

# Precision cosmology with time delay lenses: high resolution imaging requirements

**Xiao-Lei Meng,<sup>1</sup> Tommaso Treu,<sup>1,2</sup> Adriano Agnello,<sup>1,2</sup> Matthew W. Auger,<sup>3</sup> Kai Liao,<sup>1,2</sup> Philip J. Marshall<sup>4</sup>**

<sup>1</sup>Department of Physics, University of California, Santa Barbara, CA 93106, USA

<sup>2</sup>Physics and Astronomy Building, 430 Portola Plaza, Box 951547, Los Angeles, CA 90095-1547, USA

<sup>3</sup>Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

<sup>4</sup>Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA

E-mail: [xlmeng919@gmail.com](mailto:xlmeng919@gmail.com)

**Abstract.** Lens time delays are a powerful probe of cosmology, provided that the gravitational potential of the main deflector can be modeled with sufficient precision. Recent work has shown that this can be achieved by detailed modeling of the host galaxies of lensed quasars. The distortion of the images as measured over large number of pixels provides tight constraints on the difference between the gravitational potential between the quasar image positions, and thus on cosmology in combination with the measured time delay. We carry out a systematic exploration of the high resolution imaging required to **exploit** the thousands of lensed quasars that will be discovered by current and upcoming surveys with the next decade. Specifically we simulate realistic lens systems as imaged by the Hubble Space Telescope (HST), James Webb Space Telescope (JWST), ground based adaptive optics images taken with Keck or the Thirty Meter Telescope (TMT). We compare the performance of these pointed observations with that of images taken by the Euclid-VIS, Wide-Field Infrared Survey Telescope (WFIRST) and Large Synoptic Survey Telescope (LSST) surveys. We use as our metric the precision with which the slope  $\gamma'$  of the total mass density profile  $\rho_{tot} \propto r^{-\gamma'}$  for the main deflector can be measured. Ideally, we require that the statistical error on  $\gamma'$  be less than 0.02, thus being subdominant with respect to other sources of random and systematic uncertainties. We find that survey data will likely have sufficient depth and resolution to meet the target only for the brighter gravitational lens systems, comparable to those discovered by the SDSS survey. For fainter systems, that will be discovered by current and future surveys, targeted follow-up will be required. However, the exposure time per target will be of order a few minutes per system with upcoming facilities, such as JWST, the Keck Next Generation Adaptive Optics System, and TMT, thus making the follow-up of hundreds of systems a practical and efficient cosmological probe.

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Summary of simulated instrumental setups</b>	<b>3</b>
<b>3</b>	<b>The Lens Sample</b>	<b>4</b>
3.1	Sample Selection	4
3.2	Parameters of the Sample of Mock Lenses	5
<b>4</b>	<b>Description of the inference process</b>	<b>6</b>
4.1	Image generating process	6
4.1.1	Light Model	6
4.1.2	Mass Model	6
4.1.3	Point Source Model	7
4.1.4	Image Noise	7
4.2	Inferring the parameters	7
<b>5</b>	<b>Results</b>	<b>8</b>
<b>6</b>	<b>Summary</b>	<b>9</b>

---

## 1 Introduction

In the past few years, gravitational time delays (Refsdal 1964 [? ]) have emerged as a powerful and cost effective cosmological probe. Studies based on blind analysis have shown that a single system consisting of multiple images of a background quasar (at redshift  $z_s$ ) lensed by a foreground massive elliptical galaxy at redshift  $z_d$  can be used to measure the so-called time delay distance with an uncertainty of 6-7% (Suyu et al. 2010 [? ], 2013 [? ], 2014 [? ]). The time delay distance  $D_{\Delta t}$  gives a direct measurement of the Hubble constant and allows one to break some the main degeneracies in the interpretation of cosmic microwave background data, thus providing tight constraints on parameters such as curvature and dark energy equation of state (Coe & Moustakas 2009 [? ], Linder 2011 [? ], Weinberg 2013 [? ], Suyu 2012 [? ], Suyu et al. 2014 [? ]). The time delay distance measurement of  $H_0$  is comparable in terms of information content to that obtained via the cosmic distance ladder (Riess et al. 2011 [? ], Freedman et al. 2012 [? ]), in that is based on absolute distances and completely independent on the properties of the universe at  $z > z_s$ . Importantly, time delay distances are independent of the local distance ladder and thus provide a crucial test of any potential systematic uncertainties. Furthermore, being independent, the cosmic distance ladder and time delay distance constraints on  $H_0$  can be statistically combined for additional gains in precision.

From an observational point of view, the attainment of precise and accurate time delay measurements relies on a few important ingredients. First, monitoring of the lensed quasars is required to obtain time delays with a few percent uncertainties. Dedicated monitoring campaigns have shown that this precision is achievable with 1-2m class telescopes at present time (Tewes et al. 2013 [? ]), or in the radio (Fassnacht et al. 2002 [? ]). In the future, the exploitation of larger samples of lenses (Oguri & Marshall 2010 [? ]) will require monitoring campaigns on 2-4m telescopes (Treu et al. 2013 [? ]), or the deployment of long term high-cadence monitoring surveys, like the LSST (Liao

et al. 2015 [? ]). Second, the spectroscopic redshift of the source and deflector must be measured. For current samples, this is typically a relatively straightforward step, requiring short exposures on 4-10m class telescopes (Fassnacht & Cohen 1998 [? ], Eigenbrod et al. 2007 [? ]). Third, the gravitational potential of the main deflector needs to be constrained by the data so that the uncertainty on its difference between the location of the images is also of order 3%. This goal can be achieved by studying the extended structure of the lensed quasar host galaxy, and the stellar kinematics of the deflector galaxy (Treu & Koopmans 2002 [? ], Koopmans et al. 2003 [? ], Suyu et al. 2010 [? ], Suyu et al. 2013 [? ], Suyu et al. 2014 [? ]). Fourth, the combined effects of the inhomogenous mass distribution along the line of sight need to be taken into account. Recent work has shown that the line of sight effects can be sufficiently characterized by measuring the properties of galaxies and weak lensing signal in the field of the main deflector, and comparing with simulated lines of sight (Suyu et al. 2010 [? ], 2013 [? ], 2014 [? ], Greene et al. 2013 [? ], Collett et al. 2013 [? ]).

However, whereas current samples have been limited to a few lenses, current and ongoing surveys (add references to DES, PANSTARRS, HSC, ATLAS) will discover 100s of lenses. These large samples will be extremely powerful cosmographic probes (Coe & Moustakas 2009 [? ], Linder 2011 [? ]), provided that sufficient follow-up data will be available (Treu et al. 2013 [? ]). In this paper we carry out a feasibility study for future time delay surveys, focusing on the high-resolution imaging requirements. Specifically, we aim to investigate whether sufficient information will be available from survey data, or whether dedicated follow-up will be necessary. If dedicated follow-up will be necessary, we aim to provide an estimate of the amount of time that will be required per system. The multi-pronged nature of time delay lens follow-up makes it natural to follow the approach of addressing each component of the follow-up independently. The monitoring and spectroscopic follow-up requirements are described, e.g., by Eigenbrod et al. 2005 [? ], Linder 2015 [? ].

In order to make the problem tractable and the results general, we need a single simple metric to evaluate the quality of imaging data. In short, we need to quantify our ability to measure how the deflection angle (i.e. the derivative of the lensing potential) varies between the images, and thus how the extended images of the quasar host galaxy are stretched across the image. In practice, our ability to constrain the differential magnification will depend on the resolution of the data as well as on the number of pixels where the source is detected (hereafter *informative pixels*).

We choose to adopt as our metric the slope  $\gamma'$  of the total mass density profile of the form  $\rho_{\text{tot}} \propto r^{-\gamma'}$  as measured by fitting elliptical power law models to the data. This profile is the simplest one that provides a realistic description of galaxy scale lenses (Treu 2010 [? ]), and the uncertainty  $\gamma'$  has been shown to be approximately proportional to the uncertainty on the gravitational potential differences and thus the time delay distance (Kochanek 2002 [? ], Wucknitz 2002 [? ], Suyu 2012 [? ]). Intuitively,  $\gamma'$  is directly related to the variations of the lensing potential. In fact, if  $\gamma' = 2$  (the so-called isothermal profile) the deflection angle is constant across the image and therefore all the images will appear to have the same radial magnification. If the profile is steeper/shallower, radial magnification will vary across the image, thus giving rise to images of different widths.

Naturally, for a real measurement it is important to explore different profiles and the role of the choice of the profile in the uncertainties. However, our goal is to estimate minimal requirements on the data quality. If the data quality is insufficient to constrain  $\gamma'$  it will also be insufficient to constrain more flexible models. Likewise, systematic errors not considered in this paper, such as those arising from incomplete knowledge of the PSF, will only increase the error budget. Thus, in order to leave room for additional sources of uncertainty in the error budget, we set a rather stringent requirement of 0.02 uncertainty of  $\gamma'$  (corresponding approximately to 2% per system on time delay distance). In other words, our target corresponds to the requirement that the *statistical* error arising from the image quality be subdominant with respect to those arising from modeling uncertainties, time delays, and

line of sight effects.

This paper is organized as follows. In Section 2, we summarize the characteristics of the telescopes and instruments simulated as part of this work. Then, in Section 3, we describe the properties of the simulated lenses. Next, in Section 4 we describe the procedure used to carry out the inference. Our results are presented in Section 5. Finally, we discuss and summarize our work in Section 6. Throughout this paper, all magnitudes are given in the AB system. Even though our findings are independent of any cosmology, we adopt a spatially flat  $\Lambda$ CDM cosmology with  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$ , and the Hubble constant  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  when calculating distances.

## 2 Summary of simulated instrumental setups

We aim to carry out a systematic exploration of imaging requirements given current and future facilities. For a given lens system, the instrumental setup drives our ability to constrain  $\gamma'$  in a number of ways. First, the signal-to-noise ratio depends on the exposure-time adopted, sky or background noise in a chosen band, and the instrumental readout noise. Second, the pixel-size and the point-spread function (PSF) properties determine how finely one can map the system being observed, hence how robustly the deflections on either side of the lens can be quantified.

Here we summarize the main properties of instruments that will be considered in our simulations. In an effort to balance completeness with feasibility, we consider a suite of instruments that include ground and space based telescopes, and cadenced surveys where the total exposure cannot be set arbitrarily. Some of the choices are representative of a class of telescopes/instruments, for example we expect the performance of Keck Adaptive Optics to be a useful guidance for adaptive optics systems on other 8-10m class telescopes, and that of the Thirty Meter Telescope to be a guidance for other planned 20-30m class telescopes.

Table 1 and figure 1 display the main properties of the telescopes that are needed when generating images of mock gravitational lenses, taken from the instrument websites. Figure 2 shows the typical PSF of each instrument. In order to facilitate comparisons between instruments and telescopes, whenever possible we selected the filter/band for each configuration that is closer to the  $I$ -band in the optical and the  $K'$ -band in the near infrared. A brief description of each setup follows.

1. The Advanced Camera for Surveys (ACS) is chosen as our reference optical imager on board the Hubble Space Telescope (HST). In the chosen F814W filter its performance is comparable to those of WFC3 (slightly higher sensitivity and coarser pixel size). For simplicity, we neglect charge transfer inefficiency effects, assuming that they are negligible or can be corrected to the desired level. The PSF is simulated using the Tiny Tim software (Krist 1993 [?])
2. The Near Infrared Camera (NIRCAM) on board the James Webb Space Telescope (JWST) is chosen as the next space based imaging capability. We select the broad F200W filter, where the image quality is virtually diffraction limited.
3. The current and planned adaptive optics systems (hereafter LGSAO & NGAO; Wizinowich et al. [?], [?], respectively) at the W.M.Keck Observatory are chosen to represent current and upcoming AO performance on 8-10m class telescopes. We consider the current instrument NIRC2, although further gains might be possible with an instrument upgrade in conjunction with the AO system upgrade. We adopt the typical configuration used for studies of gravitational lens systems (Fassnacht & Cohen 1998 [?]):  $K'$  filter, MCDS readout mode, read number  $N = 16$ , and narrow camera mode. A real observed PSF is used for the current AO system. The simulated PSF and performance characteristics of NGAO have been kindly provided by the NGAO team.

4. The infrared imager and spectrograph IRIS (Larkin et al. 2010 [? ]) working behind adaptive optics on the thirty Meter Telescope (TMT) is selected to represent the performance of the extremely large telescopes that will be operational in the next decade. IRIS is expected to be close to diffraction limited in the  $K'$  band. The simulated PSF and performance characteristics of TMT-IRIS have been obtained from the IRIS team, from the TMT Exposure Time Calculator (ETC)<sup>1</sup> and from the paper by Do et al. (2014) [? ].
5. Euclid is a space survey telescope planned to be launched by the European Space Agency. Euclid has two instruments: a high-resolution visible imager, and a NIR imaging instrument. We study the performance of the high-resolution visible imager, the more suitable to detailed gravitational lensing work, with the standard survey parameters. The instrument properties are taken from Schweitzer et al. 2010 [? ], Penny et al. 2013 [? ] and Cropper et al. 2014 [? ]. Since the PSF is proprietary to the Euclid team, we adopt for simplicity a Gaussian function with full width half maximum  $0''.18$ .
6. We study the survey mode of the 2.4m WFIRST space mission, focusing on the HLS Imaging Survey in the F184W filter. The instrument/survey properties are based on WFIRST-Astrophysics Focused Telescope Assets Final Report (by the Science Definition Team (ADT) and WFIRST Project) and WFIRST ETC<sup>2</sup>. Since the final PSF was not available at the time of this writing we adopt a Gaussian function with FWHM  $0''.15$ .
7. The Large Synoptic Survey Telescope (LSST) is designed to imaging surveys in the optical bands (out to  $\sim 1$  micron). The survey and instrument properties are taken from Ivezić et al. 2008 [? ] and the LSST Science Requirements Document<sup>3</sup>. The PSF of LSST will vary over time. We adopt a Gaussian PSF with FWHM  $0''.7$  as representative of the image quality of LSST in the i-band.

### 3 The Lens Sample

Here we introduce our sample, four groups of lens system with galaxies at redshift  $z_s = 1.071$  and  $2.77$  being multiply imaged by massive lens galaxies lying at  $z_d = 0.351$  and  $0.783$ .



#### 3.1 Sample Selection

We have chosen four prototypical systems for this exploration, characterized as **faint** or **bright** and **double** or **quad** depending on the photometry and image configuration. This set of four main choices covers regimes with different numbers of informative pixels, which depend on the S/N ratio and on the number of images produced by the lens.

The brighter lenses are selected to be representative of the majority of currently known lenses, selected from the Sloan Digital Sky Survey, while the fainter lenses are selected to be representative of the fainter systems to be discovered in current and future surveys within the next decade (Oguri & Marshall 2010 [? ]). We expect that the systems to be used to measure time delay distances in the next decade will span approximately the range covered by our sample. Note that some known lensed quasars are significantly brighter than any of the systems simulated here [e.g., ? ]. Even though we expect more of those to be discovered in the near future and they will be useful, they will not represent

<sup>1</sup>[http://tmt.mtk.nao.ac.jp/ETC\\_readme.html](http://tmt.mtk.nao.ac.jp/ETC_readme.html)

<sup>2</sup><http://wfirst-web.ipac.caltech.edu/wfDepc/wfDepc.jsp?etc>

<sup>3</sup><http://www.lsst.org/files/docs/SRD.pdf>

the majority of the systems in future large samples, nor present any major challenge for follow-up, so we do not consider them here.

To remain as realistic as possible, the brighter mocks are based upon lenses in the Sloan Lens ACS Survey (SLACS; Bolton et al. 2006 [? ], Treu et al. 2006 [? ], Koopmans et al. 2006 [? ], Gavazzi et al. 2007 [? ], Bolton et al. 2008a [? ], Gavazzi et al. 2008 [? ], Bolton et al. 2008b [? ], Treu et al. 2009 [? ], Auger et al. 2009 [? ], Auger et al. 2010 [? ], Newton et al. 2011 [? ], Shu et al. 2015 [? ], Papers I-XII, respectively) and the fainter mocks on the Strong Lensing Legacy Survey (SL2S; More et al. 2012 [? ], Gavazzi et al. 2012 [? ], Ruff et al. 2011 [? ], Sonnenfeld et al. 2013 [? ], Sonnenfeld et al. 2013 [? ], Sonnenfeld et al. 2015 [? ], Papers I-V, respectively, except for the first citation) samples.

### 3.2 Parameters of the Sample of Mock Lenses

The structural parameters of the four systems are listed in Table 2 and Table 3. When not explicitly known from existing data, mock model parameters are assigned via plausibility arguments as specified below. The source and deflector of the *bright* lens system configurations are built along SLACS J0330-0020 as modelled by Bolton et al. 2008a (SLACS V, [? ]), Auger et al. 2009 (SLACS IX, [? ]), and Newton et al. 2011 (SLACS XI, [? ]). For the *faint* lens system configuration, the source and deflector have parameters from the Sonnenfeld et al. 2013 (SL2S III, [? ]) model of SL2S J135949+553550 (also, Sonnenfeld, A., 2015, private communication).

Source and deflector magnitudes in the *bright* case for  $I$  and  $V$  bands are from Newton et al. 2011 [? ], for  $K$ -band are estimated based on the other colors, and typical spectral energy distributions, i.e. an early-type galaxy for the deflector and a star forming galaxy for the source. Meanwhile, source magnitudes in the *faint* case are arbitrarily set to 25.0 in all bands. Deflector magnitudes in the *faint* case are not directly available from the data, so they require some extrapolations. The value for  $I$ -band is based on Sonnenfeld et al. 2013 (SL2S III, [? ]), while for  $K$  and  $V$  bands are estimated with K-corrections from typical spectral energy distributions. Unknown magnitudes in  $K$ -band are assigned from  $H$ -band ones via  $K_{AB} = H_{AB}$ .

The effective (half-light) radii to be fitted in *bright* systems are evaluated separately in different bands, to get a precise estimate of the effective radius and its uncertainties. Differently, a unique effective radius is asserted across bandpasses for the *faint* systems. Neglecting colour gradients in this case does not affect the ability to recover the masss density profile in any significant way [? ].

For both the *faint* and the *bright* systems, the source positions are assigned so as to map the source in either two or four images. This choice allows us to explore whether the number of images makes any substantial difference while keeping the other properties of the systems fixed.

In order to construct realistic systems for time delay measurements, point sources must be added to represent the lensed quasars. As described in the next section this is done in the image plane, in order to gain computational precision and efficiency. The magnification at the location of the quasar images is calculated by solving the lens equation using `gravlens` [? ]. The source-plane magnitudes of the point sources are chosen to be somewhat brighter then the host galaxy, as it is typically the case for medium luminosity AGN (e.g., Bennert et al. 2011 [? ]), and have realistic colors for AGN at the source redshifts. The adopted magnitudes are listed in Table 4.

We note that the *bright* source configuration is too bright to be practical for TMT, given its sensitivity – the exposure time scales as  $D^{-4}$ , where  $D$  is the telescope diameter, for background limited point source exposures **[AA: keep or excise?]**. Realistically, observations of such bright systems would be completely dominated by overheads related to target acquisition and will only be carried out in extremely rare circumstances. Therefore we do not simulate TMT observations for the bright systems. Even for the fainter system, we have chosen to make the source galaxies

artificially fainter ( $K=26$ ) in order to avoid saturating the central pixels while imaging deep enough to get a sufficient signal to noise ratio on the host galaxy. As will be shown below, the exposure time requirements for TMT are extremely short even in these cases.

## 4 Description of the inference process

In this section, we summarize the process of modeling the mock lens systems. By comparing the distribution of the ratio between the inferred values and the input, we obtain the desired estimate of the statistical precision with which each parameter can be determined. As motivated in the introduction, we are interested primarily in the mass density profile slope  $\gamma'$ . Therefore, we will only show the statistics for that parameter, even though all the parameters are varied simultaneously during the inference. We first describe the image generating process in Section 4.1, and then we describe the inference process in Section 4.2.

### 4.1 Image generating process

Before we fit the mock lens system we need to define the models to describe the source galaxy light, the lens galaxy light, and the lens mass. We start with the surface brightness distribution of the source and lens galaxies. The correspondence between image-plane and source-plane is given via the deflection angles, which in turn depend on the mass distribution of the lens. A PSF is used to convolve the light of lens galaxy, lensed source galaxy and point-source. Finally, we introduce counts noise, readout noise, and background noise. These steps are described in this subsection.

#### 4.1.1 Light Model

We use the Sérsic profile (Sérsic 1963 [?], 1968 [?]) to describe the surface brightness profiles of the sources and the deflectors. The Sérsic profile is described by

$$I(R) = I_e \exp \left[ -k \left( \left( \frac{R}{R_{\text{eff}}} \right)^{\frac{1}{n}} - 1 \right) \right], \quad (4.1)$$

$$R(x, y, q) = \sqrt{qx^2 + y^2/q}. \quad (4.2)$$

The amplitude  $I_e$  is the intensity at the effective radius  $R_{\text{eff}}$ , the Sérsic index  $n$  controls the degree of curvature of the radial light profile, the constant  $k$  is determined so that  $R_{\text{eff}}$  is the half light radius (Ciotti & Bertin 1999 [?]), and  $q$  denotes the axis ratio. Empirically,  $n$  increases with galaxy luminosity, most galaxies being fitted by Sérsic index in the range  $0.6 \leq n \leq 10.0$  (Merritt et al. 2006 [?]). We adopt  $n = 4$  for the deflectors, typical of massive early-type galaxies, and  $n = 4/3$  or  $n = 4$  for the sources, consistent with typical observed values for blue galaxies and AGN hosts. We checked by re-running simulations with different  $n$  for the source that our results do not depend significantly on this model assumption.

#### 4.1.2 Mass Model

The lens mass profile is assigned within a class of power-law models,

$$\Sigma(x, y) = \Sigma_{cr} \frac{2 - \gamma'}{1 + q_m} \left( \frac{\sqrt{X^2 + q_m^{-2} Y^2}}{R_E} \right)^{-\gamma'}, \quad (4.3)$$

where  $q_m$  is the axis ratio,  $\gamma'$  is the radial power-law slope, the  $\{X, Y\}$  principal axes are rotated by the lens position angle w.r.to the canonical  $x$  (increasing to the West) and  $y$  (increasing to the North).

Within this functional family, the deflections scale as  $R^{1-\gamma'}$  with distance from the lens. The *critical density*  $\Sigma_{\text{cr}}$  is defined by

$$\Sigma_{\text{cr}} = \frac{c^2 D_s}{4\pi G D_d D_{ds}} \quad (4.4)$$

in terms of the relative distances to the deflector ( $D_d$ ), to the source ( $D_s$ ), and between the deflector and the source ( $D_{ds}$ ). The Einstein radius  $R_E$  is chosen such that, in the spherical limit ( $q_m = 1$ ), it encloses a mean surface density equal to  $\Sigma_{\text{cr}}$ . This is also the radius of a ring traced by the host of the quasar when this is exactly aligned with the lens galaxy.

Power-law models of elliptical galaxies (Evans 1994 [? ]) have often been used with success over the years to model gravitational lenses. Fast methods to compute 2D deflections from power-law profiles have been given by Barkana 1998 [? ], as a special case of the formalism by Schramm 1990 [? ] for homoeoidal profiles.

Lensed image profiles are obtained by inverse ray shooting. First, each pixel position in the image plane is mapped back to the source plane via its corresponding deflection angle. Then, we exploit the fact that surface-brightness is preserved by lensing to assign surface brightness to each detector pixel in the image plane. Given the smoothly varying surface brightness distribution of the source host galaxy, a rectangular grid in the image-plane is adopted. As discussed in the next section, this approximation would be too crude to represent properly the sharp point source associated with the quasar images, which are therefore added directly to the image plane. Finally, the light models models are convolved with the PSF appropriate for each instrument.

#### 4.1.3 Point Source Model

In order to reduce computation time and complexity, the point source images are added to the light model by adding a PSF with appropriate normalization and position directly to the image plane. The normalization and position of the images for the center of the source in the source plane are computed for each configuration by solving the lens equation using `gravlens` [? ].

#### 4.1.4 Image Noise

In the final step of our simulations we account for the effect of noise arising from the counts statistics, the background, and detector read out. As usual, the variance of noise per pixel is given by

$$\text{Var}_{\text{pix}} = \sqrt{Ct + B \cdot t + N_{\text{read}}R^2} \quad (4.5)$$

where  $C$  is the signal from clean lens system in electrons per second per pixel,  $t$  is the integration time in seconds,  $B_{\text{sky}}$  is the sum of the sky background and detector dark current in electrons per second per pixel,  $N_{\text{read}}$  is the number of detector readouts, and  $R$  is the standard deviation of the read noise in electrons. Realistically, when long integration times are required, we set the number of readouts by requiring no single exposure be longer than a maximum time that depends on each instrument configuration. Considering the definite properties in different types of instruments, we pick 1000.0 seconds as the crude maximum exposure time for space telescopes (HST and JWST), 300.0 seconds for the ground based telescopes (Keck and TMT) appropriately. The cadenced surveys (Euclid, WFIRST, and LSST) have fixed maximum time: 590.0 seconds, 184.0 seconds, and 15.0 seconds, respectively. Examples of simulated images are shown in Figure 3 to 6.

## 4.2 Inferring the parameters



The same code is used to generate the mock models and to fit them. The model fits the following parameters: position, half light radius, axis ratio of light, position angle, and Sérsic index for the

source and deflector brightness model; position, Einstein radius, the axis ratio of the mass distribution, position angle, and power-law slope for the deflector mass model; and QSO positions in the image plane. In practice, variables are created to describe each parameter, including priors. Most priors are taken to be uniform. The prior on the position of the deflector mass is tied to the position of the deflector brightness, which is commonly observed in galaxy scale lenses.

The goodness of fit is assessed through the standard image-plane  $\chi^2$ , by comparing the model and mock data surface brightness profiles. Uncertainties are obtained via MCMC exploration of the posterior. The likelihood is simply given by  $\mathcal{L} \propto \exp[-\chi^2/2]$ .

The point-like images of the quasar are time variable and usually appreciably affected by microlensing, milli-lensing and dust lanes in the foreground, and therefore provide a limited amount of useful information on the differences in gravitational potential. Therefore, we model them in the image plane and do not enforce them to have self-consistent fluxes in the source plane.

The fitting images with different instruments are all  $4'' \times 4''$ , corresponding to the pixel numbers of  $80 \times 80$  for HST,  $130 \times 130$  for JWST,  $400 \times 400$  for Keck (both LGSAO and NGAO),  $1000 \times 1000$  for TMT,  $40 \times 40$  for Euclid,  $36 \times 36$  for WFIRST,  $20 \times 20$  for LSST approximately. The sampling numbers are fixed at 60000 for each inference progress. The accept ratios remain stable between 20% ~ 30%. Based on the parameters above mentioned, the inference is computationally expensive: typical run times range between 15 minutes and 4000 minutes per system on a linux desktop computer. In more specific terms,  $\sim 100$  minutes for HST,  $\sim 150$  minutes for JWST,  $\sim 300$  minutes for Keck (both LGSAO and NGAO),  $\sim 4000$  minutes for TMT,  $\sim 40$  minutes for Euclid and WFIRST,  $\sim 15$  minutes for LSST.

## 5 Results

The goal of this work is to investigate what combinations of instrument/telescope configuration and exposure time produce images of quality sufficient to determine  $\gamma'$  with 2% precision.

For the telescopes/instruments operated in observatory mode, we are able to set the exposure time to any desired level. The minimum exposure time required is defined as the “target” exposure time. In addition, in order to quantify how the precision on  $\gamma'$  depends on exposure time, we also simulated images with  $1/3$  and  $3\times$  the exposure time. This is meant to provide useful guidance for designing future experiments. However, for some combination of lens brigtheness and telescope/instrument **configuratio**, the target exposure time is too short to be adopted in realistic observations. In practice, therefore these lenses will either not be observed with this setup or be observed for longer exposure time. Taking into account typical overheads for pointing and acquisition we evaluated that the *brighter* lenses are too bright to be observed with TMT, while a minimum exposure time of 360s gives more than enough signal-to-noise ratio for the fainter systems. Likewise the brighter **systmes** are likely too bright to be observed with JWST. We still simulated them, but set the minimum exposure time to 60s.

For the telescopes/instruments to be operated in survey mode, we simulated the planned exposure time, and considered the specific question of whether the ensuing data quality is sufficient or not to meet the requirement.

Using the methods summarized in the previous section we generate 10 mock systems for each configuration, and carry out the full inference for each one using the python Markov Chain Monte Carlo sampler `pymc`. The results are summarized in Figures 7 to 10. Table 5 summarizes the “target” exposure times for HST, JWST, Keck LGSAO and NGAO, i.e. those required in order to reach 2% uncertainties on  $\gamma'$ .

First of all, we find that HST can deliver the desired image quality with exposures of a few ks at most, consistent with published work [? ]. Conversely, it seems that current AO capabilities – even setting aside the difficulties associated with the reconstruction of the PSF – are only sufficient to study the brighter systems, for reasonable exposure times. The improvement with NGAO will be substantial, cutting integration times down to HST-like for both the bright and the faint lenses. In this case, the higher background than from space is compensated by the advantage of having smaller detector pixels, and thus better sampling of the PSF, and a larger telescope aperture.

For TMT we concluded that 360s of exposure will be more than sufficient for all systems considered here and therefore TMT is likely to be used either to take a brief image before taking a spectrum for determining the deflector velocity field and source redshift, or to follow-up fainter systems, yet to be discovered. Similary, JWST can obtain images of the desired quality in just a few minutes, and therefore it is likely to be used mostly to take images before taking a spectrum or to follow-up fainter systems.

For the surveys, the conclusion is that while the Euclid survey seems to have insufficient depth and resolution to meet our requirements on the targets considered here, LSST and WFIRST should be sufficient at least for the brighter ones. Of course, these surveys will cover vast fractions of the sky and will therefore provide at no additional cost useful images for all the brighter systems [? ]. Those will be a great complement to the deeper pointed observation obtained with the other telescopes. WFIRST is planned to have a component of GO program, so it could meet the imaging requirements by integrating longer than in the survey mode. In fact, WFIRST is expected to have imaging performance superior to that of WFC3 on board HST, which has already been used to perform studies of time delay lenses, using integration times of a few orbits per system ( $\sim 10\text{ks}$ ; HST-GO-12889, PI: Suyu).

## 6 Summary

Gravitational time delays are a powerful tool for cosmography. Transforming measured time delays into time distances requires modeling the gravitational potential of the main deflector with sufficient accuracy. Recent work has shown that this is possible provided that images of sufficient resolution and signal to noise ratio are available. In anticipation of the hundreds or thousands of lensed quasars that are going to be discovered in the next decade from ground based low resolution imaging surveys [? ], we study the requirements for high resolution imaging follow-up. We consider a range of instrument/telescope configurations spanning from currently available systems (HST/ACS and Keck/LGSAO/NIRC2) to planned observatories and surveys from space (JWST/NIRCAM, Euclid, WFIRST/HLS) and from the ground (Keck/NGAO/NIRC2, TMT/IRIS). In order to carry out a generic and systematic comparison across multiple telescopes we simulate four realistic lens systems, spanning the range of magnitudes and configuration expected for lenses to be discovered in the next decade. We also define a single metric, which is the ability to measure the slope of the mass density profile of an elliptical power law mass density profile. In order to guarantee that the images be sufficient to constrain the Fermat potential to a level so that the random uncertainties are subdominant with respect to other sources of uncertainty, we consider it necessary for them to contain enough information to constrain  $\gamma'$  with 2% precision or better. We find that:

- HST can provide sufficient information with integration times of order 1-10ks, as shown by recent work based on real data

- Keck/LGSAO can provide sufficient information for only the brighter systems to be discovered. The critical issue will be the accuracy with which the PSF can be reconstructed. Significant effort is under way to overcome this **critical** obstacle [? ].
- The planned Keck/NGAO system will improve the performance of Keck to make it comparable to that of HST in terms of resolution and signal to noise ratio. **provided** that the PSF can be reconstructed with sufficient accuracy.
- JWST/NIRCAM and TMT/IRIS will be able to deliver the desired image quality with very short exposure times for any lens for which time delays are reasonably going to be available. Thus they will be ideal instruments for follow-up of time-delay lenses, especially if short imaging are followed by deeper spectroscopic observations to measure the redshift of the source and the velocity dispersion and redshift of the deflector. For TMT, and other ELTs, the ability to reconstruct accurately the PSF of the AO system [?] will be critical for time delay cosmography.
- The survey modes of Euclid/WFIRST/LSST will likely be too shallow and/or have insufficient resolution except for the very brightest systems in the sky. Dedicated follow-up with WFIRST in GO mode will be necessary for the fainter gravitational lens systems.

## Acknowledgments

We are grateful to F.Courbin, T. Do, R. Gavazzi, C. Hirata, J. Larkin, M. Meneghetti, P. Wizinowich, and S.A. Wright for useful conversations about planned instruments and telescopes. AA, and TT acknowledge support from NSF grant AST-1450141 “Collaborative Research: Accurate cosmology with strong gravitational lens time delays”. AA, and TT gratefully acknowledge support by the Packard Foundation through a Packard Research Fellowship to TT. The work of PJM was supported by the U.S. Department of Energy under contract number DE-AC02-76SF00515.

[...]

**Table 1.** Telescope and Instrument Properties

Telescope	Instrument	Filter	Zero Point	Readout Noise (e <sup>-</sup> /pixel)	Background Noise (e <sup>-</sup> /pixel/s)	Pixel Size (arcsec)
HST	ACS	F814W	25.94	4.20	0.11	0.050
JWST	NIRCAM	F200W	27.85	9.00	0.20	0.032
LGSАО	NIRC2	<i>K'</i>	28.04	5.75	26	0.010
NGAO	NIRC2	<i>K'</i>	28.04	5.75	26	0.010
TMT	IRIS	<i>K'</i>	31.10	2.00	21	0.004
Euclid	VIS	<i>r + i + z</i>	25.58	4.50	0.43	0.100
WFIRST	DRM(?)	F184	26.18	5.00	0.11	0.110
LSST	—	<i>I</i>	28.35	5.00	68	0.200

Observational facilities (telescopes and instruments) considered in this work. Zero points are given in the ABmag system. The VIS imager of Euclid spans the whole *r + i + z* wavelength range. The Keck 10-m telescope is considered both with the current (LGSАО) and next-generation (NGAO) adaptive optics capabilities.

**Table 2.** Surface Brightness Profile Models

Lens Name	$R_{\text{eff}}$ (arcsec)	$q$	P.A. (deg)	$n$	$\Delta x$ (arcsec)	$\Delta y$ (arcsec)	$m_I$	$m_K$	$m_{\text{VIS}}$	$m_H$
Parameters for the source										
fainter system <sup>a</sup>	0.23	0.92	54.0	4.0	0.0662	-0.167	25.0	25.0	25.0	25.0
fainter system <sup>b</sup>	0.23	0.92	54.0	4.0	0.008	0.298	25.0	25.0	25.0	25.0
brighter system <sup>b</sup>	0.12	0.77	120.0	1.33	-0.195	0.34	22.73	22.0	23.46	22.0
brighter system <sup>a</sup>	0.12	0.77	120.0	1.33	0.01	-0.005	22.73	22.0	23.46	22.0
Parameters for the deflector										
fainter system	1.76	0.61	-9.6	4.0	—	—	20.69	19.7	21.13	19.7
brighter system	0.91	0.81	113.2	4.0	—	—	17.99	16.5	18.84	16.5

$R_{\text{eff}}$  is the half light radius.  $q$  denotes the axis ratio. P.A. is with respect to the x-axis. The Sérsic index  $n$  controls the degree of curvature of the galaxy light profile. Magnitudes  $m$  are given in the ABmag system. The only difference among systems with sources the same brightness is in the source-position, which is set to produce either two or four images of the center.

<sup>a</sup> This configuration yields 4 QSO images.

<sup>b</sup> This configuration yields 2 QSO images.

**Table 3.** Lens Model Parameters

Lens Name	$z$	$R_{\text{Ein}}$ (arcsec)	$q$	P.A. (deg)	$\gamma'$
fainter system	0.783	1.14	0.6	14.7	2.0
brighter system	0.351	1.1	0.81	113.2	2.0

$R_{\text{Ein}}$  is the Einstein radius, according to the definition by Kormann+94.  $q$  denotes the axis ratio of the mass distribution. P.A. is anti-clockwise from the x-axis. The model deflector is a Singular Isothermal Ellipsoid ( $\gamma'=2$ ), which is then fit using power-law models with variable  $\gamma'$ .

**Table 4.** Point Source Parameters

	mag(I)	mag(K)	mag(VIS)	mag(H)	$\mu$ (I)	$\mu$ (II)	$\mu$ (III)	$\mu$ (IV)
fainter system <sup>a</sup>	24.0	24.0	24.0	24.0	2.84	3.75	6.59	2.78
fainter system <sup>b</sup>	24.0	24.0	24.0	24.0	2.78	2.48	—	—
brighter system <sup>b</sup>	21.7	21.5	22.9	21.5	1.22	6.10	—	—
brighter system <sup>a</sup>	21.7	21.5	22.9	21.5	6.09	8.22	6.86	7.53

Magnitudes (in the ABmag system) and magnifications of multiple images. The source magnitude ‘mag’ is given in the source plane so that the observed magnitude in the image plane is given by  $\text{mag} - 2.5 \log_{10} \mu$ , where  $\mu$  is the magnification of each image.

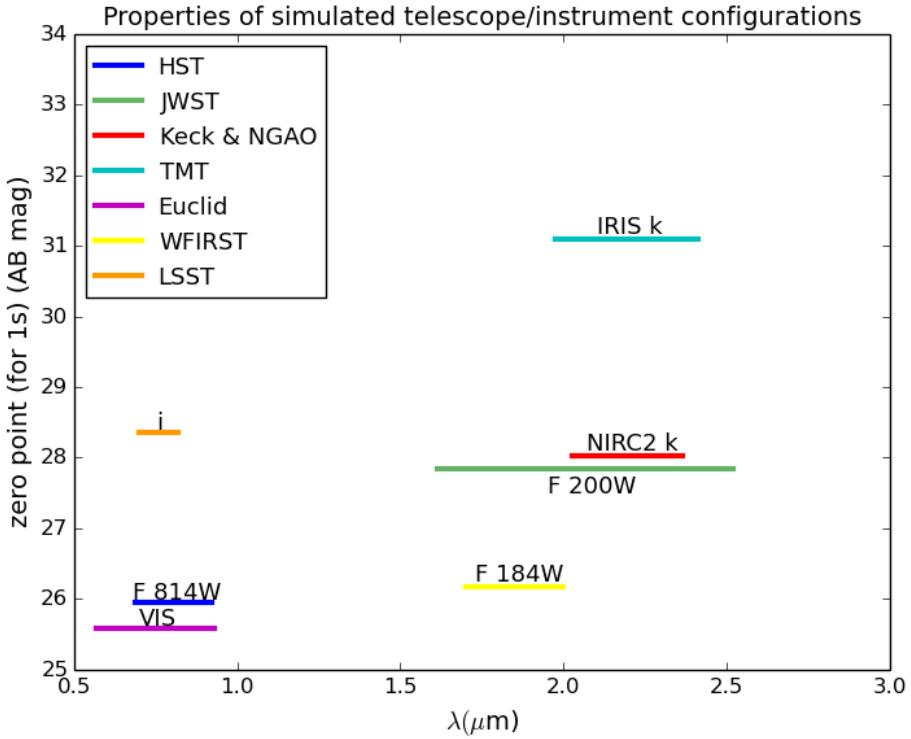
<sup>a</sup> This configuration yields 4 QSO images.

<sup>b</sup> This configuration yields 2 QSO images.

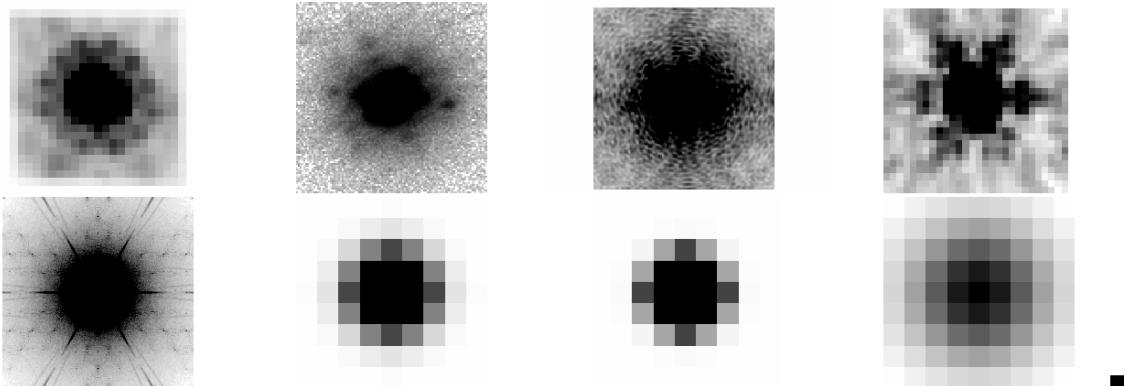
**Table 5.** Exposure time requirements

Instrument	double		quad	
	faint	bright	faint	bright
HST	$6 \times 10^3$ s	360 s	$3 \times 10^3$ s	150 s
JWST	690s	360 s	210s	<60 s
Keck (LGSAO)	$105 \times 10^3$ s	3600 s	$75 \times 10^3$ s	2400 s
Keck (NGAO)	$18 \times 10^3$ s	180 s	$3 \times 10^3$ s	150 s

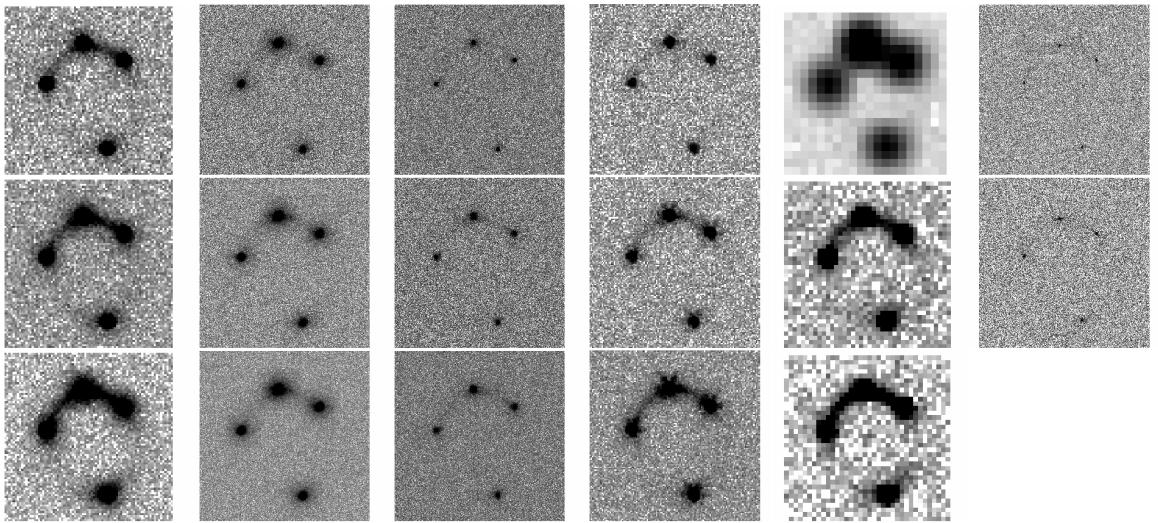
Exposure times required to attain a 2% precision on the inferred  $\gamma'$ . With 360 s of exposure, the TMT already enables percent-level precision. The Euclid, WFIRST and LSST surveys, are not quoted here because their exposure times are fixed. The default exposure times are sufficient to meet the 2% requirement in the case of LSST and WFIRST, but not for Euclid.



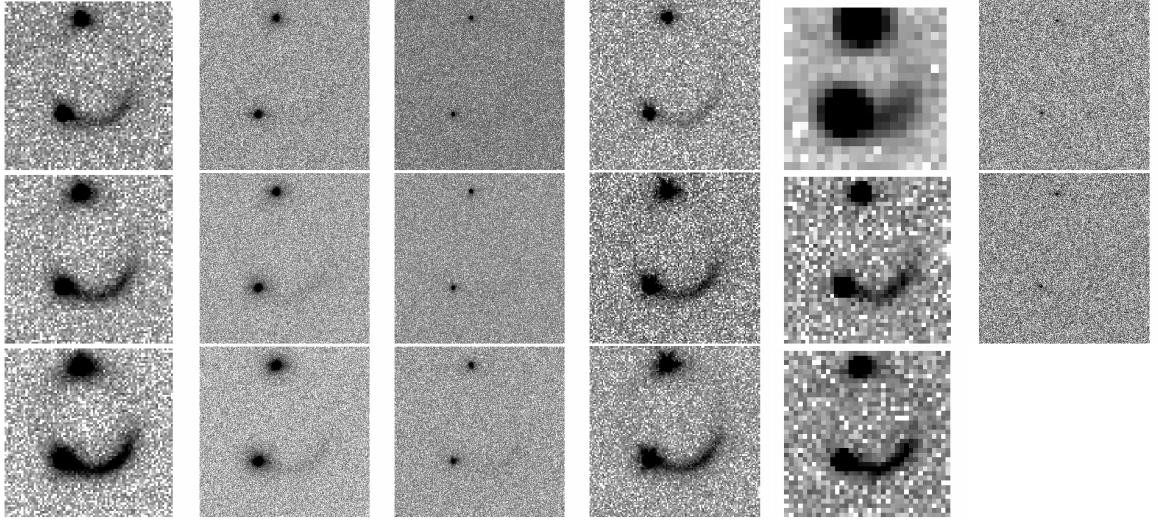
**Figure 1.** Zero points in AB magnitudes for HST/ACS (blue), JWST/NIRCAM (green), Keck NIRC2 (assumed for both LGSАО and NGAO; red), TMT/IRIS (cyan), Euclid (magenta), WFIRST (yellow) and LSST (orange), corresponding to one count per second. The colored bars indicate the wavelength range of each instrumental setup considered in this work.



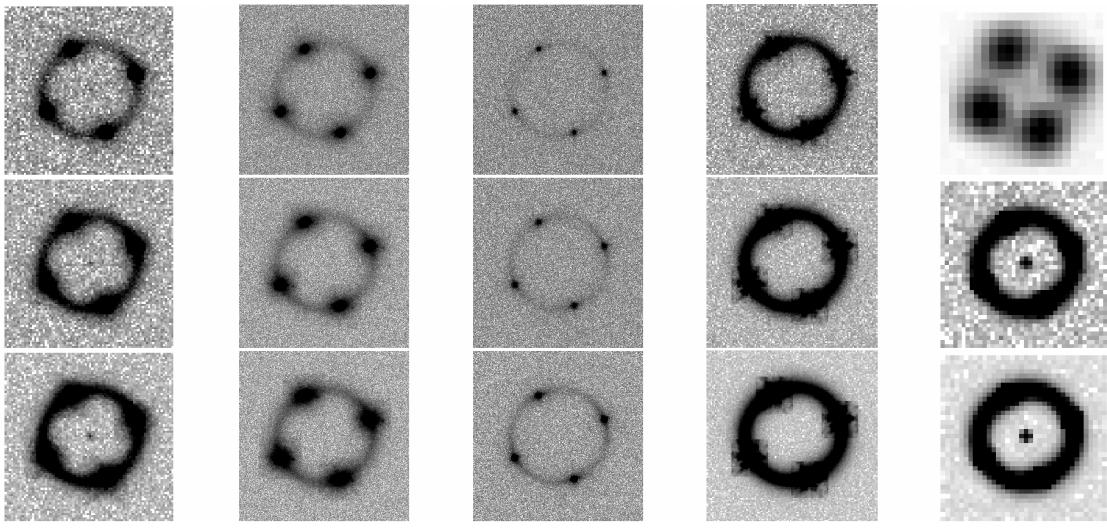
**Figure 2.** Montage of the point spread functions (PSFs) of each instrument. The upper row, from left to right, represents HST/ACS, Keck/NIRC2+LGSАО, Keck/NIRC2+NGAO, and JWST/NIRCAM. The lower row, from left to right, shows TMT/IRIS, Euclid, WFIRST, and LSST, respectively. Observed or simulated PSFs are used for HST, JWST, Keck (LGSАО & NGAO), and TMT. Gaussian PSFs are adopted for the other three survey instruments.



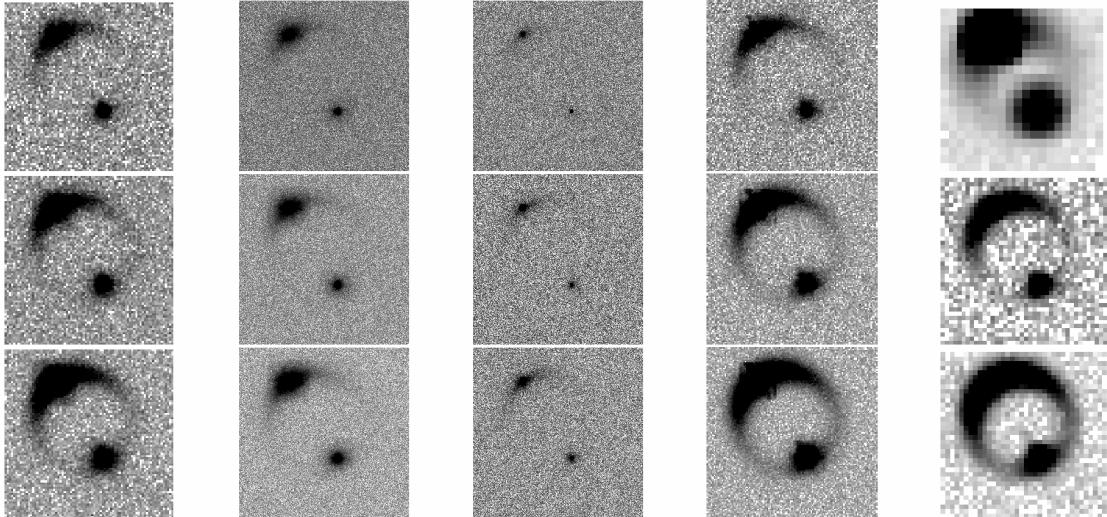
**Figure 3.** Simulations of the fainter lens system with 4 QSO images. The simulated images are all  $4'' \times 4''$ . The first 4 columns, from left to right, represent HST/ACS, Keck/NIRC2+LGS, Keck/NIRC2+NGAO, and JWST/NIRCAM; from top to bottom, the rows correspond to  $1/3 \times$  “target” exposure time, “target” exposure time, and  $3 \times$  “target” exposure time. “Target” exposure time is defined as the exposure time that yields 2% precision on the slope of the mass density profile of the mass model  $\gamma'$ . The fifth column shows simulations of 3 surveys. From top to bottom we show LSST, Euclid, and WFIRST, with the default survey exposure times (4500s, 2360s, 920s, respectively). The last column shows simulations for TMT with 2 fixed exposure times of 360 seconds and 1080 seconds.



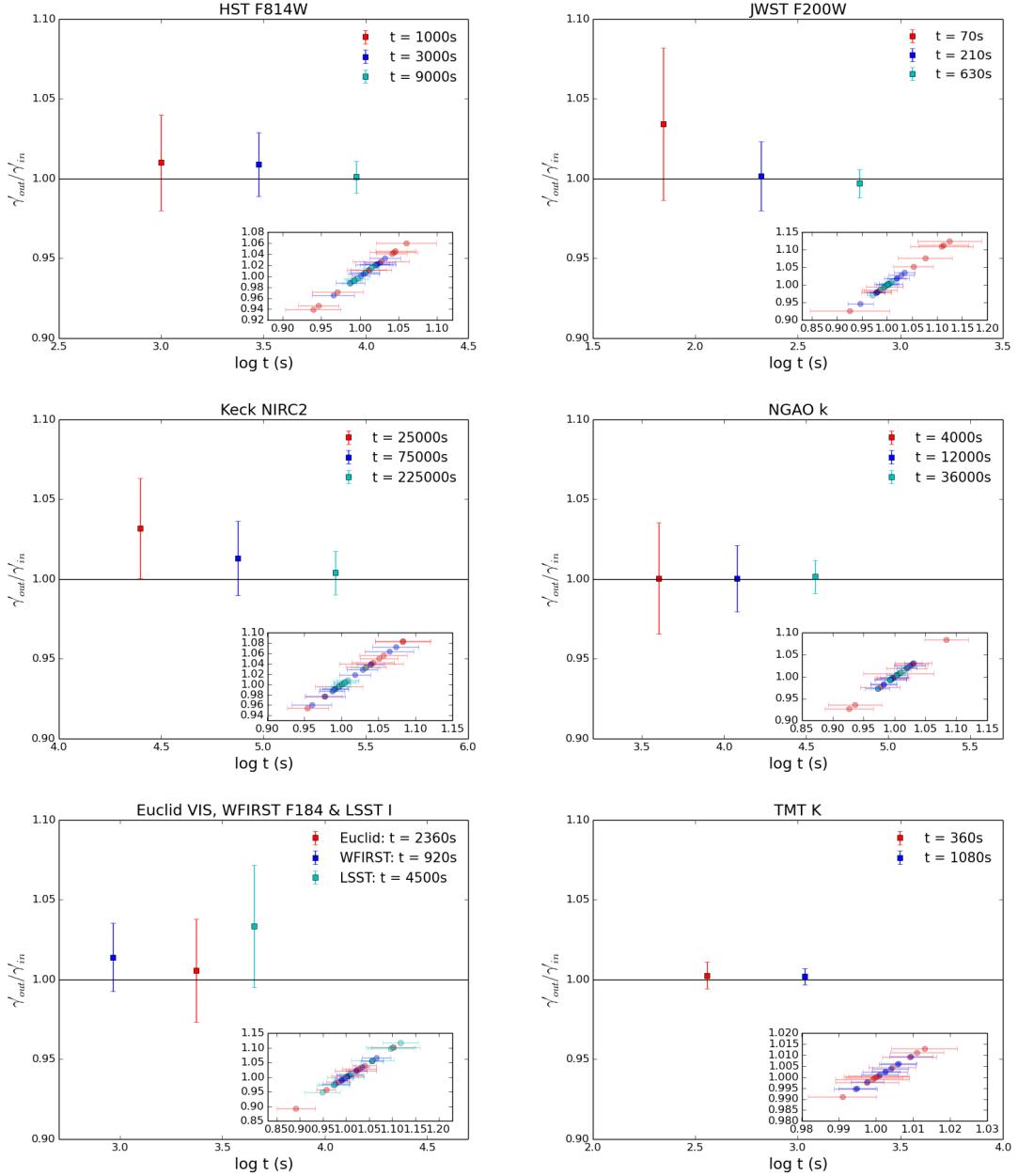
**Figure 4.** Same as Fig. 3, for the faint double system.



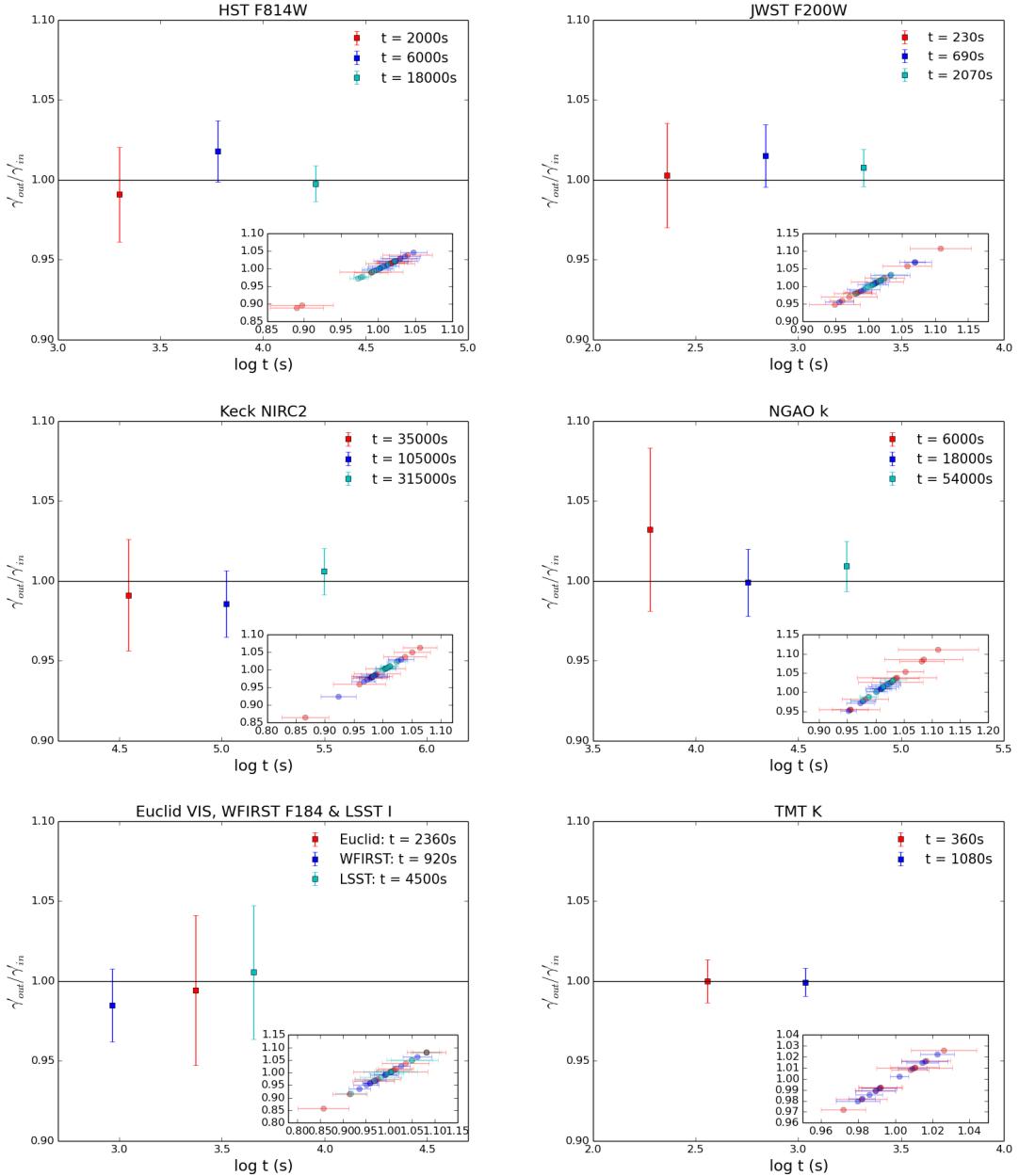
**Figure 5.** Simulations of the bright lens system with 4 QSO images. The simulated images are all  $4'' \times 4''$ . The first 3 columns, from left to right, represent HST/ACS, Keck/NIRC2+LGSAO, Keck/NIRC2+NGAO; from top to bottom, the rows correspond to  $1/3 \times$  "target" exposure time, "target" exposure time, and  $3 \times$  "target" exposure time. "Target" exposure time is defined as the exposure time that yields 2% precision on the slope of the mass density profile of the mass model  $\gamma'$ . The fourth column shows JWST simulations with 3 fixed exposure times: 60, 180, 540 seconds, from top to bottom. The fifth column shows simulations of 3 surveys. From top to bottom we show LSST, Euclid, and WFIRST, with the default survey exposure times (4500s, 2360s, 920s, respectively). TMT simulations are not shown since the system is considered too bright to be observed with this telescope in practice. **XLM to remake figure after masking black spot artifacts at the center.**



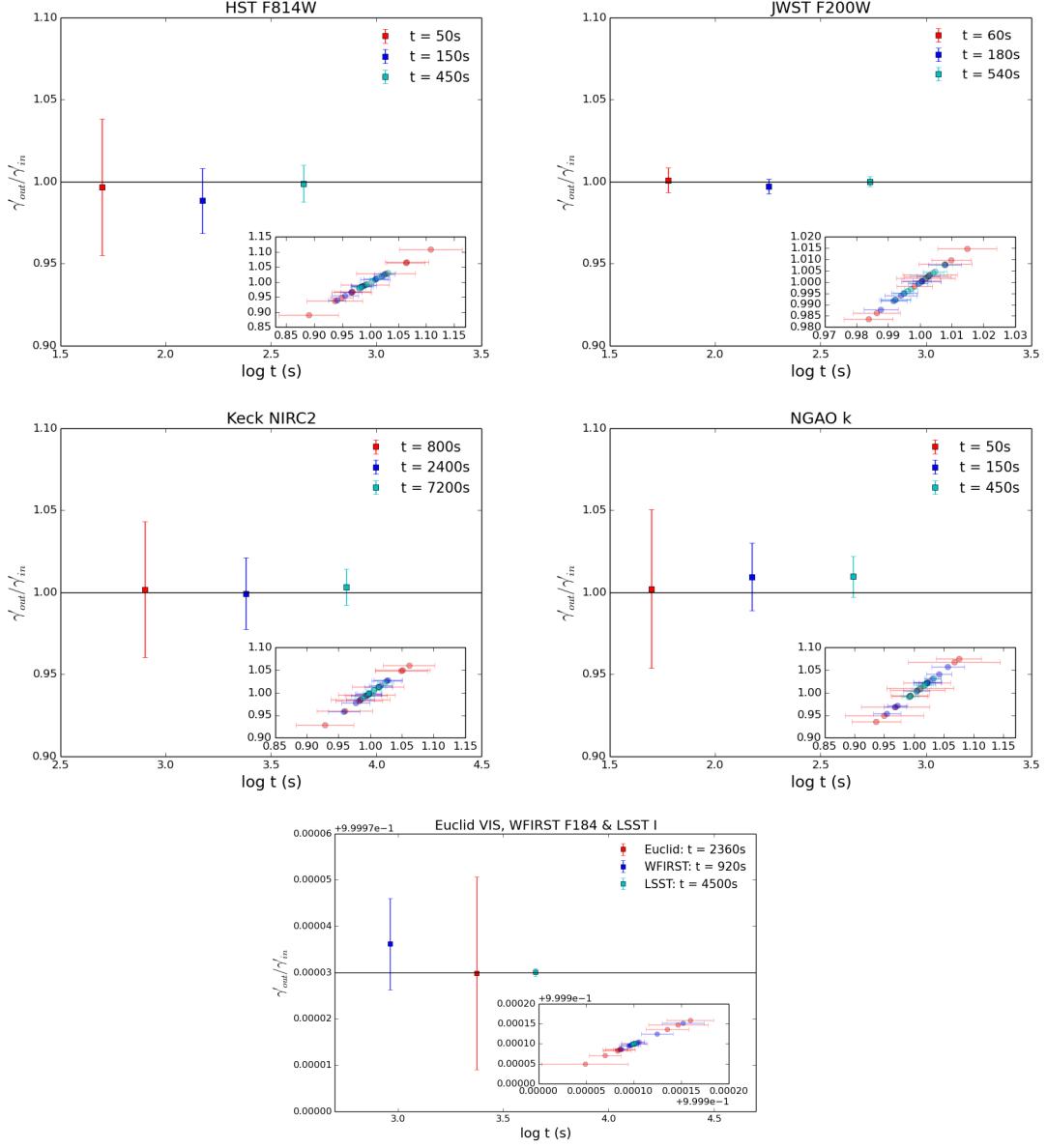
**Figure 6.** Same as Fig. 5, for the bright double imaged system.



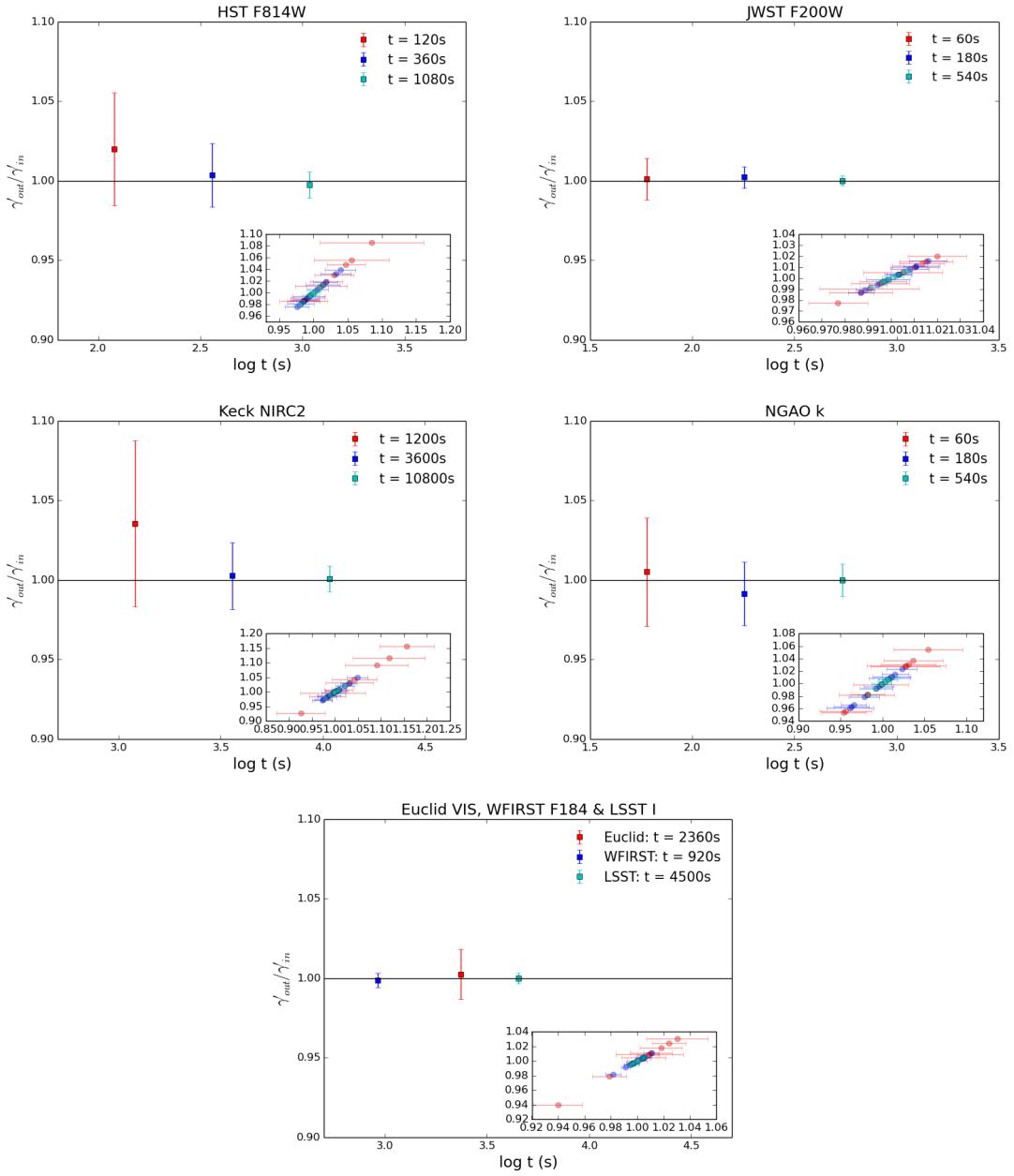
**Figure 7.** Precision on the mass density profile slope  $\gamma'$  as a function of exposure time. This figure shows the results for the fainter lens system with 4 QSO images. The quantity  $\gamma'_{in}$  is the input SIE mass slope. The quantity  $\gamma'_{out}$  is the output of the inference process. The insert in each panel shows all 10 simulation results for each exposure time with the same color coding. Note that both axes represent  $\gamma'_{out}/\gamma'_{in}$  whereas error bars are only shown on the x-axis for clarity.



**Figure 8.** Same as Fig. 7 for the fainter lens system with 2 images.



**Figure 9.** Same as Fig. 7 for the brighter lens system with 4 images.



**Figure 10.** Same as Fig. 7 for the brighter lens system with 2 images.