

Robust test of general relativity at the galactic scales by combining strong lensing systems and gravitational wave standard sirens

Tonghua Liu,¹ Marek Biesiada,² Shuxun Tian³,⁴ and Kai Liao^{4,*}

¹*School of Physics and Optoelectronic Engineering, Yangtze University, Jingzhou, 434023, China*

²*National Centre for Nuclear Research, Pasteura 7, PL-02-093 Warsaw, Poland*

³*Department of Astronomy, Beijing Normal University, 100875, Beijing, China*

⁴*School of Physics and Technology, Wuhan University, Wuhan 430072, China*



(Received 23 January 2024; accepted 10 April 2024; published 29 April 2024)

The measurement of the parametrized post-Newtonian parameter γ_{PPN} is a robust test of general relativity (GR). In some modified theories of gravity, γ_{PPN} may evolve with the redshift and deviate from one at high redshifts. This means that precise constraints on γ_{PPN} acquired in the solar system experiments could not be sufficient to test such theories and it is necessary to constrain γ_{PPN} with high precision at high redshifts. However, in many approaches aimed at extragalactic tests of GR, the results might be biased due to entanglement of various factors, such as cosmic curvature, cosmic opacity, and the Hubble constant. Strong lensing systems naturally provide a laboratory to test γ_{PPN} at galactic scales and high redshifts, but there is degeneracy between measured strength of gravity and cosmic distances in the lensing system. Gravitational waves (GWs) from binary neutron star mergers (standard sirens) provide a direct way to break this degeneracy by providing self-calibrated measurements of the luminosity distance. We investigate the possibility of estimating γ_{PPN} by combining well-measured strongly lensed systems with GW signals from coalescing neutron stars. Such combination provides a cosmological-model independent, relatively pure and unbiased method for the inference of γ_{PPN} parameter, avoiding the influence of the above factors and the mass-sheet degeneracy in the lens. Based on the simulated future 55 lensed quasar systems we demonstrated that the precision of γ_{PPN} parameter obtained by our method could be of order of $\sim 10^{-2}$. One may reasonably expect that our approach will play an increasingly important role in precise testing the validity of general relativity at galactic scales and high redshifts.

DOI: [10.1103/PhysRevD.109.084074](https://doi.org/10.1103/PhysRevD.109.084074)

I. INTRODUCTION

Over the past few decades, observations of type Ia supernovae (SNe Ia) have revealed that the expansion of the Universe is accelerating [1,2]. Based on Einstein's theory of general relativity (GR) [3], and assuming that the Universe is homogeneous and isotropic on large scales [4], it is generally accepted that the so called Λ CDM model correctly describes our Universe. Although this model is currently taken as the concordance cosmological scenario and supported by the vast majority of astronomical observations [5], it still faces some problems. First, there is a significant tension in the Hubble constant H_0 between the cosmic microwave background (CMB) measurement obtained within GR + Λ CDM using the *Planck* data, which yielded $H_0 = 67.4 \pm 0.5 \text{ km s}^{-1} \text{ Mpc}^{-1}$ at the 68% confidence level (CL) [6] and the value of $H_0 = 73.2 \pm 1.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$ at the 68% CL reported by SHOES (*Supernova H_0 for the Equation of State*) collaboration using SNe Ia calibrated by local Cepheid variable stars [7].

The H_0 Lenses in COSMOGRAIL's Wellspring (H0LiCOW) collaboration recently reported the measurement $H_0 = 73.3_{-1.8}^{+1.7} \text{ km s}^{-1} \text{ Mpc}^{-1}$ with a joint analysis of six lensed quasars for a flat Λ CDM model [8] using a technique independent of above mentioned methods. Further, the same happens with measurements of cosmic curvature [9,10] and the growth of structure S_8 parameter, which also reveal tension (see a recent review [11] for more details) between alternative approaches and analyses. These inconsistencies are the fairly challenging problem in astrophysics and cosmology. From the theoretical point of view, cosmological constant Λ responsible for the accelerating expansion of Universe should be invoked in the framework of GR, but inconsistency between its observed value and theoretical predictions from the quantum field theory is considerable [12].

Considering the fact that, the values of H_0 inferred from *Planck* data and H0LiCOW collaboration are based on GR plus Λ CDM model, it opens a discussion of whether the GR could fail at larger, cosmological scales. Although GR has passed with a very high precision all tests at the millimeter scale in the laboratory, up to the solar system

*liaokai@whu.edu.cn

scales [13,14], and the detection of gravitational waves also provided the possibility for testing the validity of GR on very extreme scales [15–19], the long-range nature of gravity on the extragalactic or cosmological scale is still relatively insufficiently tested. At present, precise constraints regarding the parametrized post-Newtonian parameter γ_{PPN} can be obtained at the Solar System or stellar scales in the local Universe, e.g., $\Delta\gamma_{\text{PPN}} \sim 10^{-5}$ given by the Cassini mission [14]. However, the extragalactic constraints on this parameter are much weaker, e.g., $\gamma_{\text{PPN}} = 0.97 \pm 0.09$ on kiloparsec scale by using a nearby lens, ESO 325-G004 given by [20]. Recent reviews of the progress in experimental testing of GR can be found in [21–23]. The issue of whether general relativity breaks down on larger cosmological scales should be further validated [24].

Independent alternative techniques could provide new perspectives. As one of the most important predictions of GR, strong gravitational lensing, especially strong lensing time-delay measurements, have become a powerful tool for studying cosmology and gravity. In addition, lenses with measured time-delay have recently been found to be more powerful cosmological probes with capability to measure the angular diameter distances to the lenses D_d^A by combining the measurements of time-delays, image configuration, and stellar kinematics (spectroscopy of lens galaxy). The possible deviations from GR enter into the lens formula, which further affects the measured D_d^A . However, the possible effects of GR deviation and measured distances are highly degenerate. In other words, the gravitational mass is inferred from the Einstein radius of the lens, but the observed Einstein radius also depends on the distances in the optical system. Fortunately, gravitational wave (GW) signals from inspiraling and merging compact binaries provide absolute distance (luminosity distances D^L) [25] and do not suffer from interstellar extinction effects, which can help to break aforementioned degeneracy. Such combination of time-delay lenses and GWs opens a new possibility to test deviations from GR at galactic scales and high redshifts in the limit of weak gravitational field. This paper is organized as follows: In Sec. II, we present the methodology and observational data. The results and discussion are given in Sec. III. Finally, we summarize our findings in Sec. IV. The natural units of $c = G = 1$ are adopted throughout this letter. The fiducial cosmological model, a flat Λ CDM with $\Omega_m = 0.30$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is assumed for simulating the lensing data and GW data.

II. METHODOLOGY

A. Background

In the weak field limit, background FLRW metric perturbed by the presence of the lens can be written as

$$ds^2 = -(1 + 2\Phi)dt^2 + a(t)^2(1 - 2\Psi)d\mathbf{r}^2, \quad (1)$$

where $a(t)$ is the scale factor, Φ is the Newtonian gravitational potential, and Ψ is the spatial curvature potential. We remark here that Φ and Ψ may be different in some modified gravity theories, i.e., Φ is gravitational potential that responds to the motion of nonrelativistic matter such as baryons and dark matter. The relationship of Ψ to the Newtonian potential is determined by the matter content (e.g., usual matter species in the concordance model plus possible extra scalar fields) and gravity itself. The deviation from GR is quantified by the ratio $\gamma_{\text{PPN}} = \Psi/\Phi$ (denoted as the PPN parameter), with $\gamma_{\text{PPN}} = 1$ in the framework of GR. In modified gravity theories, Φ and Ψ could be unequal with each other [21], and even their ratio could be a function of space [26–28] and time [29]. In this paper, we aim to construct an unbiased method to constrain γ_{PPN} at galactic scales and high redshifts. For simplicity, we assume that γ_{PPN} is independent of the space coordinates, and only depends on time. Such case can be realized in many modified gravity theories, e.g., the nonlocal Gauss-Bonnet gravity.

Strong lensing systems enables to perform two types of mass measurements: gravitational mass of the lens inferred from the lensed images of the source, and dynamical mass obtained from spectroscopic measurements of stellar kinematics of the lensing galaxy. The dynamical mass is sensitive to the Newtonian gravitational potential Ψ only. On the other hand, gravitational mass is sensitive to both potentials, i.e., to the Weyl potential (defined as $\Phi_+ = \frac{\Phi + \Psi}{2} = \frac{1 + \gamma_{\text{PPN}}}{2}\Phi$). Thus, strong lensing provides a natural laboratory to test gravity and further measure the PPN parameter γ_{PPN} by directly comparing the difference between the dynamical mass and gravitational mass or comparing Φ and Φ_+ .

B. Angular diameter distance from strong lensing

Combination of time delay and stellar kinematics measurements, can provide the measurement of the angular diameter distance D_d^A to the lens (deflector). Let us briefly outline the standard procedure for determination of D_d^A pursued in the HOLiCOW program. For a typical strong lensing system, with quasar at the redshift z_s acting as a background source, lensed by a foreground elliptical galaxy (at the redshift z_d), multiple bright images of the active galactic nucleus (AGN) are formed together with the arcs of its host galaxy. Time delays between multiple images can be measured from monitoring of variability of the AGN light curves. From theoretical point of view, lensing time delay is determined by both the geometry of the Universe (different paths of rays forming different images) as well as the Shapiro effects through [30]

$$\Delta t_{\text{AB}} = D_{\Delta t}[\phi(\theta_A, \beta) - \phi(\theta_B, \beta)] = D_{\Delta t}\Delta\phi_{\text{AB}}(\xi_{\text{lens}}), \quad (2)$$

where $\phi(\theta, \beta) = [(\theta - \beta)^2/2 - \psi(\theta)]$ is the Fermat potential at images, β is the source position, ψ is effective lensing

potential (the integral of the Weyl potential along the line-of-sight) obeying the Poisson equation $\nabla^2\psi = 2\kappa$, where κ (so called convergence) is the surface mass density of the lens in units of critical density $\Sigma_{\text{crit}} = D_s^A/(4\pi D_d^A D_{\text{ds}}^A)$, and ξ_{lens} denotes the lens model parameters. The cosmological background is reflected in the so-called ‘‘time delay distance’’ $D_{\Delta t} = (1 + z_d) \frac{D_d^A D_s^A}{D_{\text{ds}}^A}$, where subscripts d and s stand for lens (deflector) and source, and superscript A denotes the angular diameter distance. The key point here is that the Fermat potential difference $\Delta\phi_{\text{AB}}(\xi_{\text{lens}})$ can be reconstructed by high-resolution lensing imaging from space telescopes.

In the framework of parametrized post-Newtonian (PPN) formalism, the inferred lensing mass parameters are rescaled by a factor of $(1 + \gamma_{\text{PPN}})/2$. Hence, we denote the actually inferred lens model parameters in the Fermat potential as ξ'_{lens} . Adopting the notations in work [31], one can write the time-delay distance as

$$D_{\Delta t} = (1 + z_d) \frac{D_d^A D_s^A}{D_{\text{ds}}^A} = \frac{\Delta t_{\text{AB}}}{\Delta\phi_{\text{AB}}(\xi'_{\text{lens}})}. \quad (3)$$

Thus, the time delay distance can be obtained from both the measurements of time delay and the actual reconstructed Fermat potential with parameter ξ'_{lens} . The only difference in this equation is that in the case of $\gamma_{\text{PPN}} \neq 1$ under the PPN framework the inferred lens model parameters are ξ'_{lens} but not the original ξ_{lens} under GR.

On the other hand, assuming some explicit model of the lens such as the simplest singular isothermal sphere (SIS) model (or its extensions like singular ellipsoid SIE or power-law model), observations regarding stellar kinematics (dynamical mass determination) allow to obtain the following distance ratio

$$\frac{D_s^A}{D_{\text{ds}}^A} = \frac{\sigma_v^2}{c^2 J(\xi_{\text{lens}}, \xi_{\text{light}}, \beta_{\text{ani}})}, \quad (4)$$

where σ_v is the line of sight (LOS) projected stellar velocity dispersion of the lens galaxy. This distance ratio provides a valuable extra constraint. The function J captures all of the model components calculated from the lensed images and luminosity-weighted projected velocity dispersion (from the spectroscopy). Since the radial velocity dispersion σ_v can be modeled via the anisotropic Jeans equation, arguments of J function comprise lens model parameters ξ_{lens} , parameters related to luminosity distribution of the lens ξ_{light} and anisotropy parameter β_{ani} . More details concerning modeling issues related to the function J can be found in Sec. 4.6 of [32]. Stellar kinematics of lensing galaxies is sensitive to the Newtonian potential only, which is independent of PPN parameter. Thus, the inferred distance ratio D_s^A/D_{ds}^A combined with the well-measured velocity dispersion is independent of the cosmological model and

time delays, but still relies on the lens model ξ_{lens} [32,33]. The lens model parameter in J is the unrescaled ξ_{lens} . If we replace ξ_{lens} with ξ'_{lens} , the resulted distance ratio shall also be rescaled, correspondingly

$$\frac{2}{1 + \gamma_{\text{PPN}}} \frac{D_s^A}{D_{\text{ds}}^A} = \frac{\sigma_v^2}{c^2 J(\xi'_{\text{lens}}, \xi_{\text{light}}, \beta_{\text{ani}})}. \quad (5)$$

The PPN parameter is introduced here explicitly, and this formula was widely used in the works on constraining the GR parameter for strong gravitational lensing systems [34–37]. Furthermore, we can define $D_d'^A = \frac{1 + \gamma_{\text{PPN}}}{2} D_d^A$, and by combining Eqs. (3) and (5), the angular diameter distance to lens can be expressed as

$$D_d'^A = \frac{1}{1 + z_d} \frac{c \Delta t_{\text{AB}}}{\Delta\phi_{\text{AB}}(\xi'_{\text{lens}})} \frac{c^2 J(\xi'_{\text{lens}}, \xi_{\text{light}}, \beta_{\text{ani}})}{\sigma_v^2}. \quad (6)$$

Let us stress that this distance is unaffected by cosmic opacity.¹ In fact, one could use the time delay distance (Eq. (3) to test gravity theory, but this will inevitably introduce cosmic curvature parameter, because D_{ds}^A is not directly observable. More importantly, one of the main obstacle for lensing mass modeling is the mass-sheet degeneracy, which is completely circumvented by testing gravity using angular diameter distances. Many factors such as the Hubble constant, dark energy model and cosmic opacity (e.g. caused by light extinction by intergalactic dust) affect determination of cosmological distances. This would bias most straightforward tests of GR and reliable test of GR should be pure enough to ensure its validity and being unbiased. Therefore, we point out in this work that GR testing by gravitational lensing could be independent of these factors.

The current and future programs for time-delay with lensed quasars have great progress, such as H0LiCOW [8], COSMOGRAIL [38], STRIDES [39], and SHARP, which have now combined in the TDCOSMO collaboration [40,41]. Although our proposed method has many advantages, yet the existing sample of time-delay lenses is still insufficient, so we turn to a new generation of wide and deep sky observations and make appropriate simulations. Regarding the assessment of the uncertainties of D_d^A , for each lens one should followed the process presented in [42], which resulted in 1.8% precision of D_d^A for RXJ1131-like system. In a realistic situation, the analysis of each lens is very complicated, and such simulations are beyond this work. Instead, we follow the work of [43] and consider that

¹In the case of gravitational lensing observables, cosmic opacity can change the absolute intensity (magnitude) of images but not the relative intensity, thus not biasing the distance determination in strong lensing system. Besides, the velocity dispersion based on spectroscopic measurements are also unaffected by the opacity.

the main sources of uncertainties give the expected uncertainty level for D_d^A determination. These uncertainties, including the measurement of time delays, the mass fluctuation along the LOS, recovering stellar kinematics from measured velocity dispersions, the lens mass parametrization from highly resolved imaging will result in a few percent level of D_d^A determination accuracy. We take the 5% uncertainty on the measurement of D_d^A as the best case scenario, which was also used in [43,44] and as the aim of the H0LiCOW program [45]. In the conservative scenario, we also adopt 10% uncertainty on D_d^A for comparison. This conservative uncertainty is estimated from the [46], which used simulations to obtain realistic error estimates with current/upcoming instruments, e.g., OH Suppressing InfraRed Imaging Spectrograph on Keck [47], and Near-Infrared Spectrograph on *James Webb Space Telescope* [48], and InfraRed Imaging Spectrograph on the Thirty Metre Telescope [49].

In the near future, the next generation of wide and deep sky surveys, with improved depth, area and resolution, may increase the current galactic-scale lens sample sizes by orders of magnitude. For example, the upcoming Legacy Survey of Space and Time (LSST) conducted in Vera Rubin Observatory, is expected to find ten thousand lensed quasars [50]. However, considering that the metric efficiency is about 20%, the Time Delay Challenge (TDC) program showed that only 400 well-measured time delays are available [51]. More precisely, LSST will discover 10000 lensed quasars, but it has been estimated that only 2000 lensed quasar systems would have a long enough ($10 < \Delta t < 120$ days) time delays measurements. The TDC expects to be able to make time delay measurements with the high accuracy in at least 20% efficiency of an LSST sample of 2000 lenses. This would correspond to a well-measured sample of around 400 lensed quasars. In this simulation, we use the OM10 catalog of mock lenses² [52], which has the distribution of time-delay lenses expected from the LSST. Besides, one would need auxiliary data comprising high resolution imaging and spectroscopy from instruments such as the Hubble Space Telescope (HST) and ground-based observatories, in order to have all ingredients needed for accurate lens mass modeling and determination of D_d^A . It should be pointed out that LSST will discover a number of fainter, smaller-separation lenses where it is not clear that the same level of precision of lens model reconstruction can be reached. Therefore, we follow the work of [43] and set the following criteria: (1) the lensed quasar image separation should be greater than 1 *arcsec*; (2) the third brightest quasar image has *i*-band magnitude $m_i < 21$ mag; (3) the lens galaxy has $m_i < 22$ mag; (4) the lensed quasar image is quadruple imaging lens systems (this type of system provide more information to break the

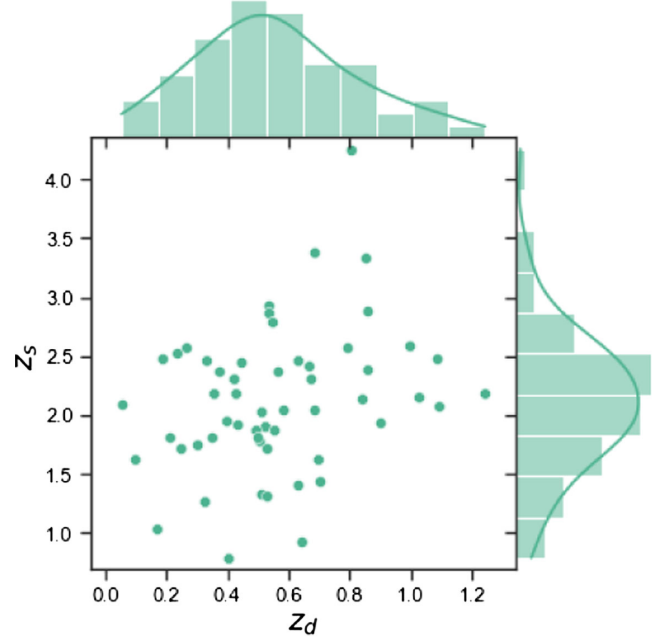


FIG. 1. Redshift distribution of lenses and sources in a sample of 55 strong lensing systems expected to be observed in current and upcoming projects.

source-position transformation [53,54]). After applying these conditions to the OM10 catalog, 55 high-quality quadruple lens systems are kept [43]. Fig. 1 shows the source and the lens redshift distributions of 55 quadruple lenses that match the selection criteria.

C. Luminosity distance from gravitational waves

From Eq. (6), it is clear that the PPN parameter and the angular diameter distance are entangled. In order to break this degeneracy and complete the test of the gravity theory, one needs independent, complementary distance measurements. For this purpose, we turn to standard sirens.

Simultaneous detections of GW signal [55] from the binary neutron star (NS-NS) merger and the electromagnetic (EM) counterparts [56] from the same transient source opened a new era of multimessenger astronomy. GW signals from coalescing compact binary systems, i.e. binary black holes (BH-BH) [57], neutron stars (NS-NS) or possible mixed neutron star–black hole (NS-BHs) systems provide us direct measurements of luminosity distances D^L . Hence, they are called standard sirens [25]. Standard sirens are self-calibrating, meaning that the luminosity distances can be directly inferred from the detected waveforms using matched-filter method. Redshift of the source z is unfortunately non-measurable in GW domain alone, and one has to identify the host galaxy in order to obtain z . There are two advantages of standard sirens: firstly, GW signals are unaffected by cosmic opacity, and propagate through the Universe without any absorption and dissipation. Second, GW signals from standard sirens provide

²<https://github.com/drphilmarshall/OM10>.

direct luminosity distance measurements instead of the relative distances as in case of SNe Ia. Thus D^L from standard sirens do not need to be anchored to the cosmic distance ladder.

Einstein Telescope (ET)³ [58,59] is the third-generation gravitational wave detector, which is designed to have a fantastic sensitivity in the frequency range of 1–10⁴ Hz. See [60] for the details of the ET conceptual design study. We simulate GW signals based on the foreseen performance of the ET. Compared with the advanced detectors such as AdLIGO and AdVirgo, such proposed third-generation detector aims for a broadband factor of 10 sensitivity improvement, especially for the characteristic distance sensitivity [61]. In this simulation, we follow the work by [62,63], and assume that the GW sources are caused by binary merger of the NS with the NS, which can generate the intense short gamma ray bursts (SGRBs) with measurable source redshifts, or binary merger of the BH with the BH, which the redshift information comes from a statistical analysis over a catalogue of potential host galaxies.

Following the strategy proposed in [62,63], the uncertainty for luminosity distance obtained from the GW signal from merger event is

$$\sigma_{\text{tot}}^2 = \sigma_{\text{inst}}^2 + \sigma_{\text{len}}^2 + \sigma_{\text{pec}}^2, \quad (7)$$

where σ_{inst} is the instrumental measurement uncertainty and calculated by

$$\sigma_{\text{inst}} \simeq \sqrt{\left\langle \frac{\partial \mathcal{H}(f)}{\partial D^{L,GW}}, \frac{\partial \mathcal{H}(f)}{\partial D^{L,GW}} \right\rangle^{-1}}, \quad (8)$$

where $\mathcal{H}(f)$ is the Fourier transform of *strain* $h(t)$ of a chirp waveform in frequency domain. Since the *strain* h and \mathcal{H} is inversely proportional to $D^{L,GW}$, we obtain $\sigma_{\text{inst}} \simeq D^{L,GW}/\rho$, where ρ denotes the combined signal-noise-ratio (SNR) for the network of independent interferometers, determined by the square root of the inner product of $\mathcal{H}(f)$. Considering the uncertainty from the inclination angle ι would also affect the SNR, the maximal inclination angle effect from $\iota = 0^\circ$ to $\iota = 90^\circ$ on the SNR, we choose to double the instrumental uncertainty of D^L as the upper limit, as was proposed in the recent study of the third-generation GW detector (ET) based on Fisher information matrix [64], i.e.,

$$\sigma_{\text{inst}} \simeq \frac{2D^{L,GW}}{\rho}. \quad (9)$$

The second uncertainty term σ_{len} is the weak gravitational lensing effect caused by the large scale structure, and

ignoring it would result in the biased distance estimation. Following the work [62,63], this uncertainty is modeled as [65,66]

$$\sigma_{\text{lens}} = 0.066 \left(\frac{1 - (1+z)^{-0.25}}{0.25} \right)^{1.8} D^L(z) F_{\text{delens}}(z), \quad (10)$$

where $F_{\text{delens}}(z) = 1 - \frac{0.3}{\pi/2} \arctan z/z_*$, with $z_* = 0.073$ [66]. The latter factor takes into account the possibility to reduce the uncertainty due to weak lensing with the future detectors such as the Extremely Large Telescope [67]. The final uncertainty term σ_{pec} is caused by the peculiar velocity of the host galaxy [68–70], and can be approximated by a fitting formula [71]

$$\sigma_{\text{pec}} = \left[1 + \frac{c(1+z)^2}{H(z)D^L(z)} \right] \frac{\sqrt{\langle v^2 \rangle}}{c} D^L(z), \quad (11)$$

where the averaged peculiar velocity $\langle v^2 \rangle$ is set as 500 km/s, in agreement with the observed values in galaxy catalogs [72].

Although the next generation ground-based detector ET is expected to detect tens or hundreds of thousand NS-NS inspiral events per year up to the redshift $z \sim 2$ and NS-BH mergers up to $z \sim 5$ with $\rho > 8$, yet SGRBs necessary for z measurements are strongly beamed, and only the nearly face-on configurations of NS-NS or NS-BH mergers are useful. Probability of their occurrence is $\sim 10^{-3}$. Assuming that the redshift distribution of simulated GW sources follows the cosmic star formation history [73], we sample the mass of neutron star and black hole within $[1, 2]M_\odot$ and $[3, 10]M_\odot$, respectively. Thus, we simulate 1000 GW events observable in the ET and their accurate redshifts up to $z_{GW} = 5$. Matching these events with the redshifts of the lens z_d needed for testing the PPN parameter, reduces the sample to 300 sources within redshift $z_{GW} < 1.25$.

III. RESULTS AND DISCUSSION

To demonstrate the performance of our method to detect the possible deviation of GR, we use three parametrized forms of γ_{PPN} : (1) $\gamma_{\text{PPN}} = \gamma_0$ being constant; (2) $\gamma_{\text{PPN}} = 1 + \gamma_1 * z$ changing with redshift; (3) $\gamma_{\text{PPN}} = 1 + \gamma_2 * a$ changing with the scale factor $a = 1/(1+z)$, which unlike the redshift is the gravitational degree of freedom. The above mentioned PPN parameters and their uncertainties are fitted by minimizing the χ^2 objective function defined in the following way:

$$\chi^2 = \sum_{i=1}^{55} \left[\frac{(D_{d,i}^{A} - D_{GW,i}^L (1+z)^{-2})^2}{\sigma_{D_{GW,i}^L}^2 + \sigma_{D_{d,i}^A}^2} \right], \quad (12)$$

where γ_{PPN} parameter is implicitly present in D_d^A . The total uncertainty consists of observational uncertainties from

³The Einstein Telescope Project, <https://www.et-gw.eu/et/>.

TABLE I. Summary of the constraints on the γ_{PPN} parameter using the combination of strong lensing systems plus GW events.

$\sigma_{D^A}/D^A = 5\%$	γ_0	γ_1	γ_2
$\gamma_{\text{PPN}} = \gamma_0$	$1.003^{+0.018}_{-0.018}$
$\gamma_{\text{PPN}} = 1 + \gamma_1 * z$...	$0.006^{+0.030}_{-0.031}$...
$\gamma_{\text{PPN}} = 1 + \gamma_2 * a$	$0.008^{+0.028}_{-0.028}$
$\sigma_{D^A}/D^A = 10\%$	γ_0	γ_1	γ_2
$\gamma_{\text{PPN}} = \gamma_0$	$1.003^{+0.032}_{-0.033}$
$\gamma_{\text{PPN}} = 1 + \gamma_1 * z$...	$0.007^{+0.050}_{-0.050}$...
$\gamma_{\text{PPN}} = 1 + \gamma_2 * a$	$0.004^{+0.050}_{-0.049}$

GW and strong lensing contributions (see Sec. II B and C). The uncertainty $\sigma_{D_{\text{GW}}} = \sigma_{\text{tot}}/(1+z)^2$ is calculated using σ_{tot} given by Eq. (7), strong lensing related uncertainties are given by $\sigma_{D_d^A} = 5\%D_d^A$ and $\sigma_{D_d^A} = 10\%D_d^A$ for the best case scenario and the conservative scenario, respectively. We simulated ten thousand realizations of the data with different random seeds and repeated the minimization process to get an unbiased estimation for the PPN parameters. We emphasize that in this work we do not attempt to constrain PPN parameter from real data, but propose an unbiased approach and discuss its ability to explore the presence of possible evolution or deviations from GR on simulated data based on future observational forecasts.

The final probability distribution functions (PDFs) for γ_{PPN} are reported in Table I and displayed in Fig. 2. Since the PDFs are approximately Gaussian-like, we calculate the standard deviations as the 1σ uncertainty levels. For the 5% uncertainty (optimistic case) on the D_d^A measurements from strong lensing systems, our results regarding $\gamma_{\text{PPN}} = \gamma_0$ constant case suggest that the uncertainty on γ_0 is 0.018, corresponding to 1.8% precision. The precision obtained by our method is > 10 times better than the one obtained from current four well-measured H0LiCOW quasars lenses, and almost 3 times better than the one from simulated future 40 lensed quasar systems [31]. Meanwhile, the precision obtained by our method is > 4 times better than expected from the simulated future 10 strongly lensed fast radio bursts (FRBs) systems [74]. This result demonstrates the superiority of our method. There are two reasons for this. First, only one parameter γ_0 is estimated in this work, whereas the previous work was simultaneously estimating H_0 and PPN parameters, and the distance information was derived there by assuming a concrete cosmological model. This affects the ability to constrain PPN parameter alone. Second, one of the main obstacle for lensing mass modeling is the mass-sheet degeneracy,⁴ which is considered the dominant source of residual modeling error in

⁴It is a mathematical degeneracy that leaves the lensing observables unchanged, while rescaling the absolute time delay, and thus the inferred H_0 .

time-delay cosmography. Our method circumvents the mass-sheet degeneracy by using angular diameter distances to perform the test of gravity, which greatly reduces systemic uncertainty and further improves the precision on PPN parameter in our work.

Some modified gravitational theories, such as the non-local Gauss-Bonnet gravity [29] require that PPN parameter may evolve with redshift or the scale factor. With this motivation, we further assume a parametrized form of γ_{PPN} changing with redshift or the scale factor. The results are also shown in Fig. 2 and reported in Table I. We obtain $\Delta\gamma_1 = 0.030$ and $\Delta\gamma_2 = 0.028$ in the optimistic scenario. Compared with PPN parameter treated as a constant, the precision of PPN parameters evolving with redshift or the scale factor is noticeably reduced. In Fig. 3, we plot the reconstruction of γ_{PPN} for the different parametrizations with their uncertainties, and a prediction from a nonlocal Gauss-Bonnet theory with specific parameters [75]. It is clearly seen that if GR does not hold and is replaced by this nonlocal gravity, our method will distinguish them with very high precision. Here, for the purple dashed line in Fig. 3, we adopt the conventions in [29] and set the model parameter $\alpha = 5.6 \times 10^{-6} c^4/H_0^4$, the integral constant $C_1 = 0$, and the matter density $\rho \propto a^{-3}$. Note that, generally nonlocal Gauss-Bonnet gravity predicts $|1 - \gamma_{\text{PPN}}| = \mathcal{O}(1)$ [29]. The α we use here is too small to explain the cosmic late-time acceleration, which is the initial motivation to propose this theory [75]. Therefore, this purple line corresponds to a toy model. This illustration is just to show that γ_{PPN} in modified gravity can be very close to 1 at low redshifts and deviate from 1 at high redshifts. Such possible behavior of γ_{PPN} highlights the importance of constraining it at high redshifts. In addition, the method proposed in this paper is independent of any specific cosmological models and gravitational theories.

The main purpose of this work is to quantify the ability of upcoming TDSL plus GW to test GR against modified gravity theories. The key question now arises: Is this combination of observations sufficient to detect possible effects of deviation from GR? They can hardly become competitive to high-precision measurements of $\gamma_{\text{PPN}} = 1 + (2.1 \pm 2.3) \times 10^{-5}$ in the solar system made by Cassini mission. Considering the 10% uncertainty (conservative case) on the D_d^A measurements from strong lensing systems, the precision on PPN parameter is almost half of that for the optimistic case. This suggests that the main source of uncertainty in our method is the observational error from the angular diameter distance measured by the gravitational lensing system. This is reasonable, considering that only 55 gravitational lensing systems have been used in this work, and the uncertainty of the PPN parameters caused by systematic errors is much greater than that caused by statistical errors. If we take a more optimistic estimation presented by [42], the future uncertainty on D_d^A can even decrease dramatically to the order of 1.8%, which indicates

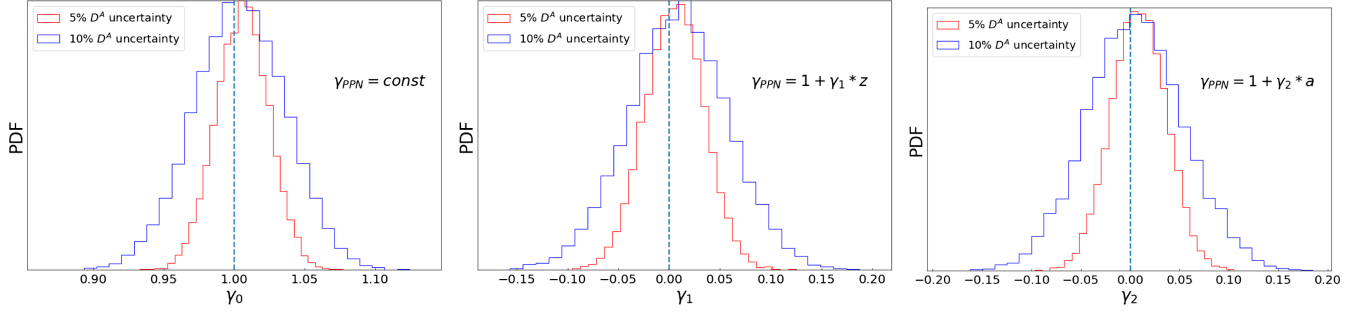


FIG. 2. Posterior distributions for PPN parameters obtained by the combination of the future LSST survey and ET detector yields. Three panels illustrate the following parametrizations: $\gamma_{\text{PPN}} = \gamma_0$ constant case (left), $\gamma_{\text{PPN}} = 1 + \gamma_1 * z$ varying with redshift (middle), and $\gamma_{\text{PPN}} = 1 + \gamma_2 * a$ varying with the scale factor (right).

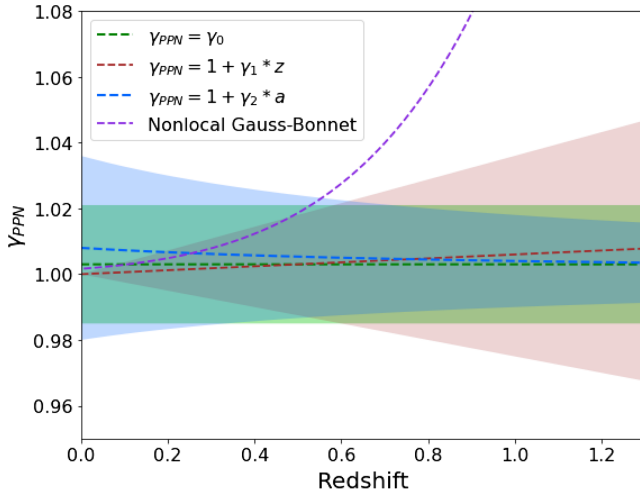


FIG. 3. The reconstructed γ_{PPN} for the three parametrizations with their 1σ uncertainties. The purple dashed line represents the nonlocal Gauss-Bonnet theory. The green, pink, and blue bands correspond to the 1σ regions of the best-fitting in the constant, evolution with the redshift, and evolution with the scale factor cases, respectively.

a more promising future of TDSL plus GW technique for testing general relativity. With the upcoming sky surveys such as LSST, the number of lensing systems will increase dramatically. Hence, one can expect that our approach will greatly improve the measurement precision on γ_{PPN} at the galactic scale.

IV. CONCLUSION

In this work, we investigated, for the first time, the possibility of estimating post-Newtonian parameter γ_{PPN} by combining well measured strongly lensed systems with gravitational wave signals. The combination of strong lensing and self calibrating standard sirens observed in GWs enables to avoid the possible bias that comes with assuming a particular cosmological model. Meanwhile, such combination provides a relatively pure and unbiased

method for GR testing at the galactic scales and high redshifts, and brings various benefits. It can reduce possible bias on GR testing induced by cosmic curvature, cosmic opacity, the dark energy, the Hubble constant (the D_d^A provided by strong lensing and the D^L provided by gravitational waves are both absolute distances and are therefore unaffected by the Hubble constant), and the mass-sheet degeneracy inside the lens. Combining measurements of time delay, stellar velocity dispersion, high resolution images, LOS environment modeling from upcoming LSST survey and GWs acting standard siren from future ET detector, we obtained the precision of $\sim 1.8\%$ (optimistic case) and 3.2% (conservative case) in the case of γ_{PPN} being a constant. Second, we considered the case of γ_{PPN} parameter displaying evolution with redshift or the scale factor and studied the precision with which this effect can be revealed. This combination of data also allows us to infer post-Newtonian parameter and detect possible deviation of GR at different redshifts. Although the precision of testing GR at the galactic scale can hardly achieve the results competitive to the measurements within the solar system, yet our technique would be able to distinguish the departures from the GR quite precisely. We also indicated a more promising future of TDSLs plus GWs for testing general relativity.

As a final remark, there are many potential ways to improve our method. For instance, current and future surveys like the Dark Energy Survey (DES) [39], the Hyper SuprimeCam Survey [76], and the Legacy Survey of Space and Time (LSST) [52,77] will bring us hundreds of thousands of lensed quasars in the most optimistic discovery scenario. Even if only some small fraction of them will have precise measurements of time delays between multiple images, the resulting statistics will outshine current catalogs. With high-quality auxiliary observations, one can use high-cadence, high-resolution and multi-filter imaging of the resolved lensed images, to derive an accurate determination of the Fermat potential, which will increase the precision of time delay distance by an order of magnitude. On the other hand, the first strongly

lensed supernova “SN Refsdal” [78] with multiple images opened a new window for astrophysics and cosmology with the concept of “lensed transients.” We also expect that the discovery of various lensed transients explosive sources like gamma ray bursts (GRBs), fast radio bursts (FRBs) and even GWs will give us new advantages over the traditional targets in studying the Universe [79,80]. At the same time, we also expect in the future, the synergies between GW and EMW observations in various bands will yield reliable cosmological probe. It is reasonable to expect that our approach will play an increasingly important role in precise testing the validity of general relativity.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China under Grants No. 12222302, No. 12203009, No. 11973034. Tonghua Liu was supported by Chutian Scholars Program in Hubei Province (No. X2023007); Kai Liao was supported by Funds for the Central Universities (Wuhan University Grant No. 1302/600460081). Shuxun Tian was supported by the Fundamental Research Funds for the Central Universities (Beijing Normal University No. 10800-310400209522). Marek Biesiada was supported by Hubei Province Foreign Expert Project (No. 2023DJC040).

-
- [1] A. G. Riess, A. V. Filippenko, P. Challis *et al.*, *Astron. J.* **116**, 1009 (1998).
- [2] S. Perlmutter, G. Aldering, G. Goldhaber *et al.*, *Astrophys. J.* **517**, 565 (1999).
- [3] R. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984).
- [4] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, edited by Steven Weinberg (Wiley-VCH, Hoboken, 1972), p. 688, ISBN: 0-471-92567-5.
- [5] P. A. R. Ade, N. Aghanim *et al.* (Planck Collaboration), *Asrton. Astrophys.* **594**, A13 (2016).
- [6] N. Aghanim, Y. Akrami *et al.* (Planck Collaboration), *Asrton. Astrophys.* **641**, A6 (2020).
- [7] A. G. Riess, S. Casertano, W. Yuan, J. B. Bowers, L. Macri, J. C. Zinn, and D. Scolnic, *Astrophys. J.* **908**, L6 (2021).
- [8] K. C. Wong, S. H. Suyu, G. C.-F. Chen *et al.*, *Mon. Not. R. Astron. Soc.* **498**, 1420 (2020).
- [9] E. Di Valentino, A. Melchiorri, and J. Silk, *Nat. Astron.* **4**, 196 (2020).
- [10] W. Handley, *Phys. Rev. D* **103**, L041301 (2021).
- [11] E. Abdalla, G. F. Abellán, A. Aboubrahim *et al.*, *J. High Energy Astrophys.* **34**, 49 (2022).
- [12] S. Weinberg, *Rev. Mod. Phys.* **61**, 1 (1989).
- [13] N. Ashby, *Phys. Today* **55**, No. 5, 41 (2002).
- [14] B. Bertotti, L. Iess, and P. Tortora, *Nature (London)* **425**, 374 (2003).
- [15] J. M. Weisberg and Y. Huang, *Astrophys. J.* **829**, 55 (2016).
- [16] L. Shao, *Phys. Online J.* **14**, 173 (2021).
- [17] M. Isi, K. Chatziioannou, and W. M. Farr, *Phys. Rev. Lett.* **123**, 121101 (2019).
- [18] H. Ding, A. T. Deller, E. Fonseca, I. H. Stairs, B. Stappers, and Andrew Lyne, *Astrophys. J. Lett.* **921**, L19 (2021).
- [19] J. R. Gair, M. Vallisneri, S. L. Larson, and J. G. Baker, *Living Rev. Relativity* **16**, 7 (2013).
- [20] T. E. Collett, L. J. Oldham, R. J. Smith, R. J. Smith, M. W. Auger, K. B. Westfall, D. Bacon, R. C. Nichol, K. L. Masters, K. Koyama, and R. van den Bosch, *Science* **360**, 1342 (2018).
- [21] C. M. Will, *Living Rev. Relativity* **17**, 4 (2014).
- [22] M. Ishak, *Living Rev. Relativity* **22**, 1 (2019).
- [23] J.-P. Uzan, *Gen. Relativ. Gravit.* **42**, 2219 (2010).
- [24] K. Koyama, *Rep. Prog. Phys.* **79**, 046902 (2016).
- [25] B. F. Schutz, *Nature (London)* **323**, 310 (1986).
- [26] W. Hu and I. Sawicki, *Phys. Rev. D* **76**, 064004 (2007).
- [27] K. Saaidi, A. Mohammadi, and H. Sheikahmadi, *Phys. Rev. D* **83**, 104019 (2011).
- [28] D. B. Thomas, T. Clifton, and T. Anton, *J. Cosmol. Astropart. Phys.* **04** (2023) 016.
- [29] S. X. Tian and Z.-H. Zhu, *Phys. Rev. D* **99**, 064044 (2019).
- [30] I. I. Shapiro, *Phys. Rev. Lett.* **13**, 789 (1964).
- [31] T. Yang, S. Birrer, and B. Hu, *Mon. Not. R. Astron. Soc.* **497**, L56 (2020).
- [32] S. Birrer, T. Treu, C. E. Rusu *et al.* *Mon. Not. R. Astron. Soc.* **484**, 4726 (2019).
- [33] S. Birrer, A. Amara, and A. Refregier, *J. Cosmol. Astropart. Phys.* **08** (2016) 020.
- [34] S. Cao, X. Li, M. Biesiada, T. Xu, Y. Cai, and Z.-H. Zhu, *Astrophys. J.* **835**, 92 (2017).
- [35] X.-H. Liu, Z.-H. Li, J.-Z. Qi, and X. Zhang, *Astrophys. J.* **927**, 28 (2022).
- [36] J.-J. Wei, Y. Chen, S. Cao, and X.-F. Wu, *Astrophys. J. Lett.* **927**, L1 (2022).
- [37] T. Liu and K. Liao, *Mon. Not. R. Astron. Soc.* **528**, 1354 (2024).
- [38] A. Eigenbrod, F. Courbin, C. Vuissoz, G. Meylan, P. Saha, and S. Dye, *Astron. Astrophys.* **436**, 25 (2005).
- [39] T. Treu, A. Agnello, M. A. Baumer *et al.*, *Mon. Not. R. Astron. Soc.* **481**, 1041 (2018).
- [40] S. Birrer and T. Treu, *Astron. Astrophys.* **649**, A61 (2021).
- [41] A. J. Shajib, P. Mozumdar, G. C.-F. Chen *et al.*, *Astron. Astrophys.* **673**, A9 (2023).
- [42] A. Yıldırım, S. H. Suyu, and A. Halkola, *Mon. Not. R. Astron. Soc.* **493**, 4783 (2020).
- [43] I. Jee, E. Komatsu, S. H. Suyu, and D. Huterer, *J. Cosmol. Astropart. Phys.* **04** (2016) 031.
- [44] E. V. Linder, *Phys. Rev. D* **84**, 123529 (2011).
- [45] S. H. Suyu, V. Bonvin, F. Courbin *et al.*, *Mon. Not. R. Astron. Soc.* **468**, 2590 (2017).

- [46] A. J. Shajib, S. Birrer, T. Treu *et al.*, *Mon. Not. R. Astron. Soc.* **494**, 6072 (2020).
- [47] J. Larkin, M. Barczys, A. Krabbe *et al.*, *New Astron. Rev.* **50**, 362 (2006).
- [48] S. M. Birkmann, P. Ferruit, G. Giardino *et al.*, *Astron. Astrophys.* **661**, A83 (2022).
- [49] S. A. Wright, G. Walth, T. Do *et al.*, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (SPIE, Berkeley, 2016), Vol. 9909, p. 990905, 10.1117/12.2233182.
- [50] Ž. Ivezić, S. M. Kahn, J. A. Tyson *et al.*, *Astron. Astrophys.* **873**, 111 (2019).
- [51] K. Liao, T. Treu, P. Marshall *et al.*, *Astron. Astrophys.* **800**, 11 (2015).
- [52] M. Oguri and P. J. Marshall, *Mon. Not. R. Astron. Soc.* **405**, 2579 (2010).
- [53] P. Schneider and D. Sluse, *Astron. Astrophys.* **559**, A37 (2013).
- [54] P. Schneider and D. Sluse, *Astron. Astrophys.* **564**, A103 (2014).
- [55] B. P. Abbott *et al.* (LIGO Scientific and the Virgo Collaborations), *Phys. Rev. Lett.* **119**, 161101 (2017).
- [56] B. P. Abbott, R. Abbott, T. D. Abbott *et al.*, *Astrophys. J. Lett.* **848**, L12 (2017).
- [57] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 061102 (2016).
- [58] N. Dalal, D. E. Holz, S. A. Hughes, and B. Jain, *Phys. Rev. D* **74**, 063006 (2006).
- [59] S. R. Taylor, J. R. Gair, and I. Mandel, *Phys. Rev. D* **85**, 023535 (2012).
- [60] M. Abernathy *et al.*, Einstein gravitational wave telescope: Conceptual design study, Document No. ET-0106A-10, 2011, available from European Gravitational Observatory.
- [61] S. R. Taylor and J. R. Gair, *Phys. Rev. D* **86**, 023502 (2012).
- [62] M. Califano, I. de Martino, D. Vernieri *et al.*, *Mon. Not. R. Astron. Soc.* **518**, 3372 (2023).
- [63] M. Califano, I. de Martino, D. Vernieri, and S. Capozziello, *Phys. Rev. D* **107**, 123519 (2023).
- [64] R.-G. Cai and T. Yang, *Phys. Rev. D* **95**, 044024 (2017).
- [65] C. M. Hirata, D. E. Holz, and C. Cutler, *Phys. Rev. D* **81**, 124046 (2010).
- [66] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein, and A. Petiteau, *J. Cosmol. Astropart. Phys.* **04** (2016) 002.
- [67] L. Speri, N. Tamanini, R. R. Caldwell, J. R. Gair, and B. Wang, *Phys. Rev. D* **103**, 083526 (2021).
- [68] J. Hjorth, A. J. Levan, N. R. Tanvir, J. D. Lyman, R. Wojtak, S. L. Schröder, I. Mandel, C. Gall, and S. H. Bruun, *Astrophys. J. Lett.* **848**, L31 (2017).
- [69] C. Howlett and T. M. Davis, *Mon. Not. R. Astron. Soc.* **492**, 3803 (2020).
- [70] S. Mukherjee, G. Lavaux, F. R. Bouchet, J. Jasche, B. D. Wandelt, S. Nissanke, F. Leclercq, and K. Hotokezaka, *Astron. Astrophys.* **646**, A65 (2021); R. Schneider, V. Ferrari, and S. Matarrese, *Mon. Not. R. Astron. Soc.* **324**, 797 (2001).
- [71] B. Kocsis, Z. Frei, Z. Haiman, and K. Menou, *Astrophys. J.* **637**, 27 (2006).
- [72] R. Cen and J. P. Ostriker, *Astrophys. J.* **538**, 83 (2000).
- [73] R. Schneider, V. Ferrari, S. Matarrese, and S. F. P. Zwart, *Mon. Not. R. Astron. Soc.* **324**, 797 (2001).
- [74] R. Gao, Z. Li, and H. Gao, *Mon. Not. R. Astron. Soc.* **516**, 1977 (2022).
- [75] S. Capozziello, E. Elizalde, S. Nojiri, and S. D. Odintsov, *Phys. Lett. B* **671**, 193 (2009).
- [76] A. More, C.-H. Lee, M. Oguri *et al.*, *Mon. Not. R. Astron. Soc.* **465**, 2411 (2017).
- [77] T. E. Collett, *Astrophys. J.* **811**, 20 (2015).
- [78] P. L. Kelly, S. A. Rodney, T. Treu *et al.*, *Science* **347**, 1123 (2015).
- [79] K. Liao, M. Biesiada, and Z.-H. Zhu, *Chin. Phys. Lett.* **39**, 119801 (2022).
- [80] K. Liao, X.-L. Fan, X. Ding, M. Biesiada, and Z.-H. Zhu, *Nat. Commun.* **8**, 1148 (2017).