The Astrophysical Journal, 950:158 (15pp), 2023 June 20 https://doi.org/10.3847/1538-4357/accea3
© 2023. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

Harnessing Unresolved Lensed Quasars: The Mathematical Foundation of the Fluctuation Curve

Satadru Bag[1], Wuhyun Sohn[1], Arman Shafieloo[1,2], and Kai Liao[3]
[1] Korea Astronomy and Space Science Institute, Daejeon 34055, Republic of Korea; satadru@kasi.re.kr
[2] University of Science and Technology, Daejeon 34113, Republic of Korea; shafieloo@kasi.re.kr
[3] School of Physics and Technology, Wuhan University, Wuhan 430072, People's Republic of China
Received 2022 October 17; revised 2023 March 19; accepted 2023 April 18; published 2023 June 20

Abstract

Strong gravitational lensed quasars (QSOs) have emerged as powerful and novel cosmic probes as they can deliver crucial cosmological information, such as a measurement of the Hubble constant, independent of other probes. Although the upcoming LSST survey is expected to discover 10^3–10^4 lensed QSOs, a large fraction will remain unresolved due to seeing. The stochastic nature of the quasar intrinsic flux makes it challenging to identify lensed quasars and measure the time delays using unresolved light-curve data only. In this regard, Bag et al. introduced a data-driven technique based on the minimization of the fluctuation in the reconstructed image light curves. In this article, we delve deeper into the mathematical foundation of this approach. We show that the lensing signal in the fluctuation curve is dominated by the autocorrelation function (ACF) of the derivative of the joint light curve. This explains why the fluctuation curve enables the detection of the lensed QSOs only using the joint light curve, without making assumptions about QSO flux variability, nor requiring any additional information. We show that the ACF of the derivative of the joint light curve is more reliable than the ACF of the joint light curve itself because intrinsic quasar flux variability shows significant autocorrelation up to a few hundred days (as they follow a red power spectrum). In addition, we show that the minimization of fluctuation approach provides even better precision and recall as compared to the ACF of the derivative of the joint light curve when the data have significant observational noise.

Unified Astronomy Thesaurus concepts: Quasars (1319); Strong gravitational lensing (1643); Astrostatistics techniques (1886)

1. Introduction

Strong gravitational lensed systems have emerged as a powerful and novel cosmic probe (see, e.g., Treu & Marshall 2016 for a review). They can deliver cosmological information independent of other probes such as the Type Ia supernovae (SNe), Baryon Acoustic Osculations (BAO), and Cosmic Microwave Background (CMB). Time-delay measurements, together with accurate lens modeling, allow us to directly estimate the present epoch value of the cosmic expansion rate, i.e., the Hubble constant (H_0) (Refsdal 1964; Refsdal & Bondi 1964; Saha et al. 2006; Oguri 2007; Bonvin et al. 2017; Birrer et al. 2020; Wong et al. 2020; Birrer & Treu 2021). Therefore, "time-delay cosmography" can play a crucial role in elucidating the ongoing H_0 tension between the local measurements (Abdalla et al. 2022; Riess et al. 2022) and early-universe probes like the CMB (Aghanim et al. 2020). Other applications of strong lensing in cosmology and astrophysics are summarized in the review by Treu (2010).

For time-delay measurements, one needs time-variable sources, such as quasars (QSOs) and supernovae (SNe). Lensed SNe (Oguri 2019; Liao et al. 2022; Suyu et al. 2023) are extremely rare as only four with multiple images have been discovered so far (Kelly et al. 2015; Goobar et al. 2017, 2022; Rodney et al. 2021). In comparison, lensed QSOs are more abundant and thus remain to be the primary source for the time-delay cosmography (however, lensed SNe could be at the forefront of time-delay cosmography in the next decade (Suyu et al. 2020)). Although hundreds of lensed QSOs are known (Lemon et al. 2023), only a few have been used for cosmology. For example, using only six "good-quality" lensed QSOs the H0LiCOW team (Suyu et al. 2017) measured the Hubble constant with 2.4% uncertainty (Wong et al. 2020), under standard assumptions about the mass density profile of the deflector, and a seventh brings the precision to 2% (Millon et al. 2020; Shajib et al. 2020). However, if one drops the assumptions and adopts density profiles that are maximally degenerate with H_0 through the mass sheet degeneracy, the uncertainty increases to 9% (Birrer et al. 2020), highlighting the need for substantially larger samples. While the uncertainty can also be reduced by additional information per lens (especially stellar kinematics), a powerful way of improving the H_0 precision is to increase the sample volume significantly (Sonnenfeld 2021). For example, observations of hundreds of lensed systems will deliver sub-percent uncertainty and accuracy, regardless of any assumption on their mass density profile (Jee et al. 2016; Birrer & Treu 2021).

The angular separation of images are typically of the order $\sim 1''$–$2''$ for galaxy-scale lenses (Narayan & Bartelmann 1996; Treu 2010). Therefore, one is required to first resolve the images by using either a sufficiently high-resolution (ground-based or space) telescope or through spectroscopy. Then the individual light curves need to be monitored at sufficient resolution for several years in order to measure the time delays (e.g., Tewes et al. 2013; Liao et al. 2015; Millon et al. 2020). This can be difficult as these observations are expensive. On the other hand, unresolved light curves may be available "for free" from synoptic surveys. For example, we expect a lot of

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

lensed QSOs to be partially resolved or completely unresolved in the wide-field surveys, such as by the Zwicky Transient Facility (ZTF) (Bellm et al. 2019) and Legacy Survey of Space and Time (LSST) (LSST Science Collaboration et al. 2009, 2017), because their angular resolution is limited by seeing. In such cases, we can observe the joint light curve of the unresolved system that is a blend of the individual light curves. A robust method of detecting lensed QSOs through unresolved light curves can take advantage of the more abundant smaller telescopes. Thus, the importance of this approach cannot be overstated for boosting the sample size of the observed lensed systems.[4]

There are multiple other advantages in working with unresolved systems. Since there is no need for resolving the images a priori, this approach can be applied to the light-curve data from ongoing time domain surveys such as ZTF (Bellm et al. 2019), Pan-STARRS1 (Chambers et al. 2016). This will become more crucial when the upcoming Vera Rubin Observatory starts the LSST (LSST Science Collaboration et al. 2009, 2017). This approach also evades any degeneracy between a binary QSO pair and a doubly lensed QSO that creates confusion in lens detection using the resolved photometry (Mortlock et al. 1999; Peng et al. 1999).

Recently, multiple different techniques have been proposed to identify the lensed systems and to measure their time delays using the joint light curves (Shu et al. 2021; Springer & Ofek 2021a, 2021b; Biggio et al. 2022). The primary challenge in detecting the lensed cases using the joint light curves is that the intrinsic QSO light curves are highly stochastic and show a vast diversity in the flux variability. Therefore, any assumption on the flux variability can lead to biased results with low precision and high false-positive detection rate when the real light curves are not well described by the assumption. Therefore, it is extremely important to be model agnostic for achieving higher recall and precision as well as for reliable unbiased results.

In Paper I (Bag et al. 2022), we introduced a novel data-driven method that can detect the lensed QSOs and simultaneously measures the time delays only using the joint light curves, most importantly neither assuming anything about the quasar flux variability nor using any additional information. The technique is based on the empirical observation that the reconstructed image light curves corresponding to incorrect time delays exhibit more fluctuation than the ones reconstructed using the correct time delay. Although Bag et al. (2022) demonstrates that this approach is successful in the presence of significant noise (e.g., ZTF-like noise) and on existing data quality, it lacks a formal explanation as to how the minimization of the fluctuation works. This article looks deeper into the mathematical formalism of this approach. We attempt to understand the mathematical reasoning behind the empirical observations that laid the foundation of this approach. A clear insight into the mechanism should allow us to explore the strengths and possible limitations of this approach.

The paper is organized as follows. Section 2 recapitulates the method introduced by Bag et al. (2022). In Section 3, we provide a detailed mathematical explanation for the method's ability to detect lensed systems. We show that the lensing signal in the fluctuation curve is dominated by the autocorrelation function (ACF) of the derivative of the joint light curve. Section 4 explains why the ACF of the derivative of the joint light curve performs better in identifying lensed systems than the ACF of the joint light curve itself. Finally, in Section 5, we compare the performance of ACF of the derivative of the joint light curves against the full fluctuation curves. We conclude our findings in Section 6.

2. Reconstructing the Underlying Light Curves and Fluctuations in Them

The joint light curve of an unresolved lensed QSO having N_I images is the sum of the image light curves,

$$F(t) = \sum_{j=1}^{N_I} a_j \mathcal{F}(t - T_j), \tag{1}$$

where the individual image light curves can be described by a common function $\mathcal{F}(t)$ but with different magnifications (a_j) and time delays (T_j). To break the degeneracy among the images, we choose the ordering such that $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_{N_I}$ without loss of generality. For simplicity, let us first consider double systems (two images) with

$$F(t) = f(t) + \hat{\mu} f(t - \hat{\Delta}t), \text{ where } f(t) = a_1 \mathcal{F}(t - T_1)\ ,\ \hat{\mu} = a_2/a_1 \text{ and } \hat{\Delta}t = T_2 - T_1. \tag{2}$$

Here, $f(t)$ is the light curve of the brighter image, which we take as a reference. The other (fainter) image's light curve is then given by $\hat{\mu} f(t - \widehat{\Delta t})$, where $\hat{\mu} \leqslant 1$ and $\widehat{\Delta t}$ correspond to the magnification ratio and time delay, respectively, with respect to the reference (brighter) image. Note that $\widehat{\Delta t}$ can be positive or negative; a positive (negative) $\widehat{\Delta t}$ implies that the fainter image arrives later (earlier) in time than the brighter image.

It is difficult to model the quasar flux variability due to its highly stochastic nature. Instead, one can reconstruct the light curve of the brighter image following Bag et al. (2022) (see also Geiger & Schneider 1996) as

$$\begin{aligned} f_{\rm rec}(t;\, \mu,\, \Delta t) &= F(t) - \mu f_{\rm rec}(t - \Delta t;\, \mu,\, \Delta t), \\ &= F(t) - \mu F(t - \Delta t) + \mu^2 F(t - 2\Delta t) \\ &\quad - \mu^3 F(t - 3\Delta t) + \mu^4 F(t - 4\Delta t) - \ldots, \end{aligned} \tag{3}$$

$$= \sum_{n=0}^{\infty} (-\mu)^n F(t - n \cdot \Delta t), \tag{4}$$

from the joint light curve $F(t)$, given Δt and $\mu < 1$ which need not be equal to the true underlying values. In this article, we denote the true magnification ratio and time delay by $\{\hat{\mu},\ \widehat{\Delta t}\}$ to avoid confusion with the free parameters of the

[4] Although this work focuses on the unresolved lensed QSOs, similar exercises for the unresolved lensed supernovae have been pursued in the literature (Bag et al. 2021; Denissenya et al. 2022; Denissenya & Linder 2022).

reconstruction, $\{\mu,\ \Delta t\}$, which are the same as $\{\mu_{\rm try},\ \Delta t_{\rm try}\}$ in Paper I (Bag et al. 2022).

Note that the higher order terms in Equation (4) require $F(t)$ outside its observed range. Since the quasar flux cannot be predicted, we assume that $F(t)$ remains flat outside its observed range. As discussed in Geiger & Schneider (1996) and Bag et al. (2022), this convention has a negligible effect on the reconstruction except near either of the boundaries (depending upon the sign of Δt).

We emphasize that, by restricting ourselves to $\mu < 1$, which ensures the convergence of the sum in Equation (4), we are essentially reconstructing the light curve of the brighter image without any loss of generality. Moreover, for any choice of $\{\mu < 1,\ \Delta t\}$, Equation (4) gives a unique solution for the brightest image light curve; the corresponding fainter image light curve would be $\mu f_{\rm rec}(t-\Delta t;\mu,\ \Delta t)$. Only when $\mu = \widehat{\mu}$ and $\Delta t = \widehat{\Delta t}$, however, Equation (4) recovers the true image light curve, $f_{\rm rec}(t) = f(t)$. Still, the equation $F(t) = f_{\rm rec}(t;\mu,\ \Delta t) + \mu f_{\rm rec}(t-\Delta t;\mu,\ \Delta t)$ is exactly satisfied for any choice of $\{\mu < 1,\ \Delta t\}$ by construction. This demonstrates the mathematical degeneracy present in this lensed detection problem as discussed in Geiger & Schneider (1996) and Bag et al. (2022); without any prior assumptions on $f(t)$, any choice of $\{\mu,\ \Delta t\}$ can yield a (unique) lensing solution to the joint light curve $F(t)$.

In Paper I (Bag et al. 2022), we introduce a completely data-driven technique to break the degeneracy in the time delay by minimizing the fluctuation in reconstructed image light curves. To quantify the amount of fluctuation in a reconstructed image light curve $f_{\rm rec}(t;\mu,\ \Delta t)$, Bag et al. (2022) uses the simple metric,

$$\epsilon(\Delta t;\ \mu) \equiv \sum_{i=0}^{N_D-1} [f_{\rm rec}(t_{i+1};\ \mu,\ \Delta t) - f_{\rm rec}(t_i;\ \mu,\ \Delta t)]^2. \quad (5)$$

Fixing the trial magnification ratio (μ) to an arbitrary value (less than unity), one reconstructs the brighter image light curve for a number of Δt and finally looks for minima in the fluctuation curve, $\epsilon(\Delta t)$.

Figure 1 demonstrates how one can detect the unresolved lensed QSOs through the fluctuation curves by considering an example of a double (two-image) system that is simulated using the damped-random-walk (DRW) process in perfect conditions (marginal observational noise and 1 day of cadence). The true magnification ratio and time delay are set to $\widehat{\mu} = 0.86$ and $\widehat{\Delta t} = 22.0$ days. The three panels show $\epsilon(\Delta\ t)$ for different (arbitrary) choices of μ. We make the following observations from Figure 1.

1. The fluctuation curve $\epsilon(\Delta t)$ is highly symmetric with respect to $\Delta t = 0$.
2. There exists a global minimum at $\Delta t = 0$; the amount of fluctuations in the reconstructed light curve is minimized when we assume that the system is not lensed. This is a generic feature of all fluctuation curves irrespective of the system being lensed or not, and hence can be ignored[5].
3. Strikingly, when $\Delta t = \pm\widehat{\Delta t}$ (shown by the dashed vertical lines in each panel), we observe a pair of prominent secondary minima that correctly identify the system as lensed and simultaneously estimate the time delay accurately. On the other hand, if the system is unlensed (i.e., $\widehat{\mu} = 0$ or $\widehat{\Delta t} = 0$) we do not find any prominent secondary minima in $\epsilon(\Delta t)$ as demonstrated in Bag et al. (2022).
4. All of the observations above are somewhat insensitive to the choice of μ. Therefore, the true magnification ratio cannot be estimated in this approach. Nevertheless, one can accurately measure the time delays from the location of the pair of secondary minima in the fluctuation curves.

In the next section, we provide the mathematical reasoning behind all the above empirical observations.

3. Why Is Fluctuation Minimized for Correct Time Delay?

Substituting Equation (4) into Equation (5), one gets

$$\epsilon(\Delta t) = \sum_{i}^{N_D-1}\Bigg[\sum_{n}^{\infty}\{(-\mu)^n F(t_{i+1} - n\Delta t)\} - \sum_{m}^{\infty}\{(-\mu)^m F(t_i - m\Delta t)\}\Bigg]^2, \quad (6)$$

$$= \sum_{i}^{N_D-1}[\{F(t_{i+1}) - F(t_i)\} - \mu\{F(t_{i+1} - \Delta t) - F(t_i - \Delta t)\} + \mu^2\{F(t_{i+1} - 2\Delta t) - F(t_i - 2\Delta t)\} - \mu^3\{F(t_{i+1} - 3\Delta t) - F(t_i - 3\Delta t)\} + \ldots]^2. \quad (7)$$

For convenience, let us define the difference in the successive observed joint flux as a separate time series,

$$H(t_i) \equiv F(t_{i+1}) - F(t_i). \quad (8)$$

Using Equation (8), one can simplify Equation (7) to

$$\epsilon(\Delta t) = \sum_i [H(t_i) - \mu H(t_i - \Delta t) + \mu^2 H(t_i - 2\Delta t) - \mu^3 H(t_i - 3\Delta t) + \ldots]^2 = \sum_i \Bigg[\sum_n^{\infty}(-\mu)^n H(t_i - n\cdot\Delta t)\Bigg]^2 \quad (9)$$

Further expanding the squared term, we get

$$\epsilon(\Delta t) = \sum_i H(t_i)^2 - 2\mu\sum_i H(t_i)H(t_i - \Delta t) + \mu^2 \times \Bigg[\sum_i H(t_i - \Delta t)^2 + 2\sum_i H(t_i)H(t_i - 2\Delta t)\Bigg] - 2\mu^3\Bigg[\sum_i H(t_i)H(t_i - 3\Delta t) + \sum_i H(t_i - \Delta t)H(t_i - 2\Delta t)\Bigg] + \ldots. \quad (10)$$

Thus we can express $\epsilon(\Delta t)$ as a power series in μ,

$$\epsilon(\Delta t) = \epsilon_0 + \mu\epsilon_1(\Delta t) + \mu^2\epsilon_2(\Delta t) + \mu^3\epsilon_3(\Delta t) + \mu^4\epsilon_4(\Delta t) + \ldots, \quad (11)$$

[5] While trying to detect the lensed QSOs through the minimal fluctuation, as proposed by Bag et al. (2022), this unlensed solution always remains to be a feasible lensing solution to the unresolved problem in the absence of any assumption on the flux variability.

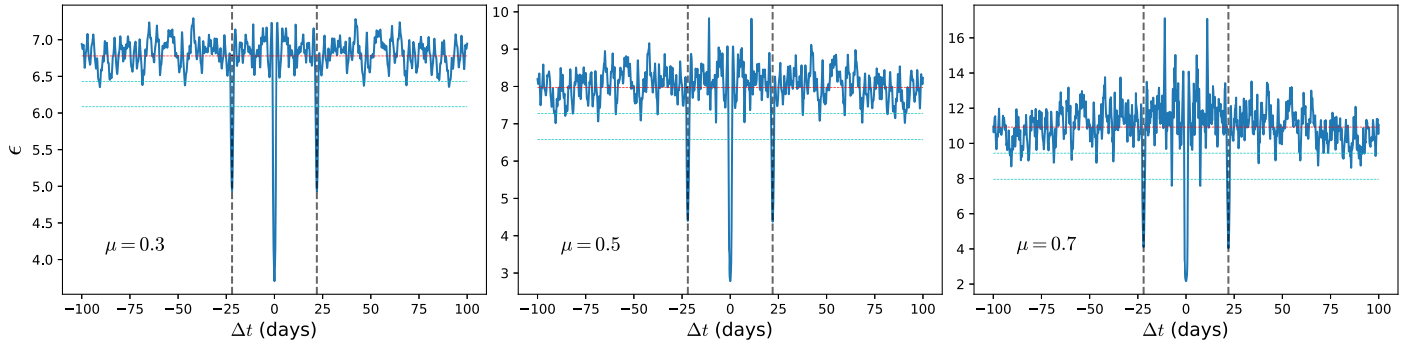


Figure 1. The fluctuation curves for an example of a double system simulated (in perfect condition with negligible observation noise) using damped-random-walk (DRW) process with time delay $\widehat{\Delta t} = 22.0$ days and the magnification ratio $\widehat{\mu} = 0.86$. The amount of fluctuation (ϵ calculated from Equation (5)) in the reconstructed image light curves is plotted against the trial Δt. The three panels correspond to three arbitrary choices of trial magnification ratio: $\mu = 0.3$, 0.5, 0.7. The dashed vertical lines in each plot mark the true time delay, $\Delta t = \pm\widehat{\Delta t}$, where we find prominent pairs of secondary minima that can be used to detect the lensed system and measure the time delay. This figure is qualitatively similar to Figure 2 of Bag et al. (2022).

where

$$\epsilon_0 = \sum_i H(t_i)^2, \tag{12}$$

$$\epsilon_1(\Delta t) = -2\sum_i H(t_i)H(t_i - \Delta t), \tag{13}$$

$$\epsilon_2(\Delta t) = \sum_i H(t_i - \Delta t)^2 + 2\sum_i H(t_i)H(t_i - 2\Delta t), \tag{14}$$

$$\epsilon_3(\Delta t) = -2\sum_i H(t_i)H(t_i - 3\Delta t) - 2\sum_i H(t_i - \Delta t)H(t_i - 2\Delta t), \tag{15}$$

and so on. Therefore, the fluctuation curve can be written in the following closed form,

$$\epsilon(\Delta t) = \sum_{n=0}^{\infty}(-\mu)^n \sum_{m=0}^{\lfloor n/2\rfloor}(2 - \delta_{m,n-m}) \times \sum_i [H(t_i - m\Delta t)H(t_i - (n-m)\Delta t)], \tag{16}$$

where the *floor function* $\lfloor n/2\rfloor$ returns the highest integer equal to or below $n/2$. Since $\mu < 1$, the contribution from the higher order terms gets suppressed rapidly.

Let us now take a closer look at the difference series $H(t)$ (defined in Equation (8)), which can be recast as

$$H(t_i) = [f(t_{i+1}) - f(t_i)] + \widehat{\mu}[f(t_{i+1} - \widehat{\Delta t}) - f(t_i - \widehat{\Delta t})], \tag{17}$$

$$= h(t_i) + \widehat{\mu}\, h(t_i - \widehat{\Delta t}), \tag{18}$$

where

$$h(t_i) \equiv f(t_{i+1}) - f(t_i) \tag{19}$$

can be regarded as another time series. For uniformly sampled data, $H(t)$ and $h(t)$ are proportional to the derivative of $F(t)$ and $f(t)$, respectively. In this article, we refer to $H(t)$ and $h(t)$ as the derivatives of $F(t)$ and $f(t)$, respectively. However, for nonuniformly sampled data, the former two are just difference series and the mathematical arguments remain intact. From Equation (18), note that $H(t)$ follows a similar lensing equation as $F(t)$ in Equation (2) but with $h(t)$ as the underlying time series.

The QSO intrinsic light curve, $f(t)$, follows a stochastic process that is "wide-sense stationary"; the mean and covariance properties remain constant over time. It is easy to check that the time series $F(t)$, $h(t)$, and $H(t)$ are also wide-sense stationary. For a generic wide-sense stationary time series $X(t)$,

$$\langle X(t)\rangle_t \equiv \frac{1}{N_D}\sum_i X(t_i) \approx \langle X(t + t_c)\rangle_t, \tag{20}$$

$$\sigma^2_{X(t)} \equiv \langle X(t)^2\rangle_t - \langle X(t)\rangle_t^2 \approx \langle X(t + t_c)^2\rangle_t - \langle X(t + t_c)\rangle_t^2 \equiv \sigma^2_{X(t+t_c)} \tag{21}$$

and

$$\langle X(t + t_{c_1})X(t + t_{c_2})\rangle_t \approx \langle X(t)X(t + [t_{c_2} - t_{c_1}])\rangle_t \approx \text{function of } |t_{c_2} - t_{c_1}| \text{ only}, \tag{22}$$

for any shift in time much smaller than the observed time range ($t_c,\ t_{c_1},\ t_{c_2} \ll t_{\rm range}$). Therefore,

$$\langle H(t)\rangle_t = \langle F(t_{i+1})\rangle_t - \langle F(t_i)\rangle_t \approx 0 \tag{23}$$

since we assume that the statistical properties of the joint light curve remain invariant under translations, i.e., $\langle F(t_{i+1})\rangle_t \approx \langle F(t_i)\rangle_t$.

Now that we have assembled all the necessary tools, we proceed to explain the characteristics of the fluctuation curve that are described in Section 2 and illustrated in Figure 1.

1. First, notice that Equation (10) is invariant under $\Delta t \to -\Delta t$ as long as $|\Delta t| \ll t_{\rm range}$, since $H(t)$ also follows Equations (20)–(22). This explains the symmetry in the $\epsilon(\Delta t)$ curve with respect to $\Delta t = 0$.
2. The first term in Equation (10) or (11), ϵ_0, is constant (independent of Δt) and hence can be ignored. The next leading order term $\epsilon_1(\Delta t)$ is proportional to the correlation coefficient between $H(t)$ and $H(t - \Delta t)$ defined as (also known as the autocorrelation function (ACF), see Equation (A1))

$$\mathrm{ACF}(H;\ \Delta t) = \frac{\sum_i [H(t) - \langle H(t)\rangle_t][H(t - \Delta t) - \langle H(t - \Delta t)\rangle_t]}{\sqrt{\sum_i [H(t) - \langle H(t)\rangle_t]^2 \cdot \sum_i [H(t - \Delta t) - \langle H(t - \Delta t)\rangle_t]^2}} \approx \frac{1}{\epsilon_0}\sum_i H(t_i)H(t_i - \Delta t), \tag{24}$$

where one can identify that $\langle H(t)\rangle_t \approx \langle H(t-\Delta t)\rangle_t \approx 0$ and the denominator simply reduces to $\sum_i H(t_i)^2 = \epsilon_0$. Therefore, Equation (10) can be approximated as

$$\epsilon(\Delta t) \approx \epsilon_0[1 - 2\mu \mathrm{ACF}(H; \Delta t)] + \mathcal{O}(\mu^2). \quad (25)$$

3. The autocorrelation function, ACF(H; Δt), is maximized to unity at $\Delta t = 0$. This in turn minimizes the $\epsilon_1(\Delta t)$ term that dominates the Δt dependence of the fluctuation curve in Equation (11) at the leading order in μ.[6] Therefore, one always observes a global minimum in the $\epsilon(\Delta t)$ curve at $\Delta t = 0$ irrespective of the system being lensed or not: $\epsilon(0) \approx \epsilon_0 - 2\mu\epsilon_0 + \mathcal{O}(\mu^2)$. In fact, it is easy to calculate the exact value $\epsilon(0) = \epsilon_0/(1+\mu)^2$ as given in Appendix B.1.
4. When the system is lensed, ACF(H; Δt) is also maximized locally at $\Delta t = \pm\widehat{\Delta t}$ as the correlation function picks up excess power due to the matching of two sets of the same intrinsic features that are separated by the true time delay in the joint light curve (see Appendix A for detailed derivations). Thus the fluctuation curve, $\epsilon(\Delta t)$, shows a pair of secondary minima at $\Delta t = \pm\widehat{\Delta t}$ following Equation (25). When $h(t)$ is pure white noise, $\mathrm{ACF}(H; \Delta t = \pm\widehat{\Delta t}) \approx \hat{\mu}/(1+\hat{\mu}^2)$ as evident from Equation (A6) for long observation ranges, the exact height of the secondary minima in $\epsilon(\Delta t)$ is calculated in Appendix B.2 under this approximation. On the other hand, when the system is unlensed ($\hat{\mu} = 0$), we get no such secondary minima. Therefore, by detecting this pair of secondary minima, which are the dominant part of the *lensing features* in the fluctuation curve, one can identify the system as lensed. Simultaneously, the location of this minima pair allows us to estimate the time delay of the system.
5. As the ϵ_i terms in Equation (11), which contain the Δt variable, are independent of the choice of μ, the locations of the lensing features (local minima and maxima due to lensing) in the fluctuation curve $\epsilon(\Delta t)$ is very much insensitive to μ. (However, note that, for a higher value of μ, higher order terms in Equation (11) also contribute, which makes the curve fluctuate more as evident from the three panels of Figure 1.) Therefore, we cannot determine the true magnification ratio in this approach, although the time delay can be recovered accurately.[7]

Figure 2 demonstrates the above arguments for the same example considered in Figure 1. The top-left panel shows the joint light curve (in perfect conditions with negligible noise and one day of cadence) of the doubly lensed system simulated using the damped-random-walk (DRW) template with $\widehat{\Delta t} = 22.0$ days and $\hat{\mu} = 0.86$. The derivative of the joint light curve, $H(t)$, is shown in the bottom-left panel. The top-right panel displays the auto-correlation function of $H(t)$ where we can clearly find the pair of secondary maxima at $\Delta t = \pm\widehat{\Delta t} = \pm 22.0$ days. Finally, we compare the full fluctuation curve ($\epsilon(\Delta t)$) with its main contributing term (linear in μ) stemming from ACF(H; Δt) for an arbitrary trial $\mu = 0.3$. We notice that the features in the fluctuation curve are dominated by the ACF(H; Δt), which is also responsible for the secondary minima in $\epsilon(\Delta t)$appearing at $\Delta t = \pm\widehat{\Delta t}$. The system is therefore identified as lensed. Note that the autocorrelation of the joint light curve itself, i.e., ACF(F; Δt), does not show the lensing peaks in this case; hence it is not reliable for lens detection as explained in Section 4 in more detail.

3.1. Contribution from the Higher Order Terms

For an arbitrarily long time series ($t_{\mathrm{range}} \to \infty$), one can keep substituting

$$\sum_i^{N_D} H(t+t_c)H(t+t_c-n\Delta t) \approx \mathrm{ACF}(H; n\Delta t)\cdot\epsilon_0, \quad (27)$$

for any n ignoring the boundary effect. Here t_c is any constant shift in time and ϵ_0 is given by (12). Therefore, Equations (13)–(16) can be simplified as

$$\begin{aligned}
\epsilon_1(\Delta t) &= -\epsilon_0[2\mathrm{ACF}(H; \Delta t)],\\
\epsilon_2(\Delta t) &= \epsilon_0[2\mathrm{ACF}(H; 2\Delta t) + 1],\\
\epsilon_3(\Delta t) &= -\epsilon_0[2\mathrm{ACF}(H; 3\Delta t) + 2\mathrm{ACF}(H; \Delta t)],\\
\epsilon_4(\Delta t) &= \epsilon_0[2\mathrm{ACF}(H; 4\Delta t) + 2\mathrm{ACF}(H; 2\Delta t) + 1],\\
&\ldots
\end{aligned} \quad (28)$$

$$\epsilon_n(\Delta t) = \begin{cases} (-1)^n\epsilon_0[2\mathrm{ACF}(H; n\Delta t) + 2\mathrm{ACF}(H; (n-2)\Delta t) + \ldots + 2\mathrm{ACF}(H; \Delta t)] & \text{if } n \text{ is odd.} \\ (-1)^n\epsilon_0[2\mathrm{ACF}(H; n\Delta t) + 2\mathrm{ACF}(H; (n-2)\Delta t) + \ldots + 2\mathrm{ACF}(H; 2\Delta t) + 1] & \text{if } n \text{ is even.} \end{cases} \quad (29)$$

Note that ACF(H; $n\Delta t$) exhibits a pair of peaks at $\Delta t = \pm\widehat{\Delta t}/n$. Thus we get pairs of local minima (maxima) in the fluctuation curve at $\pm\widehat{\Delta t}/n$ for each odd (even) n from Equation (11). We refer to these features as the *lensing signal* in the fluctuation curve. Remarkably, from each odd term, we get a pair of minima at $\Delta t = \pm\widehat{\Delta t}$ which, despite being suppressed by a factor of μ^n, enhances our target pair of secondary minima and increases its detectability. On the other hand, we find local maxima in the fluctuation curve at $\Delta t = \pm\widehat{\Delta t}/2$, $\widehat{\Delta t}/4$,… from the even terms. Therefore, as more and more terms contribute to the $\epsilon(\Delta t)$ curve, we get slight enhancement in the lensing minima (as compared to the lensing peaks in ACF(H; Δt)) but at the expense of more extrema that increase the overall oscillations in the fluctuation curve. Note that, for a finite time series, we still expect similar higher order features in the fluctuation curve up to the order $n \ll t_{\mathrm{range}}/|\widehat{\Delta t}|$.

[6] This could also be obtained from the inequality,

$$\epsilon_1(\Delta t) = -2\sum_i H(t_i)H(t_i-\Delta t) \geqslant -\sum_i H(t_i)^2 - \sum_i H(t_i-\Delta t)^2 \approx -2\sum_i H(t_i)^2 = -2\epsilon_0 \quad (26)$$

following the fact that $2xy \leqslant x^2 + y^2$. The equality holds when $H(t_i) = H(t_i - \Delta t)$ for all i requiring $\Delta t = 0$ unless $H(t)$ is a constant function. Therefore, $\epsilon_1(\Delta t)$ is minimized at $\Delta t = 0$. Note that this argument does not require $\langle H(t)\rangle_t \approx 0$.

[7] In an ideal condition where this is uncorrelated in time, the observation range is long, data have negligible noise, etc., one can in principle determine the magnification ratio from the height of the secondary minima at $\pm\widehat{\Delta t}$ using Equation (B5) (or from the height of the secondary peaks at ACF(H; $\pm\widehat{\Delta t}$)). This will not be reliable for the realistic QSO light curves with unknown correlation and especially in the presence of significant observational noise.

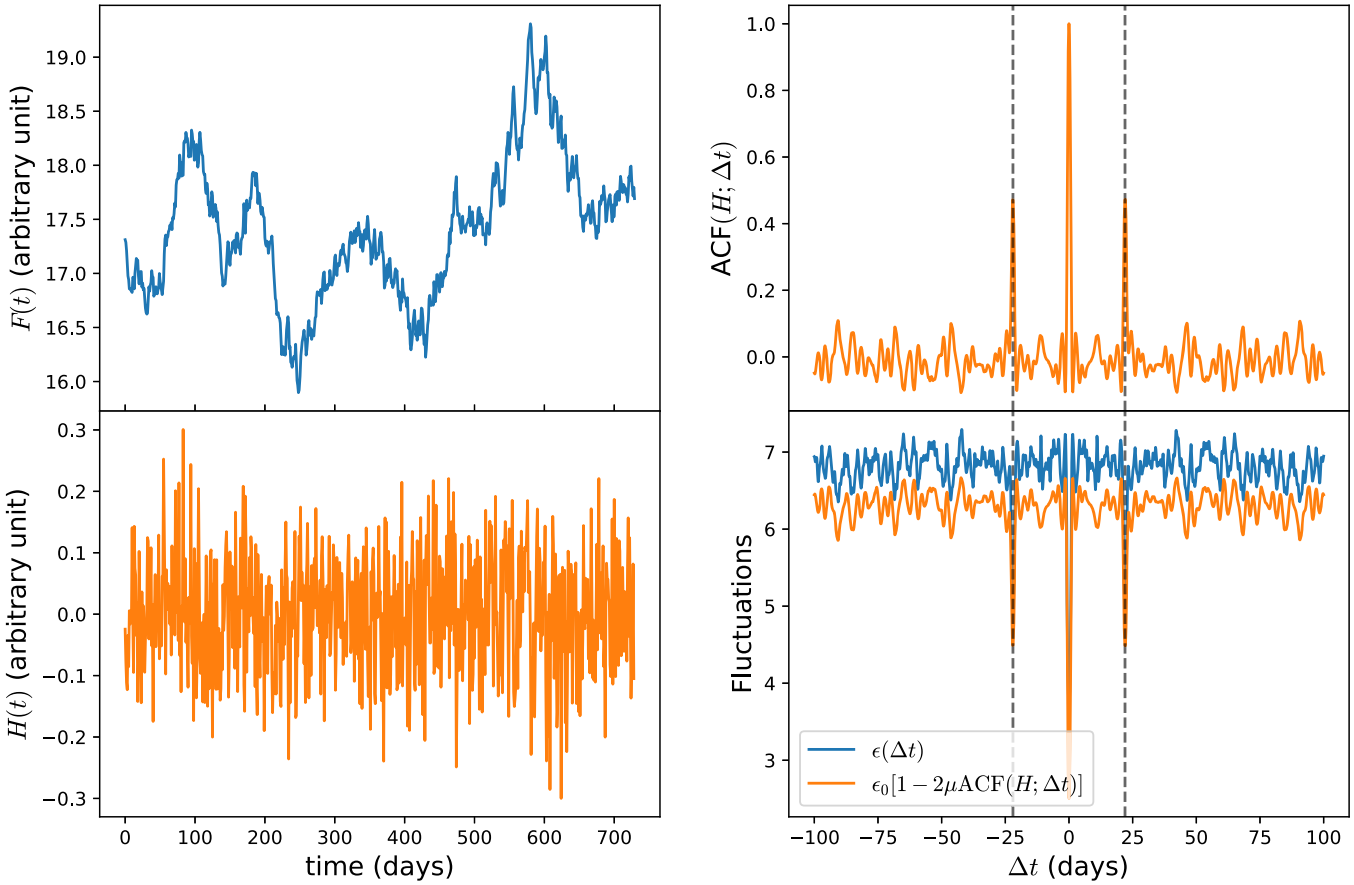


Figure 2. The top-left panel shows the joint light curve (in perfect condition with negligible observation noise) of the double system considered in Figure 1: $\widehat{\Delta t} = 22.0$ days and $\widehat{\mu} = 0.86$. The difference series for this system, $H(t)$ as defined in Equation (8) (which is simply the derivative of $F(t)$ when it is sampled uniformly), is shown in the bottom-left panel. Its autocorrelation function, ACF(H; Δt) shown in the top-right panel, exhibits a prominent and sharp pair of secondary maxima at $\Delta t = \pm\widehat{\Delta t}$ through which one can easily detect the lensed system. Finally, the bottom-right panel compares the full fluctuation curve ($\epsilon(\Delta t)$, same as in the left panel of Figure 1) using the trial $\mu = 0.3$ with the contribution from the ACF(H; Δt) to it.

3.2. Quad Systems

It is possible to generalize the image reconstruction, Equation (4), for more than two image systems. However, for a generic system with N_I images, ACF(H; Δt) exhibits $N_I(N_I-1)/2$ pairs of secondary maxima, one for each relative time delay $|T_i - T_j|$ ($i \neq j$). Note that, if the time delays between two pairs of images are too close to each other, or more explicitly, if the difference between two time delays is smaller than the observation cadence or the smoothing timescale (used to deal with noisy data), then the two corresponding peaks will merge into one in ACF(H; Δt). For example, ACF (H; Δt) for a quad system (four images blended together in reality) can show up to six pairs of secondary maxima at the time delays $\Delta t = \pm\ \Delta t_{i,j\neq i}$ given that $\Delta t_{i,j\neq i}$ are well separated from each other. As ACF(H; Δt) dominates the lensing signal in the fluctuation curve, the latter also shows the same number of secondary minima pairs in this case even if one assumes only two images in the reconstruction analysis, i.e., following Equation (4) with two images only. Thus, one can identify the quad systems using the method presented in Paper I (Bag et al. 2022) by detecting multiple (up to six) pairs of minima in the $\epsilon(\Delta t)$ curve.

This has been illustrated in Figure 3 where we consider an example of a quad unresolved lensed system simulated using DRW process with time delays: $\widehat{\Delta t} = 11.0$, 31.0, 68.0 days with respect to the first image (again in perfect condition). The top and bottom panels show the autocorrelation function for the derivative, ACF (H; Δt) and the fluctuation curve $\epsilon(\Delta t)$ (using two-image analysis), respectively. In both panels, we find six pairs of prominent extrema at $\Delta t = \pm 11$, ± 31, ± 68, ± 20, ± 57, ± 37 days. Therefore, one can detect the quad systems using both approaches. In a blind analysis, if we detect one pair of secondary minima in the $\epsilon(\Delta t)$ (or secondary maxima in ACF(H; Δt)) curve, we can identify the system as a doubly lensed QSO. On the other hand, if we find multiple such pairs (up to six), we can detect it as a quad system since three-image systems do not exist in reality.

Thus, in this section, we explain the mathematical origin of the characteristics of the fluctuation curves that play the pivotal role in developing the method introduced in Paper I (Bag et al. 2022). However, we are left with two additional questions: (i) why can one detect the lensed QSOs more reliably using ACF (H; Δt) as compared to ACF(F; Δt), and (ii) what are the advantages of using the fluctuation curve over the simpler ACF (H; Δt) in detecting lensed QSOs? These questions are addressed below in Sections 4 and 5, respectively.

4. ACF(F; Δt) versus ACF(H; Δt)

One can in principle reliably detect the lensed systems using the autocorrelation function of the joint light curve, ACF(F; Δt), if the intrinsic light curve $f(t)$ is uncorrelated in time (white noise) as explained in Appendix A. However, the QSO light curves can have long timescale correlations that violate Equation (A4). In this case, the existence of the lensing peaks

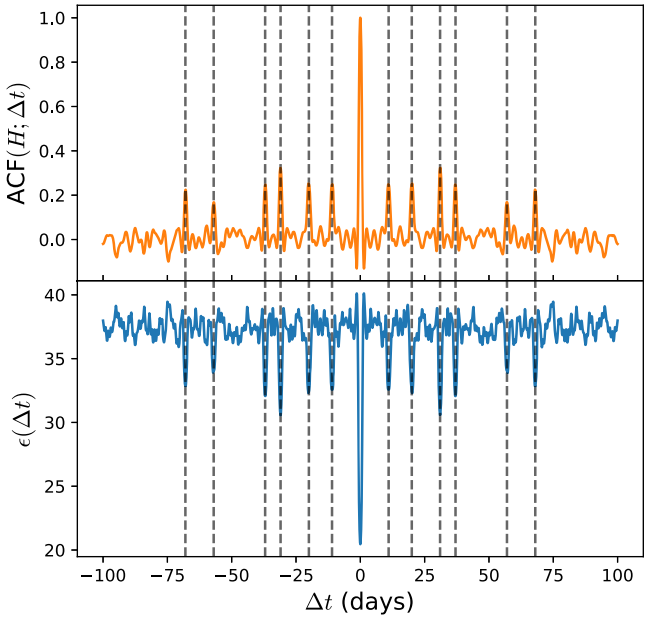


Figure 3. The top panel shows the autocorrelation function of the derivative of the joint light curve ($H(t)$, defined in Equation (8)) for a quad system (four images in reality). The bottom panel shows the fluctuation curve for the same system, analyzed assuming just two images. The six dashed vertical lines on either side of $\Delta t = 0$ in each panel represent the six relative time delays between the four actual images.

in ACF(F; Δt) depends on the characteristics of the autocorrelation function of the intrinsic light curve, ACF(f; Δt). Even when Equation (A4) is not strictly valid, we can expect excess power in ACF(F; Δt) at $\Delta t = \pm\widehat{\Delta t}$ from Equation (A3) if ACF(f; Δt) decays reasonably sharply (from its peak at $\Delta t = 0$), i.e., if $f(t)$ and $f(t \pm \Delta t)$ are uncorrelated for $|\Delta t| \gtrsim |\widehat{\Delta t}|$. In other words, one can detect the lensing peaks in ACF(F; Δt) only if ACF(f; Δt)is narrowly peaked around $\Delta t = 0$ even in the ideal condition with negligible observation noise. This is discussed in Appendix A.1 in detail (see Equation (A7) for the precise condition). Since $F(t)$ and its derivative $H(t)$ follow the same lensing equation (compare Equations (2) and (18)), the above criterion is also applicable to ACF(h; Δt) in order to find the lensing peaks in ACF(H; Δt).

However, QSO flux variability typically shows temporal correlation until a few hundred to even one thousand days (Kelly et al. 2009; MacLeod et al. 2010). The expectation value of the autocorrelation function of light curves generated using damped-random-walk (DRW) process (Kelly et al. 2009; MacLeod et al. 2010; Zu et al. 2013) decays exponentially,

$$\langle \mathrm{ACF}(f;\ \Delta t)\rangle_E = \exp(-|\Delta t|/\tau), \tag{30}$$

but not sharply, since the decay timescale, τ, is typically $\mathcal{O}(10^2$–$10^3)$ days. Here, $\langle\cdot\rangle_E$ denotes an ensemble average over all possible realizations of $f(t)$ with the same stochastic properties.

Therefore, it can be difficult to detect the secondary peaks in ACF(F; Δt), and one can also have a significant number of false-positive detections (Geiger & Schneider 1996; Shu et al. 2021). This has been demonstrated in Figure 4 using numerical simulations. Here we simulate 1000 realizations of the intrinsic QSO light curves ($f(t)$) using the DRW template. We fix the correlation timescale to $\tau = 10^{2.5} \approx 316$ days, which is consistent with the findings of Kelly et al. (2009) and MacLeod et al. (2010) (see also Dobler et al. (2015)) throughout the paper for our illustration purpose. Then we construct the joint light curves ($F(t)$) separately for each realization following Equation (2) with $\widehat{\mu} = 0.86$ and $\widehat{\Delta t} = 50$ days, kept the same across the realizations. For simplicity, we consider the perfect condition with marginal noise in the data. The solid curves in the top-left panel show ACF(f; Δt) for five randomly selected samples, whereas the dashed black curve and the shaded region represent the ensemble average of ACF(f; Δt) and the 68% quantile, respectively.[8] It is evident that there exists significant correlation in $f(t)$ until a few hundred days as the ACF(f; Δt) curves decay slowly with $|\Delta t|$. Also, notice that the 68% quantile region of ACF(f; Δt) expands with $|\Delta t|$, thus some realizations of ACF(f; Δt) can exhibit local maxima as in the case for the red and purple curves in the top-left panel. Thus, using ACF(F; Δt) one can get a substantial number of false lensed detections in true unlensed cases because of these maxima.

The ACF(F; Δt) for the same five realizations have been shown in the bottom-left panel, the dashed vertical lines represent the true time delay, $\Delta t = \pm 50$ days, for these systems. Although for some realizations (purple curve), one can find excess power in ACF (F; Δt) at $\Delta t = \pm\widehat{\Delta t}$, for others (the green, red, and blue curves), this is not true.

On the other hand, since DRW behaves like a random walk at small timescales ($t \ll \tau$), its derivative $h(t)$ behaves like white noise obeying Equation (A4) at this limit. This is evident from the top-right panel of Figure 4 where we show ACF(h; Δt) for the same five realizations with different colors. For all realizations, we find that ACF(h; Δt) becomes very small for $\Delta t \neq 0$. In fact the ensemble average, shown by the dashed black curve, follows $\langle \mathrm{ACF}(h;\ \Delta t)\rangle \approx \delta_{\Delta t,0}$. Hence, we get a prominent pair of secondary maxima in ACF(H; Δt) at $\Delta t = \pm\widehat{\Delta t}$ for all the realizations as evident from the bottom-right panel.

Last, let us consider time series with different correlation timescales τ in Equation (30), even if they do not describe the QSO light curves accurately. A smaller (larger) τ leads to a narrower (broader) peak in $\langle \mathrm{ACF}(f;\ \Delta t)\rangle$, which in turn increases (decreases) the probability of ACF(F; Δt) showing the lensing peaks, according to Appendix A.1. However, in both limits, ACF(h; Δt) has a sharp peak at $\Delta t = 0$, so that one can always find the lensing peaks in ACF(H; Δt).

This exercise using the DRW template thus demonstrates that ACF(H; Δt) outperforms ACF(F; Δt) in terms of detectability of the lensing peaks. However, this conclusion is not restricted to DRW and unbound random-walk templates. It stands valid for any "red-type" power spectrum, as argued below in Section 4.1 and in Appendix C more explicitly.

4.1. Connection to the Power Spectrum

Let us define the power spectrum, $P_f(\omega)$, of a time series as the two-point correlation function in the Fourier space;

$$\langle \tilde{f}(\omega)\tilde{f}(\omega')^*\rangle_E = P_f(\omega)\delta(\omega - \omega'), \tag{31}$$

where $\tilde{f}(\omega)$ is the Fourier transform of the time series $f(t)$. We assume that the Fourier modes are uncorrelated. To be precise,

[8] Equation (30) is true only if the observation range is much larger than correlation scale, i.e., $t_{\mathrm{range}} \gg \tau$ (τ has been set to $10^{2.5}$ days for these simulations) so that ergodicity is observed. So the black dashed curve in the top-left panel of Figure 4 coincides with Equation (30) for much longer observation range.

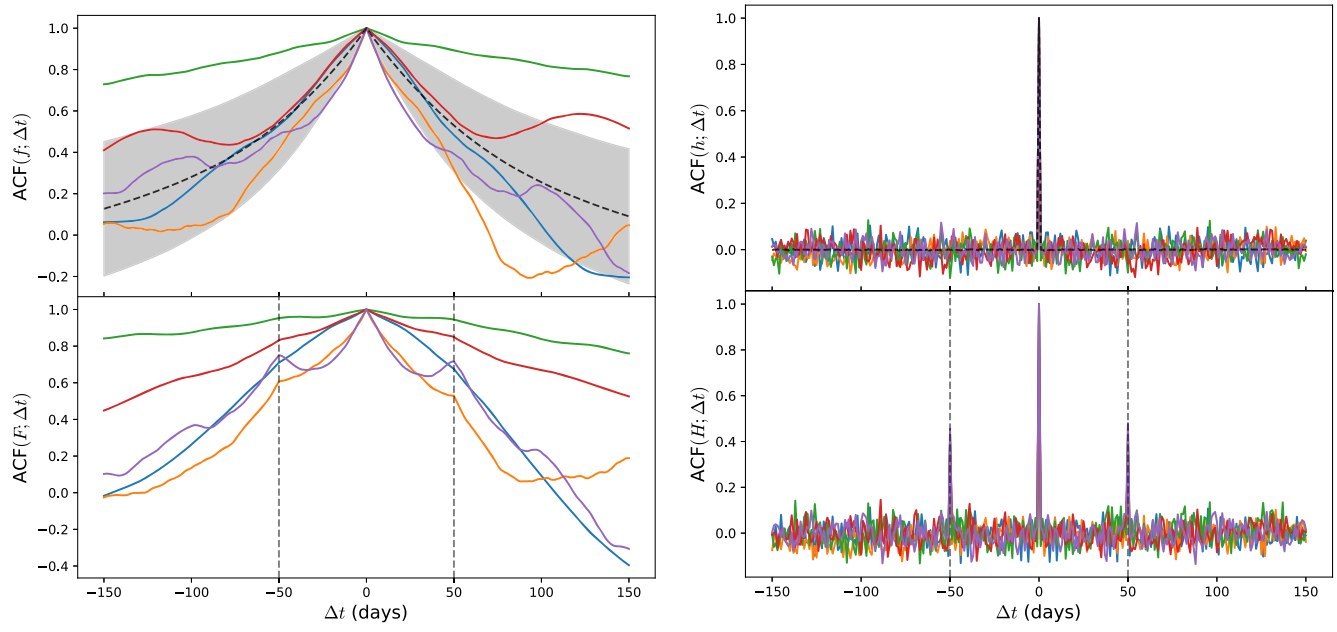


Figure 4. The top-left panel shows autocorrelation in the intrinsic QSO light curve, ACF(f; Δt), for five out of 1000 realizations (randomly chosen) simulated using the DRW template. The expectation value (ensemble average) and 68% percentile are marked by the dashed black curve and the shaded region, respectively. The top-right panel shows ACF(h; Δt) for the same five realizations along with the expected value (black dashed curve). The bottom two panels compare ACF(F; Δt) and ACF(H; Δt) for the double systems constructed using these five realizations. The dashed vertical lines in the bottom two panels represent the true time delay: $\Delta t = \pm 50$ days for these systems. Since ACF(h; Δt) is typically much narrower than ACF(f; Δt), we expect to find the lensing peaks in ACF(H; Δt) more reliably than in ACF(F; Δt), as evident from the comparison of the bottom two panels.

we assume that the QSO intrinsic light curve $f(t)$ is "wide-sense stationary" as its mean and covariance properties do not vary over time (see (20)–(22)). Under these assumptions, the Wiener–Khinchin theorem states that the (expected) autocorrelation function of $f(t)$ is given by the Fourier transform of the power spectrum (Einstein 1914; Wiener 1930; Khintchine 1934). Therefore, one can determine the autocorrelation function of a time series by studying its power spectrum. This is especially useful since a derivative in the time domain corresponds to a multiplication by $i\omega$ in the Fourier domain, and hence $P_h(\omega) = \omega^2 P_f(\omega)$.

The simplest example is when the underlying signal consists purely of white noise $f_{\rm white}(t)$ with the power spectrum given by $P_{f_{\rm white}}(\omega) = \sigma^2_{\rm white} = \text{constant}$. Taking the Fourier transform yields the expected value of the autocorrelation function ACF ($f_{\rm white}$; Δt) $\propto \delta(\Delta t)$. Since this is sharply peaked around $\Delta t = 0$, we can accurately retrieve the lensing peaks in the ACF(F; Δt) curve. Further details for this white-noise scenario are discussed in Appendix A.

For the damped-random-walk (DRW) templates often used for describing the intrinsic QSO light curves, the power spectrum takes the form

$$P_{f_{\rm DRW}}(\omega) \propto \tau/(1 + \tau^2\omega^2), \tag{32}$$

which is the Fourier transform of Equation (30) up to a normalization factor. Note that power spectra like $P_{f_{\rm DRW}}(\omega)$ that decay with $|\omega|$ give rise to red noise and hence can be classified as "red-type" power spectra. The decay timescale τ parameterizes how long it takes for the time series to forget the fluctuations that happened in the past. Since $P_{f_{\rm DRW}}(\omega) \approx$ constant for $\tau^2\omega^2 \ll 1$, the DRW is mostly independent of its past and behaves like white noise when τ is small.

To describe realistic quasar light curves, τ should typically be of the order 10^2–10^3 days. The DRW then behaves more similarly to an (undamped) Gaussian random walk with a power spectrum $P_f(\omega) \sim \omega^{-2}$ up to a timescale smaller than τ, or $\omega^2 \gg 1/\tau^2$. The Fourier transform of such a power spectrum has a broad peak at zero unlike the sharp delta-function-like peak of the white-noise one. Therefore, ACF(F; Δt) is less reliable for finding lensed cases in this case.

Meanwhile, the derivative time series $h(t)$ has a much flatter power spectrum with $P_h(\omega) \propto \omega^2/(1 + \tau^2\omega^2)$. This is nearly constant for large ω; $h(t)$ behaves like white noise. Thus, ACF (h; Δt) tends to be much narrower and is expected to decay down quickly as compared to ACF(f; Δt). This is consistent with our previous findings that ACF(H; Δt) is more reliable than ACF(F; Δt) for detecting the lensed systems. Indeed, a larger τ makes the $P_f(\omega)$ steeper (or redder) in Equation (32) and leads to a flatter $\langle$ACF(f; Δt)$\rangle$ in Equation (30). This in turn reduces the possibility of the lensing peaks appearing in ACF(F; Δt) and thus enables ACF(H; Δt) to outperform ACF (F; Δt) even to a greater extent. In fact, this is true for all red-type power spectra ($P_f(\omega)$ decreases with increasing $|\omega|$) as illustrated in Appendix C. Since the lensing features in the fluctuation curve $\epsilon(\Delta t)$ predominantly arise from ACF(H; Δt), we expect similar performance from both of these approaches.

5. Fluctuation in Reconstruction versus Autocorrelation of Derivative

Since the lensing signal in the fluctuation curve is dominated by the autocorrelation of the derivative of the joint light curve ($H(t)$ defined in Equation (8)), it is interesting to test if ACF(H;

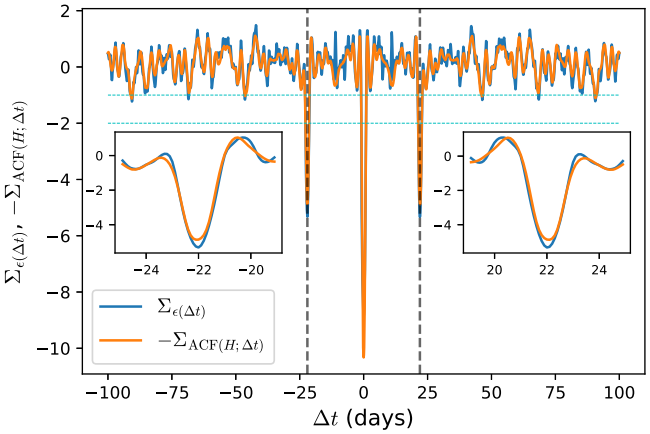


Figure 5. The fluctuation curve and ACF(H; Δt) are compared after being normalized using Equation (33) for the same system as in Figure 1 (i.e., simulated with DRW template and with $\widehat{\Delta t} = 22.0$ days and $\widehat{\mu} = 0.86$). We find secondary extrema of similar strengths in both curves as the ACF(H; Δt) dominates the lensing signal in the fluctuation curve. However, the latter still has more lensing signal due to contributions from the higher order terms in Equation (11) as illustrated by the two inset plots which zoom into the two lensing/secondary minima.

Δt) can similarly be used to detect the lensed QSOs. Note that ACF(H; Δt) would be less computationally expensive and has a better physical interpretation.

Let us begin by comparing the lensing signals (the prominence of the secondary maxima that can be used for lensed detection) in ACF(H; Δt) and the fluctuation curve ($\epsilon(\Delta t)$) for the example presented in Figures 1 (left panel) and 2. We normalize both curves using the generic transformation,

$$\Sigma_{X(\Delta t)} \equiv \frac{X(\Delta t) - \langle X(\Delta t)\rangle_{\Delta t}}{\sigma_{X(\Delta t)}}, \tag{33}$$

where $\langle X(\Delta t)\rangle_{\Delta t}$ and $\sigma_{X(\Delta t)}$ are the average and standard deviation taken over Δt of a one-dimensional function $X(\Delta t)$. (Thus $\Sigma_{X(\Delta t)}$ simply measures $X(\Delta t)$ in the units of its standard deviation.) Figure 5 compares $\Sigma_{\epsilon(\Delta t)}$ (with trial $\mu = 0.3$) and $-\Sigma_{\mathrm{ACF}(H;\ \Delta t)}$ for that example system (simulated using DRW template with $\widehat{\Delta t} = 22.0$ days, $\widehat{\mu} = 0.86$ and marginal observation noise). As expected, both curves show a prominent pair of secondary minima at $\Delta t = \pm\widehat{\Delta t} = \pm 22.0$ days (marked by dashed vertical lines) deeper than $\Sigma = -2$. However, we still find that the secondary minima in the $\Sigma_{\epsilon(\Delta t)}$ are slightly deeper than that of the $-\Sigma_{\mathrm{ACF}(H;\ \Delta t)}$ curve. This indicates that $\epsilon(\Delta t)$ contains a more enhanced lensing signal than ACF(H; Δt) in this case.

Next, we test the approach based on ACF(H; Δt) on the same validation and blind sets used in Bag et al. (2022) (simulated using the DRW template). As usual, we first consider the perfect conditions where the observational noise is marginal compared to the time variation in the intrinsic QSO light curves. Using the same selection criteria, ACF(H; Δt) can detect 19 out of 20 true lensed cases correctly. Among the 20 unlensed cases, however, it gives one false-positive case while identifying the rest of the 19 unlensed systems correctly. In comparison, the minimization of the fluctuation (in the reconstructed image light curves) approach produces 1 false-negative case but zero false-positive cases for the same data sets. We also notice that the signal in $\Sigma_{\mathrm{ACF}(H;\ \Delta t)}$ is slightly diminished as compared to $\Sigma_{\epsilon(\Delta t)}$.

When we consider uncertainty in the observed light-curve data, the ACF(H; Δt) approach suffers more from the added noise than the technique based on the fluctuation curve. We follow the same prescription given in Bag et al. (2022) for both methods to handle noisy data; we smooth the joint light curve for multiple smoothing scales and combine the fluctuation curves obtained from each smoothed light curve. For the same data sets with ZTF-like noise, the ACF(H; Δt) approach detects only eight out of the 20 lensed systems correctly. However, for two other lensed cases, it detects the lensing nature based on peaks at completely wrong time delays (hence these two should be counted as false positives). Furthermore, it detects 15 out of 20 unlensed systems correctly but gives the remaining five as false positives. Thus, in combination, it produces a precision of $8/15$ (slightly higher than 50%) and a recall of $8/20$. These numbers are significantly inferior to that of the fluctuation curve method which produces a precision of $12/13$ and a recall of $12/20$ on the same data sets (Bag et al. 2022). The recall and precision values for these two approaches have been summarized in Table 1. As before, we again notice that the lensing signal in $\Sigma_{\mathrm{ACF}(H;\ \Delta t)}$ is typically diminished as compared to $\Sigma_{\epsilon(\Delta t)}$.

We also test ACF(H; Δt) on the COSMOGRAIL system SDSS J1226-0006, which was used in Bag et al. (2022) as an example. The time delay estimated using the resolved light curves by the COSMOGRAIL team is 33.7 ± 2.7 days for this system (Millon et al. 2020). Figure 6 compares the fluctuation curve (blue curve, same as in Figure 21 of Bag et al. (2022)) and ACF(H; Δt) (orange curve) after normalization. Like the fluctuation curve, $-\Sigma_{\mathrm{ACF}(H;\ \Delta t)}$ shows a pair of prominent secondary minima symmetrically placed around $\Delta t = 0$. Specifically, the secondary minima in $-\Sigma_{\mathrm{ACF}(H;\ \Delta t)}$ occurs at $\Delta t = -28.9,\ 26.9$ days with the depths $\Sigma = -1.40$ and -1.50, respectively, leading to the final time-delay estimation of $\Delta t_{\mathrm{est}} = 27.9$ days. In comparison, $\Sigma_{\epsilon(\Delta t)}$ curve exhibits slightly overall deeper minima ($\Sigma = -1.35$, -1.68) at $\Delta t = -28.7,\ 30.5$ days that give rise to $\Delta t_{\mathrm{est}} = 29.6$ days (Bag et al. 2022) which is a better agreement with the COSMOGRAIL results.

In conclusion, it is evident that the fluctuation curve approach as introduced by Bag et al. (2022) performs better than ACF(H; Δt) although the former method is dominated by the latter. This is because the secondary minima (i.e., the lensing signal) are more prominent in the fluctuation curves, $\epsilon(\Delta t)$, than the secondary maxima in the corresponding ACF (H; Δt) curves. This is due to the fact that, in $\epsilon(\Delta t)$ curves, the secondary minima at $\Delta t = \pm\widehat{\Delta t}$ are further enhanced by all the odd terms in Equation (11) as compared to ACF(H; Δt) curves. Therefore, the minimization of fluctuation approach can be more useful for marginal detection of the unresolved lensed QSOs. However, ACF(H; Δt) can be very useful and quick crosscheck as it is computationally inexpensive.

6. Conclusion

Bag et al. (2022) introduces a data-driven technique for detecting lensed QSOs and for measuring their time delays only using the unresolved joint light-curve data by minimizing the fluctuations in the reconstructed image light curves. In this article, we provide the proof as to how this method works. We showed that the lensing signal in the simple fluctuation estimator given by Equation (5) is dominated by

Table 1
We Compare the Results from the Two Approaches—ACF(H; Δt) vs. the Full Fluctuation Curve—in Terms of Recall (Completeness) and Precision (Purity) for the Validation and Blind Sets Considered in Bag et al. (2022)

Data Sets from (Bag et al. 2022)	Recall		Precision	
	ACF(H; Δt)	Fluctuation curve (Bag et al. 2022)	ACF(H; Δt)	Fluctuation curve (Bag et al. 2022)
With marginal noise	95%	95%	95%	100%
With ZTF-like noise	40%	60%	53.33%	92.31%

Note. The results corresponding to marginal noise and ZTF-like noise in the data are presented in the top and bottom rows, respectively.

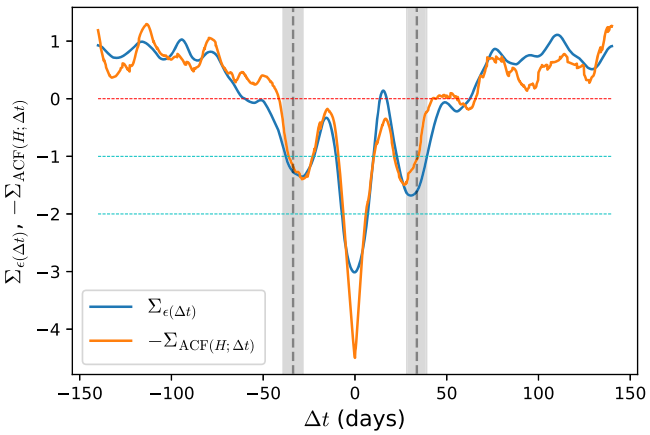


Figure 6. The normalized fluctuation curve (blue curve, from Bag et al. (2022)) and ACF(H; Δt) (shown by the orange curve) are compared for the COSMOGRAIL system SDSS J1226-0006.

the autocorrelation of the derivative (difference series in general for nonuniformly sampled data) of the joint light curve. This observation explains all the characteristics of the fluctuation curve that were used as the foundation of the technique proposed by Bag et al. (2022). Above all, ACF(H; Δt) is locally maximized at $\Delta t = \pm\widehat{\Delta t}$ and these lensing peaks manifest themselves as the secondary minima in the fluctuation curve, $\epsilon(\Delta t)$, which have been used to detect the lensed cases in Bag et al. (2022). Other interesting results are summarized below.

1. We also showed that ACF(H; Δt) is more reliable than the autocorrelation function of the joint light curve itself ACF (F; Δt), because the intrinsic flux variability of QSOs is correlated in the time domain, or in other words, the power spectra of the intrinsic quasar light curves are of the red type. Nevertheless, even if $P_f(w)$ is flat or of blue type, ACF (H; Δt) can also find the lensed cases. Since the primary contribution to the lensing signal in $\epsilon(\Delta t)$ comes from the ACF(H; Δt), the minimization of the fluctuation approach would be similarly successful in these scenarios.
2. However, the approach based on the fluctuation curve provides better recall and precision over the ACF(H; Δt) when one considers the significant amount of noise in the joint light-curve data. This is due to the higher order terms contributing in Equation (11) that further enhance the pair of secondary minima in the $\epsilon(\Delta t)$ curve (as compared to the lensing peaks in ACF(H; Δt)).
3. For a generic lensed system having N_I images, ACF(H; Δt) displays $N_I(N_I-1)/2$ pairs of lensing local maxima. Likewise, the fluctuation minimization method can be used to detect multiple imaged lensed systems even if we assume two images in the reconstruction analysis a priori. For example, by detecting one pair of prominent secondary minima in $\epsilon(\Delta t)$, one can identify a double system, whereas if there exist multiple such lensing minima pairs (up to six), the system must be a quad (having four images).

Although we choose time delays of the order of tens of days as examples in this work for demonstration purposes, one can in principle detect the lensing minima pair in the fluctuation curve for any arbitrarily small time delay as long as it is sufficiently larger than the observation cadence and the signal-to-noise ratio is sufficiently high. However, in reality, the cadence could vary from a few days to $\mathcal{O}(10)$ days depending upon the observation conditions or the observing strategy; this puts a limitation on the sensitivity of the method as the time delay needs to be larger than the cadence for the lensing minima pair to emerge in the fluctuation curve.

The fact that the primary contribution to the lensing signal in the fluctuation curves stems from ACF(H; Δt) also informs us about some key benefits of the fluctuation minimization approach. The fluctuation minimization approach should be able to handle microlensing up to a certain limit as the autocorrelation can withstand moderate microlensing effects. We plan to comprehensively investigate the effect of microlensing on the performance of this method in the follow up work. Note that microlensing can significantly alter the time-delay measurements from the resolved image light curves, up to a few days (Tie & Kochanek 2018; Liao 2020). It would be interesting to see how this affects the results of our method based on the unresolved fluxes.

To discern another crucial advantage, recall that the selection criteria for lens detection using the fluctuation curve are so far based on only the pair of minima at $\Delta t = \pm\widehat{\Delta t}$. The higher order terms further put predictable features in the fluctuation curve at certain values of Δt, e.g., a pair of minima (maxima) at $\Delta t = \pm\widehat{\Delta t}/n$ for every odd (even) n. Although the higher order features are suppressed by μ^n, the first few of these features can nevertheless be exploited to improve the selection criteria, potentially using deep learning.

Finally, we emphasize that the enhancement of fluctuations in the image light curves reconstructed using wrong time delays is a fundamental characteristic of the fluctuation curve approach. However, this article is restricted to the simple metric Equation (5) for quantifying the fluctuations. The lensing signal in this estimator $\epsilon(\Delta t)$ is found to be dominated simply by ACF (H; Δt). Nevertheless, there might be a better metric for quantifying the fluctuations in the reconstructions that delivers better results in terms of recall and precision. This remains an open question and is worth investigating further.

As stressed in Bag et al. (2022), to estimate the error in the time-delay measurements in this nonparametric approach, one needs to statistically analyze a large number of unresolved cases simulated in a variety of observational conditions; e.g., considering different flux variations, cadence distributions, noise levels, many microlensing realizations, etc. This exercise forms the focus of the follow up work.

We thank Tommaso Treu for his crucial inputs to this work. S. B. also thanks Eric V. Linder and Alex G. Kim for useful discussions. The Seondeok cluster at Korea Astronomy and Space Science Institute has been used to carry out a part of the analysis. A.S. would like to acknowledge the support by the National Research Foundation of Korea NRF-2021M3F7A1082053 and the support of the Korea Institute for Advanced Study (KIAS) grant funded by the government of Korea. K.L. was supported by the National Natural Science Foundation of China (NSFC) under grant Nos. 12222302, 11973034 and Wuhan University talent research start-up funds. A.S. and K.L. also acknowledge the support and hospitality received from Beijing Normal University.

Appendix A
Autocorrelation Function

The autocorrelation function of a generic time series $x(t)$ is defined as

$$\mathrm{ACF}(x;\,\Delta t)=\frac{\sum_i^{N_D}[x(t_i)-\langle x(t)\rangle_t][x(t_i-\Delta t)-\langle x(t-\Delta t)\rangle_t]}{\sqrt{\sum_i^{N_D}[x(t_i)-\langle x(t)\rangle_t]^2\cdot\sum_i^{N_D}[x(t_i-\Delta t)-\langle x(t-\Delta t)\rangle_t]^2}}. \tag{A1}$$

For a wide-sense stationary (bounded and long) time series $\langle x(t)\rangle_t\approx\langle x(t-\Delta t)\rangle_t$ which reduces the denominator to $N_D\sigma_x^2$. Let us define the joint time series $X(t)$ following the lensing equation (similar to $F(t)$ in Equation (2) and $H(t)$ in Equation (18))

$$X(t)=x(t)+\widehat{\mu}x(t-\widehat{\Delta t}), \tag{A2}$$

where $\widehat{\mu}$ and $\widehat{\Delta t}$ represent the (true) magnification ratio and relative time delay as in Equations (2) and (18). Thus, $\{x(t), X(t)\}$ here are proxies for $\{f(t), F(t)\}$ or their derivatives $\{h(t), H(t)\}$. Using Equation (A1) it is easy to find thatwhere the denominator is just a normalization constant (independent of Δt).

$$\mathrm{ACF}(X;\,\Delta t)=\frac{\mathrm{ACF}(x;\,\Delta t)+\frac{\widehat{\mu}}{1+\widehat{\mu}^2}[\mathrm{ACF}(x;\,\Delta t-\widehat{\Delta t})+\mathrm{ACF}(x;\,\Delta t+\widehat{\Delta t})]}{1+\frac{2\widehat{\mu}}{1+\widehat{\mu}^2}\mathrm{ACF}(x;\,\widehat{\Delta t})}, \tag{A3}$$

If the intrinsic time series $x(t)$ is uncorrelated in time (white noise), sufficiently long, and hence obeys

$$\langle x(t)x(t-\Delta t)\rangle_t=\begin{cases}\langle x(t)^2\rangle_t & \text{if } \Delta t=0,\\ \langle x(t)\rangle_t^2 & \text{if } \Delta t\neq 0,\end{cases}$$
$$\text{therefore, } \mathrm{ACF}(x;\,\Delta t)=\begin{cases}1 & \text{if } \Delta t=0,\\ 0 & \text{for all } \Delta t\neq 0,\end{cases} \tag{A4}$$

the denominator of Equation (A3) reduces to unity (in the generic cases with $\widehat{\Delta t}\neq 0$) leading to

$$\mathrm{ACF}(X;\,\Delta t)=\mathrm{ACF}(x;\,\Delta t)+\frac{\widehat{\mu}}{1+\widehat{\mu}^2}\times[\mathrm{ACF}(x;\,\Delta t-\widehat{\Delta t})+\mathrm{ACF}(x;\,\Delta t+\widehat{\Delta t})], \tag{A5}$$

and we can conclude the following.

1. When $\Delta t=0$, Equation (A5) trivially reduces to unity as only the first term contributes. Naturally, the autocorrelation is always maximized at unity for no shift in time ($\Delta t=0$).
2. On the other hand, for $\Delta t\neq 0$ in general, Equation (A4) ensures that all three terms in Equation (A5) reduce to zero and $\mathrm{ACF}(X;\ \Delta t)\approx 0$.
3. Interestingly, $\mathrm{ACF}(H;\ \Delta t)$ is locally maximized at $\Delta t=\pm\widehat{\Delta t}$. In the view of Equation (A4), when $\Delta t=\widehat{\Delta t}$ or $-\widehat{\Delta t}$ only the first or the second term in the square bracket of Equation (A5) contributes, that leads to
$$\mathrm{ACF}(X;\,\Delta t=\pm\widehat{\Delta t})\approx\frac{\widehat{\mu}}{1+\widehat{\mu}^2}\leqslant\frac{1}{2} \tag{A6}$$
for both cases. In fact, the set of intrinsic features in the underlying time series $x(t)$ appears twice in $X(t)$, separated by $\widehat{\Delta t}$ (and scaled by $\widehat{\mu}$, see Equation (A2)). Since $\mathrm{ACF}(X;\,\Delta t)$ measures the correlation between two copies of $X(t)$ with one shifted by Δt in the time domain, it gains an excess power when $\Delta t=\pm\widehat{\Delta t}$ due to matching of the two sets of the same features (the $\pm$ sign accounts for the shift in either direction). Although this argument applies to any generic $x(t)$ with a substantial amount of features, we emphasize that the above equation is valid strictly when the $x(t)$ is pure white noise following Equation (A4).
4. For unlensed cases, when $\widehat{\mu}=0$ or $\widehat{\Delta t}=0$, this pair of secondary maxima vanishes.

In summary, $\mathrm{ACF}(X;\,\Delta t)$ shows a pair of secondary maxima at $\Delta t=\pm\widehat{\Delta t}$ (the lensing peaks) and remains zero at all other $\Delta t\neq 0$ as long as the intrinsic time series, $x(t)$, is white noise (no temporal correlation) and satisfies Equation (A4). Note that the whole analysis is applicable to both $F(t)$ and $H(t)$ with the underlying functions $f(t)$ and $h(t)$, respectively.

A.1. When to Expect Lensing Peaks in ACF(F; Δt) Even If f(t) Is Not White Noise

Let us discuss the interesting case when Equation (A4) is not strictly valid. This is important for quasars since the intrinsic light curves, $f(t)$, typically possess correlations up to a timescale of a few hundred days (Kelly et al. 2009; MacLeod et al. 2010). In such cases, the autocorrelation function ACF(f; Δt) would show a broad peak at $\Delta t = 0$, unlike a delta function. ACF(f; Δt)is still expected to be symmetric around $\Delta t = 0$ and monotonically decaying with $|\Delta t|$ (the decay rate depends on the correlation timescale). For instance, see the top-left panel of Figure 4 for typical examples of ACF(f; Δt) where $f(t)$ is generated from a damped random walk (DRW) with a correlation scale of $\sim$300 days.

The lensing peaks in ACF(F; Δt) arise from the terms in the square bracket in the numerator of Equation (A3) (after replacing $\{x(t), X(t)\}$ by $\{f(t), F(t)\}$). Thus, as $|\Delta t|$ approaches $|\widehat{\Delta t}|$ from the below, the lensing peaks could emerge only if the change in the second term in the numerator of Equation (A3) dominates over the change in the first term. For all practical purposes (i.e., with the expected ACF(f; Δt) being symmetric around $\Delta t = 0$ and monotonically decreasing with $|\Delta t|$), this requirement boils down to the condition,

$$|\mathrm{ACF}'(f;\ \Delta t \sim \pm\widehat{\Delta t})| < \frac{\widehat{\mu}}{1+\widehat{\mu}^2}|\mathrm{ACF}'(f;\ \Delta t \sim 0)|, \quad \text{(A7)}$$

where ACF$'(f; \Delta t)$ is the derivative of ACF(f; Δt) with respect to Δt. The stronger the inequality is, the steeper (more prominent) the lensing maxima pair in ACF(F; Δt) would be. For our purpose, the above condition requires that ACF(f; Δt) must have sufficiently narrow peak at $\Delta t = 0$ so that it stabilizes at $|\Delta t| \gtrsim |\widehat{\Delta t}|$ by decaying down sufficiently fast. We again emphasize that all the arguments made in this section stand valid if we replace $F(t)$ and $f(t)$ by $H(t)$ and $h(t)$, respectively. Note that, here, we assumed the best-case scenario when the noise is negligible; the inclusion of observation noise brings in additional complexities.

Appendix B
Exact Values of the Global and Secondary Minima in the Fluctuation Curve

B.1. Global Minima in the Fluctuation Curve

For $\Delta t = 0$, Equation (3) reduces to $f_{\rm rec}(t) = \frac{1}{1+\mu}F(t)$ which reduces Equation (5) to

$$\epsilon(\Delta t = 0) = \frac{1}{(1+\mu)^2}\sum_{i}^{N_D}[F(t_{i+1}) - F(t_i)]^2 = \frac{\epsilon_0}{(1+\mu)^2}. \quad \text{(B1)}$$

Note that the above expression is valid even for the unlensed cases.

One can get the same expression from Equation (11) as follows,

$$\begin{aligned}\epsilon(\Delta t = 0) &= \epsilon_0[1 - 2\mu + 3\mu^2 - 4\mu^3 + 5\mu^4 - 6\mu^5 + \ldots] \\ &= \epsilon_0\frac{\partial}{\partial\mu}[\mu - \mu^2 + \mu^3 - \mu^4 + \mu^5 - \mu^6 + +\ldots], \\ &= \epsilon_0\frac{\partial}{\partial\mu}\left[\frac{\mu}{1+\mu}\right] = \frac{\epsilon_0}{(1+\mu)^2}.\end{aligned} \quad \text{(B2)}$$

Here we identify that, for $\Delta t = 0$, all ϵ_n terms become proportional to ϵ_0, and the sum converges as $\mu < 1$.

B.2. Secondary Minima When h(t) Is White Noise

If $h(t)$ is pure white noise (or in other words, if it follows Equation (A4)), we can analytically calculate the height of the secondary minima. Following Equation (28), we get

$$\begin{aligned}&\epsilon(\Delta t = \pm\widehat{\Delta t}) \\ &\quad = \epsilon_0[1 - 2\mathrm{ACF}(H;\ \pm\widehat{\Delta t})\mu + \mu^2 - 2\mathrm{ACF}(H;\ \pm\widehat{\Delta t})\mu^3 \\ &\qquad + \mu^4 - 2\mathrm{ACF}(H;\ \pm\widehat{\Delta t})\mu^5 + \mu^6 - \ldots],\end{aligned} \quad \text{(B3)}$$

$$\begin{aligned}&= \epsilon_0[(1 + \mu^2 + \mu^4 + \mu^6 + \ldots) \\ &\quad - 2\mathrm{ACF}(H;\ \pm\widehat{\Delta t})(\mu + \mu^3 + \mu^5 + \ldots)],\end{aligned} \quad \text{(B4)}$$

$$= \epsilon_0\frac{1 - 2\mu\mathrm{ACF}(H;\ \pm\widehat{\Delta t})}{1 - \mu^2}, \quad \text{(B5)}$$

where we use ACF(H; $n\cdot\widehat{\Delta t}$) ≈ 0 for $n \neq 0,\ \pm 1$. Note from Equation (A6) that ACF(H; $\pm\widehat{\Delta t}$) $\approx \widehat{\mu}/(1+\widehat{\mu}^2) \leqslant 1/2$ in this limit. It is also easily seen that for unlensed cases, when ACF(H; $\Delta t \neq 0$) $= 0$, $\epsilon(\Delta t) = \epsilon_0/(1-\mu^2) =$ constant for any $\Delta t \neq 0$.

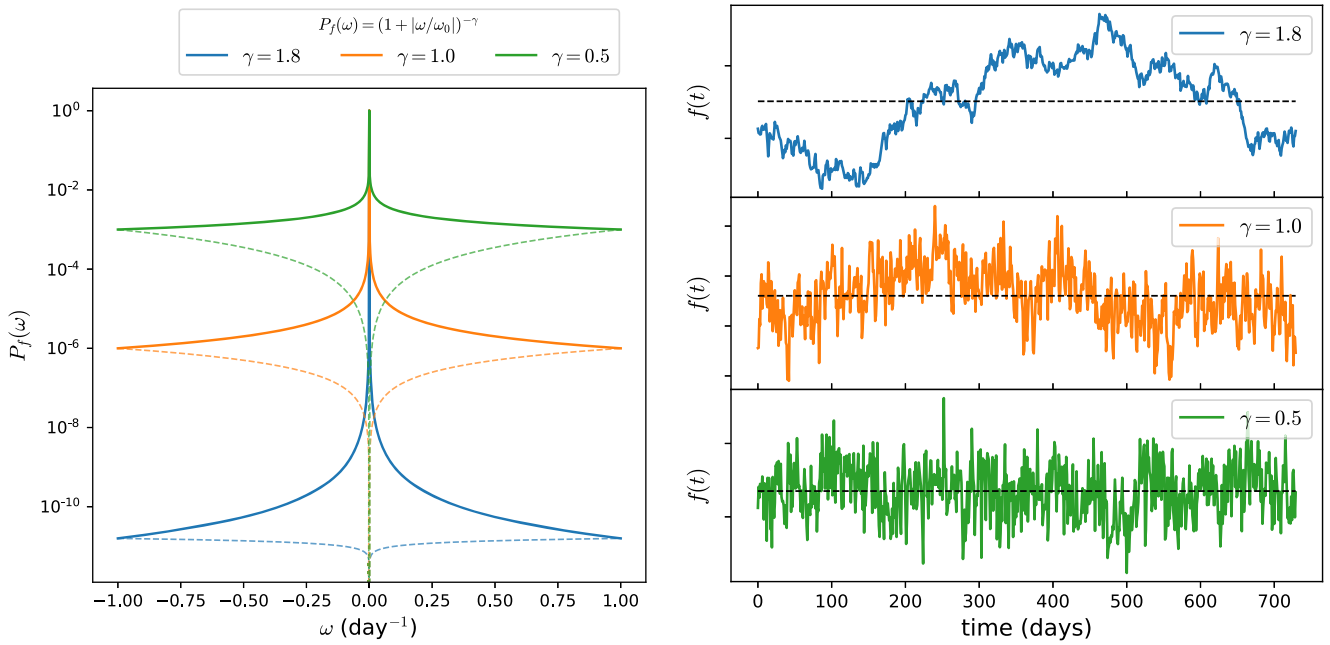
By comparing with Equation (B2), one can trivially show that

$$\begin{aligned}\frac{\epsilon(\Delta t = \pm\widehat{\Delta t})}{\epsilon(\Delta t = 0)} &= \frac{1+\mu}{1-\mu}(1 - 2\mu\mathrm{ACF}(H;\ \pm\widehat{\Delta t})) \\ &\approx 1 + 2\mu(1 - \mathrm{ACF}(H;\ \pm\widehat{\Delta t})) + \mathcal{O}(\mu^2),\end{aligned} \quad \text{(B6)}$$

and owing to ACF(H; $\pm\widehat{\Delta t}$) $\leqslant 1/2$ it is readily seen that $\epsilon(\Delta t = 0) < \epsilon(\Delta t = \pm\widehat{\Delta t})$. Thus the central minimum at $\Delta t = 0$ will always be deeper than the pair of lensing minima.

Appendix C
ACF(H; Δt) Is More Reliable Than ACF(F; Δt) for Any Red Power Spectrum

We have seen in Section 4 that the quasar intrinsic flux variability generated from a random damped walk (DRW) follows a red power spectrum given by Equation (32). We



(a) (b)

Figure 7. The left panel shows the power spectra ($P_f(\omega)$ in log scale) given by Equation (C3) for $\gamma = 1.8$, 1.0, and $1/2$ by the blue, orange, and green solid curves, respectively. Here we set $\omega_0 = 10^{-6}$ so that the power spectra behave like $|\omega|^{-\gamma}$ for the most part. The dashed curves represent the power spectra of the derivative: $P_h(\omega) = \omega^2 P_f(\omega) \approx |\omega|^{2-\gamma}$. Each of the three panels on the right displays an example of the intrinsic flux $f(t)$ (in linear scale and arbitrary unit) generated from these power spectra. Note that observational noise is marginal here. The horizontal dashed line in each right panel represents the mean of the light curve.

found that ACF(H; Δt) is more likely to exhibit the lensing peaks as compared to ACF(F; Δt) in this case due to the presence of temporal correlation in $f(t)$ up to a few hundred days realistically. Moreover, we discussed that a larger correlation scale (τ) leads to a steeper (or redder) $P_f(\omega)$ in Equation (32) and a flatter (expected) ACF(f; Δt)in Equation (30), which enables ACF(H; Δt) to outperform ACF(F; Δt) even more profoundly. In this appendix, we argue that this is not restricted to just the DRW template; for any time series with a red-type power spectrum, ACF(H; Δt) is more reliable than ACF(F; Δt) for finding the lensed cases.

Recall that the expected autocorrelation function of a "wide-sense stationary" time series is given by the Fourier transform of the power spectrum. Since $\langle \mathrm{ACF}(f;\ \Delta t)\rangle_E$ is symmetric around its peak at $\Delta t = 0$, the power spectrum should also be symmetric around $\omega = 0$; $P_f(\omega)$ must be a function of $|\omega|$ only. Recall further from Equations (A3) and (A7) that ACF(F; Δt) can show the lensing peaks only if ACF(f; Δt) decays sharply from its peak at $\Delta t = 0$ and plateaus by $|\Delta t| \gtrsim |\widehat{\Delta t}|$. In other words, the narrower the (central) peak in ACF(f; Δt), the better the chance of finding the lensing peaks in ACF(F; Δt). The same criterion applies to ACF(h; Δt) for the lensing peaks to appear in ACF(H; Δt). However, for all practical purposes, a redder (or steeper) power spectrum, $P_f(\omega)$, leads to a flatter $\langle \mathrm{ACF}(f;\ \Delta t)\rangle_E$ which diminishes the probability of finding the lensing peaks in ACF(F; Δt). A completely general proof of this statement is challenging, so here we argue using some concrete examples.

First, consider a Gaussian power spectrum:

$$P_f(\omega) \propto \exp(-a\ \omega^2/2). \tag{C1}$$

Its Fourier transform is then also Gaussian with $\langle \mathrm{ACF}(f;\ \Delta t)\rangle_E \propto \exp[-\Delta t^2/(2a)]$. Thus, having a large a in Equation (C1) gives a narrow $P_f(\omega)$ but a broad $\langle \mathrm{ACF}(f;\ \Delta t)\rangle_E$. This argument can be made more general; when we scale a function by $t \to at$, its Fourier transform gets inversely scaled,

$$\mathrm{F.T}[g(at)] = \frac{1}{|a|} G(\omega/a), \tag{C2}$$

where $G(\omega)$ is the Fourier transform of a generic function $g(t)$. Thus, if a scaling makes $P_f(\omega)$ steeper, $\langle \mathrm{ACF}(f;\ \Delta t)\rangle_E$ would be flatter and vice-versa. On the other hand, the derivative series $h(t)$ always has a bluer power spectrum than that of $f(t)$ as $P_h(\omega) = \omega^2 P_f(\omega)$. Thus ACF($h$; Δt) falls sharper from its peak at $\Delta t = 0$ as compared to ACF(f; Δt) resulting in more prominent lensing peaks in ACF(H; Δt).

Let us consider another example of red-type power spectrum,

$$P_f(\omega) \propto (1 + |\omega/\omega_0|)^{-\gamma},\ \gamma > 0, \tag{C3}$$

which behaves like a power-law equation for $\omega \gg \omega_0$ but avoids blowing up at $\omega \to 0$ by remaining stable at $|\omega| \ll \omega_0$. As the analytical solution for the Fourier transform of Equation (C3) does not exist for a generic γ, we carry out a numerical analysis. We set $\omega_0 = 10^{-6}$ day^{-1} so that $P_f(\omega)$ behaves like $|\omega|^{-\gamma}$ in the most part (except $\omega \approx 0$). In the left panel of Figure 7, we show three such power spectra with $\gamma = 1.8,\ 1.0$, and $1/2$ by the solid blue, orange, and green curves, respectively. The power spectrum of the derivative, $P_h(\omega) = \omega^2 P_f(\omega) \approx |\omega|^{2-\gamma}$, is shown by the dashed curve with respective color. The three panels on the right show examples of the intrinsic light curves, $f(t)$, generated from the power spectra with the three values of γ in Equation (C3); the dashed horizontal line in each right panel represents the mean of $f(t)$. It is evident that as $P_f(\omega)$ becomes redder with larger γ, $f(t)$ shows correlation until a longer timescale (i.e., two nearby points are

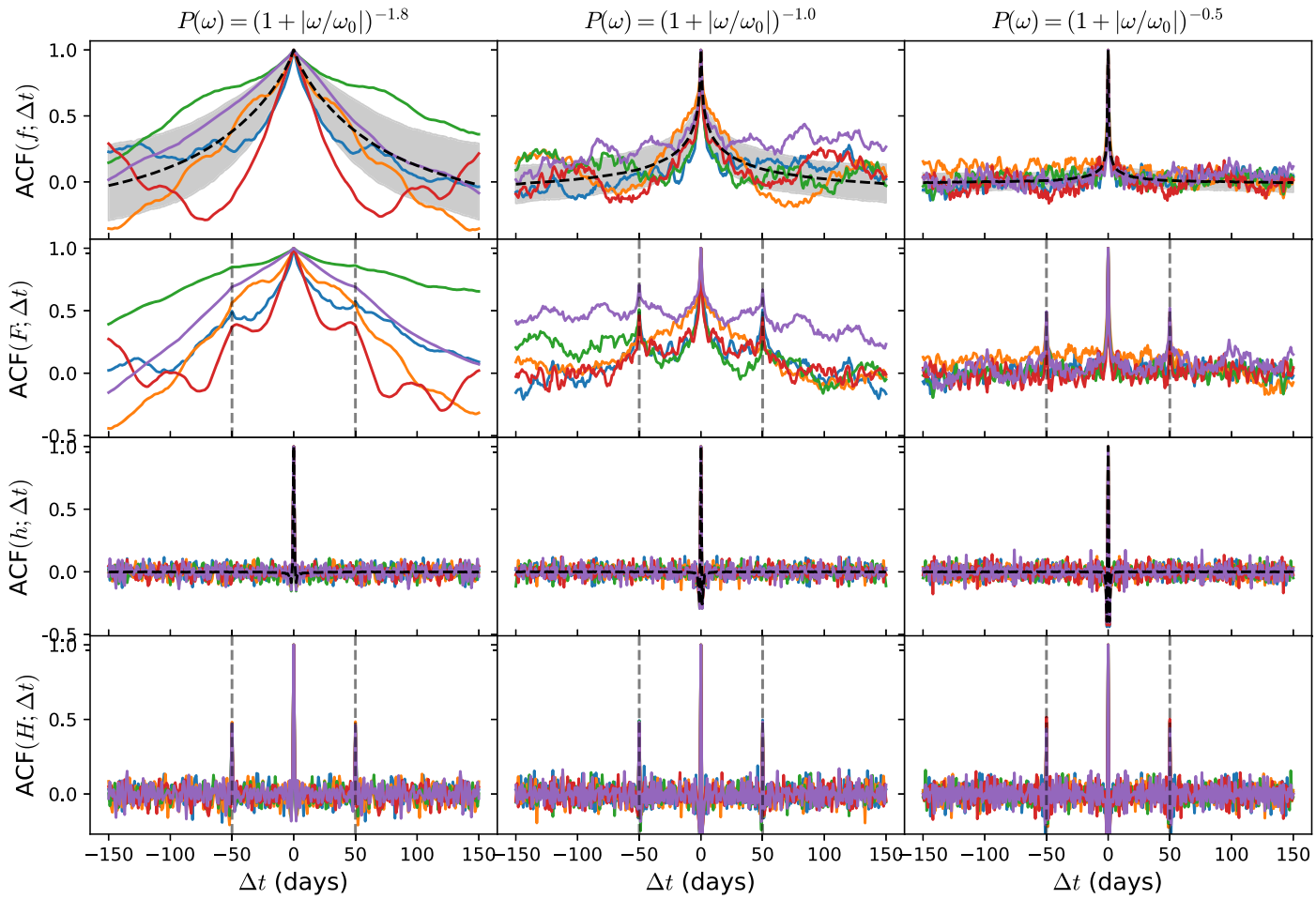


Figure 8. Comparing ACF(f; Δt), ACF(F; Δt), ACF(h; Δt) and ACF(H; Δt) for five random realizations out of 1000 for three red-type power-law spectra: $P(\omega) = (1 + |\omega/\omega_0|)^{-\gamma}$ with $\gamma = 1.8$ (left panels), 1 (middle panels), and $1/2$ (right panels). Here, we set $\omega_0 = 10^{-6}$ day^{-1} so that $P_f(\omega)$ behaves like a power law for most part of it.

more likely to be either above or below the mean unless not separated sufficiently in time).

Similar to the example illustrated in Figure 4 for the DRW template, we now simulate 1000 realizations for the intrinsic flux variability separately for each of these three power spectra; the corresponding results are arranged in the three columns of Figure 8. Then for each realization, we construct a double-lensed system using Equation (2) with $\widehat{\mu} = 0.86$ and $\widehat{\Delta t} = 50.0$ days, again we consider the perfect condition with marginal noise in the data for simplicity. The four rows in Figure 8 show ACF(f; Δt), ACF(F; Δt), ACF(h; Δt), ACF(H; Δt), respectively, from the top. Five random realizations have been shown by the solid curves in each panel. The dashed black curve and the shaded region in the first and third row (from the top) panels represent the ensemble average and 68% quantile around it. From the top-row panels, it is clearly evident that the redder the power spectrum (i.e., with larger γ) is, the slower the ACF(f; Δt) decays from unity and the broader the 68% quantile is. Therefore, a redder power spectrum in turn reduces the probability of detecting the lensing peaks in the ACF(F; Δt) as shown in the second row (from the top) panels. Also, statistically, we tend to get more false-positive cases, e.g., the blue/red curve in the top-left panel.

On the other hand, for smaller γ (flatter power spectrum), ACF (f; Δt) tends to decay faster that increases the possibility of detecting lensed systems through ACF(F; Δt) (indeed, in the limit $\gamma \to 0$, $f(t)$ becomes white noise and ACF(f; Δt) $\propto \delta(0)$). Nevertheless, the peak prominence is still inferior to that of ACF(H; Δt) (bottom panels). In contrast, for all the γs, ACF(h; Δt) decays sharply from its peak at $\Delta t = 0$ as evident from the dashed black curves in the third (from the top) row panels.[9] Thus, ACF(H; Δt) shows the sharp lensing peaks at $\Delta t = \pm\widehat{\Delta t}$ (marked by the dashed vertical lines) for all the realizations and for all the power spectra considered. Therefore, Figure 8 (along with the Figure 4 for DRW process) demonstrates that for any red-type power spectrum, ACF(H; Δt) is more reliable than ACF(F; Δt) for finding the lensed systems and the advantage of using ACF(H; Δt) over ACF(F; Δt) increases for a redder power spectrum.[10] Since the lensing signal in the $\epsilon(\Delta t)$ curve is dominated by ACF (H; Δt), the fluctuation statistics introduced in Bag et al. (2022) is also very successful in all red power-spectra scenarios.

[9] When $\gamma = 2$, $P_h(\omega) \approx |\omega|^{2-\gamma} =$ constant and $h(t)$ becomes white noise. Hence, for $\gamma < 2$, $h(t)$ possess (negative) correlation restricted to only the adjacent point that explains the local minima in ACF(h; Δt) just next to $\Delta t = 0$ on either side. But, this does not affect the detection efficiency of ACF(H; Δt).

[10] For completeness, let us discuss if the power spectrum is flat or blue ($\gamma > 0$) although it does not describe QSO light curves well. In these cases, ACF(F; Δt) shows sharp lensing peaks and so does ACF(H; Δt) since the correlation in $h(t)$ is limited to the neighboring point(s). Thus both perform well in these scenarios. This is also illustrated in Appendix C of Bag et al. (2022)

ORCID iDs

Satadru Bag https://orcid.org/0000-0003-0141-606X
Wuhyun Sohn https://orcid.org/0000-0002-6039-8247
Arman Shafieloo https://orcid.org/0000-0001-6815-0337
Kai Liao https://orcid.org/0000-0002-4359-5994

References

Abdalla, E., Abellán, G. F., Aboubrahim, A., et al. 2022, JHEAp, 34, 49
Aghanim, N., Akrami, Y., Ashdown, M., et al. 2020, A&A, 641, A6
Bag, S., Kim, A. G., Linder, E. V., & Shafieloo, A. 2021, ApJ, 910, 65
Bag, S., Shafieloo, A., Liao, K., & Treu, T. 2022, ApJ, 927, 191
Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, PASP, 131, 018002
Biggio, L., Domi, A., Tosi, S., et al. 2022, MNRAS, 515, 5665
Birrer, S., Shajib, A. J., Galan, A., et al. 2020, A&A, 643, A165
Birrer, S., & Treu, T. 2021, A&A, 649, A61
Bonvin, V., Courbin, F., Suyu, S. H., et al. 2017, MNRAS, 465, 4914
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
Denissenya, M., Bag, S., Kim, A. G., Linder, E. V., & Shafieloo, A. 2022, MNRAS, 511, 1210
Denissenya, M., & Linder, E. V. 2022, MNRAS, 515, 977
Dobler, G., Fassnacht, C. D., Treu, T., et al. 2015, ApJ, 799, 168
Einstein, A. 1914, ArS, 37, 254
Geiger, B., & Schneider, P. 1996, MNRAS, 282, 530
Goobar, A., Johansson, J., Schulze, S., et al. 2022, arXiv:2211.00656
Goobar, A., Amanullah, R., Kulkarni, S. R., et al. 2017, Sci, 356, 291
Jee, I., Komatsu, E., Suyu, S. H., & Huterer, D. 2016, JCAP, 2016, 031
Kelly, B. C., Bechtold, J., & Siemiginowska, A. 2009, ApJ, 698, 895
Kelly, P. L., Rodney, S. A., Treu, T., et al. 2015, Sci, 347, 1123
Khintchine, A. 1934, MatAn, 109, 604
Lemon, C., Anguita, T., Auger, M., et al. 2023, MNRAS, 520, 3305
Liao, K. 2020, ApJL, 899, L33
Liao, K., Biesiada, M., & Zhu, Z.-H. 2022, ChPhL, 39, 119801
Liao, K., Treu, T., Marshall, P., et al. 2015, ApJ, 800, 11
LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv:0912.0201
LSST Science Collaboration, Marshall, P., Anguita, T., et al. 2017, arXiv:1708.04058
MacLeod, C. L., Ivezić, Ž, Kochanek, C. S., et al. 2010, ApJ, 721, 1014
Millon, M., Courbin, F., Bonvin, V., et al. 2020, A&A, 640, A105
Millon, M., Galan, A., Courbin, F., et al. 2020, A&A, 639, A101
Mortlock, D. J., Webster, R. L., & Francis, P. J. 1999, MNRAS, 309, 836
Narayan, R., & Bartelmann, M. 1996, arXiv:astro-ph/9606001
Oguri, M. 2007, ApJ, 660, 1
Oguri, M. 2019, RPPh, 82, 126901
Peng, C. Y., Impey, C. D., Falco, E. E., et al. 1999, ApJ, 524, 572
Refsdal, S. 1964, MNRAS, 128, 307
Refsdal, S., & Bondi, H. 1964, MNRAS, 128, 295
Riess, A. G., Yuan, W., Macri, L. M., et al. 2022, ApJL, 934, L7
Rodney, S. A., Brammer, G. B., Pierel, J. D. R., et al. 2021, NatAs, 5, 1118
Saha, P., Coles, J., Macciò, A. V., & Williams, L. L. R. 2006, ApJL, 650, L17
Shajib, A. J., Birrer, S., Treu, T., et al. 2020, MNRAS, 494, 6072
Shu, Y., Belokurov, V., & Evans, N. W. 2021, MNRAS, 502, 2912
Sonnenfeld, A. 2021, A&A, 656, A153
Springer, O. M., & Ofek, E. O. 2021a, MNRAS, 508, 3166
Springer, O. M., & Ofek, E. O. 2021b, MNRAS, 506, 864
Suyu, S. H., Bonvin, V., Courbin, F., et al. 2017, MNRAS, 468, 2590
Suyu, S. H., Goobar, A., Collett, T., More, A., & Vernardos, G. 2023, arXiv:2301.07729
Suyu, S. H., Huber, S., Cañameras, R., et al. 2020, A&A, 644, A162
Tewes, M., Courbin, F., & Meylan, G. 2013, A&A, 553, A120
Tie, S. S., & Kochanek, C. S. 2018, MNRAS, 473, 80
Treu, T. 2010, ARA&A, 48, 87
Treu, T., & Marshall, P. J. 2016, A&ARv, 24, 11
Wiener, N. 1930, AcMa, 55, 117
Wong, K. C., Suyu, S. H., Chen, G. C. F., et al. 2020, MNRAS, 498, 1420
Zu, Y., Kochanek, C. S., Kozłowski, S., & Udalski, A. 2013, ApJ, 765, 106