Astronomy & Astrophysics

TDCOSMO

XXII. Triaxiality and projection effects in time-delay cosmography

Xiang-Yu Huang[1],[*][iD], Simon Birrer[1][iD], Michele Cappellari[2][iD], Tommaso Treu[3][iD],
Shawn Knabel[3][iD], and Dominique Sluse[4][iD]

[1] Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA
[2] Sub-Department of Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK
[3] Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA
[4] STAR Institute, University of Liège, Quartier Agora, Allée du six Août 19c, 4000 Liège, Belgium

ABSTRACT

Context. Constraining the mass-sheet degeneracy (MSD) is crucial for improving the precision and accuracy of time-delay cosmography. Joint analyses based on lensing and stellar kinematics have been widely adopted to break the MSD. A three-dimensional (3D) mass and stellar tracer population is required to accurately interpret the kinematics data.
Aims. Our forward-modeling procedure is aimed at evaluating the projection effects using strong lensing and kinematics observables and to determine an optimal model assumption for the stellar kinematics analysis leading to an unbiased interpretation of the MSD and H_0.
Methods. We numerically simulated the projection and selection effects for both a triaxial early-type galaxy (ETG) sample from the TNG100 simulation and an axisymmetric sample that matches the properties of slow-rotator galaxies representative of the strong lens galaxy population. Using the axisymmetric sample, we generated mock kinematics observables with spherically aligned axisymmetric Jeans anisotropic modeling (JAM) and assessed the kinematic recovery under different model assumptions. Using the triaxial sample, we quantified the random uncertainty introduced by modeling triaxial galaxies with axisymmetric JAM.
Results. We show that spherical JAM analysis of spatially unresolved kinematic data introduces a bias of up to 2–4% (depending on the intrinsic shape of the lens) in the inferred MSD. Our model largely corrects this bias, resulting in a residual random uncertainty in the range of 0–2.2% in the stellar velocity dispersion (0–4.4% in H_0), depending on the projected ellipticity and the anisotropy of the stellar orbits. This residual uncertainty can be further mitigated by the use of spatially resolved kinematic data, which constrain the intrinsic axis ratio. We also show that the random uncertainty in the kinematics recovery using axisymmetric JAM for axisymmetric galaxies is at the level of 0.24% in the velocity dispersion, and the uncertainty using axisymmetric JAM for triaxial galaxies is at the level of 0.17% in the velocity dispersion.

Key words. gravitational lensing: strong – galaxies: kinematics and dynamics – distance scale

1. Introduction

Strong gravitational lenses are a powerful tool for probing absolute distances in the Universe and constraining its expansion history. Time-delay measurements, which use the relative time delays between the multiple images of a time-varying source (Refsdal 1964), provide precise constraints on the Hubble constant, H_0 (e.g., Kundić et al. 1997; Schechter et al. 1997; Koopmans et al. 2003; Kochanek et al. 2006; Saha et al. 2006; Oguri 2007; Suyu et al. 2010, 2013, hereafter time-delay cosmography, TDC). Detailed reviews of TDC are presented in e.g., Treu & Marshall (2016), Birrer et al. (2024).

A key limiting factor on the constraining power of TDC is the mass sheet degeneracy (MSD; Falco et al. 1985), which leaves the relative lensing observables invariant except for the time-delay prediction. To acquire a precise and accurate measurement of the Hubble constant, the MSD must first be broken. Currently, the primary method to break the MSD is to use stellar kinematics of the deflector galaxy to constrain the 3D gravitational potential of the deflector (Treu & Koopmans 2002; Suyu et al. 2010; Barnabè et al. 2011).

For example, Birrer et al. (2020) (hereafter, TD4), exclusively constrained the MSD component of the deflector's density profile with measurements of the stellar velocity dispersion. The measured Hubble constant by TD4 of the seven TDCOSMO lensed quasars is $H_0 = 74.5^{+5.6}_{-6.1}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, consistent with the $H_0 = 73.3^{+1.7}_{-1.8}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ value measured by the H0LiCOW collaboration (Wong et al. 2020), assuming assertive mass profiles. Since TD4 used a maximally flexible parameterization for the mass model with regard to the MSD in a Bayesian hierarchical framework, the uncertainty of the Hubble constant increases from two percent to eight percent, with the major component in the error budget being the precision in the kinematic observations and related modeling uncertainties. To further improve the precision, TD4 imposed population prior from the stellar velocity dispersion profiles of 33 Sloan Lens ARCS (SLACS) lenses (Bolton et al. 2006; Shu et al. 2015; Shajib et al. 2021) and improved the uncertainty of H_0 to five percent, giving $H_0 = 67.4^{+4.1}_{-3.2}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$; however, this value should be considered illustrative of the uncertainty rather than of

* Corresponding author: xiangyu.huang.1@stonybrook.edu

the best estimate, owing to insufficient quality of the SDSS stellar velocity dispersions (Knabel et al. 2025a). The key assumption of the TDCOSMO + SLACS analysis is that the external lens sample (i.e., the SLACS lenses) are from the same parent population as the TDCOSMO lenses. Selection effects in the specific lens sample and across different samples must be studied and mitigated appropriately to the level of the stated precision in H_0 (e.g., Sonnenfeld 2024).

Another assumption in TD4 is similar to that of previous studies, namely, that the kinematics interpretation assumed a spherical dynamical model, while real galaxies are in general nonspherical. Therefore, the intrinsic shape and the orientation of the lens galaxy can be one potential source of systematic uncertainties, since the lensing and kinematics observables change with the viewing angle. In this paper, we use the term "projection effect" to refer to the phenomenon that the lensing and kinematics observables of a galaxy depend on the orientation relative to the observer. For example, an oblate lens galaxy which appears edge-on to an observer will have more projected mass-density at its center and therefore will have a larger Einstein radius, while with a face-on orientation, it will have the smallest Einstein radius. For a spherical galaxy of the same total mass, the Einstein radius falls between the maximum and minimum Einstein radius created by the oblate galaxy. For lens mass modeling, the deprojection of the lens galaxy is unnecessary since all lensing quantities are determined by the projected mass-density. In contrast, kinematics modeling, which is used to constrain the 3D mass distribution and break the MSD, requires the deprojection of the lens galaxy in light and mass (e.g., Dutton et al. 2011).

For an individual galaxy, the deprojection for the lens mass from photometric data is in general underconstrained. The deprojection for the kinematics model is possible through integral field spectroscopy (IFS, e.g., Cappellari 2008) only under certain assumptions on the stellar kinematics anisotropy. Population-level deprojection of galaxies are more applicable, since the intrinsic shape distribution can be obtained by statistically inverting the distribution of projected ellipticities of a sample of galaxies (Sandage et al. 1970; Benacchio & Galletta 1980; Binney & de Vaucouleurs 1981; Lambas et al. 1992; Ryden 1992; Vincent & Ryden 2005; Kimm & Yi 2007; Padilla & Strauss 2008; Rodríguez & Padilla 2013) under the assumption of isotropy of the sample; alternatively, another approach is to invert the distribution of the projected ellipticities and the misalignment angle between the photometric and the kinematic axes (Weijmans et al. 2014; Li et al. 2018; Ene et al. 2018), despite obtaining potentially nonunique solutions. The intrinsic shape distribution of galaxies can be used to model the projection effects and selection effects under some selection criterion. To do so, a thorough understanding of the intrinsic galaxy population acting as strong lenses is required.

Most strong gravitational lenses discovered so far are massive early-type galaxies (ETGs), which are the most massive galaxies in the Universe. This is a direct consequence of the lensing cross-section. The images of the background source must have large enough angular separations to be identified as multiply lensed images. In regard to their dynamical structures, Barnabè et al. (2011) showed that a sample of early-type galaxies from the SLACS Survey can be divided into the two usual kinematics classifications, slow rotators and fast rotators (e.g., Emsellem et al. 2007; Cappellari 2016). More specifically, Li et al. (2018) modeled the projected shape of a complete sample of all 189 massive ($M_* > 2 \times 10^{11} M_\odot$) slow rotator ETGs out of a sample of 2200 galaxies from the DR14 of the

SDSS MaNGA IFU survey. They found a weakly triaxial shape consistent with a dominant triaxial-oblate population under a wide set of model assumptions and no significant fraction of prolate-triaxial galaxies. Fast-rotator ETGs, on the other hand, are found to be much flatter and oblate-like (Weijmans et al. 2014), showing consistency with their dynamical status. For either type of ETG, kinematics modeling with state-of-the-art Jeans anisotropic modeling (JAM) method (Cappellari 2008, 2020) can aptly recover the 3D kinematics model, which is consistent (in projection) with the observational data, under certain assumptions on the intrinsic shape of the ETG (e.g., assuming spherical or axial symmetry of the gravitational potential). In the axisymmetric case, the projection effect of the 3D kinematics model will impact the interpretation of the kinematics in TDC. The projection effect in the kinematics model also leads to potential selection bias under lens selection criterion in the lens finding stage and follow-up analysis, due to the change in lensing efficiency with orientation.

The goal of this paper is to investigate the projection and selection effects introduced by the intrinsic shape of the mass distribution and the kinematics stellar tracer distribution of the lensing galaxies. As the projection effect affects the observation and interpretation of lensing and kinematics observables, we aim to assess qualitatively and quantitatively its impact on the measurement of the Hubble constant via TDC. For context, we note that uncertainty or bias in the stellar velocity dispersion σ affects the determination of the Hubble H_0 as $\delta H_0/H_0 = 2\delta\sigma/\sigma$. Therefore, if we aim to measure H_0 with a 2% accuracy, biases on σ predicted by the lens model need to be contained below 1%.

To achieve our goal, we study the projection and selection effects in the lensing observables using two galaxy samples: one is a synthetic axisymmetric sample, which resembles the SLACS lenses, and the other is a triaxial ETG sample from the IllsutrisTNG-100 simulation. We created a catalog of mock lensing observables (e.g., projected ellipticities and the Einstein radii) for both samples, assuming the viewing angle is random. We quantified the projection effect of individual galaxies with the scatters in the Einstein radius. We qualitatively analyzed the selection effect by applying a selection criterion on the Einstein radius. On the kinematics side, using the synthetic axisymmetric sample, we created mock kinematics observables and analyzed the projection and selection effects specific to axisymmetric systems. We then quantified the accuracy in the kinematics recovery under different assumptions on the galaxy model with JAM. Specifically, we explicitly derived the quantitative corrections in the interpretation of the kinematics observables as a function of measured projected ellipticity and estimated the residual uncertainty for the corrected velocity dispersion. We show that it is possible to constrain potential H_0 biases below the target 1% precision level. We review the key axisymmetric assumption in the kinematics modeling by comparing the kinematics observables from the axisymmetric kinematics models of triaxial systems. We also obtain a numerical calculation of an upper limit of the relative difference in the velocity dispersion using the triaxial TNG ETG sample, which is more elliptical and triaxial than real lensing galaxies such as the SLACS lenses at the population level.

This paper is organized as follows. Section 2 discusses the theoretical basis including lensing basics, projection formalism, and kinematics modeling. Section 3 discusses our lens sample construction. Section 4 shows the projection and selection effect in the lensing observables of a triaxial catalog of ETGs from the TNG100 simulation. In Section 5, we describe how we generated mock kinematics data using axisymmetric JAM for a syn-

thetic axisymmetric lens sample presented in Section 3. Then, we discuss the selection effect for axisymmetric systems. We also compare the kinematics model assumptions by estimating their relative biases in the kinematics recovery. In Section 6, we discuss the effect of triaxiality in the kinematics recovery of ETG using axisymmetric JAM. We present our conclusions in Section 7. Throughout this work, we assume a flat ΛCDM cosmology with $H_0 = 70 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$, where required. We note that this arbitrary choice does not affect any of our conclusions regarding the selection function and quantitative correction terms.

2. Basics of time-delay cosmography

This section reviews the theoretical basis of this work, including strong lensing observables and formalism (Section 2.1), the mass sheet degeneracy (Section 2.2), projection of triaxial systems (Section 2.3), kinematics modeling with JAM (Section 2.4), and the projection formalism of axisymmetric multi-Gaussians (Section 2.5).

2.1. Strong lensing formalism

The deflection of light follows the lensing equation,

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{1}$$

where $\boldsymbol{\beta}$ is the unlensed angular source position, $\boldsymbol{\theta}$ is the angular image position seen from the observer, and $\boldsymbol{\alpha}$ is the deflection angle as seen on the sky. The deflection angle can be expressed as the gradient of the lensing potential

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \nabla \psi(\boldsymbol{\theta}), \tag{2}$$

which is related to the total projected mass-density on the lens plane via

$$\kappa(\boldsymbol{\theta}) = \frac{1}{2} \nabla^2 \psi. \tag{3}$$

Here $\kappa(\boldsymbol{\theta})$ is the convergence, defined as

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_{\mathrm{crit}}} \tag{4}$$

under small angle approximation. $\Sigma(\boldsymbol{\theta})$ is the total projected excess mass-density compared to the cosmological background density on the lens plane, and Σ_{crit} is the critical density for the lens-source configuration

$$\Sigma_{\mathrm{crit}} = \frac{c^2}{4\pi G} \frac{D_{\mathrm{s}}}{D_{\mathrm{ds}} D_{\mathrm{d}}}. \tag{5}$$

Here, D_{s}, D_{d}, and D_{ds} are the angular diameter distances to the source, to the deflector, and from the deflector to the source. The strong lensing efficiency, characterized by the deflection angle α, is determined by the total projected mass within the Einstein radius θ_E. The Einstein radius, θ_E, is the angular radius within which the mean convergence is unity:

$$\int_{\mathcal{A}(<\theta_{\mathrm{E}})} \kappa(\boldsymbol{\theta}) d\boldsymbol{\theta} \equiv \mathcal{A}(<\theta_{\mathrm{E}}). \tag{6}$$

Here, $\mathcal{A}(<\theta_{\mathrm{E}})$ is an area with θ_E as the effective radius $\mathcal{A}(<\theta_{\mathrm{E}}) \equiv \pi \theta_{\mathrm{E}}^2$. We specify here that the definition of the Einstein radius in this paper is the circularized value over the lens plane, i.e., $\sqrt{q'} \theta_E^{\mathrm{maj}}$, where q' is the apparent axis ratio of the iso-density

contours, and θ_E^{maj} is the Einstein radius measured along the major axis of the elliptical iso-density contours.

The time delays between the different images are proportional to the difference in their Fermat potential:

$$\Delta t_{AB} = \frac{1}{c} D_{\Delta t} \left[\phi(\boldsymbol{\theta}_A, \boldsymbol{\beta}) - \phi(\boldsymbol{\theta}_B, \boldsymbol{\beta}) \right], \tag{7}$$

where c is the speed of light, A and B label two different images of the same background source, $\phi(\boldsymbol{\theta}, \boldsymbol{\beta})$ is the Fermat potential

$$\phi(\boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{(\boldsymbol{\theta} - \boldsymbol{\beta})^2}{2} - \psi(\boldsymbol{\theta}), \tag{8}$$

and $D_{\Delta t}$ is a coefficient with the dimension of distance, also called the time-delay distance

$$D_{\Delta t} = (1 + z_{\mathrm{d}}) \frac{D_{\mathrm{d}} D_{\mathrm{s}}}{D_{\mathrm{ds}}}. \tag{9}$$

Then, with a measurement of the time delay Δt and an inference on the difference of the Fermat potential, one can measure the time-delay distance. The Hubble constant H_0 scales inversely with the angular diameter distance, and thus

$$H_0 \propto D_{\Delta t}^{-1}. \tag{10}$$

2.2. The mass sheet degeneracy

The mass sheet degeneracy (MSD; Falco et al. 1985) stems from the so-called mass sheet transform (MST) on the convergence and the unknown source position

$$\kappa(\boldsymbol{\theta}) \rightarrow \kappa_\lambda(\boldsymbol{\theta}) = \lambda \kappa(\boldsymbol{\theta}) + 1 - \lambda, \tag{11}$$
$$\boldsymbol{\beta} \rightarrow \boldsymbol{\beta}_\lambda = \lambda \boldsymbol{\beta}. \tag{12}$$

Here, $\boldsymbol{\theta}$ is the angular position of the lensed image of the source and $\boldsymbol{\beta}$ is the angular position of the unlensed source, as in Eq. (1). Here $(1 - \lambda)$ acts as an infinite sheet of mass. The MST preserves the image configuration and the relative magnification between the multiple images of the source, but changes the time delay prediction under a fixed lens mass model by

$$\Delta t_\lambda = \lambda \Delta t. \tag{13}$$

The time-delay distance is then transformed via

$$D_{\Delta t, \lambda} = \lambda^{-1} D_{\Delta t}. \tag{14}$$

The inferred Hubble constant then changes via

$$H_{0, \lambda} = \lambda H_0. \tag{15}$$

The MSD effect can be divided into two components regarding the source of the over- or under-density with respect to the background: (a) the internal MSD λ_{int}, which is associated with the mass distribution of the deflector galaxy due to changes in the radial density profile; and (b) the external MSD, contributed by the projected mass of the line-of-sight (LoS) structure, κ_{ext}, outside the associated halo.

The total MSD (Schneider & Sluse 2013; Birrer et al. 2020) can be expressed as

$$\lambda \approx (1 - \kappa_{\mathrm{ext}}) \lambda_{\mathrm{int}}. \tag{16}$$

Both the internal and the external MSD have an effect on the deflector potential. To first order, the total MSD can be constrained with stellar kinematics. In this work, we mainly focus on the projection effect caused by the surface density of deflector galaxy itself; thus, we refer to $\lambda \approx \lambda_{\mathrm{int}}$ for the remainder of this paper, noting that the external component has an additional impact on the kinematics (e.g., Fassnacht et al. 2006).

2.3. Projection effect and projection formalism for triaxial systems

According to Eq. (4), all lensing quantities, including the Einstein radius, can be derived from the 2D surface mass-density,

$$\Sigma(x', y') = \int_{-\infty}^{\infty} \rho(x', y', z') \, dz'. \tag{17}$$

Here (x', y') are coordinates on the lens plane and z' is the third dimension along the LoS. If the intrinsic mass-density distribution is nonspherical, $\Sigma(x', y')$ is dependent on the LoS direction. In this paper, the term "projection effect" specifically refers to the change in the surface mass-density, light surface brightness, and the stellar kinematics under different viewing angles of the lens galaxy.

The projection formalism for density profiles stratified on similar concentric spheroids is presented in detail by Stark (1977) and Binney (1985). Here we follow the convention of Binney (1985) and briefly summarize it. We assume the mass-density profile of a triaxial galaxy has a constant radial shape, without angular twists or changes in the axis ratio as a function of radius. With this assumption, the density is only a function of the ellipsoidal radius variable, namely,

$$\rho = \rho(a_v), \tag{18}$$

with

$$a_v = Z\sqrt{x^2 + \frac{y^2}{p^2} + \frac{z^2}{q^2}}. \tag{19}$$

Here, the ellipsoid has its major, intermediate, and minor axis aligned with the x, y, and z axis of the Cartesian coordinate system, respectively; p is the axis ratio between the intermediate and major axis of the ellipsoid and q is the axis ratio between the minor and major axis of the ellipsoid. $1 \geq p \geq q$.

The coefficient Z is a normalization factor added to preserve mass when varying the shape parameters p and q. There are two choices to preserve the total integrated mass when varying the shape parameters: (a) $Z = 1$ (not rescaling the radius) and renormalize the density uniformly in $\rho(a_v)$, and (b) $Z = (pq)^{1/3}$ without any renormalization of the overall density. A detailed comparison of the two choices for the density normalization is discussed in Appendix A.

When a galaxy is viewed along a LoS characterized by the polar angle θ and the azimuthal angle ϕ in a spherical coordinate system, its projected mass-density in Eq. (17) can be written as

$$\Sigma(x', y') = \frac{2}{\sqrt{f}} \int_0^{\infty} \rho\left(z''^2 + a_s^2\right) dz''. \tag{20}$$

Here, z'' is an integration variable, and a_s is the elliptical radius variable of the iso-density contour in projection,

$$a_s^2 = \frac{1}{f}\left[Ax'^2 + Bx'y' + Cy'^2\right]. \tag{21}$$

Parameters A, B, C, and f are solely determined by the LoS direction (θ, ϕ) and the intrinsic axis ratios p and q as expressed

in Eqs. (6, 11) in Binney (1985), namely,

$$A = Z^2 \left(\frac{\cos^2\theta}{q^2}\left(\sin^2\phi + \frac{\cos^2\phi}{p^2}\right) + \frac{\sin^2\theta}{p^2}\right), \tag{22}$$

$$B = Z^2 \cos\theta \sin 2\phi \left(1 - \frac{1}{p^2}\right)\frac{1}{q^2}, \tag{23}$$

$$C = Z^2 \left(\frac{\sin^2\phi}{p^2} + \cos^2\phi\right)\frac{1}{q^2}, \tag{24}$$

$$f = Z\left(\sin^2\theta\left(\cos^2\phi + \frac{\sin^2\phi}{p^2}\right) + \frac{\cos^2\theta}{q^2}\right). \tag{25}$$

The apparent axis ratio $q'(q' \leq 1)$ of the projected elliptical iso-density contour is

$$q'(\theta, \phi, p, q) = \sqrt{\frac{A + C - \sqrt{(A-C)^2 + B^2}}{A + C + \sqrt{(A-C)^2 + B^2}}}. \tag{26}$$

We define the projected ellipticity of the surface mass-density as $e(\theta, \phi, p, q) = (1 - q')/(1 + q')$.

The projection formalism also works for the stellar luminosity component of the galaxy, trading the intrinsic mass-density, $\rho(a_v)$, for the intrinsic luminosity distribution, $l(a_v)$.

2.4. Stellar kinematics modeling

Stellar kinematics data provide an independent measurement of the gravitational potential of the deflector galaxy and thus can break the MSD. The luminosity-weighted LoS velocity dispersion σ^P is effective in constraining the mass-density slope of elliptical galaxies via joint lensing and kinematics analysis under certain assumptions on the stellar anisotropy profile (Treu & Koopmans 2002). For a given stellar tracer distribution, the squared velocity dispersion is proportional to the mass-density or, equivalently, the mass-to-light ratio, M/L. For $\lambda \approx 1$ and near the galaxy center $\theta \approx 0$ where the surface density is large, the first term in Eq. (11) dominates over the second and thus $\kappa_\lambda(\theta) \approx \lambda\kappa(\theta)$. In this limit, the MST has the same effect as varying the total M/L, and the velocity dispersion is transformed via

$$\sigma_\lambda^P \approx \sqrt{\lambda}\sigma^P. \tag{27}$$

This is equivalently $\sigma_*^P \approx \sqrt{\lambda}\sigma_l^P$, where σ_*^P is the stellar velocity dispersion and σ_l^P is the modeled velocity dispersion derived from the best fit lens model of the imaging data. The inferred λ from comparing observed and predicted velocity dispersion can then be applied to correct the lensing potential and, hence, H_0 in the inference (Birrer et al. 2020). The relative uncertainty in constraining H_0 that comes from the uncertainty in the velocity dispersion (model or measurement) is transformed via

$$\frac{\delta H_0}{H_0} = \frac{\delta\lambda}{\lambda} = 2\frac{\delta\sigma^P}{\sigma^P}. \tag{28}$$

The kinematic modeling methods for galaxies can be grouped into three categories. First, we have particle-based methods (e.g., made-to-measure; Syer & Tremaine 1996) and orbit-based methods (e.g., Schwarzschild; Schwarzschild 1979) are highly flexible and can describe complex systems, including triaxial and time-dependent configurations. They fit the data with sets of orbital or particle weights. These approaches are particularly powerful when high-quality kinematic data provide

higher-order Gauss-Hermite moments beyond the mean velocity and velocity dispersion. The main drawbacks are computational cost, especially for joint lensing and kinematics inference, and discreteness noise. Most importantly, the stellar distribution function (DF) is represented implicitly by the fitted weights. This implicit form makes the methods unsuitable for tasks that require an explicit, parameterized DF, such as constructing mock datasets. Second, we have analytic DF-based solutions (e.g., Mamon & Łokas 2005) provide explicit functional forms. They are well suited to direct analysis and forward modeling, but their applicability is usually limited to simple symmetries, such as spherical or basic axisymmetric configurations. Third, we have solutions of the Jeans equations (Jeans 1922), in particular, the Jeans anisotropic modeling (JAM) method (Cappellari 2008, 2020), which offer an effective alternative. JAM solves the Jeans equations for axisymmetric or spherical systems and allows for velocity anisotropy. It uses multi-Gaussian expansion (MGE; Emsellem et al. 1994; Cappellari 2002) to approximate the mass and tracer densities. Because JAM provides an explicit, parameterized description of the kinematic moments and the anisotropy, it enables direct computations of the observables and the efficient construction of mock datasets, as required for the present work.

In this work, we use the spherically aligned axisymmetric JAM framework (Cappellari 2020), which has the velocity ellipsoid aligned with a spherical coordinate system, for the dynamical modeling of stellar kinematics observables. Compared with the cylindrically aligned JAM proposed in Cappellari (2008), the spherical-aligned JAM provides a better estimate of the stellar kinematics for slow-rotator type ETGs, as most of the strong gravitational lenses are of this type (Knabel et al. 2025b), because the slow-rotator ETG population are found to have a generally rounder or weakly triaxial shape (Cappellari 2016; Li et al. 2018).

We present a brief description of the formalism of spherically aligned JAM. The stellar motion in a galaxy under equilibrium can be described with a six-dimensional (6D) distribution function (DF) $f(\boldsymbol{x}, \boldsymbol{v})$, which follows the steady-state, collisionless Boltzmann equation under a smooth gravitational potential $\Phi(\boldsymbol{x})$ (Eqs. (4)–(13b), Binney & Tremaine 1987), expressed as

$$\sum_{i=1}^{3} \left(v_i \frac{\partial f}{\partial x_i} - \frac{\partial \Phi}{\partial x_i} \frac{\partial f}{\partial v_i} \right) = 0. \tag{29}$$

Equation (29) can be solved under some simplifying assumptions: here we assume axial symmetry, namely, $\partial \Phi / \partial \phi = 0$, $\partial f / \partial \phi = 0$, where ϕ is the polar angle in a spherical coordinate, with the azimuthal angle, $\theta = 0$, aligned with the symmetry axis. We also assume that the velocity ellipsoid is aligned with the spherical coordinate, i.e., $\langle v_r v_\theta \rangle = 0$. Eq. (29) is then reduced to (de Zeeuw et al. 1996; Cappellari 2020)

$$\frac{\partial \left(\rho_* \langle v_r^2 \rangle \right)}{\partial r} + \frac{(1+\beta) \rho_* \langle v_r^2 \rangle - \rho_* \langle v_\phi^2 \rangle}{r} = -\rho_* \frac{\partial \Phi}{\partial r}, \tag{30}$$

$$(1-\beta) \frac{\partial \left(\rho_* \langle v_r^2 \rangle \right)}{\partial \theta} + \frac{(1-\beta) \rho_* \langle v_r^2 \rangle - \rho_* \langle v_\phi^2 \rangle}{\tan \theta} = -\rho_* \frac{\partial \Phi}{\partial \theta}, \tag{31}$$

which are the so-called Jeans (Jeans 1922) equations under axial symmetry. Here ρ_* is the density of the dynamical tracer,

$$\rho_* = \int f(\boldsymbol{x}, \boldsymbol{v}) d\boldsymbol{v}, \tag{32}$$

$\rho_* \langle v_i v_j \rangle$ are the elements of the stress tensor,

$$\rho_* \langle v_i v_j \rangle = \int v_i v_j f(\boldsymbol{x}, \boldsymbol{v}) d\boldsymbol{v}, \tag{33}$$

and β is defined as the anisotropy parameter in the spherical coordinate,

$$\beta = 1 - \frac{\langle v_\theta^2 \rangle}{\langle v_r^2 \rangle}. \tag{34}$$

Using a set of multi-Gaussian decomposition (Emsellem et al. 1994; Cappellari 2002) for the mass-density generating the gravitational potential and the intrinsic density of the dynamical tracer, and by assuming a spatially constant anisotropy parameter β and a constant mass-to-light ratio, M/L, for each Gaussian, JAM solves Eqs. ((30)–(31)) and predicts the LoS luminosity-weighted second moment of the velocity

$$\Sigma \langle (\sigma^P)^2 \rangle (x', y') = \int \rho_* \langle v_z'^2 \rangle dz', \tag{35}$$

where $\Sigma(x', y')$ is the surface brightness of the galaxy.

The observed second moment of velocity is affected by the atmospheric seeing and the instrumental response. The prediction for the observed second moment within an aperture, $\mathcal{A}$, is then the aperture-integrated, luminosity-weighted, seeing-convolved dispersion (Suyu et al. 2010)

$$\langle (\sigma^P)^2 \rangle_{\mathcal{A}} = \frac{\int_{\mathcal{A}} \left[\Sigma(x', y') \langle (\sigma^P)^2 \rangle * \mathcal{P} \right] dx' dy'}{\int_{\mathcal{A}} \left[\Sigma(x', y') * \mathcal{P} \right] dx' dy'}, \tag{36}$$

where $* \mathcal{P}$ denotes a convolution with the PSF $\mathcal{P}$.

2.5. Axisymmetric MGE projection formalism

The projection formalism for axisymmetric systems, where the axisymmetric Jeans equations can be solved using the MGE parameterization of the stellar light profile, can be simplified from the projection formalism described in Sect. 2.3 by taking $p = 1$ for oblate systems, and $p = q$ for prolate systems. However, when using the MGE parameterization, the projected axis ratio follows a simpler calculation and the projected profile can be calculated using the conservation of total luminosity instead of the brute-force numerical integrals. Here, we briefly summarize the projection formalism, following the derivation in Cappellari (2020).

The intrinsic luminosity profile of a galaxy $\nu(r, \psi)$ can be expressed with

$$\nu(r, \psi) = \sum_{k=1}^{N} \nu_{0k} \exp\left[-\frac{r^2}{2\sigma_k^2} \left(\sin^2 \psi + \frac{\cos^2 \psi}{q_k^2} \right) \right], \tag{37}$$

where ψ is the angle between the field point (r, ψ) and the symmetry axis of the ellipsoid, r is the distance to the galaxy center, ν_{0k}, σ_k and q_k are the peak, dispersion, and the intrinsic axis ratio of the k-th MGE component. The intrinsic axis ratio q_k is $q_k < 1$ for oblates and $q_k > 1$ for prolates. The surface brightness profile $\Sigma(x', y')$ is then

$$\Sigma(x', y') = \sum_{k=1}^{N} \Sigma_{0k} \exp\left[-\frac{1}{2\sigma_k^2} \left(x'^2 + \frac{y'^2}{q_k'^2} \right) \right], \tag{38}$$

where x' is aligned with the photometric major axis. Similar to the notation of q_k, q_k' has $q_k' < 1$ for oblates, and $q_k' > 1$ for prolates.

The orientation of an axisymmetric system is characterized by the inclination angle, i, which is the angle between the LoS and the symmetry axis of the axisymmetric ellipsoid. For

isotropic inclination angles, $\cos i \in \mathcal{U}[0, 1]$. The the projected axis ratio, q', for an oblate galaxy is computed directly with

$$q'^2_k = q_k^2 \sin^2 i + \cos^2 i, \tag{39}$$

where the subscript k labels the k-th MGE component. For a prolate galaxy,

$$\frac{1}{q'^2_k} = \frac{\sin^2 i}{q_k^2} + \cos^2 i. \tag{40}$$

The total mass or total luminosity of each MGE component is conserved in the projection, leading to

$$\Sigma_{0k} = \frac{q_k \nu_{0k} \sigma_k \sqrt{2\pi}}{q'_k}. \tag{41}$$

3. Lens sample construction

In this section, we discuss our lens sample construction. We used two different lens sets for the analysis of the projection effect and the selection effect in lensing and kinematics. For the analysis on the kinematics, we used a synthetic axisymmetric lens sample similar to the SLACS lenses in the $R_e - \sigma_v$ space, with empirical intrinsic shape distributions. For the lensing analysis, we used a triaxial ETG sample from the TNG100 simulation. We also used this triaxial sample for the analysis of the triaxiality in axisymmetric kinematics models (Section 6). In Section 3.1, we describe these two samples. In Section 3.2, we describe our choice of the mass, light, and anisotropy profiles for the lens samples. In Section 3.3, we list the simplifying assumptions on the kinematics models of our sample.

3.1. Sample properties

3.1.1. The synthetic lens set

We generated an axisymmetric synthetic lens sample using priors from the SLACS lenses (Bolton et al. 2008). We used the velocity dispersion of the SLACS as a description of the total mass normalization, along with the effective radius as a description of the shape of the stellar tracer profile. In practice, we sample velocity dispersion and effective radius from the 2D kernel density estimator (KDE) constructed using the grade "A" lenses in Table 4 of Bolton et al. (2008). Figure 1 shows the 2D KDE, the original data points of the SLACS lenses in the $\sigma_v - R_e$ space, and the synthetic lens sample. The size of the synthetic lens sample was set as 600.

We then assigned to each lens an intrinsic axis ratio drawn from $\mathcal{N}[\mu = 0.74, \sigma = 0.08]$ for the oblate sample and $\mathcal{N}[\mu = 0.84, \sigma = 0.04]$ for the prolate sample. These values are inferred from the SDSS MaNGA slow-rotator ETG (Table 1 of Li et al. 2018, line 1 for the oblate shape distribution, and line 4 for the prolate shape distribution) using the relation between the misalignment angles between the kinematics major axis and the photometric major axis, and the observed ellipticities.

3.1.2. The TNG100 triaxial ETG

We also used a triaxial ETG catalog from the TNG100 simulation to illustrate the projection effect of triaxial systems. The IllustrisTNG is a set of large-scale, gravity+magnetohydrodynamical simulations for the study of galaxy formation and evolution (Pillepich et al. 2018), including three
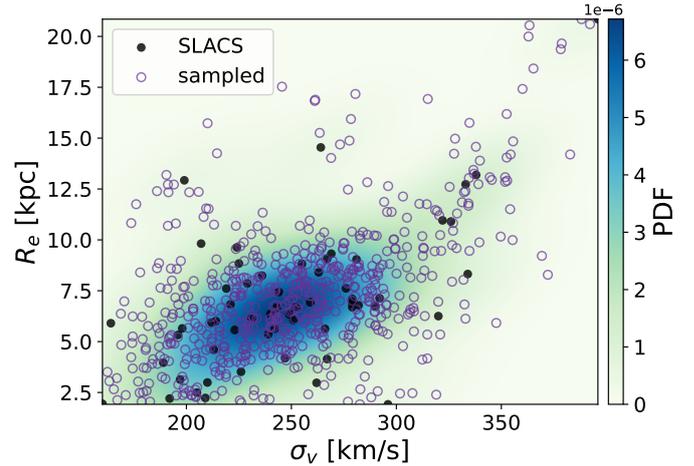
Fig. 1. Axisymmetric synthetic lens set in the $\sigma_v - R_e$ space. Black dots are the SLACS grade "A" lenses reported in Bolton et al. (2008). The purple circles are drawn from the 2D KDE of the original data points.

primary runs with different volumes and resolutions, the TNG300, the TNG100, and the TNG50, with side length ~300, 100, and 50 Mpc, respectively. The catalog we used was also selected in Pulsoni et al. (2020), spanning a stellar mass range of $10^{10.3} - 10^{12}$ $M_\odot$, used to study the relation between the stellar kinematics and the intrinsic shape at different radial scales. The selection was performed on snapshots at $z = 0$ in the color-stellar mass diagram, in which the selected ETGs are confirmed to match with the observed ETGs in IFU surveys such as the SDSS MaNGA (Bundy et al. 2015) and SAMI (Croom et al. 2012). For a detailed description of their selection, we refer to Section 5 of Pulsoni et al. (2020). For the selected sample, the physical properties were extracted from the simulated photometric and IFU data. We used the following in this work:

- Effective radius in projection, R_e: half-mass radius of the total bound stellar mass on the projected major axis when the projection is along one random LoS. For the remainder of this paper, we assume that R_e from the TNG catalog is the circularized half-light radius in projection, namely, $R_e^{\text{circ}} = \sqrt{q'} R_e^{\text{major}}$, where q' is the projected apparent axis ratio.
- Velocity dispersion at effective radius, σ_{rm}: velocity dispersion map is obtained for a random LoS. The mean velocity and velocity dispersion of stellar particles are calculated within a radial bin (annulus) around R_e. We specify here that we did not adopt σ_{rm} as the "true" velocity dispersion of the sample galaxies, but as a quantity to represent the total mass normalization depending on the mass-density profile of the lens galaxies.
- Intrinsic axis ratios: the intrinsic intermediate-to-major axis ratio, p, and minor-to-major axis ratio q characterizing the iso-density contours of each galaxy. It was calculated from the inertia tensor, $I_{ij} = \sum_n x_{n,i} x_{n,j} M_n / \sum_n M_n$, in each radial bin, where the summation $\sum_n$ was performed on the 50% nearest stellar particles, $x_{n,i}$ are the coordinates, and M_n are their mass. For simplicity of our strong lensing mock tests, we ignored the radial variation of p and q in the simulation and used the p and q value at $1R_e$ as constant axis ratios throughout all radii. We also assumed a perfect alignment between the stellar mass distribution, stellar luminosity, and total mass distribution; therefore, in the next steps of the
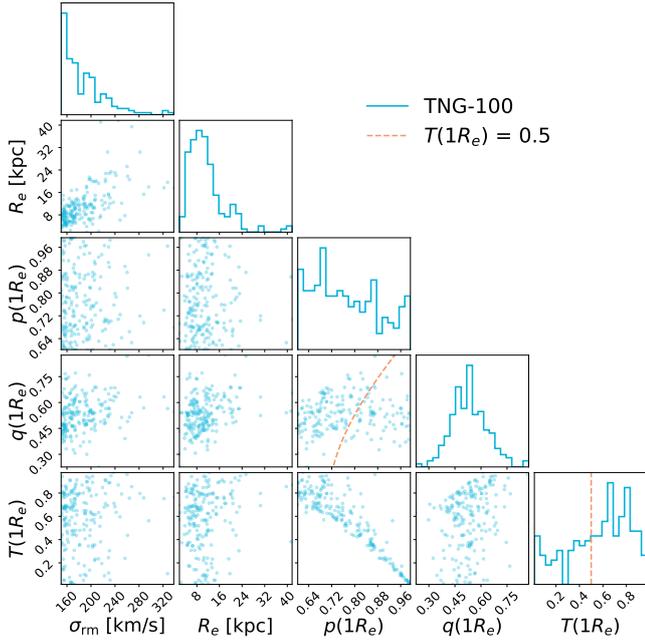
Fig. 2. Distribution of the velocity dispersion on a random LoS $\sigma_{\rm rm}$, the effective radius R_e, the intrinsic axis ratios p and q and the triaxiality parameter T for the selected ETG sample from them TNG100 simulation. Dashed line represent $T = 0.5$ in the $p − q$ space and in the 1D histogram of T, which we used to separate the full sample into the "oblate" subsample and the "prolate" subsample.

study, p and q were used to characterize the shape of the stellar tracer as well as the total density.

For the TNG sample, we used $\sigma_{\rm rm}$ as an additional mass cut, keeping only galaxies with $\sigma_{\rm rm} \geq 150$ km/s. This is to construct a more realistic lens sample as lens galaxies are generally the most massive ones with velocity dispersion in the range of 200–350 km/s (e.g., Bolton et al. 2006). We note however, that the distribution of stellar velocity dispersions and ellipticities for the TNG sample is much different than that found for SLACS galaxies. Even after the cut, the TNG sample peaks at the low end of the velocity dispersion distribution and has virtually no galaxies above 240 km/s. The projected ellipticity of the TNG sample under random projection (simulated in Section 4.2) is also generally larger than that of the SLACS lenses. Therefore, we consider the results based on the TNG sample to be representative of the maximum possible effects, since lens galaxies will generally be rounder and less triaxial, due to selection effect. We can use the triaxiality parameter,

$$T = \frac{1 - p^2}{1 - q^2}, \tag{42}$$

to characterize the intrinsic shape of the sample ETGs. For prolate ellipsoids, $T = 1$; for oblate ellipsoids, $T = 0$; for triaxial ellipsoids, $0 < T < 1$. Figure 2 shows the intrinsic axis ratios and triaxiality distribution of the TNG100 sample. The full TNG ETG sample is further divided into two subsamples: the oblate sample with $T \leq 0.5$ and the prolate sample with $T > 0.5$, since galaxies of different intrinsic shapes have different projection effects. Finally, among the TNG100 sample, the oblate subsample has 65 ETGs and the prolate subsample has 126 ETGs.

3.2. Mass-density, luminosity, and anisotropy assumptions

For our sample galaxies, we adopted an intrinsic galaxy model composed of a dark matter halo and a stellar component. The overall density profile is assumed to be a deformed (triaxial or axisymmetric) singular isothermal sphere (SIS), which approximates the population level density profile of massive ETGs (Koopmans et al. 2009). The Einstein radius of a SIS halo is directly linked to its velocity dispersion by

$$\theta_{\rm E} = 4\pi \left(\frac{\sigma_v}{c}\right)^2 \frac{D_{\rm ds}}{D_{\rm s}}. \tag{43}$$

To better conserve the mass when varying the intrinsic shape of the nonspherical density profiles, we manually add a characteristic "truncation" radius, r_c, with $r_c \gg \theta_E$, such that the overall density profile follows

$$\rho(r) \propto \frac{\sigma_v^2}{\left(\frac{r}{r_c}\right)^2 \left[1 + \left(\frac{r}{r_c}\right)^2\right]}. \tag{44}$$

In this work, r_c is chosen to be 200 times θ_E, far outside the strong lensing regime. This choice is arbitrary and it does not affect our conclusions. Adopting 100 or 300 would have yielded the same results. This is because strong lensing measurements do not have constraining power for the density profile far outside the Einstein radius. We then deformed the spherical iso-density contours into triaxial and axisymmetric ellipsoids, as described in more detail in Section 2.3 and Section 2.5. We chose the normalization factor $Z = (pq)^{1/3}$, such that the spherically averaged density profile is closer to that of a spherical model of the same mass (discussed in detail in Appendix A).

For the stellar luminosity, we adopted the Jaffe profile (Jaffe 1983), which approximates the de Vaucouleurs profile of elliptical galaxies in projection. The intrinsic luminosity profile follows

$$l(r) \propto \left(\frac{r}{r_s}\right)^{-2} \left[1 + \left(\frac{r}{r_s}\right)\right]^{-2}, \tag{45}$$

where $r_s = R_e/0.763$ with R_e being the half-light radius. The absolute normalization of the stellar luminosity is arbitrary, since the stellar motion in a galaxy is only dependent on the underlying total mass distribution and the shape of the tracer component. The predicted LoS velocity dispersion is weighted by the tracer density, whose absolute value is canceled out in the solution of the Jeans equation, expressed in Eqs. ((30)–(31)).

For the anisotropy in the stellar orbits, we adopted uniformly isotropic orbits, i.e., $\beta = 0$ at any radius. This choice falls in the range of β inferred using Schwarzschild's models from the SAURON project (Cappellari et al. 2007), and is shown in Figure 7 that it does not drastically affect the qualitative results. Without any loss of generality, we could make some simplifying assumptions about our sample of lenses:

1. We assumed all the lenses are at $z = 0.5$ and all the lensing sources are at $z = 1.5$.
2. For simplicity and to isolate the projection effect, we assumed that the density profile and luminosity profile of the galaxy are perfectly aligned.

3.3. Kinematics model

In the kinematics model construction with JAM, we used the galaxy model with an axisymmetrically deformed truncated SIS profile as the total mass-density profile and an axisymmetric

Jaffe profile as the stellar luminosity profile. For simplicity and to isolate the effect of the intrinsic shape and the inclination angle of the lens galaxy, we chose an isotropic anisotropy parameter (i.e., $\beta = 0$) at all radii, assuming no black hole mass contribution to the kinematics. We use the stellar dynamical modeling software[1] JAMPY (Cappellari 2008, 2020), which solves the Jeans equation in Eqs. ((30)–(31)) and computes the velocity dispersion map in projection. In Section 5.1, we present our mock velocity dispersion data, generated by projecting the 3D kinematics models onto random directions. We then analyze the projection effect and the selection effect.

4. Projection effects in strong lensing observables

In this section, we discuss the projection and selection effect in the strong lensing observables for triaxial lens galaxies. In Section 4.1, we describe the projection effect in the lensing observables for an individual lens galaxy. In Section 4.2, we first present the projection effect for a sample of triaxial lens galaxies from the TNG100 simulation and then apply lensing selections to model the selection effect for lens galaxies of different shapes.

4.1. Projection effect for a single triaxial lens galaxy

In this section, we present an example of the projection effect in the strong lensing observables of individual galaxies. We use the deformed truncated SIS profile in Eq. (44) as the total mass-density profile. We assume the viewing angle (θ, ϕ) is isotropic on a sphere, namely, $\cos\theta \in \mathcal{U}[0, 1]$ and $\phi \in \mathcal{U}[0, \pi]$. We project the mass-density profile 800 times and calculate the circularized Einstein radius θ_E and the projected ellipticity e. Figure 3 is an example with one triaxial ETG. Due to the projection effect, the mean of the Einstein radius is slightly different with the Einstein radius of a spherical lens of the same mass.

4.2. Projection effect for the TNG ETG sample

In this section, we model the projection effect of a triaxial lens galaxy sample. We randomly project the TNG100 ETG sample and calculate the circularized Einstein radius θ_E from the radial density profile. To generate more data points, we project each ETG in the sample along 4 random orientations. We quantify the scatter in the Einstein radius for the sample introduced by the projection effect using the mean of the relative standard deviation (RSD) as

$$\sum_j^N \frac{\sigma_{\theta_{E,j}}}{\langle \theta_{E,j} \rangle} \frac{1}{N}, \tag{46}$$

where j represents each individual lens, N is the total number of lenses, $\sigma_{\theta_{E,j}}$ is the standard deviation of the Einstein radius in different projections, and $\langle \theta_{E,j} \rangle$ is the mean Einstein radius in the projections. The mean scatter for the triaxial TNG100 ETG sample is calculated to be 7.8%. The projection effect in the Einstein radius then leads to the uncertainty in the deprojection of the best fit lens mass model for the kinematics prediction.

We define a "bias" term in the inferred SIS-model-equivalent velocity dispersion:

$$b_\sigma = \frac{\sigma_{SIS}}{\sigma_{rm}} - 1, \tag{47}$$
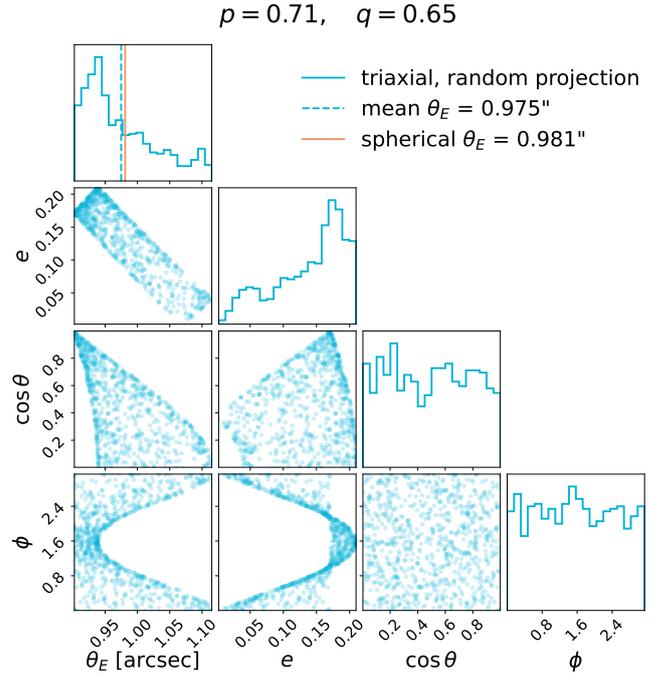
$$p = 0.71, \quad q = 0.65$$

Fig. 3. Illustration of the projection effect in the strong lensing observables. We start with 3D density profiles of lens galaxies and project onto random directions 800 times. Dashed blue line marks the mean of θ_E under random projections. Orange solid line marks the θ_E of a spherical lens of the same mass. The projection effect in lensing is reflected by the scattering of θ_E around the spherical value.

where σ_{SIS} is calculated by inverting Eq. (43) with θ_E. If the mass distribution within a lens is spherical, $b_\sigma = 0$. Assuming the SIS mass model for every lens, and taking σ_{rm} as the observed velocity dispersion, the mass sheet parameter λ is simply constrained by (from Eq. (27))

$$\lambda = \frac{1}{(b_\sigma + 1)^2}. \tag{48}$$

Figure 4 presents the lensing observables under the projection effect for the oblate subsample and the prolate subsample, with each blue dot representing a projection. Due to the projection effect, the means of the bias, b_σ, for both samples are overall biased low comparing to the spherical case.

4.3. Selection effect

Starting from the lensing observables distribution in Figure 4, we model the lensing selection effect for the lens galaxies with different shapes. We apply a lensing cross-section weighting proportional to θ_E^2 to the projected lensing observables (Sonnenfeld 2024). The distribution of the lensing selected quantities are shown with green and purple lines and dots in Figure 4. For the oblate sample, the lensing selection prefers more elliptical objects in projection, and higher θ, which is the inclination angle i if the galaxy is a pure oblate ($p = 1$). For the prolate sample, which is relatively rounder, the lensing selection slightly favors less elliptical objects in projection. The inclination angle i of a pure prolate ($p = q$) has $\cos i = \sin\theta \cos\phi$, and thus the selection in the viewing angles (θ, ϕ) is not directly visible from the corner plot.

For both samples, we observe that the bias b_σ is relieved with the lensing selection. This is because the projection effect
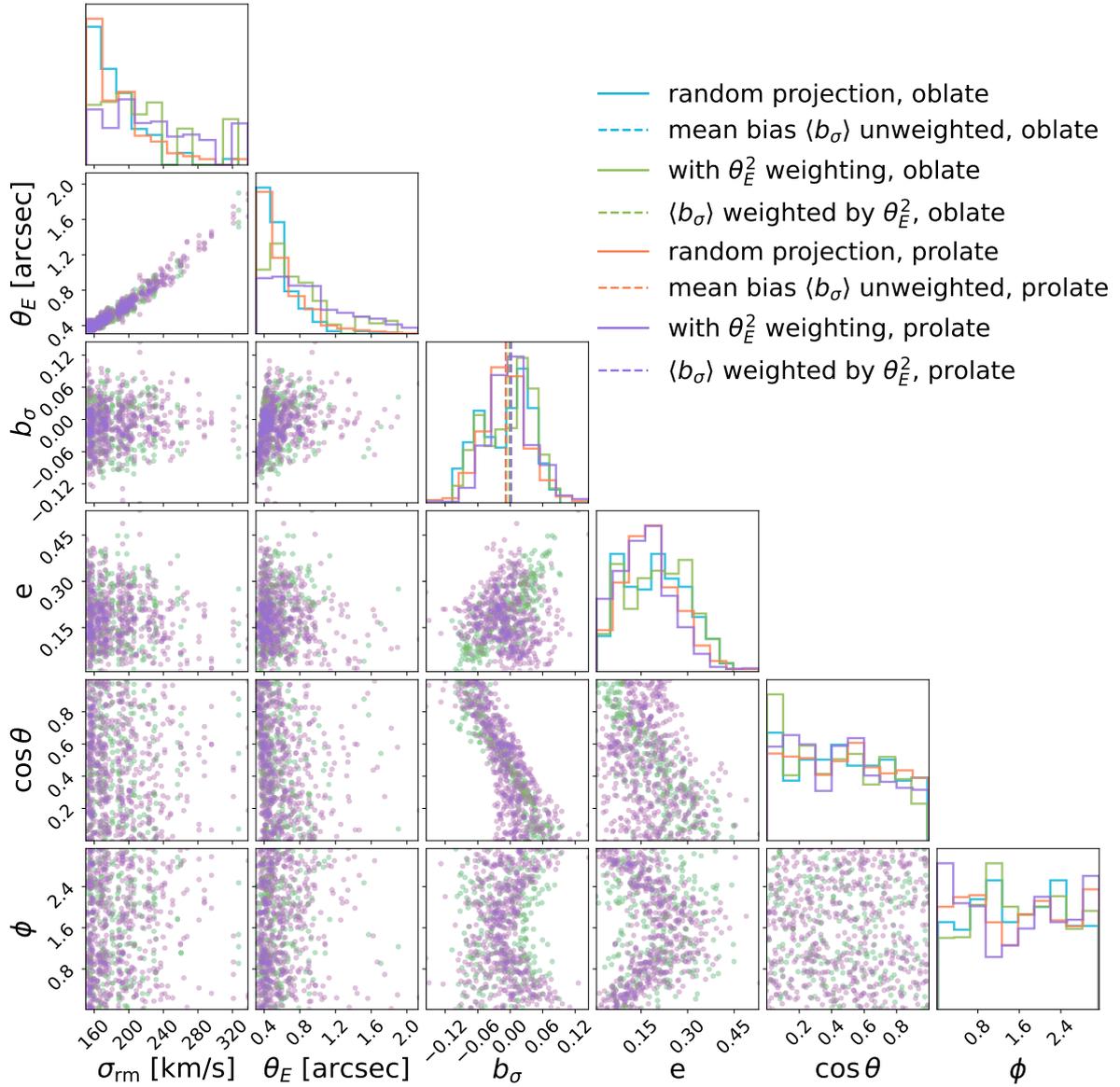
TNG100 ETG sample

Fig. 4. Projection effect of the TNG-100 ETG sample. Each ETG is projected 4 times onto random directions (blue and coral histograms). Each dot in the 2D histograms represents a projection. We apply a lensing cross-section weighting proportional to θ_E^2, as represented with the green and purple histograms. We model the lensing selection on the projected ellipticity e, showing that for the oblate sample, lensing selection favors more elliptical galaxies in projection, while for the prolate sample, the lensing selection favors rounder galaxies. For the oblates, the viewing angle θ equals the inclination angle i for pure oblates, and thus under lensing selection $\cos\theta$ is also inclined to the lower end, i.e., towards higher inclination angles and consequently higher ellipticities. The "bias" in the kinematics under the assumption of SIS lens models $b_\sigma = \sigma_{\mathrm{SIS}}/\sigma_{\mathrm{rm}} - 1$ is labeled with dashed lines. The mean bias $\langle b_\sigma \rangle$ for the oblate sample, indicated with blue dashed lines, is directly under the coral dashed line representing $\langle b_\sigma \rangle$ for the prolate sample and thus invisible from the plot.

of each individual galaxy leads to lower stacked surface density around the center, deviating b_σ from 0. The lensing selection favors more surface density, and therefore the bias b_σ moves in the direction of the spherical case (i.e., 0).

5. Projection effects in kinematics

In Section 4, we use the most simplified kinematics model (i.e., the SIS density profile) to explain how the projection effects in the lensing observables affect the interpretation of the mass-density parameter λ. In this section, we adopt the more flexible JAM for the stellar kinematics modeling and investigate the projection and selection effect.

In Section 5.1, we present our mock kinematics data using the axisymmetric synthetic lens sample based on assumed intrinsic mass-density profile and stellar luminosity profile. We apply the lensing selection proportional to the lensing cross-section to model the selection effect in the kinematic observables. In Section 5.2 we test the recovery of the velocity dispersion for the mock data sample using spherical Jeans modeling. In Section 5.3, we test the recovery of the velocity dispersion for the mock data using axisymmetric JAM.
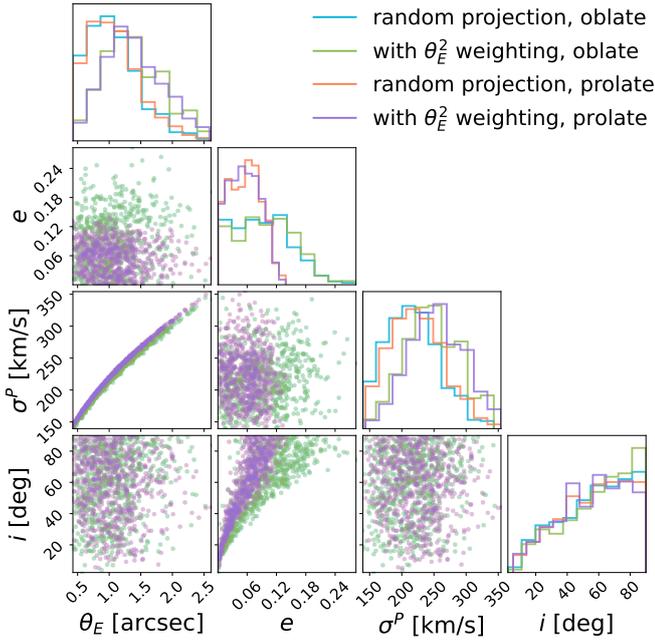
Fig. 5. Corner plot of the kinematics and lensing observables for the mock data, generated by projecting the axisymmetric synthetic lens sample assuming random LoS. Each ETG is projected one time. We applied a lensing selection weighted by the cross-section area, $\propto \theta_E^2$, to illustrate the lensing selection effect on the projected ellipticity and the inclination angle.

Throughout this section, we base our discussion of the projection effect on axisymmetric kinematics models, where we use q to represent the intrinsic axis ratio of the ETGs, and q' to represent the projected axis ratio. For oblates, $\{q, q'\} < 1$; for prolates, $\{q, q'\} > 1$.

5.1. Mock kinematics data and the lensing selection

We use the dynamical modeling software JAMPY (Cappellari 2008, 2020) to generate a set of mock kinematics data with the axisymmetric synthetic lens sample. We start from the intrinsic MGE components of the mass-density and stellar luminosity profiles, project onto random directions, and model the observed velocity dispersion. The projection formalism of axisymmetric systems under MGE parameterization has been summarized in Section 2.5. The generation of the mock kinematics data was performed as follows.

1. We started with the 3D "deformed, truncated SIS" mass-density profile and the 3D Jaffe stellar luminosity profile of the synthetic lens sample and fit the analytic profile with `mge_fit_1d` from the MGEFIT software package[2] (Cappellari 2002) to obtain their intrinsic MGE components.
2. We projected the intrinsic, axisymmetric MGE components according to Eqs. ((39)–(41)) to obtain their projected counterparts.
3. JAMPY was used to de-project the projected MGE components and compute their Jeans solutions. As a result, maps of the projected second velocity moments were generated. We used some simplifying assumptions: no PSF convolution in the projected velocity moments map, no black hole at the galaxy center, isotropic velocity components throughout all
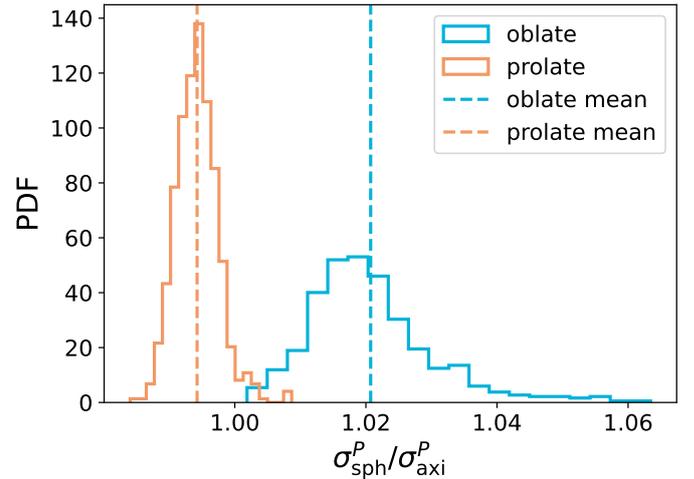

Fig. 6. Distribution of the ratio of the velocity dispersion recovered using spherical JAM and generated using axisymmetric JAM, starting from the same 3D models. For oblates, this ratio is larger than unity, while for prolates most values are smaller than unity, showing that the two populations have different bias in the kinematics recovery when spherical model is assumed. The mean of this ratio for oblates is 1.020, and for prolates is 0.994.

radii ($\beta = 0$), and perfect alignment between the stellar halo and the total mass profile.

Specifically, for the prolate galaxies, we assumed the symmetry axis lies along the projected major axis and rotate the input coordinate by 90 degrees. These galaxies have orthogonal photometric and kinematics major axes (Tsatsi et al. 2017; Li et al. 2018). Figure 5 shows the distribution of the mock data. The circularized Einstein radius, θ_E, is computed from the projected mass-density MGE components. We further applied a lensing cross-section weighting proportional to θ_E^2 to model the lensing selection effect for the projected ellipticity and the inclination angle. Similar to what we conclude from Figure 4, the lensing selection prefers higher projected ellipticity and higher inclination angle for the oblate population, while for the prolate population, the lensing selection prefers lower projected ellipticity, and lower inclination angle.

5.2. Kinematics recovery with spherical JAM

In this section, we focus on the impact from the spherical assumption in the kinematics recovery on the interpretation of the mass-density profile. We start with the mock kinematics data we generated (see Section 5.1) and recovered their 3D kinematics model using spherical JAM. We calculate the spherically averaged MGE for the mass-density profile and the stellar luminosity profile. The total mass (luminosity) expressed with MGE can be written as

$$L_k = \sum_{k=1}^{N} \nu_{0k} (\sigma_k \sqrt{2\pi})^3 q_k, \tag{49}$$

and, therefore, the sphericalized MGE has $\sigma_k^{\mathrm{sph}} = \sqrt{q'} \sigma_k$. Assuming perfect measurement, we compare the velocity dispersion recovered using spherical JAM and the mock data generated using axisymmetric JAM. In Figure 6 we show the distribution of the ratio $\sigma_{\mathrm{sph}}^P / \sigma_{\mathrm{axi}}^P$ and calculate their mean values. We find that the recovered velocity dispersion distribution is biased 2% on

average for the oblate sample, and is biased 0.6% for the prolate sample. To explain the sign of the percentage error, we refer the reader to Appendix A, where we compare the stacked axisymmetric surface density profiles and their spherical counterparts. We use the normalization factor $Z = (pq)^{1/3}$, and as presented in the lower right corner of Figure A.1, the stacked surface density profile of an oblate density profile is slightly lower than that of a spherical profile at small radius, while the prolate case is the opposite. Therefore, the oblate profile has more mass around the center compared to the spherical case, and therefore contribute more to the velocity dispersion within the effective radius. The prolate case is the opposite, having less mass around the center and, hence, contributing less to σ^P.

We specify here that the ratio $\sigma^P_{\rm sph}/\sigma^P_{\rm axi}$ is independent of the velocity dispersion used to characterize the truncated SIS profile, and thus independent of the overall mass of the galaxy. The shape of the tracer profile will only slightly impact the value of the $\sigma^P_{\rm sph}/\sigma^P_{\rm axi}$. We test using Hernquist profile as stellar tracer, with the same scale radius as the Jaffe profiles. The change in the value of the ratio is less than 0.5%. We also test whether the value of $\sigma^P_{\rm sph}/\sigma^P_{\rm axi}$ is impacted by the value of the scale radius. Assuming Jaffe profile as the tracer, when the scale radius is in the range of [3.28, 26.21] kpc (corresponding to half-light radius of [2.5, 20] kpc), the relative change is less than 0.2%. Therefore, we conclude that the choice of the tracer profile and its scale radius are subdominant influencing factors to $\sigma^P_{\rm sph}/\sigma^P_{\rm axi}$. The major parameter impacting $\sigma^P_{\rm sph}/\sigma^P_{\rm axi}$ is the intrinsic axis ratio of the lens population. We then use a fixed velocity dispersion which characterizes our truncated SIS profile (in terms of mass-density), a fixed scale radius which characterizes the Jaffe profile (stellar tracer), and more realizations of the intrinsic axis ratio drawn from the intrinsic shape prior (Li et al. 2018), to generate a larger sample of modeled $\sigma^P_{\rm sph}/\sigma^P_{\rm axi}$.

As an example, we drew 2500 intrinsic axis ratios for both the oblate and prolate population and calculate $\sigma^P_{\rm sph}/\sigma^P_{\rm axi}$, under three choices of the anisotropy parameter $\beta \in \{-0.2, 0, 0.2\}$. For the purpose of making a correction to the JAM modeling, depending on the observed axis ratio of the lens, we inverted $\sigma^P_{\rm sph}/\sigma^P_{\rm axi}$ to obtain a multiplicative correction term, $\sigma^P_{\rm axi}/\sigma^P_{\rm sph}$, and also to condition the distribution of $\sigma^P_{\rm axi}/\sigma^P_{\rm sph}$ on the observed axis ratio. The distribution of $\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1$ (the -1 is for visualization) versus the projected axis ratio q' is shown in Figure 7. While $\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1$ for the prolate population is relatively flat for different q', the oblate curve shows a tight correlation between $\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1$ and q'. Both features make it possible to estimate the systematic uncertainty in the model prediction of the velocity dispersion under spherical assumption. The prolate curve has a smaller dispersion since the population is intrinsically rounder. For the oblates, the residual uncertainty (the 1σ interval) increases at both ends of q' due to the sparsity of data points. For the more elliptical end ($q' \sim 0.5$), the lack of data points is due to the intrinsic shape distribution of the population, since the most elliptical projection has $q' = q$, $i = 90°$. For the rounder end, the lack of data points has two causes: the intrinsic shape distribution and the distribution of isotropic inclination angles. The isotropic inclination angle satisfies $\cos i \in \mathcal{U}[0, 1]$, and therefore galaxies tend to have higher inclination angles if no selection effect is considered, resulting in less round galaxies in projection. To sum up, assuming an axisymmetric galaxy population, one can precisely correct for the spherical Jeans modeling for galaxies given solely the intrinsic axis ratio distribution of the population.
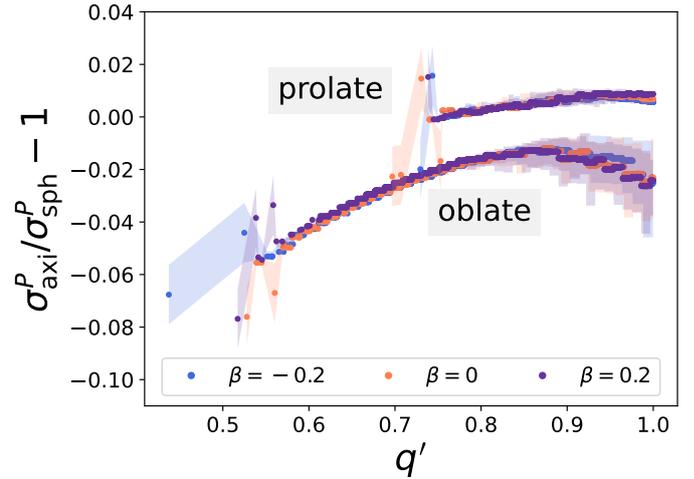

Fig. 7. $\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1$ conditioned on the observed axis ratio q' for the oblate intrinsic shape distribution $q \in \mathcal{N}[\mu = 0.74, \sigma = 0.08]$ and the prolate $1/q \in \mathcal{N}[\mu = 0.84, \sigma = 0.04]$. For this test, the velocity dispersion used to characterize the truncated SIS profile is set to be 200 km/s, and the half-light radius used to characterize the Jaffe profile is set to be 7.5 kpc. The number of draws of the intrinsic axis ratio is 2500, while each data point is projected once assuming isotropic inclination angles. The discreteness of the curve is due to the binning from the 2D histogram of q' vs. $\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1$. The solid dots are the median of the conditional probability $P(\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1 | q')$, while the shaded area is the 1σ interval.

Furthermore, we can combine the correction $\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1$ for the oblate and prolate population assuming that the real lens population is a mix of prolates and oblates. We use the kinematics misalignment angle measured from the spatially resolved spectra of a subsample of the SLACS lenses from Knabel et al. (2025b) to determine the fraction of oblates in our population, and marginalize over the fraction of oblates to combine the two $\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1$ curves for the oblate and prolate samples. We refer the readers to Appendix B for a detailed description of the determination of the fraction of oblates in our population. The combined correction term is shown in Figure 8 in the lower panel. For a lens population with unknown intrinsic shape distribution, the combined $\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1$ versus q' curve can be used to apply correction to the velocity dispersion modeled with spherical JAM. The uncertainty in the correction term $\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1$ is propagated to the velocity dispersion after correction (effectively $\sigma^P_{\rm axi}$) as a residual uncertainty. In our exercise, depending on the choice of the anisotropy parameter and the projected ellipticity, the residual uncertainty is in the range of 0–2.2%, resulting in an uncertainty of 0–4.4% in the inferred mass sheet parameter λ, and therefore H_0. In the upper panel of Figure 8, we plot the histogram of the projected axis ratios of the early-type SLACS lenses from Bolton et al. (2008), to illustrate the range of q' from observed data. Marginalizing over all the q' of the SLACS ETG plotted in Figure 8, and under the choice of $\beta = 0$, we obtain the median correction term $(\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1)_{\rm SLACS} = -0.016^{+0.009}_{-0.007}$. Translating the median correction $(\sigma^P_{\rm axi}/\sigma^P_{\rm sph} - 1)_{\rm SLACS}$ into the median "bias" $(\sigma^P_{\rm sph}/\sigma^P_{\rm axi} - 1)_{\rm SLACS}$, we obtain $(\sigma^P_{\rm sph}/\sigma^P_{\rm axi} - 1)_{\rm SLACS} = 0.017^{+0.007}_{-0.009}$. In other words, in the absence of spatially resolved kinematics, treating the lenses as spherical in the JAM modeling could bias the modeled velocity dispersion up by
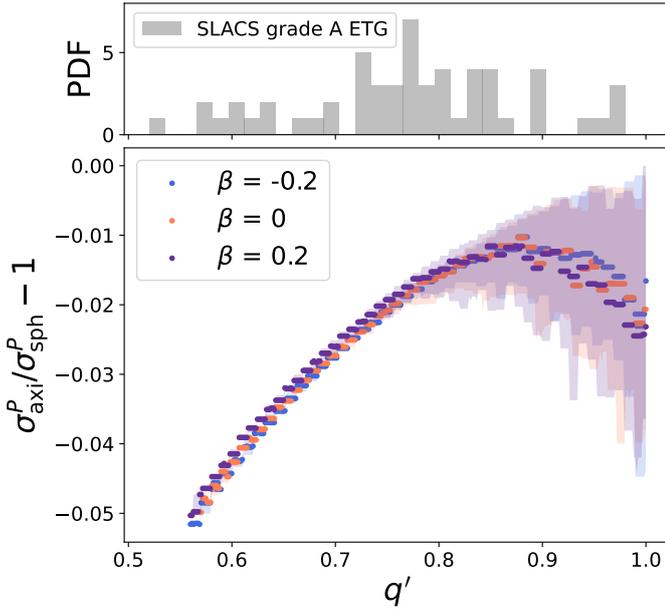
Fig. 8. Kinematics correction term $\sigma_{axi}^P/\sigma_{sph}^P - 1$ versus the projected axis ratio q' as a result of combining the two curves in Figure 7 (lower panel). We use the misalignment angle between the kinematics major axis and the photometric major axis of the SLACS lenses measured in Knabel et al. (2025b), and a simple model for the fraction of oblates in a population (see Appendix B), to determine the weight used to combine the two curves. The solid dots in the lower panel are the median of the conditional probability $P(\sigma_{axi}^P/\sigma_{sph}^P - 1|q')$, while the shaded area is the 1σ interval. The upper panel is to illustrate the distribution of the projected axis ratios of the early-type SLACS lenses from Bolton et al. (2008). Averaging over all these SLACS lenses, and under the choice of $\beta = 0$, we obtain the median correction to be $(\sigma_{axi}^P/\sigma_{sph}^P - 1)_{SLACS} = -0.016^{+0.009}_{-0.007}$.

1.7%, if not corrected for. After correction, the random uncertainty in σ^P is 0.8%.

We conclude that the relative uncertainty introduced by using spherical models rather than axisymmetric models is non-negligible in the population analysis of the kinematics of ETG with JAM. Therefore, axisymmetric dynamical models should be used to accurately recover the mass distribution of ETGs.

5.3. Kinematics recovery using axisymmetric JAM

In this section, we discuss the effect of the intrinsic shape of the lens galaxies in the recovery of their kinematics. The question we want to address is how to best model the velocity dispersion of the lens galaxies under axial symmetry, given their observed ellipticities, the stellar luminosity profiles, the mass-density profiles from lens mass modeling, and an assumed anisotropy profile. We assume that there is no information from spatially resolved spectra, and therefore the inclination angle cannot be directly obtained for an individual lens galaxy. We investigate whether the projected kinematics of a sample of galaxies can be recovered, either with or without the prior knowledge of the intrinsic shape of the population. These two approaches are described in Section 5.3.1.

We describe how we used the mock data from Section 5.1, where we project the axisymmetric synthetic lens sample and recorded the projected ellipticities, the MGE of the projected

stellar luminosity profile, and the MGE of the mass-density profile. We used these observables as input for axisymmetric JAM and compare the recovered velocity dispersion distribution σ^P with the mock data. The results are presented in Section 5.3.2.

5.3.1. Deproject kinematics with and without the intrinsic shape information of the population

We first describe a framework of proposing a distribution of the inclination angle based on the projected ellipticity and the intrinsic axis ratios of the lens galaxies.

The inclination angle is not isotropic for a galaxy with projected ellipticity, e. According to Eq. (39), there is a lower limit of the inclination angle, $i_{min} = \arccos q'$, where q' is the apparent axis ratio of the projected elliptical isophote. At the minimum inclination angle i_{min}, the intrinsic axis ratio, q, deprojected with apparent axis ratio q' is 0 (i.e., the intrinsic shape of the galaxy's mass density-and-stellar luminosity distribution is extremely elongated). To avoid such an unphysical situation, we set a lower limit of the deprojected intrinsic axis ratio to be $q_{min} = 0.05$, noting that this is a very wide prior. The inclination angle distribution for each individual lens galaxy then has a lower bound, i_{min}, which satisfies

$$q_{min}^2 \sin^2 i_{min} + \cos^2 i_{min} = q'^2. \tag{50}$$

The distribution of the inclination angle for the sample satisfies

$$P(i) = \int de P(i|e)P(e), \tag{51}$$

where $P(e)$ is the distribution of the projected ellipticity and $P(i|e)$ is the conditional probability of the inclination angle i given a projected ellipticity, e. Here, $P(e)$ is an observable from the data, while i is not. We modeled $P(i|e)$ with the intrinsic axis ratio distribution of the sample, assuming that the inclination angle is isotropic. We created a 2D histogram of the inclination angle i versus the projected ellipticity, e, with a large number of sample points by drawing from the intrinsic shape distribution. Then, for each q, we used isotropic inclination angles project for 1000 times using Eq. (39) and calculated the projected ellipticity $e = (1 - q')/(1 + q')$. The 2D histogram of the inclination angle i versus the projected ellipticity, e, is provided in Figure 9. In the dimension of the projected ellipticity, we use 30 bins. For the deprojection of each individual galaxy with ellipticity e, we take a slice $P(i|e \in e^i)$, where e^i is the ellipticity bin. An example of $P(i|e = 0.08)$ is included in Figure 9. We then drew the inclination angle from $P(i|e)$ and modeled its projected velocity dispersion with axisymmetric JAM, using the projected MGE for the stellar luminosity profile and the mass-density profile. Using the conditional probability $P(i|e)$, we recovered the inclination angle distribution for the galaxy sample. We then used the recovered inclination angle to recover the velocity dispersion distribution.

Another approach to recover the inclination angles is to assume galaxies are randomly oriented and the intrinsic shape is only constrained by their projected ellipticities. In other words, in such a case, we would not include any intrinsic shape prior of the population in the inclination angle recovery. Then, for each ETG, we draw uniformly from the random inclination angle distribution, and reject the inclination angles which are physically forbidden by the projected ellipticity, according to Eq. (50).
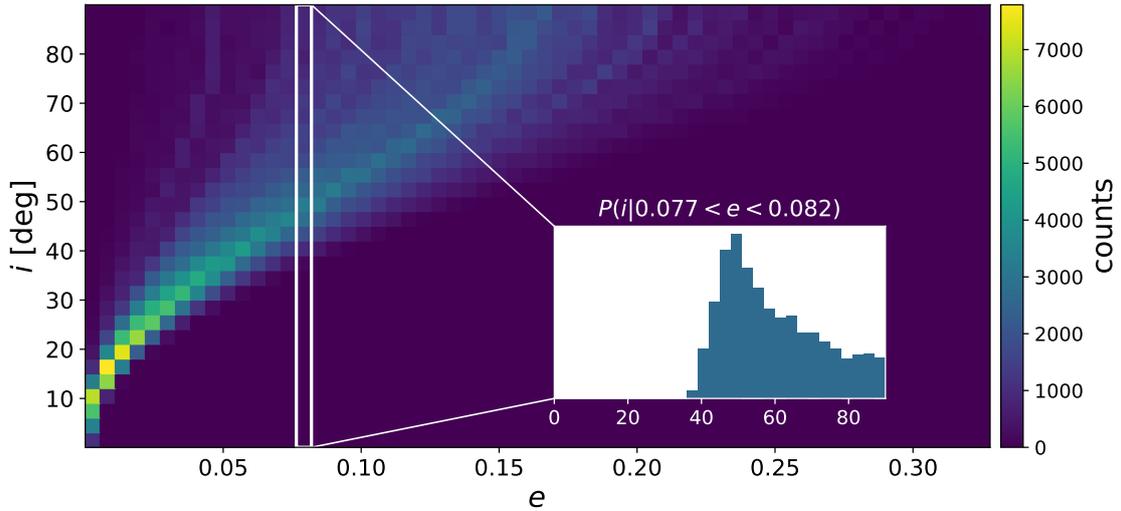
Fig. 9. 2D histogram of the inclination angle i versus the projected ellipticity, e, modeled for the oblate synthetic sample assuming isotropic inclination angles. In the lower right corner is a slice of $P(i|e)$ at $e = 0.08$ as an illustration of the shape of the inclination angle distribution.

5.3.2. Results of axisymmetric kinematics recovery with and without intrinsic shape prior

In Section 5.3.1 we described two methods to recover the inclination angle for a population of ETGs in the absence of spatially resolved spectra. We apply both approaches to the mock data to obtain a comparison between the recovered inclination angle distribution and the velocity dispersion distribution. The results are shown in Figure 10, in which we calculate the relative error between the mean of the recovered velocity dispersion and the mock data. When the intrinsic shape prior is included (denoted in Figure 10 by "recovered from $P(i|e)$"), the distribution of the recovered inclination angle is very similar to the inclination angle distribution from the mock data, which is sampled isotropically on a sphere. The inclination angle recovered without the intrinsic shape prior (denoted in Figure 10 as "recovered from isotropic") are overall larger than the mock data for both the oblate and prolate sample, due to the fact that the inclination angle at any projected ellipticity has a distribution that is the shape of the isotropic one but with a lower cutoff. Meanwhile, the actual conditional probability $P(i|e)$, as illustrated in Figure 9, is peaked at a smaller value. Having understood the bias in the recovered inclination angle, we find that the recovered velocity dispersion is not impacted. The relative error between the mean of the recovered velocity dispersion and the mock data is less than 0.24% with or without the inclination angle prior.

We conclude that for a ETG population with an underlying intrinsic shape distribution of $q \in \mathcal{N}[\mu = 0.74, \sigma = 0.08]$ for oblates and $1/q \in \mathcal{N}[\mu = 0.84, \sigma = 0.04]$ for prolates, the kinematic recovery using axisymmetric JAM is accurate with or without knowledge of the intrinsic shape distribution. However, we note that this conclusion is impacted by the choice of the intrinsic shape distribution, and therefore our conclusion is applicable for an ETG population which is overall less elliptical than our sample.

6. Triaxiality in kinematics

In Section 5, we discuss the projection and selection effect in the kinematics assuming axisymmetric kinematics models; namely, galaxies, if they are nonspherical, are either oblate or prolate.

Real galaxies, however, can be triaxial in both the mass distribution and the stellar tracer distribution. Therefore, when we perform axisymmetric kinematic recovery using the observed stellar tracer profile, we ought to bear in mind that the ellipticity, the amplitude, and the shape of the tracer profile is actually the projection of a triaxial ellipsoid. In this section, we assess the uncertainty in the kinematics recovery from assuming axisymmetric mass and tracer distributions for systems with axisymmetric mass distribution and triaxial tracer distribution, noting that the kinematic model of a system with both triaxial mass and tracer distribution is beyond the scope of this work and needs further mathematical derivation or simulation. We then compare the velocity dispersion recovered using the ellipticity projected by triaxial tracer profiles to the velocity dispersion modeled using axisymmetric mass and tracer distributions. Practically, we project the triaxial and the axisymmetric galaxy light profiles of the same total luminosity under the same inclination angle. For the axisymmetric model, we record the projected velocity dispersion. For the triaxial model, we record the projected ellipticity and the MGE components, then use them to recover the projected velocity dispersion assuming axisymmetric models.

6.1. Axisymmetric kinematics models for the TNG sample

In Section 3.1.2, we introduce a triaxial ETG catalog from the TNG100 simulation. In the present section, we use them to generate approximate axisymmetric JAM models and an ellipticity distribution, e, which matches the random projection of the triaxial sample. We start by describing how we approximate the triaxial ETG with axisymmetric ones. For the oblate sample, which has $T \leq 0.5$, we take $p \rightarrow 1$, and $q \rightarrow q(1 + p)/2$. For the prolate sample with $T > 0.5$, we take $p = q \rightarrow (p + q)/2$. Then, we project them randomly, each 10 times to increase data size, and record the projected ellipticity, the projected velocity dispersion, the MGE parameters, and the viewing angles (θ, ϕ).

6.2. Axisymmetric kinematics recovery for a triaxial population

Using the viewing angles (θ, ϕ), we return to the triaxial models to calculate the projected ellipticity e^{tri} using Eq. (26). We
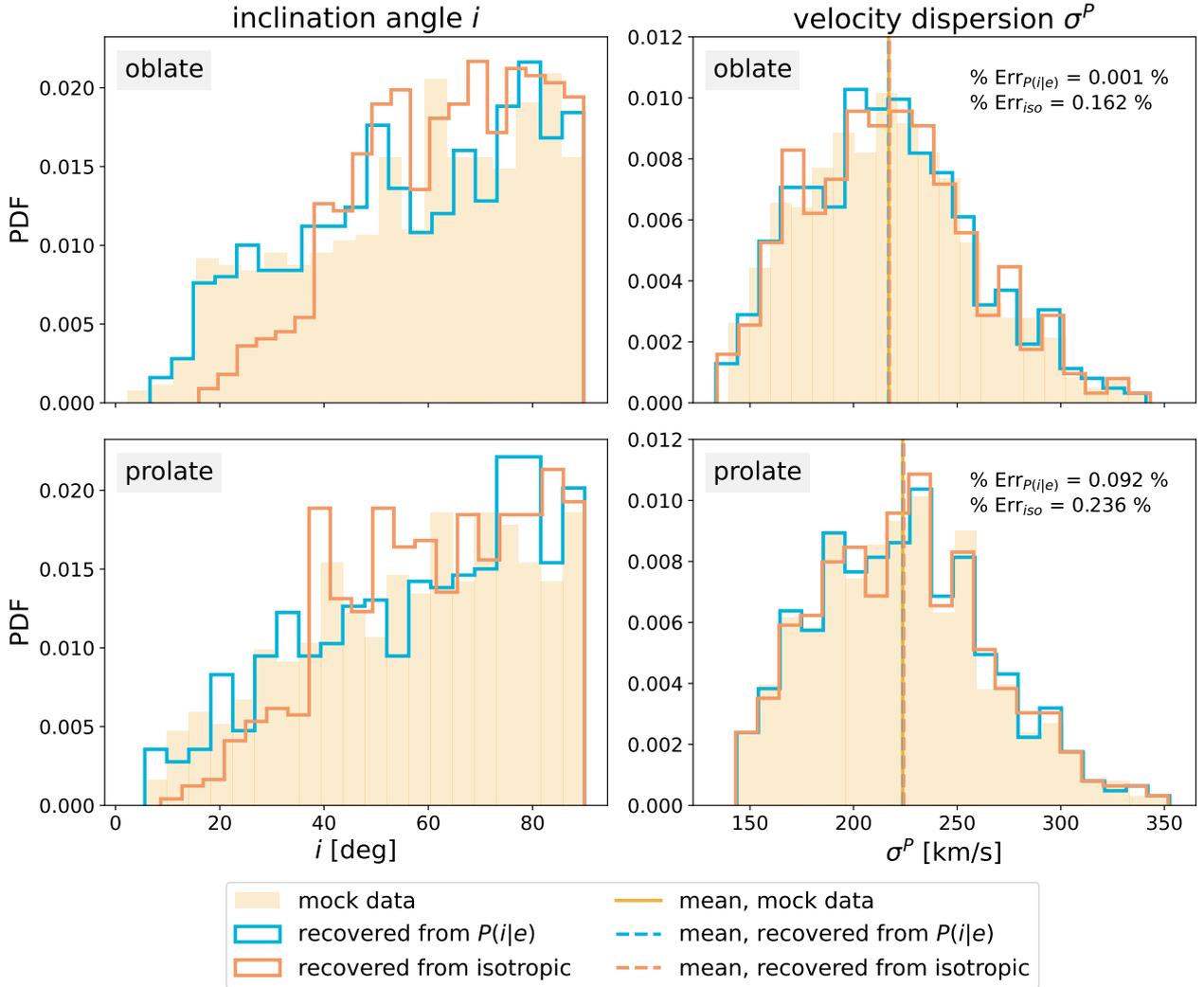
Fig. 10. Distributions of the recovered inclination angles and the velocity dispersions using axisymmetric JAM, with (denoted by "recovered from $P(i|e)$") and without intrinsic shape prior (denoted by "recovered from isotropic"). The upper and lower rows show the results for the oblate and the prolate sample, respectively. The inclination angle recovered with the intrinsic shape prior has a similar distribution as the inclination angle from the mock data which are sampled isotropically on a sphere, while the inclination angle recovered without the intrinsic shape prior are overall larger than the mock data. However, the deviation in the recovered inclination angle does not affect the recovered velocity distribution. For both the oblate and the prolate sample, the relative error in the mean of the recovered velocity dispersion is less than 0.24%. This result shows that for a ETG population with the intrinsic shape distribution similar to the one in Li et al. (2018), the axisymmetric JAM kinematics recovery can be accurate even without the intrinsic shape prior.

also modify the MGEs to preserve the total mass and the total luminosity. Following Eq. (49), the MGE for the triaxial ellipticities has $\sigma_k^{\rm tri} = \sqrt{q'/q'_{\rm tri}}\sigma_k$, where $q'_{\rm tri}$ is the projected axis ratio corresponding to $e^{\rm tri}$. We then use the modified MGE and the triaxial ellipticities as input for JAM to recover the velocity dispersion σ^P, and compare with the observables of the axisymmetric models. We use $P(i|e)$ to recover the velocity dispersion for the sample, based on the conclusion in Section 5.3.1. The results are shown in Figure 11. The percentage difference between the mean of the recovered velocity dispersion and the mock data is smaller than 0.85%, from which we validate that the kinematics recovery is unbiased, when the projected ellipticity is contributed by the intrinsic triaxiality of the sample.

We further derive a percentage difference in the mean of the velocity dispersions weighted by the ellipticity of the SLACS early-type grade "A" lenses, as the TNG sample we use is more elliptical and thus the relative difference is not typical for observed lenses. We use the ratio between the density function of the ellipticity of the SLACS lenses and the triaxial ellipticities $w(e) = P^{\rm SLACS}(e)/P^{\rm tri}(e)$ as a weight term, to calculate the mean of the velocity dispersions of our axisymmetric galaxy models and the mean of the velocity dispersions recovered using triaxial ellipticities. The percentage difference for the velocity dispersion is 0.053% for the oblate TNG sample and 0.17% for the prolate TNG sample. We note that the underlying assumption here is that the SLACS lenses have the same triaxial intrinsic shape distribution as the TNG ETGs, which we cannot yet verify. However, our TNG ETG sample is more triaxial than nearby population of observed ETGs (Li et al. 2018). Therefore, we can conclude that axisymmetric kinematic models can aptly recover the velocity dispersion for a galaxy with axisymmetric mass distribution and triaxial stellar light distribution.
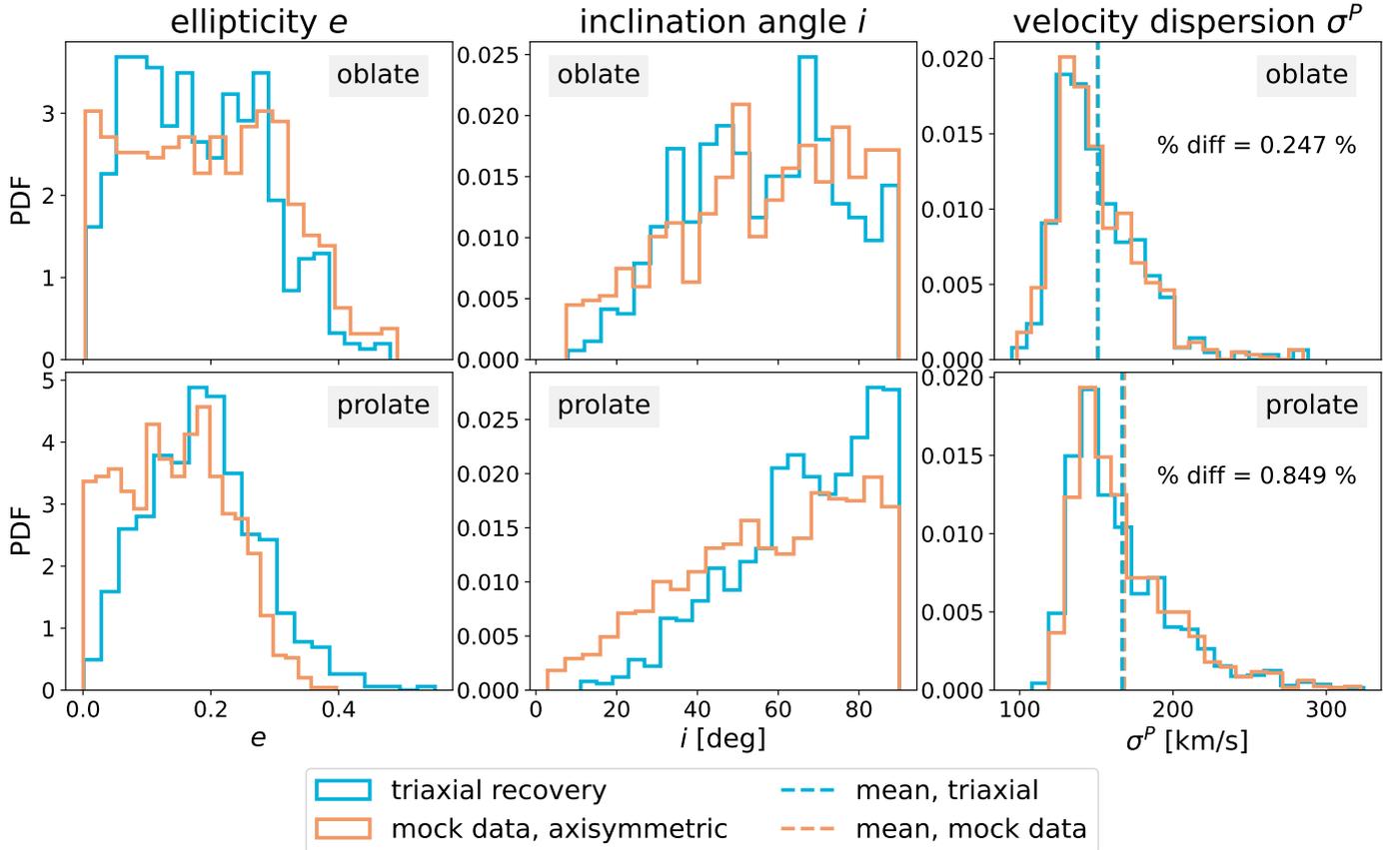
Fig. 11. Ellipticity, recovered inclination angle and recovered velocity dispersion for triaxial tracer + axisymmetric mass models and comparison with the axisymmetric tracer + mass models for the triaxial TNG-100 ETG sample. The inclination angle distribution is recovered using the conditional probability $P(i|e)$, and the velocity dispersion is computed with axisymmetric JAM. We conclude that under the assumption of axisymmetric kinematics models, when the ellipticity is contributed by a triaxial tracer component, the kinematics recovery for the sample is unbiased.

7. Conclusions

The intrinsic shapes of strong lensing galaxies lead to projection effects in both lensing and stellar kinematics observables. The latter is used to break the mass-sheet degeneracy (MSD), which is crucial for accurate measurements of the Hubble constant with time-delay cosmography. Since the projection effect affects the lensing convergence, it can lead to selection effects; for example, when the probability of observing a lens is approximately proportional to the lensing cross-section, $\propto \theta_E^2$. In this work, we quantitatively investigated the projection and selection effects introduced by the intrinsic shape of lens galaxies, using an axisymmetric synthetic lens sample similar to the SLACS lenses and a selected triaxial ETG sample from the TNG100 simulation.

In this work, we present numerical simulations of the projection effect for the lensing observables for individual triaxial galaxies. Based on the projection effect and by comparing to spherical lens models, the lensing observables of nonspherical lenses end up scattered around that of the spherical lens of the same mass for 7.8%. We were able to quantitatively analyze the selection effect introduced by the projection effect of a triaxial galaxy sample by forward-modeling the TNG100 sample and applying selection function in the form of a lensing cross-section weighting. We investigated the selection effects in the projected ellipticity of the samples as a function of intrinsic shape distributions, finding that the more oblate galaxies prefer higher projected ellipticity, and the more prolate galaxies prefer lower

projected ellipticity. We demonstrated that the lensing mass estimates can be biased when constrained with nonspherical parameterized mass-density profiles, if the projection effects are not accounted for.

We discuss the projection and selection effect in the stellar kinematics of lens galaxies, based on our use of the spherically aligned axisymmetric JAM to construct stellar kinematic models for our axisymmetric synthetic lens sample as a set of mock data. Using the set of mock data, we model and illustrate the projection effect in the luminosity-weighted velocity dispersion on both individual galaxies and the galaxy sample.

We quantified the bias in the deprojection of the stellar kinematics under different assumptions on the intrinsic shape with the mock data. One of our main conclusions is that assuming spherical JAM solution biases velocity dispersion modeled with the best-fit lens mass model of imaging data. For our oblate subsample, the mean bias is 2%. For our prolate sample, the mean bias is -0.6%. This bias is then doubled in the inference of the MSD parameter λ and H_0. Moreover, we find that this bias is a function of the projected axis ratio, which enables us to construct a correction term of $\sigma_{\mathrm{axi}}^P / \sigma_{\mathrm{sph}}^P - 1$ for the velocity dispersion modeled spherical JAM. The residual uncertainty in the correction term is in the range of 0–2.2%. This residual uncertainty contributes to the overall uncertainty on the mass sheet parameter, λ, and, therefore, H_0. As a more specific example, we calculated the median of the correction term using the projected axis ratio of the early-type SLACS lenses from Bolton et al. (2008) to be

$(\sigma_{\mathrm{axi}}^P/\sigma_{\mathrm{sph}}^P-1)_{\mathrm{SLACS}} = -0.016^{+0.009}_{-0.007}$, resulting in a residual uncertainty in the corrected velocity dispersion of 0.8%.

Conversely, using axisymmetric models to deproject the kinematics, we can accurately recover the 3D kinematic models. We presented an upper limit on the overall ellipticity of the ETG sample at which the kinematics recovery using axisymmetric JAM will not be impacted by the inclusion of the intrinsic shape prior.

We also find that assuming axisymmetric JAM models for systems with triaxial stellar light profiles and axisymmetric mass distributions can accurately retrieve the velocity dispersion distribution, using the TNG100 ETG sample. Since the TNG sample is generally more elliptical than lensing galaxies, the conclusion applies to observed lenses, which have an underlying rounder intrinsic shape distribution.

In summary, we developed the machinery for forward-modeling the projection and selection effect for the lensing and kinematics observables of nonspherical galaxies that act as strong lenses. We showed that the projection effects have non-negligible impact on the lensing and kinematic observables, finding that an accurate analysis of lensing and kinematic data requires accurate deprojections. Furthermore, when looking at a population of deflectors, treating the deflectors with spherical assumption does not average out biases on the population level. Hence, to accurately recover H_0 when using a lensing + kinematics analysis to break the MSD requires an explicit treatment of the projection effects on lensing and kinematics data in the analysis.

The framework and quantitative characterization of the lensing + kinematic selection and the axisymmetric versus spherical kinematic treatment presented in this work have been incorporated in the TDCOSMO2025 cosmological results paper (Birrer et al. 2025) and will also be included in forthcoming time-delay cosmography analysis (e.g., Paic et al. 2025).

References

Barnabè, M., Czoske, O., Koopmans, L. V. E., Treu, T., & Bolton, A. S. 2011, MNRAS, 415, 2215
Benacchio, L., & Galletta, G. 1980, MNRAS, 193, 885
Binney, J. 1985, MNRAS, 212, 767
Binney, J., & de Vaucouleurs, G. 1981, MNRAS, 194, 679
Binney, J., & Tremaine, S. 1987, Galactic Dynamics
Birrer, S., & Amara, A. 2018, Phys. Dark Universe, 22, 189
Birrer, S., Shajib, A. J., Galan, A., et al. 2020, A&A, 643, A165
Birrer, S., Shajib, A., Gilman, D., et al. 2021, J. Open Source Softw., 6, 3283
Birrer, S., Millon, M., Sluse, D., et al. 2024, Space Sci. Rev., 220, 48
Birrer, S., Buckley-Geer, E. J., Cappellari, M., et al. 2025, A&A, 704, A63
Bolton, A. S., Burles, S., Koopmans, L. V. E., Treu, T., & Moustakas, L. A. 2006, ApJ, 638, 703
Bolton, A. S., Burles, S., Koopmans, L. V. E., et al. 2008, ApJ, 682, 964
Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7
Cappellari, M. 2002, MNRAS, 333, 400
Cappellari, M. 2008, MNRAS, 390, 71
Cappellari, M. 2016, ARA&A, 54, 597
Cappellari, M. 2020, MNRAS, 494, 4819
Cappellari, M., Emsellem, E., Bacon, R., et al. 2007, MNRAS, 379, 418
Croom, S. M., Lawrence, J. S., Bland-Hawthorn, J., et al. 2012, MNRAS, 421, 872
de Zeeuw, P. T., Evans, N. W., & Schwarzschild, M. 1996, MNRAS, 280, 903
Dutton, A. A., Brewer, B. J., Marshall, P. J., et al. 2011, MNRAS, 417, 1621
Emsellem, E., Monnet, G., & Bacon, R. 1994, A&A, 285, 723
Emsellem, E., Cappellari, M., Krajnović, D., et al. 2007, MNRAS, 379, 401
Ene, I., Ma, C.-P., Veale, M., et al. 2018, MNRAS, 479, 2810
Falco, E. E., Gorenstein, M. V., & Shapiro, I. I. 1985, ApJ, 289, L1
Fassnacht, C. D., Gal, R. R., Lubin, L. M., et al. 2006, ApJ, 642, 30
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Jaffe, W. 1983, MNRAS, 202, 995
Jeans, J. H. 1922, MNRAS, 82, 132
Kimm, T., & Yi, S. K. 2007, ApJ, 670, 1048
Knabel, S., Mozumdar, P., Shajib, A. J., et al. 2025a, A&A, 703, A117
Knabel, S., Treu, T., Cappellari, M., et al. 2025b, ApJ, 990, 51
Kochanek, C. S., Morgan, N. D., Falco, E. E., et al. 2006, ApJ, 640, 47
Koopmans, L. V. E., Treu, T., Fassnacht, C. D., Blandford, R. D., & Surpi, G. 2003, ApJ, 599, 70
Koopmans, L. V. E., Bolton, A., Treu, T., et al. 2009, ApJ, 703, L51
Kundić, T., Turner, E. L., Colley, W. N., et al. 1997, ApJ, 482, 75
Lambas, D. G., Maddox, S. J., & Loveday, J. 1992, MNRAS, 258, 404
Li, H., Mao, S., Cappellari, M., et al. 2018, ApJ, 863, L19
Mamon, G. A., & Łokas, E. L. 2005, MNRAS, 363, 705
Oguri, M. 2007, ApJ, 660, 1
Padilla, N. D., & Strauss, M. A. 2008, MNRAS, 388, 1321
Paic, E., Courbin, F., Fassnacht, C. D., et al. 2025, A&A, in press https://doi.org/10.1051/0004-6361/202556411
Pillepich, A., Springel, V., Nelson, D., et al. 2018, MNRAS, 473, 4077
Pulsoni, C., Gerhard, O., Arnaboldi, M., et al. 2020, A&A, 641, A60
Refsdal, S. 1964, MNRAS, 128, 307
Rodríguez, S., & Padilla, N. D. 2013, MNRAS, 434, 2153
Ryden, B. 1992, ApJ, 396, 445
Saha, P., Coles, J., Macciò, A. V., & Williams, L. L. R. 2006, ApJ, 650, L17
Sandage, A., Freeman, K. C., & Stokes, N. R. 1970, ApJ, 160, 831
Schechter, P. L., Bailyn, C. D., Barr, R., et al. 1997, ApJ, 475, L85
Schneider, P., & Sluse, D. 2013, A&A, 559, A37
Schwarzschild, M. 1979, ApJ, 232, 236
Shajib, A. J., Treu, T., Birrer, S., & Sonnenfeld, A. 2021, MNRAS, 503, 2380
Shu, Y., Bolton, A. S., Brownstein, J. R., et al. 2015, ApJ, 803, 71
Sonnenfeld, A. 2024, A&A, 690, A325
Stark, A. A. 1977, ApJ, 213, 368
Suyu, S. H., Marshall, P. J., Auger, M. W., et al. 2010, ApJ, 711, 201
Suyu, S. H., Auger, M. W., Hilbert, S., et al. 2013, ApJ, 766, 70
Syer, D., & Tremaine, S. 1996, MNRAS, 282, 223
Treu, T., & Koopmans, L. V. E. 2002, ApJ, 575, 87
Treu, T., & Marshall, P. J. 2016, A&ARv, 24, 11
Tsatsi, A., Lyubenova, M., van de Ven, G., et al. 2017, A&A, 606, A62
Vincent, R. A., & Ryden, B. S. 2005, ApJ, 623, 137
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nat. Methods, 17, 261
Weijmans, A.-M., de Zeeuw, P. T., Emsellem, E., et al. 2014, MNRAS, 444, 3340
Wong, K. C., Suyu, S. H., Chen, G. C. F., et al. 2020, MNRAS, 498, 1420

Appendix A: Normalization factor, Z

In this appendix, we discuss the choice of the normalization factor Z in Eq. (19). Z is introduced to preserve the total mass of the mass-density profile $\rho(a_v)$, when varying the intrinsic shapes of these profiles. There are two choices of Z, (a) $Z = 1$ and (b) $Z = (pq)^{1/3}$, which maps to different effective spherical profiles, i.e., different monopole moments. The total mass within the effective radius, r_s, is

$$L_k = \int_{a_v = r_s} dV \rho(a_v), \tag{A.1}$$

where dV is the volume element in a deformed spherical coordinate with a_v as the effective radius. For $Z = 1$, L_k is not conserved with varying the intrinsic axis ratios p and q, and thus need to be normalized. For $Z = (pq)^{1/3}$, L_k is conserved at any effective radius r_s.

Without the loss of generality, we numerically test the effect of these two choices of Z on the surface density profile under axial symmetry and MGE form. We express the intrinsic mass-density profile with MGEs as in Eq. (37). The total mass is then

$$L_k = \sum_{k=1}^{N} \nu_{0k} (\sigma_k \sqrt{2\pi})^3 q_k. \tag{A.2}$$

We require the total luminosity to be conserved when the intrinsic axis ratio q_k changes to q'_k, and thus we can choose (a) $\nu'_{0k} = \nu_{0k} q_k / q'_k$, or (b) $\sigma'_k = \sigma_k (q_k / q'_k)^{1/3}$. The former corresponds to a choice of $Z = 1$, and the latter corresponds to the choice of $Z = (pq)^{1/3}$.

We test whether the surface density profile of an axisymmetric density profile is biased, comparing to a spherical profile with $q_k = 1$. We assume the orientation of the axisymmetric profile is isotropic on a sphere, namely, with the viewing angle of $\cos\theta \in \mathcal{U}[0, 1]$ and the inclination angle $i = \theta$. We project the intrinsic density profile 1000 times, and stack the surface density profiles. We compare the stacked surface density profile with the surface density profile of a intrinsically spherical density profile. Here we use an example NFW halo with a scale radius of 50 kpc and an overall normalization of 1 $M_\odot/\text{kpc}^3$. We use an intrinsic axis ratio $q_k = 0.5$ for the oblate case and an intrinsic axis ratio $q_k = 1.6$ for the prolate case. The results are shown in Figure A.1. The left column is a 3D NFW density profile with different intrinsic axis ratios. The right column is the stacked surface density profiles. Under the choice of $Z = (pq)^{1/3}$, the stacked surface density profile of an axisymmetric profile is closer to the spherical case. We conclude that choosing $Z = (pq)^{1/3}$ yields more accurate surface density profile, if we assume the stacked surface density profile of a nonspherical profile should approach the spherical surface density of the same mass.

Appendix B: The fraction of oblate ETG under a simple model

Pure oblate galaxies have their photometric major axis and their kinematics major axis fully aligned. Pure prolate galaxies, being very rare in nature, have their photometric major axis and kinematics major axis perpendicular to each other. Galaxies not falling into these two categories are thought to be triaxial. Here we consider a population composed of oblate and prolate galaxies. The distribution of offsets between the photometric and kinematic axis is

$$P(\Delta\phi|f) = f\delta(\Delta\phi) + (1 - f)\delta(\Delta\phi - \pi/2), \tag{B.1}$$

Table B.1. Misalignment angle between the kinematics major axis and the photometric major axis of the lens galaxies from K24.

Object	$\Delta\phi$ (°)	$\sigma_{\Delta\phi}$ (°)
SDSSJ0029-0055	31.0	30.0
SDSSJ0037-0942	28.0	4.0
SDSSJ0330-0020	34.0	10.0
SDSSJ1112+0826	3.0	3.0
SDSSJ1204+0358	14.0	11.0
SDSSJ1250+0523	71.0	7.0
SDSSJ1306+0600	35.0	13.0
SDSSJ1402+6321	19.0	8.0
SDSSJ1531-0105	9.0	30.0
SDSSJ1538+5817	19.0	30.0
SDSSJ1621+3931	9.0	4.0
SDSSJ1627-0053	14.0	13.0
SDSSJ1630+4520	5.0	9.0
SDSSJ2303+1422	12.0	30.0

where f is the fraction of oblate galaxies, and $\Delta\phi$ is the misalignment angle. Taking into account the measurement uncertainty on $\Delta\phi$, the likelihood function is

$$L(\Delta\phi|f) = \prod_{i}^{N} \frac{1}{\sqrt{2\pi\sigma_{\Delta\phi_i}^2}} \times$$
$$\left[f\exp(-\frac{\Delta\phi_i^2}{2\sigma_{\Delta\phi_i}^2}) + (1 - f)\exp(-\frac{(\Delta\phi_i - \pi/2)^2}{2\sigma_{\Delta\phi_i}^2}) \right], \tag{B.2}$$

where N is the total number of lenses with measured misalignment angles.

Feeding in the misalignment angle measured in Knabel et al. (2025b) using the Keck Cosmic Web Imager (KCWI) IFU data of a subsample of the SLACS lenses, as listed in Table B.1, we model the likelihood as a function of f for individual lenses (dashed lines) and the entire dataset using Eq. (B.2), as shown in Figure B.1. Marginalizing over the combined distribution of f as a weight, we then add the two curves in Figure 7 to obtain Figure 8, from which one obtains a prediction of the correction term in the kinematics modeling.

We note that the choice of f does not impact significantly the distribution of the correction term $\sigma_{\text{axi}}^P/\sigma_{\text{sph}}^P$ under a certain projected axis ratio q'. To test this, we model the combined correction term under a randomly chosen projected axis ratio using different single values of f instead of integrating over f. Here we choose $\beta = 0$, and $q' = 0.9$. For f, we choose $f = \{0.79, 0.88, 0.95\}$, corresponding to the 16-th, 50-th and 84-th percentile of our f shown in Figure B.1. The lower panel of Figure B.2 shows the distribution of the correction term at the chosen q', while the 16-th, 50-th and 84-th percentiles of the distributions are plotted in the upper panel. We calculate the difference in the median of $\sigma_{\text{axi}}^P/\sigma_{\text{sph}}^P$ at $f = 0.79$ and $f = 0.95$. We find $\sigma_{\text{axi}}^P/\sigma_{\text{sph}}^P(f = 0.79) - \sigma_{\text{axi}}^P/\sigma_{\text{sph}}^P(f = 0.95) = 0.105\%$. We conclude that when the oblate fraction f changes by one standard deviation, the change in the inferred mass-sheet parameter λ and the Hubble constant H_0 is 0.105%, which is negligible under the current expected measurement uncertainty of H_0.
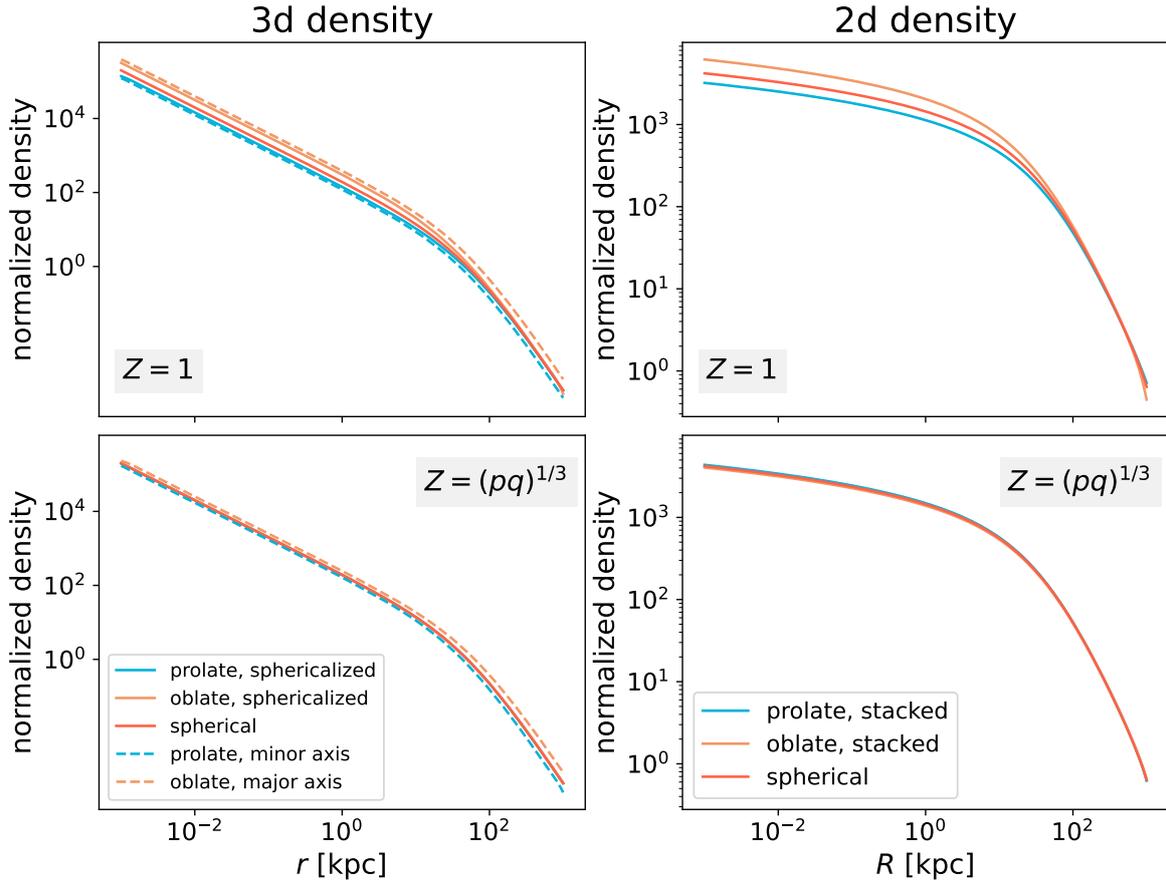
Fig. A.1. Comparison between the 3D and 2D densities using $Z = 1$ and $Z = (pq)^{1/3}$, under axial symmetry and MGE formalism. Left: 3D density of a NFW profile with different intrinsic axis ratios. The oblate profile has an intrinsic axis ratio of 0.5, and the prolate profile has an intrinsic axis ratio of 1.6. The solid lines are the sphericalized (averaged on a sphere) density profiles. The dashed lines are the 3D density profiles along the symmetry axis of axisymmetric profiles. Right: Stacked surface density obtained by projecting the 3D densities on the left column 1000 times.
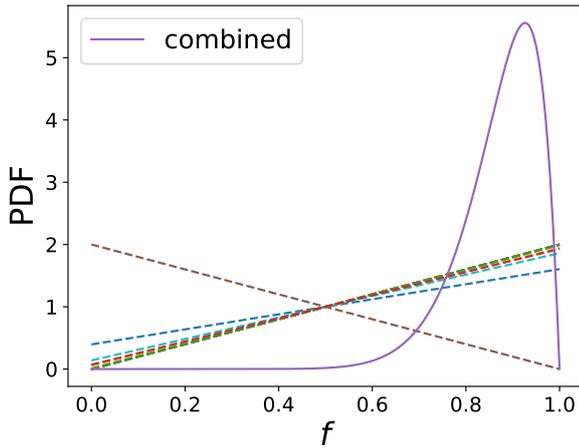

Fig. B.1. Likelihood function of measured misalignment angles from K24 as a function of the oblate fraction f assuming the model distribution in Eq. (B.1).
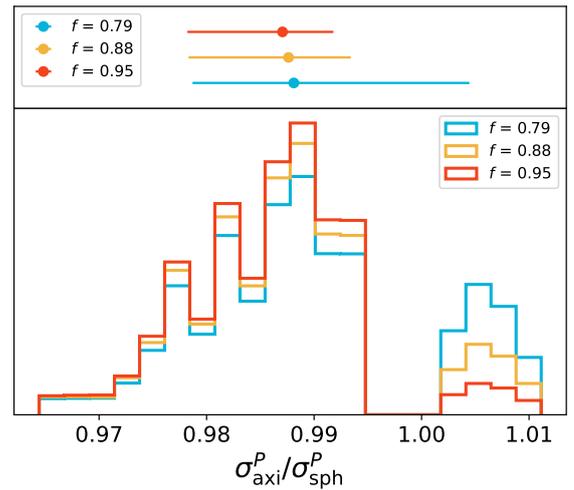

Fig. B.2. Lower panel: Normalized distribution of the correction term $\sigma^P_{\mathrm{axi}}/\sigma^P_{\mathrm{sph}}$ at $q' = 0.9$ under different values of the oblate fraction f. Upper panel: 16th (left end of the errorbar), 50th (the central marker), and 84th (right end of the errorbar) percentile corresponding to the distribution of the correction term. The difference in the median $\sigma^P_{\mathrm{axi}}/\sigma^P_{\mathrm{sph}}$ at $f = 0.79$ and $f = 0.95$ is 0.105%, indicating that the relative change in the corrected velocity dispersion when f changes by two standard deviations is 0.105%. Therefore, the change in the inferred λ and H_0 when f changes by one standard deviation is 0.105%.