





# Lomb–Scargle Periodogram Struggles with Non-sinusoidal Supermassive Black Hole Binary Signatures in Quasar Lightcurves

Allison Lin<sup>1</sup>, Maria Charisi<sup>2,3,6</sup> , and Zoltán Haiman<sup>1,4,5</sup> <sup>1</sup> Department of Astronomy, Columbia University, New York, NY 10027, USA<sup>2</sup> Department of Physics and Astronomy, Washington State University, Pullman, WA 99163, USA<sup>3</sup> Institute of Astrophysics, FORTH, GR-71110, Heraklion, Greece<sup>4</sup> Department of Physics, Columbia University, New York, NY 10027, USA<sup>5</sup> Institute of Science and Technology Austria (ISTA), Am Campus 1, Klosterneuburg 3400, Austria

Received 2025 May 13; revised 2025 November 30; accepted 2025 December 5; published 2026 January 29

## Abstract

Supermassive black hole binary (SMBHB) systems are expected to form as a consequence of galaxy mergers. At subparsec separations, SMBHBs can be identified as quasars with periodic variability, with previous periodicity searches uncovering significant candidates. However, these searches focused primarily on sinusoidal signals, while theoretical models and hydrodynamical simulations predict that binaries produce more complex non-sinusoidal pulse shapes. Here we examine the efficacy of the Lomb–Scargle periodogram (LSP; one of the most popular tools for periodicity searches in unevenly sampled lightcurves) to detect periodicities with a sawtooth shape mimicking results of hydrodynamical simulations. We simulate idealized well-sampled lightcurves, lightcurves that mimic the data in the Palomar Transient Factory (PTF) analyzed in M. Charisi et al. (2016), and lightcurves that resemble our expectations for single-band data in the upcoming Legacy Survey of Space and Time (LSST) of the Rubin Observatory. We approximate quasar variability with a damped random walk (DRW) model, inject sinusoidal and sawtooth pulse shapes, and assess their statistical significance. We find that in the presence of red noise, the LSP detects a relatively low fraction of the sinusoidal signals ( $\sim 45\%$ ,  $\sim 24\%$ , and  $\sim 23\%$ , in the PTF-like, idealized, and LSST-like lightcurves, respectively). The fraction is significantly reduced for sawtooth periodicity (with only  $\sim 9\%$  in PTF-like and  $\sim 1\%$  in idealized and LSST-like lightcurves). These low recovery rates imply that previous searches have missed the large majority of binaries. They also have significant implications for the detection of SMBHBs in upcoming LSST necessitating the development of advanced tools that go beyond the simple LSP.

*Unified Astronomy Thesaurus concepts:* [Quasars \(1319\)](#)

## 1. Introduction

Observational evidence suggests that supermassive black holes (SMBHs) exist at the center of every massive galaxy (D. Richstone et al. 1998; J. Kormendy & L. C. Ho 2013). SMBHs play an important role in shaping their host galaxies, since the mass of SMBHs is correlated with several properties of the host. In addition, frequent mergers of galaxies naturally deliver two SMBHs to the nucleus of the postmerger galaxies, which can eventually form bound SMBH binaries (SMBHBs; M. C. Begelman et al. 1980). After the galaxies merge, the SMBHs sink toward the center of the postmerger galaxy through dynamical friction. As the SMBHs spiral closer toward each other, they form a near-Keplerian system. When they reach separations of a few parsecs, gravitational interactions with nearby stars and the gaseous circumbinary disk can continue to shrink the binary orbits. If these mechanisms are sufficiently effective to overcome the “final-parsec problem,” gravitational-wave (GW) emission takes over, accelerating the inspiral toward the merger.

The most massive systems can be detected at nanohertz GW frequencies with pulsar timing arrays (PTAs; S. Burke-Spolaor et al. 2019; S. R. Taylor 2021). In fact, all major PTA

collaborations, like the North American Nanohertz Observatory for Gravitational Waves, the European Pulsar Timing Array (EPTA) along with the Indian PTA, the Parkes PTA, the Chinese PTA and the MeerKAT PTA have recently detected evidence for a GW background (G. Agazie et al. 2023a; EPTA Collaboration et al. 2023; D. J. Reardon et al. 2023; H. Xu et al. 2023; M. T. Miles et al. 2025), which is likely produced by a population of SMBHBs (G. Agazie et al. 2023b; EPTA Collaboration et al. 2024). Individual binaries are expected to be resolved on top of this background in a few years (L. Z. Kelley et al. 2018; S. R. Taylor et al. 2020; B. Bécsy et al. 2022).

In addition to strong GWs, binaries are also expected to produce bright electromagnetic (EM) emission, since SMBHBs reach small separations surrounded by large amounts of gas, which accretes to the SMBHs and produces bright EM signatures. Therefore, binaries are promising sources for multi-messenger observations combining EM+GW data (B. Kocsis et al. 2006; Z. Haiman et al. 2009; L. Kelley et al. 2019; M. Charisi et al. 2022).

Given that binary systems are expected to form in gas-rich environments, several hydrodynamic simulations have explored such systems embedded in gaseous disks. These studies suggest that binaries may manifest as quasars with periodic variability. More specifically, simulations demonstrate that torques exerted by the binary expel the gas from the central-most region and create a low-density cavity. As the orbital motion of the binary perturbs the cavity edge, the SMBHBs pull streams of gas inwards, leading to periodic mass

<sup>6</sup> Author to whom correspondence should be addressed.



transfer from the disk onto the binary components, which translates into periodic brightness fluctuations (e.g., J. R. Westernacher-Schneider et al. 2022). Periodic variability can also arise due to Doppler boost from gas bound to the individual SMBHs, and from self-lensing effects (see T. Bogdanovic et al. 2021 and D. J. D’Orazio & M. Charisi 2023 for recent reviews).

Many observational studies have searched for this EM signature in time-domain surveys and identified several candidate SMBHBs as quasars with periodically varying optical emission. Note that this approach associates the bright quasar phases with SMBHBs, which is justified as quasars are often thought to be triggered by mergers (G. Kauffmann & M. Haehnelt 2000). However, the search requires large surveys because the binary lifetime due to GW inspiral would be of order  $10^5$  yr, which is  $\sim 10^{-3}$  of the typical expected bright quasar phase (Z. Haiman et al. 2009). This implies that even if most quasars are binaries, we would expect to find them only in a small fraction  $\lesssim 10^{-3}$  of quasars. Suitably large systematic searches in the Catalina Real-Time Transient Survey (CRTS; M. J. Graham et al. 2015), the Palomar Transient Factory (PTF; M. Charisi et al. 2016), the Panoramic Survey Telescope and Rapid Response System (PanSTARRS; T. Liu et al. 2019), the Dark Energy Survey (Y.-C. Chen et al. 2020), the Zwicky Transient Facility (ZTF; Y.-J. Chen et al. 2024), the Wide-field Infrared Survey Explorer (D. Luo et al. 2025) and Gaia (P. Huijse et al. 2025) have revealed over 200 candidates to date. Additional candidates were also identified individually in lightcurves from long-term observing campaigns (B.-Z. Dai et al. 2006; E. Bon et al. 2016; Y.-R. Li et al. 2019).

However, quasars exhibit red noise variability, which can mimic periodic signals, introducing false positives, while at the same time hindering the identification of genuine periodicity (S. Vaughan et al. 2016; C. A. Witt et al. 2022; M. C. Davis et al. 2024; J. Robnik et al. 2024). This is exacerbated by limitations in our understanding of the stochastic variability of quasars, which in previous periodicity searches was modeled with a damped random walk (DRW) model (B. C. Kelly et al. 2009; S. Kozłowski et al. 2010; C. L. MacLeod et al. 2010; C. J. Burke et al. 2021; K. L. Suberlak et al. 2021) and the fact that binaries typically have relatively long periods, allowing only two to three cycles to be observed (S. Vaughan et al. 2016). Therefore, additional signatures, such as spectral line shifts indicative of radial motion (M. Eracleous et al. 2012), multiwavelength Doppler boost (D. J. D’Orazio et al. 2015; M. Charisi et al. 2018; C. Xin et al. 2020), distorted radio jets (N. Roos et al. 1993; A. Caproni et al. 2013), or atypical X-ray spectra (M. L. Saade et al. 2020), are required to boost the significance of each candidate. We refer the reader to A. De Rosa et al. (2019) and D. J. D’Orazio & M. Charisi (2023) for reviews on binary signatures and observational searches.

To date, most searches for periodic variability have focused on sinusoidal periodicity, which is the simplest periodic signal. A widely used method for detecting periodic signals in unevenly sampled data is the Lomb–Scargle periodogram (LSP; N. R. Lomb 1976; J. D. Scargle 1982 or J. T. VanderPlas 2018 for a recent review). The LSP (along with other Fourier-based techniques) is particularly effective in identifying sinusoidal variations, but may be less sensitive to signals that deviate from this shape. Such signals can be common in binaries, e.g., when the periodicity arises from periodic accretion or the binary is Doppler-modulated along an eccentric

orbit (P. C. Duffell et al. 2020; B. X. Hu et al. 2020; J. Zrake et al. 2021; J. R. Westernacher-Schneider et al. 2022). In addition, the LSP may be a suboptimal statistic for periodic signals in the presence of nonwhite stochastic noise, like in quasars (J. Robnik et al. 2024).

In this paper, we investigate the ability of the LSP to detect non-sinusoidal periodicity in quasar lightcurves under different survey conditions, focusing specifically on sawtooth periodic patterns, which approximate the pulse shapes predicted by hydrodynamic simulations (P. C. Duffell et al. 2020; J. R. Westernacher-Schneider et al. 2022). To explore this effect, we generate idealized and realistic PTF-like mock lightcurves, exhibiting sawtooth periodic signals, with realistic quasar noise modeled as a DRW, as in previous quasar periodicity searches. Since the upcoming Legacy Survey of Space and Time (LSST) of the Rubin Observatory is expected to play a significant role in quasar periodicity searches (L. Z. Kelley et al. 2021; C. Xin & Z. Haiman 2021; F. Cocchiara et al. 2024), we also examine single-band LSST-like lightcurves. We also emphasize that the underlying variability of quasars is modeled with a DRW to reflect the common practice in previous periodicity searches, but this model does not fully capture quasar variability and should be viewed as an approximation of quasar noise. For comparison, we repeat the analysis for sinusoidal signals with similar properties.

This paper is outlined as follows: In Section 2, we describe the mock quasar lightcurves and the statistical methods used to determine the significance of the periodicity. In Section 3, we report the fraction of detected periodic (sinusoidal and sawtooth) signals in idealized, PTF-like, and LSST-like lightcurves. In Section 4, we discuss our findings, and in Section 5, we summarize our results and offer our conclusions.

## 2. Methods

Our objective is to study the recovery of sawtooth periodicity and compare it with the recovery of sinusoidal signals with similar periods and amplitudes under different survey conditions. We explore the detection of such signals in three cases: highly idealized lightcurves with continuous regular sampling, lightcurves that resemble the ones observed in PTF, and lightcurves that represent our expectations for single-band, well-sampled LSST lightcurves, which are expected to play a crucial role in quasar periodicity searches in the upcoming years. For the detection, we follow the statistical analysis in M. Charisi et al. (2016), which, like nearly all similar periodicity searches, is based on the LSP. Below, we describe in detail our quasar lightcurve simulations and our methodology for detecting periodic signals and assessing their statistical significance.

### 2.1. Quasar Sample

In order to evaluate the ability of the LSP to detect non-sinusoidal signals in a realistic setting, we examine the sample of quasars detected by PTF and analyzed in M. Charisi et al. (2016). By closely following their statistical analysis for periodicity detection, this study provides one additional benefit; it allows us to assess whether non-sinusoidal periodicity would have been detected by the pipeline if present in the dataset, or such signals would have been missed, and to what extent.

We remind the reader that PTF was one of the early time-domain surveys (2009–2012), the precursor to the ongoing ZTF. It covered around  $3000 \text{ deg}^2$  with a  $5\sigma$  limiting

magnitude of approximately 20.5 in the  $R$  band. Its observing strategy was designed to capture transient phenomena, and thus, the PTF lightcurves are characterized by epochs of very dense temporal sampling (e.g., when an interesting transient was detected in the field) followed by large gaps. For a detailed discussion about the survey and the scientific goals of PTF, see N. M. Law et al. (2009) and A. Rau et al. (2009).

In this paper, we start from the sample of 35,383 quasars analyzed in M. Charisi et al. (2016). This sample was extracted from the Half Million Quasars catalog, and then focusing only on the spectroscopically confirmed quasars with relatively good lightcurves, i.e., lightcurves that have at least 50 nights of observations. Of those, we select a subset of 12,400 that have virial SMBH mass estimates measured from their broad emission lines with spectra from the Sloan Digital Sky Survey (SDSS) Data Release 7 (K. N. Abazajian et al. 2009). This selection is necessary for the estimation of the DRW parameters, as explained below (Section 2.2.1).

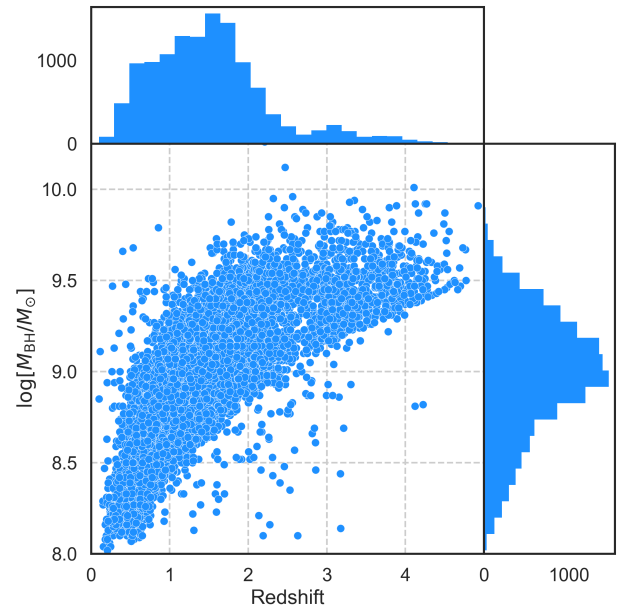
The PTF lightcurves for this subset of quasars show a variety in terms of baselines and total number of observations (e.g., see Figure 1 in M. Charisi et al. 2016). On average, the baseline for these lightcurves is  $\sim 3.2$  yr with a mean cadence of  $\sim 12$  days, a mean magnitude of 19.1, and mean photometric errors of  $\sim 0.05$ .

## 2.2. Mock Lightcurves

The lightcurves we simulate are composed of realistic quasar noise described by a DRW model (Section 2.2.1),<sup>7</sup> periodic signals of two types to emulate two broad categories of binary behavior (Section 2.2.2), and photometric errors. For each of the 12,400 observed quasars in the sample, we generate a set of six mock lightcurves exploring two periodic signals: (a) sinusoidal periodicity and (b) sawtooth periodicity, and three observing scenarios: (1) idealized lightcurves, (2) PTF-like lightcurves, and (3) LSST-like lightcurves. For each set of six simulations, we inject the exact same parameters for the periodic signal (same period, amplitude, and phase) and the same noise (i.e., the same DRW parameters  $\sigma$  and  $\tau$  and even the same noise realization).

For the idealized lightcurves, we simulate a time series with a baseline equal to the respective PTF lightcurve and photometric errors drawn from a Gaussian distribution with zero mean and standard deviation equal to the mean photometric error of the specific PTF lightcurve, but we adopt a daily cadence. Ground-based time-domain surveys cannot achieve such high-quality lightcurves, because uneven sampling and gaps are inevitable (e.g., missing observations due to weather or proximity of the source to the Sun, etc). However, this set of idealized simulations (and the comparison with more realistic PTF-like lightcurves) allows us to isolate the effect the data quality has in detecting various periodic signals from other factors like the intrinsic quasar noise.

For the PTF-like lightcurves, we take the continuous time series (which includes periodicity and DRW noise, but not photometric errors) from the above step and downsample them at the observed times of the respective PTF source. In this case, we add photometric errors drawn from a Gaussian



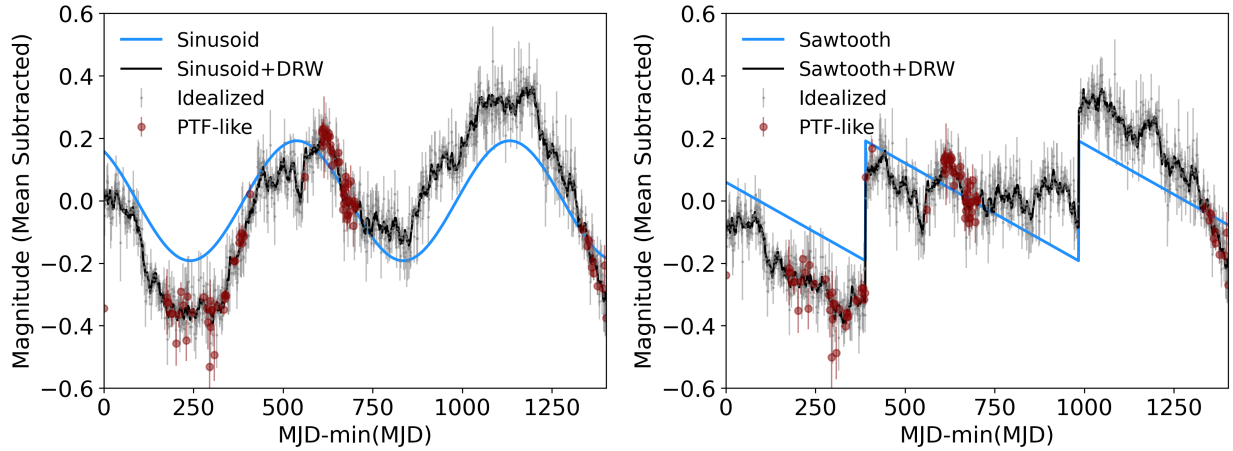
**Figure 1.** SMBH mass vs. redshift for the 12,400 quasars in our sample. The marginalized distributions of the two quantities are also shown. The quasars in our sample span a wide range in SMBH mass and redshift, representative of the entire quasar population.

distribution with zero mean and standard deviation equal to the photometric uncertainty of each point in the PTF lightcurve.

Finally, we explore the periodicity recovery in longer lightcurves, which resemble the best-sampled single-band LSST lightcurves (e.g., what we expect for lightcurves in the  $r$  or  $i$  bands). We emphasize that LSST, with its unprecedented data quality and a sample of quasars exceeding 20 million, will be a key driver in the discovery of SMBHB candidates. For those, we set the baseline to 10 yr, equal to the nominal duration of the survey. We choose a 5 day cadence, as expected for the Wide-Fast-Deep mode of LSST, but instead of evenly sampled data, we perturb the regular temporal sampling with a Gaussian of zero mean and standard deviation of 1 day. This choice allows us to more closely match the observing strategy of LSST, while avoiding strong aliasing effects from very regular sampling. We also exclude 4 months of observations every year to mimic the inevitable annual gaps. We add photometric errors drawn from a Gaussian distribution with zero mean and standard deviation of 0.01, which corresponds to the level of photometric uncertainty for sources with a mean magnitude of 21 (Ž. Ivezić et al. 2019). Finally, we inject the same periodic signals and the same DRW realizations (creating them with the same random seed) as in the PTF lightcurves. We note that since the baselines are longer in the LSST-like lightcurves, we could explore longer periods. By simulating the same periods as in the shorter PTF lightcurves, we can explore the effect of observing more cycles within the data.

Therefore, for each observed lightcurve, we generate six mock lightcurves with the same injected periodicity and noise properties. We analyze each of those lightcurves independently with a periodicity detection pipeline similar to the one presented in M. Charisi et al. (2016; Section 2.3). Figure 2 shows such a set of four simulations, idealized and PTF-like, and the components that produce the final lightcurve, with sinusoidal periodicity on the left panel and sawtooth periodicity on the right. With blue lines, we demonstrate the periodic signals, and with black lines, the aggregate time series after the DRW noise

<sup>7</sup> The DRW model, albeit successful, does not capture the full range of quasar variability. However, it provides a valuable benchmark for the effect of the red noise. In Section 4, we discuss some limitations of this choice of quasar variability.



**Figure 2.** Examples of the mock periodic lightcurves we generate to test the detectability of such signals with the LSP, with a sinusoid injected on the left panel and a sawtooth signal injected on the right panel. The noiseless periodic signals are shown with solid blue lines. On top of those, we add DRW noise (the same realization in both panels), shown with the black continuous curve. Idealized lightcurves sampled daily with photometric errors are shown with gray points and the respective error bars, while the PTF-like lightcurves shown in red are obtained by downsampling the black curve to the observed times and adding PTF-like photometric errors. The simulations are sampled at the observed times of quasar SDSS J085234.09+330100.7 with an injected period of 595 days, amplitude 0.19 mag, and DRW variability with parameters  $\sigma = 0.27$  mag and  $\tau = 1246$  days.

is included. Finally, the gray points with the error bars represent the idealized lightcurves after the photometric errors have been added, while the red points with the respective error bars show the PTF-like lightcurves produced from downsampling the black curve at the PTF time stamps and subsequently adding Gaussian errors from the PTF measurements.

### 2.2.1. Quasar Variability

In order to model realistic quasar lightcurves, we need to incorporate the intrinsic stochastic variability of quasars. The optical variability of quasars is successfully described by a DRW process (B. C. Kelly et al. 2009; S. Kozłowski et al. 2010; C. L. MacLeod et al. 2010; C. J. Burke et al. 2021; K. L. Suberlak et al. 2021). We emphasize that, albeit successful and widely used in previous periodicity searches, this model is likely an incomplete description of quasar variability (e.g., see Section 4). The choice of quasar noise model can affect our results, and thus, future studies should consider a more flexible description of the underlying variability (such as higher-order continuous auto-regressive mean average (CARMA) models; B. C. Kelly et al. 2014), but the DRW model provides a useful baseline for red noise.

The power spectral density (PSD) of the DRW is frequency dependent (red noise) at high frequencies, and becomes flat (white noise) at low frequencies. The PSD as a function of frequency  $f$  can be described as follows:

$$\text{PSD}(f) = \frac{4\sigma^2\tau}{(1 + 4\pi^2f^2\tau^2)}, \quad (1)$$

where  $\sigma^2$  is the variance of the lightcurve and  $\tau$  is the damping timescale, roughly corresponding to the autocorrelation length of the lightcurve or to the (inverse of) the break frequency in the PSD.

For each noise simulation, we first need to determine a set of DRW parameters,  $\sigma$  and  $\tau$ . In M. Charisi et al. (2016), the best-fit  $\sigma$  and  $\tau$  were determined for each lightcurve by maximizing the DRW likelihood using *Javelin*, a public software, to sample the likelihood with Markov Chain Monte Carlo methods (Y. Zu et al. 2013). Since these calculated  $\sigma$  and  $\tau$

values are no longer accessible, and that method resulted in biased estimates of the  $\tau$  parameter (see Figure 8 and Section 4.1 in M. Charisi et al. 2016), we adopt a simpler alternative method to estimate these parameters.

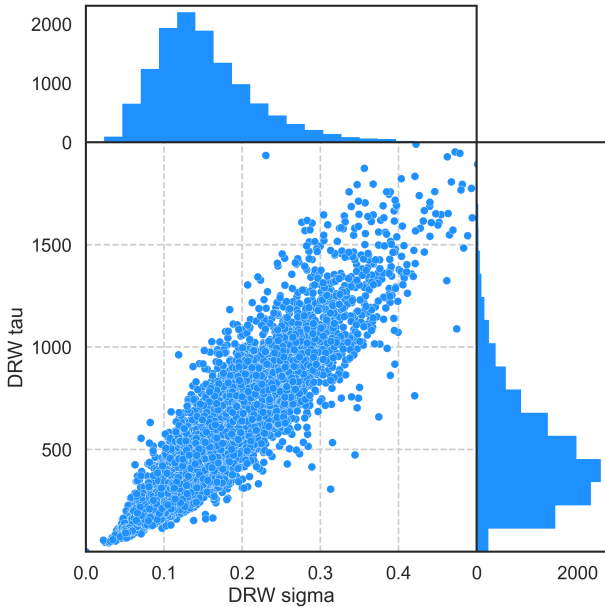
In particular, we use empirical scaling relationships from C. L. MacLeod et al. (2010), yielding  $\sigma$  and  $\tau$  from the properties of each individual quasar, as follows:

$$\log(\tau_{\text{RF}}) = 2.4 + 0.17 \log\left(\frac{\lambda_{\text{RF}}}{4000 \text{ \AA}}\right) + 0.03(M_i + 23) + 0.21 \log\left(\frac{M_{\text{BH}}}{10^9 M_{\odot}}\right), \quad (2)$$

$$\log(\sqrt{2\sigma}) = -0.51 - 0.479 \log\left(\frac{\lambda_{\text{RF}}}{4000 \text{ \AA}}\right) + 0.131(M_i + 23) + 0.18 \log\left(\frac{M_{\text{BH}}}{10^9 M_{\odot}}\right), \quad (3)$$

where  $\tau_{\text{RF}}$  represents the characteristic timescale  $\tau$  in the quasar's rest frame,  $\lambda_{\text{RF}}$  is the effective wavelength of the  $R$  band ( $\lambda_R = 6516.05 \text{ \AA}$ )<sup>8</sup> in the rest frame of the source,  $M_i$  is the absolute  $i$ -band magnitude,  $k$ -corrected to  $z = 2$ , and  $M_{\text{BH}}$  is the SMBH mass. Note that these relationships were used as priors when fitting the DRW parameters with *Javelin* in M. Charisi et al. (2016). For the above  $\sigma$  and  $\tau$  calculations, the black hole mass and the  $i$ -band luminosity of each quasar are required. For this reason, as explained in Section 2.1, we only analyze the subset of 12,400 quasars that have SMBH mass estimates and absolute  $i$ -band magnitudes, using a catalog from SDSS Data Release 7 (K. N. Abazajian et al. 2009). Figure 3 illustrates the distribution of the calculated  $\tau$  and  $\sigma$  values using Equations (2) and (3), respectively, for the 12,400 PTF quasars.

<sup>8</sup> PTF observed in the optical  $g$  and  $R$  bands, but we only analyze the  $R$ -band lightcurves, as in M. Charisi et al. (2016), because the  $g$ -band lightcurves are typically very sparse.



**Figure 3.** Values of the DRW parameters  $\sigma$  vs.  $\tau$  estimated based on their SMBH mass and  $i$ -band magnitude (C. L. MacLeod et al. 2010) for the 12,400 quasars in our sample. The marginalized distributions of the two quantities are also shown.

For the mock quasar lightcurves, we follow the procedure from J. Timmer & M. König (1995), as implemented in the `astroML` package (Ž. Ivezić et al. 2014), and generate evenly sampled lightcurves with a temporal spacing of  $\Delta t = 1$  day. As mentioned in Section 2.1, the baseline of each lightcurve is set to match the baseline of the respective observed lightcurve. For the PTF-like lightcurves, we follow a similar procedure as in M. Charisi et al. (2016) to downsample the idealized lightcurves by removing data points to match the observed pattern. For the LSST-like lightcurves, we generate the same realization of noise (using the same random seed), initially sampled daily for 10 yr and then downsampled at the expected observing pattern of LSST (e.g., see Section 2.2).

### 2.2.2. Binary Signals

For the simulations of the null hypothesis, required to assess the statistical significance of the injected periodic signals (see Section 2.3), the simulations include only DRW variability, as outlined in the steps above. To simulate binary signals, we then add periodic variability. We inject two types of periodic signals, sinusoidal and sawtooth, in the following forms:

$$s_{\text{sine}}(t) = A \cdot \sin(2\pi \cdot t/P + \phi_0) \quad (4)$$

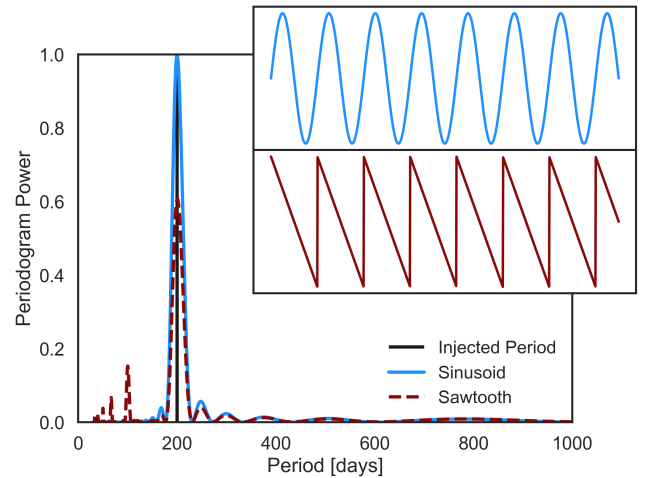
and

$$s_{\text{sawtooth}}(t) = A \cdot \text{sawtooth}(2\pi \cdot t/P + \phi_0), \quad (5)$$

where  $A$  is the amplitude in magnitude,  $\phi_0$  is the phase, and  $P$  is the period. The sawtooth waveform is characterized by an asymmetric sharp rise and linear decrease.

For the injected periodic signals, we explore a broad but realistic parameter space. In particular, for the periodic signals, we draw parameters from uniform distributions, from the following range of values:

1. Amplitude  $A$ :  $[0, 0.5)$ .
2. Period  $P$ :  $[0, P_{\text{max}}]$ .
3. Phase  $\phi_0$ :  $[0, 2\pi)$ .



**Figure 4.** The LSP of a noiseless sinusoid (blue) and a sawtooth (red), both with a period of 200 days, and an amplitude of 0.4 shows peak power of 1 and 0.61, respectively.

Here,  $P_{\text{max}}$  is the maximum period we can probe in each lightcurve, defined by its baseline so that at least 1.5 cycles of periodicity are observed, as in M. Charisi et al. (2016). That is,  $P_{\text{max}} = T_{\text{data}}/1.5$ , where  $T_{\text{data}}$  is the baseline of the respective PTF lightcurve.<sup>9</sup> Since we randomly draw the parameters for each lightcurve and the baselines vary, the high end of the period distribution is not as well sampled.

### 2.3. Periodicity Detection

Our goal is to assess whether the LSP is effective in detecting periodic signals that deviate from sinusoidal (e.g., sawtooth signals) in quasar data, exploring several scenarios for lightcurve quality. For our analysis, we use the generalized LSP from the `astroML` package (Ž. Ivezić et al. 2014). We follow the same procedure as in M. Charisi et al. (2016) and calculate the LSP for 1000 trial frequencies on a logarithmic grid with a minimum frequency of  $f_{\text{min}} = 1.5/T_{\text{data}}$ , and a maximum frequency of  $f_{\text{max}} = 1/T_{\text{min}}$ , with  $T_{\text{min}} = 30$  days. In Figure 4, we illustrate how the LSP performs on two noiseless idealized lightcurves with sinusoidal and sawtooth-shaped periodic signals, with an amplitude of 0.4 mag and a period of 200 days. We see that the peak of the periodogram is reduced by  $\sim 40\%$  for the sawtooth signal, while some secondary peaks at lower periods appear. These peaks are not from aliasing, because they are not present in the LSP of the sinusoidal signal, but rather simply additional Fourier components of the sawtooth pattern.

Even though the false alarm probability of an LSP peak can be calculated analytically in the presence of white noise (e.g., see R. V. Baluev 2008; C. Koen 2021), these approximations are not applicable in the presence of other types of stochastic variability (M. Süveges et al. 2015; S. Vaughan et al. 2016). Here, we assess the periodicity significance of the injected signals, again following M. Charisi et al. (2016). We simulate DRW-only lightcurves representing the null hypothesis, with  $\sigma$  and  $\tau$  chosen as described above, and compute the distribution of the periodogram peak heights in these mock lightcurves.

<sup>9</sup> Based on the requirement of observing 1.5 cycles within the data, in the LSST-like lightcurves, we could have explored longer periods. However, we chose to explore the same periods as in the PTF-like data to isolate the effect of observing more cycles.

Note that in a realistic scenario, the values of the DRW parameters would not be known, but our goal here is to determine the limitations of using a suboptimal statistic (like the peak of the LSP), and so we assume we know the correct DRW parameters to simplify the analysis. We compare the observed peaks of the injected sine+DRW and saw+DRW lightcurves against the observed peaks of DRW-only simulations.

To account for the fact that the DRW noise is frequency dependent and thus the peaks in the periodogram do not arise with the same probability (or power) at every frequency, the test statistic needs to be compared with the null distribution in a small range of frequencies (i.e., we cannot observe a peak at low frequencies and compare with the null distribution at high frequencies and vice versa). For this, we divide our frequency grid into 25 different bins, and for every bin, we calculate the false alarm probability of the maximum observed peak in that bin with the distribution of maximum peak heights of the simulated noise in the same bin. For each quasar, we simulate 100,000 DRW-only time series to account for the look-elsewhere effect, from the fact that the candidates were selected from a sample of 12,400 lightcurves and from using 25 frequency bins (see M. Charisi et al. 2016 for more details).<sup>10</sup> We consider the periodicity to be statistically significant if and only if a peak in any of the 25 frequency bins is higher than any of the peaks over the 100,000 DRW realizations in the same frequency bin.

### 3. Results

We assess the ability of the LSP to detect sinusoidal and sawtooth periodicity in quasar data. For this, we simulate 12,400 mock lightcurves with idealized, PTF-like, and LSST-like observing patterns, injecting periodic signals of both types with a variety of parameters and realistic quasar noise. Then we assess the statistical significance of identified periodograms peaks with DRW simulations. We select as significant the ones that had periodogram peaks higher than the ones produced by 100,000 noise realizations. Below, we report our findings.

#### 3.1. Periodic Signal Recovery Rates

Among the 12,400 simulated sinusoidal signals, 3020 (24.4%) are detected in the idealized lightcurves, 5540 (44.7%) in the PTF-like simulations, and 2809 (22.6%) in the LSST-like lightcurves. Interestingly, and somewhat counterintuitively, the recovery rate of sinusoidal signals is higher in the PTF-like lightcurves. This is likely attributable to the underlying stochastic variability of quasars, as we discuss further in Section 4 below. We see that even though the LSP is one of the most prominent tools to detect sinusoidal periodicity in sparse lightcurves, it underperforms in quasar periodicity searches, missing  $\sim 55\%$  of the sinusoidal signals in the PTF-like lightcurves and an even larger majority of  $>75\%$  in the idealized and LSST-like lightcurves.

The reason for this poor performance has been identified by J. Robnik et al. (2024), who recently demonstrated that the periodogram peak is a suboptimal test statistic for this case. This is because its underlying assumption is white noise, while

<sup>10</sup> In M. Charisi et al. (2016), the false alarm probability was assessed with 250,000 simulations of DRW noise. However, since here we only analyze approximately one-third of that sample, we adjust the number of mock realizations to 100,000 to roughly reflect the smaller size analyzed.

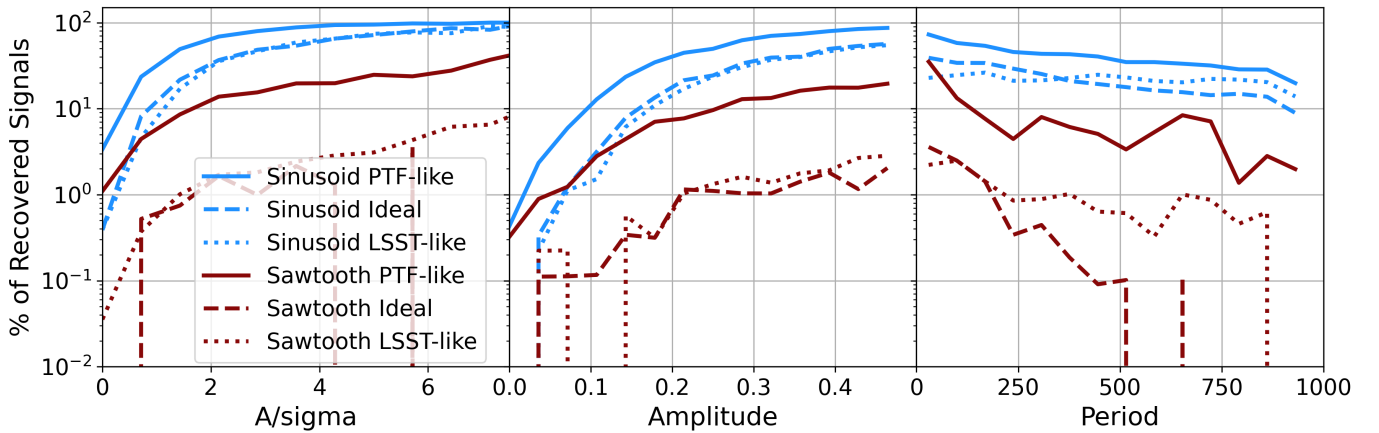
**Table 1**  
Recovery Rate for Sinusoidal and Sawtooth Periodic Signals and for Idealized, PTF-like, and LSST-like Lightcurves

	Sinusoid	Sawtooth
Idealized	$\sim 24\%$	$\sim 1\%$
PTF-like	$\sim 45\%$	$\sim 9\%$
LSST-like	$\sim 23\%$	$\sim 1\%$

quasars have red noise variability. As a result, the best-fit Fourier coefficients, which are selected based on  $\chi^2$  values that assume the incorrect noise, tend to be biased. We emphasize that the stochastic variability can still be properly taken into account, e.g., by using red noise simulations to calculate the false alarm probability, as in M. Charisi et al. (2016), but the test statistic itself is less sensitive in identifying these signals. For instance, using the Bayes factor (which explicitly incorporates the red noise) as the test statistic results in a higher true positive rate for the same false positive rate (see Figure 6 in J. Robnik et al. 2024). Similarly, using the Bayesian information criterion to select sinusoidal periodicity on top of DRW variability (i.e., including the red noise in the test statistic), C. A. Witt et al. (2022) found higher overall recovery rates of  $\sim 60\%$  in both CRTS-like and LSST-like simulated lightcurves.

The recovery rate of the sawtooth signals is significantly worse, with only 103 of the 12,400 (0.8%) injections recovered in the idealized simulations, 1150 (9.3%) in the PTF-like lightcurves, and 139 (1.1%) in the LSST-like data. These rates are summarized in Table 1, along with the results for sinusoidal signals. As with the sinusoidal signals, we detect more of the injected sawtooth signals in the PTF-like sample compared to the idealized and LSST-like case. The large reduction in the recovery rate between sinusoids and sawtooth signals is not surprising, given that a single peak in the LSP effectively uses a sinusoidal template and thus is optimized only for such signals (e.g., see Figure 4). Our results demonstrate that the LSP is very ineffective at finding non-sinusoidal periodicity. It also implies that if sawtooth signals were present in the PTF sample, it is highly unlikely that the previous analysis by M. Charisi et al. (2016) would have uncovered them. Although we do not explicitly demonstrate this here, the same conclusion very likely applies to the other previous periodicity searches listed in the Introduction, since most of those were also based on LSP. The exception is M. J. Graham et al. (2015), where a combination of wavelets and the autocorrelation function (ACF) was employed. However, we suspect that their analysis would also have been insensitive to strongly non-sinusoidal periodicities, because of the chosen shape of wavelets (modulated plane waves) and the ACF (corresponding to a truncated Fourier series).

Finally, we examine whether the period of the detected signals is correctly identified. For this, we check whether the recovered signals have periods within 15% of the simulated value. For the sinusoids, we find that the vast majority, i.e., 98.8% in the idealized lightcurves, 97.0% in the PTF-like lightcurves, and 99.7% in the LSST-like lightcurves, have periods similar to the injected value. The fraction is slightly lower for the sawtooth signals, with 91.3% in the idealized lightcurves, 79.6% in the PTF-like lightcurves, and 87.8% in the LSST-like lightcurves. As seen in Figure 4, the



**Figure 5.** Recovery rates of periodic signals as a function of their properties. The left panel shows the dependence on the SNR of the signal, quantified as the ratio of the amplitudes of the injected periodic signal vs. the DRW noise ( $A/\sigma$ ). The middle panel shows the dependence on the amplitude, and the right panel shows the dependence on the period. The blue lines show the recovery rates for sinusoidal signals, while the red lines show the recovery rates for sawtooth signals. Solid lines represent samples with PTF-like lightcurves, dashed lines are for idealized lightcurves, and dotted lines are for LSST-like simulations.

periodogram of the sawtooth signals shows additional peaks beyond the injected period, which in combination with the red noise, can lead to misidentification of the periods. In the PTF-like lightcurves, the uneven sampling can also contribute to this transfer of power between frequencies (known as aliasing). However, the misidentification of periods is overall relatively rare, especially for the sinusoids, for which the LSP is optimized.

### 3.2. Dependence on Signal Parameters

We next explore how the recovery rates depend on the parameters of the injected periodic signals, in combination with the DRW noise properties. First, we quantify the signal strength with the ratio of the amplitude of the periodic signal over the variance of the DRW noise,  $A/\sigma$ , which provides a rough estimate of the signal-to-noise ratio (SNR). In the left panel of Figure 5, we show the recovery rates as a function of their SNR, with blue lines for sinusoidal signals and red for sawtooth signals. Solid lines represent the PTF-like simulations, while dashed lines indicate the idealized lightcurves, and dotted lines the LSST-like simulations. We see that the recovery rate increases as the SNR increases, as expected, i.e., stronger signals are easier to detect compared to signals buried in the DRW noise. The trend is roughly monotonic for both types of signals and quality of lightcurves, with the exception of sawtooth signals in idealized lightcurves, where we observe some fluctuations at the relatively poorly sampled high SNR regime due to small number statistics.

Next, we consider the recovery rates in terms of the properties of the periodic signal alone. The middle panel of Figure 5 shows the recovery rates as a function of the injected amplitude of the periodic signal, with the same colors and line styles as above. Again, we observe monotonically increasing rates for both types of periodic signals and for all three types of lightcurves, i.e., signals with higher amplitudes are significantly easier to detect compared to low-amplitude signals, as expected. Even in the best-case scenario, that is, very strong signals with  $A = 0.5$ , the algorithm does not detect all the signals, reaching a maximum recovery rate of 80%. This also demonstrates the limitations of the LSP to detect periodic signals in the presence of red noise.

Finally, we explore the recovery rates as a function of the period. We observe decreasing trends with the period in all six cases we examined. In particular, for the sinusoids, the decrease is monotonic, whereas for the sawtooth signals, we see a sharp decline and the signals become practically undetectable for periods longer than 250 days for the idealized lightcurves, with fluctuations at the level of a few percent for the PTF-like and the LSST-like simulations. This trend is also expected, since a shorter period signal allows for more cycles to be observed within the available baseline, thus resulting in stronger periodogram peaks and a higher probability of detection, whereas longer-period signals, with only two to three observed cycles, are very challenging to detect. In this regime, the periodogram peaks are also relatively broad, resulting in significant uncertainty in the period estimation. Additionally, the DRW noise amplitude is lower at higher frequencies (see Equation (1)). This limitation is significant because in a population of binaries, we expect many more long-period binaries than short-period ones. Therefore, the samples of candidates in this very interesting part of the parameter space are likely highly incomplete. Again, even in the best-case scenario, of periods of 30 days, which is the minimum searched period in this study, the recovery rate does not reach 100%, underscoring the challenges of detecting real signals. Similar trends for the recovery rates with respect to the amplitude (C. A. Witt et al. 2022; M. C. Davis et al. 2024) and the period of the signal (C. A. Witt et al. 2022) have been seen before.

## 4. Discussion

Our study focuses on the detection of non-sinusoidal periodic signals in quasar lightcurves using the LSP under different scenarios for lightcurve quality. Such signals are expected to be common in binary lightcurves, but may have been missed in previous searches, which focused primarily on sinusoidal periodicity. To quantify the expected detection rates, we injected sinusoidal and sawtooth-shaped periodic signals in realistic PTF-like, idealized, and LSST-like lightcurves, which are expected to be crucial in the near future. Using a statistical analysis similar to the one employed in M. Charisi et al. (2016), which is based on the LSP, we

compute the recovery rate for both signals and for all three types of lightcurve data quality.

Our main conclusion is that the LSP struggles to detect the injected periodicity, detecting a maximum of  $\sim 45\%$  of the sinusoidal signals and only up to  $\sim 9\%$  of the sawtooth signals in the PTF-like lightcurves. We conclude that previous searches, including the one in M. Charisi et al. (2016), must have missed a significant fraction of periodic signals. This is the case even for sinusoids, but the searches likely missed the vast majority of binaries if they produce sawtooth-shaped periodicity, as hydrodynamical simulations suggest. Caution is required in generalizing this conclusion, because the recovery rates depend on the specific periodicity detection method, precise pulse shape, as well as on the data quality. However, all previous searches used Fourier algorithms, which, like the LSP, are optimized for sinusoidal signals.

Our analysis also returned a surprising (to us at least) finding; the LSP performs better in the PTF-like sample compared to the idealized or LSST-like ones, both for sinusoidal and sawtooth periodicity, although with significantly lower overall recovery rates for the latter. One would expect that the recovery rate would be higher in lightcurves with better quality and longer baselines (for the LSST-like simulations). However, as illustrated in Figure 2, the inclusion of DRW noise results in lightcurves that significantly deviate from the simple periodic behavior. The LSP, which performs a weighted least-squares fit for sinusoid amplitudes at each searched frequency, will struggle to produce a good fit, resulting in a lower/less significant peak. However, if one considers the sparser PTF-like lightcurve, it is easier to construct a sinusoidal curve (e.g., by adjusting the mean or the amplitude) that passes through the red points. We conclude that in the PTF-like lightcurves, the sparse sampling can in many cases be advantageous, i.e., if by chance the observed epochs sample close to the maxima and minima of the lightcurve. On the other hand, we caution that here we considered only real signals and did not consider the case of false detections, which presents a major challenge in quasar periodicity searches. This is likely exacerbated by the lower quality of lightcurves (e.g., see the comparison of CRTS-like versus LSST-like simulations in C. A. Witt et al. 2022).

Moreover, our study showed that for LSST-like lightcurves, the LSP detects  $<25\%$  of the sinusoidal signals and  $\sim 1\%$  of the sawtooth periodicity. This is a significant finding for periodicity searches in the upcoming LSST data, clearly demonstrating that the development of tailored periodicity tools that explicitly account for red noise variability is necessary. As we approach the era of large time-domain surveys, with the Rubin Observatory (Ž. Ivezić et al. 2019) set to begin operations in just a few months, it is evident that new methods for detecting periodicity in quasars are required. LSST will provide an unprecedented dataset with  $\sim 20\text{--}100$  million quasars in which periodic signals are expected to be abundant (L. Z. Kelley et al. 2021; C. Xin & Z. Haiman 2021; F. Cocchiara et al. 2024). Our ability to detect those relies on developing methods that have a high true positive rate (unlike the LSP), while simultaneously reducing the false positive rate as much as possible (e.g., J. Robnik et al. 2024). We have demonstrated that even in LSST-like lightcurves, which have improved quality and longer baselines compared to PTF, the recovery rates are low. Therefore, more sophisticated periodicity search algorithms, specifically targeting non-

sinusoidal periodicity, e.g., methods that employ matched filtering or search for periodicity in phase-folded lightcurves, may offer significant advantages compared to the LSP (e.g., see M. J. Graham et al. 2013 for a comparison of methods). Here, we demonstrated that one of the main tools, the LSP, performs very poorly, especially with sawtooth periodicity.

As we already mentioned, the stochastic nature of quasar variability is a major limitation. In this study, we used the DRW model to simulate the underlying noise in quasars. This model provides a successful description of optical variability (B. C. Kelly et al. 2009; S. Kozłowski et al. 2010; C. L. MacLeod et al. 2010; C. J. Burke et al. 2021; K. L. Suberlak et al. 2021), but recent work has found deviations from this simple model (V. P. Kasliwal et al. 2015; K. L. Smith et al. 2018; J. Robnik et al. 2024). A higher-order stochastic model, such as the damped harmonic oscillator (J. Moreno et al. 2019; W. Yu et al. 2022) or higher-order CARMA models (B. C. Kelly et al. 2014) may provide more flexibility and a better description of quasar variability. Improvements in our understanding of quasar variability will need to be incorporated in future periodicity searches to increase their sensitivity and to reduce the false detections. It is worth noting that our assumptions about the underlying noise may impact the reported recovery rates of periodic signals, but the overall trends will likely remain unaffected. We plan to explore additional noise models in a future study.

In addition, throughout our analysis, we adopted a simple sawtooth-shaped periodicity. This cartoon model is sufficient for the purposes of our exercise and illustrates the large impact of non-sinusoidal pulse shapes. However, this shape was merely inspired by the results of hydrodynamical simulations, and the actual shapes of lightcurves produced by binary SMBHs will differ from this shape—indeed, it is poorly understood and will likely vary from source to source. Developing better templates of the expected lightcurve shapes (e.g., see D. J. D’Orazio et al. 2024) can be used to better characterize the completeness and purity of binary candidate samples, and will likely allow improvements in recovery rates.

Finally, another possible limitation in the search for quasar periodicity is unmodeled photometric effects that can introduce spurious signals at the annual or lunar cycle. Such artificial periodicities, leftover from the data reduction process, can be convolved with real periodic signals from SMBHBs impacting their detection (e.g., introducing false positives, moving power to other frequencies, hindering the detection, or leading to misidentification of the real periodicity, etc). Previous quasar periodicity searches have not explored this effect, but discarded periods very close to 1 yr and have completely avoided periods  $<1$  month. However, it is important to understand and mitigate such effects. In a future study, we plan to inject periodic signals in observed quasar data and explore this in several photometric surveys, like CRTS, PTF, ZTF, PanSTARRS, etc., also including some of the active galactic nuclei that have long-term monitoring (e.g., B. M. Peterson et al. 2002; S. Kaspi et al. 2021; Y.-J. Chen et al. 2023).

## 5. Conclusions

In this study, we explore the efficacy of the LSP (one of the most popular tools to search for periodic signals) in detecting non-sinusoidal periodicity in quasars. Periodic signals that significantly deviate from sinusoids, like sawtooth periodicity,

are often seen in hydrodynamical simulations of SMBHBs and may be common among quasars, but would likely have been missed in observational searches to date.

We start from the sample of 35,383 quasar lightcurves from PTF, previously analyzed in M. Charisi et al. (2016), and select a subset of 12,400 for which the SMBH masses are measured. This allows us to estimate the DRW parameters from scaling relations (C. L. MacLeod et al. 2010). Since LSST will soon provide a large dataset of high-quality quasar lightcurves for SMBHB searches, we also consider LSST-like lightcurves. We simulate quasar lightcurves with DRW noise, but note that this is an approximation to quasar noise. We inject two types of periodicity: (a) sinusoidal signals and (b) sawtooth signals. We explore three types of lightcurves to examine the limitations of LSP under different survey conditions: (1) PTF-like lightcurves, with properties (photometric errors, sampling, baselines) resembling the observational data, (2) idealized lightcurves with daily cadence, baselines equal to the respective PTF lightcurves and Gaussian photometric errors with standard deviations equal to the mean photometric error of each source, and (3) LSST-like lightcurves with a baseline of 10 yr, semiregular cadence of 5 days, annual gaps of 4 months, and photometric errors, as expected for a source with a magnitude  $\sim 21$ . For the detection of periodicity, we follow the statistical analysis of M. Charisi et al. (2016), simulating DRW lightcurves to assess the false alarm probability of identified signals.

Our main findings are summarized as follows:

1. The periodogram peaks are reduced by  $\sim 40\%$  when comparing noiseless sawtooth versus sinusoidal signals of the same period and amplitude injected in idealized lightcurves.
2. Of the sinusoidal signals,  $\sim 45\%$  are recovered in PTF-like lightcurves,  $\sim 25\%$  in the idealized lightcurves, and  $\sim 23\%$  in the LSST-like simulations. The stochastic variability of quasars prevents us from identifying the majority of the injected sinusoids.
3. The fraction of recovered signals is significantly worse for sawtooth periodicity, with only  $\sim 9\%$  of the signals detected in the PTF-like lightcurves,  $\sim 1\%$  for the idealized simulations, and  $\sim 1\%$  for the LSST-like simulations. This suggests that the LSP misses the vast majority of binaries and is not an appropriate tool to search for this type of periodicity.
4. The recovered periods are within 15% of the injected values for the vast majority of detected signals, with  $\sim 99\%$  ( $\sim 91\%$ ) for sinusoidal (sawtooth signals) in the idealized lightcurves,  $\sim 97\%$  ( $\sim 80\%$ ) in the PTF-like data, and  $\sim 100\%$  ( $\sim 88\%$ ) in the LSST-like lightcurves.
5. The recovery rates depend on the signal and DRW noise properties as expected. They increase monotonically with the SNR (quantified as  $A/\sigma$ ) and the amplitude of the periodic signals, and thus, strong signals are relatively easy to detect, even in the presence of stochastic variability.
6. The recovery rates decrease as a function of the injected period, which means that the rarer short-period binaries are easier to detect compared to the more abundant long-period ones. For sawtooth signals, binaries with periods  $> 250$  days become almost undetectable (with a detection rate  $> 1\%$  for idealized and LSST-like lightcurves),

primarily due to the DRW noise, which is more significant at these low frequencies.

In this study, we found that non-sinusoidal periodic signals, particularly those with a sawtooth-shaped periodicity predicted in hydrodynamical simulations, can be easily missed in periodicity searches that rely on the LSP. The LSP struggles even with the sinusoidal signals, detecting less than 45% in all three types of lightcurves. This is likely due to the stochastic variability of quasars, which makes the peak of the LSP a suboptimal test statistic (J. Robnik et al. 2024). Based on the above, we can conclude that previous searches would struggle to detect sawtooth signals, even if they were present in the samples. They would have missed a significant fraction of binaries even in the limit that they produced pure sinusoidal signals. This, in combination with the fact that the stochastic variability can introduce false detections, underscores the challenges in identifying periodicity in quasars. Such challenges must be addressed with improved periodicity detection methods in future searches, in order to take full advantage of the upcoming time-domain data.

## Acknowledgments

M.C. acknowledges support by the European Union (ERC; MMMonsters, 101117624). This work was also supported in part by NASA grants 80NSSC24K0440 and 80NSSC22K0822. This research used the resources of the Center for Institutional Research Computing at Washington State University.

## ORCID iDs

Maria Charisi  <https://orcid.org/0000-0003-3579-2522>  
Zoltán Haiman  <https://orcid.org/0000-0003-3633-5403>

## References

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *ApJS*, **182**, 543
- Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2023a, *ApJL*, **951**, L8
- Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2023b, *ApJL*, **952**, L37
- Baluev, R. V. 2008, *MNRAS*, **385**, 1279
- Bécsy, B., Cornish, N. J., & Kelley, L. Z. 2022, *ApJ*, **941**, 119
- Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, *Natur*, **287**, 307
- Bogdanovic, T., Miller, M. C., & Blecha, L. 2021, *LRR*, **25**, 3
- Bon, E., Zucker, S., Netzer, H., et al. 2016, *ApJS*, **225**, 29
- Burke, C. J., Shen, Y., Blaes, O., et al. 2021, *Sci*, **373**, 789
- Burke-Spolaor, S., Taylor, S. R., Charisi, M., et al. 2019, *A&ARv*, **27**, 5
- Caproni, A., Abraham, Z., & Monteiro, H. 2013, *MNRAS*, **428**, 280
- Charisi, M., Bartos, I., Haiman, Z., et al. 2016, *MNRAS*, **463**, 2145
- Charisi, M., Haiman, Z., Schiminovich, D., & D’Orazio, D. J. 2018, *MNRAS*, **476**, 4617
- Charisi, M., Taylor, S. R., Runnoe, J., Bogdanovic, T., & Trump, J. R. 2022, *MNRAS*, **510**, 5929
- Chen, Y.-J., Bao, D.-W., Zhai, S., et al. 2023, *MNRAS*, **520**, 1807
- Chen, Y.-C., Liu, X., Liao, W.-T., et al. 2020, *MNRAS*, **499**, 2245
- Chen, Y.-J., Zhai, S., Liu, J.-R., et al. 2024, *MNRAS*, **527**, 12154
- Cocchiararo, F., Franchini, A., Lupi, A., & Sesana, A. 2024, *A&A*, **691**, A250
- Dai, B.-Z., Zhang, B.-K., Xiang, Y., et al. 2006, *IJMPD*, **15**, 261
- Davis, M. C., Grace, K. E., Trump, J. R., et al. 2024, *ApJ*, **965**, 34
- De Rosa, A., Vignali, C., Bogdanovic, T., et al. 2019, *NewAR*, **86**, 101525
- D’Orazio, D. J., & Charisi, M. 2023, arXiv:2310.16896
- D’Orazio, D. J., Duffell, P. C., & Tiede, C. 2024, *ApJ*, **977**, 244
- D’Orazio, D. J., Haiman, Z., & Schiminovich, D. 2015, *Natur*, **525**, 351
- Duffell, P. C., D’Orazio, D., Derdzinski, A., et al. 2020, *ApJ*, **901**, 25
- EPTA CollaborationInPTA Collaboration, Antoniadis, J., et al. 2023, *A&A*, **678**, A50
- EPTA CollaborationInPTA Collaboration, Antoniadis, J., et al. 2024, *A&A*, **685**, A94
- Eracleous, M., Boroson, T. A., Halpern, J. P., & Liu, J. 2012, *ApJS*, **201**, 23

- Graham, M. J., Drake, A. J., Djorgovski, S. G., et al. 2013, *MNRAS*, **434**, 3423
- Graham, M. J., Djorgovski, S. G., Stern, D., et al. 2015, *MNRAS*, **453**, 1562
- Haiman, Z., Kocsis, B., & Menou, K. 2009, *ApJ*, **700**, 1952
- Hu, B. X., D’Orazio, D. J., Haiman, Z., et al. 2020, *MNRAS*, **495**, 4061
- Huijse, P., Davelaar, J., De Ridder, J., Janssen, N., & Aerts, C. 2025, arXiv:2505.16884
- Ivezić, Ž., Connolly, A., Vanderplas, J., & Gray, A. 2014, *Statistics, Data Mining and Machine Learning in Astronomy* (Princeton Univ. Press)
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, **873**, 111
- Kasliwal, V. P., Vogeley, M. S., & Richards, G. T. 2015, *MNRAS*, **451**, 4328
- Kaspi, S., Brandt, W. N., Maoz, D., et al. 2021, arXiv:2106.00691
- Kauffmann, G., & Haehnelt, M. 2000, *MNRAS*, **311**, 576
- Kelly, B. C., Bechtold, J., & Siemiginowska, A. 2009, *ApJ*, **698**, 895
- Kelly, B. C., Becker, A. C., Sobolewska, M., Siemiginowska, A., & Uttley, P. 2014, *ApJ*, **788**, 33
- Kelley, L., Charisi, M., Burke-Spolaor, S., et al. 2019, *BAAS*, **51**, 490
- Kelley, L. Z., Blecha, L., Hernquist, L., Sesana, A., & Taylor, S. R. 2018, *MNRAS*, **477**, 964
- Kelley, L. Z., D’Orazio, D. J., & Di Stefano, R. 2021, *MNRAS*, **508**, 2524
- Kocsis, B., Frei, Z., Haiman, Z., & Menou, K. 2006, *ApJ*, **637**, 27
- Koen, C. 2021, *AJ*, **161**, 281
- Kormendy, J., & Ho, L. C. 2013, *ARA&A*, **51**, 511
- Kozłowski, S., Kochanek, C. S., Udalski, A., et al. 2010, *ApJ*, **708**, 927
- Law, N. M., Kulkarni, S. R., Dekany, R. G., et al. 2009, *PASP*, **121**, 1395
- Li, Y.-R., Wang, J.-M., Zhang, Z.-X., et al. 2019, *ApJS*, **241**, 33
- Liu, T., Gezari, S., Ayers, M., et al. 2019, *ApJ*, **884**, 36
- Lomb, N. R. 1976, *Ap&SS*, **39**, 447
- Luo, D., Jiang, N., & Liu, X. 2025, *ApJ*, **978**, 86
- MacLeod, C. L., Ivezić, Ž., Kochanek, C. S., et al. 2010, *ApJ*, **721**, 1014
- Miles, M. T., Shannon, R. M., Reardon, D. J., et al. 2025, *MNRAS*, **536**, 1489
- Moreno, J., Vogeley, M. S., Richards, G. T., & Yu, W. 2019, *PASP*, **131**, 063001
- Peterson, B. M., Berlind, P., Bertram, R., et al. 2002, *ApJ*, **581**, 197
- Rau, A., Kulkarni, S. R., Law, N. M., et al. 2009, *PASP*, **121**, 1334
- Reardon, D. J., Zic, A., Shannon, R. M., et al. 2023, *ApJL*, **951**, L6
- Richstone, D., Ajhar, E. A., Bender, R., et al. 1998, *Natur*, **385**, A14
- Robnik, J., Bayer, A. E., Charisi, M., et al. 2024, *MNRAS*, **534**, 1609
- Roos, N., Kaastra, J. S., & Hummel, C. A. 1993, *ApJ*, **409**, 130
- Saade, M. L., Stern, D., Brightman, M., et al. 2020, *ApJ*, **900**, 148
- Scargle, J. D. 1982, *ApJ*, **263**, 835
- Smith, K. L., Mushotzky, R. F., Boyd, P. T., et al. 2018, *ApJ*, **857**, 141
- Suberlak, K. L., Ivezić, Ž., & MacLeod, C. 2021, *ApJ*, **907**, 96
- Süveges, M., Guy, L. P., Eyer, L., et al. 2015, *MNRAS*, **450**, 2052
- Taylor, S. R. 2021, arXiv:2105.13270
- Taylor, S. R., van Haasteren, R., & Sesana, A. 2020, *PhRvD*, **102**, 084039
- Timmer, J., & König, M. 1995, *A&A*, **300**, 707
- VanderPlas, J. T. 2018, *ApJS*, **236**, 16
- Vaughan, S., Uttley, P., Markowitz, A. G., et al. 2016, *MNRAS*, **461**, 3145
- Westernacher-Schneider, J. R., Zrake, J., MacFadyen, A., & Haiman, Z. 2022, *PhRvD*, **106**, 103010
- Witt, C. A., Charisi, M., Taylor, S. R., & Burke-Spolaor, S. 2022, *ApJ*, **936**, 89
- Xin, C., Charisi, M., Haiman, Z., et al. 2020, *MNRAS*, **496**, 1683
- Xin, C., & Haiman, Z. 2021, *MNRAS*, **506**, 2408
- Xu, H., Chen, S., Guo, Y., et al. 2023, *RAA*, **23**, 075024
- Yu, W., Richards, G. T., Vogeley, M. S., Moreno, J., & Graham, M. J. 2022, *ApJ*, **936**, 132
- Zrake, J., Tiede, C., MacFadyen, A., & Haiman, Z. 2021, *ApJL*, **909**, L13
- Zu, Y., Kochanek, C. S., Kozłowski, S., & Udalski, A. 2013, *ApJ*, **765**, 106