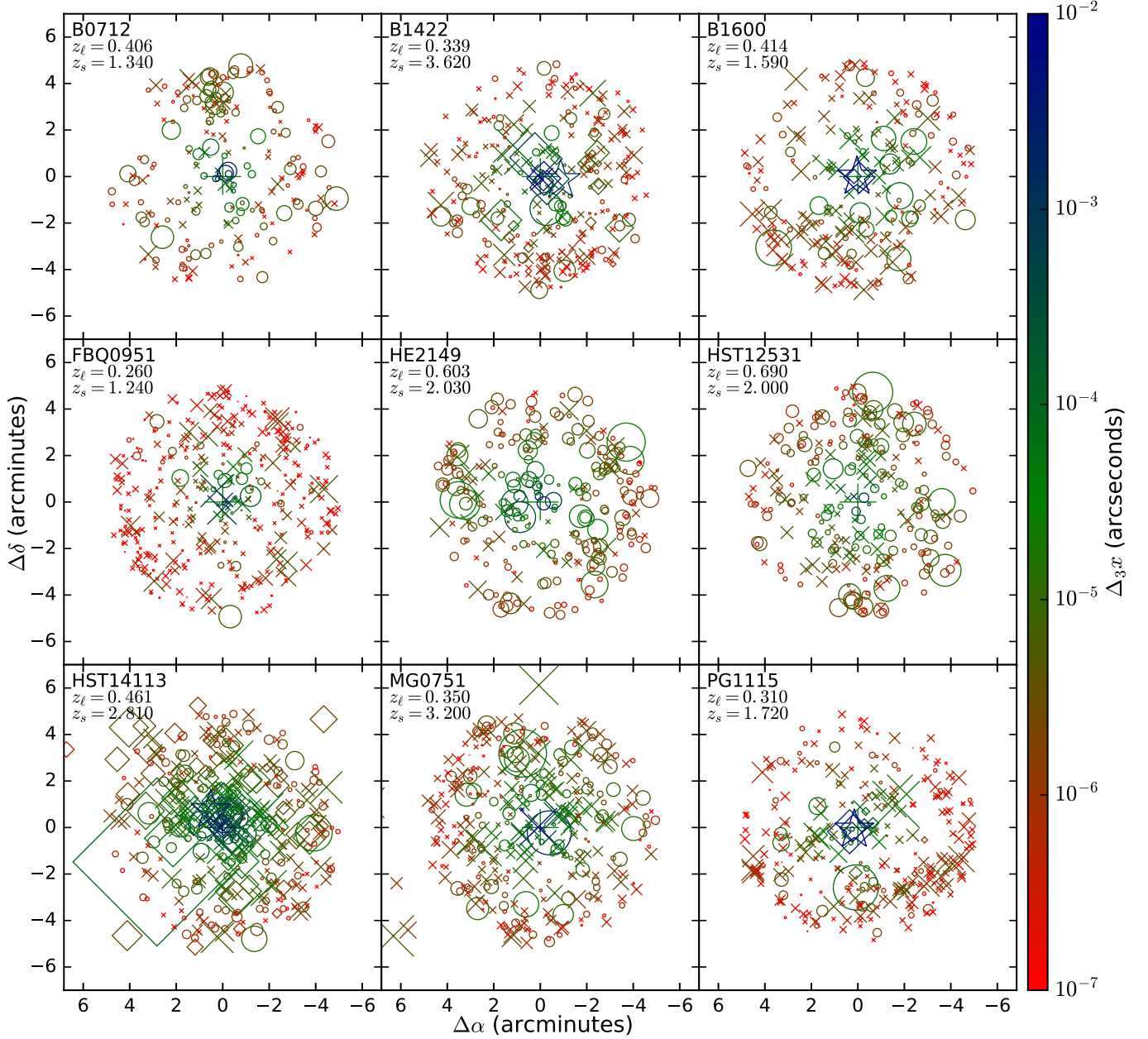


FIG. 8.— Illustration of the flexion shifts, Δ_3x , for perturbing galaxies in nine lens fields (similar to the left panel in Fig. 5). Each galaxy in the mass model is shown in projection on the sky. The area of each point is proportional to the mass of the galaxy. X's are behind the main lens galaxy, while O's are in the foreground. Diamonds represent group members, while stars indicate the locations of group halos. The color of the points represents the strength of the higher order terms characterized by the flexion shift, Δ_3x . There is dramatic variation in the ENV/LOS strengths across the different fields, implying that each field needs to be treated on an individual basis. For all of the fields shown here, there is no simple radial cut that would capture all of the important galaxies: a more complicated quantity like the flexion shift is necessary.

others. One important factor may be whether the main lens galaxy is part of a group. To make a simple quantitative comparison, we add the flexion shifts (Δ_3x values) in a given beam. The higher order effects due to flexion terms combine in a more complicated way (as the flexion is a tensor), but the simple sum gives an a good estimate of the importance of the ENV/LOS of the lens. If there are many nearby, massive perturbers then the sum will be large, implying a strong ENV/LOS contribution (such as HST14113 in Figure 8). Conversely, in a sparser field

the sum will be smaller (such as B0712). Figure 9 shows a histogram of the sums for our set of 23 lens field. While the sample size is not large, it is apparent that lenses that are members of a group tend to have stronger ENV/LOS contributions than those that are not.

3.3. Fitting Lens Observables

Now that we have characterized our mass models, we consider whether different methods for treating ENV/LOS effects can reliably recover lens galaxy parameters and the Hubble constant. We generate mock lens

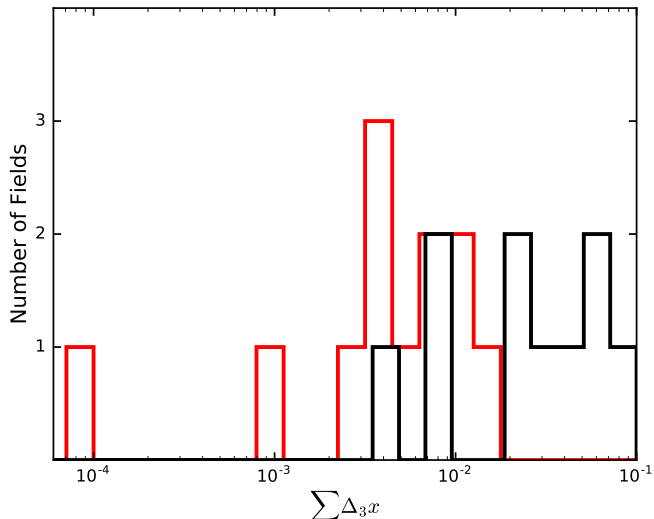


FIG. 9.— Distribution of ENV/LOS strengths for our sample of 3-D mass models. Beams for which the lens galaxy is a member of a group are shown in black, while non-group lenses are shown in red. Lenses in groups typically have stronger ENV/LOS contributions than those that are not group members. However, there is substantial overlap so each field needs to be treated individually.

observables using the full 3-D mass models in the simulations described in Section 2.4; throughout this analysis we combine results for 300 mock quad lenses with different source positions and azimuthal angles for the main lens galaxy. We then fit the mock lensing data using the three types of models discussed in Section 2.5. We first test whether simple Lens-Only and Lens+Shear models can produce robust constraints on the Hubble constant. We then examine the 3-D Lens models to ascertain how well we do when we apply the tidal approximation to ENV/LOS galaxies. We also examine the parameter degeneracies in the 3-D Lens models.

We begin with Lens-Only models that do not explicitly account for any ENV/LOS effects. Figure 10 shows the recovered lens model parameters including the mass normalization (which is related to the Einstein radius; see eq. 35), ellipticity, power law index, and Hubble constant for the RXJ1131 field. The posterior distributions of the recovered lens parameters are broad and often multimodal. All of the parameters show significant bias, and are not able to reproduce the input parameters. These results reiterate (not surprisingly) that ENV/LOS effects cannot be ignored.

Next we consider the Lens+Shear models. These models have two extra free parameters that can be used to capture some ENV/LOS effects. The results of the fits are shown in Figure 11. The posterior distributions of the recovered lens parameters are considerably tighter than for Lens-Only models. However, these models still have biases that are most dramatic in the Einstein radius and Hubble constant. The bias is associated with external convergence: the total mass in the simulated lenses includes contributions from the environment and LOS; if we assign all of that mass to the main lens galaxy (as in Lens+Shear and Lens-Only models), we overestimate the galaxy’s mass and Einstein radius. In this approach, external convergence must be included through post-processing (Collett et al. 2013; Suyu et al. 2010). The biases that we find are larger than those found in

Jaroszynski & Skowron (2016) for random sight lines in the Millenium Simulation. However, they point out that their sight lines, on average, are not as dense as that of RXJ1131. This is in agreement with Collett & Cunningham (2016) who find that lenses tend to be in (slightly) overdense lines-of-sight.

Wong et al. (2011) found that fitted shear parameters from the Lens+Shear models did not always match the shears computed directly from 3-D mass models in real lenses, including RXJ1131. We reexamine this result using our controlled simulations. Figure 12 compares the distributed of fitted shear parameters with the known value of the effective shear for this ENV/LOS. The fitted shear parameters disagree with the true effective shear at the $\sim 2\sigma$ level, broadly consistent with the results found in Wong et al. (2011). While a deeper analysis of the difference is beyond the scope of this work, the results in Figure 12 suggest that the offsets found by Wong et al. (2011) are associated with the lens modeling procedure; they may arise because the fitted shear is attempting to account for non-linear effects due to foreground perturbing galaxies.

Lastly, we consider our 3-D lens models. We initially treat all perturbing galaxies in the tidal approximation. This allows us to disentangle the non-linear effects, which are included in our models because we use the same mass model that was used to generate the lensing observables, and higher-order terms, which are included when generating the lens observables (see Section 2.4), but not in the fits. Figure 13 shows posterior distributions of the recovered lens parameters and Hubble constant. There is no bias in any of the parameters recovered from these models. Our 3-D Lens models avoid the biases seen in other models because, rather than treating shear and convergence separately, we build a physical model and then extract both shear and convergence self-consistently. We favor this approach because shear and convergence arise from the same underlying mass distribution and are therefore not independent.

The main difference between the Lens+Shear and the 3-D Lens models is the inclusion of non-linear effects in the 3-D Lens models. Therefore, the broader posterior distributions that we observe for the Lens+Shear models must be driven by non-linear effects.

However, there are strong correlations among the recovered ellipticity, power law index, and Hubble constant for our 3-D lens models. These correlations involving the Hubble constant are associated with the “lens profile degeneracy”: there can be a range of density profiles that all yield a good fit but lead to different values for h . Kochanek (2002) shows that to lowest order the degeneracy is characterized by the scaling $h \propto 1 - \langle \kappa \rangle$ where $\langle \kappa \rangle$ is the average convergence from the main lens in the annulus between the image positions. If we assume that the image annulus is narrow and centered on the Einstein radius, we can evaluate the angle average as

$$\langle \kappa \rangle = \frac{1}{2\pi} \int_0^{2\pi} \kappa(R_E, \theta) d\theta. \quad (63)$$

With our power law model (eq. 35), we can evaluate the integral by making a Taylor series expansion in the ellip-

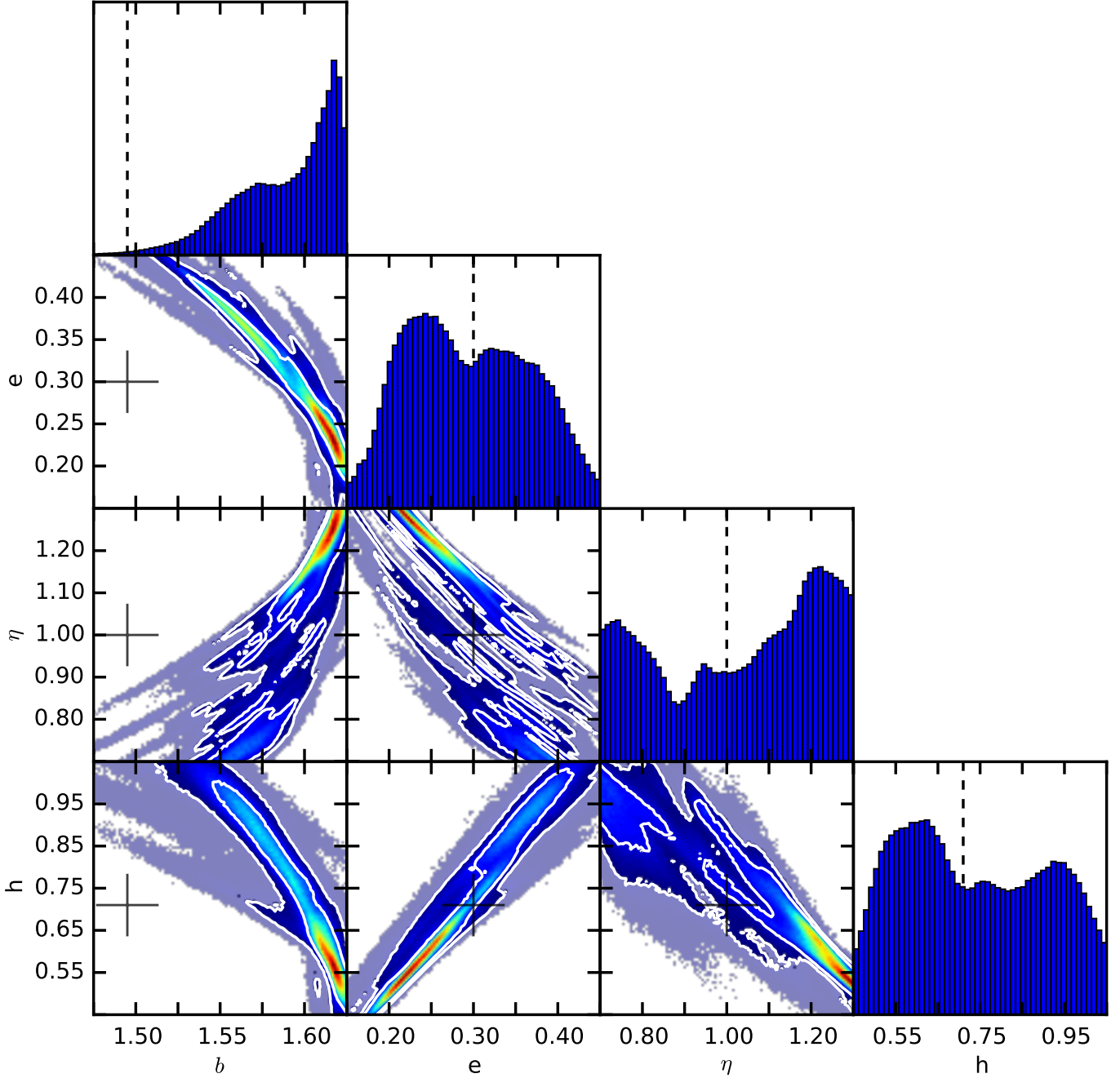


FIG. 10.— Posterior distributions for model parameters from Lens-Only models fit to mock quad lenses in the field of RXJ1131. The true input values are marked with crosses in the contour plots and with a dashed line in the histograms. We combine MCMC results for 300 mock lenses with different source positions and azimuthal angles for the main lens galaxy. Contours containing 68% and 95% of the MCMC trials are shown in white. Many of the parameters show significant biases and the distributions are often degenerate and multi-modal. These results suggest that ENV/LOS cannot be ignored. The recovered values of the Hubble constant have a scatter of $\sim 20\%$, more than an order of magnitude larger than our goal for “Precision Lensing” of $< 1\%$. The plotting ranges for corresponding parameters are the same as Figures 11 and 13 to facilitate direct comparison.

ticity and using

$$\int_0^{2\pi} \frac{d\theta}{(1 - 2e \cos^2(\theta) + e^2 \cos^4(\theta))^{1-\eta/2}} \approx 1 + e - \frac{e\eta}{2}. \quad (64)$$

For typical parameter values of interest, we find numerically that this approximation is good to $\sim 1\%$. Then we can approximate the average convergence at the Einstein

radius as

$$\langle \kappa \rangle \approx \frac{1}{2} \eta (1 - e)^{2-\eta} \left(1 + e - \frac{e\eta}{2} \right). \quad (65)$$

We test this analysis by plotting $h/(1 - \langle \kappa \rangle)$ in Figure 14. Most of the correlation of h with η and e has been accounted for. Remaining correlations are likely due to higher order terms in the expression for h as a function

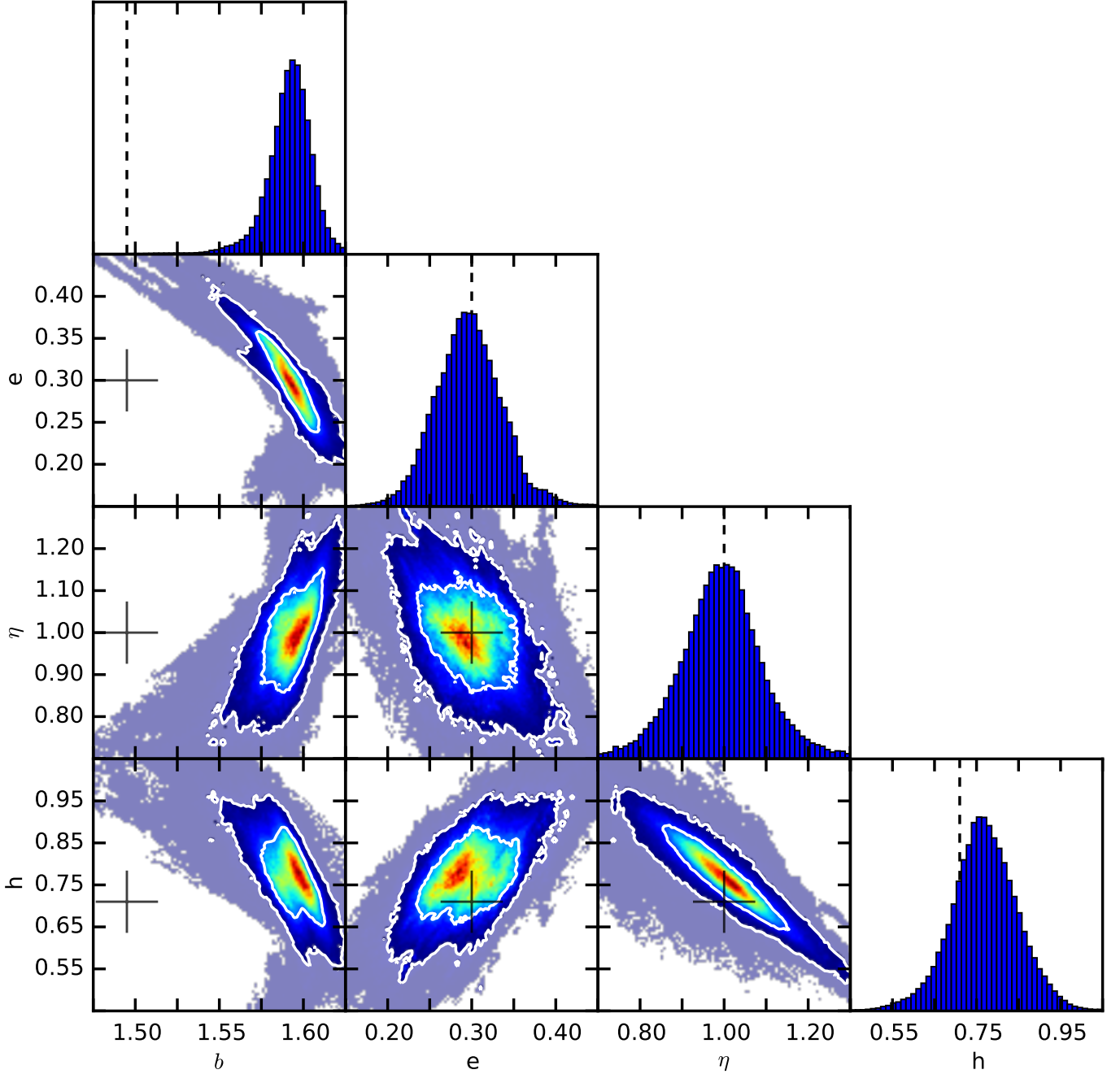


FIG. 11.— Similar to Figure 10, but for Lens+Shear models. Including an external shear dramatically improves the fits over ignoring the ENV/LOS, but the distribution of recovered parameters are non-Gaussian and may have multiple families of solutions. Many of the parameters show biases; most notably, the Einstein radius parameter (b) is strongly biased because Lens+Shear models do not explicitly include external convergence (such models must add external convergence in post-processing).

of profile parameters (Kochanek 2002) and the Taylor series expansion in equation 64.

The fact that these correlations are seen in our 3-D lens models suggests that the ENV/LOS does not break the lens profile degeneracy. This is consistent with the idea that ENV/LOS effects are produced by structure on much larger spatial scales than is relevant for the lens profile degeneracy (Xu et al. 2016; Schneider & Sluse 2013). The lens profile must be broken instead with kinematic information (e.g., Suyu et al. 2013) or with extended source modeling (Suyu 2012).

Our next goal is to determine when it is valid to use the tidal approximation. We want to find a sweet spot that speeds up the modeling without introducing systematic uncertainties that are larger than the measurement uncertainties. Figure 15 shows results from 3-D Lens model fits with different cuts on the flexion shift, $\Delta_3 x$. Each point on the horizontal axis corresponds to a different *threshold* for $\Delta_3 x$: galaxies with larger flexion shifts are treated exactly, while galaxies with smaller values are treated with the tidal approximation. (Note that the threshold *decreases* from left to right in the plot; at

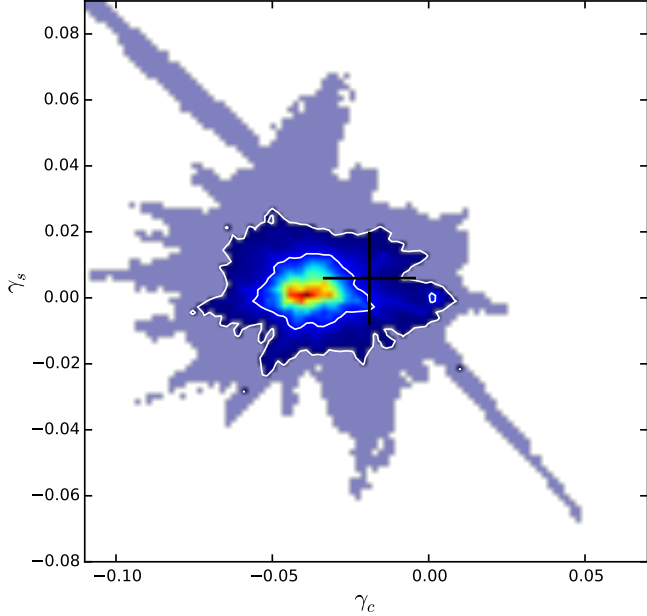


FIG. 12.— Marginalized posterior distribution of fitted external shear values from the Lens+Shear models. 68% and 95% contours are shown in white. The effective shear calculated from the input 3-D mass model is shown with an orange cross. The peak of the distribution from the MCMC models is $\sim 2\sigma$ away from the truth, similar to what was found by Wong et al. (2011). The offset appears to be inherent to the lens modelling process, possibly due to non-linear effects from foreground perturbing galaxies.

a given point, galaxies to the left are exact perturbers while galaxies to the right are tidal perturbers.) The error bars mark the 16th and 84th percentiles of the posterior distribution for the lens model parameters. Our modeling uses MCMC methods to obtain posterior distributions, so the error bars correspond to marginalized single-parameter posteriors.

The first result from Figure 15 is that 3-D Lens models do not show any bias in key parameters, even if all of the ENV/LOS galaxies are treated in the tidal approximation, at least for the RXJ1131 field. The scatter in model parameters decreases a little as the Δ_3x threshold is reduced and the strongest perturbers are incorporated into the model explicitly. This will be especially important if some extra constraint on the radial profile, e.g., kinematics, is available. We believe that it is worthwhile to set a conservative threshold of $\Delta_3x \sim 10^{-4}$ arcseconds (i.e., 1.5 dex smaller than our astrometric uncertainties), which amounts to including the strongest 10–15 perturbers explicitly as main planes. The computational complexity of the full multi-plane lens equation scales as $\sim N^2$, the reduction from ~ 300 galaxies to treating all but 10–15 using the tidal approximation produces an improvement in the performance by a factor of 400–900.

The Lens-Only and Lens+Shear models do show a bias, as mentioned above due to external convergence. Introducing a few of the strongest perturbers (i.e., moving to the right in the figure) means that we begin to explicitly include the most important ENV/LOS galaxies. Thus as we move to the right in the figure, the need for an “external” convergence is no longer necessary because it already included in the models, which in turn reduces the bias in the recovered parameters. In principle, there is a point at which the sources of convergence are properly

accounted for and the bias is removed. In practice, that point is not easy to determine *a priori*. Going to very small values of Δ_3x (i.e., far to the right in the figure) actually leads to a negative bias. This is because Lens-Only and Lens+Shear models do not include the void correction discussed in Section 2.3. This implies that the void correction is approximately 2% in the Einstein radius for the RXJ1131 field.

So far, we have presented results for a single lens field. In order to understand whether our conclusions are general, we need to examine how ENV/LOS effects vary from one lensing field to another. One simple test is to compare models that use the same fiducial main lens galaxy ($R_E = 1.5''$ and $e = 0.3$) but different fields. Figure 16 shows results from simulations using the field of B0712+472 (hereafter B0712; Jackson et al. 1998). The trends are similar to what we saw for RXJ1131 in Figure 15, but the deviations are somewhat smaller here because the ENV/LOS effects are not as strong in this field (as we saw in Section 3.2). The correction due to voids is larger for B0712 at $\sim 4\%$. Therefore, the qualitative trends of our results appear to be robust, but the quantitative differences indicate that each lens field needs to be treated on an individual basis.

In this analysis we have placed the same lens galaxy in different 3-D mass models to see how the ENV/LOS effects vary. The complementary step is to fix the 3-D model and vary the main lens galaxy to understand how that changes the sensitivity to ENV/LOS effects.

3.4. What Lens Properties Lead to the Strongest Constraints on Cosmology?

In our final set of simulations, we choose one ENV/LOS mass model (for the RXJ1131 field) and test how the constraints on the Hubble constant change when we vary the parameters of the main lens galaxy and the position of the source behind the lens.

One important parameter of the main lens galaxy is the ellipticity. Figure 17 compares results from simulations with different values of e . As the ellipticity increases, the scatter in the recovered Hubble constant decreases. The image configurations of very asymmetric lens galaxies can only be produced for a smaller range of lens models, leading to stronger constraints on both the ellipticity and the power law index. Since those parameters are correlated with the Hubble constant through the lens profile degeneracy, reducing the scatter of fitted e and η values leads to a narrower range for h as well.

A second key parameter of the main lens galaxy is the Einstein radius. Figure 18 shows model results for different values of the Einstein radius parameter b . As the Einstein radius increases, the constraints on the Hubble constant get stronger. Lenses with large Einstein radii produce stronger constraints on the ellipticity limiting the lens profile degeneracy, yielding a tighter distribution on the Hubble constant.

One possible approach to using gravitational lensing for precision cosmology is to search for one or a few “golden lenses” to use to measure cosmological parameters. One consideration when choosing these lenses needs to be the environment. The results in Figures 17 and 18 show that asymmetric image configurations, e.g., those produced by a highly elliptical main lens galaxies, are less sensitive to the lens profile degeneracy leading to

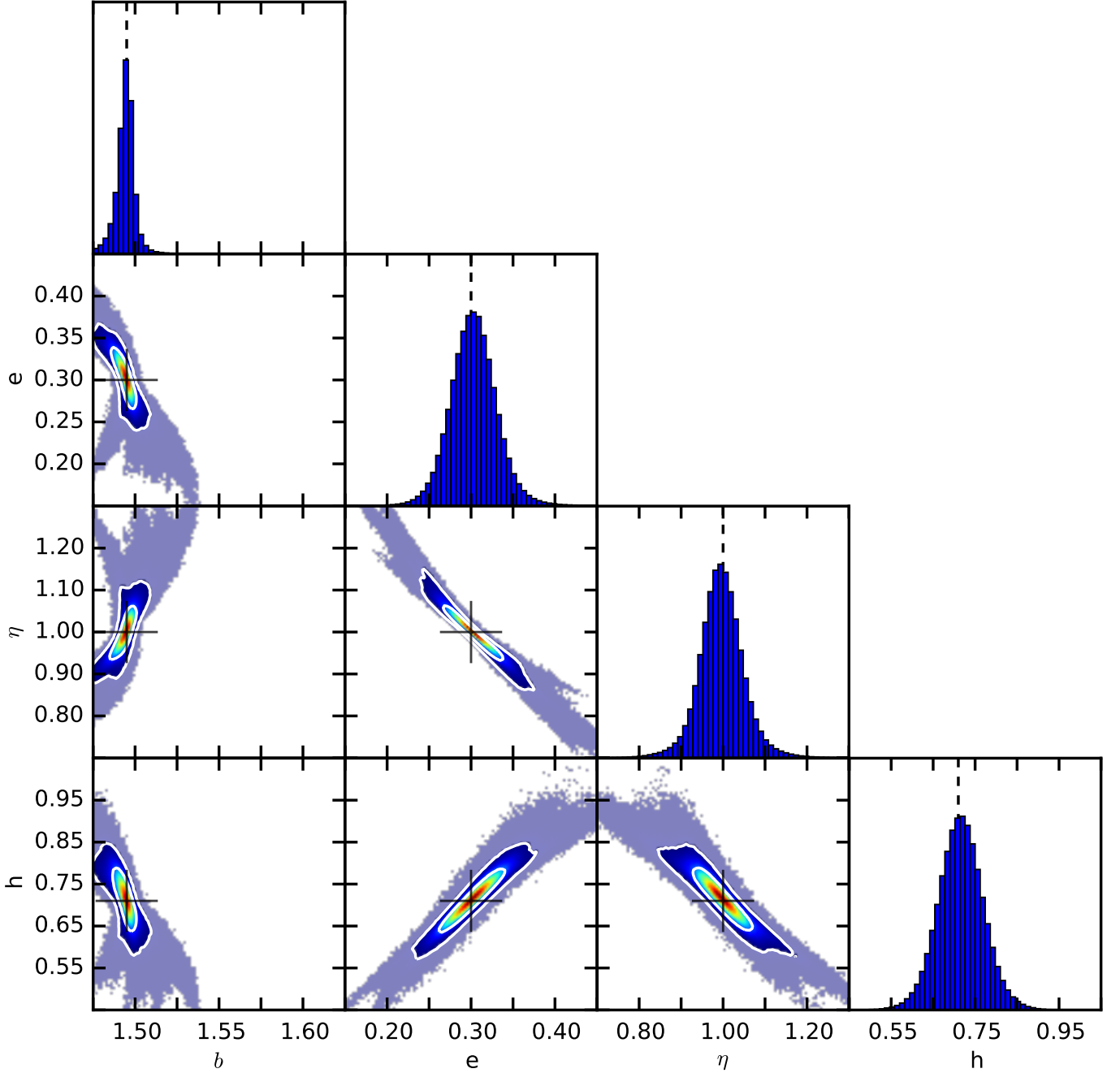


FIG. 13.— Similar to Figure 10, but for 3-D Lens models. Here all perturbing galaxies are treated with the 3-D tidal approximation. There is no bias in the recovered parameters, implying that omitting higher order terms will not bias our measurements of the Hubble constant. The scatter in the recovered parameters is much smaller than that of the Lens+Shear and Lens-Only models. However, there are still strong correlations between ellipticity, the power law index, and the Hubble constant. This is a manifestation of the lens profile degeneracy (Kochanek 2002).

improved constraints on the Hubble constant.

The final system parameter we probe is the source position. In previous simulations, we have marginalized over the orientation of the main lens galaxy, which would wash out any trends we would see in the source plane. Instead, we now fix the lens system to match the real RXJ1131. Suyu et al. (2013) find the best fit parameters of RXJ1131 to be $b = 1.64''$, $\eta = 1.05$, $e = 0.237$, and $\theta_e = 115.8^\circ$. We generate mock data for this configuration and then perform a modeling analysis as above,

again treating all ENV/LOS galaxies in the tidal approximation.

Figure 19 shows the median and scatter in the best fit Hubble constant as a function of source position. The models do well when the source is near a caustic. However, near the center of the caustic, both sets of models have larger scatter in the recovered values for the Hubble constant. These source positions correspond to more symmetric image configurations, with the exact center producing an Einstein cross. Symmetric image configura-

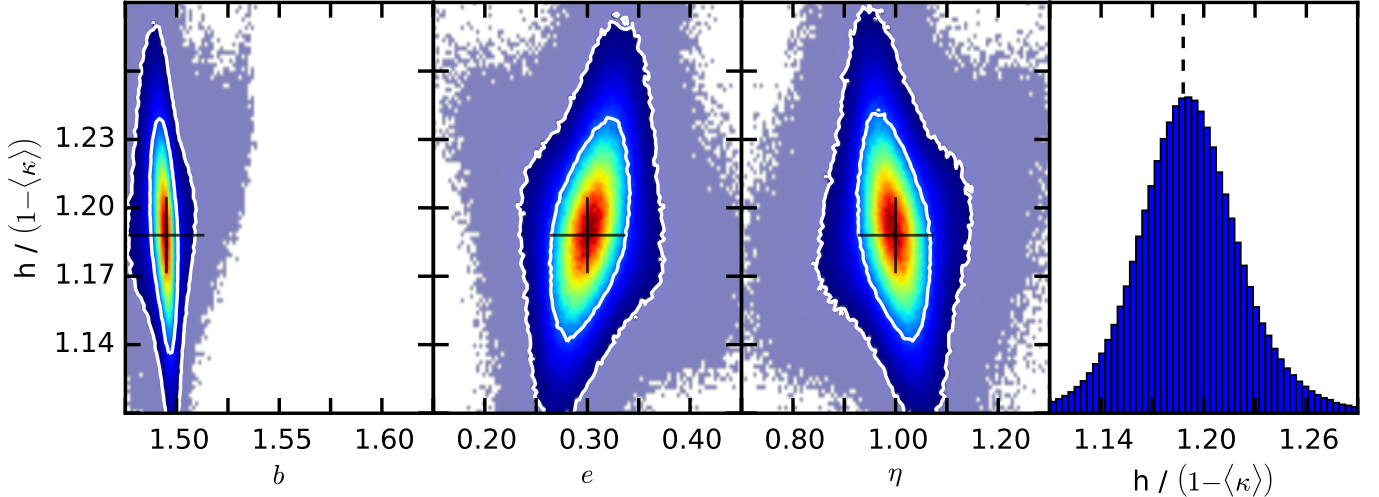


FIG. 14.— Recovered parameter distributions for 3-D Lens models (q.v. Fig. 13), but now with the Hubble constant, h , scaled to account for the lens profile degeneracy. We approximate the average convergence at the Einstein radius using equation 65. The correlations among the Hubble constant, ellipticity, and power law index are mostly removed. The residual correlations are likely due to higher order terms in the expression for h (Kochanek 2002) and the Taylor series expansion for $\langle \kappa \rangle$.

rations produce weaker constraints on the ellipticity than asymmetric image configurations and are therefore more susceptible to lens profile degeneracy as discussed above.

4. CONCLUSIONS

As lensing data improve, it becomes more important to take into account systematic effects like the perturbations due to galaxies in the environment and along the LOS. Our results show that if we want to do “precision lensing,” the environment and LOS cannot be ignored.

To understand and quantify how the ENV/LOS affects lensing, we have generated 3-D mass distributions based on photometric and spectroscopic observations. We then ray trace through these mass models to generate lensing observables. In these calculations, we self-consistently include the contribution from a smooth mass background, for the first time allowing us to account for mass (in)completeness and voids. We then fit the lensing observables with three types of models: Lens-Only models that ignore the environment, Lens+Shear models that take the common approach of fitting external shear as a free parameter, and our 3-D Lens models that treat ENV/LOS galaxies in the tidal approximation but include all non-linear effects from having mass at different redshifts.

We find that perturbers in the foreground of the lens affect the lens potential more than those in the background for two different reasons. The first is that background perturbers are downweighted; a background perturber must be closer in projection to have the same effect as a foreground perturber. The second is that while background perturbers can be mimicked by a shear in the lens plane, foreground perturbers create non-linear effects that cannot be fit with a simple external shear.

Using our results based on individual perturbing galaxies, we define a quantity that we have termed the “flexion shift”, $\Delta_3 x$ (eqs. 55 and 62), that estimates the deviations in lensed image positions due to third order (flexion) terms and can be used to characterize the ENV/LOS contributions to the lens potential for an individual perturbing galaxy.

Using this quantity, we find that the importance of environment/LOS effects varies significantly from field to field. Therefore, we argue that each field needs to be modeled individually. Even accounting for the uncertainties in building the 3-D mass models, we find that directly calculating the external convergence from the 3-D mass models produces a narrower distribution than that from ray tracing through N-body simulations. This translates to a stronger prior on measuring the Hubble constant. We find that lens galaxies that are in groups tend to have a stronger contribution from the ENV/LOS than those that are not in groups.

We show that fitting lens models that ignore the ENV/LOS does not reproduce the input lens system parameters or the Hubble constant. Models that fit an external shear, the Lens+Shear models, overpredict the Hubble constant and the Einstein radius of the main lens galaxy; these quantities need to be corrected using some other constraint on the external convergence, which is typically done in post processing. Our 3-D lens models do not produce this bias (or biases in any other model parameters) because we explicitly include the convergence in the lens models. It is still possible that there may be some unexpected bias from how we construct mass models, but that is beyond of the scope of this paper and will be considered in more detail in future works.

We show that all models, including both the Lens+Shear and 3-D Lens models, are subject to the lens profile degeneracy in agreement with Xu et al. (2016); Schneider & Sluse (2013); Sluse et al. (2012). Extra information like kinematic measurements or extended source reconstruction techniques are necessary.

For either methodology of fitting an external shear (Lens+Shear) or using our 3-D Lens model, we still must choose which galaxies are treated exactly and which are treated in the tidal approximation. Our tests suggest that a conservative cutoff in $\Delta_3 x$ is 10^{-4} arcseconds, a factor of 30 smaller than the assumed uncertainties on the observed positions. This cut typically requires 10-15 galaxies of ~ 300 galaxies within 5 arcminutes to be

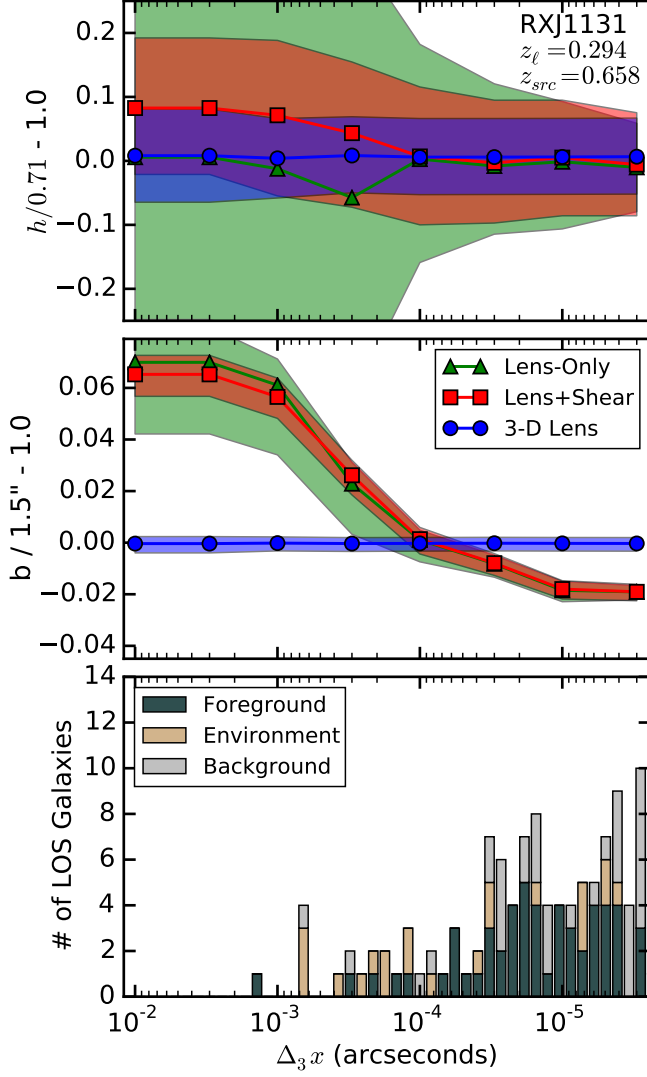


FIG. 15.— Results for lens models with different thresholds for the tidal approximation, using the RXJ1131 field. For a given value of $\Delta_3 x$, we run models in which all perturbing galaxies with a larger value of the flexion shift (i.e., those toward the left in the bottom panel) are treated exactly, while perturbing galaxies with a smaller value of the flexion shift (i.e., those toward the right) are treated with the tidal approximation. Starting at the left, all perturbors are tidal (as in Figs. 10, 11, and 13); moving to the right increases the number of perturbors that are treated explicitly instead. The top panel shows the median (points) and 68% confidence (colored band) interval for the Hubble constant recovered from 3-D Lens models (blue), Lens+Shear models (red), and Lens-Only models (green). The scatter is driven by the lens profile degeneracy. The middle panel shows similar results for the Einstein radius parameter. Lens-Only and Lens+Shear models tend to be biased because of external convergence effects which are typically added in post-processing. For RXJ1131, this accounts for a 7% correction, consistent with the peak of the distribution shown in Figure 7. As we include more galaxies exactly, moving to the right of the plot, the bias begins to disappear because we are taking into account more of the convergence explicitly. However, as we take many of the ENV/LOS galaxies into account exactly, the Lens-Only and Lens+Shear models dip below the truth. This is due to the smooth mass correction for voids which is not included in the Lens-Only and Lens+Shear models. Our simulation results suggest that the void correction is $\sim 2\%$ for RXJ1131.

treated exactly, making the calculation ~ 400 – 900 times more efficient than the full multi-plane lens equation.

LSST and Euclid will find an immense number of new

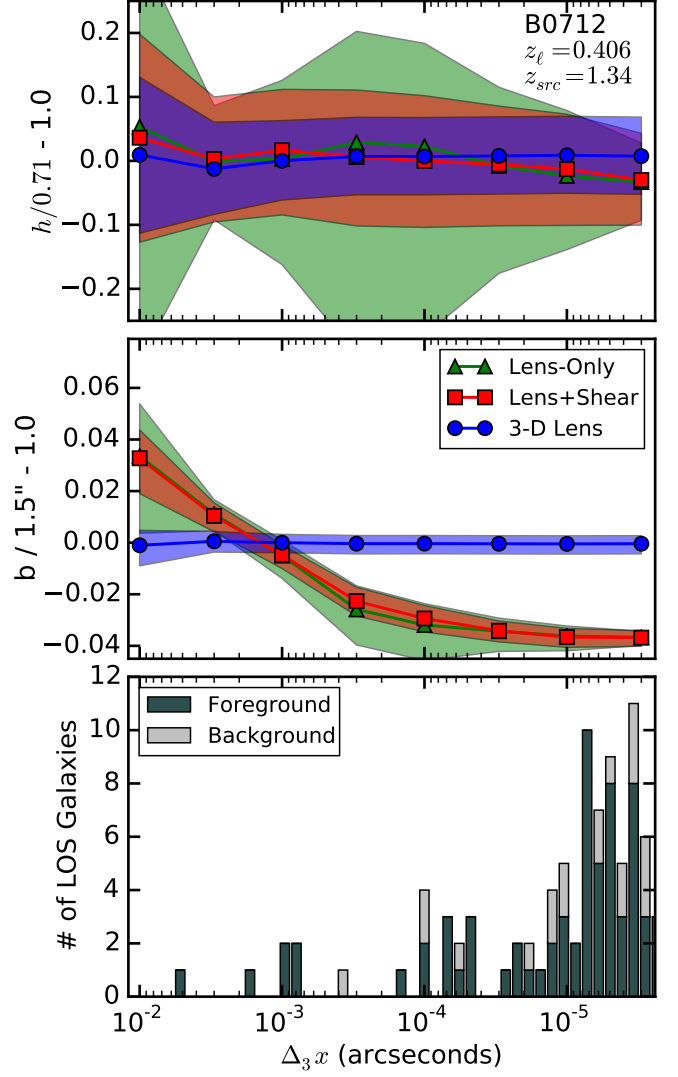


FIG. 16.— Similar to Figure 15, but for the B0712 field. This lens is not in a group in our mass models, but it does have a group of galaxies along the LOS. The qualitative trends from the ENV/LOS effects are similar to what we saw for RXJ1131, but the quantitative details differ. The correction for external convergence (seen at the farthest left point) is smaller for B0712, at $\sim 4\%$, than what we saw for RXJ1131 at $\sim 7\%$. This is consistent with the results from Figure 9 that group lenses typically have stronger ENV/LOS effects than non-group lenses. The correction for the smooth mass density (e.g., voids; inferred from the farthest right point) is larger for B0712, at $\sim 4\%$, than for RXJ1131 which only had a correction of 2% .

strong lens systems. There are various strategies about how to use this upcoming data set for cosmology. One possibility is to use all of the lenses to beat down uncertainties using statistics. However, if we are entering the systematics-limited regime, this approach will succeed only if one can account for systematic uncertainties (e.g., with our 3-D Lens models). An alternative strategy is to use the large number of lenses discovered by LSST and Euclid to search for a few rare, “golden” lenses that have small systematic uncertainties. One possible criterion for a “golden” lens could be to have a weaker contribution from the environment/LOS. Based on our analysis, to minimize environment/LOS effects, we should search for lenses with a large Einstein radius and high ellipticity.

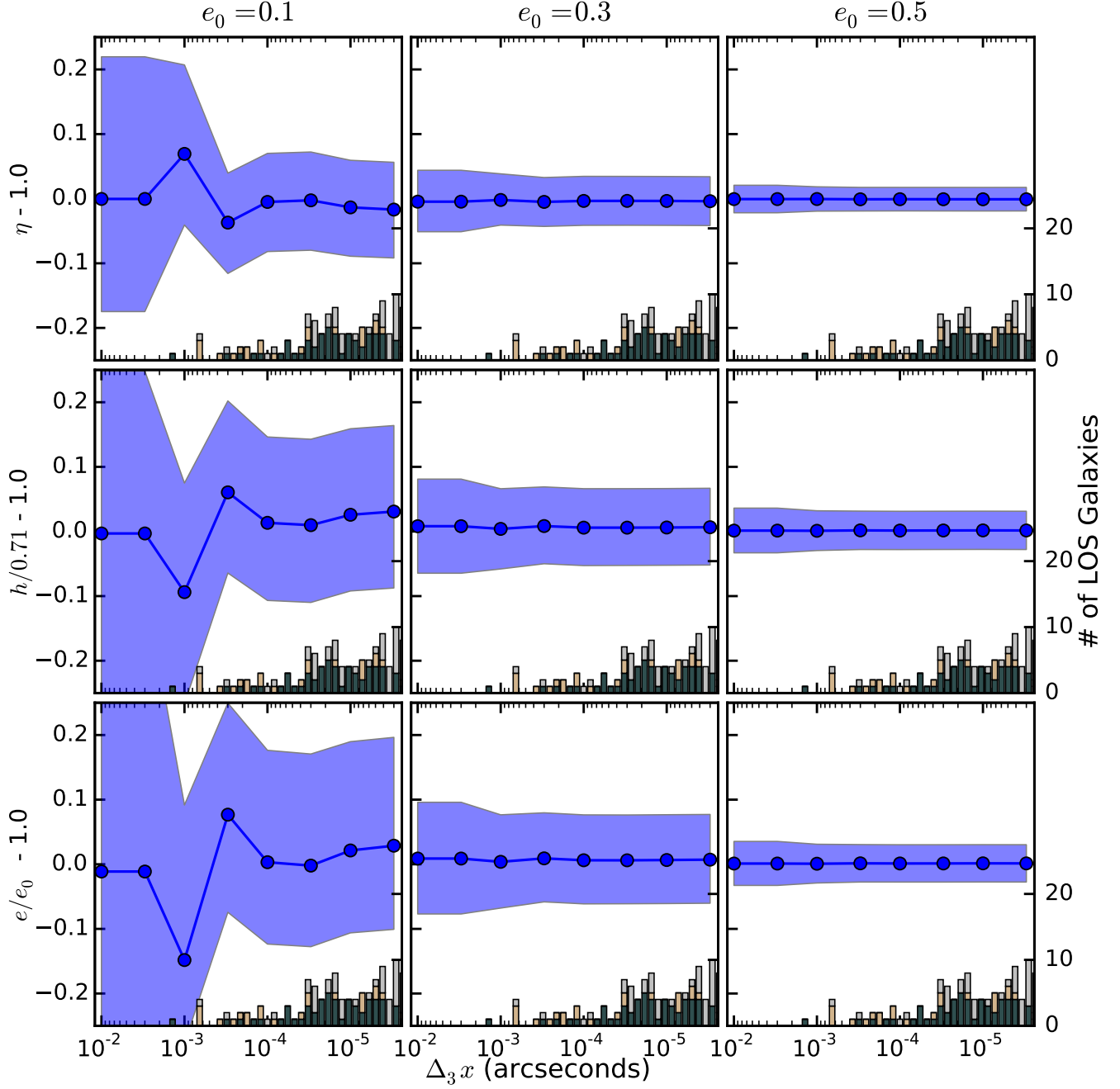


FIG. 17.— Recovered model parameters for main lens galaxies with different ellipticities, using the RXJ1131 field. The columns correspond to $e = 0.1, 0.3, 0.5$ from left to right. Systems with larger e have less scatter in the power law index, η , and the Hubble constant, h . Elongated lenses tend to produce more asymmetric image configurations. More asymmetric image configurations span a wider range of radii and provide strong constraints on the ellipticity, breaking the lens profile degeneracy and producing tighter constraints on the Hubble constant.

These large, asymmetric lenses will be less sensitive to the lens profile degeneracy. While suggestive, these results merit further investigation to get the most out of future surveys like LSST and Euclid.

Throughout this work we have assumed that our 3-D models perfectly describe the environment/LOS, but there is also uncertainty in generating the environment mass models that is only briefly addressed here. This source of uncertainty will be explored in a forthcoming

paper (Wong et al., in prep.).

We thank the Anonymous Referee for their comments and careful consideration. We thank Edi Rusu, Phil Marshall, Sherry Suyu, Chris Fassnacht, and Stefan Hilbert for enlightening discussions about 3-D lensing. We thank Iair Arcavi and Griffin Hosseinzadeh for discussions about presentation of these results. CVM acknowledges support from NSF grant AST-1313484. CRK ac-

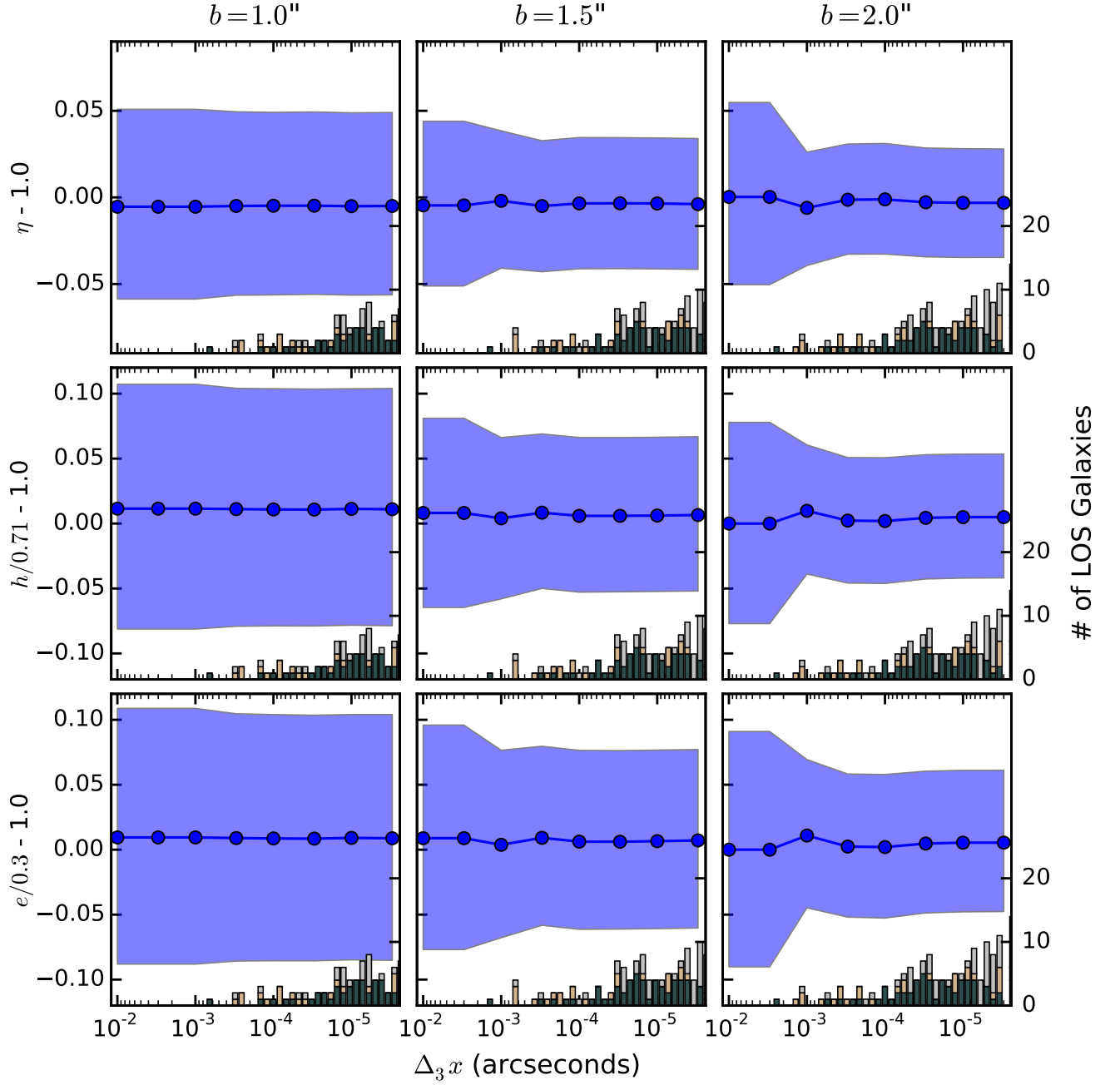


FIG. 18.— Similar to Figure 17 but for different values of the Einstein radius parameter: $b = 1.0, 1.5, 2.0$ from left to right. Systems with a smaller Einstein radius have more scatter in the Hubble constant. Constraints on the ellipticity are somewhat stronger for lenses with larger Einstein radii, especially once the few strongest perturbers are included in models. This implies that lenses with a large Einstein radius should produce the strongest constraints on the Hubble constant.

knowledge funding from NSF grant AST-1211385. AIZ acknowledges funding from NSF grant AST-1211874 and NASA grant ADP-10AE88G. She also thanks the John Simon Guggenheim Foundation and the Center for Cosmology and Particle Physics at NYU for their support. 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.

REFERENCES

- Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, *ApJS*, 203, 21
- Ammons, S. M., Wong, K. C., Zabludoff, A. I., & Keeton, C. R. 2014, *ApJ*, 781, 2

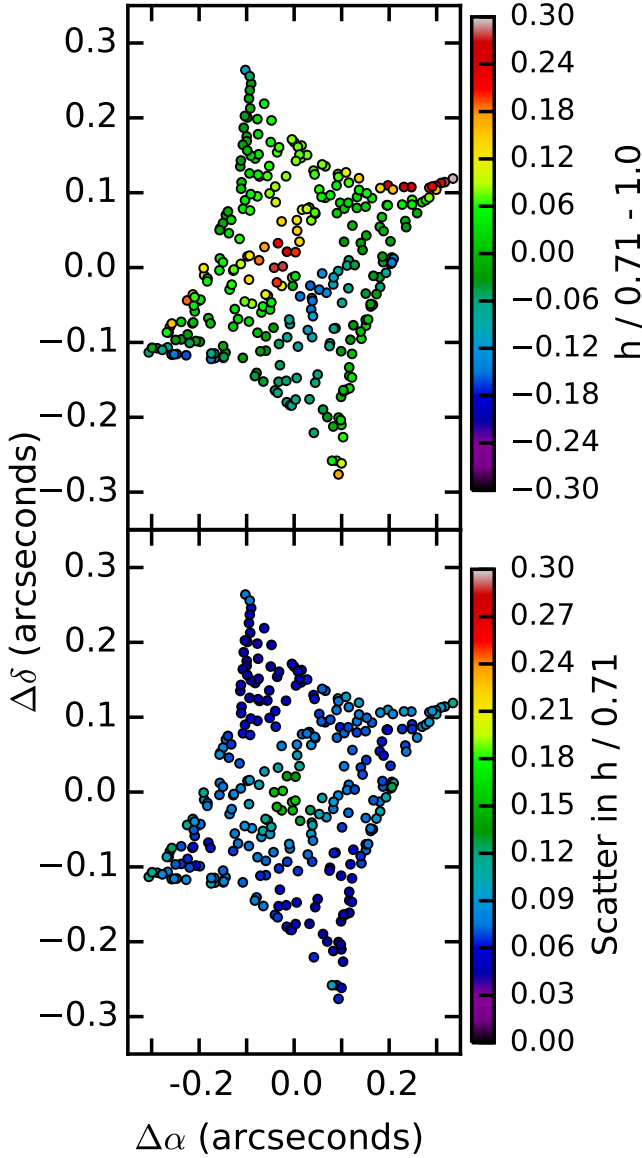


FIG. 19.— Median value (top) and scatter (bottom) for the recovered Hubble constant as a function of source position. The models have little bias and small scatter near caustics, but show an increased scatter/bias near the center. These central source positions correspond to more symmetric image configurations which produce weaker constraints on the ellipticity and are therefore more susceptible to the lens profile degeneracy. More asymmetric image configurations produce stronger constraints on the lens galaxy ellipticity and in turn on the Hubble constant.

Auger, M. W., Treu, T., Bolton, A. S., et al. 2010, *ApJ*, 724, 511
 Bacon, D. J., Goldberg, D. M., Rowe, B. T. P., & Taylor, A. N. 2006, *MNRAS*, 365, 414
 Bar-Kana, R. 1996, *ApJ*, 468, 17
 Barnabè, M., Czoske, O., Koopmans, L. V. E., et al. 2009, *MNRAS*, 399, 21
 Belokurov, V., Evans, N. W., Hewett, P. C., et al. 2009, *MNRAS*, 392, 104
 Birrer, S., Amara, A., & Refregier, A. 2015, *ArXiv e-prints*, arXiv:1511.03662
 Bolton, A. S., Burles, S., Koopmans, L. V. E., et al. 2008, *ApJ*, 682, 964
 Bonvin, V., Courbin, F., Suyu, S. H., et al. 2016, *ArXiv e-prints*, arXiv:1607.01790
 Brimiouille, F., Seitz, S., Lerchster, M., Bender, R., & Snigula, J. 2013, *MNRAS*, 432, 1046

Bruderer, C., Read, J. I., Coles, J. P., et al. 2015, *arXiv*, 1511.01097
 Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, *MNRAS*, 321, 559
 Calanog, J. A., Fu, H., Cooray, A., et al. 2014, *ApJ*, 797, 138
 Cao, S., Covone, G., & Zhu, Z.-H. 2012, *ApJ*, 755, 31
 Chen, G. C. F., Suyu, S. H., Wong, K. C., et al. 2016, *ArXiv e-prints*, arXiv:1601.01321
 Coe, D., & Moustakas, L. A. 2009, *ApJ*, 706, 45
 Collett, T. E. 2015, *ArXiv e-prints*, arXiv:1507.02657
 Collett, T. E., & Cunningham, S. D. 2016, *ArXiv e-prints*, arXiv:1605.08341
 Collett, T. E., Marshall, P. J., Auger, M. W., et al. 2013, *MNRAS*, 432, 679
 Csabai, I., Budavári, T., Connolly, A. J., et al. 2003, *AJ*, 125, 580
 Dalal, N., & Kochanek, C. S. 2002, *ApJ*, 572, 25
 Eigenbrod, A., Courbin, F., Vuissoz, C., et al. 2005, *A&A*, 436, 25
 Fadely, R., & Keeton, C. R. 2012, *MNRAS*, 419, 936
 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., & Lubin, L. M. 2002, *AJ*, 123, 627
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
 Gavazzi, R., Treu, T., Rhodes, J. D., et al. 2007, *ApJ*, 667, 176
 Girardi, M., Giuricin, G., Mardirossian, F., Mezzetti, M., & Boschin, W. 1998, *ApJ*, 505, 74
 Greene, Z. S., Suyu, S. H., Treu, T., et al. 2013, *ApJ*, 768, 39
 Hezaveh, Y., Dalal, N., Holder, G., et al. 2014, *ArXiv e-prints*, arXiv:1403.2720
 Hezaveh, Y. D., Marrone, D. P., Fassnacht, C. D., et al. 2013, *ApJ*, 767, 132
 Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2009, *A&A*, 499, 31
 Jackson, N., Nair, S., Browne, I. W. A., et al. 1998, *MNRAS*, 296, 483
 Jaroszynski, M., & Kostrzewa-Rutkowska, Z. 2012, *MNRAS*, 424, 325
 —. 2014, *MNRAS*, 439, 2432
 Jaroszynski, M., & Skowron, J. 2016, *ArXiv e-prints*, arXiv:1607.06256
 Keeton, C. R. 2001, *ArXiv: astro-ph/0102340*
 —. 2003, *ApJ*, 584, 664
 Keeton, C. R., & Kochanek, C. S. 1997, *ApJ*, 487, 42
 Keeton, C. R., Kochanek, C. S., & Falco, E. E. 1998, *ApJ*, 509, 561
 Keeton, C. R., & Zabludoff, A. I. 2004, *ApJ*, 612, 660
 Kelly, P. L., Rodney, S. A., Treu, T., et al. 2015, *Science*, 347, 1123
 —. 2016, *ApJ*, 819, L8
 King, L. J., Browne, I. W. A., Marlow, D. R., Patnaik, A. R., & Wilkinson, P. N. 1999, *MNRAS*, 307, 225
 Kochanek, C. S. 2002, *ApJ*, 578, 25
 —. 2003, *ApJ*, 583, 49
 Kochanek, C. S. 2006, in *Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro*, ed. G. Meylan, P. Jetzer, & P. North, 91–268
 Kolatt, T. S., & Bartelmann, M. 1998, *MNRAS*, 296, 763
 Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, *ApJ*, 649, 599
 Koopmans, L. V. E., Bolton, A., Treu, T., et al. 2009, *ApJ*, 703, L51
 Lagattuta, D. J., Vegetti, S., Fassnacht, C. D., et al. 2012, *MNRAS*, 424, 2800
 Lehar, J., Falco, E. E., Kochanek, C. S., et al. 2000, *ApJ*, 536, 584
 Linder, E. V. 2004, *Phys. Rev. D*, 70, 043534
 —. 2011, *Phys. Rev. D*, 84, 123529
 LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, *ArXiv*, 0912.0201
 Mao, S., & Schneider, P. 1998, *MNRAS*, 295, 587
 McCully, C., Keeton, C. R., Wong, K. C., & Zabludoff, A. I. 2014, *MNRAS*, 443, 3631
 Metcalf, R. B., & Madau, P. 2001, *ApJ*, 563, 9
 Metcalf, R. B., & Petkova, M. 2014, *MNRAS*, 445, 1942
 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
- Myers, S. T., Jackson, N. J., Browne, I. W. A., et al. 2003, *MNRAS*, 341, 1
- Nakajima, R., Bernstein, G. M., Fadely, R., Keeton, C. R., & Schrabback, T. 2009, *ApJ*, 697, 1793
- Oguri, M. 2007, *ApJ*, 660, 1
- Oguri, M., & Marshall, P. J. 2010, *MNRAS*, 405, 2579
- Oguri, M., Inada, N., Pindor, B., et al. 2006, *AJ*, 132, 999
- Patel, B., McCully, C., Jha, S. W., et al. 2014, *ApJ*, 786, 9
- Petkova, M., Metcalf, R. B., & Giocoli, C. 2014, *MNRAS*, 445, 1954
- Petters, A. O., Levine, H., & Wambsganss, J. 2001, *Singularity theory and gravitational lensing* (Birkhäuser, Boston)
- Refregier, A., Amara, A., Kitching, T. D., et al. 2010, *ArXiv e-prints*, arXiv:1001.0061
- Refsdal, S. 1964, *MNRAS*, 128, 307
- Rodney, S. A., Strolger, L.-G., Kelly, P. L., et al. 2016, *ApJ*, 820, 50
- Rusin, D., Kochanek, C. S., & Keeton, C. R. 2003, *ApJ*, 595, 29
- Rusu, C. E., Fassnacht, C. D., Sluse, D., et al. 2016, *ArXiv e-prints*, arXiv:1607.01047
- Saha, P., Coles, J., Macciò, A. V., & Williams, L. L. R. 2006, *ApJ*, 650, L17
- Schneider, P. 1997, *MNRAS*, 292, 673
- . 2014, *ArXiv e-prints*, arXiv:1409.0015
- Schneider, P., & Sluse, D. 2013, *A&A*, 559, A37
- Sluse, D., Chantry, V., Magain, P., Courbin, F., & Meylan, G. 2012, *A&A*, 538, A99
- Sluse, D., Surdej, J., Claeskens, J.-F., et al. 2003, *A&A*, 406, L43
- Sluse, D., Sonnenfeld, A., Rumbaugh, N., et al. 2016, *ArXiv e-prints*, arXiv:1607.00382
- Sonnenfeld, A., Gavazzi, R., Suyu, S. H., Treu, T., & Marshall, P. J. 2013, *ApJ*, 777, 97
- Suyu, S. H. 2012, *MNRAS*, 426, 868
- Suyu, S. H., Marshall, P. J., Auger, M. W., et al. 2010, *ApJ*, 711, 201
- Suyu, S. H., Auger, M. W., Hilbert, S., et al. 2013, *ApJ*, 766, 70
- Suyu, S. H., Bonvin, V., Courbin, F., et al. 2016, *ArXiv e-prints*, arXiv:1607.00017
- Tonry, J. L., & Kochanek, C. S. 1999, *AJ*, 117, 2034
- Treu, T., Auger, M. W., Koopmans, L. V. E., et al. 2010, *ApJ*, 709, 1195
- Treu, T., Koopmans, L. V. E., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, *ApJ*, 650, 1219
- Treu, T., Marshall, P. J., Cyr-Racine, F.-Y., et al. 2013, *ArXiv:1306.1272*, arXiv:1306.1272
- Turner, E. L. 1990, *ApJ*, 365, L43
- Vegetti, S., & Vogelsberger, M. 2014, *MNRAS*, 442, 3598
- Williams, K. A., Momcheva, I., Keeton, C. R., Zabludoff, A. I., & Lehar, J. 2006, *ApJ*, 646, 85
- Wong, K. C., Keeton, C. R., Williams, K. A., Momcheva, I. G., & Zabludoff, A. I. 2011, *ApJ*, 726, 84
- Wong, K. C., Tran, K.-V. H., Suyu, S. H., et al. 2014, *ApJ*, 789, L31
- Wong, K. C., Suyu, S. H., Auger, M. W., et al. 2016, *ArXiv e-prints*, arXiv:1607.01403
- Xu, D., Sluse, D., Schneider, P., et al. 2016, *MNRAS*, 456, 739
- Zhao, D. H., Jing, Y. P., Mo, H. J., & Börner, G. 2009, *ApJ*, 707, 354