Prepared for submission to JCAP

Measuring angular diameter distances of strong gravitational lenses

I. Jee[a], **E. Komatsu**[a,b], **S. H. Suyu**[c]

[a]Max-Planck-Institut für Astrophysik, Karl-Schwarzschild Str. 1, 85741 Garching, Germany
[b]Kavli Institute for the Physics and Mathematics of the Universe, Todai Institutes for Advanced Study, the University of Tokyo, Kashiwa, Japan 277-8583 (Kavli IPMU, WPI)
[c]Institute of Astronomy and Astrophysics, Academia Sinica, P.O. Box 23-141, Taipei 10617, Taiwan

E-mail: ijee@mpa-garching.mpg.de

Abstract. The distance-redshift relation plays a fundamental role in constraining cosmological models. In this paper, we show that measurements of positions and time delays of strongly lensed images of a background galaxy, as well as those of the velocity dispersion and mass profile of a lens galaxy, can be combined to extract the angular diameter distance of the lens galaxy. Physically, as the velocity dispersion and the time delay give a gravitational potential (GM/r) and a mass (GM) of the lens, respectively, dividing them gives a physical size (r) of the lens. Comparing the physical size with the image positions of a lensed galaxy gives the angular diameter distance to the lens. A mismatch between the exact locations at which these measurements are made can be corrected by measuring a local slope of the mass profile. We expand on the original idea put forward by Paraficz and Hjorth, who analyzed singular isothermal lenses, by allowing for an arbitrary slope of a power-law spherical mass density profile, an external convergence, and an anisotropic velocity dispersion. We find that the effect of external convergence cancels out when dividing the time delays and velocity dispersion measurements. We derive a formula for the uncertainty in the angular diameter distance in terms of the uncertainties in the observables. As an application, we use two existing strong lens systems, B1608+656 ($z_{\rm L} = 0.6304$) and RXJ1131−1231 ($z_{\rm L} = 0.295$), to show that the uncertainty in the inferred angular diameter distances is dominated by that in the velocity dispersion, σ^2, and its anisotropy. We find that the current data on these systems should yield about 16% uncertainty in D_A *per object*. This improves to 13% when we measure σ^2 at the so-called sweet-spot radius. Achieving 7% is possible if we can determine σ^2 with 5% precision.

Contents

1 Introduction

Individual strong gravitational lens systems can be used to measure cosmological parameters via a combination of the cosmological distances [1–5]. Recently, a particular combination of the distances called the "time-delay distance" of strongly lensed time delay systems has yielded precise determinations of the Hubble constant [6–8]. The time-delay distance is the angular diameter distance to the lens from Earth, $D_A(EL)$, multiplied by the distance to the source, $D_A(ES)$, divided by the distance between the lens and the source, $D_A(LS)$. While this combination is sensitive to the Hubble constant, it is less so to the other cosmological parameters than $D_A(EL)$ itself [9]. See figure 1 for the definition of these distances.

To extract more cosmological information acquirable from the strong lens time delay systems, Paraficz and Hjorth [10] have shown that, by assuming the density profile of the lens galaxy, one can obtain $D_A(EL)$ from time-delay lenses. The basic physics behind this idea is simple: the velocity dispersion gives the depth of the potential at the point where it is measured, and the time delay gives the mass of the lens galaxy enclosed within the position at which images are formed. Thus, dividing them gives the physical size of the system. We can then estimate $D_A(EL)$ by dividing the physical size by the angular separation of lensed

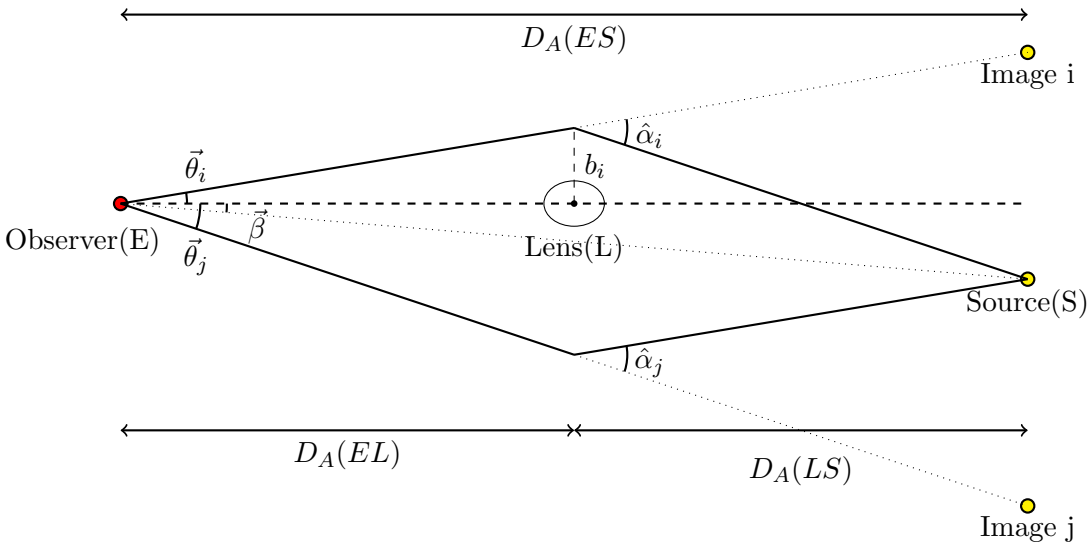

Figure 1: Configuration of a strong lens system, with definition of the variables used throughout this paper. All angles are measured with respect to the center of the lens galaxy; $\vec{\theta}$ is the angular position of the image; $\vec{\beta}$ is the angular position of the source in the absence of the lens; $\vec{\alpha}$ is the scaled deflection angle; $\hat{\alpha}$ is the deflection angle at the lens plane; and $\vec{b}$ is the physical separation to the closest approach at the lens plane.

image positions. Their analysis was limited to the singular isothermal sphere (SIS) density profile, as well as to an isotropic velocity dispersion. In this paper, we show that this simple physical picture holds even when we extend the analysis by including an arbitrary power-law profile, the effect of external convergence, and an anisotropic velocity structure. We show explicitly how to extract $D_A(EL)$ from the observational data, and provide an estimate of its associated uncertainty.

The rest of the paper is organized as follows. In section 2, we present the basic idea using a simplified SIS model, following ref. [10]. In section 3, we expand on ref. [10] by allowing for an arbitrary slope of a power-law spherical mass density profile and external convergence. In section 4, we derive an analytical formula relating the uncertainty in $D_A(EL)$ to the uncertainties in the observable quantities, and apply the formula to the observed strong lens time delay systems, B1608+656 and RXJ1131−1231. In section 5, we use Monte-Carlo simulations to study the effect of anisotropic velocity dispersion on the uncertainty in $D_A(EL)$. We conclude in section 6. In the appendix, we show how General Relativity allows us to calculate the deflection angle at the lens plane.

2 Basics of the analysis

2.1 The idea: a simple analysis using singular isothermal spheres

We review the basic idea with the simplest case in which the mass density profile of a lens galaxy is given by an SIS. This case has been worked out by Paraficz and Hjorth in 2009 [10]. The density distribution of an SIS lens, $\rho_{\rm SIS}$, is given by

$$\rho_{\rm SIS}(r) = \frac{\sigma^2}{2\pi G r^2}, \tag{2.1}$$

where σ^2 is the three-dimensional isotropic velocity dispersion. The Einstein ring radius, θ_E, is related to σ^2 via

$$\sigma^2 = \theta_{\rm E}\frac{c^2}{4\pi}\frac{D_A(ES)}{D_A(LS)}. \tag{2.2}$$

Clearly, the relation between the two observable quantities, $\theta_{\rm E}$ and σ, depends on the distance *ratio*.

To extract the actual angular diameter distance to the lens, $D_A(EL)$, instead of the ratio, we need to include the lensing time delay [11]. The presence of intervening mass between the observer and the source, usually galaxies and/or clusters of galaxies, causes two different components on time delay: the *geometrical time delay* and the *potential time delay*. Strongly lensed systems show multiple images as photons coming from the source take different paths: images are located at the closest approach to the lens of each path. The geometrical part of the time delay is caused by the fact that the total path lengths differ, while the potential part is caused by the difference in the depths of potential at each image position of the path.

In a SIS lens, the time delay between two images can be written as

$$\Delta t_{i,j} \equiv t_i - t_j = \frac{1+z_{\rm L}}{2c}\frac{D_A(EL)D_A(ES)}{D_A(LS)}(\theta_j^2 - \theta_i^2), \tag{2.3}$$

where θ_i is the angular separation between the i-th image and the center of the lens galaxy, and t_i is the absolute time delay of the i-th image, i.e., the delay in comparison to the case where the lens is absent [12]. The light ray that comes closer to the lens (i.e., smaller θ_i) is deflected more, resulting in a larger time delay. The distance ratio that appears in this relation is the time-delay distance, $D_{\Delta t} \equiv (1+z_{\rm L})D_A(EL)D_A(ES)/D_A(LS)$, which depends primarily on H_0 and has a limited sensitivity to the other cosmological parameters, such as the equation of state of dark energy.

Remarkably, when we combine the above equation with equation (2.2) and $\theta_E = (\theta_i + \theta_j)/2$, we obtain the angular diameter distance to the lens:

$$D_A(EL)(\theta_j - \theta_i) = \frac{c^3\Delta t_{i,j}}{4\pi\sigma^2(1+z_{\rm L})}. \tag{2.4}$$

The physical interpretation of the above analysis is as follows: the velocity dispersion is determined by the gravitational potential of the lens, GM/r. The time delay gives the mass of the lens system, GM, and thus dividing them gives the physical size of the system, r. Since the angular scale of the system is directly observable via $\theta_j - \theta_i$, one can estimate the angular diameter distance to the lens. Equation (2.4) indeed gives the angular diameter distance as $D_A(EL) \propto \Delta t_{i,j}/[\sigma^2(\theta_j - \theta_i)]$; thus, the uncertainty in $D_A(EL)$ is given by the quadrature sum of the uncertainties in the time delay, velocity dispersion, and image position measurements.

As the velocity dispersion uncertainty is usually the biggest of all uncertainties, the uncertainty in $D_A(EL)$ is expected to be dominated by the velocity dispersion uncertainty. The goal of this paper is to extend this analysis to more general lenses.

2.2 Lensing theory and equations

Before we proceed, let us review some of general equations for strong lensing, following ref. [7]. Let the angular position of the image be $\vec{\theta}$ and that of the source be $\vec{\beta}$, as shown in fig. 1.

The absolute time delay can be written as

$$t(\vec{\theta},\vec{\beta}) = \frac{1}{c}(1+z_{\rm L})\frac{D_A(EL)D_A(ES)}{D_A(LS)}\phi(\vec{\theta},\vec{\beta}), \tag{2.5}$$

where ϕ is the so-called Fermat potential, which is defined as

$$\phi(\vec{\theta},\vec{\beta}) \equiv \frac{(\vec{\theta}-\vec{\beta})^2}{2} - \psi(\vec{\theta}). \tag{2.6}$$

The first and the second terms in equation (2.5) are geometrical and potential time-delay terms, respectively. Here, ψ is the lens potential, which is calculated as

$$\psi(\vec{\theta}) = \frac{1}{\pi}\int d^2\theta' \kappa(\vec{\theta'}) \ln|\vec{\theta}-\vec{\theta'}|, \tag{2.7}$$

where the lensing convergence field, κ, is defined by

$$\kappa(\vec{\theta}) \equiv \frac{\Sigma(\vec{\theta})}{\Sigma_{\rm cr}}. \tag{2.8}$$

The projected surface mass density, Σ, is

$$\Sigma(\vec{\theta}) = \int_{-\infty}^{\infty} \rho[D_A(EL)\vec{\theta},\ell]\ d\ell, \tag{2.9}$$

where ℓ denotes the line-of-sight coordinate, and

$$\Sigma_{\rm cr} \equiv \frac{c^2}{4\pi G}\frac{D_A(ES)}{D_A(EL)D_A(LS)}, \tag{2.10}$$

is the critical surface mass density. Physically, when $\kappa > 1$, the system satisfies the sufficient condition to form multiple images.

The absolute time delay, t, is not an observable as we cannot directly observe the source without the lens, or the time difference between lensed and un-lensed images. However, if we have multiple images, we can compare the relative time delay between image pairs to calculate the time delay between two (or more) lensed images. Also, ϕ can be modeled to satisfy observational constraints such as image positions, flux ratios and time-delay differences between multiple pairs of images; thus, we can obtain the time-delay distance.

In a differential form, the lens potential is related to the convergence field via

$$\kappa(\vec{\theta}) = \frac{1}{2}\nabla^2\psi(\vec{\theta}), \tag{2.11}$$

where ∇ is a derivative in $\vec{\theta}$ coordinates. Now we can write the lens equation which relates the observed image position to the source position in terms of the lens potential,

$$\vec{\theta} - \vec{\beta} = \nabla\psi(\vec{\theta}) = \vec{\alpha}, \tag{2.12}$$

where $\vec{\alpha}$ is the scaled deflection angle.

3 More realistic lenses

The analysis in section 2.1 assumes the simplest possible lens system: an SIS density profile with an isotropic velocity dispersion. While the SIS profile is widely used to model lens galaxies and is considered as a good approximation, several studies have shown that slopes of density profiles of individual galaxies show a non-negligible scatter from the SIS [13–16]. In this section, we consider an arbitrary power-law density profile (section 3.1) to show that, in such a model, we can still extract $D_A(EL)$ from $\Delta t_{i,j}$, σ^2, and image positions. We then show that the external convergence cancels out (section 3.2).

3.1 Arbitrary slope of the lens mass profile

We allow the mass density of a lens to follow a general power-law:

$$\rho = \rho_0 \left(\frac{r}{r_0}\right)^{-\gamma}. \tag{3.1}$$

The lens potential also has a power-law form, $\psi(\theta) \propto \theta^l$, with $l = 3-\gamma$. The scaled deflection, $\vec{\alpha}$, which is given by $\nabla\psi = \vec{\alpha}$, and the lens equation, $\vec{\beta} = \vec{\theta} - \vec{\alpha}$, gives

$$\psi = \frac{1}{l}\ \vec{\theta}\cdot(\vec{\theta}-\vec{\beta}). \tag{3.2}$$

Using this result in equation (2.5), we obtain the time delay between two images as

$$\Delta t_{i,j} = \frac{1+z_{\rm L}}{2c}\frac{D_A(EL)D_A(ES)}{D_A(LS)}\left\{(\vec{\theta_i}-\vec{\beta})^2 - (\vec{\theta_j}-\vec{\beta})^2 - \frac{2}{l}\left[\vec{\theta_i}\cdot(\vec{\theta_i}-\vec{\beta}) - \vec{\theta_j}\cdot(\vec{\theta_j}-\vec{\beta})\right]\right\}. \tag{3.3}$$

From the geometry of the system, the lens equation and the definition of the angular diameter distance, the following relation between $\vec{\theta}$, $\vec{\beta}$, and $\hat{\alpha}$ holds:

$$\vec{\theta} - \vec{\beta} = \vec{\alpha} = \frac{D_A(LS)}{D_A(ES)}\hat{\alpha}, \tag{3.4}$$

where $\hat{\alpha}$ is the deflection angle at the lens plane. We substitute $\vec{\theta}-\vec{\beta}$ in equation (3.3) for $\hat{\alpha}$, and write

$$\Delta t_{i,j} = D_A(EL)\frac{(1+z_{\rm L})}{2c}\left[(\hat{\alpha}_i+\hat{\alpha}_j)\cdot(\vec{\theta_i}-\vec{\theta_j}) - \frac{2}{l}(\vec{\theta_i}\cdot\hat{\alpha}_i - \vec{\theta_j}\cdot\hat{\alpha}_j)\right]. \tag{3.5}$$

The remaining task is to relate $\hat{\alpha}$ to observables. As the potential of a spherically symmetric system only has a radial component with respect to the center, $\vec{\alpha}$, $\vec{\beta}$ and $\vec{\theta}$ have only radial components. Let us define $\alpha \equiv |\hat{\alpha}|$, which is the magnitude of the deflection angle at the lens plane. Under the power-law density profile model, α is given by

$$\alpha(b) = \frac{2GM(b)}{c^2 b}F(\gamma) \propto b^{-\gamma+2}, \tag{3.6}$$

where b is the physical separation between the lens and the point of the closest approach of the light ray, and

$$F(\gamma) \equiv \frac{\sqrt{\pi}\Gamma\left[\frac{1}{2}(-1+\gamma)\right]}{\Gamma(\frac{\gamma}{2})}. \tag{3.7}$$

The derivation of this formula is given in appendix A.

Using the virial theorem, we obtain the radial velocity dispersion at a given radius r as

$$\sigma_r^2(r) = \frac{1}{2(\gamma-1)}\frac{GM(r)}{r} \propto r^{-\gamma+2}. \tag{3.8}$$

If the velocity dispersion is isotropic, $\sigma_r^2(r) = \frac{1}{3}\sigma^2(r)$, and the radial velocity dispersion is the same as the line-of-sight velocity dispersion, which is observable. As both α and $\sigma_r^2(r)$ scale with radii in the same way, we can write $\alpha(b)$ as

$$\alpha = \frac{4(\gamma-1)}{c^2}F(\gamma)\sigma_r^2(b) = \frac{4(\gamma-1)}{c^2}F(\gamma)\sigma_r^2(r)\left(\frac{b}{r}\right)^{-\gamma+2}. \tag{3.9}$$

We then obtain $D_A(EL)$ from equation (3.5) with α given by equation (3.9),

$$D_A(EL) = \frac{c^3\Delta t_{i,j}}{4\pi\sigma_r^2(r)(1+z_{\rm L})}(\Delta\tilde{\theta}_{i,j})^{-1}, \tag{3.10}$$

where[1]

$$(\Delta\tilde{\theta}_{i,j})^{-1} \equiv \frac{2\pi\left\{\frac{2}{-\gamma+3}\left[\theta_j\left(\frac{\theta_j}{\Theta}\right)^{-\gamma+2} - \theta_i\left(\frac{\theta_i}{\Theta}\right)^{-\gamma+2}\right] + (\theta_i+\theta_j)\left[\left(\frac{\theta_i}{\Theta}\right)^{-\gamma+2} - \left(\frac{\theta_j}{\Theta}\right)^{-\gamma+2}\right]\right\}^{-1}}{F(\gamma)(\gamma-1)}, \tag{3.11}$$

and Θ is the angular position at which the velocity dispersion is measured, i.e., $r = \Theta D_A(EL)$. For $\gamma = 2$, we obtain $\Delta\tilde{\theta}_{i,j} = \theta_j - \theta_i$, and thus we can reproduce the result of the SIS model (equation (2.4)).

Equation (3.10) still supports the basic physical picture that the ratio of $\Delta t_{i,j}$ and σ_r^2 gives some effective physical size of the lens, and dividing it by the appropriate angular separation in the sky, $\Delta\tilde{\theta}_{i,j}$, gives the angular diameter distance. The main difference between the SIS and the power law density profiles is that, in the latter case, the velocity dispersion is a function of radii. In general, image positions are different from the points at which the velocity dispersion is measured. Thus, we need to correct for the mismatch of the exact locations of the velocity dispersion measurement and the image positions. This is why the $\left(\frac{\theta}{\Theta}\right)^{-\gamma+2}$ term appears in the final expression of $D_A(EL)$: it scales the velocity dispersion such that we can get the potential at the image position. This requires us to measure (or model) the density slope, γ, as well.

3.2 External convergence

In modeling realistic lens systems, one important factor to consider is the so-called "mass-sheet transformation (MST)". MST is a subset of the source-position transformation [17]. Degeneracy exists, such that there are many mass models of the lens galaxy that can simultaneously reproduce most of the lensing observables, such as image positions and flux ratios, with different source positions [18]. This degeneracy constitutes one of the dominant sources of uncertainty in measuring the time-delay distance [7, 8, 17, 19]. In this subsection, we show that the effect of MST cancels out, leaving no effect on the inferred $D_A(EL)$.

[1] We use $\hat{\theta}_i \cdot \hat{\theta}_j = -1$ in reducing the vector dot products in equation (3.5) to the scalar products in equation (3.11).

Once we choose a model for the convergence field, $\kappa_{\rm model}(\vec{\theta})$, that matches the observations, we transform $\kappa_{\rm model}$ and $\vec{\alpha}$ to obtain a new convergence field, $\kappa_{\rm MST}(\vec{\theta})$, and a new scaled deflection, $\vec{\alpha}_{\rm MST}$, as

$$\kappa_{\rm MST}(\vec{\theta}) = \lambda + (1-\lambda)\kappa_{\rm model}(\vec{\theta}), \tag{3.12}$$

$$\vec{\alpha}_{\rm MST}(\vec{\theta}) = \lambda\vec{\theta} + (1-\lambda)\vec{\alpha}_{\rm model}(\vec{\theta}) \tag{3.13}$$

$$= \lambda\vec{\theta} + \vec{\alpha}_{\rm MST,lens}(\vec{\theta}), \tag{3.14}$$

where λ is a constant which physically corresponds to the scaled convergence of a uniform sheet of mass external to the lens galaxy. In equation (3.14), we decompose the transformed deflection into two parts; a deflection from the lens, and that from the external convergence. We define $\vec{\alpha}_{\rm MST,lens} \equiv (1-\lambda)\vec{\alpha}_{\rm model}$, whose meaning will be explained later in this subsection. To satisfy the lens equation (2.12) while leaving the image positions invariant, the source position must transform as

$$\vec{\beta}_{\rm MST} = (1-\lambda)\vec{\beta}_{\rm model}, \tag{3.15}$$

which is why this transformation is a part of the family of transformation called the *source-position* transformation.

Considering the following relation among κ , ϕ and ψ,

$$\phi = \frac{1}{2}(\vec{\theta}-\vec{\beta})^2 - \psi, \tag{3.16}$$

$$\nabla^2\psi = 2\kappa, \tag{3.17}$$

the transformed Fermat potential of the i-th image, $\phi_{{\rm MST},i}$, becomes

$$\phi_{{\rm MST},i} = (1-\lambda)\phi_{{\rm model},i} - \frac{\lambda(1-\lambda)}{2}|\vec{\beta}|^2. \tag{3.18}$$

Since the source position $\vec{\beta}$ is the same for all the images, the second term in equation (3.18) cancels out if we calculate the difference in the Fermat potential between two images i and j. Thus, the difference, $\Delta\phi_{i,j}$, transforms as

$$\Delta\phi_{{\rm MST},i,j} = (1-\lambda)\Delta\phi_{{\rm model},i,j}. \tag{3.19}$$

As the time delay is directly proportional to the Fermat potential, we find that $\Delta t_{i,j}$ is simply increased by a factor of $1-\lambda$ after the MST for fixed distances/cosmology.

If we assume that the physical origin of MST is an effective external convergence due to mass structures along the line of sight, $\kappa_{\rm ext}$, we can identify λ with $\kappa_{\rm ext}$. In the following, we apply the MST to the power-law mass model and show that the inferred $D_A(EL)$ remains unaffected by $\kappa_{\rm ext}$. We start first with the special case of SIS to gain intuition before considering the general power-law profile.

3.2.1 Singular isothermal sphere

Here we follow the steps from section 2.1, but with MST applied to it. From equation (3.12), the transformed density profile of the lens is

$$\rho_{\rm SIS,MST} = (1-\kappa_{\rm ext})\rho_{\rm SIS,model}. \tag{3.20}$$

Note that the original transformation equation (3.12) is written in terms of the convergence, κ; however, as the convergence and the density profile are proportional to each other (equation (2.8)), we transform the density in the same way as the convergence. To satisfy equation (2.1), the velocity dispersion must transform as

$$\sigma^2_{\rm MST} = (1-\kappa_{\rm ext})\sigma^2. \tag{3.21}$$

Equation (2.2) then becomes

$$\sigma^2_{\rm MST} = (1-\kappa_{\rm ext})\theta_{\rm E}\frac{c^2}{4\pi}\frac{D_A(ES)}{D_A(LS)}. \tag{3.22}$$

From equation (3.19), the time-delay equation (2.3) transforms as

$$\Delta t_{{\rm MST},i,j} = (1-\kappa_{\rm ext})\frac{1+z_{\rm L}}{2c}\frac{D_A(EL)D_A(ES)}{D_A(LS)}(\theta_j^2-\theta_i^2), \tag{3.23}$$

and by combining the above two equations, we get

$$\Delta t_{{\rm MST},i,j} = \frac{4\pi}{c^3}\sigma^2_{\rm MST}(1+z_{\rm L})D_A(EL)(\theta_i-\theta_j), \tag{3.24}$$

in which $\kappa_{\rm ext}$ cancels out. This equation is identical to equation (2.4), but with the transformed quantities, $\Delta t_{{\rm MST},i,j}$ and $\sigma^2_{\rm MST}$.

The reason is as follows. Suppose that we have a lens system which has a velocity dispersion of σ^2 and the time-delay difference of Δt. We then try to model this system by a lens plus an external convergence, $\kappa_{\rm ext}$. Then, the modeled σ and Δt would be different from the original ones by a factor of $1-\kappa_{\rm ext}$, but the ratio of the two is invariant. As $D_A(EL)$ is proportional to the ratio of the two, we can measure the same $D_A(EL)$ as before, regardless of the existence of the external convergence.

3.2.2 Power-law density profile

Now we study the effect of MST on the power-law density profile lens galaxy model, following section 3.1. Combining the time-delay transformation with equation (3.5) yields

$$\begin{aligned}\Delta t_{{\rm MST},i,j} =& (1-\kappa_{\rm ext})\Delta t_{{\rm model},i,j}\\ =& (1-\kappa_{\rm ext})D_A(EL)\frac{(1+z_{\rm L})}{2c}\\ &\times\left[(\hat{\alpha}_{{\rm model},i}+\hat{\alpha}_{{\rm model},j})\cdot(\vec{\theta}_i-\vec{\theta}_j)-\frac{2}{l}(\vec{\theta}_i\cdot\hat{\alpha}_{{\rm model},i}-\vec{\theta}_j\cdot\hat{\alpha}_{{\rm model},j})\right].\end{aligned} \tag{3.25}$$

Again, the density normalization of the lens galaxy, ρ_0, transforms as

$$\rho_{0,{\rm MST}} = (1-\kappa_{\rm ext})\rho_{0,{\rm model}}, \tag{3.26}$$

and thus among the total deflection angle α, only a $(1-\kappa_{\rm ext})$ fraction of it is from the lens, which is why we denoted this contribution as $\vec{\alpha}_{\rm MST,lens} = (1-\lambda)\vec{\alpha}_{\rm model}$ in equation (3.14). Using this in equation (3.25) yields

$$\begin{aligned}\Delta t_{{\rm MST},i,j} =& D_A(EL)\frac{(1+z_L)}{2c}\\ &\times\left[(\hat{\alpha}_{{\rm MST,lens},i}+\hat{\alpha}_{{\rm MST,lens},j})\cdot(\vec{\theta}_i-\vec{\theta}_j)-\frac{2}{l}(\vec{\theta}_i\cdot\hat{\alpha}_{{\rm MST,lens},i}-\vec{\theta}_j\cdot\hat{\alpha}_{{\rm MST,lens},j})\right].\end{aligned} \tag{3.27}$$

As the measured velocity dispersion of the lens gives the estimate of the lens potential only, the relation between the deflection angle from the lens and the velocity dispersion does not change after the MST:

$$|\hat{\alpha}_{\rm MST,lens}| = \frac{4(\gamma-1)}{c^2}\sigma_r^2(R)F(\gamma)\left(\frac{b}{R}\right)^{-\gamma+2}. \tag{3.28}$$

Thus, $D_A(EL)$ can be calculated from the original equation (3.10) even after the MST.

This is an important finding. In the previous studies of the time-delay distance to measure the Hubble constant, $\kappa_{\rm ext}$ was the main obstacle in measuring H_0 precisely [7]. On the other hand, we have shown that $D_A(EL)$ measured from strong lensing, which combines the time-delay, the image position, and the velocity dispersion data, does not suffer from the effect of $\kappa_{\rm ext}$.

4 Error formula and implications for B1608+686 and RXJ1131−1231

4.1 Aperture-averaged line of sight velocity dispersion

We do not measure the radial component of the velocity dispersion, $\sigma_r^2(r)$. Rather, we measure the luminosity-weighted line-of-sight velocity dispersion, $\sigma_p^2(R)$. We relate them using the following equation:

$$\sigma_p^2(R) \equiv I_H(R)\sigma_s^2(R) = 2\int_R^\infty \left(1-\beta_{\rm ani}\frac{R^2}{r^2}\right)\frac{\rho_*(r)\sigma_r^2(r)rdr}{\sqrt{r^2-R^2}}. \tag{4.1}$$

Here, r denotes the three-dimensional radius, while R denotes the projected radius. We shall use these two different radii notations for the rest of the paper. $\beta_{\rm ani}$ is the effect of the velocity dispersion anisotropy, which will be studied in detail in section 5. In this section, we set $\beta_{\rm ani}=0$. The other functions are: $I_H(R)$ is the projected stellar distribution function, $\sigma_s(R)$ is the projected velocity dispersion and $\sigma_r(r)$ is given by equation (3.8). For the stellar density profile, ρ_*, we use a Hernquiest distribution given by [20]

$$\rho_*(r) = \frac{I_0 a}{2\pi r(r+a)^3}, \tag{4.2}$$

where I_0 is a normalization and a is a scale radius. The projected Hernquist distribution, $I_H(R)$, is known to provide a good fit for the stellar distribution of elliptical galaxies that follow the de Vaucouleurs law,

$$I_H(R) = \frac{I_0}{2\pi a^2(1-s^2)^2}[(2+s^2)X(s)-3], \tag{4.3}$$

where $s\equiv R/a$ is a scaled projected radius, and $X(s)$ is defined as

$$X(s) \equiv \begin{cases} \dfrac{1}{\sqrt{1-s^2}}\,\mathrm{sech}^{-1}\,s & \text{for } 0\le s\le 1 \\ \dfrac{1}{\sqrt{s^2-1}}\,\sec^{-1}\,s & \text{for } 1\le s<\infty \end{cases}. \tag{4.4}$$

Ideally, we wish to measure the line-of-sight velocity dispersion profile as a function of projected radii. In practice, however, most of the observations do not allow us to spatially resolve

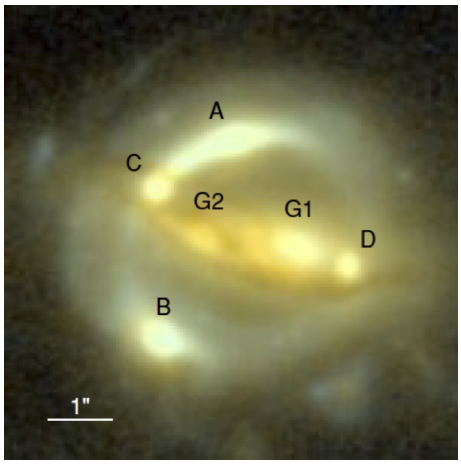

Figure 2: Image of B1608+656, adopted from figure 1 of [7].

the galaxy; rather they allow us to measure the luminosity-weighted, aperture-averaged velocity dispersion inside an aperture of a fixed size [21]. We calculate the luminosity-weighted aperture-averaged projected velocity dispersion, $\langle\sigma_p^2\rangle_{\rm ap}$, as follows:

$$\langle\sigma_p^2\rangle_{\rm ap} \equiv \frac{\int_{\rm ap} I_H \sigma_s^2 R\, dR\, d\theta}{\int_{\rm ap} I_H R\, dR\, d\theta}. \tag{4.5}$$

4.2 Analytic formula

In this section, we relate the statistical uncertainty in D_A to those of the observables, i.e., Δt, σ_p^2, and γ. (The effect of an anisotropic velocity dispersion will be discussed in detail in section 5.) Assuming that these observables are independently measured, we write the total uncertainty in $D_A(EL)$, hereafter S_{D_A}, as

$$\begin{aligned} S_{D_A} =& \sqrt{\left(\frac{\partial D_A}{\partial \Delta t_{i,j}}\right)^2 S_{\Delta t}^2 + \left(\frac{\partial D_A}{\partial \sigma_p^2}\right)^2 S_{\sigma_p^2}^2 + \left(\frac{\partial D_A}{\partial \gamma}\right)^2 S_\gamma^2} \\ =& D_A \sqrt{\left(\frac{1}{\Delta t_{i,j}}\right)^2 S_{\Delta t}^2 + \left(\frac{1}{\sigma_p^2}\right)^2 S_{\sigma_p^2}^2 + \frac{1}{D_A^2}\left(\frac{\partial D_A}{\partial \gamma}\right)^2 S_\gamma^2}, \end{aligned} \tag{4.6}$$

where S_x is the measurement uncertainty in the variable x. Since image positions, $\theta_{i,j}$, are precisely measured, we do not include their uncertainties in this formula. In the following sections 4.3 and 4.4, we shall apply this formula to two lens systems, B1608+656 and RXJ1131−1231, respectively.

4.3 B1608+656

Figure 2 shows the image configuration of B1608+656 [22]. The information on image configuration is important as our formula applies only to a circularly symmetric case. Thus, the only image pairs we can use in this paper are the ones that are on the opposite sides of the lens center. More thorough analysis using all the data will be presented elsewhere (Suyu et al., in preparation). The data of B1608+656 are mostly from ref. [7], but the image positions

are calculated from the data given in ref. [6], the time delays are from ref. [23], and the redshifts are from refs. [24, 25]. For this system, the origin of the coordinates is set at the image A. The data are summarized as :

$$\begin{aligned}
z_{\rm L} =& 0.6304 \\
z_{\rm s} =& 1.394 \\
\vec{\theta}_{\rm A} =& (0.0'', 0.0'') \\
\vec{\theta}_{\rm B} =& (-0.7380'', -1.9612'') \\
\vec{\theta}_{\rm C} =& (-0.7446'', -0.4537'') \\
\vec{\theta}_{\rm D} =& (1.1284'', -1.2565'') \\
R_{\rm eff} =& 0.58'' \\
a =& 0.551 R_{\rm eff} \\
\gamma =& 2.08 \pm 0.03 \\
\langle \sigma_p^2 \rangle_{\rm ap}^{1/2} =& 260 \pm 15 \text{ km/s} \\
\Delta t_{\rm AB} =& 31.5^{+2.0}_{-1.0} \text{ days} \\
\Delta t_{\rm CB} =& 36.0^{+1.5}_{-1.5} \text{ days} \\
\Delta t_{\rm DB} =& 77.0^{+2.0}_{-1.0} \text{ days} \\
\Delta t_{\rm CD} =& \Delta t_{\rm CB} - \Delta t_{\rm DB} = -41.0^{+2.5}_{-1.8} \text{ days}.
\end{aligned} \tag{4.7}$$

We use the CD pair. Also, as we write $D_A(EL)$ in terms of $\sigma_r(r)$ (e.g. 3.10), we normalize the radial velocity dispersion profile, $\sigma_r(r)$, using $\langle \sigma_p^2(R) \rangle_{\rm ap}$ given by the observation and using equations (4.1) and (4.5). With these values, we find $D_A(EL) = 1485.7$ Mpc. For comparison, $D_A(EL)$ from the best-fit WMAP 7-year parameters is $D_A(EL) = 1423.3$ Mpc. We now use equation (4.6) to compute S_{D_A}:

$$S_{D_A} = D_A \sqrt{3.72 \times 10^{-3} + 1.33 \times 10^{-2} + 2.36 \times 10^{-3}}, \tag{4.8}$$

where from the first term, each number indicates the fractional uncertainty in D_A contributed by the time-delay measurement $\Delta t_{i,j}$, the line-of-sight velocity dispersion measurement σ_p^2, and the density profile index γ. (Note that $S_{\sigma_p^2}/\sigma_p^2 = 2S_{\sigma_p}/\sigma_p$.) With this value, the total uncertainty, including all the terms in equation (4.8), is $S_{D_A} = 0.14 D_A$, i.e., 14% uncertainty. The dominant contribution comes from the uncertainty in σ_p, which gives $S_{D_A} = 0.12 D_A$.

4.4 RXJ1131−1231

In this section we repeat the same analysis as above, but with another well-studied strong lensing time-delay system, RXJ1131−1231, using data from refs. [26, 27] for the time delays and the redshifts, respectively, and from ref. [8] for the other quantities. The data for this

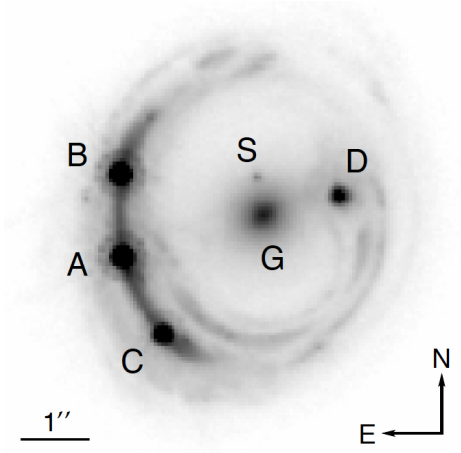

Figure 3: Image of RXJ1131−1231, adopted from figure 1 of [8].

system are

$$\begin{aligned}
z_{\rm L} &= 0.295 \\
z_{\rm s} &= 0.658 \\
\vec{\theta}_{\rm G} &= (4.411'', 4.011'') \\
\vec{\theta}_{\rm A} &= (2.383'', 3.412'') \\
\vec{\theta}_{\rm D} &= (5.494'', 4.288'') \\
R_{\rm eff} &= 1.85'' \\
\gamma &= 1.95^{+0.05}_{-0.04} \\
\langle \sigma_p^2 \rangle^{1/2}_{\rm ap} &= 323 \pm 20 \text{ km/s} \\
\Delta t_{\rm AB} &= 0.7 \pm 1.4 \text{ days} \\
\Delta t_{\rm DB} &= 91.4 \pm 1.5 \text{ days} \\
\Delta t_{\rm AD} &= \Delta t_{\rm AB} - \Delta t_{\rm DB} = -90.7 \pm 2.1 \text{ days.}
\end{aligned} \tag{4.9}$$

We use the AD pair. Using these values, we find $D_A(EL) = 813.33$ Mpc, and $D_A(EL)$ from the best-fit WMAP 7-year parameters is $D_A(EL) = 876.5$ Mpc. The total uncertainty in $D_A(EL)$ is

$$S_{D_A} = D_A \sqrt{5.36 \times 10^{-4} + 1.53 \times 10^{-2} + 1.46 \times 10^{-3}} = 0.13 D_A. \tag{4.10}$$

The velocity dispersion alone gives $S_{D_A} = 0.12 D_A$.

Therefore, we expect the *existing* data on these systems to yield $D_A(EL)$ with $13-14\%$ precision per object, assuming the isotropic velocity dispersion. In the next section, we shall study the effect of the largest source of systematic uncertainty in our method: an anisotropic velocity dispersion, and how to reduce its effect in the estimation of $D_A(EL)$.

5 Anisotropic velocity dispersion

The anisotropic stellar motion changes the relation between the potential and the observed line-of-sight velocity dispersion. As our method crucially relies upon knowing the potential

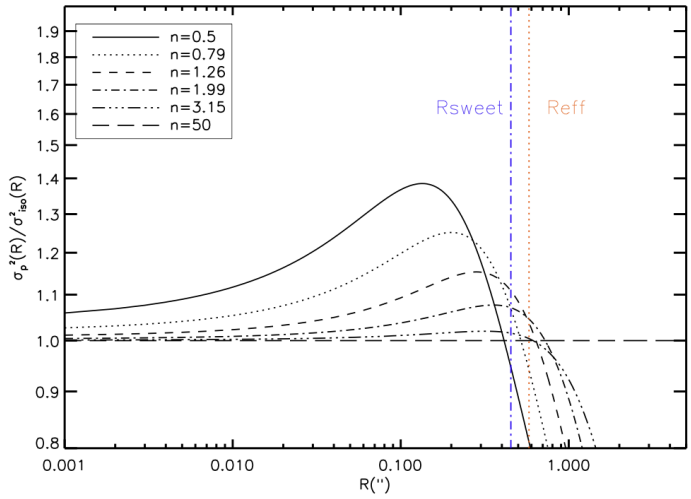

Figure 4: Ratio of $\sigma_p^2(R)$ to $\sigma_{\rm iso}^2(R)$, as a function of the projected radius R, and $n \equiv r_{\rm ani}/R_{\rm eff}$. The former is observable, while the latter is related more directly to GM/R. Two vertical lines show the effective radius ($R_{\rm eff}$) and the sweet-spot radius ($R_{\rm sweet}$) defined in section 5.2.

depth, we must take into account the anisotropic velocity dispersion of stars. We do this by following ref. [7], which uses spherical Jeans modeling to relate the observed line-of-sight velocity dispersion to the mass distribution, assuming that the stellar distribution follows a Hernquist profile (section 5.1). We then study the effect of anisotropy on the aperture-averaged value of the velocity dispersion (section 5.1) as well as on the velocity dispersion measured at the so-called "sweet radius" (section 5.2). Finally, we use Monte Carlo simulations to compute the effect of anisotropy on the uncertainty in $D_A(EL)$ (section 5.3).

5.1 Spherical Jeans equation

We solve the spherical Jeans equation for a given mass distribution (i.e., a power-law density profile) to obtain the three-dimensional radial velocity dispersion σ_r,

$$\frac{1}{\rho_*}\frac{d(\sigma_r^2\rho_*)}{dr} + 2\beta_{\rm ani}\frac{\sigma_r^2}{r} = -\frac{GM(\leq r)}{r^2}. \tag{5.1}$$

Here, the anisotropy function, $\beta_{\rm ani}(r)$, is chosen as the Osipkov-Merritt anisotropy [28, 29],

$$\beta_{\rm ani}(r) \equiv \frac{r^2}{r_{\rm ani}^2 + r^2} = 1 - \frac{\sigma_T^2(r)}{\sigma_r^2(r)}, \tag{5.2}$$

where $\sigma_T(r)$ and $\sigma_r(r)$ are the velocity dispersions in the tangential and radial directions, respectively. Although the anisotropy is parametrized by a single variable, $r_{\rm ani}$, under this specific model, we can model almost any velocity structures by linearly superimposing the solutions [29]. We then calculate $\sigma_p^2(R)$ from $\sigma_r^2(r)$ using equation 4.1, and $\langle\sigma_p^2\rangle_{\rm ap}$ using equation 4.5.

In figure 4, we show the ratio of $\sigma_p^2(R)$ to the isotropic velocity dispersion, $\sigma_{\rm iso}^2(R)$, with $a = 0.551R_{\rm eff}$ and $R_{\rm eff} = 0.58''$. The isotropic velocity dispersion is a solution to the Jeans

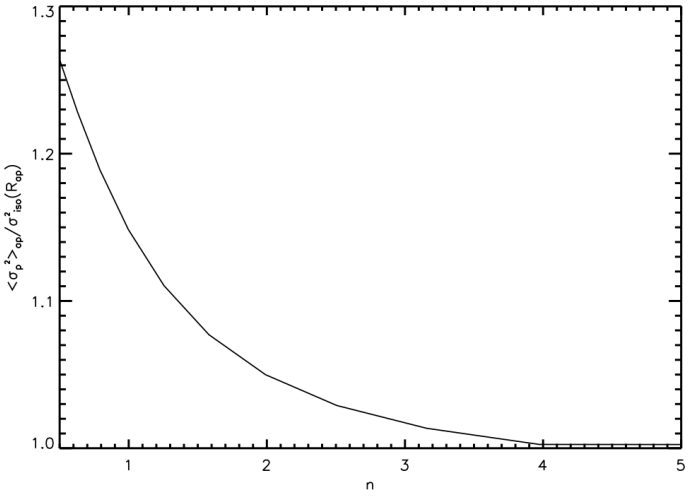

Figure 5: Ratio of $\langle\sigma_p^2\rangle_{\rm ap}$ to $\sigma^2_{\rm iso}(R_{\rm ap})$, as a function of n. The size of the aperture is fixed at $R_{\rm ap} = 0.42''$. The upper limit of n, 5, is chosen since the velocity dispersion does not differ much from the isotropic case beyond n of 5, while the lower limit, 0.5, is determined by observations [e.g. [30]] and radial instability arguments [e.g. [31, 32]].

equation (5.1) with no anisotropy, $\beta_{\rm ani} \equiv 0$; thus, it is related more directly to GM/R. We have one free parameter, n, which parametrizes the anisotropic radius as

$$r_{\rm ani} \equiv nR_{\rm eff}. \tag{5.3}$$

For a given mass distribution of the lens, $\sigma_p^2(R)$ depends on n. We vary n from 0.5 to 50 in logarithmic spacing. We find $\sigma_p^2(R)/\sigma^2_{\rm iso}(R) \approx 1$ to within 10% at $R = R_{\rm eff}$, except for the highly anisotropic case of $n = 0.5$.

In figure 5, we show the ratio of $\langle\sigma_p^2\rangle_{\rm ap}$ to $\sigma^2_{\rm iso}(R_{\rm ap})$ as a function of n, where $R_{\rm ap}$ is fixed to 0.42". This ratio reaches 26% for $n = 0.5$, and decreases as n increases. Since the inferred $D_A(EL)$ is proportional to the inverse of the isotropic velocity dispersion, having a large variation in the inferred isotropic velocity dispersion can cause a large uncertainty in $D_A(EL)$. Unfortunately, anisotropy is not directly observable, unless we have a three-dimensional velocity dispersion measurement. Clearly, a better approach is needed.

5.2 Sweet-spot method

It has been pointed out that, when the observations of the surface brightness profile and the velocity dispersion profile are available, one can find the so-called *sweet radius*, $R_{\rm sweet}$, at which the effect of the anisotropic velocity dispersion on the mass determination is minimized [33]. Also see [34, 35]. The Osipkov-Merritt anisotropy model has an isotropic core and a radial envelope. However, as we observe the projected velocity dispersion, there are two components that play roles in the estimation of the observed velocity dispersion. The anisotropy changes the ratio between tangential and radial components of the velocity dispersion at a given radius, while the projection changes the magnitude of contributions from radial and tangential components. Quantitatively, at a fixed radius of observation R, $\sigma_p(R)$ has contributions from infinitely many shells with radii $r = R/\cos x$, where $x = [0, \pi/2]$.

At each radius r we can decompose the contributions to the projected velocity dispersion into tangential and radial components as $\sigma_T(r)\cos x$ and $\sigma_r(r)\sin x$, respectively. Due to the weighting by the trigonometric functions, at small x, contributions from the tangential component is bigger than that from the radial component, and vice versa at large x. Now, let us assume that the total velocity dispersion, $\sigma^2(r) = \sigma_T^2(r) + \sigma_r^2(r)$, is the same for the isotropic and anisotropic model for a given galaxy mass, as it is proportional to the total kinetic energy. Then, when $r = R/\cos x$ is large, the tangential component is suppressed while the radial component is enhanced compare to the isotropic case, due to the anisotropy (since $\sigma_T(r)$ becomes small for large r in equation (5.2)). As a result, in comparison to the isotropic case where $\sigma_T(r) = \sigma_r(r)$, anisotropic velocity dispersion shows $\sigma_p(R) > \sigma_{\rm iso}(R)$ at small R, and $\sigma_p(R) < \sigma_{\rm iso}(R)$ at large R. Thus, if we observe an anisotropic system, there exists a projected radius R at which the transition from one to the other occurs, as we increase R from the center of a galaxy to the outskirt of it. This transition radius is the sweet spot.

While the analytical derivation of $R_{\rm sweet}$ has been done assuming a constant $\beta_{\rm ani}$, the further study [36] shows that the method works for systems with a non-constant $\beta_{\rm ani}$ as well. The sweet spot can be determined from the brightness profile of a massive elliptical galaxy [36]. It is close to the projected radius at which R satisfies $d\ln I(R)/d\ln R = -2$. For a Hernquist surface brightness profile, we find $R_{\rm sweet} \approx 0.78R_{\rm eff}$. In figure 4, the sweet radius is shown as the left vertical line. We find that the difference between projected velocity dispersions with various anisotropy parameters is minimum around $R = 0.78R_{\rm eff} = 0.45''$. It particularly reduces the effect of a highly anisotropic case with $n = 0.5$, compared to using $\sigma_p^2(R)$ at the effective radius or the aperture-averaged σ_p^2. The uncertainty in the mass of massive ellipticals estimated from the sweet-spot method is claimed to be 5-7 per cent. Therefore, the best approach is to use spatially-resolved spectroscopic data of lens galaxies to obtain the velocity dispersion at the sweet radius.

5.3 Monte Carlo simulation

We use Monte Carlo simulations to study how much the velocity anisotropy inflates the uncertainty in $D_A(EL)$, and how well we can mitigate it by using the sweet-spot method.

For two time-delay systems B1608+656 and RXJ1131−1231, we generate 11 discrete radial profiles of anisotropic velocity dispersion by solving the Jeans equation, with logarithmically spaced $n = [0.5, 50]$. The effective radius and the density profile index γ are fixed at the best-fit values given in sections 4.3 and 4.4. We randomly choose one profile from the set of different anisotropy parameters to create a mock galaxy. We then compute σ_p^2 from each mock galaxy in three ways: the aperture-averaged value $\langle\sigma_p^2\rangle_{\rm ap}$ with the aperture size of $0.42''$ for both systems, $\sigma_p^2(R)$ with $R = R_{\rm eff}$, and $\sigma_p^2(R)$ with $R = R_{\rm sweet}$. As the uncertainty in $D_A(EL)$ is dominated by that of σ_p^2, we add a Gaussian random noise to σ_p^2 with variance of $S^2_{\sigma_p^2} = 2S^2_{\sigma_p}(S^2_{\sigma_p} + 2\sigma_p^2)$. We then compute $D_A(EL)$ from these simulated data with the best-fit values of the time-delay data and image positions given in sections 4.3 and 4.4. (We do not add noise to time delays or image positions.) While our simulated galaxies have anisotropic velocity dispersions, we use the isotropic velocity dispersion model to calculate $D_A(EL)$. In this way we can quantify the effect of our ignoring anisotropic velocity dispersion by marginalizing over it.

Figures 6, 7, and 8 show the distributions of $D_A(EL)$ obtained from mock B1608+656 realizations using $\langle\sigma_p^2\rangle_{\rm ap}$, $\sigma_p^2(R_{\rm eff})$, and $\sigma_p^2(R_{\rm sweet})$, respectively. The solid and dashed histograms in each panel show the realizations with the isotropic and anisotropic velocity dis-

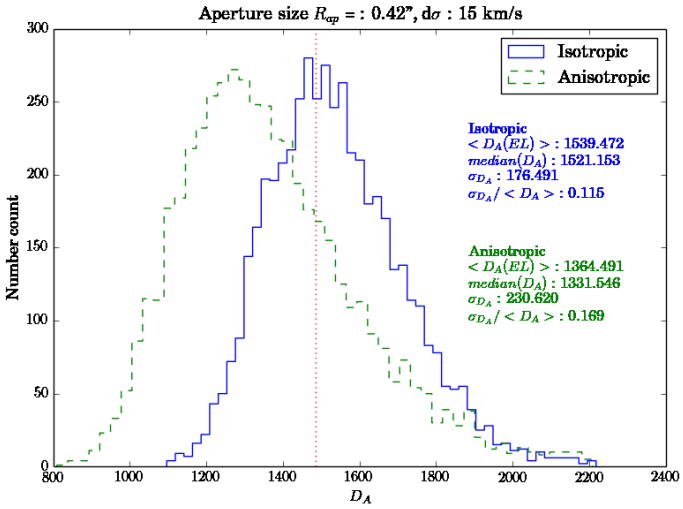

Figure 6: Simulated distribution of $D_A(EL)$ to B1608+656. The solid and dashed histograms show the distributions with the isotropic and anisotropic simulations, interpreted by the isotropic model. We use the aperture averaged velocity dispersion, $\langle\sigma_p^2\rangle_{\rm ap}$, with the aperture size of $0.42''$. The standard deviation of the velocity dispersion used in simulations is 15 km/s.

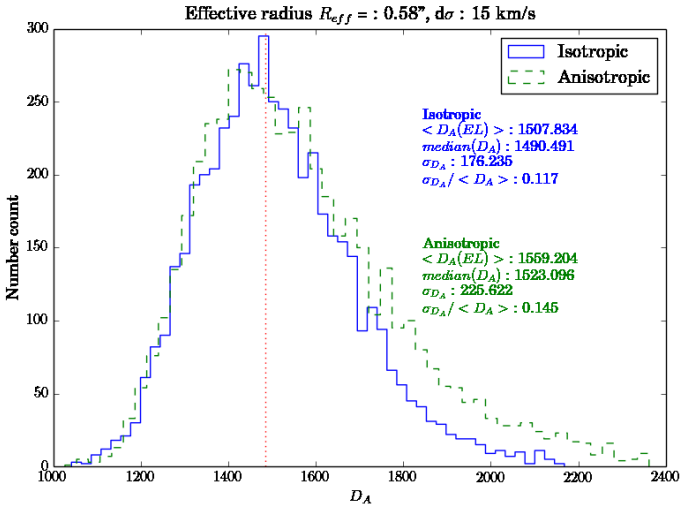

Figure 7: Same as figure 6, but with $\sigma_p^2(R_{\rm eff})$ and $R_{\rm eff} = 0.58''$.

persions, respectively. The former realizations are used to check validity of our simulations, as well as to make a direct assessment of the effect of anisotropy. The vertical dotted lines show $D_A(EL) = 1485.7$ Mpc that we obtained in section 4.3.

We summarize the results from the analysis on B1608+656 and RXJ1131−1231 in tables 1 and 2, respectively. The uncertainties in $D_A(EL)$ from isotropic simulations (interpreted by the isotropic model) agree with the analytical estimates given in sections 4.3 and 4.4. On the other hand, those from anisotropic simulations (again interpreted by the isotropic model) show significantly larger uncertainties when $\langle\sigma_p^2\rangle_{\rm ap}$ or $\sigma_p^2(R_{\rm eff})$ is used. Fortunately,

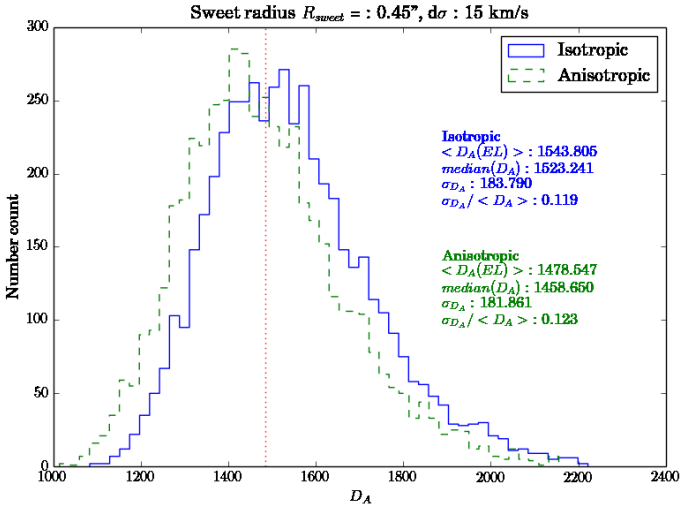

Figure 8: Same as figure 6, but with $\sigma_p^2(R_{\rm sweet})$ and $R_{\rm sweet} = 0.45''$.

Table 1: Expected fractional uncertainty in $D_A(EL)$ to B1608+656

	Isotropic	**Anisotropic**
$R_{\rm ap}$	11.5%	16.9%
$R_{\rm eff}$	11.7%	14.5%
$R_{\rm sweet}$	11.9%	12.3%

using $\sigma_p^2(R_{\rm sweet})$ eliminates most of the inflation of the uncertainty due to velocity anisotropy. Figure 8 shows that the peak is shifted in the anisotropic case in comparison to the isotropic case, while in figure 7 the peak remains at the same position. This is due to the marginalization of the anisotropy. In figure 4, we choose 6 different n values that are spaced logarithmically, and choose two radii ($R_{\rm eff}$ and $R_{\rm sweet}$) to calculate the D_A distributions for both the isotropic and anisotropic cases. At $R_{\rm sweet}$, the scatter between the curves is smaller compare to that at $R_{\rm eff}$; however, at $R_{\rm sweet}$, the curves are also shifted toward higher velocity dispersions compared to the isotropic case. As a result, the whole distribution of D_A is shifted toward lower values. On the other hand, at $R_{\rm eff}$, while the scatter is larger, there is no systematic change in σ_p^2 relative to $\sigma_{\rm iso}$ (i.e. among 6 values of n, two give σ_p^2 larger than the $\sigma_{\rm iso}^2$, two give smaller, and the other two give σ_p^2 almost identical to the $\sigma_{\rm iso}^2$ value). As a result, the peak position does not change, while we get an extended tail towards higher D_A value. This does not mean that using $R_{\rm sweet}$ gives a biased D_A, as we cannot assume that the velocity dispersion structure is isotropic. Also, as the width of the distribution is much bigger than the shift of the peak, at the moment the effect of this shift is negligible. As the distribution of D_A depends on the choice of the anisotropy model as well as on the range/selection of n, we should apply other anisotropy parametrizations to further study and improve the anisotropy model to achieve more precise measurements of D_A.

Table 2: Expected fractional uncertainty in $D_A(EL)$ to RXJ1131−1231

	Isotropic	**Anisotropic**
$R_{\rm ap}$	12.5%	15.1%
$R_{\rm eff}$	12.4%	14.8%
$R_{\rm sweet}$	12.5%	12.6%

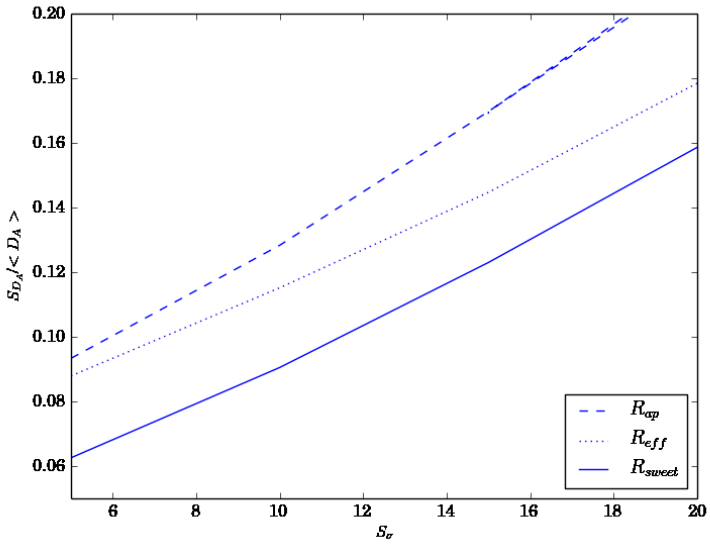

Figure 9: Expected fractional uncertainty in $D_A(EL)$ to B1608+656 as a function of the uncertainty in σ in units of km/s. The dashed, dotted, and solid lines are for $\sigma = \langle\sigma_p^2\rangle_{\rm ap}^{1/2}$, $\sigma_p(R_{\rm eff})$, and $\sigma_p(R_{\rm sweet})$, respectively.

6 Conclusion

In this paper, we have shown that we can determine $D_A(EL)$ to strong lens systems with time delays. The underlying physics is simple; thus, this method offers a robust determination of $D_A(EL)$ to individual systems. The key advantage of this method is that the external convergence does not affect the distance determination. The uncertainty in the inferred $D_A(EL)$ is dominated by that in the velocity dispersion and its anisotropy. The effect of anisotropy can be minimized by measuring the velocity dispersion at the sweet radius.

The *existing* data on B1608+656 and RXJ1131−1231 should yield $D_A(EL)$ with 17% and 15% precision, respectively. If we use the velocity dispersions at the sweet radii, the precision improves to about 13%. In figure 9, we show the expected fractional uncertainty in $D_A(EL)$ to B1608+656 as a function of the uncertainty in the velocity dispersions, σ. The σ at the sweet radius measured with 260 ± 7 km/s corresponds to σ^2 measured with 5% precision. This yields $D_A(EL)$ with 7% precision, after marginalizing over velocity anisotropy. However, to quantify the uncertainties from the anisotropy, further study on other anisotropy parametrizations and their prior ranges is needed.

This paper describes the basic idea and presents an estimate of what we can do with the

existing data. Since we assumed spherical density profiles, our analysis is not precise enough to yield the best determinations of $D_A(EL)$ to B1608+656 and RXJ1131−1231. The method presented in this paper has been implemented in the full analysis pipeline used by refs. [7, 8], and the results will be reported in a future publication (Suyu et al., in preparation).

Acknowledgments

IJ would like to thank Karl Gebhardt for his support and encouragement, and Adriano Agnello, Matthew Auger, Simona Vegetti, Stefan Hilbert, Chiara Spiniello, Natalia Lyskova, and Edward Robinson for useful discussions. EK would like to thank Phil Marshall and Masamune Oguri for useful discussions, and Stefan Hilbert for his clear lecture on strong lenses. SHS would like to thank Adriano Agnello, Matthew Auger and Matteo Barnabè for helpful discussions. Funding for this work has been provided in part by Texas Cosmology Center (TCC). TCC is supported by the College of Natural Sciences and the Department of Astronomy at the University of Texas at Austin and the McDonald Observatory.

References

[1] M. Im, R. E. Griffiths, and K. U. Ratnatunga, *A Measurement of the Cosmological Constant Using Elliptical Galaxies as Strong Gravitational Lenses*, *Astrophys.J.* **475** (1997) 457–461, [astro-ph/9611105].

[2] T. Futamase and T. Hamana, *Constraint on the Cosmological Constant by Einstein Ring System 0047-2808*, *Progress of Theoretical Physics* **102** (1999) 1037–1041.

[3] T. Futamase and S. Yoshida, *Possible measurement of quintessence and density parameter using strong gravitational lensing events*, *Progress of Theoretical Physics* **105** (2001) 887–891, [gr-qc/0011083].

[4] K. Yamamoto and T. Futamase, *Possible method to reconstruct the cosmic equation of state from strong gravitational lensing systems*, *Progress of Theoretical Physics* **105** (2001) 707–716, [astro-ph/0104253].

[5] E. V. Linder, *Strong gravitational lensing and dark energy complementarity*, *Phys.Rev.* **D70** (2004) 043534, [astro-ph/0401433].

[6] L. Koopmans, T. Treu, C. Fassnacht, R. Blandford, and G. Surpi, *The Hubble Constant from the gravitational lens B1608+656*, *Astrophys.J.* **599** (2003) 70–85, [astro-ph/0306216].

[7] S. Suyu, P. Marshall, M. Auger, S. Hilbert, R. Blandford, et al., *Dissecting the Gravitational Lens B1608+656. II. Precision Measurements of the Hubble Constant, Spatial Curvature, and the Dark Energy Equation of State*, *Astrophys.J.* **711** (2010) 201–221, [arXiv:0910.2773].

[8] S. Suyu, M. Auger, S. Hilbert, P. Marshall, M. Tewes, et al., *Two accurate time-delay distances from strong lensing: Implications for cosmology*, *Astrophys.J.* **766** (2013) 70, [arXiv:1208.6010].

[9] M. Fukugita, T. Futamase, and M. Kasai, *A Possible Test for the Cosmological Constant with Gravitational Lenses*, *Mon.Not.Roy.Astron.Soc.* **246** (1990) 24P.

[10] D. Paraficz and J. Hjorth, *Gravitational lenses as cosmic rulers: Ω_m, Ω_Λ from time delays and velocity dispersions*, *Astron. Astrophys.* **507** (2009) L49–L52, [arXiv:0910.5823].

[11] S. Refsdal, *On the possibility of determining Hubble's parameter and the masses of galaxies from the gravitational lens effect*, *Mon.Not.Roy.Astron.Soc.* **128** (1964) 307.

[12] H. Witt, S. Mao, and C. Keeton, *Analytic time delays and h_0 estimates for gravitational lenses*, *Astrophys.J.* **544** (2000) 98–103, [astro-ph/0004069].

[13] L. Koopmans, A. Bolton, T. Treu, O. Czoske, M. Auger, et al., *The Structure & Dynamics of Massive Early-type Galaxies: On Homology, Isothermality and Isotropy inside one Effective Radius*, *Astrophys.J.* **703** (2009) L51–L54, [arXiv:0906.1349].

[14] M. W. Auger, T. Treu, A. S. Bolton, R. Gavazzi, L. V. E. Koopmans, P. J. Marshall, L. A. Moustakas, and S. Burles, *The Sloan Lens ACS Survey. X. Stellar, Dynamical, and Total Mass Correlations of Massive Early-type Galaxies*, *Astrophys.J.* **724** (2010) 511–525, [arXiv:1007.2880].

[15] M. Barnabè, O. Czoske, L. V. E. Koopmans, T. Treu, and A. S. Bolton, *Two-dimensional kinematics of SLACS lenses - III. Mass structure and dynamics of early-type lens galaxies beyond $z \simeq 0.1$*, *Mon.Not.Roy.Astron.Soc.* **415** (Aug., 2011) 2215–2232, [arXiv:1102.2261].

[16] A. Sonnenfeld, T. Treu, R. Gavazzi, S. H. Suyu, P. J. Marshall, et al., *The SL2S Galaxy-scale Lens Sample. IV. The dependence of the total mass density profile of early-type galaxies on redshift, stellar mass, and size*, *Astrophys.J.* **777** (2013) 98, [arXiv:1307.4759].

[17] P. Schneider and D. Sluse, *Source-position transformation – an approximate invariance in strong gravitational lensing*, *Astron. Astrophys.* **564** (2014) A103, [arXiv:1306.4675].

[18] E. E. Falco, M. V. Gorenstein, and I. I. Shapiro, *On model-dependent bounds on H(0) from gravitational images Application of Q0957 + 561A,B*, *Astrophys.J* **289** (1985) L1–L4.

[19] S. Suyu, T. Treu, S. Hilbert, A. Sonnenfeld, M. Auger, et al., *Cosmology from gravitational lens time delays and Planck data*, *Astrophys.J.* **788** (2014) L35, [arXiv:1306.4732].

[20] L. Hernquist, *An analytical model for spherical galaxies and bulges*, *Astrophys.J.* **356** (1990) 359.

[21] M. Cappellari, R. Bacon, M. Bureau, M. Damen, R. L. Davies, et al., *The SAURON Project. 4. The Mass-to-light ratio, the virial mass estimator and the fundamental plane of elliptical and lenticular galaxies*, *Mon.Not.Roy.Astron.Soc.* **366** (2006) 1126, [astro-ph/0505042].

[22] S. Suyu, P. Marshall, R. Blandford, C. Fassnacht, L. Koopmans, et al., *Dissecting the Gravitational Lens B1608+656: Lens Potential Reconstruction*, *Astrophys.J.* **691** (2009) 277–298, [arXiv:0804.2827].

[23] C. Fassnacht, E. Xanthopoulos, L. Koopmans, and D. Rusin, *A Determination of H(O) with the class gravitational lens B1608+656. 3. A Significant improvement in the precision of the time delay measurements*, *Astrophys.J.* **581** (2002) 823–835, [astro-ph/0208420].

[24] C. D. Fassnacht, D. S. Womble, G. Neugebauer, I. W. A. Browne, A. C. S. Readhead, K. Matthews, and T. J. Pearson, *1608+656: A gravitationally lensed poststarburst radio galaxy*, *Astrophys.J.* **460** (Apr., 1996) L103.

[25] S. T. Myers, C. D. Fassnacht, S. G. Djorgovski, R. D. Blandford, K. Matthews, G. Neugebauer, T. J. Pearson, A. C. S. Readhead, J. D. Smith, D. J. Thompson, D. S. Womble, I. W. A. Browne, P. N. Wilkinson, S. Nair, N. Jackson, I. A. G. Snellen, G. K. Miley, A. G. de Bruyn, and R. T. Schilizzi, *1608+656: A Quadruple-Lens System Found in the CLASS Gravitational Lens Survey*, *Astrophys.J.* **447** (July, 1995) L5.

[26] M. Tewes, F. Courbin, G. Meylan, C. S. Kochanek, E. Eulaers, N. Cantale, A. M. Mosquera, P. Magain, H. Van Winckel, D. Sluse, G. Cataldi, D. Vörös, and S. Dye, *COSMOGRAIL: the COSmological MOnitoring of GRAvItational Lenses. XIII. Time delays and 9-yr optical monitoring of the lensed quasar RX J1131-1231*, *Astron.Astrophys.* **556** (aug, 2013) A22, [arXiv:1208.6009].

[27] D. Sluse, J. Surdej, J. Claeskens, D. Hutsemekers, C. Jean, et al., *A Quadruply imaged quasar with an optical Einstein ring candidate: 1RXS J113155.4-123155*, *Astron.Astrophys.* **406** (2003) L43–L46, [astro-ph/0307345].

[28] L. P. Osipkov, *Spherical systems of gravitating bodies with an ellipsoidal velocity distribution*, *Soviet Astronomy Letters* **5** (1979) 42–44.

[29] D. Merritt, *Spherical stellar systems with spheroidal velocity distributions*, *Astronomical.J.* **90** (1985) 1027–1037.

[30] A. Kronawitter, R. P. Saglia, O. Gerhard, and R. Bender, *Orbital structure and mass distribution in elliptical galaxies*, *Astron.Astrophys.* **144** (may, 2000) 53–84.

[31] D. Merritt and L. A. Aguilar, *A numerical study of the stability of spherical galaxies*, *Mon.Not.Roy.Astron.Soc.* **217** (dec, 1985) 787–804.

[32] M. Stiavelli and L. S. Sparke, *Influence of a dark halo on the stability of elliptical galaxies*, *Astron.Astrophys.* **382** (dec, 1991) 466–474.

[33] E. Churazov, S. Tremaine, W. Forman, O. Gerhard, P. Das, et al., *Comparison of an approximately isothermal gravitational potentials of elliptical galaxies based on X-ray and optical data*, *Mon.Not.Roy.Astron.Soc.* **404** (2010) 1165–1185, [arXiv:1001.3435].

[34] J. Wolf, G. D. Martinez, J. S. Bullock, M. Kaplinghat, M. Geha, R. R. Muñoz, J. D. Simon, and F. F. Avedo, *Accurate masses for dispersion-supported galaxies*, *Mon.Not.Roy.Astron.Soc.* **406** (Aug., 2010) 1220–1237, [arXiv:0908.2995].

[35] M. G. Walker, M. Mateo, E. W. Olszewski, J. Peñarrubia, N. Wyn Evans, and G. Gilmore, *A Universal Mass Profile for Dwarf Spheroidal Galaxies?*, *Astrophys.J.* **704** (Oct., 2009) 1274–1287, [arXiv:0906.0341].

[36] N. Lyskova, E. Churazov, I. Zhuravleva, T. Naab, L. Oser, et al., *Testing a simple recipe for estimating galaxy masses from minimal observational data*, *Mon.Not.Roy.Astron.Soc.* **423** (2012) 1813–1824, [arXiv:1204.0465].

A Deflection angle of an arbitrary power-law density profile

We derive the expression for a deflection angle near a galaxy with the density profile following a power-law with arbitrary density profile index. When the density profile is given as equation (3.1), the mass contained within a radius r is

$$M(r) = \int_0^r 4\pi r^2 \rho_0 r_0^\gamma r^{-\gamma} dr = \frac{4\pi\rho_0 r_0^\gamma}{-\gamma+3} r^{-\gamma+3}, \tag{A.1}$$

which yields an acceleration given by

$$\vec{g}(\vec{r}) = -\frac{4\pi G\rho_0 r_0^\gamma}{3-\gamma} r^{-\gamma}\vec{r}, \tag{A.2}$$

on the test mass located at the radius r. According to the post-Newtonian approximation in General Relativity, the rate of change of the direction of the velocity vector of the test mass, $\vec{u}$, is given as

$$c^2 \frac{d\vec{u}}{dt} = -2\vec{u} \times (\vec{u} \times \vec{g}). \tag{A.3}$$

We define a new parameter α to be the angle by which the light is deflected as it passes near the lens galaxy. In the cases we consider, the deflection angle will be small. Thus, we can choose coordinates such that the path of the light is roughly along the x-axis, and the line connecting the center of the lens galaxy to the point of the closest approach is along the y axis. Again, because the deflection angle is small, we use the thin lens approximation, namely,

light is bent sharply at the closest approach to the lens. Thus, the separation from the center of lens to the light path, r, becomes $r^2 = b^2 + x^2$, and, more importantly, $\vec{u} \times (\vec{u} \times \vec{r}) = -c^2\vec{b}$.

We define the deflection angle at the lens plane, $\hat{\alpha}$, as the total change in the photon propagation direction, and the magnitude of the deflection angle as α. Then,

$$\hat{\alpha} \equiv \frac{1}{c} \int d\vec{u} = -\alpha \frac{\vec{r}}{r}, \tag{A.4}$$

where the minus sign indicates that the deflection happens toward the lens center. Then α becomes

$$\begin{aligned} \alpha &= \frac{8\pi G\rho_0 r_0^\gamma}{c^2(3-\gamma)} \int_{-\infty}^{\infty} b r^{-\gamma} dx \\ &= \frac{8\pi G\rho_0 r_0^\gamma}{c^2(3-\gamma)} b \int_{-\infty}^{\infty} \frac{dx}{(x^2+b^2)^{\gamma/2}} \\ &= \frac{8\pi G\rho_0 r_0^\gamma b^{2-\gamma}}{c^2(3-\gamma)} \frac{\sqrt{\pi}\ \Gamma[\frac{1}{2}(-1+\gamma)]}{\Gamma(\frac{\gamma}{2})}, \end{aligned} \tag{A.5}$$

assuming $\gamma > 1$.