Self-Lensing Flares from Black Hole Binaries:
Observing Black Hole Shadows via Light Curve Tomography

Jordy Davelaar$^{1,2,3,*}$ and Zoltán Haiman$^1$

$^1$Department of Astronomy, Columbia University, 550 W 120th Street, New York, New York 10027, USA
$^2$Columbia Astrophysics Laboratory, Columbia University, 550 W 120th Street, New York, New York 10027, USA
$^3$Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, New York 10010, USA

(Received 27 December 2021; accepted 15 March 2022; published 9 May 2022)

Supermassive black hole (BH) binaries are thought to produce self-lensing flares (SLFs) when the two BHs are aligned with the line of sight. If the binary orbit is observed nearly edge-on, we find a distinct feature in the light curve imprinted by the relativistic shadow around the background (“source”) BH. We study this feature by ray tracing in a binary model and predict that 1% of the current binary candidates could show this feature. Our BH tomography method proposed here could make it possible to extract BH shadows that are spatially unresolved by high-resolution very long base line interferometry (VLBI).

Background.—Supermassive black hole binaries (SMBBHs) are thought to reside in the cores of many galaxies, as a result of galaxy mergers [1]. Their immediate surroundings consist of a circumbinary disk that transfers material to two minidisks, each orbiting one BH component. If the viewing angle with respect to the orbital plane is close to edge-on, a minidisk can get lensed by the foreground BH resulting in a self-lensing flare (SLF) [2,3], especially in the case of ultracompact binaries close to merger, in the regime where their gravitational wave (GW) emission is detectable by LISA [4].

Observational evidence of SLFs is sparse, a first candidate was identified by [5] in the optical Kepler data, KIC-11606854 (dubbed “Spikey”). A second candidate, identified in x rays, was discussed by [3]. More generally, quasiperiodic modulations in the light curve of active galactic nuclei (AGN) indicating the presence of a SMBHB have recently been found in large optical time-domain surveys [6–10].

The first attempts of modeling SLFs were made by using a point source approximation for the lens and the source, with the amplification factor derived from microlensing [2,4,5]. These models are computationally cheap, but they lack general relativistic (GR) effects, and do not take finite source sizes into account. Reference [2] also studied the impact of finite source size, but their emission morphology, lacks the strong bending of light close to the BH since GR effects are not taken into account. References [3] and [11] used general-relativistic ray tracing (GRRT) to study SLFs, but this was either limited to a single mass ratio $q = 0.01$ [11] or aimed at a single source and only considered circular equal-mass binaries [3]. Reference [11] used a superimposed binary metric, while [3] used an image of a BH as a faraway image lensed by a single Kerr BH. See also for various applications of lensing in compact binaries [12–17].

In single BH emission models, there is typically a flux depression present in the apparent image. In the optically thin case, this “hole” in the image coincides with the BH shadow (BHS) [18–20], while in the optically thick case, the hole coincides with the direct lensed image of the event horizon or the “inner shadow” [21,22]. The shadow is a flux depression in the apparent surface brightness, caused by photons being trapped by the event horizon. The first direct observation of such a flux depression or BHS was made by the Event Horizon Telescope (EHT) [23]. They used very-long baseline interferometry to construct an image of the immediate surrounding of the SMBH in the nucleus of M87.

A limitation of direct imaging is that an extremely high spatial resolution of 20 $\mu$arc sec was needed to resolve M87*. This limits the potential number of BHs measurable by the current EHT array, although this number is expected to grow with future EHT upgrades [24,25].

In our companion paper, Ref. [26], we present a comprehensive suite of SLF models, obtained via GRRT calculations with our modified version of the GRRT code RAPTOR [27,28]. The code includes a superposed binary metric where the BHs are on Keplerian orbits. The emission is generated by two Novikov-Thorne–like [29] minidisks. The temperature of the disks is set such that the spectra are in agreement with hydro simulations of shock heated minidisks, which have the peak of their thermal emission around the 1–10 keV band, see, e.g., [30–33]. Each minidisk extends from the horizon or the innermost stable circular orbit (ISCO) to the tidal truncation radius. For full details on the code and model, see Ref. [26].

In this Letter, we investigate the substructure found in the SLFs for a subset of models in Ref. [26]. These “dips” occur when the BHS passes behind the lens. We focus on high-energy emission, which is concentrated in a small region around each BH component, and shares the BH
gravitational radius), we find that this phase difference corresponds to a size of \( d \approx 10.4 \, R_g \), which is the size of the projected shadow. For guidance, the vertical lines mark the phase duration of the ISCO and photon rings centered on the SLF.

In the optically thin case, model M2, presented in Fig. 2, the main subpeaks trace the photon ring, since this is the dominant feature in the image. The spacing in phase between the subpeaks is equal to that for the subpeaks produced by the photon ring in model M1. We again mark the phase durations of the ISCO and photon rings centered on the middle of the SLF. The dip inside the flare coincides in phase with the shadow’s passage behind the lens.

In models M3a–b, we vary the spin of the BHs (assumed to be the same for both components for simplicity) to \( a_{\text{BH}} = 0.5 \) and 0.95, respectively, compared to the non-spinning fiducial case. The light curves of these models are shown in Panel 3 of Fig. 2. The addition of spin results in a slightly smaller and more asymmetric shadow. This is also visible in the light curves; the spacing between the peaks shrinks for increasing spin. However, as the figure shows, the effect of the spins is modest.

In models M4a–f we study the dependence on binary separation by changing \( a_{\text{orb}} \) to 200, 300, 400, 500, and 1000 \( R_g \), compared to 100 \( R_g \) in the fiducial model. Increasing the separation narrows the dip in the SLF, since the source spends a smaller fraction of the orbit within the Einstein radius. In M5a–f we alter the inclination to 89°, 88°, 87°, 86°, 85°, and 80° respectively, compared to the fiducial 90°. These light curves are shown in Panel 5 of Fig. 2. The dip is clearly present in models M5a–c, which puts a limit on the range of inclinations for which it can be observed, \( i_{\text{orb}} = 90° \pm 3° \). In model M6a–f we increase the binary separation to 200 \( R_g \) and cover the same span of inclinations. The light curves of these models can be seen in

---

**TABLE I.** Model parameters. M0 is the fiducial model. The parameters that are varied with respect to the fiducial model M0 are shown in bold font.

|     | \( R_{\text{inner}} \) | \( a_{\text{orb}} \) | \( i_{\text{orb}} \) | \( a_{\text{BH}} \) | \( \tau \)      |
|-----|----------------------|----------------|----------------|----------------|----------------|
| M0  | \( r_h \)            | 100            | 90             | 0              | Thick          |
| M1  | \( r_{\text{ISCO}} \) | 100            | 90             | 0              | Thick          |
| M2  | \( r_h \)            | 100            | 90             | 0              | Thin           |
| M3a | \( r_h \)            | 100            | 90             | 0.5            | Thick          |
| M3b | \( r_h \)            | 100            | 90             | 0.95           | Thick          |
| M4a | \( r_h \)            | 200            | 90             | 0              | Thick          |
| M4b | \( r_h \)            | 300            | 90             | 0              | Thick          |
| M4c | \( r_h \)            | 400            | 90             | 0              | Thick          |
| M4d | \( r_h \)            | 500            | 90             | 0              | Thick          |
| M4e | \( r_h \)            | 1000           | 90             | 0              | Thick          |
| M5a | \( r_h \)            | 100            | 89             | 0              | Thick          |
| M5b | \( r_h \)            | 100            | 88             | 0              | Thick          |
| M5c | \( r_h \)            | 100            | 87             | 0              | Thick          |
| M5d | \( r_h \)            | 100            | 86             | 0              | Thick          |
| M5e | \( r_h \)            | 100            | 85             | 0              | Thick          |
| M5f | \( r_h \)            | 100            | 80             | 0              | Thick          |
| M6a | \( r_h \)            | 200            | 89             | 0              | Thick          |
| M6b | \( r_h \)            | 200            | 88             | 0              | Thick          |
| M6c | \( r_h \)            | 200            | 87             | 0              | Thick          |
| M6d | \( r_h \)            | 200            | 86             | 0              | Thick          |
| M6e | \( r_h \)            | 200            | 85             | 0              | Thick          |
| M6f | \( r_h \)            | 200            | 80             | 0              | Thick          |

---

**FIG. 1.** Fiducial model. The top row shows three snapshots during the self-lensing flare. The bottom row shows the light curve at 2.5 keV. Numbers shown on the light curve correspond with numbers in the top panels.
FIG. 2. Various light curves are shown at 2.5 keV for different model parameters. Panel 1: dependence on the inner radius, which is at the horizon in model M0 and at ISCO in model M1. Vertical black lines (also in Panel 2) mark the time interval corresponding to the BHS size and gray lines the ISCO. Panel 2: dependence on the minidisk optical depth, which is assumed to be thick in model M0, and thin in model M2. Panel 3: dependence on the BH spin parameter, which is assumed to be $a = 0$, 0.5 and 0.95 (the same for both BHs) in models M0, M3a, and M3b, respectively. Panel 4: dependence on binary separation, which is 100, 200, 300, 400, 500 and 1000 $R_g$ in models M0 and M4a−e, respectively. Panel 5: dependence on inclination of the binary orbital plane, which is 90° (edge-on), 89°, 88°, 87°, 86°, 85°, and 80° in models M0 and M5a−f, respectively. Panel 6: same as Panel 5, but for a binary separation of 200 $R_g$.

Panel 6 of Fig. 2. As the binary separation increases, the inclination window shrinks to $i_{\text{orb}} = 90° \pm 2°$.

Analytic expectations.—In the case of a perfectly edge-on circular binary, we can derive the expected phase spacing between the two peaks in the optically thin case by taking the ratio between the diameter of the shadow and the circumference of the orbit,

$$\Delta \phi = \frac{d_{\text{shadow}}}{2\pi a_{\text{orb}}}, \quad (1)$$

where $a_{\text{orb}}$ is the orbital radius, and $d_{\text{shadow}} = 2\sqrt{27GM_{\text{source}}/c^2}$ is the diameter of the BHS of the source [18,19,35,36], where $M_{\text{source}}$ is related to the full binary mass via $M_{\text{bin}} = qM_{\text{source}}/(1 + q)$. This assumes the source is the secondary BH, which in general is expected to out-accrete and out-shine the primary BH [37,38].

The orbital radius is given by $a_{\text{orb}} = \sqrt{(4\pi^2T^2/GM_{\text{bin}})}$, where $T$ is the orbital period and $M_{\text{bin}} = M_1 + M_2$ is the total binary mass. Combined with Eq. (1) and using geometrized units ($G = c = 1$), yields

$$\Delta \phi = \frac{2\sqrt{27}}{(4\pi^2)^{1/6}} \frac{qM_{\text{bin}}^{2/3}}{(1 + q)T^{2/3}} \approx 5.63 \frac{qM_{\text{bin}}^{2/3}}{(1 + q)T^{2/3}}. \quad (2)$$

For our fiducial model, $M_{\text{bin}} = 2 \times 10^7 M_\odot$, $q = 1$ and $T = 4443 R_g/c$, which gives $\Delta \phi = 0.0165$, identical to what we found from our light curves. We note that this relation is less trivial for an eccentric binary, since the velocity depends on the nodal angle at which the BHs align and produce the SLF.

Inserting the fiducial model values, we find $\Delta i = 3°$, confirming our numerical result from Panel 5 of Fig. 2. In Fig. 3 we show the relations for the time interval $\Delta T = \Delta \phi T$ and the inclination window $\Delta i$ for a binary with mass $M_{\text{bin}} = 2 \times 10^7 M_\odot$ where we vary the period. As $T$ increases, which is equivalent to increasing the binary separation, the time interval increases, and the inclination window decreases. The spacing of the dip in our fiducial model is $\Delta \phi = 0.0165$. Converting this to physical time for a binary with mass $M_{\text{bin}} = 2 \times 10^7 M_\odot$, we find a spacing of 30 min. This puts a limit on the time cadence required by an observation (although repeated observations could be phase folded and reveal the dip in a more coarsely sampled flare). As a function of mass, the orbital period $T$ and $\Delta T$ both scale linearly with mass, making more massive binaries easier to resolve in time.

Observational limitations.—Ideally, individual SLFs should outshine the accretion-induced variability and...
The visibility of these features. A recent work [16] ray accretion flow, e.g., from hydrodynamic or GR magneto-less prominent (Ref.[26]). Using a physical model for the e.g., strongly Doppler-deboosted on one side of the image, their dips. If the image of the source is highly asymmetric, the function of the azimuthal position of an observer at and examined the time-averaged bolometric flux as a traced the emission from a binary in GRMHD simulations, and the duration of the dip is on the order of hours to days depending on the BH mass. The flux variability in AGN typically shows a stochastic red-noise power spectrum [43–46]. This is beneficial since the amplitude of the variability on the short timescales of the SLFs is lower. Additionally, phase folding can be used to average out stochastic variability.

Large time-domain surveys discovered dozens of candidate SMBHBs based on periodicity in optical light curves [6,7], although with sparsely sampled light curves. The estimated BH mass distribution of these candidates is skewed towards the higher BH masses of $10^8$--$10^{10}$ $M_\odot$. This widens the dip's duration from approximately an hour in our fiducial model to a few days (scaling only the BH mass). By increasing the period of the orbit, this interval would further widen (see Fig. 3) and exceed the five-day cadence for some candidates in [7].

There are some caveats to the detectability of SLFs and their dips. If the image of the source is highly asymmetric, e.g., strongly Doppler-deboosted on one side of the image, the secondary peak we observe in the light curve will be less prominent (Ref. [26]). Using a physical model for the accretion flow, e.g., from hydrodynamic or GR magneto-hydrodynamic (GRMHD) simulations, should help clarify the visibility of these features. A recent work [16] ray traced the emission from a binary in GRMHD simulations, and examined the time-averaged bolometric flux as a function of the azimuthal position of an observer at different latitudes. This is an estimate of the phase-folded light curve, and a hint of a dip is visible for nearly edge-on viewing angles (their Fig. 11). The deleterious effects of asymmetric Doppler boosting could also be mitigated if the plane of one (or both) minidisks is misaligned with the binary's orbital plane. Such misalignment, or related lateral tearing of the minidisks, is possible, if the circumbinary disk's angular momentum and the BH spin axis are misaligned [47–49], although this configuration is expected to be less common [50,51].

Another possible limitation is that the edge-on circumbinary disk can block the view to the event horizon. However, the binary orbital plane can be misaligned with respect to the circumbinary disk [48,52,53]. Furthermore, for thin AGN disks with accretion rates near but below the Eddington limit, with a separation of $\approx100$ $R_\text{g}$, the disk aspect ratio is $\lesssim0.03$ [54], so that the circumbinary disk would not obscure the flare from a coplanar binary.

In our Letter, we used temperature profiles that are in agreement with hydrodynamical simulations of a MBHB system (see, e.g., [30,31,32,33]). In these Letters, shocks are present which are shown to be dominant over viscous heating. The higher temperatures cause a peak in our spectra in the 1–10 keV band, which is higher than single BH systems which typically peak below 0.1 keV. If the effect of shock heating is overestimated and temperatures of the minidisks are in reality colder, then the ideal frequency to observe the dip would move to lower frequencies. In general, one would expect to see the dip as long as the spectral slope at the observed frequency is close to zero. This is because if the spectral slope is too steep it would result in strong Doppler boosting and deboosting in the minidisks which would make the emission region highly asymmetric, without a clear central shadow. Future hydrodynamical or MHD simulations should help resolve the uncertainty in the temperature profiles.

Finally, compact binaries exhibiting SLFs may be rare because they are short lived. For instance the GW-driven inspiral time in our fiducial model is a mere 1.5 yr ([55], see Eq. 28 in [54]). This is mitigated for larger BH masses since the inspiral time scales linearly with BH mass. For example, a $10^9$ $M_\odot$ BH would increase the inspiral time by a factor of 100. Additionally, wider binaries are substantially more long-lived since the inspiral time scales with $a_{\text{orb}}^4$. Wider orbits, which also have wider dips, should ensure that a substantial population is present that might have the right condition for future observations to detect SLFs with BH shadows.

Search strategies.—An estimation of the probability of observing SLFs was performed by [2], who estimated that 10% of the currently known SMBHB candidates potentially produce SLFs. These candidates typically have orbital periods of a year and masses of $\sim10^9$ $M_\odot$. The visibility of the dip is strongly limited by the orbital inclination, and decreases the probability to observe it. From our formula we
find, for the above high-mass candidates, a window of $\Delta i \approx 1^\circ$, corresponding to a 1% chance assuming isotropic inclinations. Since 150 candidates are known, one of these could be in the right inclination regime.

Since the probability of detection is low, we propose two strategies to find SLFs with BHS features. First, SLFs could be observable by current and future optical and near-infrared ground-based telescopes, designed for time-domain surveys. One option would be to perform follow-up observations of the already known SMBHB candidates with an approximately daily cadence. Additionally, the Rubin Observatory’s LSST is expected to identify 20–100 million quasars [56,57] down to a BH mass of a few $10^5 M_\odot$. The sheer number of potential sources increases the probability of finding AGNs with the right conditions to measure BHSs, Ref. [58] estimates that hundreds of potential SLFs could be found. The light curves of these quasars will also be sampled at a cadence of a few days, which should be suitable to identify many periodic sources and SLFs. Candidates identified in a large time-domain survey, such as LSST, could then be followed-up by x-ray telescopes, such as IXPE, XMM-Newton or ATHENA, to study the emission expected to be strongly dominated by the minidisks.

A second strategy is to perform electromagnetic followups of on-going SMBHB mergers discovered with LISA [59]. LISA is expected to detect SMBHBs at their late inspiral phase starting at periods of a few hours. At this moment, the LISA is expected to detect SMBHBs at their late inspiral of on-going SMBHB mergers discovered with LISA [59].

The light curves of these quasars will also be sampled at a cadence of a few days, which should be suitable to identify many periodic sources and SLFs. Candidates identified in a large time-domain survey, such as LSST, could then be followed-up by x-ray telescopes, such as IXPE, XMM-Newton or ATHENA, to study the emission expected to be strongly dominated by the minidisks.

Summary.—We presented numerical models for SLFs in the light curves of SMBHBs, based on GRRT simulations of thin disc emission in an approximate binary metric. If such a SMBHB system is observed close to edge-on, we find a double-peaked substructure in the SLF, where the spacing of the subpeaks is set by the angular size of the BHS. For our fiducial model the probability of detecting this dip is $\approx 3\%$, assuming isotropic inclinations. The BH tomography method proposed here would open a new window to discovering BHSs and characterizing their size. The features in the SLFs also yield independent constraints on the BH masses, properties of their emission morphology and kinematics, and the binary’s orbital parameters. Since this method does not require spatially resolving the source, it allows probing BHs which are out of reach for the EHT.

This research has made use of NASA’s Astrophysics Data System. Software: PYTHON [61,62], SCIPY [63], NUMPY [64], MATPLOTLIB [65], RAPTOR [27,28].

The authors thank Christiana Brinkerink, Anastasia Gvozdenko, Jeremy Schnittman, and Daniel D’Orazio for valuable comments and discussions during this project. J. D. is supported by a Joint Columbia and Flatiron Postdoctoral Fellowship. Research at the Flatiron Institute is supported by the Simons Foundation. We acknowledge support by NASA Grant NO. NNX17AL82G and NSF Grants NO. AST-1715661 and NO. AST-2006176. Computations were performed at SDSC maintained by Flatiron Institute’s SCC.
In principle, the periodically recurring SLFs could be discovered via a blind phase folding of the light curve, even if they are individually below the noise.