Quasars with periodic variability as sub-parsec supermassive black hole binary candidates

Maria Charisi

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2017

ABSTRACT

Quasars with periodic variability as sub-parsec supermassive black hole binary candidates

Maria Charisi

Supermassive black hole binaries (SMBHBs) are the natural consequence of galaxy mergers and should form frequently in galactic nuclei. Especially at sub-parsec separations, where the binary evolution is slow, SMBHBs should be fairly abundant. However, the observational evidence remains elusive.

In this thesis, we focus on periodic variability of quasars as a potential signature of compact SMBHBs. First, we present a systematic search for periodic variability in the photometric database of the Palomar Transient Factory. Our search in a large sample of $\sim$35,000 quasars returned 50 candidates, 33 of which remain significant after the reanalysis of extended light curves including data from the Catalina Real-Time Transient Survey and the intermediate PTF. Our candidates have periods of a few hundred days.

Next, we focus on independent signatures that could verify the binary nature of the candidates. We present a case study of quasar PKS 1302-102, the first candidate that emerged from the large time-domain surveys. We search for multiple periodic components in the variability with a characteristic frequency pattern predicted by hydrodynamical simulations of circumbinary disks. We do not find compelling evidence for a secondary period.

Additionally, in compact SMBHBs, relativistic Doppler boost should be significant and may dominate the variability. This model was suggested as a smoking-gun signature for quasar PKS 1302-102, since it is not expected in quasars with a single BH and it offers a robust prediction, which can be tested with multi-wavelength data. With a control sample of non-periodic quasars, we test whether this signature is distinct from the intrinsic multi-wavelength variability of quasars. We concluded that the Doppler boost does not provide a sharp test for SMBHBs.

Table of Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I would like to express my immense gratitude to my advisors and mentors Zoltán Haiman, Szabi Márka and Imre Bartos. This thesis would not have been possible without your continuous encouragement, support and guidance. You have had such a positive influence on my scientific and personal growth and I value every moment. It has been a pleasure and a privilege and I hope our collaboration will continue for many years. Additionally, I thank my thesis committee Jules Halpern and Greg Bryan for the useful suggestions for improvement, and David Schiminovich for our recent collaboration and for being on my defense committee. I am also grateful to my undergraduate mentor, Prof. Loukas Vlahos, who believed in me from my first steps in physics and encouraged me to pursue a PhD in astronomy.

The astronomy department has been a second family for me over the past six years. I couldn't thank enough my friends Jia Liu, Miao Li and Kohei Inayoshi for the deep understanding, the emotional safety and for all the precious moments we have shared. I am also grateful to my wonderful office mates Aleksey Generozov, Jose Zorrilla and Mihir Kulkarni for the heated debates on every possible topic, scientific and not. I was fortunate to interact and learn from many great fellow graduate students and post-docs. Many thanks to Ximena Fernandez, Jenna Lemonias, Alejandro Nunez, Stephanie Douglas, Nick Stone, Kohei Ichikawa and Robyn Sanderson.

I would also like to thank all my friends outside of astronomy, both in New York and in Greece, for giving me perspective and keeping me sane. Special thanks to Pantelis for reading and editing many of my writings and to Tassos for his thoughtful advice on my best and worst days. Of course, I wouldn't have accomplished anything without the love and the everyday support from my family, my parents, my grandparents, my aunts and my sisters Vassiliki, Stergiani and Alexandra.

Στους γονείς μου Βασίλη και Μαίρη

Chapter 1

Introduction

1.1 Theory of Supermassive Black Hole Binaries

In 1963, [Schmidt, 1963] discovered the first quasar; a bright star-like object at redshift 0.158 associated with the radio source 3C 273. Several similar discoveries followed soon after. The tremendous energy release, required to explain the brightness of these sources at cosmological distances, along with their compactness, inferred from the observed rapid variability, led to the conclusion that quasars are powered by accreting supermassive black holes (SMBHs), i.e. BHs with mass $10^6 - 10^9 M_\odot$ [Salpeter, 1964].

Today, SMBHs are believed to be ubiquitous in the universe; they can be found not only at the hearts of quasars, but also at the centers of most (if not all) quiescent galaxies [Lynden-Bell, 1969], including our own Milky Way [Ghez *et al.*, 1998]. The mass of the SMBH correlates with several properties of the host galaxy, e.g., the velocity dispersion of the bulge, the mass and the luminosity of the bulge, etc [Kormendy and Ho, 2013]. This suggests that the central SMBH is an important component of the galaxy and potentially plays a key role in the evolution of the galaxy.

Additionally, in the standard Λ-CDM cosmology, the larger structures we see in the universe are formed through hierarchical merging of smaller structures [Peebles, 1980]. Low-mass proto-galaxies are formed when baryonic matter falls into collapsing dark matter halos, which under the attraction of gravity undergo mergers and build-up higher-mass galaxies. The mergers continue throughout the cosmic history, until we arrive at the massive galaxies

we observe today in the Local Universe. The mergers are fairly frequent, although the merger rate has declined compared to the past; for instance, the Hubble Deep and Ultra-Deep fields contain a higher percentage of perturbed or interacting galaxies compared to the local universe [Conselice, 2014].

In a seminal paper, [Begelman *et al.*, 1980] proposed that the natural consequence of galaxy mergers is the formation of bound SMBH binaries (SMBHBs) and argued that many Active Galactic Nuclei (AGNs) may host, in their centers, two SMBHs orbiting each other. They also detailed three main evolutionary stages of the BHs in the post-merger galaxy.

Initially, dynamical friction against the stellar bulge acts efficiently on each BH and pushes them towards the center of the common gravitational potential. The BHs form a Keplerian binary when the mass of the stars enclosed in the binary orbit is roughly equal to the total mass of the binary (Phase I). Subsequently, the binary decays through close three-body encounters by scattering nearby stars (Phase II). At the final stage, the binary evolution is dominated by the emission of gravitational waves (Phase III).

If the emission of Gravitational Waves (GWs) is asymmetric, the GWs carry away linear momentum. The conservation of momentum dictates that the center of mass of the system (and eventually the final BH) should receive a non-zero kick [Bekenstein, 1973; Komossa, 2012]. The velocity of the recoil depends on the parameters of the binary (such as the mass ratio, or the spin orientation) and in extreme cases it can exceed the escape velocity of the merged galaxy [Gualandris and Merritt, 2008; Lousto and Zlochower, 2011].

The transition between Phase II and Phase III is currently under debate. If the loss cone is depleted when the binary is at large separation, the evolution of the binary due to GWs is slow and the binary does not merge within a Hubble time. This issue is known as the Final Parsec Problem (FPP, see [Colpi, 2014] for a review). In fact, the stalling of the binary at parsec separations has been robustly shown in N-body simulations, when the binary is insparalling in an isotropic, spherical galaxy [Milosavljević and Merritt, 2001; Preto *et al.*, 2011].

More recent simulations have challenged this picture; if the binary evolves in an axisymmetric or triaxial potential, a more likely outcome of a galaxy merger, the loss cone (i.e. the domain, in phase-space, populated by stars with sufficiently low angular momentum to

interact with the binary) is replenished, and the binary decays to small separations [Berczik *et al.*, 2006; Khan *et al.*, 2013; Wang *et al.*, 2014]. However, [Vasiliev *et al.*, 2015] cautioned that the hardening of the binary in a non-spherical potential may be an artefact of poor mass resolution in the N-body simulations. Even if the stellar interactions cannot extract adequate energy, a stalled binary may be driven to coalescence in a subsequent galaxy merger, due to perturbations in the binary orbit from the third massive BH [Blaes *et al.*, 2002; Hoffman and Loeb, 2007].

Another possible solution to the FPP arises from the existence of gas in the nuclear region following a gas-rich merger. Numerical simulations indicate that dense gaseous clouds present in the pre-merger galaxies also sink to the nuclear regions of the post-merger galaxy [Barnes and Hernquist, 1992; Mayer *et al.*, 2007], and therefore the binary is surrounded by large amounts of gas. This has also been confirmed observationally; for instance, the luminosity of Ultra Luminous Infra-Red Galaxies (ULIRGs), which are likely advanced mergers, arises in massive clouds of dusty molecular gas [Sanders and Mirabel, 1996]. The gas around a bound binary is expected to cool and settle in a rotationally supported circumbinary disk [Barnes, 2002].

Extensive theoretical studies have explored the interaction of a binary with a gaseous disk (applications of such studies extend beyond the SMBHB case, e.g., protoplanetary disks, binary stars, etc.). The gas can extract angular momentum from the binary, catalyzing its evolution and pushing the binary to the GW regime [Begelman *et al.*, 1980; Gould and Rix, 2000; Escala *et al.*, 2004; Roedig *et al.*, 2012; Tang *et al.*, 2017]. It can also produce significant electromagnetic signatures [Chang *et al.*, 2010; Baruteau *et al.*, 2012] and potentially even affect the GW waveforms [Armitage and Natarajan, 2005; Yunes *et al.*, 2011].

Early work on circumbinary disks showed that if the binary is embedded in a thin disk, coplanar with the orbit of the binary, there are no stable circular orbits in the central regions of the disk (to a distance $2a$ from the centre of mass of the system, where a is the separation of the binary) and thus the inner disk is truncated [Artymowicz and Lubow, 1994; Milosavljević and Phinney, 2005]. The binary resides in a central cavity of very low density gas and the accretion on the binary is halted. The formation of the cavity was observed for

a variety of mass ratios.

Hydrodynamical simulations that followed suggested that the cavity is not totally devoid of gas [Artymowicz and Lubow, 1996]. Gaseous streams are tidally stripped from the rim of the cavity and can reach the SMBHs facilitated by shocks, which redistribute the angular momentum [Shi and Krolik, 2015]. The gas inflow is significant, since it affects the inward migration of the BHs [Kocsis *et al.*, 2012a; Kocsis *et al.*, 2012b; Rafikov, 2013; Rafikov, 2016]. Additionally, it has been shown that it can lead to high near-Eddington luminosities, therefore SMBHBs can have bright quasar-like electromagnetic properties. A common conclusion of many hydrodynamical simulations, and the main motivation for the work presented in this thesis, is that the mass accretion rate onto the BHs is periodically modulated [Hayasaki *et al.*, 2007; MacFadyen and Milosavljević, 2008; Cuadra *et al.*, 2009; Roedig *et al.*, 2012; Shi *et al.*, 2012; Noble *et al.*, 2012; D'Orazio *et al.*, 2013; Farris *et al.*, 2014; Gold *et al.*, 2014].

Additionally, when the binary decays due to the circumbinary disk, the evolution is expected to be slow. [Haiman *et al.*, 2009] combined the expectations for periodic variability with simple disc migration models and calculated that a large fraction of quasars could host binaries with orbital periods of weeks to months. They suggested that a search for quasars with periodic variability in the large samples of the current time domain surveys could reveal several candidate SMBHBs. In Chapter 2, we describe such a search in the photometric database of Palomar Transient Factory (PTF).

Another generic conclusion of hydrodynamical simulations is that the central cavity is typically elongated [MacFadyen and Milosavljević, 2008; Roedig *et al.*, 2012; Shi *et al.*, 2012; Noble *et al.*, 2012; D'Orazio *et al.*, 2013; Farris *et al.*, 2014; Gold *et al.*, 2014]. For certain mass ratios ($q > 0.3$), an over-density is formed at the edge of the lopsided cavity. The accretion rate into the cavity is again periodically modulated, but in this case the strongest periodic component does not coincide with the orbital period of the binary, but is several times longer, equal to the orbital period of the hotspot. This means that the variability of SMBHBs may consist of multiple periodic components with a characteristic frequency pattern, e.g., 1:5 (in a range of 1:8 to 1:3 depending on the binary and disk parameters). This provides an additional EM signature of SMBHBs, which can be used to

identify binary candidates. In Chapter 3 of this work, we explore the existence of multiple periodic components for a binary candidate identified from its periodicity.

Additionally, simulations predict that the infall of gas into the cavity results in the formation of mini-disks around each BH [Ryan and MacFadyen, 2017]; the mini-disc of the secondary BH dominates the luminosity of the binary. When the binary is at sub-parsec separations, the orbital velocity of the BHs reaches a significant fraction of the speed of light. In this regime, the emission of the gas bound to the BHs is strongly affected by special relativity.

In detail, the binary will appear brighter when the most luminous BH (typically the less massive secondary) is moving towards the observer and dimmer when it is receding, even if the rest-frame luminosity is constant. This could lead to the quasi-sinusoidal variability. Additionally, this model has a robust prediction, which can be tested with multi-wavelength data. For instance, if the optical and UV variability originate in the mini-disks, they should vary in tandem, but with different amplitudes. The relative amplitudes in the two bands depend on the respective spectral indices (e.g., see [D'Orazio *et al.*, 2015b]). In Chapter 4, we explore this potential signature of compact SMBHBs; in particular, we present a crucial 'null test' required for a detection of this signature.

1.2 Observations of Supermassive Black Hole Binaries

From an observational standpoint, galaxy mergers and the different stages of the pairing of their SMBHs have been the focus of many studies (see [Dotti *et al.*, 2012; Komossa and Zensus, 2016] for reviews). The initial stages of galaxy mergers, when galaxies are still distinct, but gravitationally bound at $\sim$100 kpc, have been explored mainly by searching for quasar clustering and quasar pairs [Hennawi *et al.*, 2006; Myers *et al.*, 2008; Foreman *et al.*, 2009; Farina *et al.*, 2011].

At smaller separations of $\sim$10 kpc to 1 kpc, where SMBHs are driven by dynamical friction and have not formed bound Keplerian binaries, such systems can be identified as dual AGN (both of the merging SMBHs are active) or offset AGNs (only one of the SMBHs is active, with the second BH either inactive or obscured).

Several hundred candidates of both dual and offset AGN have been selected from spectra with double-peaked [Comerford *et al.*, 2009; Wang *et al.*, 2009; Liu *et al.*, 2010; Smith *et al.*, 2010; Ge *et al.*, 2012; Barrows *et al.*, 2013; Comerford *et al.*, 2013] or single-peaked narrow emission lines shifted compared to the lines of the host galaxy [Comerford and Greene, 2014]. These spectral features can be produced by the motion of the narrow line regions along the orbit of the AGNs in a binary system, but they can also be produced by a single AGN with outflows, jets, or a rotating gaseous disk [Crenshaw *et al.*, 2010; Fischer *et al.*, 2011].

Follow-up observations of these candidates have confirmed the existence of dozens of binaries through direct imaging of the two BHs in infrared [McGurk *et al.*, 2011], optical [Comerford *et al.*, 2009; Shields *et al.*, 2012], X-rays [Komossa *et al.*, 2003; Liu *et al.*, 2013; McGurk *et al.*, 2015; Comerford *et al.*, 2015; Koss *et al.*, 2016] and radio [Fu *et al.*, 2011; Fabbiano *et al.*, 2011]. A few triple systems, involving binaries at kpc separations have also been reported [Barth *et al.*, 2008; Liu *et al.*, 2011; Schawinski *et al.*, 2011; Deane *et al.*, 2014].

At even smaller separations of order of 10 pc, when the binary becomes bound, individual BHs can only be resolved in the radio, through Very Long Baseline Interferometry (VLBI). The record holding binary was identified in the radio galaxy 0402+379 with a projected separation of 7.2 pc [Rodriguez *et al.*, 2006]. The direct detection of binaries with this method becomes increasingly challenging, as the binary decays to smaller separations. For reference, a nearby binary at a distance of 100 Mpc and a separation of 1 pc subtends an angle of only 2 mas.

Despite our inability to directly image individual BHs in very compact binaries (although the Event Horizon Telescope [Fish *et al.*, 2016], which has achieved a resolution better than 60 μas would be able to resolve a SMBHB with milli-parsec separation in M87), the search for SMBHBs at sub-parsec separations is an area of very active research. In this regime, one has to rely on indirect methods, i.e. detecting the effects of a binary on its surroundings. Several authors have attempted systematic searches for different potential signatures of sub-pc binaries, which have returned several candidates. A few additional candidates were discovered serendipitously. Below is a short and likely incomplete summary of the methods

and the findings.

Similarly to dual AGN, searches for sub-parsec binaries have focused on identifying either single-peaked broad lines with significant velocity offset from the narrow lines or double-peaked broad lines, assuming that each BH has a bound broad line region [Shen and Loeb, 2010; Popović, 2012]. Several candidates have been identified from single epoch spectroscopy, mainly from systematic searches in SDSS [Boroson and Lauer, 2009; Decarli *et al.*, 2010; Tsalmantza *et al.*, 2011; Eracleous *et al.*, 2012]. If the spectral shifts are indeed due to the presence of a binary, long-term monitoring should reveal the characteristic Doppler-shifts of the emission lines reflecting the orbital motion.

Follow-up observations of double-peaked AGN have not revealed variability for most of the candidates and therefore it seems that this subset of candidates are unlikely to be binaries [Halpern and Eracleous, 2000; Liu *et al.*, 2016a], although some recent papers have reported periodic variability in the broad lines of two nearby AGNs, NGC 4151 [Bon *et al.*, 2012] and NGC 5548 [Li *et al.*, 2016; Bon *et al.*, 2016]. Also, in the case of candidates with offset single-peaked broad lines, given the long timescales required for monitoring several orbital periods, it is typically easier to disprove the candidates due to lack of variability rather than confirm them [Runnoe *et al.*, 2017]. A related but slightly different approach uses multi-epoch spectroscopy to select binary candidates taking advantage of the aforementioned expectation for spectral variability and has returned a few additional candidates [Ju *et al.*, 2013; Shen *et al.*, 2013; Wang *et al.*, 2017].

Multiple candidates have also been proposed from observations of radio jets with intriguing morphologies [Villata and Raiteri, 1999; Romero *et al.*, 2000; Romero *et al.*, 2003; Lobanov and Roland, 2005; Caproni *et al.*, 2013; Kun *et al.*, 2014]. For instance, the orbital motion of the SMBH at the base of the jet can lead to quasi-periodic deviations from a straight line [Kaastra and Roos, 1992; Roos *et al.*, 1993]. Additionally, the presence of a binary could also result in precession of the jet angle, leading to a jet with conical morphology [Begelman *et al.*, 1980; Gower *et al.*, 1982]. The acceleration of the binary at the last stages before the merger could also have observable consequences on the morphology of the jet, i.e. produce a helical "chirping" jet with a widening opening angle and increasingly dense wiggles [Kulkarni and Loeb, 2016].

Another potential spectral signature was explored in the case of Mrk 231, the spectrum of which shows a flux deficit at ∼3000 A and flat continuum both in optical and UV [Yan *et al.*, 2015]. This was interpreted as a manifestation of the opening of a gap in the circumbinary disk, mentioned above, with the optical and UV continuum originating from emission in the circumbinary and the mini-disk of the secondary BH, respectively. However, as pointed out by [Farris *et al.*, 2015b], the shock heated mini-disks and the streams that penetrate the cavity possibly prevent the formation of the "notch" in the spectrum. Additionally, this candidate was disputed, because this spectrum can also be explained by dust reddening [Veilleux *et al.*, 2016; Leighly *et al.*, 2016].

Additionally, if a star is tidally disrupted by a SMBH in a binary, the presence of the second BH could be imprinted in the the light curve of the tidal disruption event (TDE). The secondary can temporarily halt the accretion to the primary causing a dip in the light curve [Liu *et al.*, 2009]. This model was invoked to explain the atypical TDE flare in SDSSJ120136.02+300305.5 [Liu *et al.*, 2014] and it requires an unequal mass binary ($q \sim 0.1$) at a separation of 0.6 mpc.

As was already mentioned in Section 1.1, another method to identify compact SMBHBs (and the main focus of this thesis) is through their periodic variability. Several claims can be found in the literature with periods ranging from a few days to decades [Raiteri *et al.*, 2001; Fan *et al.*, 2002; Rieger, 2004; Xie *et al.*, 2008; Kudryavtseva *et al.*, 2011], but many of them were disproved, since the periodic variability was not persistent.

One prominent example is the case of blazar OJ287. This bright blazar has been observed for over a century and shows twin outbursts every 11.86 yr separated by 1 yr. The variability can be explained, if the binary mass ratio is low ($q \sim 0.01$) and the secondary BH orbits in a highly eccentric orbit with a period of 11.86 years. The periodic outbursts are observed when the secondary plunges through the accretion disk surrounding the primary (see [Valtonen and Wiik, 2012] and references therein).

The advent of time-domain astronomy with the massive datasets has recently allowed more systematic searches for periodic variability. [Graham *et al.*, 2015a] analyzed quasar light curves from the Catalina Real-Time Transient Survey (CRTS), an all-sky time domain survey, which has continuously monitored the sky for 9 yr. In a large sample of ∼250,000

quasars, they identified 111 candidates with periods of a few years.

Our work, which is described in Chapter 2 used data from the Palomar Transient Factory (PTF), a more recent time-domain survey with a baseline of four years, which covered the Northern sky to deeper magnitudes. We searched for periodic variability in a sample of ∼35,000 quasars and identified additional 33 candidates, typically with shorter periods (compared to the CRTS sample) of a few hundred days.

[Liu *et al.*, 2015] analyzed a small sample of ∼300 quasars from Pan-STARRS1 Medium Deep Survey and reported the detection of a quasar with a period of 550 d, at a separation of 6 mpc. Follow-up monitoring of this source has shown that the periodicity is not persistent [Liu *et al.*, 2016b]. [Zheng *et al.*, 2016] analyzed CRTS light curves for a sample of ∼350 low-redshift quasars in Stripe 82 and identified an additional candidate with two periodic components in the variability (∼750 and 1500 d). The characteristic period ratio is predicted in hydrodynamical simulations for a range of mass ratios (e.g., $0.05 < q < 0.8$). A few individual candidates have also been reported recently [Bon *et al.*, 2016; Li *et al.*, 2017].

It is important to recognize that identifying periodicity in quasars is challenging; this is already demonstrated by the candidates that have been already disproved. First, the intrinsic variability of quasars is stochastic. It is best described by a Damped Random Walk (DRW) model, i.e. red noise variability at high frequencies and white noise at lower frequencies [Kelly *et al.*, 2009; MacLeod *et al.*, 2010; Kozłowski *et al.*, 2010]. Most of the above searches assessed the statistical significance of the findings assuming DRW variability. [Vaughan *et al.*, 2016] pointed out that the incomplete knowledge of the variability can lead to false detections of periodicity. In fact, more recent studies have indicated deviations from the DRW model, with steeper power spectral density at very high frequencies [Kasliwal *et al.*, 2015].

Another limitation is that the identified periods are typically long compared to the baselines of the available light curves and for most of the candidates very few cycles have been observed. Of course, if the observed periodicity is indeed due to a binary, it should be persistent. So, if we wait long enough (which in the case of the longest periods means a few decades), we would be able to prove the binary nature of the source.

A complementary approach is to search for multiple independent pieces of evidence.

Identifying unique signatures that are associated with SMBHBs can increase the credibility of the candidates. As mentioned in Section 1.1, hydrodynamical simulations predict that the most pronounced periodic component in a binary system may not correspond to the orbital period, but to a longer period, which is related to the orbital period of a hotspot in the circumbinary disk. This may be translated in multiple periodic components in the variability with a characteristic frequency pattern, e.g., 1:5. In Chapter 3, we describe a search for this characteristic binary signature in the periodogram of the PKS1302-102 (one of the first periodic candidates that emerged from the current time-domain surveys [Graham *et al.*, 2015b]).

Another suggested signature, as described in Section 1.1 is to identify Doppler boosting of the emission due to the high orbital velocity. This model was proposed to explain the optical and the UV variability of PKS 1302-102 [D'Orazio *et al.*, 2015b]. In Chapter 4, we explore this proposed multi-wavelength signature, which is inevitable in compact SMBHBs. More specifically, we test whether this prediction of the relativistic Doppler boost can be robustly distinguished from intrinsic multi-wavelength variability of quasars.

It is also worth noting that even though we still don't have a confirmed sub-pc SMBHB, there is circumstantial evidence that SMBHBs do merge within a Hubble time. For instance, X-shaped radio galaxies can be explained by a rapid re-orientation of the jet axis at the final stages of the merger (see [Gopal-Krishna *et al.*, 2012] for a review). The non-axisymmetric emission of GWs will lead to a velocity kick of the merged SMBH and several candidates of recoiling SMBHs have been identified [Komossa, 2012]. The deficit of stellar light in the core of elliptical galaxies has also been suggested as evidence of mergers. The stars at the central regions are expelled by a decaying SMBHB [Dullo and Graham, 2014].

1.3 Gravitational Waves from Supermassive Black Hole Binaries

The most conclusive evidence that SMBHBs indeed overcome the FPP and coalesce may come from the detection of GWs. Compact SMBHBs at the final stages of their evolution, which are the main focus of this thesis, are expected to be the strongest sources of GWs.

With the groundbreaking discovery of high-frequency GWs from merging stellar-mass BHs by LIGO [Abbott *et al.*, 2016], we have just stepped into the era of GW astronomy and the detection of low-frequency GWs form merging SMBHs is the next major goal. Even though the frequencies of these waves are inaccessible from the ground, low-frequency GWs are targeted by experiments like the Laser Interferometer Space Antenna (LISA) and Pulsar Timing Arrays (PTAs).

More specifically, LISA [Amaro-Seoane *et al.*, 2017] is an approved space-based interferometer, planned to launch in the early 2030s. According to the proposed design, which is not finalized yet, the observatory will consist of three identical spacecrafts in a triangular configuration, separated by 2.5 million km. The three arms will be linked with six laser beams and will be trailing the Earth's orbit around the Sun. LISA will be sensitive to milli-Hz GWs (from a few 10^{-5} to 10^{-1} Hz), which will allow to probe mergers of SMBHBs with mass of $10^4 - 10^7 M_\odot$ up to redshift z$\sim$10.

The lowest frequencies of the GW spectrum (nano-Hz) are currently probed by PTAs [Manchester and IPTA, 2013]. These experiments rely on pulsars, the nature's most precise clocks. A passing GW will change the relative path between the pulsar and the Earth, resulting in deviations in the times of arrivals (TOAs) of the pulses; GWs can be detected from the coherent deviations of TOAs from a population of pulsars. PTAs are sensitive to GWs from nearby massive SMBHBs ($10^8 - 10^9 M_\odot$). Additionally, they may detect the stochastic GW background from the entire population of merging SMBHs. In fact, the current upper limits on the GW background][Shannon *et al.*, 2015] already provide meaningful constraints on astrophysical models of galaxy mergers, including constraints on the population of periodic quasars [Sesana *et al.*, 2017].

Chapter 2

A Population of Short-Period Variable Quasars from PTF as Supermassive Black Hole Binary Candidates

2.1 Introduction

Strong observational evidence suggests that every massive galaxy hosts a supermassive black hole in its nucleus [Kormendy and Ho, 2013]. The central black hole (BH) is an important component of the galaxy, since the BH mass is correlated with the global properties of the host galaxy, e.g., dispersion velocity, bulge luminosity, or bulge mass. Moreover, hierarchical models of structure formation predict that galaxies merge frequently [Haehnelt and Kauffmann, 2002], which naturally leads to the the formation of Supermassive black hole binaries (SMBHBs).

Following the merger, the BHs rapidly sink towards the centre of the common gravitational potential, under the effect of dynamical friction, and form a bound Keplerian binary. Subsequently, the binary orbit decays, as the BHs expel nearby stars in close three-body interactions, and/or as they interact with a gaseous circumbinary disc. At close separations,

the binary is driven to coalescence by the emission of gravitational radiation [Begelman *et al.*, 1980]. Eventually, the emerging BH recoils [Komossa, 2012] to counterbalance the net linear momentum transferred by the gravitational waves (GWs). The strong gravitational radiation, emitted during the final stages of the inspiral, may be detectable by Pulsar Timing Arrays (North American Nanohertz Observatory for Gravitational waves; [Jenet *et al.*, 2009], Parkes Pulsar Timing Array; [Manchester, 2008], European Pulsar Timing Array; [Janssen *et al.*, 2008]) or future space-based interferometers, such as eLISA [Amaro-Seoane *et al.*, 2013].

Whether three-body interactions can extract adequate energy to bring the binary to the GW regime is still unclear and it has been suggested that the binary may stall at parsec separations (*final parsec problem*; see [Khan *et al.*, 2013] or [Colpi, 2014], for a recent review). However, if cold gas exists in the central regions, the final parsec problem may be circumvented. This is motivated by simulations of gas-rich mergers, which indicate that copious amounts of gas are funneled into the nuclear regions [Barnes and Hernquist, 1992]. The gas, following the formation of a Keplerian binary, settles in a rotationally supported disc [Barnes, 2002], and the binary hardens, as it dissipates energy and angular momentum to the gaseous disc. The exchange of angular momentum at this stage is slow and the binary is expected to spend a significant fraction of the fiducial $10^7 - 10^8$ yr lifetime at sub-parsec separations (e.g., [Haiman *et al.*, 2009]; see also [Kocsis *et al.*, 2012a; Kocsis *et al.*, 2012b; Rafikov, 2013; Rafikov, 2016] for the long-term evolution of a system of a SMBHB with a circumbinary disc).

Given the central role mergers play in galaxy formation and evolution and the uncertainties in binary evolution, identifying SMBHBs is of major importance. It would allow us to observationally constrain several key questions regarding the galaxy merger rate, as well as the processes and the environments involved in the path of SMBHBs to coalescence. It would also provide an estimate of the population of SMBHBs emitting GWs. Other important questions include the connection of SMBHBs with increased AGN activity [Gaskell, 1985] or triggering starburst activity [Taniguchi and Wada, 1996] and the growth of SMBHs through mergers. The significance of the above questions has prompted intensive efforts to detect SMBHBs (see [Komossa, 2006], for a review).

Several wide binaries (at kpc separation) have been spatially resolved in different bands, from X-rays to radio [Owen *et al.*, 1985; Komossa *et al.*, 2003; Bianchi *et al.*, 2008; Green *et al.*, 2010; Koss *et al.*, 2011; Fabbiano *et al.*, 2011; Comerford *et al.*, 2011; Fu *et al.*, 2015], including a few in triple systems [Liu *et al.*, 2011; Deane *et al.*, 2014]. At parsec separations, SMBHBs can be resolved only with radio interferometers (e.g., the Very Long Baseline Array; VLBA), if both BHs are radio loud. [Rodriguez *et al.*, 2006], for instance, identified a pair of active SMBHs separated by 7.3 pc (the smallest separation that has been resolved) in the radio galaxy B3 0402+379. At sub-parsec separations, where binaries are expected to spend non-negligible time, resolving individual black holes is practically impossible, especially at cosmological distances. Therefore, observational efforts have focused on the effects of the binary on its environment as indirect probes of binaries.

If either of the BHs is associated with a radio jet, the orbital motion and/or the precession of the spin axis will be imprinted on the geometry of the emitted jet, resulting in radio jets with wiggles or knots [Kaastra and Roos, 1992], or in helical jets with conical geometry [Gower *et al.*, 1982], respectively. At the closest separations, when the orbital decay leads to an observable decrease in orbital period, the jets may exhibit a "chirp" behavior [Kulkarni and Loeb, 2016] Helical or wiggled radio jets have been attributed to SMBHBs in several cases [Roos *et al.*, 1993; Romero *et al.*, 2000; Britzen *et al.*, 2001; Lobanov and Roland, 2005; Valtonen and Wiik, 2012; Caproni *et al.*, 2013; Kun *et al.*, 2015]. Additionally, if gas is bound to one or both BHs, the spectral lines are expected to be noticeably Doppler shifted, reflecting the high orbital velocities in a close binary. Double peaked Balmer lines and displaced Broad Line Regions (BLR), relative to the galaxy's rest frame, have been identified in AGN and quasar spectra and were linked to SMBHBs (see [Popović, 2012], for a review). However, the above candidates remain controversial, since alternative scenarios can also provide feasible explanations. For example, the morphology of a radio jet can be distorted due to Kelvin-Helmholtz instability, and spectral signatures can be produced from a single SMBH if the BLR has a complex geometry (e.g., [Chornock *et al.*, 2010; Liu *et al.*, 2016a]).

Another proposed signature of SMBHBs is periodic modulation of the luminosity in AGNs or quasars (in optical, UV, or X-ray bands), induced by the orbital motion of the

binary. From theoretical work on circumbinary discs, we expect that, if a SMBHB is embedded in a thin accretion disc, the torques exerted by the binary will expel the gas from the central region, leaving a central cavity almost devoid of gas [Artymowicz and Lubow, 1994]. The orbital motion of the binary perturbs the inner edge of the cavity, pulling gaseous streams towards the BHs. The mass accretion rate, and possibly the brightness, is periodically modulated at the orbital period of the binary; a generic conclusion from several hydrodynamical simulations [Hayasaki *et al.*, 2007; MacFadyen and Milosavljević, 2008; Noble *et al.*, 2012; Roedig *et al.*, 2012; D'Orazio *et al.*, 2013; Gold *et al.*, 2014; Farris *et al.*, 2014].

Periodic variability is a promising method to detect SMBHBs at very close separations.[1] Multiple claims for periodic variability in blazars, AGNs and quasars have been reported in the literature, with periods ranging from a few days to a few decades [Komossa, 2006]. A very compelling case is the BL Lac Object OJ287, which has been monitored for over a century. The optical light curve shows persistent twin outbursts with a period of 11.86 yr. The outbursts are separated by ∼1 yr, and the second outburst is also accompanied by enhanced radio emission [Pursimo *et al.*, 2000].

Nevertheless, the individual claims for detection of periodicity (and thus the binary nature of the individual sources) have been disputed. An alternative approach is to statistically identify a population of periodic AGNs and quasars. The period distribution of such a population can trace the evolution of SMBHBs, enabling us to study the physics of the orbital decay. More specifically, at shorter periods, the distribution is expected to follow a steep ($\propto t_{\rm orb}^{8/3}$) power law, indicating a rapidly decaying population dominated by the emission of GWs, whereas at longer periods, the distribution is less steep, signifying the slower evolution and slower acceleration, during which the binary exchanges angular momentum with the circumbinary disc [Haiman *et al.*, 2009]. The discovery of a population with the characteristic slope of GW decay would serve as the first indirect detection of low-frequency GWs. At the same time, it would show that SMBHBs can produce

[1] In this regime, the separation of the binary is smaller than the size of the BLR and the gas is bound to both BHs. Therefore, the line profiles are very complex and discovery of SMBHBs through line diagnosis is challenging.

bright electromagnetic emission even at the late stages of the merger [Noble *et al.*, 2012; Farris *et al.*, 2015a].

Recently, [Graham *et al.*, 2015a][hereafter G15] reported the detection of a population of SMBHB candidates with optical periodicity, and suggested that the period distribution is consistent with a population of near-equal-mass SMBHBs in the GW-dominated regime. The search was conducted with data from the Catalina Real-Time Transient Survey (CRTS; [Drake *et al.*, 2009; Mahabal *et al.*, 2011; Djorgovski *et al.*, 2012]), an all-sky, time domain survey, in unfiltered visible light, calibrated to Jonhson V-band (e.g., see § 3 in [Drake *et al.*, 2013]), with a limiting magnitude of 19-21.5.[2] A particularly compelling member of this sample, PG1302-102, is a bright (median V-band mag ∼15) quasar at redshift $z = 0.2784$ with an observed period of 5.2 ± 0.2 yr [Graham *et al.*, 2015b].

[Liu *et al.*, 2015] also identified a SMBHB candidate in the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS; [Kaiser *et al.*, 2010]) Medium Deep Survey with an observed period of 542 ± 15 d and an estimated separation of 7 Schwarzschild radii. PSO J3334.2028+01.4075 was not confirmed as a periodic quasar in G15, although the photometric precision of the two surveys is not comparable. We also note that, according to population models [Haiman *et al.*, 2009], the discovery of such a close binary is extremely unlikely, for the small sample of quasars they analysed, due to the short (< 100 year) lifetime of a massive binary at this separation.

[Zheng *et al.*, 2016] reported the detection of a SMBHB candidate in the radio quiet quasar SDSS J0159+0105. The source was selected from the analysis of CRTS light curves in a small sample of ∼350 quasars in Stripe 82. This candidate shows two periodic components at ∼1500 d and ∼740 d, which were attributed to the redshifted orbital period of a putative SMBHB with separation of 0.013 pc and half of the orbital period, respectively, as expected from hydrodynamical simulations. The quasar was not identified as a periodic source in the sample of G15; it is possible that the existence of the multiple periodic components

[2] CRTS combines data streams from three distinct Schmidt telescopes: (1) the 0.7 m Catalina Sky Survey (CSS) telescope, (2) the 1.5 m Mount Lemmon Survey (MLS) telescope (both located in Arizona) and (3) the 0.5 m Siding Springs Survey (SSS) telescope, which was located in Australia and operated until July 2013. In 30 s exposures, the telescopes achieve nominal 5σ detection limits of 19.5, 21.5 and 19.0, respectively.

decreased the significance of the primary component in the search developed by G15.

In the present paper, we perform a systematic search for periodically varying quasars in the photometric dataset from the Palomar Transient Factory (PTF), a large synoptic survey well suited for this search. PTF has a few advantages over the aforementioned surveys. With a 5σ limiting magnitude of $\sim$20.5, it allows the detection of fainter and hence more distant quasars compared to CRTS, while the higher cadence allows a search for periodicity at shorter time-scales.[3] Furthermore, PTF covered a much larger fraction of the sky ($\sim$3,000 deg^2) compared to the area from Pan-STARRS Medium Deep Survey (80 deg^2), thus offering the possibility to analyse a significantly larger sample. We identify periodic quasars via unusually high peaks in the Lomb-Scargle periodograms of their optical light curves. We then assess the statistical significance of periodic variability by simulating time series that exhibit stochastic damped random walk (DRW) variability. The DRW gives a good description of quasar variability in general [Kelly *et al.*, 2009; Kozłowski *et al.*, 2010], but our statistical analysis improves on previous work, by finding the best-fit DRW model to each individual quasar.

This paper is organized as follows: In § 2.2, we describe the PTF survey, the quasar sample, and the algorithm we use for the detection of periodic variability. In § 2.3, we present our findings. A discussion of our results follows in § 2.4, and our findings and their implications are summarized in § 2.5.

2.2 AGN Sample and Methodology

2.2.1 Palomar Transient Factory

The Palomar Transient Factory (PTF) was an optical time-domain survey designed to explore the transient and variable sky. The scientific phase of the survey lasted from 03/2009 to 12/2012. The observations were made at Palomar Observatory with the 48-inch Samuel Oschin Schmidt telescope, equipped with the CHF12K camera.[4] With 11 (out of initially 12)

[3] G15 imposed a minimum requirement for the period at 400 d.

[4] The camera was previously used at the Canada-Hawai-France Telescope and was modified for PTF [Rahmer *et al.*, 2008].

light-sensitive CCDs (2 k×4 k pixels), the camera provided a wide field-of-view of 7.26 $\deg^2$ and median seeing of $2''$. The scientific goals and the technical aspects of the PTF project are detailed in [Rau *et al.*, 2009] and [Law *et al.*, 2009]. Here we briefly recapitulate the main features of the survey relevant for the present paper.

The PTF survey was conducted mainly in Mould-R and SDSS-g bands. More specifically, initially, the majority of the images were taken using the R-filter, whereas, from 01/2011, the two filters were alternated between dark (g- band) and bright (R-band) nights. In 60 s exposures, the camera achieved 5σ limiting magnitudes of $m_R \approx 20.6$ and $m_g \approx 21.3$. PTF covered a total footprint of $\sim$8,000 $\deg^2$, consisting of the entire Northern sky with declination $> -30^\circ$ with the exception of the Galactic plane. The observing time was mainly devoted to the *transient search experiment*, which covered a large part of the sky ($\sim$2700 $\deg^2$) with an average 5 d cadence, and the *dynamic cadence experiment*, which was activated at the detection of interesting transients for intensive follow-up.

The data were stored and processed at the Infrared Processing and Analysis Centre (IPAC). For each image, the source positions were identified using `SExtractor`, a standard algorithm to generate source catalogues from images in large scale-surveys [Bertin and Arnouts, 1996]. The flux of each source was calculated using aperture photometry, and the photometric measurements were calibrated by comparing the PTF magnitudes for a set of standard stars to the relevant SDSS values (for details, see [Ofek *et al.*, 2012a; Ofek *et al.*, 2012b; Laher *et al.*, 2014]). When retrieving light curve data, we use a magnitude cut in both filters to remove saturated sources and sources well below the single-exposure, 5σ detection limit. We additionally exclude data points with problematic photometry identified either by `SExtractor` or by the IPAC reduction software (e.g., see flags below). In detail, we select only the measurements that meet the following criteria:

1. The magnitude (in either filter) is in the range $12 < \mathrm{mag} < 22$.
2. The photometric error is less than 1 mag ($\sigma < 1$).
3. All `SExtractor` flags are off, except for BIASED_PHOTOM or SATURATED.
4. All IPAC photometry flags are off, except for HALO or GHOST.

2.2.2 Light curves and sample selection

We used the *Half Million Quasars catalogue (HMQ)*[5] as the input catalogue to select the sample of targets for our analysis. HMQ includes all the type-I quasars (QSOs), AGNs, and BL Lac objects published in the literature prior to January 15, 2015, including the most recent data release from SDSS (DR12; http://www.sdss.org/dr12/).

We selected sources within a radius of 3″ from the input positions. For each source, we extracted two light curves (R-band and g-band), when available. From the 424,748 spectroscopically confirmed quasars in the catalogue, 278,740 were observed at least once in PTF. Among those, 99,630 were observed at least once in both filters, 172,829 (and 6,281) were observed in R-band, but not in g-band (and vice versa).

Although the selection of data points, described above, guarantees high-quality light curves, it is not surprising that some outliers are not flagged by the automated pipeline. We remove the remaining outliers with the following procedure: We apply a 3-point-median filter to the light curve, and subsequently fit the filtered light curve with a 5^{th} order polynomial. We discard points that deviate by $\pm 3\sigma$ from the polynomial fit [Palanque-Delabrouille *et al.*, 2011].

As a result of the observing strategy, described in § 2.2.1 (i.e. alternating between the regular 5 d cadence survey and the intensive follow-up of transients), the extracted light curves show a very large diversity in terms of the total number of observations and the sampling rate. For instance, several fields were covered only a few times in the course of the 4 yr survey (∼25% of the R-band light curves have between 1-10 data points), whereas fields in which a transient was detected were covered with high cadence, resulting in light curves with sampling rate as high as one observation every 2 min and a maximum number of ∼6000 observations (see Figs. 2 and 3 in [Price-Whelan *et al.*, 2014] for illustrations).

Additionally, since the high-frequency (intra-night) variability is not relevant for our periodicity search,[6] we bin the observations taken within the same night. More specifically,

[5] This is a subsample of the Million Quasar Catalogue v.4.4; https://heasarc.gsfc.nasa.gov/W3Browse/all/milliquas.html (HMQ; [Flesch, 2015])

[6] We search for binaries with periods of several weeks. Given our sample size, the detection of any source with a shorter period (i.e. few days) would be extremely unlikely; see § 2.3.3 below.

we replace all data points y_{ij} taken during the j^{th} night ($i = 1, 2, ...N_j$, where N_j is the total number of data points during the j^{th} night), with their average Y_j, weighted by the inverse variance of the measurements:

$$Y_j = \sum_{i=1}^{N_j} w_{ij} y_{ij}. \tag{2.1}$$

Here $j = 1, 2, ...k$, with k denoting the total number of nights, during which at least one observation was taken. We calculate the weights w_{ij} as follows:

$$w_{ij} = \frac{1}{W_j}\frac{1}{\sigma_{ij}^2} \qquad \text{and} \qquad W_j \equiv \sum_{i=1}^{N_j} \frac{1}{\sigma_{ij}^2}, \tag{2.2}$$

where σ_{ij} is the photometric error for the i^{th} datapoint during the j^{th} night. We apply the same weights to the times t_{ij} of individual observations, whereas the photometric errors on the binned fluxes are calculated via error propagation,

$$T_j = \sum_{i=1}^{N_j} w_{ij} t_{ij} \qquad \text{and} \qquad \Sigma_j = \left(\sum_{i=1}^{N_j} w_{ij}^2 \sigma_{ij}^2 \right)^{1/2}. \tag{2.3}$$

The final light curve consists of k data points $\{T_j, Y_j, \Sigma_j\}$.

In Fig. 2.1, we illustrate the diversity of the extracted light curves by showing the total number of data points in the binned light curves (which practically represents the total number of nights each quasar was observed) in R-band (N_{R-band}) versus the baseline of the light curves ($\text{MJD}_{max} - \text{MJD}_{min}$, where MJD is the Modified Julian Date of the observation). We also illustrate the temporal sampling of a few representative light curves in the embedded panel. We did not analyse light curves with fewer than 50 data points (i.e. quasars that were observed for fewer than a total of 50 nights). This resulted in a large cut, excluding $\sim$80% of the initial quasar sample, as shown in the hatched region of the main figure and below the dashed-dotted line in the embedded panel. While the loss is significant, we found that, for a reliable periodicity search, it is necessary to impose such a cut, since the number of independent frequencies in the periodogram is defined by the number of points in the light curve.

As mentioned above, most of the fields were more extensively covered in R-band, whereas the coverage in g-band is typically quite sparse.[7] In Fig. 2.2, we present the histogram of the

[7]This is also apparent from the number of quasars observed at least once in only one filter, stated above.

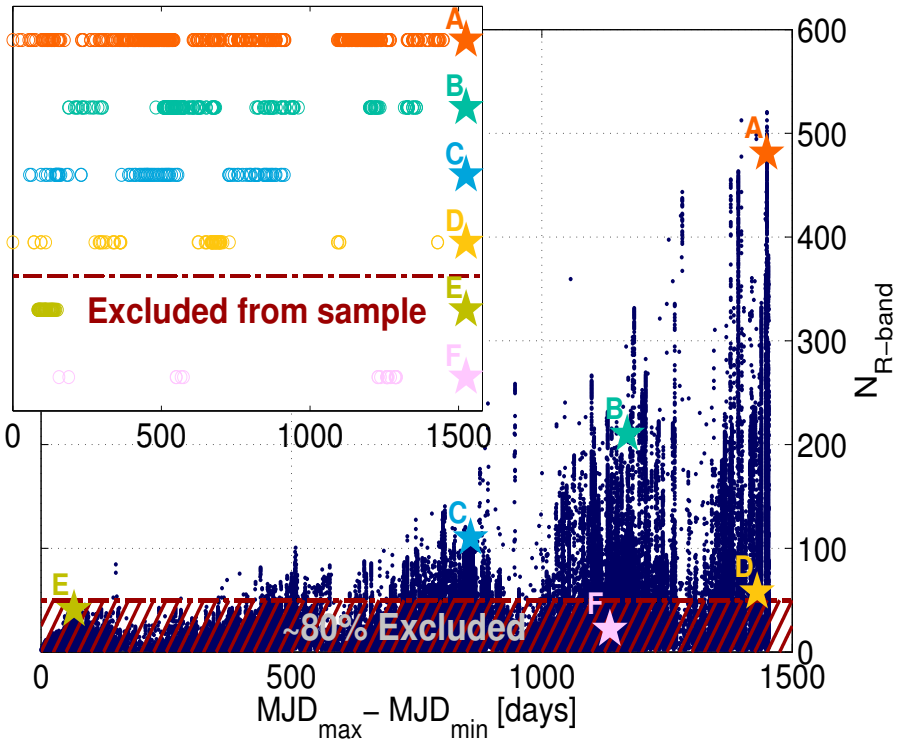


Figure 2.1: Total number of data points (i.e. number of nights) versus the total length of the baseline of each binned R-band light curve. The hatched region shows the light curves that were excluded from the final sample, because the source was observed for fewer than 50 different nights. The time sampling of a few representative light curves is shown in the embedded panel.

fraction of the total number of data points in the binned light curve in R-band (N_{R-band}) over the relevant number for g-band (N_{g-band}), for the population of QSOs that were covered in both filters. We note that less than 5% of the sources have more observations in g-band, compared to R-band. For this subsample, we plot the histogram of the number of data points in the binned g-band light curve, in the embedded panel; only a small number of sources (111 quasars) is in compliance with our minimum requirement of 50 data points (the quasars that are not consistent with our minimum requirement are shown in the hatched region of the embedded histogram). We conclude that it is more advantageous to focus only on the R-band light curves for the periodicity search. Therefore, our final sample consists of 35,383 QSOs, which have at least 50 data points in the binned R-band light curve.

We emphasize that there is no obvious selection effect, in terms of magnitude or redshift, for the sample we analysed, compared to the entire quasar sample. To illustrate this, in Fig. 2.3 we show the redshift-magnitude distribution for all the quasars in the HMQ (light green points), the quasars that were observed at least once in R-band (light blue points) and the final sample of quasars with well-sampled light curves (orange). The side panels show the respective histograms of redshifts and magnitudes with the same colour coding. The most distant and faint quasars are outside of the detection capabilities of PTF and were not included in our sample.

Finally, in Fig. 2.4 we show the distribution of quasars on the sky, with the same colour-coding as before. Most of the spectroscopically confirmed quasars were identified in the SDSS database; the green points therefore roughly trace the SDSS footprint. Similarly, the blue points show the overlap of SDSS with the PTF footprint, and the orange points trace the PTF fields with significant coverage (i.e. our final sample of quasars with at least 50 nights of observations).

2.2.3 Statistical search methodology

We develop an automated algorithm to systematically analyse light curves and detect periodic variability, based on the generalized version of the Lomb-Scargle periodogram [Zechmeister and Kürster, 2009].[8] The Lomb-Scargle periodogram ([Lomb, 1976]; [Scargle, 1982])

[8]For the analysis, we use the astroML python package [Vanderplas *et al.*, 2012; Ivezić *et al.*, 2014].

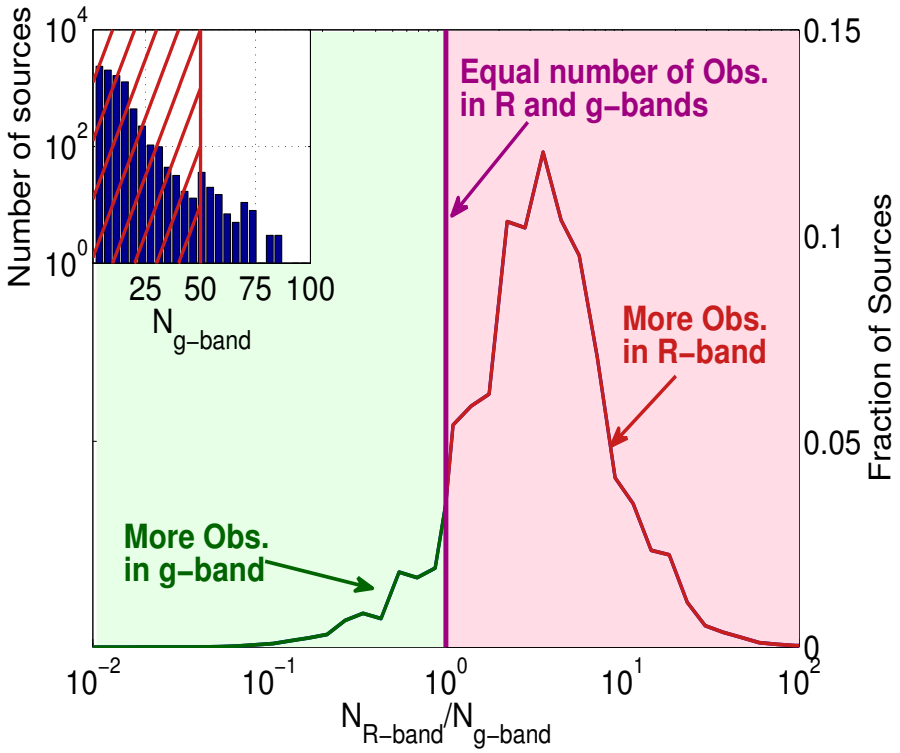


Figure 2.2: Histogram of the total number of data points in the binned light curve in R-band versus g-band. For the small subsample of quasars that were observed more times in g-band, the histogram of the total number of data points (in the binned g-band light curves) is shown in the embedded panel, with the hatched region representing quasars that are not consistent with the minimum requirement of 50 distinct nights.

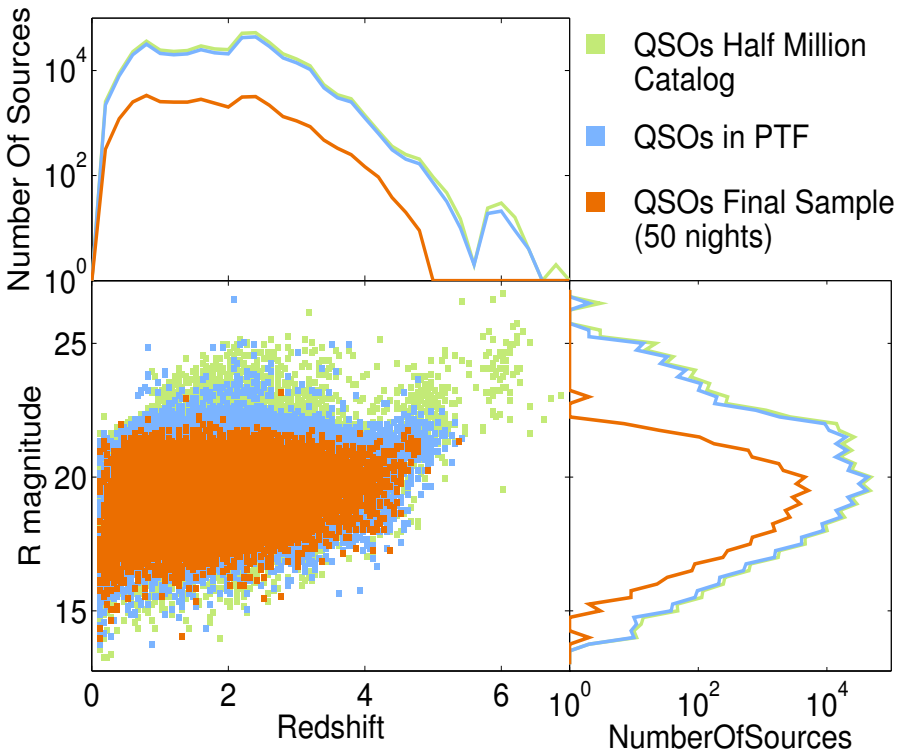


Figure 2.3: Redshift and magnitude distribution for all the quasars in the HMQ (light green), the quasars that were observed at least once in R-band (light blue) and the quasars that were observed for at least 50 distinct nights and constitute our final sample (orange). The respective histograms are also shown with the same colour coding. The figure shows that the sample we analysed is representative of the entire population of quasars.

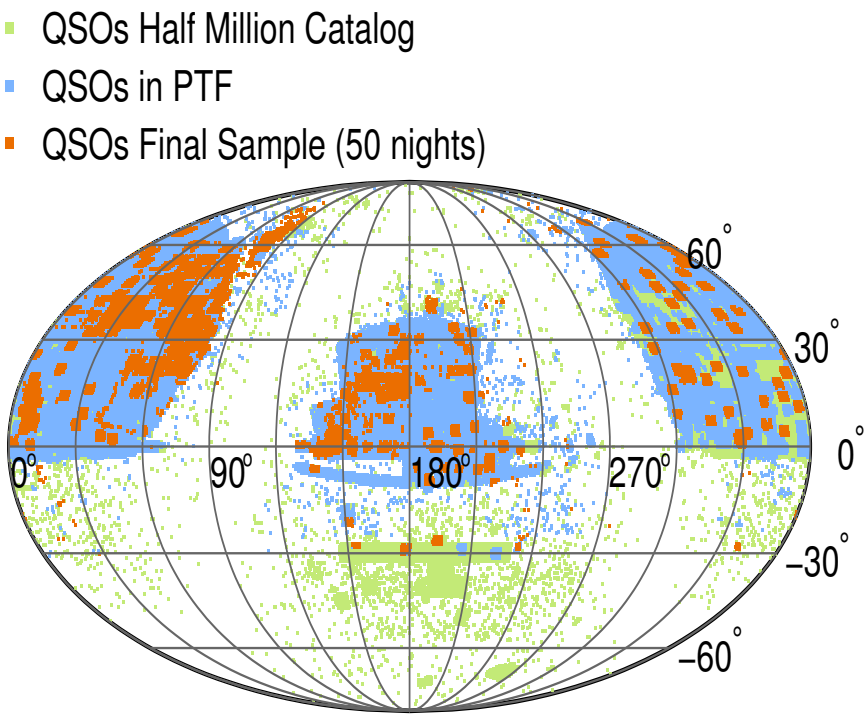


Figure 2.4: Distribution of quasars on the sky in equatorial coordinates. The colour-coding is the same as in Fig. 2.3.

is a standard method for detecting periodic signals in light curves with non-uniform temporal sampling. We calculate the periodogram for 1000 trial frequencies, uniformly distributed on a logarithmic frequency grid from $f_{min} = 1/T_{data}$ to $f_{max} = 1/T_{min}$, with $T_{data} = MJD_{max} - MJD_{min}$, the baseline of the light curve and $T_{min} = 60\,d$.[9]

A periodic signal is detected when a peak with significant power is identified in the periodogram. We use the power of the peaks as the statistic to test the null hypothesis of pure noise. More specifically, for every identified peak, we calculate the probability that a peak of similar power arises from the background by simulating light curves that mimic the quasar variability, and computing periodograms of repeated realizations of the simulated data.

Several authors [Kelly *et al.*, 2009; MacLeod *et al.*, 2010; Kozłowski *et al.*, 2010] have suggested that the optical variability of quasars is successfully modeled by a damped random walk (DRW) process, described by an exponential covariance matrix

$$S_{ij} = \sigma^2 \exp\left(-|t_i - t_j|/\tau\right), \tag{2.4}$$

where σ is the long-term variance of the variability, τ a characteristic time-scale and t_i, t_j the different observing times. The power spectral density (PSD) of this model is given by

$$PSD(f) = 4\sigma^2\tau/(1 + 4\pi^2 f^2 \tau^2). \tag{2.5}$$

In this model, the power decreases with frequency for high frequencies, whereas for low frequencies ($f \ll 1/\tau$), the power spectrum becomes flat.[10]

For each quasar, we identify the best-fit parameters (σ and τ) for the DRW model by directly fitting the light curve in time domain. For this purpose, we make use of the publicly available code Javelin v3.1 [Zu *et al.*, 2013]. The algorithm employs a Gaussian likelihood associated with the covariance in eq. (2.4) and samples the posterior distribution function

[9]The minimum period (or equivalently the maximum frequency) we probe is set to 60 d. Given the sample size, the detection of a SMBHB with such a short period is unlikely, see § 2.3.3 below.

[10]Recent work on quasar variability has shown that the variability may deviate from the DRW model, at high frequencies [Graham *et al.*, 2014; Kasliwal *et al.*, 2015]. However, the deviations are expected to occur at frequencies outside our range of interest, and therefore are not significant for this work; see also § 2.4.2 below.

of σ and τ with a Markov Chain Monte Carlo (MCMC) sampler[11] (see also [Zu *et al.*, 2013] and the documentation of the code[12] for a detailed description). Moreover, we use log-normal priors for σ and τ taking into account the scaling relations found by [MacLeod *et al.*, 2010].

In more detail, we use a prior distribution for each individual quasar according to its observed properties (e.g., magnitude, BH mass). The mean σ and τ of these distributions are estimated from the fitting formulae [MacLeod *et al.*, 2010]:

$$\begin{aligned}\log(\tau_{RF}) =& 2.4 + 0.17\log\left(\frac{\lambda_{RF}}{4000\mathring{A}}\right) + 0.03\,(M_i + 23) \\ &+ 0.21\log\left(\frac{M_{\rm BH}}{10^9 M_\odot}\right), \end{aligned} \tag{2.6}$$

$$\begin{aligned}\log\left(\sqrt{2}\sigma\right) =& -0.51 - 0.479\log\left(\frac{\lambda_{RF}}{4000\mathring{A}}\right) + 0.131\,(M_i + 23) \\ &+ 0.18\log\left(\frac{M_{\rm BH}}{10^9 M_\odot}\right) \end{aligned} \tag{2.7}$$

where τ_{RF} is the characteristic time-scale τ in the rest frame of the quasar and λ_{RF} is the effective wavelength of the R-band filter, $\lambda = 6516.05\,\mathring{A}$, in which the observations were made, in the rest frame of the source (i.e. for a quasar at redshift z, $\tau_{RF} = \tau(1+z)^{-1}$ and $\lambda_{RF} = \lambda(1+z)^{-1}$). We calculate the absolute i-band magnitude M_i, k-corrected to $z = 2$, from the mean apparent R-band magnitude in the HMQ, adopting the mean quasar spectral energy distribution (SED) from [Elvis *et al.*, 1994] and an opacity model for the Lyα forest from [Madau *et al.*, 1996]. We adopt the virial black hole mass $M_{\rm BH}$ estimated from the width of broad lines [Shen *et al.*, 2008]. For quasars that do not have a mass estimate, we draw the mass from the expected Gaussian distribution given the absolute i-band magnitude (e.g., [MacLeod *et al.*, 2010] based on the results from [Shen *et al.*, 2008]),

$$p\,(\log M_{\rm BH}|M_i) = \frac{1}{\sqrt{2\pi}s}\exp\left[-\frac{(\log M_{\rm BH} - m)^2}{2s^2}\right], \tag{2.8}$$

with $m = 2.0 - 0.27M_i$ and $s = 0.58 + 0.011M_i$. The variance of the prior σ and τ distributions, for each quasar, is determined by propagating the uncertainties of the fitting

[11] The MCMC chain consists of 20,000 iterations in the σ-τ parameter space (10,000 iterations for the burn-in process and 10,000 iterations for the actual chain).

[12] http://www.andrew.cmu.edu/user//yingzu/codes.html

coefficients in eq. (2.6) and (2.7),

$$Var(\tau) = \tau^2 \log(10)^2 Var(\log \tau), \tag{2.9}$$

$$Var(\log \tau) = 0.2^2 + 0.02^2 \log\left(\frac{\lambda_{RF}}{4000\mathring{A}}\right)^2 + 0.04^2 \left(M_i + 23\right)^2 + 0.07^2 \log\left(\frac{M_{\rm BH}}{10^9 M_\odot}\right)^2, \tag{2.10}$$

and similarly for σ (Table 1 in [MacLeod *et al.*, 2010]). We also note that the uncertainty in the BH mass measurement is included in the uncertainty of the fitting coefficients.

Following the prescription from [Timmer and Koenig, 1995], we generate evenly sampled light curves (temporal resolution Δt=1 day) that exhibit DRW variability (with the PSD in eq. 2.5), fixing the values of σ and τ at the median of the respective posterior distribution. Next, we downsample the time series to match the observing times of the quasar light curve under consideration.[13] We add Gaussian deviates with zero mean and standard deviation equal to the photometric uncertainty of each point to incorporate the measurement errors, and shift the generated light curve by a constant to match the observed mean magnitude. Finally, we calculate the periodogram using the same frequency grid as for the actual observed time series.

Since the noise spectrum is frequency-dependent, it is more meaningful to assess the statistical significance of the identified peaks compared to the local background (i.e. within a relatively narrow frequency range). Hence, we divide the frequency grid into 25 logarithmically spaced frequency bins, each containing 40 trial frequencies.[14] For each frequency bin, we identify the peak with the maximum power and compare it to the distribution of maxima (within the same frequency bin), from the periodograms of the generated DRW light curves.

For each quasar, we simulate 250,000 DRW time series to account for the trial factors

[13] Downsampling at the observation times ensures that aliasing peaks from the uneven sampling will not be falsely detected as periodic signals (see, e.g., [Charisi *et al.*, 2015] for a discussion of aliasing peaks in the periodogram of the quasar PG1302-102).

[14] We choose relatively narrow frequency bins to ensure a fair comparison with the neighbouring frequencies. Within a narrow bin, the frequency dependence of the noise is not very pronounced and the noise locally resembles white noise.

introduced by the number of frequency bins, and the sample size (*look-elsewhere effect*). We define the P-value of a peak as the number of realizations with at least one peak with power exceeding the power of the peak under consideration divided by the total number of background realizations. A quasar is considered to show significant periodicity, when at least one peak (in any of the 25 frequency bins) is above our significance threshold (P-value$<$1/250,000). Finally, we fit a sine wave with frequency around the frequency of the significant peak to the observed light curve, and exclude candidates that are not observed for at least 1.5 cycles within our baseline.

2.3 Results

2.3.1 Quasars with significant periodicity

We analysed the periodograms of the 35,383 quasars which have been observed for at least 50 distinct nights. We identified 67 quasars with significant peaks, as defined in the previous section. Of these, 50 were consistent with our requirement for a minimum of 1.5 cycles within the data. If we increase the minimum requirement to at least 2 or 3 cycles within the baseline, the number of sources decreases to 42 and 25, respectively. Note that with the P-value threshold defined above, under the null-hypothesis of pure DRW noise, we expect to find, on average, $25 \times 35,383/250,000 \sim 3.5$ such peaks by random chance.

For the identified population, it is crucial to assess the statistical significance for the ensemble of the sources, rather than the significance of individual findings. For this purpose, we generated mock DRW light curves, with σ and τ drawn from the distributions described in Section 2.2.3, downsampled at the observations times of quasars and repeated the entire automated analysis, from identifying the best–fit DRW parameters to calculating the P-value of the peaks. We identified 7 significant peaks in the DRW periodograms, which is a factor of two higher than the theoretically expected number of false positives (see § 2.4.1 for a possible explanation), but with only one having at least 1.5 periods within our data. It is clear that the DRW model alone cannot reproduce the set of significant periodicities found in the quasar sample and, thus, the identified sample of periodic quasars is statistically significant. We emphasize again that the statistical significance refers to the population of

~ 50 candidates, and not necessarily to any of the individual quasars.

In Fig. 2.5, we show the number of sources, both for the observed quasar data (solid curve) and for the mock DRW realizations (dashed curve), which would be identified by our procedure as periodic, as a function of the significance threshold. The figure shows that at low thresholds ($P10^{-2}$), the number of peaks in the quasar sample matches the expectations from pure DRW noise. However, for higher thresholds, we detect an increasingly larger excess of periodic sources in the real quasar sample, compared to the simulated DRW data. The shaded purple region on the right side of the figure highlights findings that are significant after 250,000 iterations (i.e. above the final threshold we considered), whereas the shaded orange region represents the final population (at the same significance level), after we excluded periodic sources with fewer than 1.5 observed cycles.

In Table 2.1, we present the names, coordinates and the observed properties (redshift, average R-band apparent magnitude from HMQ) of the 50 quasars that were identified to show significant periodicity in PTF. The BH mass measurements from [Shen *et al.*, 2008] are shown. When the mass estimate was not available in the catalogue, we include the value drawn from eq. (2.8), as well as the mean and standard deviation of the expected distribution, given the quasar magnitude. In the Table, we also note whether the quasar has been associated with X-ray or radio emission, and the relevant catalogue in which the source was identified.

For the 50 sources with significant periodicity, we extract additional photometric data from the intermediate-PTF (iPTF) and the CRTS, extending the available baseline by at least a factor of 2 (see § 2.4.4 for a discussion). In light curves from both iPTF and CRTS, we bin the observations taken within the same night, as described in § 2.2.3. Additionally, since the data are obtained in different photometric systems (Mould-R for PTF and iPTF and unfiltered V-band for CRTS), we calibrate the different datasets as follows: For each object, we first identify the maximum interval of temporal overlap between the two light curves $[MJD_0, MJD_1]$.[15] In this interval, we interpolate each light curve using a non-parametric model (LOWESS regression) and calculate the offset between the light curves in 100 distinct points evenly spaced within the interval $[MJD_0, MJD_1]$. Next, we shift the

[15]For the calibration, we consider the data from PTF and iPTF as a single extended light curve.

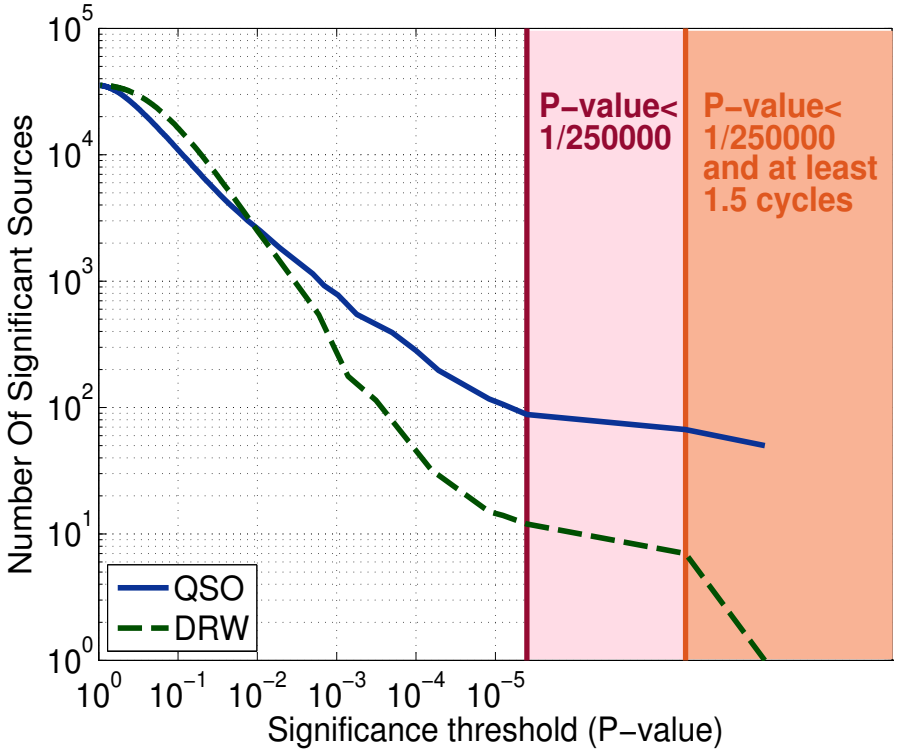


Figure 2.5: Number of sources that would be identified as periodic versus the significance threshold (P-value). The solid line represents the observed quasar data and the dashed line represents the simulated DRW light curves downsampled at the observation times. The purple shaded region shows the number of candidates consistent with our final significance threshold (P-value<1/250,000). The shaded orange region shows the number of candidates that remain in the sample after we exclude sources that were not observed for at least 1.5 periods.

PTF light curve by a constant value defined as the median of the local offsets in the discrete points.

We analyze the periodograms of the extended light curves as before, using the same frequency grid (and frequency bins) as in PTF. We calculate the P-value of the previously identified period, which we present in Table 2.2. In the Appendix, we show the light curves of these candidates (red points for PTF observations, i.e. the light curves we analyzed initially and identified the periodicity, black points for iPTF observations and blue points for observations from CRTS), along with their best-fit sinusoids. In the right column of the figure, we include the light curves phase folded at the observed period. The P-values calculated from the analysis of the extended light curves are also included in the figures. (We note that three of the sources do not have additional data outside of PTF.)

2.3.2 Periodic quasars as SMBHBs candidates

Assuming that the observed periodicity is the redshifted orbital period of a SMBHB, we calculate several properties of the tentative binaries, which we include in Table 2.2. For instance, the separation is given from Kepler's law (assuming, for simplicity, a circular orbit):

$$R = \left(\frac{GMP_{orb}^2}{4\pi^2}\right)^{1/3}, \tag{2.11}$$

where G is the gravitational constant, and M and P_{orb} are the total mass and the orbital period of the binary ($P_{orb} = P_{obs}(1+z)^{-1}$). For reference, we also calculate the projected angular separation,

$$\theta \simeq R/D_A, \tag{2.12}$$

where D_A is the angular size distance for the standard cosmological parameters $H_0 = 67.8\,\mathrm{km\ s^{-1}Mpc^{-1}}$ and $\Omega_m = 0.308$, and a spatially flat universe [Ade *et al.*, 2015].

The gravitational waves emitted by a close SMBHB affects the arrival time of radio-pulses, when the waves cross the line of sight between pulsars and the Earth, inducing a time residual t_r in the arrival time. We estimate these time-residuals (in ns) for the fiducial

inferred SMBHB parameters of each quasar, following [Sesana, 2015]:

$$t_r \simeq 30 \left(\frac{M}{10^9 M_\odot}\right)^{5/3} \left(\frac{D_L}{100\,\mathrm{Mpc}}\right)^{-1} \left(\frac{f}{5\times10^{-8}\,\mathrm{Hz}}\right)^{-1/3}. \quad (2.13)$$

Here D_L is the luminosity distance to the quasar, $M = (M_1 M_2)^{3/5}/(M_1+M_2)^{1/5}$ is the chirp mass of a binary with individual BHs of mass M_1, M_2 and $f = 2/P_{orb}$. In Table 2.2, we include the maximum expected value for the time residuals induced by equal mass binaries ($q \equiv M_2/M_1 = 1$). We note that the estimated residuals are quite small compared to the sensitivity of current PTAs, for all of the tentative binaries. Even though individually undetectable, the binaries still contribute to the stochastic GW background, which was recently constrained by PTAs (e.g., [Shannon *et al.*, 2015]).

For each quasar, we also compute the inspiral time of the orbit due to losses to gravitational waves [Peters, 1964]

$$t_{GW} = \frac{5}{256}\frac{c^5}{G^3}\frac{R^4}{(M_1+M_2)(M_1 M_2)}. \quad (2.14)$$

For reference, we show the inspiral time both for equal ($q = 1$) and unequal mass ($q = 0.01$) binaries.

Finally, for each SMBHB candidate, given its observed period, we calculate the *residence time*, i.e. the time a binary is expected to spend at a given orbital period, or equivalently at a specific orbital separation. The residence time is determined by the rate of orbital decay $t_{\rm res} \equiv -R(dR/dt)^{-1}$. Following [Haiman *et al.*, 2009], we assume that, at large separations, the orbital decay of the binary is dominated by the tidal-viscous exchange of angular momentum with a gaseous circumbinary disc, whereas at small separation the decay is dominated by the emission of gravitational waves. We adopt a standard geometrically thin, optically thick, radiatively efficient, steady-state accretion disc model for the circumbinary disc, coplanar with the orbit of the binary. For the disc parameters, we use the values of the fiducial model in [Haiman *et al.*, 2009] (e.g, viscosity parameter $\alpha = 0.3$, and accretion rate at 10% of the Eddington accretion rate $\dot{m} = \dot{M}/\dot{M}_{\rm Edd} = 0.1$). In the above model, for fixed orbital period, the residence time depends only on the mass of the binary (which is either measured or estimated from the apparent magnitude of the quasar) and the (unknown) mass ratio q of the binary.

2.3.3 Expected SMBHB population

The residence time is a useful quantity to assess the feasibility of a tentative population of SMBHBs. If we attribute the bright phase of quasars to SMBHBs, we can derive the theoretically expected distribution of residence times for the analysed sample of quasars. Since the bright phase of quasars ($t_Q \simeq$ few$\times 10^7$ yr, e.g., [Martini, 2004]) is comparable to the fiducial time-scale for the binary evolution from the outer edge of the circumbinary disc to coalescence [Haiman *et al.*, 2009], quasars will harbor binaries with separations in this entire range, distributed according to their residence time (i.e. a larger fraction of sources at longer residence times, and a smaller fraction at shorter residence times). Therefore, we can express the expected fraction of SMBHBs at residence time $t_{\rm res}$, as a linear function of the residence time, $f(t_{\rm res}) \sim t_{\rm res}/t_Q$. Accordingly, the total number of quasars N, identified as periodic in a sample with size N_{sample}, should scale linearly with the residence time. More specifically,

$$N(t_{\rm res}) = f(t_{\rm res}) \times N_{sample} = t_{\rm res}/t_Q \times N_{sample} \tag{2.15}$$

For the sample in this paper ($N_{sample} = 35,383$) and assuming $t_Q \sim 3.5 \times 10^7$ yr, we expect about 1 quasar with $t_{\rm res} = 10^3$ yr.[16]

The expectation above (eq. 2.15) is an upper limit, as it is derived for an idealized survey, without taking into account any observational limitations. An obvious such limitation is imposed by the finite baseline of the light curves. The cadence and length of the time series defines the range of periods we can identify, which in turn defines the range of residence times we can probe with the available data. This effect can be incorporated into our calculation by estimating the residence time intervals for each light curve. More specifically, if the baseline of the time series is T_{data}, we can search for signals with periods in the interval

$$\left[P_{obs}^{min}, P_{obs}^{max}\right] = \left[T_{min}, \frac{2}{3}T_{data}\right] \tag{2.16}$$

For a quasar at redshift z, this interval is translated into a range of orbital periods

$$\left[P_{orb}^{min}, P_{orb}^{max}\right] = (1+z)^{-1} \times \left[T_{min}, \frac{2}{3}T_{data}\right] \tag{2.17}$$

[16] It follows from the above that we should expect $\sim$0.1 SMBHBs with $t_{\rm res} = 10^2$ yr and $\sim$0.01 with $t_{\rm res} = 10$ yr. Therefore, it is unlikely that any identified periodicity corresponding to $t_{\rm res} 10^3$ yr, is related to SMBHBs (see § 2.4.6 below).

Given the models of binary evolution, discussed in [Haiman *et al.*, 2009], this interval corresponds to a range of residence times $\left[t_{\rm res}^{min}, t_{\rm res}^{max}\right]$.

Fig. 2.6 illustrates the above process for one of the quasars in our sample. The lines trace the evolution of a binary (i.e. the evolution of its residence time), with total BH mass $M \sim 10^8 M_{\odot}$, as the orbit decays from longer to shorter orbital periods, assuming three different mass ratios (blue solid line for $q = 1$, green dashed line for $q = 0.1$ and purple dashed-dotted line for $q = 0.01$). The segments with the different slopes signify the distinct stages of the binary evolution. At long orbital periods, the binary evolution is slow and is dominated by angular momentum exchange with the circumbinary disc (shallower part of the evolutionary tracks), whereas at short orbital periods, the binary enters the GW-driven regime and the evolution is faster (steeper part). The transition between the two regimes occurs at different orbital periods, depending on the mass ratio. The orange shaded region highlights the parts of the binary evolution that are accessible for study, given the cadence and limited baseline of the observed data. The corresponding residence time window is highlighted on the vertical axis (colour-coded according to the mass ratio, as before).

We calculate the range of residence times $\left[t_{\rm res}^{min}, t_{\rm res}^{max}\right]$ for each quasar in the sample, repeating the process described above. Next, for each value of the residence time $t_{\rm res}$, we calculate the fraction of the sources in the sample, for which the specific value of $t_{\rm res}$ is within the observable limits $\left[t_{\rm res}^{min}, t_{\rm res}^{max}\right]$ and define this as the *observable fraction* of the residence time $f_o(t_{\rm res})$. Given the observable fraction, we can calculate the expected number of binaries N_e, accounting for observational limitations, by multiplying eq. (2.15) with $f_o(t_{\rm res})$,

$$N_e(t_{\rm res}) = N(t_{\rm res}) \times f_o(t_{\rm res}). \tag{2.18}$$

In Fig.2.7, we present the distributions of residence times for the sample of quasars with significant periodicity, calculated for the three different mass ratios $q = 1$ (top panel), $q = 0.1$ (middle panel), and $q = 0.01$ (bottom panel). From the distribution we exclude 17 sources, the peaks of which have high P-value in the periodograms of the extended light curves, e.g., P-value$>$ 1% (see § 2.4.4). The corresponding expectations from eq. (2.18) are shown by solid curves in each panel. For reference, the naively expected populations without taking into account the observable fractions (eq. 2.15) are also shown, by the dotted

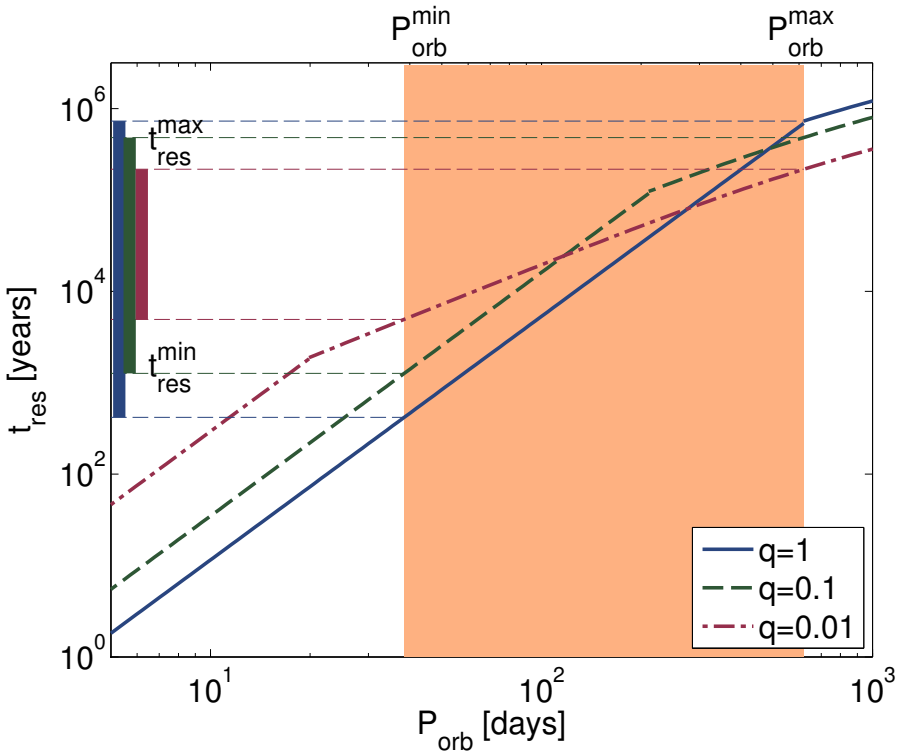


Figure 2.6: Residence time of a SMBHB with a total mass of $M \approx 10^8 M_\odot$, as a function of the orbital period, based on the models in Haiman et al 2009, for mass ratio $q = 1$ (blue solid line), $q = 0.1$ (green dashed line) and $q = 0.01$ (purple dashed-dotted line). The region highlighted in orange shows the accessible orbital periods, given the cadence and finite baseline of the observed light curve. The corresponding residence time range is highlighted on the vertical axis (for different q values, following the colour-coding of the lines).

curves. The vertical dashed line represents $t_{\rm res} = 10^3\,{\rm yr}$, below which it is unlikely to identify SMBHBs in our sample. The figure shows that the observed periodic candidates match the theoretical expectation for the unequal-mass case ($q = 0.01$) better. Additionally, in this case, the number of unlikely findings (with $t_{\rm res} < 10^3\,{\rm yr}$) is smaller compared to $q = 1$ or $q = 0.1$. We also note that, for $q = 0.01$, all of the candidates are in the gas-driven regime, whereas for the equal-mass case ($q = 1$), the majority of the candidates would be dominated by the emission of GWs. We emphasize that in reality, binaries will have a distribution of mass ratios, and not a fixed value; nevertheless, the distribution of the detected residence times favors a low typical mass ratio.

2.4 Discussion

2.4.1 DRW parameter biases

As discussed above, previous work has found that the optical variability of quasars, in general, is successfully described by a two-parameter stochastic noise model (DRW, eq. 2.5). We assessed the statistical significance of the identified peaks in the periodograms by generating mock time series using this noise model. For each quasar, we identify the best-fit DRW parameters (σ and τ), using a Gaussian likelihood and sampling the posterior distribution with an MCMC sampler. Since the DRW model is an essential part of assigning a significance to observed periodicities, we tested the efficiency of recovering input DRW parameters in the PTF sample. For this purpose, we adopted the fitting algorithm from [Zu *et al.*, 2013], used to investigate light curves from the Optical Gravitational Lensing Experiment (OGLE). We generated DRW light curves with known input parameters σ and τ, downsampled at the observation times of the PTF light curves, and used these as inputs to the fitting algorithm. We found that the algorithm is very successful in recovering the input σ, whereas it typically underestimates τ. This low-τ bias will generally tend to underestimate the noise, and over-estimate the significance of peaks. It is therefore a plausible explanation for why we identify slightly more (roughly twice as many) false positives than theoretically expected, when we analyse DRW light curves (e.g., see Fig. 2.5).

In Fig. 2.8, we show the output σ and τ versus the input values of these parameters,

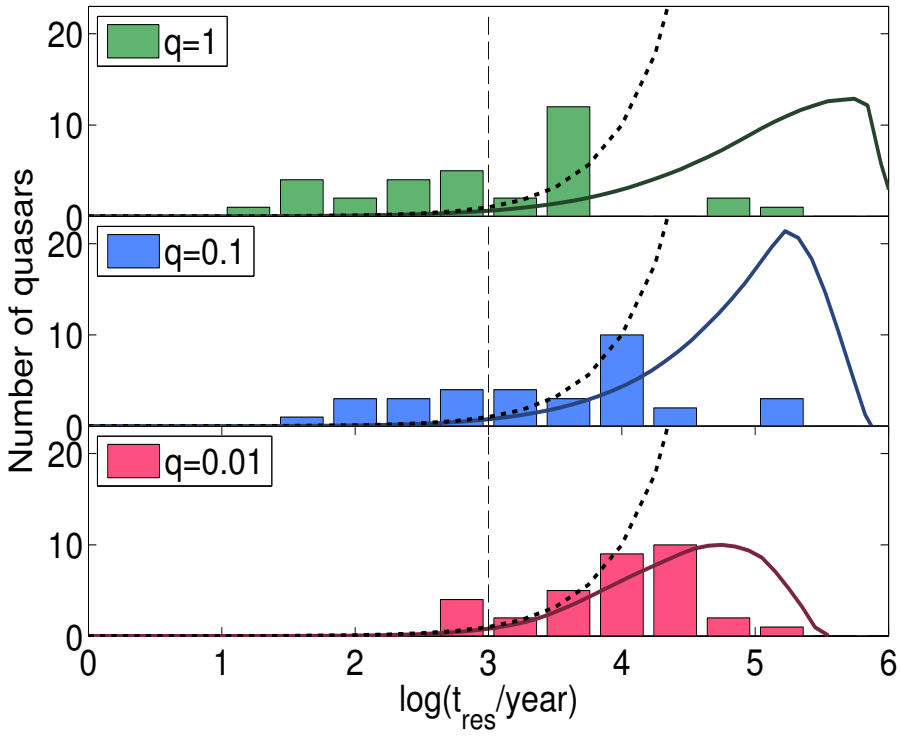


Figure 2.7: Histograms of inferred residence time for the 33 periodic SMBHB candidates (out of the 50 candidates identified in PTF, 33 remain significant when we analyzed the extended light curves with data from iPTF and CRTS), assuming different mass ratios: $q = 1$ (top panel), $q = 0.1$ (middle panel) and $q = 0.01$ (bottom panel). The dotted curves show the distribution expected without considering observational effects, and the solid curves indicate these distributions, after taking into account the limitations imposed by the cadence and baseline of each quasar's observed light curve. The dashed vertical line corresponds to $t_{\rm res} = 10^3\,{\rm yr}$, below which it is unlikely to identify binaries in our sample.

for 10,000 DRW realizations sampled at the observation times of quasars in our sample. The unbiased line is drawn in both cases for comparison. We see that for the PTF light curves, the estimates of τ are biased (a similar bias could of course arise when fitting any light curve with photometric errors and uneven sampling comparable to those in PTF). This bias is possibly the consequence of the following factors: (1) the light curves from PTF have relatively short baselines – in the majority of cases, the input τ is a large fraction of the baseline, (2) the PTF observations are characterized by very uneven sampling, with periods of dense coverage and extensive gaps, (3) the photometric uncertainty can be comparable to the long-term variance σ, making it almost impossible to differentiate between DRW and white noise, which naturally results in underestimation of the parameter τ, and (4) the DRW parameters are drawn from the prior distributions in [MacLeod *et al.*, 2010], which rely on estimates of the absolute i-band magnitude and BH mass of each quasar.

The impact of the temporal sampling is also present in previous papers on quasar variability. For instance, [MacLeod *et al.*, 2010] analysed light curves from SDSS Stripe82, which have long baselines and relatively sparse sampling and found that τ is overestimated for a fraction of their light curves (e.g., see their Fig. 11). They also identified typically longer time-scales τ, compared to the sample in [Kozłowski *et al.*, 2010]. The latter study analysed a sample of quasars from OGLE, with baselines comparable to Stripe82 and better sampling. [MacLeod *et al.*, 2010] suggested that the lack of shorter τ in their sample could be explained either by the different sampling rate of the light curves in the two samples, or by a potential contamination of stars in the sample from [Kozłowski *et al.*, 2010]. The above results, in combination with the underestimation of τ in the PTF light curves, shown in Fig. 2.8, provide a hint that the sampling may introduce a bias in the τ estimation, which may explain the discrepancy between the τ distributions from [Kozłowski *et al.*, 2010] and [MacLeod *et al.*, 2010], although further investigation is required.

We emphasize, however, that the biases in the inferred DRW parameters do not alter our conclusions about the identified periodicity. We guard against this by analysing mock pure-noise realizations with the exact same algorithm that we apply to the quasar light curves: in other words, our null-hypothesis of pure noise will suffer from a similar bias. To show this more explicitly, in Fig. 2.9, we show the distributions of the best-fit σ (left panel)

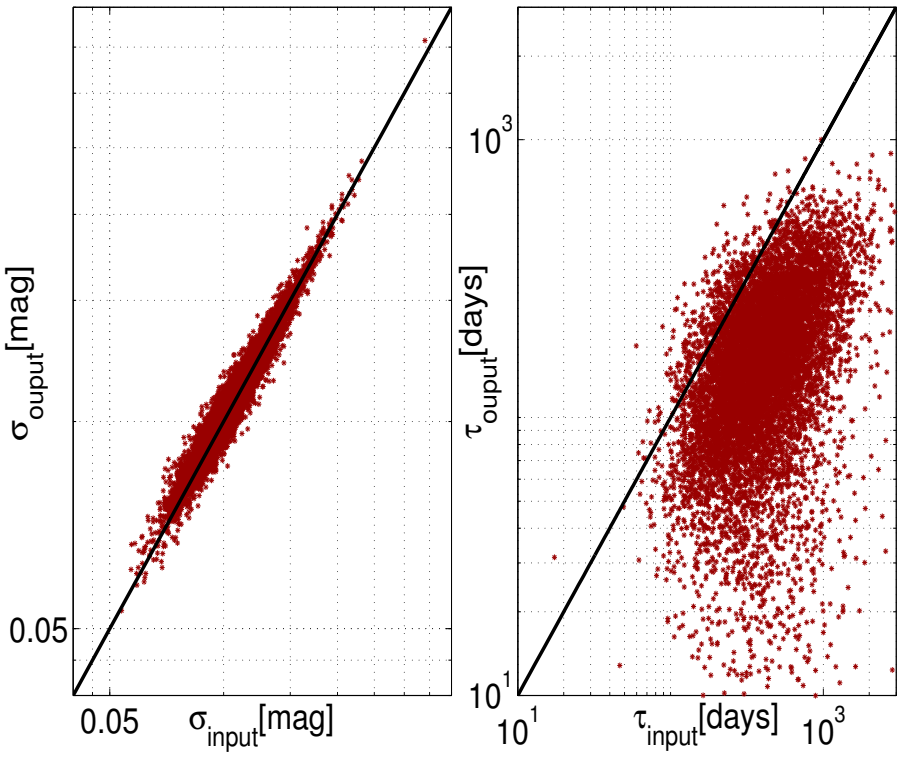


Figure 2.8: Output parameters of the DRW fitting algorithm (σ in the left panel and τ in the right panel) versus the relevant input values, for DRW realizations sampled at the observation times of 10,000 quasar light curves. The solid lines show the unbiased output values.

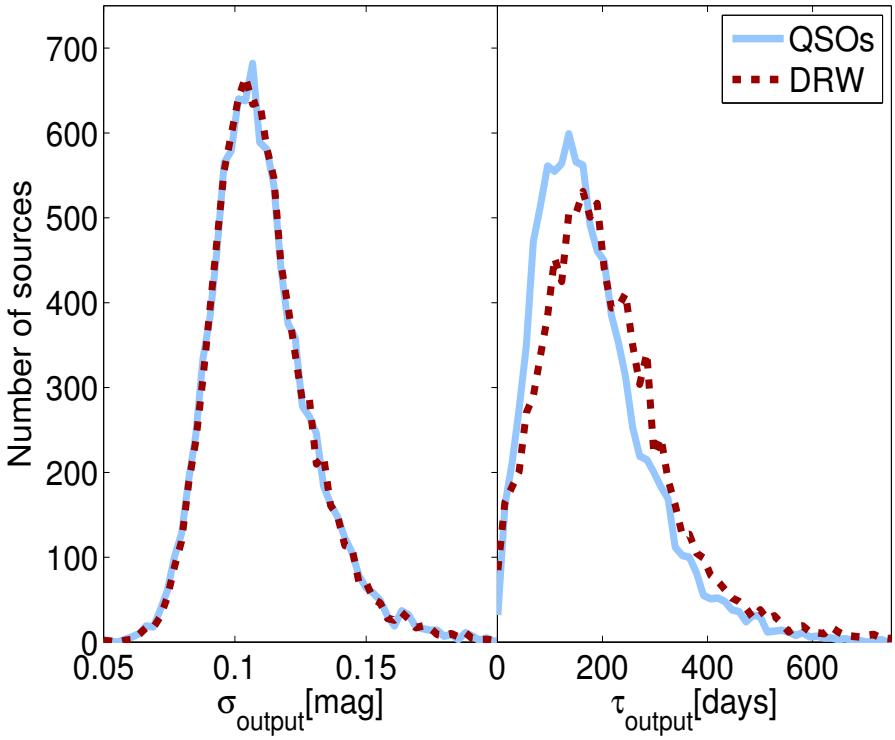


Figure 2.9: Histograms of the best-fit parameters σ (left) and τ (right) for 10,000 quasars (solid curves) and for mock DRW light curves sampled at the same observation times as the quasars (dotted curves). The parameters of the DRW realizations were drawn from the distributions in MacLeod et al 2010.

and τ (right panel) for the real quasars (solid line) and for DRW realizations (dotted line). These distributions are clearly similar. This also means that the distributions from which the input parameters for the DRW realizations were drawn [MacLeod *et al.*, 2010] must be representative of the population, since the bias in the estimation is common both for the simulated and the real data. This also justifies our choice to use the distributions from [MacLeod *et al.*, 2010] as priors in the estimation of the DRW parameters.

Most importantly, as shown in Fig. 2.5, the mock data consisting of pure noise cannot produce the periodicities observed in the sample of quasars. Therefore, we conclude that the population of periodic quasars we identified is statistically significant compared to the DRW model.

2.4.2 Departures from DRW quasar variability

We statistically detected significant periodicity in a population of quasars, compared to the DRW variability. Even though we adopted the most widely accepted model for optical variability [Kelly *et al.*, 2009; Kozłowski *et al.*, 2010; MacLeod *et al.*, 2010], it is possible that a different process may provide a better description of the variability and then the identified periodicity may be less significant. For instance, [Andrae *et al.*, 2013] performed a Bayesian comparison of ∼20 different stochastic and deterministic models in a sizable sample of 6304 quasars from Stripe82. Their results indicate that for a large fraction (∼25%) of these quasars, a combined model of DRW plus a sine wave is favored over pure DRW noise. Additionally, in a small number of quasars (29 out of 6304), they found decisive evidence against stochastic variability and a sinusoidal model is strongly preferred.

Some recent studies have also suggested that the DRW model may be a simplistic description of quasar variability [Mushotzky *et al.*, 2011; Zu *et al.*, 2013; Graham *et al.*, 2014; Simm *et al.*, 2016]. These studies have reported deviations from the DRW model mainly on short time-scales, ranging from a few days up to ∼100 d. Prompted by the above findings, [Kelly *et al.*, 2014] introduced the Continuous-time Auto-Regressive Moving Average (CARMA) models to capture the variability features in quasar light curves. These models offer an extension to the DRW model, since they include higher-order derivatives in the differential equation that describes a stochastic process[17] and therefore allow greater flexibility overall (for details, see [Kelly *et al.*, 2014] and G15 for a discussion of CARMA models and periodic variability).

We note, however, that these reported deviations from DRW may not affect our results significantly, since they typically occur outside of the temporal window we analysed. We restricted our search to periods longer than 60 d and identified only 7 candidates with periods shorter than 300 d. Nevertheless, more generally, it is worth reiterating here that our statistical findings depend on the underlying variability model, and if quasars are proven to follow a more complex stochastic process, our results will need to be validated taking into account the new variability model and calculating the false alarm probability using the

[17]The DRW model is also known as Continuous Auto-Regressive CAR(1) model, and considers only the first order terms of the stochastic differential equation.

new variability model as the null hypothesis for pure noise.

2.4.3 Preference for low mass-ratio SMBHBs

We have shown that the quasars with significant periodic variability are consistent with a population of SMBHBs with a low mass ratio. Fig. 2.7 indicates that the model with $q = 0.01$ is preferred over models with $q = 0.1$ or $q = 1$. G15 detected a sample of 111 periodic quasars in CRTS and suggested that their findings are consistent with a population of equal-mass ($q = 1$) SMBHBs, the evolution of which is dominated by the emission of gravitational waves. In order to see if there is any discrepancy between these two samples, we computed the distribution of residence times for their sample for $q = 0.01$. The important point is that if the quasars are indeed SMBHBs with this low mass ratio, then none of them are in the GW regime, which modifies the expected period distributions. This exercise reveals that the G15 sample also prefers a low mass ratio, similarly to the one identified here.

More specifically, in Fig. 2.10, we present the distribution of residence times for a subset (98 out of 111)[18] of the periodically varying quasars from G15 for mass ratios $q = 1, 0.1$ and 0.01 (top, middle and bottom panel, respectively). We also show, with dotted curves, the theoretically expected distribution from eq. (2.15), for N_{sample}= 243,500. Given that G15 searched for periods between 400 d and 6 yr (at least 1.5 cycles within the 9-year baseline), and given the redshift, magnitude distributions of quasars analyzed in their sample, shown in Fig. 5 and Fig. 6 in G15, we can estimate the observable fraction $f_o(t_{\rm res})$ for this sample. The residence time distributions accounting for the finite baseline, expected from eq. (2.18), are shown by the solid curve in each panel. Moreover, as G15 pointed out, they identify $\sim$25% of the theoretically expected quasars (which could be attributed to only a quarter of all quasars being activated by mergers). We scale the expected distribution by this factor to facilitate the comparison, shown by the dashed curves. We see that the unequal-mass case ($q = 0.01$) indeed fits the observed distributions better.

The figure above also reveals a discrepancy between the results from the two studies, in

[18]We considered only the quasars for which BH mass estimates were available, following G15, see their Fig. 9.

terms of the fraction of quasars that host a SMBHB. Our results indicate that all quasars may harbor a SMBHB, whereas the findings in G15 suggest that this fraction is ∼25%. Nevertheless, the periodicity was identified in two distinct datasets using completely different search algorithms, making a direct comparison challenging. For instance, we note that there is a potential selection effect favoring the brighter quasars in the sample of G15. More specifically, if G15 had limited their analysis to the brighter end of the sample (e.g., quasars with mag<19), they would have identified 104 candidates in a sample of ∼78,000 quasars (see Fig. 5 and Table 2 in G15), resulting in a similar fraction of quasars hosting SMBHBs as in our sample. Therefore, it is possible that the decreased occurrence rate of periodic quasars in the sample of G15, and thus the discrepancy in the two samples, can be explained due to the limited photometric accuracy of CRTS at fainter magnitudes compared to PTF. We will address the question of the fraction of quasars with SMBHBs in a future study.

Mass ratios of SMBHBs have been discussed for a handful of individual candidates in the past, and have been inferred to be low. For instance, the variability of the well-studied SMBHB candidate OJ287 can be explained under the assumption of a very massive ($\sim 10^{10} M_{\odot}$) primary BH with a ∼100 times smaller secondary BH on a highly eccentric orbit, perpendicular to the accretion disc of the primary BH [Valtonen *et al.*, 2012]. Furthermore, [D'Orazio *et al.*, 2015b] proposed that the observed periodic variability of the recently identified SMBHB candidate PG1302-102 may be due to relativistic boosting of a steady accretion flow onto the rapidly moving secondary BH in a highly unequal-mass system, with $q0.05$ favoured. An unequal mass binary (with $q = 0.026$) has also been proposed for the quasar Mrk 231 [Yan *et al.*, 2015]. This candidate was identified from its peculiar spectrum, which the above authors interpreted as a signature of a circumbinary disc with a central cavity, cleared by the motion of the secondary BH (although we note that for such a low mass ratio, a cavity may not be present; [Farris *et al.*, 2015b; D'Orazio *et al.*, 2016]). The recently detected SMBHB candidate SDSS J0159+0105 [Zheng *et al.*, 2016], which shows two periodic components in the optical variability, also requires the existence of an unequal-mass binary with $0.05 < q < 0.8$.

The above results raise the intriguing possibility that SMBHBs with low mass ratios may be more common than near-equal-mass binaries. This is not entirely surprising, in

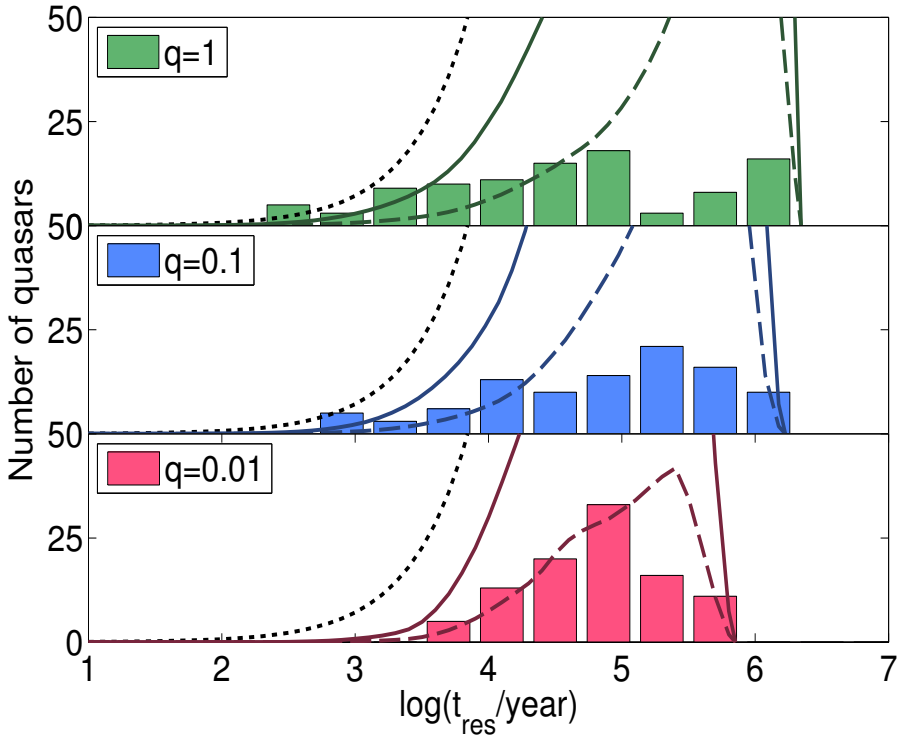


Figure 2.10: Histograms of residence time inferred for 98 of the 111 SMBHB candidates identified by G15, assuming different mass ratios: $q = 1, 0.1$ and 0.01 (top, middle and bottom panel, respectively). The dotted curves show the distribution expected without accounting for observational limitations; the solid curves incorporate these observational constraints. The dashed curves are the same as the solid curves, scaled down by 25%, which is the fraction of all quasars inferred to host SMBHB.

light of cosmological galaxy merger models, which predict that at moderate redshift (e.g., $z < 10$), most BH pairs will have unequal masses [Volonteri *et al.*, 2003]. We note that unequal-mass binaries are fairly unexplored from a theoretical point of view. In particular, this low mass-ratio regime is typically ignored in hydrodynamical simulations of binaries with circumbinary gas discs, with only a handful of exceptions ([D'Orazio *et al.*, 2013; Farris *et al.*, 2015a; D'Orazio *et al.*, 2016]; see also [Shi and Krolik, 2015] for simulations of a $q = 0.1$ binary).

The above conclusion should be considered with caution, since it could be the result of a strong selection effect. Both the algorithm developed here and the one employed by G15 are optimized for detecting sinusoidal variations. Hence, they may be preferentially sensitive to binaries with significant Doppler boosting (e.g., see [D'Orazio *et al.*, 2015b]), which is more prominent for unequal-mass binaries. More nearly equal-mass binaries may produce more "bursty" light curves, as a result of periodic fluctuations in the accretion rate onto the BHs [Hayasaki *et al.*, 2007; MacFadyen and Milosavljević, 2008; Roedig *et al.*, 2012; Noble *et al.*, 2012; D'Orazio *et al.*, 2013; Farris *et al.*, 2014; Farris *et al.*, 2015a; Shi and Krolik, 2015; D'Orazio *et al.*, 2016]. The latter are likely to remain undetected with the current searches for periodicity.

2.4.4 Extended light curves

For the candidates identified in PTF, we extended the light curves adding points from iPTF and CRTS. The iPTF light curves are practically an extension of PTF, since the data are obtained with the same telescope and filter, following a similar observing strategy as in PTF. On the contrary, CRTS is a distinct survey; CRTS covers up to $\sim$2500 deg^2 per night, with 4 exposures per visit, separated by 10 min, over 21 nights per lunation. The observations are obtained in unfiltered visual light and the depth of the survey is typically limited compared to PTF (as mentioned in § 4.1, several telescopes are used for the survey, resulting in different limiting magnitudes for different parts of the sky). For simplicity, here we consider these additional data only for validation of the periodic candidates selected from PTF alone,

rather than attempting to identify periodic candidates from a full combined dataset.[19]

From the extended light curves, we calculated the P-value of the period identified in the PTF light curves. We emphasize, however, that this P-value constitutes only a rough estimate. First, the photometric accuracy of the CRTS dataset is reduced compared to PTF. In several cases, the photometric errors, are comparable to the amplitude of the identified periodicity. Furthermore, despite our efforts to calibrate the PTF-iPTF photometry in order to match the CRTS, systematic effects (e.g., due to different photometric systems) are likely to be present. Third, our full analysis, in the case of the PTF data above, has demonstrated that the P-values can not be directly interpreted as true false alarm probabilities. For instance, when our analysis was applied to random DRW realizations, the number of false positives we identified was by a factor of two higher than the theoretically expected number given the trial factors (see Fig. 2.5 and the related discussion above).

Therefore, at this step, we only excluded sources with high P-values (P-value$\geq$1%). This tentative significance threshold is justified, given that we only analyzed the selected 50 candidates and we restricted the search into one frequency bin (we calculate the significance of the previously identified period, not every possible peak in the periodogram). The number of trial factors is thus reduced compared to the initial PTF search. However, given the sparse PTF light curves from which the periodicity was selected and the fact that the extended data do not always provide a conclusive answer regarding the periodicity, it is crucial that the identified SMBHB candidates are further monitored with similar photometric precision in order to confirm that the periodicity persists for several cycles.

We note that, in a few cases, even though the P-value from the extended light curve is below our threshold, the folded light curve looks inconsistent with the stated period. A possible explanation for this is that we detect higher harmonics of a true periodic signal with a longer period than the one identified in PTF. It is well known that the periodogram of a sine wave convolved with the periodogram of sampling function can introduce peaks at non-trivial frequencies [Roberts *et al.*, 1987]. We stress that the aliasing peaks from

[19]We intend to search for periodicity in the composite light curves from PTF, iPTF and CRTS in a future paper, since such a search will offer the combined benefits of the two surveys, e.g., the long CRTS baseline and the high temporal resolution from PTF and iPTF.

the sampling pattern alone are taken into account in our analysis by generating mock light curves with the exact same time stamps as in the observed data. In Fig. 2.11, we illustrate an example of this effect for quasar PKS 2203-215. The best fit sinusoid corresponding to the most significant periodogram peak within PTF has a period of 497 d (see Appendix), whereas if we consider the extended light curve the period of the best fit sine wave is 2480 d (5 times longer than the period in PTF). This low frequency peak is significant (P-value$<$1/250,000), if we examine the entire frequency range allowed by the composite light curve instead of limiting the search within the PTF baseline. We have also seen this effect in the periodogram of PG1302 [Charisi *et al.*, 2015]. The periodogram shows a significant peak at 300 d, which coincides with a peak from a noiseless sinusoid with period of 1884 d sampled at the observed times. However, if PG1302 was identified in a survey with a shorter baseline, comparable to PTF, the peak at 300 d would be identified as the actual periodicity of the source. Although, falsely identifying a harmonic of the actual period can affect the interpretation of the population of quasars, from visual inspection of the phase folded light curves in our sample, we conclude that this effect is not dominant.

2.4.5 Comparison with periodic quasars in CRTS

G15 analysed a large sample of ($\sim 250,000$) quasars from CRTS. Although CRTS is an all-sky survey, most of the spectroscopically confirmed quasars are from SDSS and are spread mostly over the Northern Hemisphere (see Fig. 2.4). Therefore, there is a significant overlap between the quasars analysed in the two samples (75% of the quasars in our sample were also included in the sample of G15). Despite this overlap, we do not identify any common periodic candidates. Here we clarify the reasons for this.

Among the 111 SMBHB candidates in the G15 sample, 101 were in the initial catalogue we used to extract sources from PTF. We examined a more recent version of the HMQ, and we did not include any low-luminosity AGN and blazars in our sample, since it is uncertain whether the DRW model can describe the variability of these sources. As a result, of the 101 objects, only 77 were covered in PTF, and only 15 of these were consistent with our minimum requirement for 50 nights and were included in our final sample. The baseline of the PTF light curves is shorter than the periods identified by G15 for all of these 15

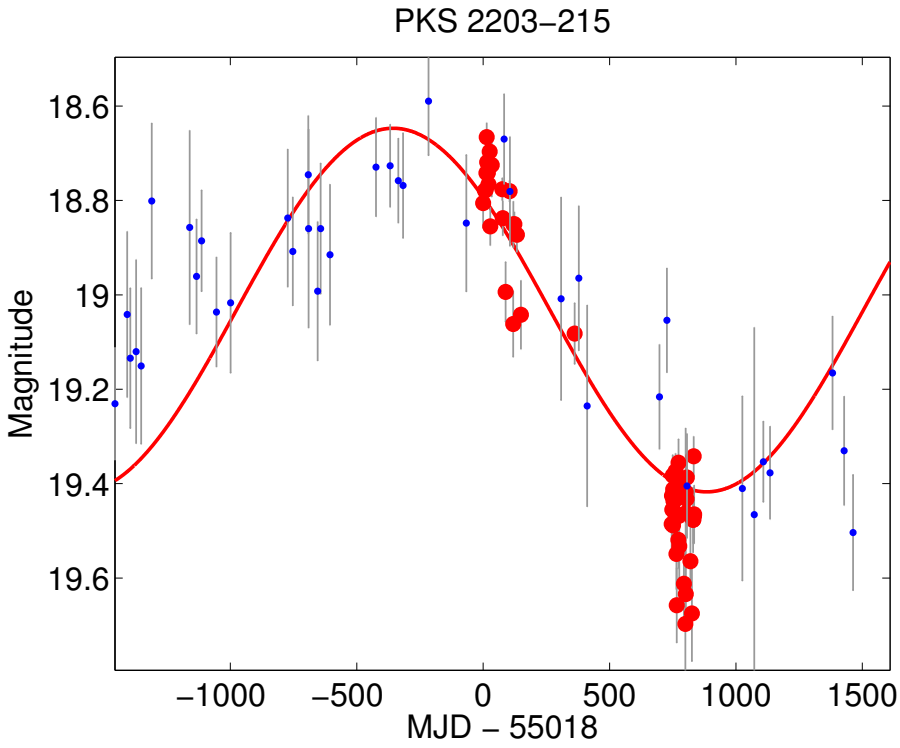


Figure 2.11: The composite light curve of SMBHB candidate PKS 2203-103. Red points indicate PTF observations, blue points observations from CRTS and the red line shows the best fit sinusoid. A sinusoid with period 5 times longer than the one identified in PTF provides a better fit to the combined PTF and CRTS data.

candidates. Therefore, it would be impossible to identify any of these objects as periodic in the sample we analysed.

Likewise, the fact that none of our candidates were identified in G15 is unsurprising. A large fraction of the sources we identified (28 out of 50) have periods below the 400 d cutoff imposed by G15. Also the majority of the candidates are too faint for CRTS. There are only 7 out of 111 candidates with magnitude fainter than 19 mag in G15, even though the vast majority of the sources they analysed is below this magnitude. As mentioned in § 2.4.4, the measurement uncertainty of the CRTS data points in several cases is comparable to the amplitudes of the sinusoids in the PTF sample. This is also obvious from the composite light curves in the Appendix.

2.4.6 Periodic variability in quasars

We have equated the observed optical periods with the (redshifted) orbital period of SMBHBs. According to several hydrodynamical simulations, the mass accretion rate onto the BHs is modulated at the orbital period of the binary [Hayasaki *et al.*, 2007; MacFadyen and Milosavljević, 2008; Noble *et al.*, 2012; Roedig *et al.*, 2012; D'Orazio *et al.*, 2013; Farris *et al.*, 2014; Gold *et al.*, 2014]. In particular, our finding that low mass ratios are favoured support the identification of the optical period as the redshifted orbital period [D'Orazio *et al.*, 2013; Farris *et al.*, 2014; D'Orazio *et al.*, 2016]. A different scenario involves Doppler boosting of the emission arising in the mini-disc around the secondary BH, as it orbits with relativistic velocities. [D'Orazio *et al.*, 2015b] proposed this model to explain the optical and UV variability of PG1302. In this case, the optical and orbital periods would again coincide.

Although the above is reassuring, it is worth noting that the optical periodicity does not necessarily reflect the orbital period. Hydrodynamical simulations of SMBHBs with higher mass ratios ($q > 0.3$) predict the existence of several periodic components in the variability [Roedig *et al.*, 2012; Shi *et al.*, 2012; D'Orazio *et al.*, 2013; Farris *et al.*, 2014]. For instance, [D'Orazio *et al.*, 2015a] associated the observed period of PG1302-102 with the longer orbital period of a lump in the lopsided accretion disc and predicted that the orbital period may be 5-8 shorter than the observed, although the analysis of the periodogram did not reveal additional peaks [Charisi *et al.*, 2015]. More generally, periodic variability of

quasars does not necessarily require the presence of a binary: quasi-periodic modulations can arise around a single BH due to, e.g., Lense-Thiring precession, a warped accretion disc, or the precession of a jet (see G15 for an extended discussion).

2.4.7 Selection Effects

We have analyzed a highly heterogeneous sample of light curves. Here, we explore the role of some potential selection effects and biases among the periodic sources we have identified, which are likely present in our sample. For instance, in Fig. 2.12, we show the histograms of the observed period, P_{obs}, in dark blue and the orbital period $P_{orb} = (1+z)^{-1} P_{obs}$, in light blue, for the periodic quasars in our sample. In the histogram of the observed period, we notice two prominent peaks at $\sim$300 d and $\sim$450 d, while a clear deficit of candidates with periods of $\sim$1 yr is also present. A possible explanation is that our algorithm is less sensitive at specific timescales (e.g., for periods of $\sim$1 yr or for long periods > 500 d) and we miss some genuine periodic sources, leading to the observed scarcity of specific frequencies. Another explanation is that the peaks at $\sim$300 d and $\sim$450 d may reflect the identification of some aliasing peaks from real periodic signals with periods longer than the PTF baseline. An example of this is the periodogram of quasar PG1302-102; in our previous work, we have explicitly shown that at the particular time sampling of that source, a genuine 5.2 yr sinusoid would introduce strong peaks at $\sim$300 d and $\sim$500 d in the periodogram (see Figure 1 in [Charisi *et al.*, 2015]). It is likely that similar aliasing results in the misidentification of some of our periods, as already shown for one source in Fig. 2.11.

Another possible selection effect is illustrated in Fig. 2.13, in which we present the histogram of the phase of all the observations in the light curves of our periodic sample. We notice that there is an excess of observations with phases of $\sim$0.25 and $\sim$0.75 (i.e. at the maxima and minima of the sinusoid, respectively). Although the phase distribution should be uniform in the ideal case of densely and uniformly sampled light curves, the observed distribution could be explained by the significant deviations from the ideal sampling of the PTF light curves (Fig 2.1). For instance, we consider how the two extreme cases of unfortunate sampling, which may be present in our data, can affect the phase distribution: (1) If a true sinusoid is sampled only around the mean (phases 0, 0.5 and 1), the Lomb-

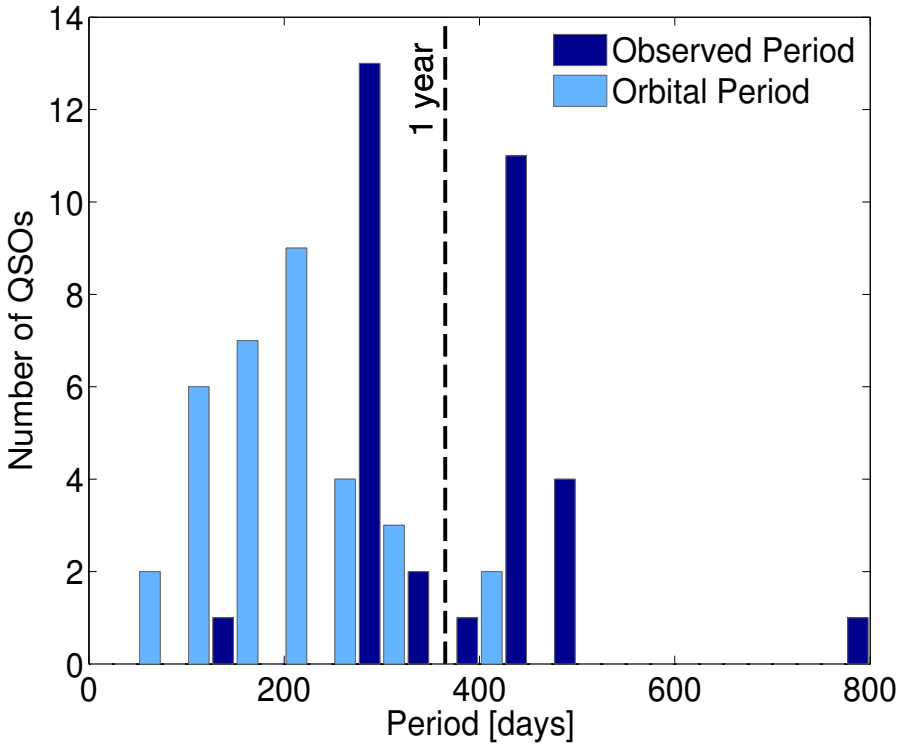


Figure 2.12: Histogram of observed periods (in dark blue) and redshifted orbital periods (in light blue). The dashed line indicates a period of 1 yr; a clear deficit of periods is observed at this timescale.

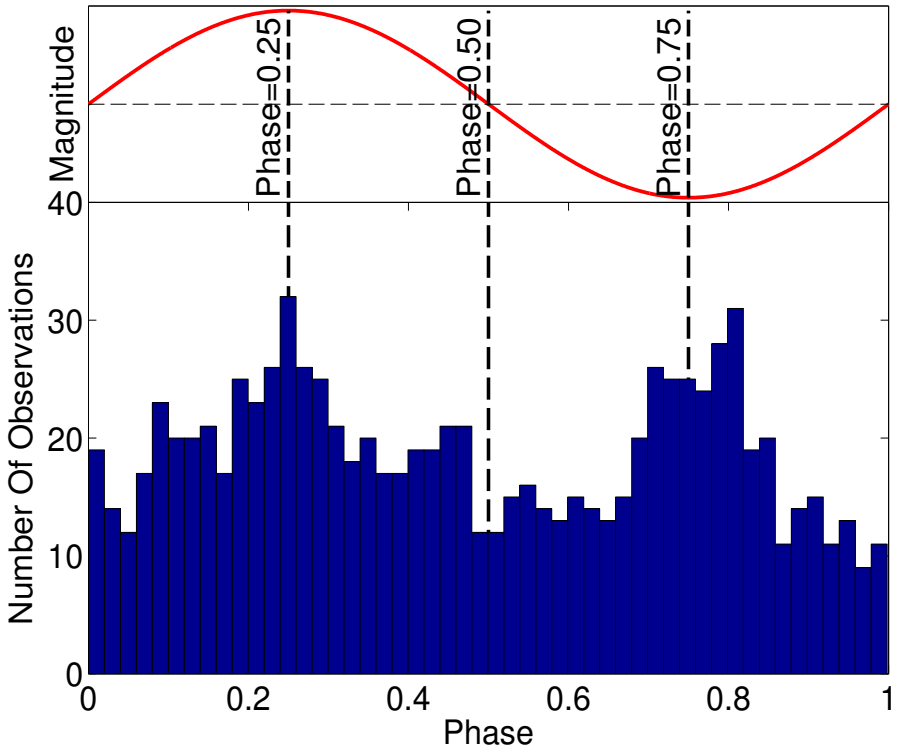


Figure 2.13: Histogram of the phase for all the data points in the sample of quasars with significant periodicity. The histogram shows deviations from the expected uniform distribution. In the top panel, we illustrate a phase folded sinusoid.

Scargle periodogram would fail to detect a peak, and the periodicity would be completely missed, leading to a scarcity of the relevant phases in the overall phase distribution, and (2) If only the minima and the maxima of the sinusoid were sampled, the power of the periodogram peak would be significant and the likelihood of detection is increased, leading to a potential excess of observations with phases $\sim$0.25 and $\sim$0.75. The combination of these effects would, therefore, translate to a deficit of phases near 0, 0.5, and 1, and an excess of phases near 0.25 and 0.75, matching the phase distribution we observe.

On the other hand, [Vaughan *et al.*, 2016] generated mock DRW realizations to reassess the statistical significance of PG1302. They found that false positives, with periods of $\sim$1 yr or less, were detected at an increased rate in cases when the simulated maxima and minima aligned with the quasi-periodic sampling. If a significant number of false detections was present in our sample due to the above effect, it could result in a phase distribution similar

to Fig. 2.13. However, this is unlikely the case in our sample, because we have included the effects of the irregular sampling in the calculation of the false alarm probability. If the DRW model in combination with the irregular sampling could mimic the periodic variability, it would create peaks with significant power in the periodogram and thus the quasar would not be identified as periodic. Additionally, when we simulated DRW light curves and conducted the periodogram analysis over the entire sample (e.g., see Fig 2.5), we did not detect false periodicity that shows clustering around the maxima and minima. In detail, among the 7 DRW realizations that were above our detection threshold, 6 had long periods and were not consistent with our requirement for a minimum of 1.5 cycles within the data. The only case that was consistent with all our requirements for detection and would be identified as a short period quasar did not show clustering of the observations around the minima and maxima of the best fit sinusoid. Finally, we note that the sampling of the PTF light curves is irregular and not quasi-periodic, like the sampling of PG1302.

In order to understand the underlying population of SMBHBs, it is important to understand and correct for the selection effects and biases discussed above. While this is beyond the scope of the present paper, these effects will be carefully explored in a separate follow-up study.

2.5 Conclusions

We developed a statistical search to identify periodicity in the optical variability of quasars. We analysed the data of 35,383 spectroscopically confirmed quasars from the Palomar Transient Factory, and assessed the statistical significance of our findings by simulating stochastic time series that mimic the quasar variability, which we modeled as a damped random walk process. Our main conclusions are the following:

- We detected a statistically significant population of 50 periodic quasars with at least 1.5 cycles within the PTF baseline. Of these periods, 33 remain significant even with the re-analysis of light curves including data from iPTF and CRTS. This identified population is significant as an ensemble of sources rather than as individual detections.

- The periodic sources are characterized with typically short periods of a few hundred

days and fainter magnitudes compared to the previous study in CRTS (G15). Our findings reflect the better photometry at fainter magnitudes and the higher temporal resolution of the PTF light curves compared to CRTS.

- If the identified periodicities correspond to the redshifted orbital periods of SMBHBs, then the period-distribution of this population favors SMBHBs with a low mass ratio ($q \approx 0.01$).

- We found a similar conclusion about the mass ratio for the population of periodic quasars from G15, which was identified in a separate sample with a different selection algorithm. Unequal masses ($q = 0.01$) provide a better fit to their observed period-distributions, as well.

In this paper, we identified quasars with short-period optical variability as promising candidates for SMBHBs. However, given the relatively short PTF baseline and the limited photometric accuracy of the extended data from CRTS, it is essential to further monitor the candidates to confirm that the periodicity persists for several cycles. It is also important to search for additional signatures for the binary nature of the sources (e.g., wiggled radio jets, signs of relativistic boosting, etc). We also note that we have assessed the significance of the detected periodicity, compared to a null-hypothesis assuming that quasar variability is described by the DRW model. Even though this model has proven successful as a general description of quasar variability, the significance of our findings would need to be validated within any other variability model.

Table 2.1: Observational properties of quasars with significant periodicity.

Name	Ra	Dec	z	R-mag	log $(M/M_{\odot})$	X-ray/radio
UM 269*	10.8322440	0.8542910	0.308	17.8	(8.4079) 8.5337±0.3138**	(X), (R)
SDSS J005158.83-002054.1	12.9951310	-0.3483820	1.047	19.3	8.783***	
SDSS J005453.30-003258.3	13.7221040	-0.0.5495320	0.961	19.8	(7.9341) 8.6412±0.3094	
SDSS J023050.06+005843.1	37.7086065	0.9786625	1.447	17.9	(10.02) 9.2635±0.2841	
SDSS J024442.77-004223.2	41.1782445	-0.7064567	0.628	19.7	7.926	
2QZ J095344.7+010354	148.4366915	1.0651954	0.994	19.3	(8.3814) 8.7953±0.3032	
SDSS J104648.62+513912.6	161.7026015	51.6535127	0.587	18.7	(8.853) 8.6646±0.3085	
SDSS J132815.49+361715.9	202.0645650	36.2877730	1.088	17	(9.6039) 9.4686±0.2757	(X)
SDSS J133254.51+371735.5	203.2271569	37.2932211	2.46	19.3	(8.7384) 9.3584±0.2802	
SDSS J133840.66+315936.4	204.6694431	31.9934505	2.944	19.6	(9.4063) 9.3868±0.2791	
SDSS J134553.57+334336.0	206.4732442	33.7266720	0.886	19.3	(9.0644) 8.7308±0.3058	(R)
SDSS J134556.16+343224.5	206.4840056	34.5401540	0.874	19.7	(8.6547) 8.6152±0.3105	
SDSS J140929.76+535930.2	212.3740072	53.9917232	0.863	19	(8.5757) 8.7972±0.3031	
SDSS J141004.41+334945.5	212.5184030	33.8293160	0.63	18.2	8.962	
SDSS J141244.09+421257.6	213.1837413	42.2160210	0.805	18.7	(9.6913) 8.8397±0.3013	(X), (R)
SDSS J142339.44+471240.8	215.9143459	47.2113337	1.24	19.7	(8.5389) 8.8182±0.3022	
TEX 1428+370	217.6691030	36.8177490	0.566	17.2	(8.5293) 9.0492±0.2928	(X), (R)
SDSS J143637.44+090155.5	219.1560102	9.0320841	0.568	18.7	(7.706) 8.6461±0.3092	
SDSS J145713.26+140334.1	224.3052882	14.0594797	2.926	19.3	(9.2106) 9.4641±0.2759	
SDSS J145859.07+153144.7	224.7461508	15.5290922	2.898	17.2	(9.9355) 10.0252±0.2530	
SDSS J150900.70+175114.3	227.2529518	117.8539974	0.742	19.6	(8.6112) 8.5516±0.3131	
FBQS J150911.2+215508	227.2968330	21.919110	0.438	16.4	(8.5381) 9.1195±0.2899	(R)
SDSS J150912.07+204004.6	227.3003130	20.6679610	0.339	18.7	(7.8945) 8.3484±0.3214	(X)
SDSS J151053.24+240943.3	227.7218528	24.1620337	0.807	19.3	(8.3276) 8.6791±0.3079	
SDSS J151243.67+195845.1	228.1819980	19.9792200	0.808	18.1	(8.8351) 9.0038±0.2947	
SDSS J151646.10+221724.7	229.1921240	22.2902190	0.599	18.9	(8.6375) 8.6219±0.3102	
SDSS J152739.97+413234.6	231.9165420	41.5429560	1.014	17.8	(9.232) 9.0332±0.2935	(X)
SDSS J152903.11+223623.8	232.2629760	22.6066240	0.506	18.5	(8.4734) 8.6349±0.3097	
SDSS J153051.79+503440.1	232.7158316	50.5778224	0.928	19.4	(8.8238) 8.7296±0.3058	
SDSS J153251.06+335852.2	233.2127605	33.9811873	1.889	18	8.964	
PDS 898	234.2429720	34.5304110	0.886	18	(8.737) 8.8062±0.3027	
SDSS J155308.65+501436.5	238.2860757	550.2434900	2.774	19.4	(9.2423) 9.4045±0.2783	
SDSS J160322.68+200535.2	240.8445362	20.0931253	2.337	19	(9.872) 9.4081±0.2782	
SDSS J160454.57+315733.5	241.2273914	31.9593278	3.159	20	(9.2593) 9.3223±0.2817	
SDSS J162634.15+325032.6	246.6422999	32.8423962	0.858	19.9	(8.4876) 8.5510±0.3131	
SDSS J170942.58+342316.2	257.4274310	34.387850	1.734	19.4	9.228	(X)
SDSS J171122.67+342658.9	257.8444609	34.4497191	2.132	20.2	(8.7648) 9.0275±0.2937	
SDSS J171617.49+341553.3	259.0728780	34.2648190	1.149	18.5	(9.1003) 9.0959±0.2909	
SDSS J171909.93+344001.3	259.7914157	34.6670451	2.155	18.2	(9.3335) 9.5742±0.2714	
SDSS J212939.60+004845.5	322.4150072	0.8126513	2.266	19.8	(9.2628) 9.1732±0.2878	
SDSS J214036.77+005210.1	325.1532346	0.8694920	0.92	20.2	(8.8142) 8.5087±0.3148	
SDSS J214225.29+001643.2	325.6054161	0.2786872	1.26	19.5	(8.497) 8.7333±0.3057	
SDSS J214357.98+003349.6	325.9916181	0.5638052	2.338	20.2	9.355	
PKS 2203-215	331.6725000	-21.3277778	0.577	18.8	(8.907) 8.6280±0.3100	(X), (R)
SDSS J224008.39+263928.4	340.0349874	26.6579024	2.827	19	(9.2615) 9.5240±0.2735	
SDSS J231733.66+001128.3	349.3902730	0.1912120	0.841	18.2	8.898	(X)
SDSS J232135.73+173916.5	350.3988982	17.6546025	0.842	19.4	(9.1391) 8.6755±0.3080	(X)
2QZ J235800.2-281429	359.5011112	-28.2413889	1.598	19.1	(8.9538) 9.1415±0.2891	
SDSS J235928.99+170426.9	359.8708046	17.0741612	0.714	18.6	(8.995) 8.8004±0.3029	
SDSS J235958.72+003345.3	359.9946880	0.5625920	1.694	19.1	9.076	

*We emphasize with bold, the candidates that remain significant after the re-analysis of the composite light curves.

** The BH mass in the parenthesis is the mass drawn from the distribution in eq. (2.8); The mean and the standard deviation of the distribution are also shown.

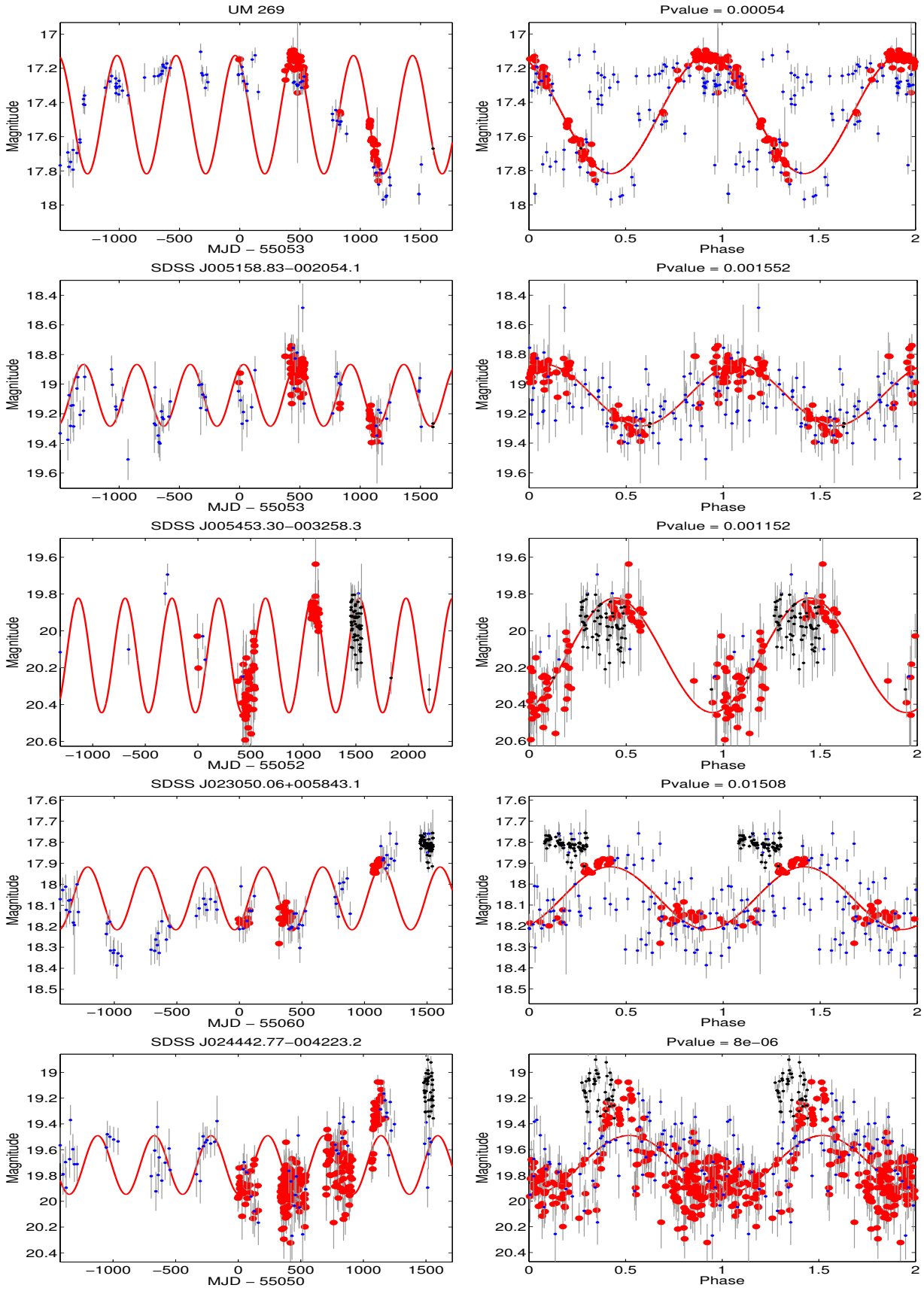
*** BH mass without a parenthesis is taken from the catalog in [Shen *et al.*, 2008], measured from the width of broad lines in quasar spectra.

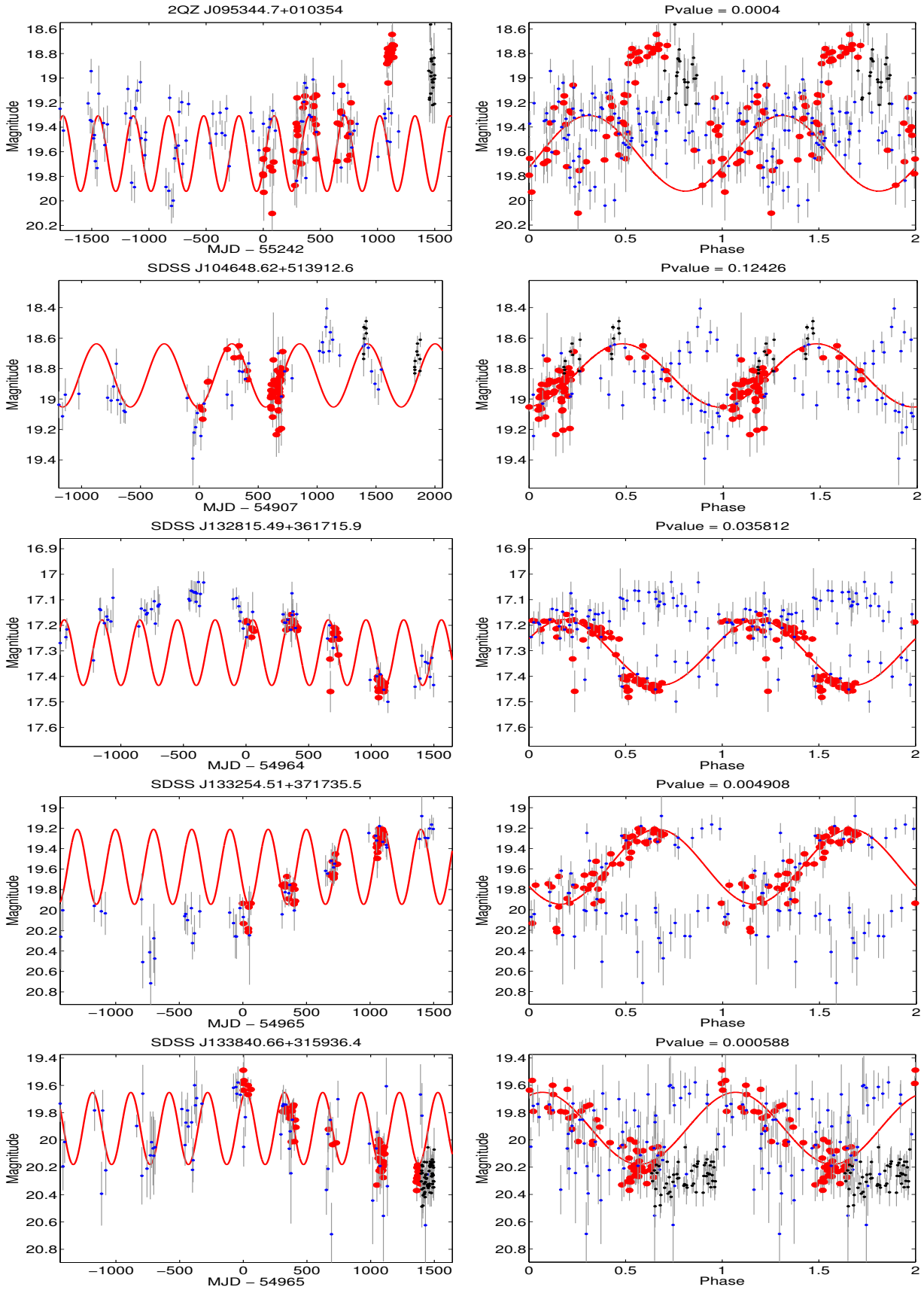
Table 2.2: SMBHB properties

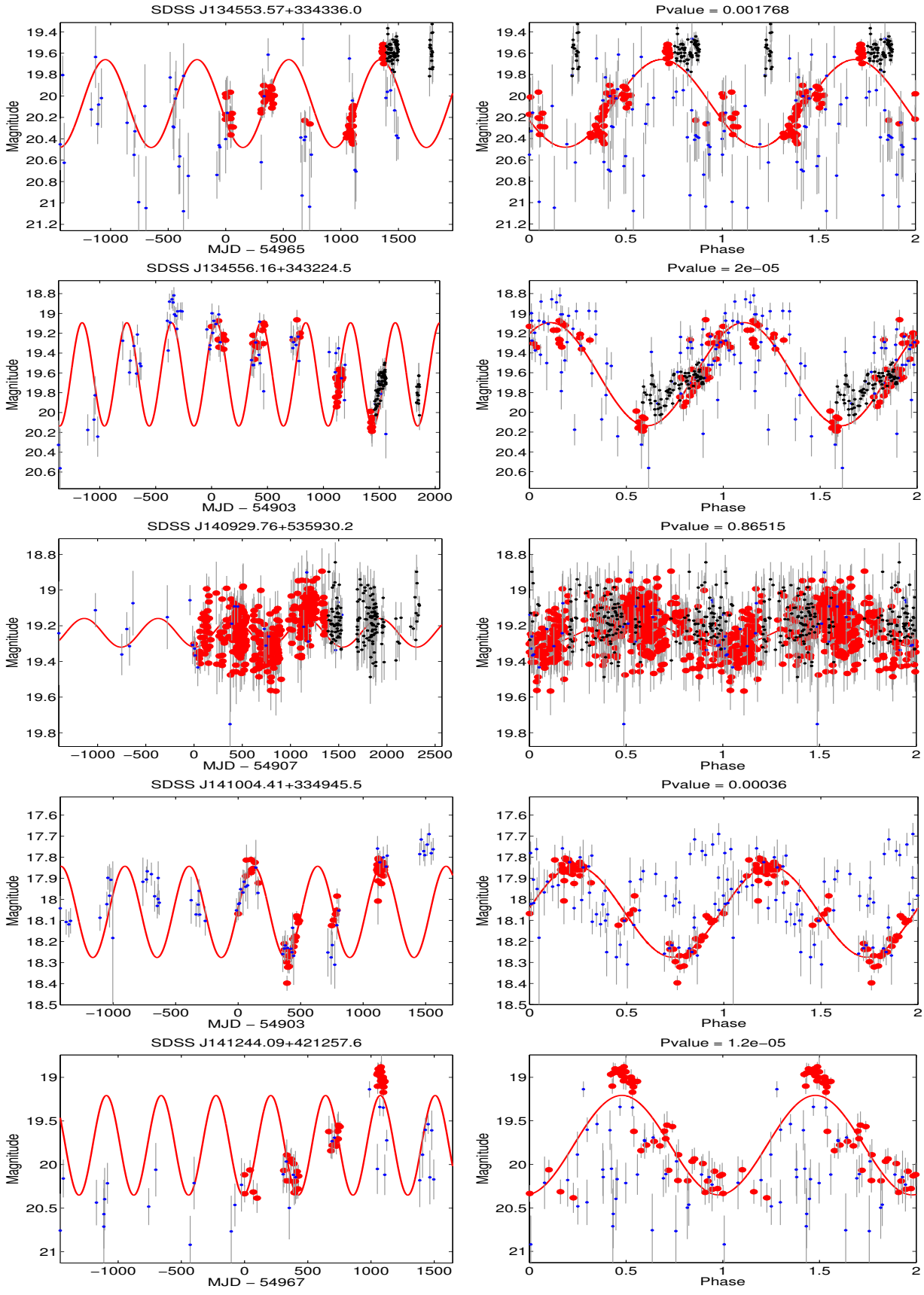
Name	P-value*	Observed Period [days]	Variability Fraction** [%]	Separation [mpc]	Angular Separation [μas]	Δt [ns]	t_{insp} ($q=1$) [years]
UM 269	5.4×10^{-4}	490.5	16.0	3.13	0.67	0.044	1.3×10^{4}
SDSS J005158.83-002054.1	1.5×10^{-3}	443.1	9.6	2.89	0.35	0.035	7.4×10^{2}
SDSS J005453.30-003258.3	1.2×10^{-3}	444.7	14.4	1.55	0.19	0.002	2.2×10^{4}
SDSS J023050.06+005843.1	1.5×10^{-2}	445.2	6.9	6.65	0.77	2.570	4.0×10^{0}
SDSS J024442.77-004223.2	8.0×10^{-6}	452.9	10.5	1.77	0.25	0.003	3.9×10^{4}
2QZ J095344.7+010354	4.0×10^{-4}	308.0	14.2	1.70	0.22	0.007	1.4×10^{3}
SDSS J104648.62+513912.6	2.8×10^{-1}	575.0	9.6	4.30	0.63	0.111	2.2×10^{3}
SDSS J132815.49+361715.9	3.5×10^{-2}	301.1	5.9	4.14	0.49	0.685	1.1×10^{1}
SDSS J133254.51+371735.5	4.9×10^{-3}	300.2	17.0	1.52	0.18	0.008	7.7×10^{1}
SDSS J133840.66+315936.4	5.9×10^{-4}	301.3	12.2	2.33	0.29	0.077	4.2×10^{0}
SDSS J134553.57+334336.0	1.8×10^{-3}	796.6	19.0	5.60	0.70	0.159	1.5×10^{3}
SDSS J134556.16+343224.5	2.0×10^{-5}	400.1	24.2	2.60	0.33	0.027	1.2×10^{3}
SDSS J140929.76+535930.2	5.6×10^{-1}	881.3	4.5	4.15	0.53	0.026	1.3×10^{4}
SDSS J141004.41+334945.5	3.6×10^{-4}	509.4	10.0	4.24	0.60	0.147	9.9×10^{2}
SDSS J141244.09+421257.6	1.2×10^{-5}	433.4	26.6	6.22	0.80	1.645	3.0×10^{1}
SDSS J142339.44+471240.8	$<4.0\times10^{-6}$	298.8	13.6	1.74	0.20	0.009	5.2×10^{2}
TEX 1428+370	7.3×10^{-1}	288.3	12.9	2.14	0.32	0.027	1.3×10^{3}
SDSS J143637.44+090155.5	2.6×10^{-2}	319.1	11.5	1.21	0.18	0.001	3.9×10^{4}
SDSS J145713.26+140334.1	2.0×10^{-1}	321.1	7.2	2.10	0.26	0.037	1.1×10^{1}
SDSS J145859.07+153144.7	1.0×10^{-2}	317.4	6.0	3.65	0.46	0.609	6.5×10^{-1}
SDSS J150900.70+175114.3	5.0×10^{-3}	317.8	14.3	2.26	0.30	0.026	9.1×10^{2}
FBQS J150911.2+215508	6.5×10^{-2}	314.4	3.7	2.41	0.41	0.040	1.9×10^{3}
SDSS J150912.07+204004.6	2.1×10^{-2}	315.3	9.6	1.55	0.31	0.005	2.8×10^{4}
SDSS J151053.24+240943.3	3.2×10^{-2}	712.2	16.5	3.04	0.39	0.010	2.1×10^{4}
SDSS J151243.67+195845.1	1.5×10^{-1}	308.6	5.3	2.57	0.33	0.055	3.2×10^{2}
SDSS J151646.10+221724.7	3.9×10^{-4}	309.7	14.7	2.40	0.35	0.038	9.6×10^{2}
SDSS J152739.97+413234.6	4.7×10^{-3}	438.3	7.2	4.10	0.50	0.206	1.3×10^{2}
SDSS J152903.11+223623.8	1.2×10^{-2}	310.4	14.4	2.21	0.35	0.026	2.1×10^{3}
SDSS J153051.79+503440.1	$<4.0\times10^{-6}$	429.2	28.9	3.04	0.38	0.048	6.8×10^{2}
SDSS J153251.06+335852.2	3.1×10^{-3}	436.2	7.8	2.61	0.30	0.030	1.4×10^{2}
PDS 898	3.0×10^{-4}	436.2	12.1	2.92	0.37	0.037	1.1×10^{3}
SDSS J155308.65+501436.5	1.2×10^{-2}	438.2	13.1	2.72	0.34	0.050	2.4×10^{1}
SDSS J160322.68+200535.2	8.0×10^{-1}	237.5	15.0	3.18	0.38	0.592	5.8×10^{-1}
SDSS J160454.57+315733.5	4.4×10^{-5}	307.3	16.5	2.03	0.26	0.040	6.7×10^{0}
SDSS J162634.15+325032.6	4.6×10^{-2}	297.5	13.5	1.89	0.24	0.013	1.0×10^{3}
SDSS J170942.58+342316.2	3.6×10^{-4}	455.2	7.2	3.42	0.39	0.096	6.7×10^{1}
SDSS J171122.67+342658.9	3.4×10^{-4}	285.5	13.6	1.60	0.19	0.010	7.9×10^{1}
SDSS J171617.49+341553.3	4.4×10^{-3}	130.7	5.5	1.58	0.19	0.070	7.4×10^{0}
SDSS J171909.93+344001.3	7.4×10^{-4}	292.6	5.6	2.50	0.29	0.090	9.3×10^{0}
SDSS J212939.60+004845.5	—-	313.0	9.8	2.43	0.29	0.066	1.3×10^{1}
SDSS J214036.77+005210.1	7.4×10^{-3}	315.8	7.2	2.47	0.32	0.042	3.1×10^{2}
SDSS J214225.29+001643.2	8.3×10^{-3}	316.7	13.3	1.74	0.20	0.008	7.0×10^{2}
SDSS J214357.98+003349.6	7.1×10^{-3}	456.0	10.1	3.30	0.39	0.101	2.4×10^{1}
PKS 2203-215	1.6×10^{-3}	497.0	17.7	4.08	0.60	0.133	1.3×10^{3}
SDSS J224008.39+263928.4	5.4×10^{-1}	314.1	3.8	2.19	0.27	0.047	8.9×10^{0}
SDSS J231733.66+001128.3	5.7×10^{-3}	467.3	10.8	3.51	0.45	0.076	7.3×10^{2}
SDSS J232135.73+173916.5	2.2×10^{-4}	337.4	22.1	3.40	0.43	0.171	1.2×10^{2}
2QZ J235800.2-281429	—-	306.1	13.5	2.20	0.25	0.033	7.6×10^{1}
SDSS J235928.99+170426.9	8.8×10^{-5}	330.3	15.4	3.15	0.43	0.122	2.4×10^{2}
SDSS J235958.72+003345.3	—-	486.3	10.3	3.21	0.37	0.056	1.5×10^{2}

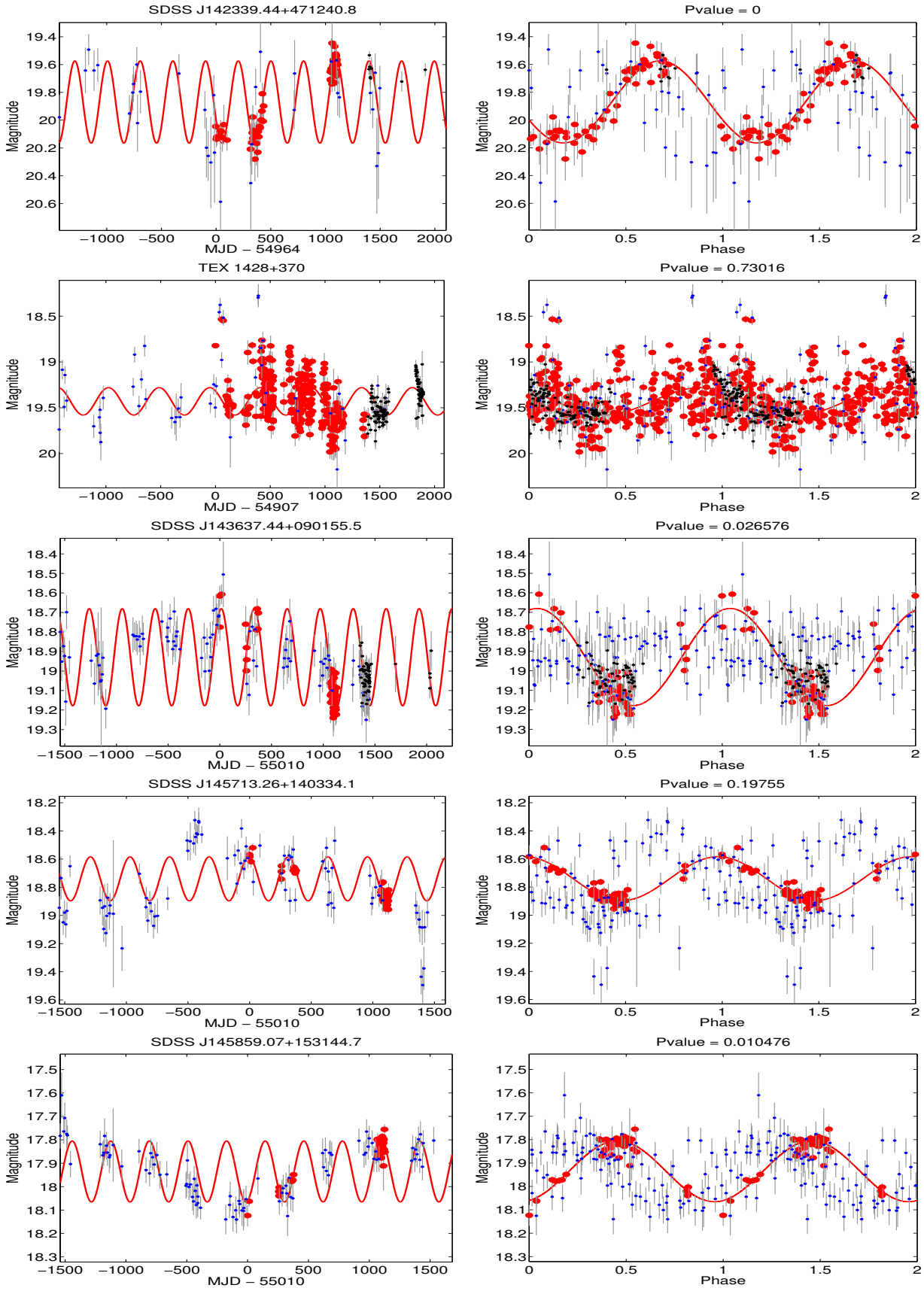
* The P-values shown here are calculated from the composite light curves (with data from PTF, iPTF and CRTS). We note that all the quasars shown in the tables have P-value<1/250,000, when only the PTF data are taken into account.

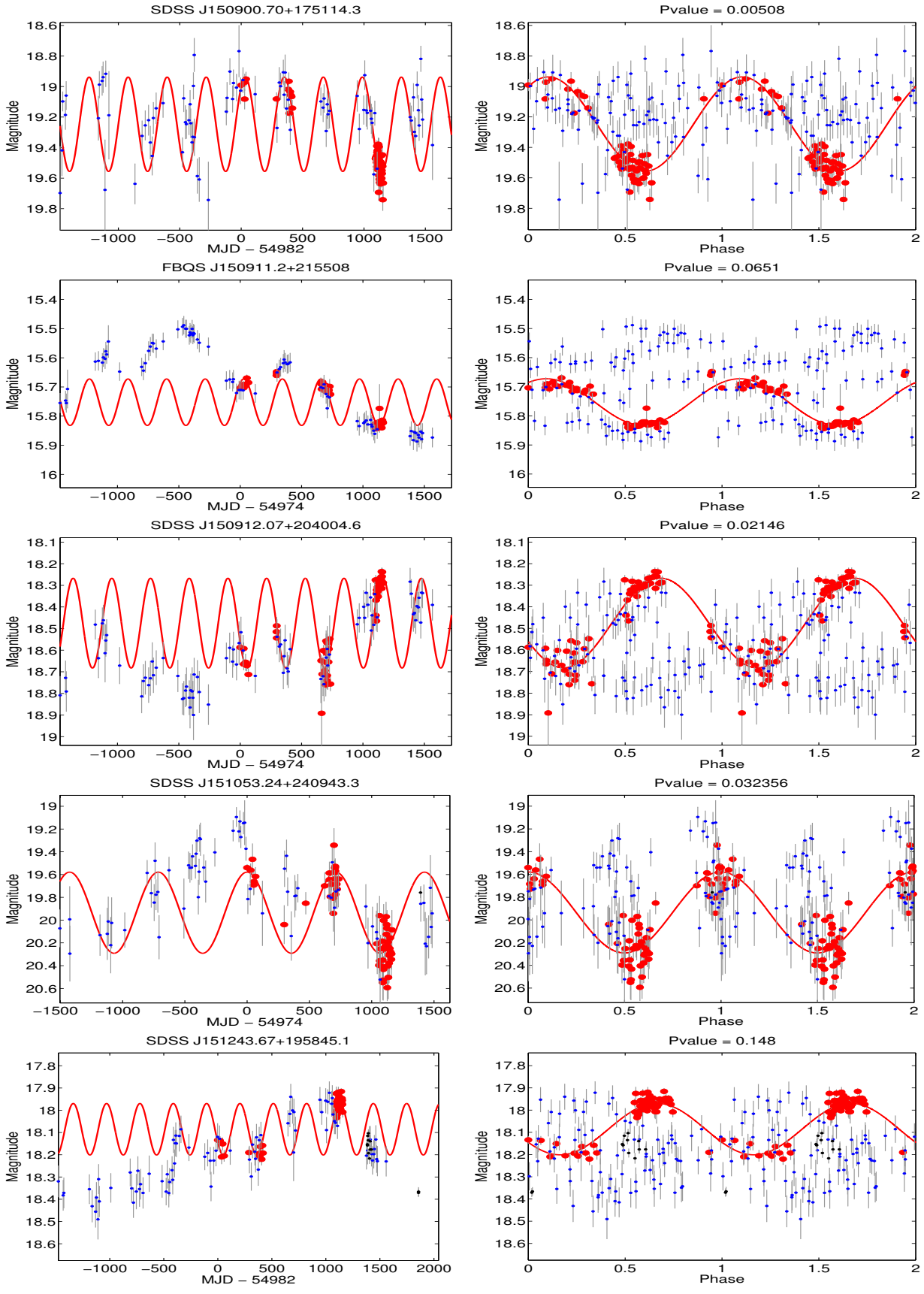
**We calculated the variability fraction as $\frac{F_{max}-F_{min}}{2F_{mean}} \times 100\% = \frac{10^{-\frac{m-A/2}{2.5}} - 10^{-\frac{m+A/2}{2.5}}}{2\times10^{-\frac{m}{2.5}}} \times 100\%$, where m is the mean magnitude of the quasar and A the amplitude of the best fit sinusoid.

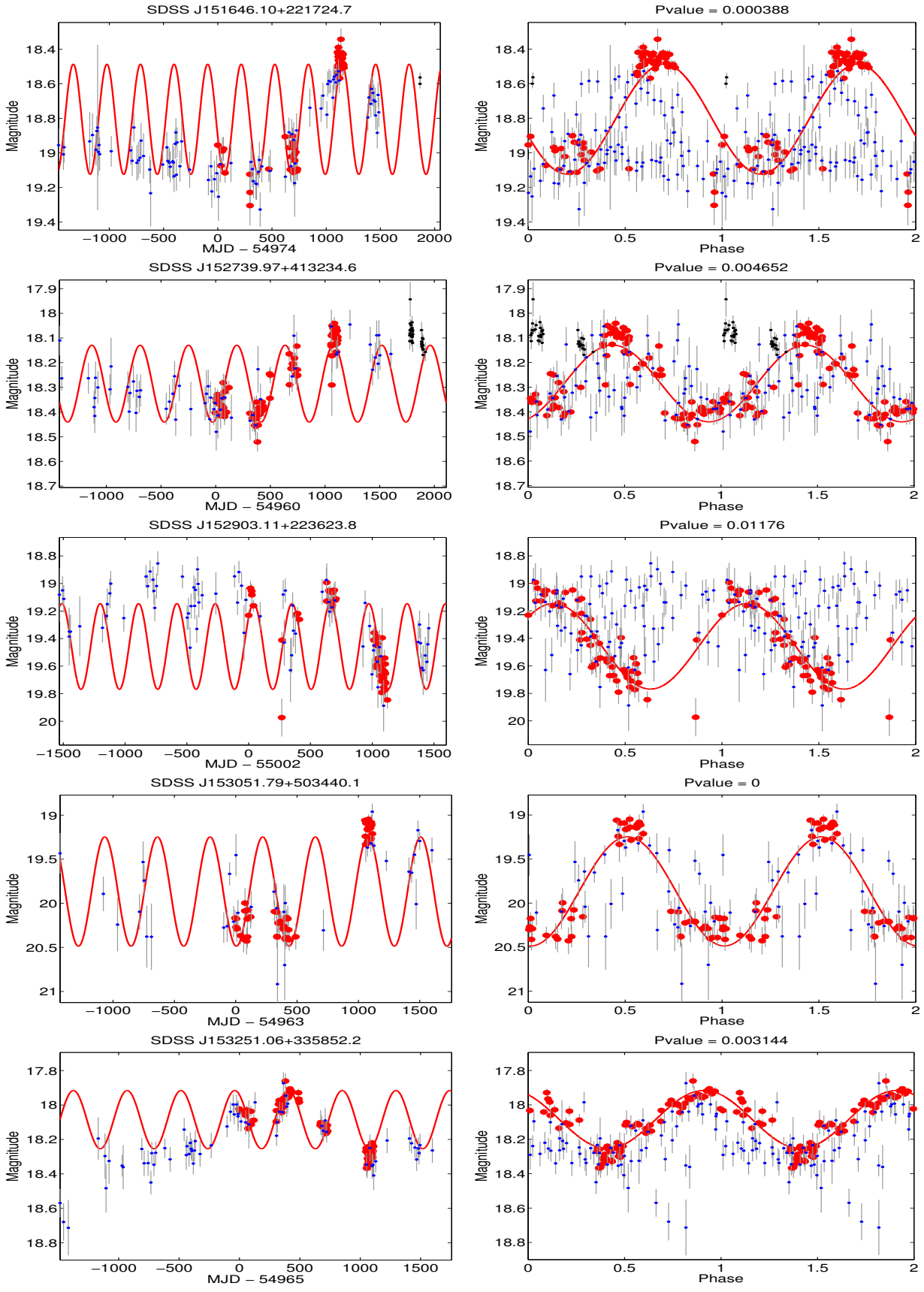

UM 269
Pvalue = 0.00054
Magnitude
MJD − 55053
Phase
SDSS J005158.83−002054.1
Pvalue = 0.001552
SDSS J005453.30−003258.3
MJD − 55052
Pvalue = 0.001152
SDSS J023050.06+005843.1
MJD − 55060
Pvalue = 0.01508
SDSS J024442.77−004223.2
MJD − 55050
Pvalue = 8e−06

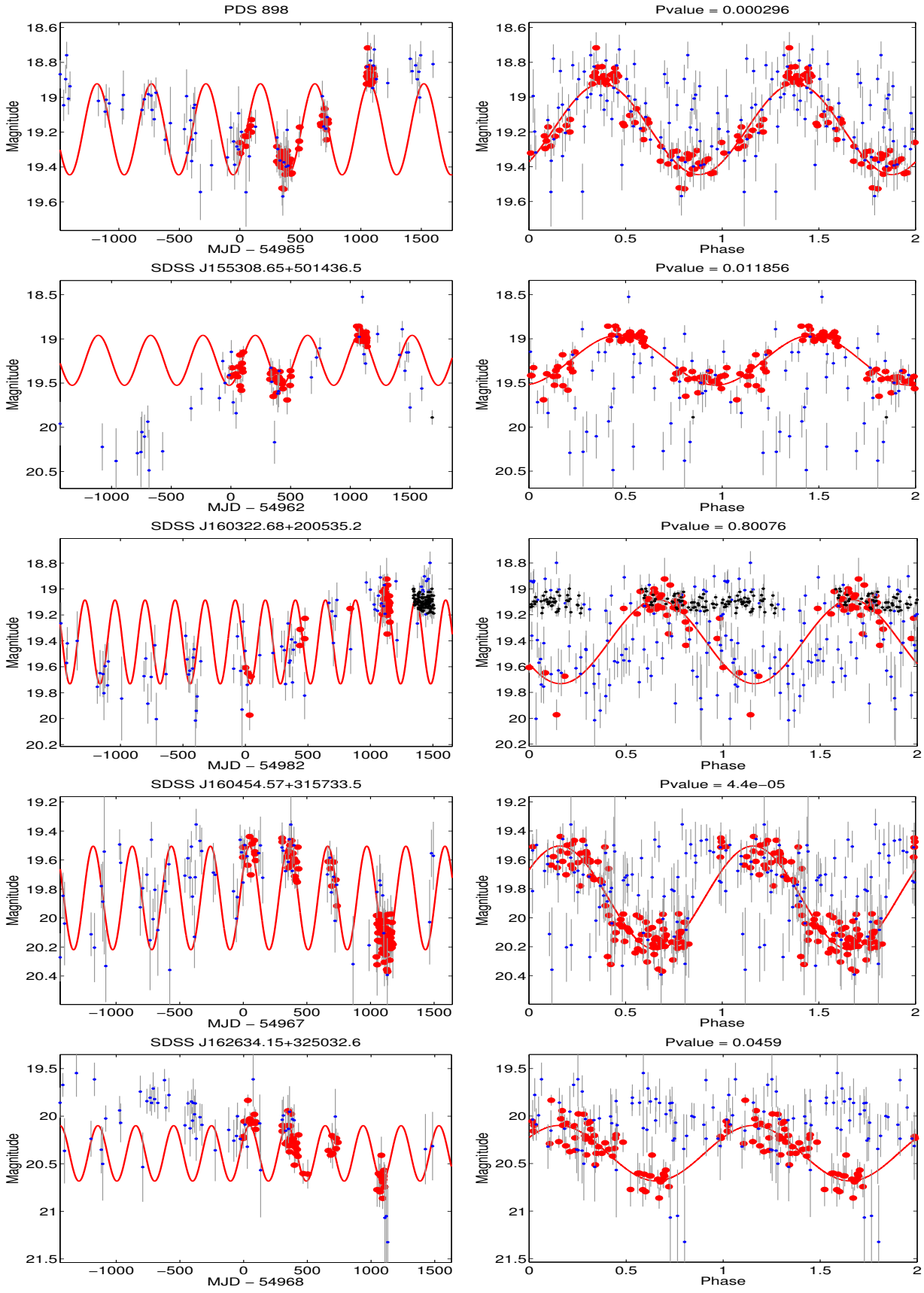

2QZ J095344.7+010354
Pvalue = 0.0004
SDSS J104648.62+513912.6
Pvalue = 0.12426
SDSS J132815.49+361715.9
Pvalue = 0.035812
SDSS J133254.51+371735.5
Pvalue = 0.004908
SDSS J133840.66+315936.4
Pvalue = 0.000588
Magnitude
MJD − 55242
MJD − 54907
MJD − 54964
MJD − 54965
MJD − 54965
Phase

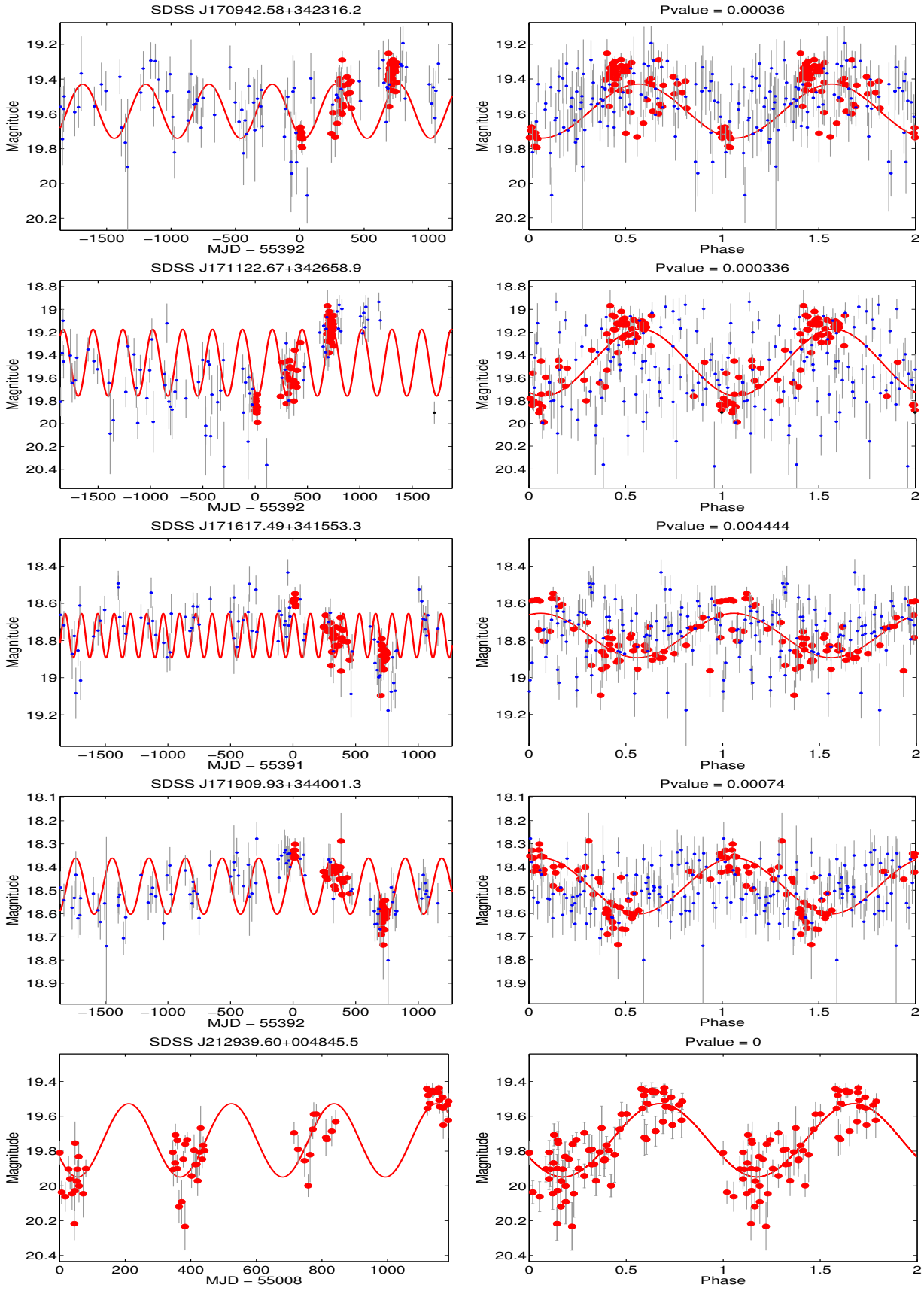

SDSS J134553.57+334336.0
Pvalue = 0.001768
SDSS J134556.16+343224.5
Pvalue = 2e−05
SDSS J140929.76+535930.2
Pvalue = 0.86515
SDSS J141004.41+334945.5
Pvalue = 0.00036
SDSS J141244.09+421257.6
Pvalue = 1.2e−05
Magnitude
MJD − 54965
MJD − 54903
MJD − 54907
MJD − 54967
Phase

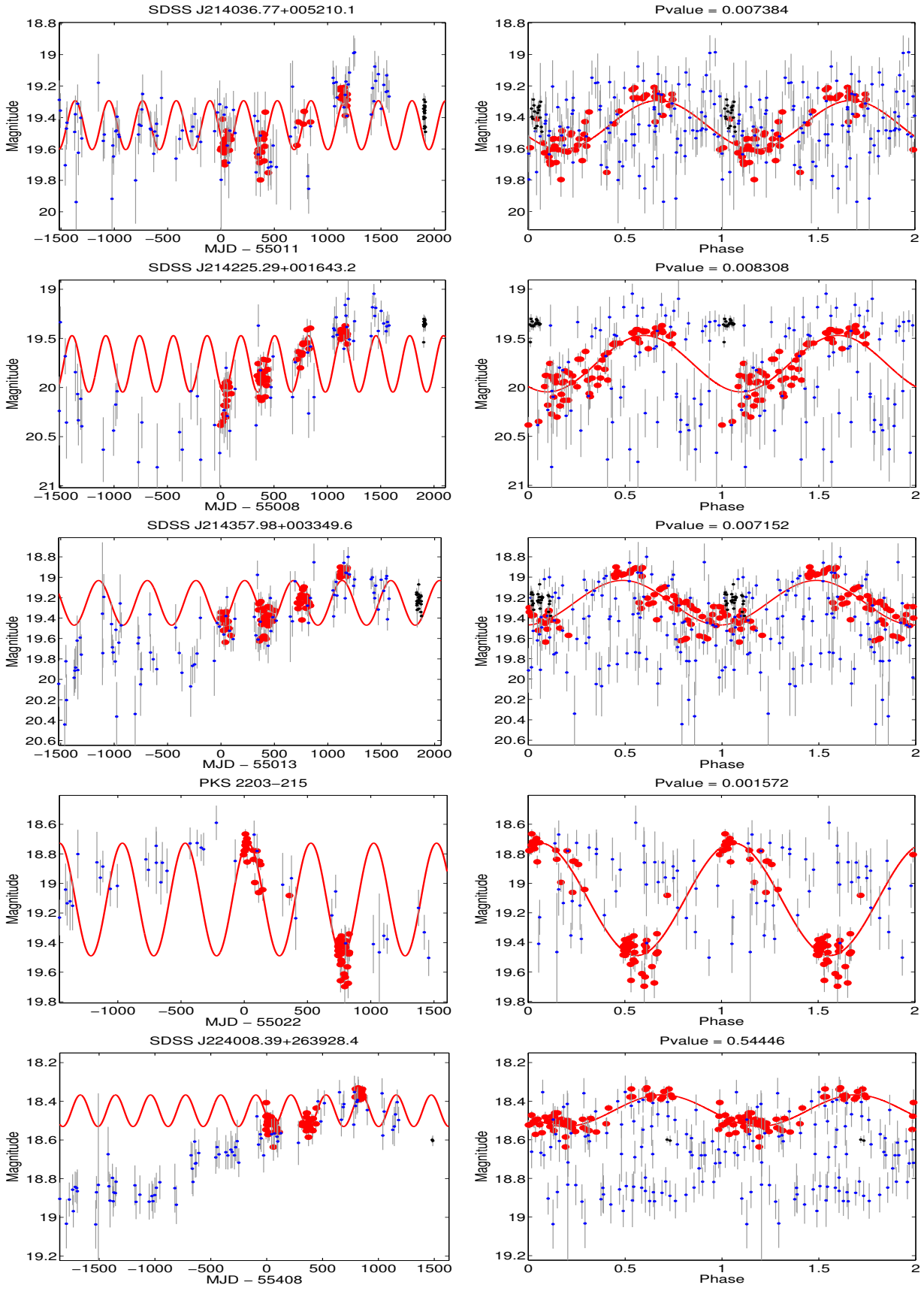

SDSS J142339.44+471240.8
Pvalue = 0
TEX 1428+370
Pvalue = 0.73016
SDSS J143637.44+090155.5
Pvalue = 0.026576
SDSS J145713.26+140334.1
Pvalue = 0.19755
SDSS J145859.07+153144.7
Pvalue = 0.010476
Magnitude
MJD − 54964
MJD − 54907
MJD − 55010
Phase

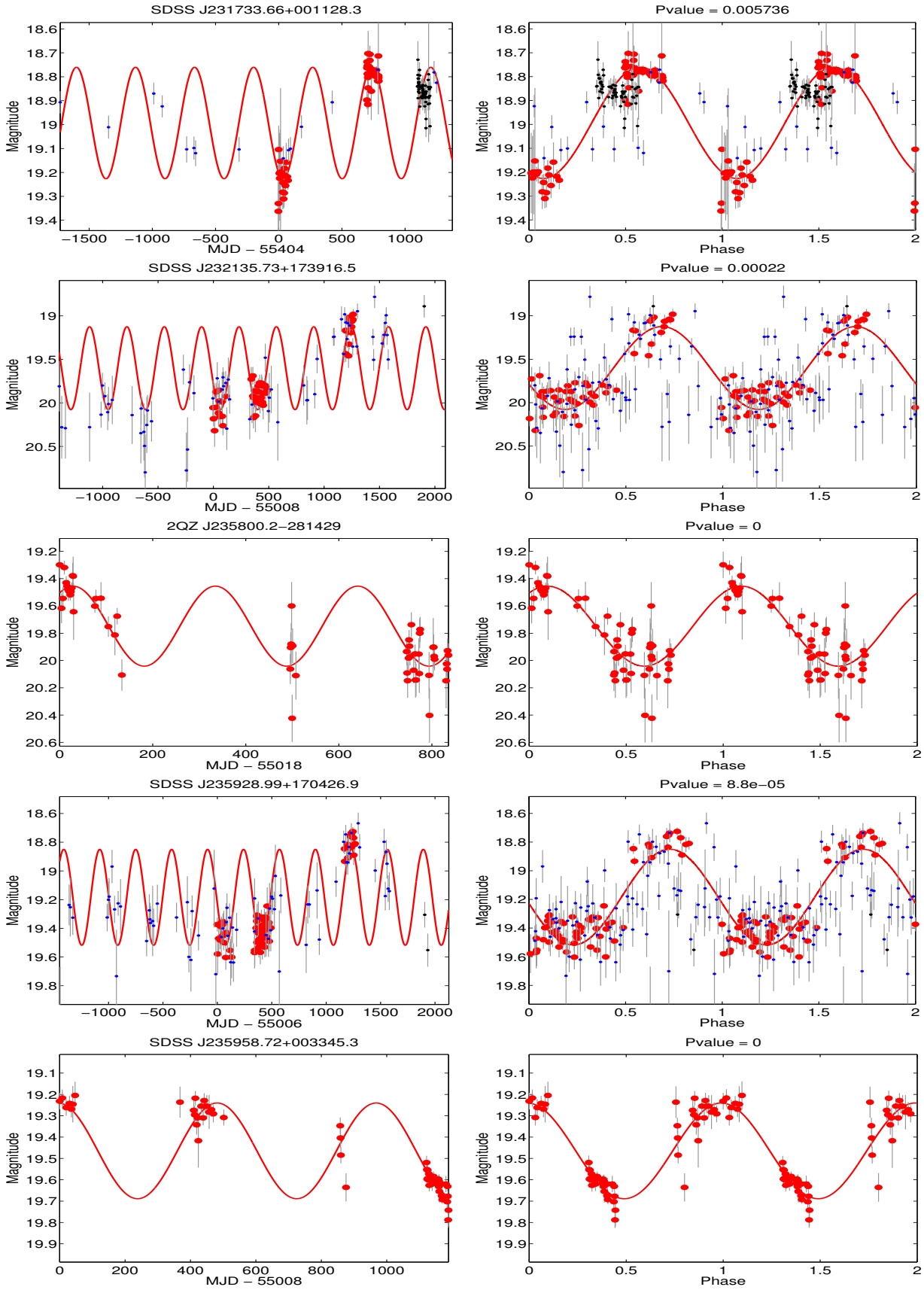

SDSS J150900.70+175114.3
Pvalue = 0.00508
FBQS J150911.2+215508
Pvalue = 0.0651
SDSS J150912.07+204004.6
Pvalue = 0.02146
SDSS J151053.24+240943.3
Pvalue = 0.032356
SDSS J151243.67+195845.1
Pvalue = 0.148
Magnitude
MJD − 54982
MJD − 54974
Phase


SDSS J151646.10+221724.7
Pvalue = 0.000388
SDSS J152739.97+413234.6
Pvalue = 0.004652
SDSS J152903.11+223623.8
Pvalue = 0.01176
SDSS J153051.79+503440.1
Pvalue = 0
SDSS J153251.06+335852.2
Pvalue = 0.003144
Magnitude
MJD − 54974
MJD − 54960
MJD − 55002
MJD − 54963
MJD − 54965
Phase


PDS 898
Pvalue = 0.000296
SDSS J155308.65+501436.5
Pvalue = 0.011856
SDSS J160322.68+200535.2
Pvalue = 0.80076
SDSS J160454.57+315733.5
Pvalue = 4.4e−05
SDSS J162634.15+325032.6
Pvalue = 0.0459
Magnitude
MJD − 54965
MJD − 54962
MJD − 54982
MJD − 54967
MJD − 54968
Phase


SDSS J170942.58+342316.2
Pvalue = 0.00036
SDSS J171122.67+342658.9
Pvalue = 0.000336
SDSS J171617.49+341553.3
Pvalue = 0.004444
SDSS J171909.93+344001.3
Pvalue = 0.00074
SDSS J212939.60+004845.5
Pvalue = 0
Magnitude
MJD − 55392
MJD − 55391
MJD − 55008
Phase


SDSS J214036.77+005210.1
Pvalue = 0.007384
SDSS J214225.29+001643.2
Pvalue = 0.008308
SDSS J214357.98+003349.6
Pvalue = 0.007152
PKS 2203−215
Pvalue = 0.001572
SDSS J224008.39+263928.4
Pvalue = 0.54446
Magnitude
MJD − 55011
MJD − 55008
MJD − 55013
MJD − 55022
MJD − 55408
Phase


SDSS J231733.66+001128.3
Pvalue = 0.005736
SDSS J232135.73+173916.5
Pvalue = 0.00022
2QZ J235800.2−281429
Pvalue = 0
SDSS J235928.99+170426.9
Pvalue = 8.8e−05
SDSS J235958.72+003345.3
Pvalue = 0
Magnitude
MJD − 55404
MJD − 55008
MJD − 55018
MJD − 55006
Phase

Chapter 3

Multiple periods in the variability of the supermassive black hole binary candidate quasar PG1302-102?

3.1 Introduction

It is well established that massive galaxies harbour supermassive black holes in their centres, with masses tightly correlated with the global properties of the host galaxy [Kormendy and Ho, 2013]. According to cosmological models of hierarchical structure formation, galaxies grow through mergers ([Springel *et al.*, 2005]; [Robertson *et al.*, 2006]). The formation of SMBHBs is a natural consequence of the above process. Combining the high expected rates of galaxy mergers with the expectation that SMBHBs spend a large fraction of their lifetimes at close separation [Haehnelt and Kauffmann, 2002], SMBHBs at sub-parsec separations should be fairly common, despite the scarcity of observational evidence.

Recently, [Graham *et al.*, 2015b][hereafter G15] reported the detection of strong periodic variability in the optical flux of quasar PG1302-102. PG1302 is a bright (median V-band magnitude $\sim$15), radio-loud quasar at redshift $z = 0.2784$, with inferred black hole (BH)

mass of $10^{8.3-9.4} M_{\odot}$. The light curve in optical bands shows a quasi-sinusoidal modulation, with a best-fit period of (5.2±0.2) yr and amplitude of ≈0.14 mag. G15 suggest that PG1302 may be a SMBHB at close separation (∼0.01 pc), interpreting the observed periodicity as the (redshifted) orbital period of the binary.

Theoretical work on circumbinary disks predicts that SMBHBs can excite periodic enhancements of the mass accretion rate that could translate into periodic luminosity enhancements, not only at the orbital period, but also on longer and shorter timescales. More specifically, if the binary is embedded in a thin disk, the gas will be expelled from the central region due to torques exerted by the binary, creating a cavity [Artymowicz and Lubow, 1994]. Several hydrodynamical simulations [Hayasaki *et al.*, 2007; MacFadyen and Milosavljević, 2008; Noble *et al.*, 2012; Shi *et al.*, 2012; Roedig *et al.*, 2012; D'Orazio *et al.*, 2013; Farris *et al.*, 2014; Gold *et al.*, 2014] indicate that the interaction of the individual BHs with the inner edge of the accretion disk can pull gaseous streams into the cavity, resulting in periodic modulation of the accretion rate at timescales corresponding to ≈1/2 and 1 times the binary's orbital period. For high BH mass ratios ($q \equiv M_1/M_2 0.3$), the cavity is lopsided, leading to the formation of a hotspot in the accretion disk. The strongest modulation in the accretion rate is observed at the orbital period of the overdense region, a few (∼3-8) times the orbital period of the binary.

These results imply that the observed 5.2 yr period in PG1302 may not reflect the orbital period of the binary. If the orbital period is shorter/longer than the period of optical valiability t_{opt}, there are major implications for the expected quasar-binary fraction, and the probability of detecting SMBHBs [Haiman *et al.*, 2009], as well as the physics of the orbital decay. [D'Orazio *et al.*, 2015a] showed that, under the assumption that the optical period corresponds to the longer period of the hotspot in the accretion disk, PG1302 would be in the gravitational inspiral regime. This would confirm that SMBHBs can produce bright electromagnetic emission even at these late stages of the merger [Noble *et al.*, 2012; Farris *et al.*, 2015b].

The significance of the above considerations motivates us to search for multiple periodicity in the optical variability of PG1302. In this paper, we search for secondary peaks in the Lomb-Scargle periodogram. We assess the statistical significance of the identified peaks

by generating mock time series that show correlated noise, as expected for quasars [Kelly *et al.*, 2009; Kozłowski *et al.*, 2010]. The false alarm probability of the most significant secondary peak is 3%, below a reliable detection threshold. We set upper limits on the amplitude of secondary sinusoid variations that could be detectable with high significance, in the presence of correlated noise, as well as the 5.2 yr periodic signal. Although the current data do not allow the detection of weak secondary periodic terms, the addition of three years of observations, can improve the sensitivity up to a factor of 2.

3.2 Data Analysis

We investigate the possibility of multiple periodic terms in the photometric variability of PG1302 by analysing the light curve published in G15[1]. The light curve consists of data from: (1) the Catalina Real-time Transient Survey (CRTS; [Drake *et al.*, 2009; Mahabal *et al.*, 2011; Djorgovski *et al.*, 2012]); (2) the Lincoln Near-Earth Asteroid Research (LINEAR; [Sesar *et al.*, 2011]) program; (3) the All Sky Automated Surveys (ASAS; [Pojmanski, 1997]) and; (4) other archival data ([Garcia *et al.*, 1999]; [Eggers *et al.*, 2000]), providing a baseline of ∼20 yr. The different datasets were calibrated to account for the differences between the various photometric systems, as detailed in G15.

For our analysis, we use the generalised Lomb-Scargle (LS) periodogram [Lomb, 1976; Scargle, 1982; Zechmeister and Kürster, 2009], a standard method for detecting periodicity in unevenly sampled data[2]. We compute the periodogram for 1000 trial frequencies on a uniform logarithmic grid, spanning from $f_{min} = 1/T_{data}$ to $f_{max} = 1/T_{min}$, where T_{data} is the baseline of the light curve and $T_{min} = 100\,d$.

For each peak in the periodogram, we then calculate the probability that a peak of equal power can arise by chance (false alarm probability; hereafter FAP). For this purpose, we generate time series that mimic the optical variability of PG1302. Quasars show correlated stochastic variability, best described as a damped random walk [Kelly *et al.*, 2009; Kozłowski

[1]We are grateful to M. Graham for providing us with their calibrated photometric data in electronic form.

[2]Throughout the analysis, we make use of the astroML python package [Ivezić *et al.*, 2014].

et al., 2010]. At high frequencies, the power spectral density decreases with frequency ($PSD \sim 1/f^2$, or red noise), whereas, at low frequencies, it becomes flat ($PSD \sim f^0$, or white noise).

We generate evenly sampled stochastic time series with a dense temporal resolution ($dt = 2\,h$) that exhibit red noise variability [Timmer and Koenig, 1995]. To account for the flattening of the power spectrum at low frequencies, we consider two additional varibility models; pink noise with PSD $\sim$ 1/f and white noise with flat power spectrum.

For each model, we simulate time series with baseline significantly longer than the observed ($500T_{data}$) to include the effects of long-term variability (i.e. *red noise leak*; [Uttley *et al.*, 2002]). We divide the time series into 500 segments spanning T_{data} and subsample to match the observation times in the light curve of PG1302. We incorporate the measurement uncertainty by adding standard normal deviates of the photometric errors. Next, we multiply the generated time series with a scaling factor to ensure that its variance equals the variance of the observed data.

We generate 20,000 (500×40) stochastic time series and calculate the periodogram for each realisation. At each trial frequency, we calculate the FAP by comparing the power in the periodogram of PG1302 to the distribution of power *at the same frequency* in the periodograms of the mock time series. We define FAP_f, the false alarm probability *per frequency*, as the fraction of the 20,000 simulated periodograms with power exceeding the value observed for PG1302.

Since the frequency bins have non-trivial correlations, we compute the effective number of independent bins, N_{eff}, for each FAP_f, by: (i) applying the same analysis as above, but replacing the PG1302 data with each individual simulated mock periodogram, and (ii) determining the fraction of the 20,000 mock periodograms that show FAP_f (or lower) at *any* frequency. For $\mathrm{FAP}_f = 0.5 \times 10^{-4}$, we find $N_{\mathrm{eff}} = 150$, and the total FAP$\equiv N_{\mathrm{eff}} \times \mathrm{FAP}_f$=0.0075[3].

We proceed to search for putative secondary periodic components by repeating the analysis above. In this case, however, we assume that the variability of PG1302 consists

[3] N_{eff} is decreased to $\approx$80 for $\mathrm{FAP}_f=10^{-3}$ (and $\approx$40 for $\mathrm{FAP}_f=10^{-2}$), but is relatively insensitive to the noise model.

of stochastic variability with an additional sinusoid component with period of $1,884 \pm 88$ d and amplitude of 0.14 mag, as reported in G15. The generated time series are again scaled to match the variance of PG1302.

3.3 Results

3.3.1 Main Periodic Component

First, we examine the statistical significance of the reported periodic component in PG1302 using an independent statistical analysis[4] and considering red, pink and white noise, along with a damped random walk model (DRW with PSD $\sim \hat{\sigma}^2\tau^2/(1+4\pi^2 f^2\tau^2)$). For the latter, we maximize the Gaussian likelihood of the data (marginalized over the unkown mean flux, e.g., eq. (A8) in [Kozłowski *et al.*, 2010]) to identify the best fit parameters. With the best fit $\tau \sim$100 d and $\hat{\sigma} \sim$0.3 mag/yr$^{1/2}$, we generate time series [Timmer and Koenig, 1995] and perform the periodogram analysis.

The peak at 5.2 yr is identified with a FAP<0.75% for every noise model, except for red noise (FAP=4.5%)[5]. We stress that the FAP refers to a single source. However, PG1302 was selected from a large sample of 247,000 quasars, which needs to be considered in order to assign the FAP for the entire survey ("look elsewhere effect"). In practice, this requires to repeat the periodogram analysis 3.7×10^9 times (247,000×150×100; to account for the sample size, the frequency bins, and for detection at 1% level). This is beyond the scope of this paper, but since this effect was not included in G15, the overall significance of the 5.2 yr period in PG1302 remains uncertain and further analysis is required.

3.3.2 Secondary Periodic Components

If the optical variability of PG1302 is the superposition of multiple periodic terms, it is expected that significant secondary peaks will appear in the periodogram. Moreover, following the predictions of hydrodynamical simulations, we expect the peaks roughly at ratios 1:2

[4]The periodicity in G15 was identified using wavelet and autocorrelation function based techniques.

[5]We note that, since $\tau \sim$100 d for the DRW model, the peak at 5.2 yr corresponds to the flat part of the power spectrum and thus the red noise model is too conservative.

(1 and 1/2 times the orbital period) or from 1:3:6 to 1:8:16 (hotspot period, orbital period and half of the orbital period of the binary).

More specifically, with the limitations imposed by the currently available data, we were able to explore three possibilities: (A) If the optical period t_{opt} is the binary orbital period $t_{\rm bin}$, is there a secondary peak near $0.5t_{\rm bin}$? (B) If t_{opt} is $0.5t_{\rm bin}$, is there a periodic component at $t_{\rm bin}$? (C) If t_{opt} is the long orbital period $\approx (3-8)t_{\rm bin}$ of the hotspot in the disk, as hypothesised by [D'Orazio *et al.*, 2015a], is there a secondary period corresponding to $\approx 0.5t_{\rm bin}$ or $t_{\rm bin}$? Note that in scenario (A) and (B), periodicity is predicted at $\sim (3-8)t_{\rm bin}$. However, the 20 yr baseline of the light curve does not allow us to probe this low-frequency range.

The detection of weaker periodic terms in the presence of a significant periodic component is in general challenging. The periodogram of an unevenly sampled time series is the convolution of the actual power spectrum of the signal and the window function, defined by the irregular time sampling. A periodic signal can introduce aliasing peaks, and hide the secondary peaks. To illustrate this, in Fig. 3.1, we show the LS periodogram of PG1302, superimposed with the periodogram of a pure sine wave, with period of 5.2 yr, sampled at the observation times of PG1302. The periodogram of the sine wave shows a set of aliasing peaks, which coincide with many of the observed features in the periodogram.

We analysed the LS periodogram of PG1302 searching for secondary peaks, focusing on the frequencies, within a factor of ∼10 around the peak of the main period. We assessed the statistical significance of the observed secondary peaks using the method discussed in § 3.2 above. In Fig. 3.2, we present the periodogram of PG1302 with the average and the maximum periodogram power from 20,000 realisations of the simulated variability (red noise plus a sine wave), with the latter corresponding to our FAP $< 1\%$ significance threshold. In the bottom panel, we show the FAP at each frequency. A peak at ∼300 d (∼40 nHz) is identified with the lowest FAP ($< 1\%$). This peak is most likely artificial, since it coincides with one of the strongest aliasing peaks, as demonstrated in Fig. 3.1. The most significant peak that does not overlap with aliasing peaks appears at ∼155 d (∼77.5 nHz), with a FAP∼3%.

Our conclusions are similar for the different noise models. More specifically, we identify

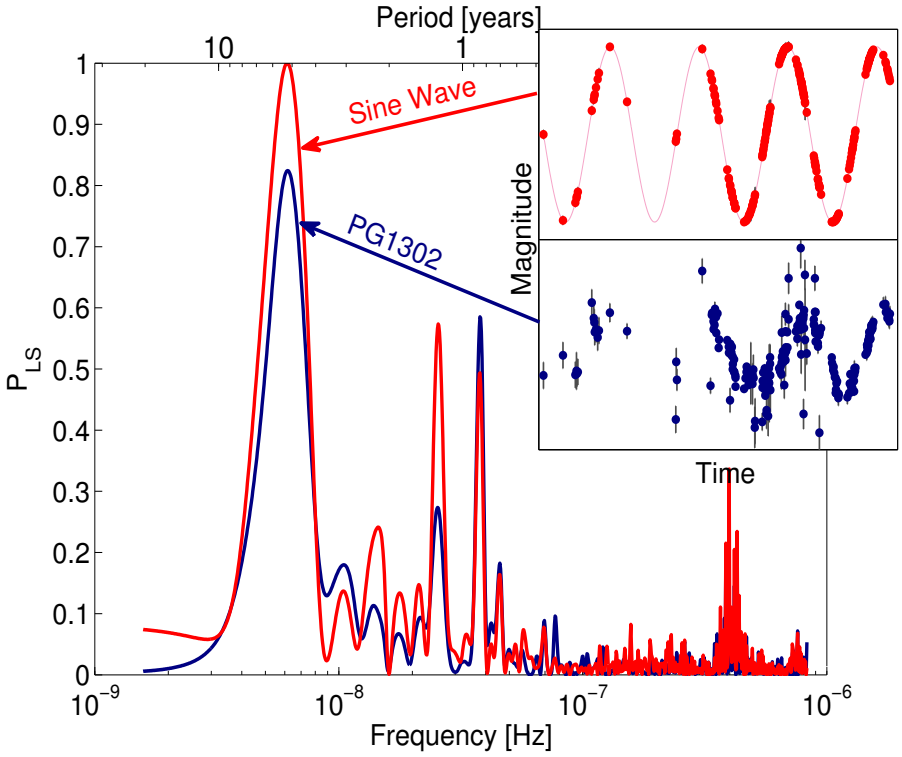


Figure 3.1: LS periodogram of PG1302 (blue) and of a sine wave sampled at the observation times of PG1302 (red). The embedded plots illustrate the respective light curves, spanning a baseline of ∼20 yr. The irregular sampling of the data produces the artificial secondary peaks seen in the red curve.

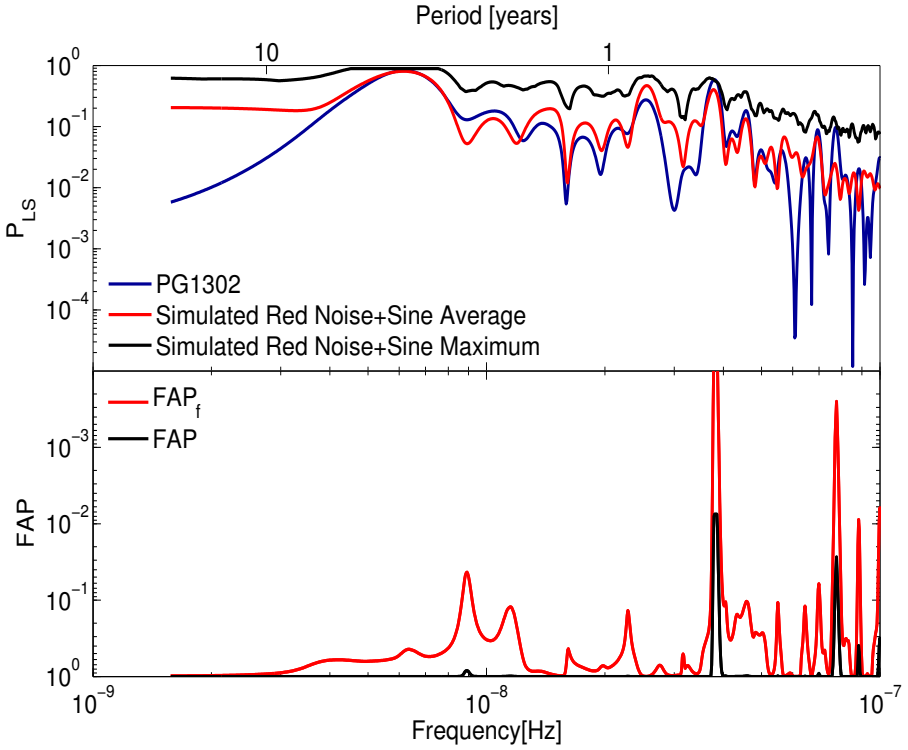


Figure 3.2: *Top panel*: LS periodogram of PG1302 (blue), superimposed with the average of 20,000 realisations of a model that includes stochastic red noise variability and a 5.2 yr-sinusoid. The maximum power among these 20,000 realisations is also shown (black), corresponding to 1% false alarm probability threshold. *Bottom panel*: FAP_f (red) and total FAP (black) of the peaks in the periodogram of PG1302's, as a function of frequency. Note that the y-axis in the bottom panel decreases upwards.

the same peak at ∼300 d (which we discard as aliasing) with FAP $< 1\%$ both for white and pink noise; we also find the peak at ∼155 d with FAP=0.25 and 0.1 for pink and white noise, respectively.

3.3.3 UpperLimits

We next derive upper limits for putative secondary periodic terms with the current data. For this purpose, we simulate time series as before, but with an additional sinusoid. We identify the minimum amplitude of the second component at which the power of the new

peak in the LS periodogram exceeded the FAP $< 1\%$ detection threshold at least 90% of the time. We repeat this process for different frequencies within the interval of interest. In Fig. 3.3, we present the minimum relative amplitude (corresponding also to V-band magnitude) of a secondary sinusoid that would be detectable at different frequencies. The shaded areas indicate frequencies relevant to the three possibilities (A)-(C) discussed above.

As Fig. 3.3 shows, we can set weak limits for scenario (B); a second periodic component would need a higher amplitude than the one identified in G15 to be detectable. The reduced sensitivity in this frequency range is reasonable, since the light curve is well sampled only in the last $\sim$9 yr, hindering the detection of a $\sim$10 yr period. Moreover, at this frequency range the effect of red noise is significant. For scenario (A) and for the majority of frequencies relevant to scenario (C), the secondary term needs to have amplitude comparable to the main 5.2 yr modulation to be detectable. Finally, for a few specific frequencies relevant to scenario (C), we can probe secondary sinusoids with amplitudes down to ~ 0.06 mag. For the peak with the lowest FAP (155 d), the minimum detectable amplitude would be 0.09 mag.

The upper limits for the different noise models are also shown in Fig. 3.3. For white noise, the minimum detectable amplitude is almost constant ($\sim$0.08 mag) for all frequencies, with fluctuations from the irregular sampling similar to the case of red noise. The results are similar for pink noise, with the sensitivity falling in-between red and white noise.

3.4 Discussion

3.4.1 Historic Data

Historic photometry of PG1302 is available from the digitisation of old photographic plates, as part of the Digital Access to a Sky Century @ Harvard (DASCH; [Grindlay *et al.*, 2012]) project. The measurements were in unfiltered B photographic magnitude, which is close to wide-band Johnson B magnitude and were calibrated with the AAVSO APASS catalog [Henden *et al.*, 2012]. We extracted the B-band light curve from the online database (`http://dasch.rc.fas.harvard.edu/lightcurve.php`) and excluded measurements with high astrometric errors or with magnitudes within 0.5 of the limiting magnitude, and data points

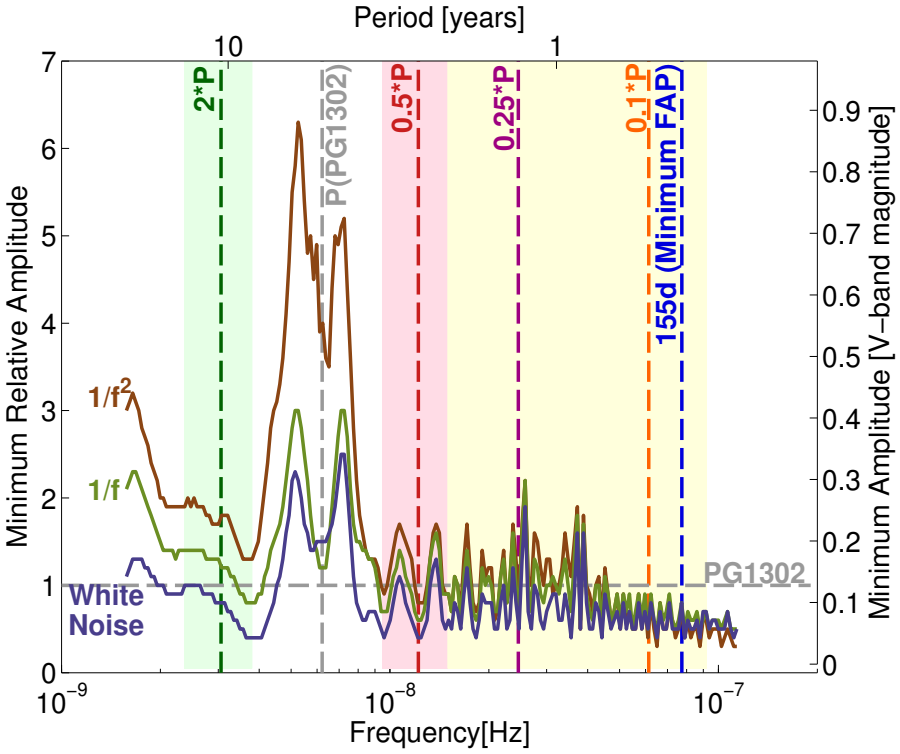


Figure 3.3: The minimum amplitude of a secondary sinusoid term that would be detectable at each frequency. The y axis shows the amplitude relative to that of the main 5.2 yr peak (left), and also in V mag (right). The three curves correspond to noise models with three different power–law power spectra: $\propto 1/f^2$ (brown), $\propto 1/f$ (olive), and constant white noise (dark blue). Frequency ranges relevant to different scenarios are highlighted with light pink (A), light green (B) and light yellow (C). Vertical lines mark specific frequencies of interest (see text).

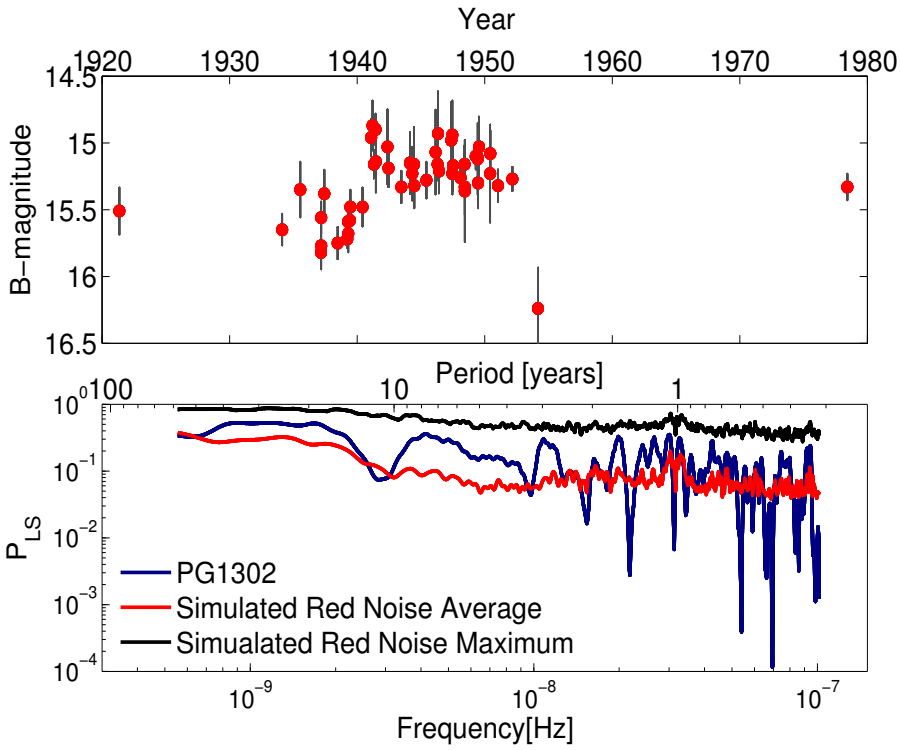


Figure 3.4: *Top Panel*: B-band light curve of PG1302 from the DASCH project. *Bottom Panel*: LS periodogram of the B-band light curve (blue), superimposed with the average red noise model (red) and the maximum power of the periodogram from 20,000 realisations of the noise (black), which corresponds to our significance threshold.

that were flagged as plate defects. We analysed the periodogram, as before, assuming red noise variability (without the periodic component) and scaling with the observed variance.

In Fig. 3.4, we present the B-band light curve and the LS periodogram, along with the average and maximum power of the LS periodogram from 20,000 realisations of red noise. All of the identified peaks (including the 5.2 yr period from G15) are below the significance threshold. This result, however, must be viewed in light of the photometric accuracy of the DASCH dataset. The average photometric error is 0.18 mag, exceeding the 0.14 mag amplitude of the periodic variability in G15. Interestingly, the two most significant peaks, within the interval of interest, are observed at similar frequencies ($\sim$42 nH and $\sim$76 nH) as the peaks in the periodogram from the G15 data, although their significance remains low.

3.4.2 Future Data

We investigated how the sensitivity for detection of secondary peaks would increase with three years of weekly follow-up observations. In order for PG1302 to be observable throughout the year, a network of telescopes is required (e.g., Las Cumbres Observatory; http://lcogt.net/). Given that PG1302 is a bright quasar, even small telescopes can observe it with relatively small photometric errors. To allow for a flexible, realistic (affected by weather conditions, resources etc.) observation schedule as well as to avoid strong aliasing from an exactly periodic time sampling, we generate a hypothetical follow-up data set, consisting of one randomly chosen day each week, and allowing an additional scatter of ± 3 h in the observation times. We also assume photometric errors of 10 mmag, comparable to the ones in G15.

We analyse the extended light curve as before: we generate 20,000 time series that exhibit red noise variability with a periodic component as in G15, sampled at the observation times (including both existing and future data) and calculate the LS periodogram. We set our significance threshold to the maximum power of the periodogram from the 20,000 realisations (FAP$<$1%). Next, as done above, we include a secondary sinusoidal term and identify the minimum amplitude that would result in a peak with power above the threshold 90% of the time.

Fig. 3.5 shows the upper limits for detection with the extended data, compared to the upper limits with only the current data. For low frequencies, the sensitivity does not improve with the addition of three-years of data. A longer monitoring campaign is required for this purpose. For frequencies relevant to scenarios (A) and (C), the improvement varies from marginal to a factor of 2, depending on the precise frequency. Significant improvement is observed for frequencies where the aliasing peaks occur in the current data.

3.4.3 Implications for SMBH Binary Models

The lack of a significant secondary peak can place constraints on the physical parameters of the SMBHB PG1302. In particular, models that predict two peaks, comparable in height, and separated in frequency by a factor of few, can be ruled out. While this is not a generic prediction of hydrodynamical simulations, for specific mass ratios, within the range

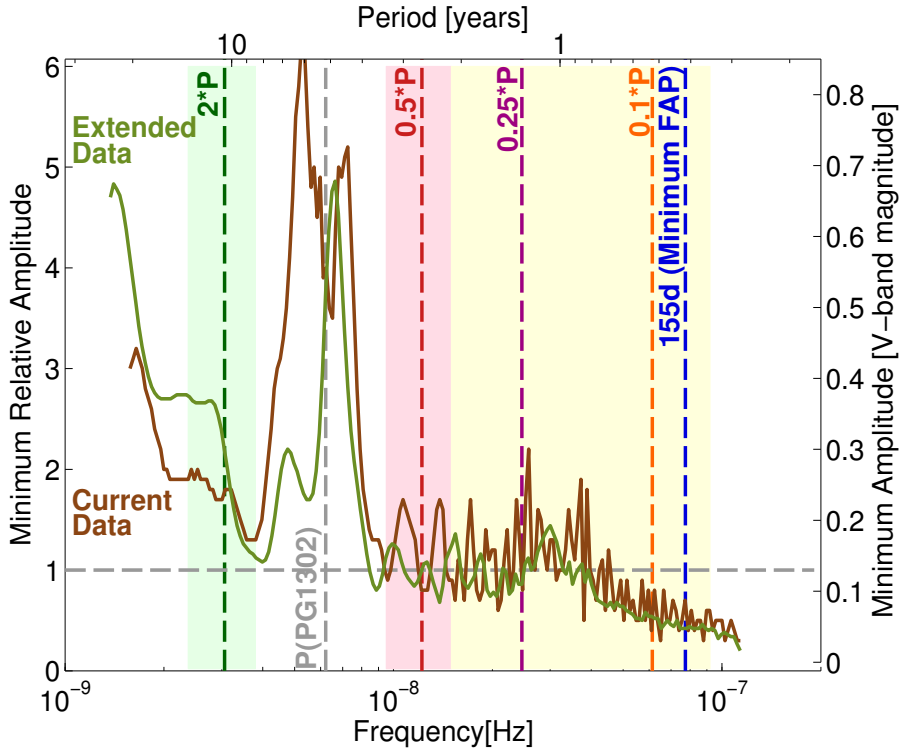


Figure 3.5: Minimum relative amplitude (also in V mag) of a secondary sinusoid versus frequency of secondary sinusoid that could be detectable in the current data (brown) and in data extended with three years of weekly follow-up observations (olive). The shaded regions show the frequencies of interest, with color coding as before.

of $0.25 < q < 0.45$, two near-equal peaks are predicted in the periodograms of the mass accretion rate onto the BHs (see their Fig. 9 in [Farris *et al.*, 2014]). The specific value of the disfavoured ratio q depends on disk parameters (viscosity, temperature) and also on the definition of accretion rate. If the accretion rate is defined based on the total mass accreting into the central cavity, this q value would be somewhat above $q = 0.5$ (see Fig. 6 in [D'Orazio *et al.*, 2013]).

3.5 Conclusions

The presence of multiple periodic components in the variability of SMBHBs is predicted by hydrodynamical simulations of circumbinary disks. The detection of multiple periodicity would provide additional indication that PG1302 is a SMBHB. We analysed the LS periodogram of the optical light-curve of the SMBHB candidate quasar PG1302 to search for multiple periodic components, and assessed the statistical significance of secondary peaks by simulating the variability of PG1302. Our analysis returned a peak with FAP=3%, below our detection threshold. By injecting fake secondary sinusoids into our models, we found that the current data would only allow the detection of secondary periodic modulations comparable in amplitude to the main peak. Future observations (e.g., with a baseline of a few years and weekly sampling), and/or a more sophisticated analysis, could significantly improve the detection capability and help uncover secondary peaks expected for SMBH binaries.

Chapter 4

Testing the relativistic boost hypothesis in Supermassive Black Hole Binaries

4.1 Introduction

It is well established that every massive galaxy harbors a Supermassive Black Hole (SMBH) in its nucleus. The mass of the central BH is correlated with several properties of the host galaxy (e.g., dispersion velocity, bulge luminosity, etc.), which suggests that the SMBH and the host galaxy may co-evolve [Kormendy and Ho, 2013]. Additionally, according to cosmological models of galaxy formation, galaxies are built-up from lower mass progenitors through frequent mergers [Haehnelt and Kauffmann, 2002]. Each merger delivers a Supermassive Black Hole Binary (SMBHB) at the center of the newly formed galaxy along with significant amounts of gas [Barnes and Hernquist, 1992].

In the post-merger galaxy, the orbit of the binary shrinks initially due to dynamical friction and subsequently by scattering nearby stars [Begelman *et al.*, 1980]. At smaller (sub-parsec) separations, the orbital decay is dominated by the emission of low-frequency Gravitational Waves (GWs), which could be detectable in the future by Pulsar Timing Arrays [Manchester and IPTA, 2013] and eLISA[Amaro-Seoane *et al.*, 2017]. It is currently

unclear whether three-body interactions can extract adequate energy to bring SMBHBs to the GW regime and the binaries may stall at parsec separations ("final parsec problem", see [Colpi, 2014]). However, the ambient gas, which settles in a circumbinary disk [Barnes, 2002], can also play significant role in accelerating the binary evolution, while, at the same time, it can accrete onto the BHs providing bright electromagnetic emission.

The orbital decay of the binary due to the interactions with the circumbinary disk is expected to be slow [Haiman *et al.*, 2009; Kocsis *et al.*, 2012a; Kocsis *et al.*, 2012b; Rafikov, 2013; Rafikov, 2016] and SMBHBs should spend significant time at sub-pc separations and thus should be fairly common. Despite their expected abundance, SMBHBs have only been detected at large separations, from several kpc [Komossa *et al.*, 2003; Bianchi *et al.*, 2008; Green *et al.*, 2010; Koss *et al.*, 2011; Fabbiano *et al.*, 2011; Comerford *et al.*, 2011; Fu *et al.*, 2015], down to a few pc [Rodriguez *et al.*, 2006]. At sub-pc separations, it is nearly impossible to spatially resolve the individual BHs and one has to rely on the effects of the binary on its environment. Several candidates have been identified, from large velocity offsets in quasar spectra, helical morphology of radio jets, etc (see [Runnoe *et al.*, 2017; Wang *et al.*, 2017; Kun *et al.*, 2014], and references therein).

SMBHBs can also be identified as quasars with periodic variability. This is a natural expectation since: (1) Quasars are thought to be activated by galaxy mergers [Gaskell, 1985] and thus may host SMBHBs. (2) SMBHBs are expected to produce bright quasar-like luminosities periodically modulated at the orbital period of the binary; several hydrodynamical simulations have concluded that the mass accretion rate onto the BHs and potentially the luminosity of the binary varies periodically [Hayasaki *et al.*, 2007; MacFadyen and Milosavljević, 2008; Noble *et al.*, 2012; Shi *et al.*, 2012; Roedig *et al.*, 2012; D'Orazio *et al.*, 2013; Farris *et al.*, 2014; Gold *et al.*, 2014]. In a binary system, periodic variability can also arise from a precessing jet, as the viewing angle of the jet varies periodically.

Recently, a large number of quasars with significant periodicity has emerged, mainly from systematic searches in the photometric databases of the large time-domain surveys. [Graham *et al.*, 2015a] (hereafter G15) identified 111 candidates in a sample of ∼245,000 quasars from the Catalina Real-Time Transient Survey (CRTS). [Charisi *et al.*, 2016] (hereafter C16) analyzed a sample of ∼35,000 quasars from the Palomar Transient Factory (PTF)

and identified 33 additional candidates, with typically shorter periods and dimmer magnitudes. Another candidate, quasar SDSS J0159+0105, was also identified to have two periodic components in its variability with a characteristic pattern 1:2 from a smaller sample of $\sim$ 350 low-redshift quasars in CRTS [Zheng *et al.*, 2016].[1]

Identifying periodic variability in quasars is extremely challenging, because quasars show stochastic variability. The variability is successfully described by a Damped Random Walk (DRW) model [Kelly *et al.*, 2009; Kozłowski *et al.*, 2010; MacLeod *et al.*, 2010] although some recent studies have suggested that an advanced description of quasar variability may be required [Mushotzky *et al.*, 2011; Zu *et al.*, 2013; Graham *et al.*, 2014; Simm *et al.*, 2016].[2] Additionally, the optical light curves are typically sparse with seasonal gaps and relatively short baselines compared to the periods (the identified candidates were observed only for a limited number of 2-3 cycles). As pointed out by [Vaughan *et al.*, 2016], our incomplete knowledge of quasar variability, in combination with the limited quality of the optical light curves can lead to false detections of periodicity. In fact, this was demonstrated in the case of quasar PSO J3334.2028+01.4075 [Liu *et al.*, 2015], which was proven to not have persistent periodicity in follow-up monitoring [Liu *et al.*, 2016b].[3]

It is, thus, crucial to explore additional signatures that could support the binary nature of the periodic sample. Several such studies followed the discovery of the first quasar with periodic variability (PKS 1302-102; [Graham *et al.*, 2015b]), including the search for multiple periodic components in the optical variability, as expected from hydrodynamical simulations [Charisi *et al.*, 2015; D'Orazio *et al.*, 2015a], the analysis of the helical structure in the radio jet of PKS 1302-102 [Kun *et al.*, 2015; Mohan *et al.*, 2016], and the detection of periodic variability in the infrared [Jun *et al.*, 2015; D'Orazio and Haiman, 2017].

[1]This quasar was also included in the sample analyzed in G15, but was not identified as periodic. The shortest periodic component of $\sim$740 d was identified with both the wavelet and the autocorrelation function analysis in G15, but did not have enough power to be classified as periodic, potentially due to the second periodic component (M. Graham; private communication).

[2]The periodicity search algorithms and statistical analyses in the papers above were significantly different. However, in all cases, the underlying assumption for the quasar variability was the DRW model.

[3]We note, however, that the initial detection in [Liu *et al.*, 2015] did not explicitly take into account the stochastic variability of quasars.

Another proposed signature of SMBHBs is to detect evidence for relativistic Doppler Boost. Assuming that the observed periodicity corresponds to the redshifted orbital period of the binary, we can infer that most of the candidates are at sub-pc separations, orbiting with mildly relativistic velocities (a few percent of the speed of light). If the optical emission arises in gas bound to the individual BHs, e.g., in mini-disks seen in hydrodynamical simulations [Farris *et al.*, 2014; Ryan and MacFadyen, 2017], the effects of special relativity are inevitable and may dominate the variability.

In detail, the perceived frequency of the emitted photons will change due to the relativistic motion. Additionally, the number of photons, which is proportional to F_ν/ν^3, with F_ν the flux at a specific frequency ν, is Lorentz invariant (all the observers see the same number of photons). Assuming that the emitted radiation has a power-law spectrum $F_\nu \sim \nu^\alpha$, it can be shown (see [D'Orazio *et al.*, 2015b] for the detailed derivation) that the variability due to Doppler boost to a first order in velocity is

$$\frac{\Delta F_\nu}{F_\nu} = (3-\alpha)\frac{v}{c}\cos\phi\sin i \tag{4.1}$$

where v is the orbital velocity of the most luminous BH (typically the less massive secondary BH), i is the inclination of the orbit to the line of sight and ϕ the phase of the orbit.

Therefore, even if the optical luminosity in the mini-disks is constant, the binary will appear blue-shifted and brighter, when the most luminous BH is moving towards the observer, and vice-versa. The relativistic Doppler boost may naturally explain the observed light curves, which show smooth quasi-sinusoidal periodicity. We note that the periodic accretion rate, is expected to produce more "bursty" periodic profiles.

Moreover, the relativistic Doppler boost provides a robust prediction, which can be tested with multi-wavelength data. For instance, if the UV emission also arises in the mini-disks, then the UV variability should also follow eq. (4.1). This means that the UV light curve should track the optical, but with a different variability amplitude, which depends on the spectral index in the respective band. The relative amplitudes can be derived from eq. (4.1) as follows

$$\frac{A_{UV}}{A_{opt}} = \frac{3-\alpha_{UV}}{3-\alpha_{opt}} \tag{4.2}$$

where, A_{opt}, A_{UV} are the amplitudes of the optical, and UV variability, respectively and

α_{opt}, α_{UV} are the spectral indices in the two bands.[4]

[D'Orazio *et al.*, 2015b] proposed this model to explain the variability of the periodic candidate PKS 1302-102. In particular, they found that if PKS 1302-102 hosts an unequal mass binary ($q \equiv M_1/M_2 < 0.05$), orbiting nearly face-on, the Doppler boost should dominate the variability. The model can successfully fit the optical light curve and is also consistent with the variability in the near-UV (nUV) and far-UV (fUV) bands, tested with archival data from the Galaxy Evolution Explorer (GALEX) and the Hubble Space Telescope (HST).

In this paper, we examine whether the proposed Doppler boost test is indeed a smoking gun signature for SMBHBs. For this, we test whether we can robustly differentiate the prediction of the relativistic Doppler boost from generic multi-wavelength variability of quasars using a control sample of non-periodic quasars. The paper is organized as follows: In Section 4.2, we present the sample selection and the data analysis. In Section 4.3 we show the results of our analysis.

4.2 Sample and Methods

4.2.1 Sample

We want to test the null hypothesis that the Doppler boost signature arises from intrinsic multi-wavelength variability of quasars. In other words, we want to test whether the relative amplitude of optical and UV variability is by chance consistent with the prediction from eq. (4.2). For this purpose, we assembled a control sample of quasars and Active Galactic Nuclei (AGNs) that do not exhibit periodic variability.

In order to calculate the Doppler boost test, the following data are necessary for each source in the control sample: (1) an optical light curve, (2) a UV light curve, temporally coincident with the optical light curve,[5] (3) an optical spectrum, and (4) a UV spectrum.

[4]The X-ray luminosity likely arises in gas bound to the SMBHs and the test could be expanded in the X-rays, as well. In this paper, we will focus only in the optical/UV variability.

[5]We only considered UV light curves consisting of at least two distinct epochs, separated by 100 d, within the optical baseline (see below).

Since the availability of UV spectroscopy is limited, we maximized the sample size, starting from a sample of sources with available UV spectra.[6]

More specifically, we made use of two HST spectroscopic catalogs, which provide calibrated co-added UV spectra: (1) the Atlas of Recalibrated HST Faint Object Spectrograph (FOS) Spectra of AGN and Quasars [Evans and Koratkar, 2004] and (2) the Cosmic Origin Spectrograph (COS) quasar catalog of fUV spectra, provided by the Hubble Spectroscopic Legacy Archive (HSLA; [Peeples *et al.*, 2016]).[7] The FOS and COS catalogs include spectra from 204 and 564 sources, respectively, among which 56 are common in both catalogs.

From the COS catalog, we eliminated five sources that we could not identify in the astronomical database Simbad[8] and two X-ray binaries. We also excluded from both catalogs sources classified as Bl Lac Objects, since the variability of these objects is dominated by the jet. These sources were also excluded from the searches for periodicity in G15 and C16. We excluded 5 additional sources from the COS catalog, and one source (PKS 1302-102) which was included in both catalogs, because they were identified as periodic.

For the remaining sources, we extracted optical spectra from the Sloan Digital Sky Survey (SDSS),[9] optical light curves from CRTS[10] and UV light curves from GALEX.[11] Since, as mentioned above, light curves and spectra in both bands are necessary for the Doppler test we kept in the sample only the sources that had all the necessary information. We note that GALEX obtained simultaneously photometric measurements with two UV filters (covering the nUV and fUV bands), but the fUV light curves are typically more sparse, because quasars are fainter in fUV. Moreover, the UV spectra usually do not extend to both bands. So if, for instance, the UV spectrum covers only the fUV band, we checked

[6] Light curves in optical and UV bands, as well as optical spectra are available for very large samples of quasars from CRTS, GALEX and SDSS, respectively.

[7] To the best of our knowledge, the above are the only available catalogs with high-level quasar and AGN spectra. Calibrated nUV spectra from COS, and from the Space Telescope Imaging Spectrograph (STIS) will be released in the future in the HSLA (`https://hla.stsci.edu`).

[8] `http://simbad.u-strasbg.fr/simbad/`

[9] `https://dr13.sdss.org`

[10] `http://nesssi.cacr.caltech.edu/DataRelease/`

[11] `http://galex.stsci.edu`

only for the availability of the fUV light curve (with the additional requirement mentioned before, i.e. two distinct epochs and coincidence with the optical data).

In order to obtain reliable estimates of the spectral slopes, we imposed a redshift cut at $z \leq 0.5$ for the COS catalog and at $z \leq 1.3$ for the FOS catalog. Additionally, we did not consider wavelengths shorter than the Lyman limit. This requirement ensures that the spectra are not strongly affected by absorption from the interstellar medium. [12] For the same reason, we excluded a few spectra with low signal-to-noise ratio or with very strong absorption features. Since we estimated the spectral slopes based only on the line-free regions of the spectra (see below), we further required that the wavelength range of those regions covers a significant fraction of the respective GALEX band. As for the GALEX bands, we considered the wavelength range at FWHM of the filter transmition curve extended by 10% at both sides; this is typically the amount of Doppler redshift/blueshift in the periodic sample.

Given all the above necessary constraints, the test of the null hypothesis was feasible for 30 sources in the nUV band (all the sources were from the FOS catalog) and for 18 sources in the fUV band (12 from the COS catalog, five from the FOS catalog and one common in both catalogs).

4.2.2 Optical and UV Spectra

We assume that the continuum of quasar/AGN spectra can be approximated by a single power-law $F_\lambda \sim \lambda^\beta$, where $\beta = -\alpha - 2$. For this, it is essential to avoid the broad emission lines and fit on line-free regions of the spectrum.

For the optical spectra, we automated the fit taking advantage of the spectral line information provided by the SDSS pipeline. More specifically, each line is fit with a Gaussian profile and for lines with a valid fit (as indicated by the pipeline) we removed the main part of the line by interpolating between $-\sigma$ to $+\sigma$ from the central wavelength. Subsequently, we smoothed the spectra (to remove potential contribution of the parts of the broad lines that were not removed by the interpolation) by calculating the moving average over a wide

[12] We imposed different redshift cuts for the two catalogs, because the spectra from COS cover only the fUV band, whereas spectra from FOS extend to nUV, as well.

window of 400 wavelength bins and fit a power-law on the moving average.

For the UV spectra, on the other hand, the spectral line information is not included in the catalogs. Therefore we manually identify the line free regions and fit the continuum based on the line free regions. In both cases, we confirmed with visual inspection that the validity of the fit.

If more than one spectrum was available, we used the average slope from all the spectra.

4.2.3 Photometric Fits

We extracted the optical light curves from the public CRTS database (Data Release 2). The survey combines data from three different telescopes and thus the lightcurves may consist of multiple data streams, which are not calibrated as a single light curve. In order to avoid systematic effects, we selected the light curve with the highest number of observations. Additionallly, the CRTS light curves contain four data points per night (four observations per visit separated by 10 min). Since the short-term variability is not significant for this study, we binned the observations taken within the same night. Next, we extracted UV light curves from GALEX. These light curves, especially in the fUV band, are typically sparse with only a handful of epochs.

The next step consisted of measuring the relative amplitudes in optical and UV (left side of eq. 4.2). This step is challenging, because quasars exhibit stochastic variability and the optical and UV data were not taken simultaneously. For this, we approximated the optical variability with an n^{th} degree polynomial. We chose the order of the polynomial between 5 and 20 based on the reduced chi square (we chose the smallest polynomial order which gives $\chi^2_{dof} \sim 1$) combined with visual inspection of the fitted light curves to ensure that the chosen fit reasonably represents the variability. For instance, in some cases, the reduced chi square is smaller than unity ($\chi^2_{dof} < 1$) for all the tested polynomials, which may suggest that the provided error bars are overestimated, whereas, in other cases, it is always higher than one ($\chi^2_{dof} > 1$), because the polynomial fit cannot entirely capture the short-term variability.

The relative amplitude is calculated by rescaling the polynomial fit of the optical light curve to fit the UV light curve. We emphasize that the polynomial fit, is sufficient for

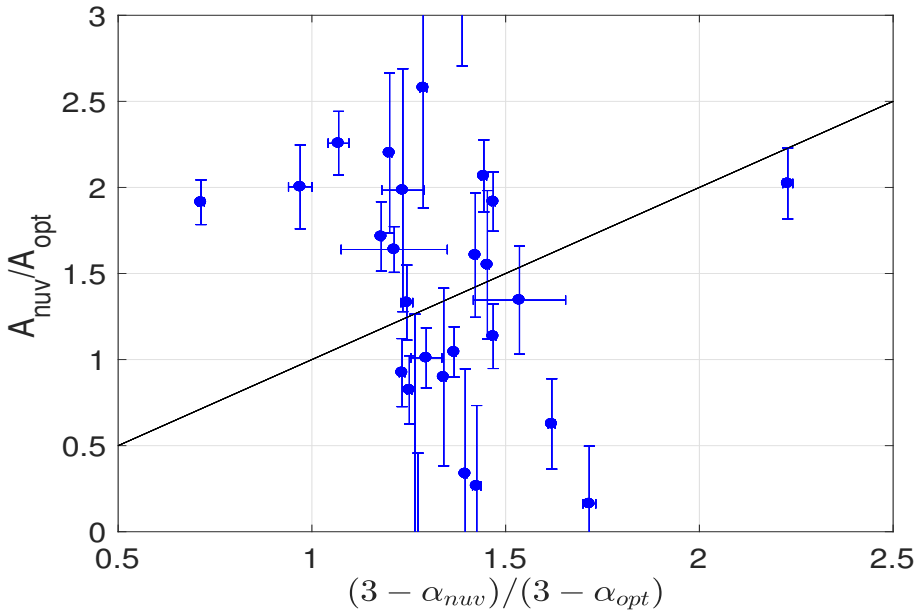


Figure 4.1: Doppler boost test in the nUV band for a control sample of 30 non-periodic quasars.

interpolating the light curves, but it does not have predictive power. Therefore, UV data points that do not overlap with the optical were excluded from the fit.

4.3 Results

We tested the null hypothesis that the Doppler boost signature arises by multi-wavelength variability of quasars. We see that the variability is consistent with the Doppler prediction in 7 out of 30 quasars, when we consider data in the nUV band. This means that the Doppler boost signature can be observed by chance in $\sim$25% of the quasars in the sample.

Chapter 5

Conclusions

Supermassive black hole binaries (SMBHBs) at sub-parsec separations should be common in galactic nuclei, as a result of frequent galaxy mergers. Hydrodynamical simulations of circumbinary discs predict strong periodic modulation of the mass accretion rate on time-scales comparable to the orbital period of the binary. As a result, SMBHBs may be recognized by the periodic modulation of their brightness. In this thesis:

- We conducted a statistical search for periodic variability in a sample of 35,383 spectroscopically confirmed quasars in the photometric database of the Palomar Transient Factory (PTF). We analysed Lomb-Scargle periodograms and assessed the significance of our findings by modeling each individual quasar's variability as a damped random walk (DRW).

- We identified 50 quasars with significant periodicity beyond the DRW model, typically with short periods of a few hundred days. We find 33 of these to remain significant after a re-analysis of their periodograms including additional optical data from the intermediate-PTF and the Catalina Real-Time Transient Survey (CRTS).

- Assuming that the observed periods correspond to the redshifted orbital periods of SMBHBs, we conclude that our findings are consistent with a population of unequal-mass SMBHBs, with a typical mass ratio as low as $q \equiv M_2/M_1 \approx 0.01$.

- We explored additional independent signatures that could verify the binary nature

of the periodic candidates. First, we focused on a case study of quasar PKS 1302-102; [Graham *et al.*, 2015b] discovered a supermassive black hole binary (SMBHB) candidate and identified the detected 5.2 yr period of the optical variability as the orbital period of the binary. Hydrodynamical simulations predict multiple periodic components for the variability of SMBHBs, thus raising the possibility that the true period of the binary is different from 5.2 yr.

- We analysed the periodogram of PKS 1302-102 and find no compelling evidence for additional peaks. We also pointed out that, despite the 5.2 yr peak being significant if a single source is considered, further analysis is required to account for the fact that PG1302 was selected among a large sample of 247,000 quasars.

- We derived upper limits on any additional periodic modulations in the available data, by modeling the light-curve as the sum of stochastic noise and the known 5.2 yr periodic component, and injecting additional sinusoidal signals. We found that, with the current data, we would be able to detect with high significance (false alarm probability $< 1\%$) secondary periodic terms, with periods in the range predicted by the simulations, if the amplitude of the variability was at least $\sim$0.06 mag (compared to 0.14 mag for the main sinusoid).

- A three-year follow-up monitoring campaign with weekly observations can increase the sensitivity for detecting secondary peaks by $\approx$50%, and would allow a more robust test of predictions from hydrodynamical simulations.

- We examined the relativistic Doppler boost, as a potential signature for SMBHBs. In compact binaries, the relativistic orbital velocity of the BHs could cause the smooth sinusoidal variability, we observe. This model also predicts that the UV luminosity will track the optical periodicity, but with a different amplitude. The relative amplitudes depend on the spectral indices in the respective bands.

- We analyzed data from the HST and the GALEX for a control sample of non-periodic quasars. We tested the null hypothesis that the Doppler signature is produced by intrinsic multi-wavelength variability of quasars. We found that the Doppler signature

arises by chance in ∼25% of the quasars in the sample. We concluded that the Doppler boost is not a sharp test for the binarity of the periodic quasars.

Bibliography

[Abbott *et al.*, 2016] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters*, 116(6):061102, February 2016.

[Ade *et al.*, 2015] P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *ArXiv e-prints*, February 2015.

[Amaro-Seoane *et al.*, 2013] P. Amaro-Seoane, S. Aoudia, S. Babak, P. Binétruy, E. Berti, A. Bohé, C. Caprini, M. Colpi, N. J. Cornish, K. Danzmann, J.-F. Dufaux, J. Gair, I. Hinder, O. Jennrich, P. Jetzer, A. Klein, R. N. Lang, A. Lobo, T. Littenberg, S. T. McWilliams, G. Nelemans, A. Petiteau, E. K. Porter, B. F. Schutz, A. Sesana, R. Stebbins, T. Sumner, M. Vallisneri, S. Vitale, M. Volonteri, H. Ward, and B. Wardell. eLISA: Astrophysics and cosmology in the millihertz regime. *GW Notes, Vol. 6, p. 4-110*, 6:4–110, May 2013.

[Amaro-Seoane *et al.*, 2017] P. Amaro-Seoane, H. Audley, S. Babak, J. Baker, E. Barausse, P. Bender, E. Berti, P. Binetruy, M. Born, D. Bortoluzzi, J. Camp, C. Caprini, V. Cardoso, M. Colpi, J. Conklin, N. Cornish, C. Cutler, K. Danzmann, R. Dolesi, L. Ferraioli, V. Ferroni, E. Fitzsimons, J. Gair, L. Gesa Bote, D. Giardini, F. Gibert, C. Grimani, H. Halloin, G. Heinzel, T. Hertog, M. Hewitson, K. Holley-Bockelmann, D. Hollington, M. Hueller, H. Inchauspe, P. Jetzer, N. Karnesis, C. Killow, A. Klein, B. Klipstein, N. Korsakova, S. L Larson, J. Livas, I. Lloro, N. Man, D. Mance, J. Martino, I. Ma-

teos, K. McKenzie, S. T McWilliams, C. Miller, G. Mueller, G. Nardini, G. Nelemans, M. Nofrarias, A. Petiteau, P. Pivato, E. Plagnol, E. Porter, J. Reiche, D. Robertson, N. Robertson, E. Rossi, G. Russano, B. Schutz, A. Sesana, D. Shoemaker, J. Slutsky, C. F. Sopuerta, T. Sumner, N. Tamanini, I. Thorpe, M. Troebs, M. Vallisneri, A. Vecchio, D. Vetrugno, S. Vitale, M. Volonteri, G. Wanner, H. Ward, P. Wass, W. Weber, J. Ziemer, and P. Zweifel. Laser Interferometer Space Antenna. *ArXiv e-prints*, February 2017.

[Andrae *et al.*, 2013] R. Andrae, D.-W. Kim, and C. A. L. Bailer-Jones. Assessment of stochastic and deterministic models of 6304 quasar lightcurves from SDSS Stripe 82. *A&A*, 554:A137, June 2013.

[Armitage and Natarajan, 2005] P. J. Armitage and P. Natarajan. Eccentricity of Supermassive Black Hole Binaries Coalescing from Gas-rich Mergers. *ApJ*, 634:921–927, December 2005.

[Artymowicz and Lubow, 1994] P. Artymowicz and S. H. Lubow. Dynamics of binary-disk interaction. 1: Resonances and disk gap sizes. *ApJ*, 421:651–667, February 1994.

[Artymowicz and Lubow, 1996] P. Artymowicz and S. H. Lubow. Mass Flow through Gaps in Circumbinary Disks. *ApJL*, 467:L77, August 1996.

[Barnes and Hernquist, 1992] J. E. Barnes and L. Hernquist. Dynamics of interacting galaxies. *ARA&A*, 30:705–742, 1992.

[Barnes, 2002] J. E. Barnes. Formation of gas discs in merging galaxies. *MNRAS*, 333:481–494, July 2002.

[Barrows *et al.*, 2013] R. S. Barrows, C. H. Sandberg Lacy, J. Kennefick, J. M. Comerford, D. Kennefick, and J. C. Berrier. Identification of Outflows and Candidate Dual Active Galactic Nuclei in SDSS Quasars at z = 0.8-1.6. *ApJ*, 769:95, June 2013.

[Barth *et al.*, 2008] A. J. Barth, M. C. Bentz, J. E. Greene, and L. C. Ho. An Offset Seyfert 2 Nucleus in the Minor Merger System NGC 3341. *ApJL*, 683:L119, August 2008.

[Baruteau *et al.*, 2012] C. Baruteau, E. Ramirez-Ruiz, and F. Masset. No snowplough mechanism during the rapid hardening of supermassive black hole binaries. *MNRAS*, 423:L65–L69, June 2012.

[Begelman *et al.*, 1980] M. C. Begelman, R. D. Blandford, and M. J. Rees. Massive black hole binaries in active galactic nuclei. *Nature*, 287:307–309, September 1980.

[Bekenstein, 1973] J. D. Bekenstein. Gravitational-Radiation Recoil and Runaway Black Holes. *ApJ*, 183:657–664, July 1973.

[Berczik *et al.*, 2006] P. Berczik, D. Merritt, R. Spurzem, and H.-P. Bischof. Efficient Merger of Binary Supermassive Black Holes in Nonaxisymmetric Galaxies. *ApJL*, 642:L21–L24, May 2006.

[Bertin and Arnouts, 1996] E. Bertin and S. Arnouts. SExtractor: Software for source extraction. *A& AS*, 117:393–404, June 1996.

[Bianchi *et al.*, 2008] S. Bianchi, M. Chiaberge, E. Piconcelli, M. Guainazzi, and G. Matt. Chandra unveils a binary active galactic nucleus in Mrk 463. *MNRAS*, 386:105–110, May 2008.

[Blaes *et al.*, 2002] O. Blaes, M. H. Lee, and A. Socrates. The Kozai Mechanism and the Evolution of Binary Supermassive Black Holes. *ApJ*, 578:775–786, October 2002.

[Bon *et al.*, 2012] E. Bon, P. Jovanović, P. Marziani, A. I. Shapovalova, N. Bon, V. Borka Jovanović, D. Borka, J. Sulentic, and L. Č. Popović. The First Spectroscopically Resolved Sub-parsec Orbit of a Supermassive Binary Black Hole. *ApJ*, 759:118, November 2012.

[Bon *et al.*, 2016] E. Bon, S. Zucker, H. Netzer, P. Marziani, N. Bon, P. Jovanović, A. I. Shapovalova, S. Komossa, C. M. Gaskell, L. Č. Popović, S. Britzen, V. H. Chavushyan, A. N. Burenkov, S. Sergeev, G. La Mura, J. R. Valdés, and M. Stalevski. Evidence for Periodicity in 43 year-long Monitoring of NGC 5548. *ApJS*, 225:29, August 2016.

[Boroson and Lauer, 2009] T. A. Boroson and T. R. Lauer. A candidate sub-parsec supermassive binary black hole system. *Nature*, 458:53–55, March 2009.

[Britzen *et al.*, 2001] S. Britzen, J. Roland, J. Laskar, K. Kokkotas, R. M. Campbell, and A. Witzel. On the origin of compact radio sources. The binary black hole model applied to the gamma-bright quasar PKS 0420-014. *A& A*, 374:784–799, August 2001.

[Caproni *et al.*, 2013] A. Caproni, Z. Abraham, and H. Monteiro. Parsec-scale jet precession in BL Lacertae (2200+420). *MNRAS*, 428:280–290, January 2013.

[Chang *et al.*, 2010] P. Chang, L. E. Strubbe, K. Menou, and E. Quataert. Fossil gas and the electromagnetic precursor of supermassive binary black hole mergers. *MNRAS*, 407:2007–2016, September 2010.

[Charisi *et al.*, 2015] M. Charisi, I. Bartos, Z. Haiman, A. M. Price-Whelan, and S. Márka. Multiple periods in the variability of the supermassive black hole binary candidate quasar PG1302-102? *MNRAS*, 454:L21–L25, November 2015.

[Charisi *et al.*, 2016] M. Charisi, I. Bartos, Z. Haiman, A. M. Price-Whelan, M. J. Graham, E. C. Bellm, R. R. Laher, and S. Márka. A population of short-period variable quasars from PTF as supermassive black hole binary candidates. *MNRAS*, 463:2145–2171, December 2016.

[Chornock *et al.*, 2010] R. Chornock, J. S. Bloom, S. B. Cenko, A. V. Filippenko, J. M. Silverman, M. D. Hicks, K. J. Lawrence, A. J. Mendez, M. Rafelski, and A. M. Wolfe. The Quasar SDSS J1536+0441: An Unusual Double-peaked Emitter. *ApJL*, 709:L39–L43, January 2010.

[Colpi, 2014] M. Colpi. Massive Binary Black Holes in Galactic Nuclei and Their Path to Coalescence. *SSRv*, 183:189–221, September 2014.

[Comerford and Greene, 2014] J. M. Comerford and J. E. Greene. Offset Active Galactic Nuclei as Tracers of Galaxy Mergers and Supermassive Black Hole Growth. *ApJ*, 789:112, July 2014.

[Comerford *et al.*, 2009] J. M. Comerford, B. F. Gerke, J. A. Newman, M. Davis, R. Yan, M. C. Cooper, S. M. Faber, D. C. Koo, A. L. Coil, D. J. Rosario, and A. A. Dut-

ton. Inspiralling Supermassive Black Holes: A New Signpost for Galaxy Mergers. *ApJ*, 698:956–965, June 2009.

[Comerford *et al.*, 2011] J. M. Comerford, D. Pooley, B. F. Gerke, and G. M. Madejski. Chandra Observations of a 1.9 kpc Separation Double X-Ray Source in a Candidate Dual Active Galactic Nucleus Galaxy at z = 0.16. *ApJL*, 737:L19, August 2011.

[Comerford *et al.*, 2013] J. M. Comerford, K. Schluns, J. E. Greene, and R. J. Cool. Dual Supermassive Black Hole Candidates in the AGN and Galaxy Evolution Survey. *ApJ*, 777:64, November 2013.

[Comerford *et al.*, 2015] J. M. Comerford, D. Pooley, R. S. Barrows, J. E. Greene, N. L. Zakamska, G. M. Madejski, and M. C. Cooper. Merger-driven Fueling of Active Galactic Nuclei: Six Dual and Offset AGNs Discovered with Chandra and Hubble Space Telescope Observations. *ApJ*, 806:219, June 2015.

[Conselice, 2014] C. J. Conselice. The Evolution of Galaxy Structure Over Cosmic Time. *ARA&A*, 52:291–337, August 2014.

[Crenshaw *et al.*, 2010] D. M. Crenshaw, S. B. Kraemer, H. R. Schmitt, Y. L. Jaffé, R. P. Deo, N. R. Collins, and T. C. Fischer. The Geometry of Mass Outflows and Fueling Flows in the Seyfert 2 Galaxy MRK 3. *AJ*, 139:871–877, March 2010.

[Cuadra *et al.*, 2009] J. Cuadra, P. J. Armitage, R. D. Alexander, and M. C. Begelman. Massive black hole binary mergers within subparsec scale gas discs. *MNRAS*, 393:1423–1432, March 2009.

[Deane *et al.*, 2014] R. P. Deane, Z. Paragi, M. J. Jarvis, M. Coriat, G. Bernardi, R. P. Fender, S. Frey, I. Heywood, H.-R. Klöckner, K. Grainge, and C. Rumsey. A close-pair binary in a distant triple supermassive black hole system. *Nature*, 511:57–60, July 2014.

[Decarli *et al.*, 2010] R. Decarli, M. Dotti, C. Montuori, T. Liimets, and A. Ederoclite. The Peculiar Optical Spectrum of 4C+22.25: Imprint of a Massive Black Hole Binary? *ApJL*, 720:L93–L96, September 2010.

[Djorgovski *et al.*, 2012] S. G. Djorgovski, A. A. Mahabal, A. J. Drake, M. J. Graham, C. Donalek, and R. Williams. Exploring the Time Domain with Synoptic Sky Surveys. In E. Griffin, R. Hanisch, and R. Seaman, editors, *New Horizons in Time Domain Astronomy*, volume 285 of *IAU Symposium*, pages 141–146, April 2012.

[D'Orazio and Haiman, 2017] D. J. D'Orazio and Z. Haiman. Lighthouse in the Dust: Infrared Echoes of Periodic Emission from Massive Black Hole Binaries. *ArXiv e-prints*, February 2017.

[D'Orazio *et al.*, 2013] D. J. D'Orazio, Z. Haiman, and A. MacFadyen. Accretion into the central cavity of a circumbinary disc. *MNRAS*, 436:2997–3020, December 2013.

[D'Orazio *et al.*, 2015a] D. J. D'Orazio, Z. Haiman, P. Duffell, B. D. Farris, and A. I. MacFadyen. A reduced orbital period for the supermassive black hole binary candidate in the quasar PG 1302-102? *MNRAS*, 452:2540–2545, September 2015.

[D'Orazio *et al.*, 2015b] D. J. D'Orazio, Z. Haiman, and D. Schiminovich. Relativistic boost as the cause of periodicity in a massive black-hole binary candidate. *Nature*, 525:351–353, September 2015.

[D'Orazio *et al.*, 2016] D. J. D'Orazio, Z. Haiman, P. Duffell, A. MacFadyen, and B. Farris. A transition in circumbinary accretion discs at a binary mass ratio of 1:25. *MNRAS*, 459:2379–2393, July 2016.

[Dotti *et al.*, 2012] M. Dotti, A. Sesana, and R. Decarli. Massive Black Hole Binaries: Dynamical Evolution and Observational Signatures. *Advances in Astronomy*, 2012:940568, 2012.

[Drake *et al.*, 2009] A. J. Drake, S. G. Djorgovski, A. Mahabal, E. Beshore, S. Larson, M. J. Graham, R. Williams, E. Christensen, M. Catelan, A. Boattini, A. Gibbs, R. Hill, and R. Kowalski. First Results from the Catalina Real-Time Transient Survey. *ApJ*, 696:870–884, May 2009.

[Drake *et al.*, 2013] A. J. Drake, M. Catelan, S. G. Djorgovski, G. Torrealba, M. J. Graham, V. Belokurov, S. E. Koposov, A. Mahabal, J. L. Prieto, C. Donalek, R. Williams,

S. Larson, E. Christensen, and E. Beshore. Probing the Outer Galactic Halo with RR Lyrae from the Catalina Surveys. *ApJ*, 763:32, January 2013.

[Dullo and Graham, 2014] B. T. Dullo and A. W. Graham. Depleted cores, multicomponent fits, and structural parameter relations for luminous early-type galaxies. *MNRAS*, 444:2700–2722, November 2014.

[Eggers *et al.*, 2000] D. Eggers, D. B. Shaffer, and D. Weistrop. Optical Variability of Radio-Luminous PG Quasars. *AJ*, 119:460–468, February 2000.

[Elvis *et al.*, 1994] M. Elvis, B. J. Wilkes, J. C. McDowell, R. F. Green, J. Bechtold, S. P. Willner, M. S. Oey, E. Polomski, and R. Cutri. Atlas of quasar energy distributions. *ApJS*, 95:1–68, November 1994.

[Eracleous *et al.*, 2012] M. Eracleous, T. A. Boroson, J. P. Halpern, and J. Liu. A Large Systematic Search for Close Supermassive Binary and Rapidly Recoiling Black Holes. *ApJS*, 201:23, August 2012.

[Escala *et al.*, 2004] A. Escala, R. B. Larson, P. S. Coppi, and D. Mardones. The Role of Gas in the Merging of Massive Black Holes in Galactic Nuclei. I. Black Hole Merging in a Spherical Gas Cloud. *ApJ*, 607:765–777, June 2004.

[Evans and Koratkar, 2004] I. N. Evans and A. P. Koratkar. A Complete Atlas of Recalibrated Hubble Space Telescope Faint Object Spectrograph Spectra of Active Galactic Nuclei and Quasars. I. Pre-COSTAR Spectra. *ApJS*, 150:73–164, January 2004.

[Fabbiano *et al.*, 2011] G. Fabbiano, J. Wang, M. Elvis, and G. Risaliti. A close nuclear black-hole pair in the spiral galaxy NGC3393. *Nature*, 477:431–434, September 2011.

[Fan *et al.*, 2002] J. H. Fan, R. G. Lin, G. Z. Xie, L. Zhang, D. C. Mei, C. Y. Su, and Z. M. Peng. Optical periodicity analysis for radio selected BL Lacertae objects (RBLs). *A& A*, 381:1–5, January 2002.

[Farina *et al.*, 2011] E. P. Farina, R. Falomo, and A. Treves. A study of six low-redshift quasar pairs. *MNRAS*, 415:3163–3167, August 2011.

[Farris *et al.*, 2014] B. D. Farris, P. Duffell, A. I. MacFadyen, and Z. Haiman. Binary Black Hole Accretion from a Circumbinary Disk: Gas Dynamics inside the Central Cavity. *ApJ*, 783:134, March 2014.

[Farris *et al.*, 2015a] B. D. Farris, P. Duffell, A. I. MacFadyen, and Z. Haiman. Binary black hole accretion during inspiral and merger. *MNRAS*, 447:L80–L84, February 2015.

[Farris *et al.*, 2015b] B. D. Farris, P. Duffell, A. I. MacFadyen, and Z. Haiman. Characteristic signatures in the thermal emission from accreting binary black holes. *MNRAS*, 446:L36–L40, January 2015.

[Fischer *et al.*, 2011] T. C. Fischer, D. M. Crenshaw, S. B. Kraemer, H. R. Schmitt, R. F. Mushotsky, and J. P. Dunn. Hubble Space Telescope Observations of the Double-peaked Emission Lines in the Seyfert Galaxy Markarian 78: Mass Outflows from a Single Active Galactic Nucleus. *ApJ*, 727:71, February 2011.

[Fish *et al.*, 2016] V. Fish, K. Akiyama, K. Bouman, A. Chael, M. Johnson, S. Doeleman, L. Blackburn, J. Wardle, and W. Freeman. Observingand ImagingActive Galactic Nuclei with the Event Horizon Telescope. *Galaxies*, 4:54, October 2016.

[Flesch, 2015] E. W. Flesch. The Half Million Quasars (HMQ) Catalogue. *PASA*, 32:10, March 2015.

[Foreman *et al.*, 2009] G. Foreman, M. Volonteri, and M. Dotti. Double Quasars: Probes of Black Hole Scaling Relationships and Merger Scenarios. *ApJ*, 693:1554–1562, March 2009.

[Fu *et al.*, 2011] H. Fu, Z.-Y. Zhang, R. J. Assef, A. Stockton, A. D. Myers, L. Yan, S. G. Djorgovski, J. M. Wrobel, and D. A. Riechers. A Kiloparsec-scale Binary Active Galactic Nucleus Confirmed by the Expanded Very Large Array. *ApJL*, 740:L44, October 2011.

[Fu *et al.*, 2015] H. Fu, A. D. Myers, S. G. Djorgovski, L. Yan, J. M. Wrobel, and A. Stockton. Radio-selected Binary Active Galactic Nuclei from the Very Large Array Stripe 82 Survey. *ApJ*, 799:72, January 2015.

[Garcia *et al.*, 1999] A. Garcia, L. Sodré, F. J. Jablonski, and R. J. Terlevich. Optical monitoring of quasars - I. Variability. *MNRAS*, 309:803–816, November 1999.

[Gaskell, 1985] C. M. Gaskell. Galactic mergers, starburst galaxies, quasar activity and massive binary black holes. *Nature*, 315:386, May 1985.

[Ge *et al.*, 2012] J.-Q. Ge, C. Hu, J.-M. Wang, J.-M. Bai, and S. Zhang. Double-peaked Narrow Emission-line Galaxies from the Sloan Digital Sky Survey. I. Sample and Basic Properties. *ApJS*, 201:31, August 2012.

[Ghez *et al.*, 1998] A. M. Ghez, B. L. Klein, M. Morris, and E. E. Becklin. High Proper-Motion Stars in the Vicinity of Sagittarius A*: Evidence for a Supermassive Black Hole at the Center of Our Galaxy. *ApJ*, 509:678–686, December 1998.

[Gold *et al.*, 2014] R. Gold, V. Paschalidis, Z. B. Etienne, S. L. Shapiro, and H. P. Pfeiffer. Accretion disks around binary black holes of unequal mass: General relativistic magnetohydrodynamic simulations near decoupling. *PhRvD*, 89(6):064060, March 2014.

[Gopal-Krishna *et al.*, 2012] Gopal-Krishna, P. L. Biermann, L. Á. Gergely, and P. J. Wiita. On the origin of X-shaped radio galaxies. *Research in Astronomy and Astrophysics*, 12:127–146, February 2012.

[Gould and Rix, 2000] A. Gould and H.-W. Rix. Binary Black Hole Mergers from Planet-like Migrations. *ApJL*, 532:L29–L32, March 2000.

[Gower *et al.*, 1982] A. C. Gower, P. C. Gregory, W. G. Unruh, and J. B. Hutchings. Relativistic precessing jets in quasars and radio galaxies - Models to fit high resolution data. *ApJ*, 262:478–496, November 1982.

[Graham *et al.*, 2014] M. J. Graham, S. G. Djorgovski, A. J. Drake, A. A. Mahabal, M. Chang, D. Stern, C. Donalek, and E. Glikman. A novel variability-based method for quasar selection: evidence for a rest-frame

54 d characteristic time-scale. *MNRAS*, 439:703–718, March 2014.

[Graham *et al.*, 2015a] M. J. Graham, S. G. Djorgovski, D. Stern, A. J. Drake, A. A. Mahabal, C. Donalek, E. Glikman, S. Larson, and E. Christensen. A systematic search

for close supermassive black hole binaries in the Catalina Real-time Transient Survey. *MNRAS*, 453:1562–1576, October 2015.

[Graham *et al.*, 2015b] M. J. Graham, S. G. Djorgovski, D. Stern, E. Glikman, A. J. Drake, A. A. Mahabal, C. Donalek, S. Larson, and E. Christensen. A possible close supermassive black-hole binary in a quasar with optical periodicity. *Nature*, 518:74–76, February 2015.

[Green *et al.*, 2010] P. J. Green, A. D. Myers, W. A. Barkhouse, J. S. Mulchaey, V. N. Bennert, T. J. Cox, and T. L. Aldcroft. SDSS J1254+0846: A Binary Quasar Caught in the Act of Merging. *ApJ*, 710:1578–1588, February 2010.

[Grindlay *et al.*, 2012] J. Grindlay, S. Tang, E. Los, and M. Servillat. Opening the 100-Year Window for Time-Domain Astronomy, April 2012.

[Gualandris and Merritt, 2008] A. Gualandris and D. Merritt. Ejection of Supermassive Black Holes from Galaxy Cores. *ApJ*, 678:780–797, May 2008.

[Haehnelt and Kauffmann, 2002] M. G. Haehnelt and G. Kauffmann. Multiple supermassive black holes in galactic bulges. *MNRAS*, 336:L61–L64, November 2002.

[Haiman *et al.*, 2009] Z. Haiman, B. Kocsis, and K. Menou. The Population of Viscosity- and Gravitational Wave-driven Supermassive Black Hole Binaries Among Luminous Active Galactic Nuclei. *ApJ*, 700:1952–1969, August 2009.

[Halpern and Eracleous, 2000] J. P. Halpern and M. Eracleous. The End of the Lines for OX 169: No Binary Broad-Line Region. *ApJ*, 531:647–653, March 2000.

[Hayasaki *et al.*, 2007] K. Hayasaki, S. Mineshige, and H. Sudou. Binary Black Hole Accretion Flows in Merged Galactic Nuclei. , 59:427–441, April 2007.

[Henden *et al.*, 2012] A. A. Henden, S. E. Levine, D. Terrell, T. C. Smith, and D. Welch. Data Release 3 of the AAVSO All-Sky Photometric Survey (APASS). *Journal of the American Association of Variable Star Observers (JAAVSO)*, 40:430, June 2012.

[Hennawi *et al.*, 2006] J. F. Hennawi, M. A. Strauss, M. Oguri, N. Inada, G. T. Richards, B. Pindor, D. P. Schneider, R. H. Becker, M. D. Gregg, P. B. Hall, D. E. Johnston, X. Fan,

S. Burles, D. J. Schlegel, J. E. Gunn, R. H. Lupton, N. A. Bahcall, R. J. Brunner, and J. Brinkmann. Binary Quasars in the Sloan Digital Sky Survey: Evidence for Excess Clustering on Small Scales. *AJ*, 131:1–23, January 2006.

[Hoffman and Loeb, 2007] L. Hoffman and A. Loeb. Dynamics of triple black hole systems in hierarchically merging massive galaxies. *MNRAS*, 377:957–976, May 2007.

[Ivezić *et al.*, 2014] Ž. Ivezić, A.J. Connolly, J.T. Vanderplas, and A. Gray. *Statistics, Data Mining and Machine Learning in Astronomy*. Princeton University Press, 2014.

[Janssen *et al.*, 2008] G. H. Janssen, B. W. Stappers, M. Kramer, M. Purver, A. Jessner, and I. Cognard. European Pulsar Timing Array. In C. Bassa, Z. Wang, A. Cumming, and V. M. Kaspi, editors, *40 Years of Pulsars: Millisecond Pulsars, Magnetars and More*, volume 983 of *American Institute of Physics Conference Series*, pages 633–635, February 2008.

[Jenet *et al.*, 2009] F. Jenet, L. S. Finn, J. Lazio, A. Lommen, M. McLaughlin, I. Stairs, D. Stinebring, J. Verbiest, A. Archibald, Z. Arzoumanian, D. Backer, J. Cordes, P. Demorest, R. Ferdman, P. Freire, M. Gonzalez, V. Kaspi, V. Kondratiev, D. Lorimer, R. Lynch, D. Nice, S. Ransom, R. Shannon, and X. Siemens. The North American Nanohertz Observatory for Gravitational Waves. *ArXiv e-prints*, September 2009.

[Ju *et al.*, 2013] W. Ju, J. E. Greene, R. R. Rafikov, S. J. Bickerton, and C. Badenes. Search for Supermassive Black Hole Binaries in the Sloan Digital Sky Survey Spectroscopic Sample. *ApJ*, 777:44, November 2013.

[Jun *et al.*, 2015] H. D. Jun, D. Stern, M. J. Graham, S. G. Djorgovski, A. Mainzer, R. M. Cutri, A. J. Drake, and A. A. Mahabal. Infrared Time Lags for the Periodic Quasar PG 1302-102. *ApJL*, 814:L12, November 2015.

[Kaastra and Roos, 1992] J. S. Kaastra and N. Roos. Massive Binary Black-Holes and Wiggling Jets. *A& A*, 254:96, February 1992.

[Kaiser *et al.*, 2010] N. Kaiser, W. Burgett, K. Chambers, L. Denneau, J. Heasley, R. Jedicke, E. Magnier, J. Morgan, P. Onaka, and J. Tonry. The Pan-STARRS wide-

field optical/NIR imaging survey. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7733 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 0, July 2010.

[Kasliwal *et al.*, 2015] V. P. Kasliwal, M. S. Vogeley, and G. T. Richards. Are the variability properties of the Kepler AGN light curves consistent with a damped random walk? *MNRAS*, 451:4328–4345, August 2015.

[Kelly *et al.*, 2009] B. C. Kelly, J. Bechtold, and A. Siemiginowska. Are the Variations in Quasar Optical Flux Driven by Thermal Fluctuations? *ApJ*, 698:895–910, June 2009.

[Kelly *et al.*, 2014] B. C. Kelly, A. C. Becker, M. Sobolewska, A. Siemiginowska, and P. Uttley. Flexible and Scalable Methods for Quantifying Stochastic Variability in the Era of Massive Time-domain Astronomical Data Sets. *ApJ*, 788:33, June 2014.

[Khan *et al.*, 2013] F. M. Khan, K. Holley-Bockelmann, P. Berczik, and A. Just. Supermassive Black Hole Binary Evolution in Axisymmetric Galaxies: The Final Parsec Problem is Not a Problem. *ApJ*, 773:100, August 2013.

[Kocsis *et al.*, 2012a] B. Kocsis, Z. Haiman, and A. Loeb. Gas pile-up, gap overflow and Type 1.5 migration in circumbinary discs: general theory. *MNRAS*, 427:2660–2679, December 2012.

[Kocsis *et al.*, 2012b] B. Kocsis, Z. Haiman, and A. Loeb. Gas pile-up, gap overflow and Type 1.5 migration in circumbinary discs: general theory. *MNRAS*, 427:2660–2679, December 2012.

[Komossa and Zensus, 2016] S. Komossa and J. A. Zensus. Compact object mergers: observations of supermassive binary black holes and stellar tidal disruption events. In Y. Meiron, S. Li, F.-K. Liu, and R. Spurzem, editors, *Star Clusters and Black Holes in Galaxies across Cosmic Time*, volume 312 of *IAU Symposium*, pages 13–25, February 2016.

[Komossa *et al.*, 2003] S. Komossa, V. Burwitz, G. Hasinger, P. Predehl, J. S. Kaastra, and Y. Ikebe. Discovery of a Binary Active Galactic Nucleus in the Ultraluminous Infrared Galaxy NGC 6240 Using Chandra. *ApJL*, 582:L15–L19, January 2003.

[Komossa, 2006] S. Komossa. Observational evidence for binary black holes and active double nuclei. *MmSAI*, 77:733, 2006.

[Komossa, 2012] S. Komossa. Recoiling Black Holes: Electromagnetic Signatures, Candidates, and Astrophysical Implications. *Advances in Astronomy*, 2012:364973, 2012.

[Kormendy and Ho, 2013] J. Kormendy and L. C. Ho. Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. *ARA&A*, 51:511–653, August 2013.

[Koss *et al.*, 2011] M. Koss, R. Mushotzky, E. Treister, S. Veilleux, R. Vasudevan, N. Miller, D. B. Sanders, K. Schawinski, and M. Trippe. Chandra Discovery of a Binary Active Galactic Nucleus in Mrk 739. *ApJL*, 735:L42, July 2011.

[Koss *et al.*, 2016] M. J. Koss, A. Glidden, M. Baloković, D. Stern, I. Lamperti, R. Assef, F. Bauer, D. Ballantyne, S. E. Boggs, W. W. Craig, D. Farrah, F. Fürst, P. Gandhi, N. Gehrels, C. J. Hailey, F. A. Harrison, C. Markwardt, A. Masini, C. Ricci, E. Treister, D. J. Walton, and W. W. Zhang. NuSTAR Resolves the First Dual AGN above 10 keV in SWIFT J2028.5+2543. *ApJL*, 824:L4, June 2016.

[Kozłowski *et al.*, 2010] S. Kozłowski, C. S. Kochanek, A. Udalski, Ł. Wyrzykowski, I. Soszyński, M. K. Szymański, M. Kubiak, G. Pietrzyński, O. Szewczyk, K. Ulaczyk, R. Poleski, and OGLE Collaboration. Quantifying Quasar Variability as Part of a General Approach to Classifying Continuously Varying Sources. *ApJ*, 708:927–945, January 2010.

[Kudryavtseva *et al.*, 2011] N. A. Kudryavtseva, S. Britzen, A. Witzel, E. Ros, M. Karouzos, M. F. Aller, H. D. Aller, H. Teräsranta, A. Eckart, and J. A. Zensus. A possible jet precession in the periodic quasar B0605-085. *A& A*, 526:A51, February 2011.

[Kulkarni and Loeb, 2016] G. Kulkarni and A. Loeb. Radio crickets: chirping jets from black hole binaries entering their gravitational wave inspiral. *MNRAS*, 456:3964–3971, March 2016.

[Kun *et al.*, 2014] E. Kun, K. É. Gabányi, M. Karouzos, S. Britzen, and L. Á. Gergely. A spinning supermassive black hole binary model consistent with VLBI observations of the S5 1928+738 jet. *MNRAS*, 445:1370–1382, December 2014.

[Kun *et al.*, 2015] E. Kun, S. Frey, K. É. Gabányi, S. Britzen, D. Cseh, and L. Á. Gergely. Constraining the parameters of the putative supermassive binary black hole in PG 1302-102 from its radio structure. *MNRAS*, 454:1290–1296, December 2015.

[Laher *et al.*, 2014] R. R. Laher, J. Surace, C. J. Grillmair, E. O. Ofek, D. Levitan, B. Sesar, J. C. van Eyken, N. M. Law, G. Helou, N. Hamam, F. J. Masci, S. Mattingly, E. Jackson, E. Hacopeans, W. Mi, S. Groom, H. Teplitz, V. Desai, D. Hale, R. Smith, R. Walters, R. Quimby, M. Kasliwal, A. Horesh, E. Bellm, T. Barlow, A. Waszczak, T. A. Prince, and S. R. Kulkarni. IPAC Image Processing and Data Archiving for the Palomar Transient Factory. *PASP*, 126:674–710, July 2014.

[Law *et al.*, 2009] N. M. Law, S. R. Kulkarni, R. G. Dekany, E. O. Ofek, R. M. Quimby, P. E. Nugent, J. Surace, C. C. Grillmair, J. S. Bloom, M. M. Kasliwal, L. Bildsten, T. Brown, S. B. Cenko, D. Ciardi, E. Croner, S. G. Djorgovski, J. van Eyken, A. V. Filippenko, D. B. Fox, A. Gal-Yam, D. Hale, N. Hamam, G. Helou, J. Henning, D. A. Howell, J. Jacobsen, R. Laher, S. Mattingly, D. McKenna, A. Pickles, D. Poznanski, G. Rahmer, A. Rau, W. Rosing, M. Shara, R. Smith, D. Starr, M. Sullivan, V. Velur, R. Walters, and J. Zolkower. The Palomar Transient Factory: System Overview, Performance, and First Results. *PASP*, 121:1395–1408, December 2009.

[Leighly *et al.*, 2016] K. M. Leighly, D. M. Terndrup, S. C. Gallagher, and A. B. Lucy. The Binary Black Hole Model for Mrk 231 Bites the Dust. *ApJ*, 829:4, September 2016.

[Li *et al.*, 2016] Y.-R. Li, J.-M. Wang, L. C. Ho, K.-X. Lu, J. Qiu, P. Du, C. Hu, Y.-K. Huang, Z.-X. Zhang, K. Wang, and J.-M. Bai. Spectroscopic Indication of a Centi-parsec

Supermassive Black Hole Binary in the Galactic Center of NGC 5548. *ApJ*, 822:4, May 2016.

[Li *et al.*, 2017] Y.-R. Li, J.-M. Wang, Z.-X. Zhang, K. Wang, Y.-K. Huang, K.-X. Lu, C. Hu, P. Du, L. C. Ho, J.-M. Bai, W.-H. Bian, and Y.-F. Yuan. A Possible ˜20-Year Periodicity in Long-term Variations of the Nearby Radio-Quiet Active Galactic Nucleus Ark 120. *ArXiv e-prints*, May 2017.

[Liu *et al.*, 2009] F. K. Liu, S. Li, and X. Chen. Interruption of Tidal-Disruption Flares by Supermassive Black Hole Binaries. *ApJL*, 706:L133–L137, November 2009.

[Liu *et al.*, 2010] X. Liu, Y. Shen, M. A. Strauss, and J. E. Greene. Type 2 Active Galactic Nuclei with Double-Peaked [O III] Lines: Narrow-Line Region Kinematics or Merging Supermassive Black Hole Pairs? *ApJ*, 708:427–434, January 2010.

[Liu *et al.*, 2011] X. Liu, Y. Shen, and M. A. Strauss. Cosmic Train Wreck by Massive Black Holes: Discovery of a Kiloparsec-scale Triple Active Galactic Nucleus. *ApJL*, 736:L7, July 2011.

[Liu *et al.*, 2013] X. Liu, F. Civano, Y. Shen, P. Green, J. E. Greene, and M. A. Strauss. Chandra X-Ray and Hubble Space Telescope Imaging of Optically Selected Kiloparsec-scale Binary Active Galactic Nuclei. I. Nature of the Nuclear Ionizing Sources. *ApJ*, 762:110, January 2013.

[Liu *et al.*, 2014] F. K. Liu, S. Li, and S. Komossa. A Milliparsec Supermassive Black Hole Binary Candidate in the Galaxy SDSS J120136.02+300305.5. *ApJ*, 786:103, May 2014.

[Liu *et al.*, 2015] T. Liu, S. Gezari, S. Heinis, E. A. Magnier, W. S. Burgett, K. Chambers, H. Flewelling, M. Huber, K. W. Hodapp, N. Kaiser, R.-P. Kudritzki, J. L. Tonry, R. J. Wainscoat, and C. Waters. A Periodically Varying Luminous Quasar at z = 2 from the Pan-STARRS1 Medium Deep Survey: A Candidate Supermassive Black Hole Binary in the Gravitational Wave-driven Regime. *ApJL*, 803:L16, April 2015.

[Liu *et al.*, 2016a] J. Liu, M. Eracleous, and J. P. Halpern. A Radial Velocity Test for Supermassive Black Hole Binaries as an Explanation for Broad, Double-peaked Emission Lines in Active Galactic Nuclei. *ApJ*, 817:42, January 2016.

[Liu *et al.*, 2016b] T. Liu, S. Gezari, W. Burgett, K. Chambers, P. Draper, K. Hodapp, M. Huber, R.-P. Kudritzki, E. Magnier, N. Metcalfe, J. Tonry, R. Wainscoat, and C. Waters. A Systematic Search for Periodically Varying Quasars in Pan-STARRS1: An Extended Baseline Test in Medium Deep Survey Field MD09. *ApJ*, 833:6, December 2016.

[Lobanov and Roland, 2005] A. P. Lobanov and J. Roland. A supermassive binary black hole in the quasar 3C 345. *A& A*, 431:831–846, March 2005.

[Lomb, 1976] N. R. Lomb. Least-squares frequency analysis of unequally spaced data. *Ap&SS*, 39:447–462, February 1976.

[Lousto and Zlochower, 2011] C. O. Lousto and Y. Zlochower. Hangup Kicks: Still Larger Recoils by Partial Spin-Orbit Alignment of Black-Hole Binaries. *Physical Review Letters*, 107(23):231102, December 2011.

[Lynden-Bell, 1969] D. Lynden-Bell. Galactic Nuclei as Collapsed Old Quasars. *Nature*, 223:690–694, August 1969.

[MacFadyen and Milosavljević, 2008] A. I. MacFadyen and M. Milosavljević. An Eccentric Circumbinary Accretion Disk and the Detection of Binary Massive Black Holes. *ApJ*, 672:83–93, January 2008.

[MacLeod *et al.*, 2010] C. L. MacLeod, Ž. Ivezić, C. S. Kochanek, S. Kozłowski, B. Kelly, E. Bullock, A. Kimball, B. Sesar, D. Westman, K. Brooks, R. Gibson, A. C. Becker, and W. H. de Vries. Modeling the Time Variability of SDSS Stripe 82 Quasars as a Damped Random Walk. *ApJ*, 721:1014–1033, October 2010.

[Madau *et al.*, 1996] P. Madau, H. C. Ferguson, M. E. Dickinson, M. Giavalisco, C. C. Steidel, and A. Fruchter. High-redshift galaxies in the Hubble Deep Field: colour selection and star formation history to z˜4. *MNRAS*, 283:1388–1404, December 1996.

[Mahabal *et al.*, 2011] A. A. Mahabal, S. G. Djorgovski, A. J. Drake, C. Donalek, M. J. Graham, R. D. Williams, Y. Chen, B. Moghaddam, M. Turmon, E. Beshore, and S. Larson. Discovery, classification, and scientific exploration of transient events from the Catalina Real-time Transient Survey. *Bulletin of the Astronomical Society of India*, 39:387–408, September 2011.

[Manchester and IPTA, 2013] R. N. Manchester and IPTA. The International Pulsar Timing Array. *Classical and Quantum Gravity*, 30(22):224010, November 2013.

[Manchester, 2008] R. N. Manchester. The Parkes Pulsar Timing Array Project. In C. Bassa, Z. Wang, A. Cumming, and V. M. Kaspi, editors, *40 Years of Pulsars: Millisecond Pulsars, Magnetars and More*, volume 983 of *American Institute of Physics Conference Series*, pages 584–592, February 2008.

[Martini, 2004] P. Martini. QSO Lifetimes. *Coevolution of Black Holes and Galaxies*, page 169, 2004.

[Mayer *et al.*, 2007] L. Mayer, S. Kazantzidis, P. Madau, M. Colpi, T. Quinn, and J. Wadsley. Rapid Formation of Supermassive Black Hole Binaries in Galaxy Mergers with Gas. *Science*, 316:1874, June 2007.

[McGurk *et al.*, 2011] R. C. McGurk, C. E. Max, D. J. Rosario, G. A. Shields, K. L. Smith, and S. A. Wright. Spatially Resolved Spectroscopy of SDSS J0952+2552: A Confirmed Dual Active Galactic Nucleus. *ApJL*, 738:L2, September 2011.

[McGurk *et al.*, 2015] R. C. McGurk, C. E. Max, A. M. Medling, G. A. Shields, and J. M. Comerford. Spatially Resolved Imaging and Spectroscopy of Candidate Dual Active Galactic Nuclei. *ApJ*, 811:14, September 2015.

[Milosavljević and Merritt, 2001] M. Milosavljević and D. Merritt. Formation of Galactic Nuclei. *ApJ*, 563:34–62, December 2001.

[Milosavljević and Phinney, 2005] M. Milosavljević and E. S. Phinney. The Afterglow of Massive Black Hole Coalescence. *ApJL*, 622:L93–L96, April 2005.

[Mohan *et al.*, 2016] P. Mohan, T. An, S. Frey, A. Mangalam, K. É. Gabányi, and E. Kun. Parsec-scale jet properties of the quasar PG 1302-102. *MNRAS*, 463:1812–1821, December 2016.

[Mushotzky *et al.*, 2011] R. F. Mushotzky, R. Edelson, W. Baumgartner, and P. Gandhi. Kepler Observations of Rapid Optical Variability in Active Galactic Nuclei. *ApJL*, 743:L12, December 2011.

[Myers *et al.*, 2008] A. D. Myers, G. T. Richards, R. J. Brunner, D. P. Schneider, N. E. Strand, P. B. Hall, J. A. Blomquist, and D. G. York. Quasar Clustering at 25 h^{-1} kpc from a Complete Sample of Binaries. *ApJ*, 678:635–646, May 2008.

[Noble *et al.*, 2012] S. C. Noble, B. C. Mundim, H. Nakano, J. H. Krolik, M. Campanelli, Y. Zlochower, and N. Yunes. Circumbinary Magnetohydrodynamic Accretion into Inspiraling Binary Black Holes. *ApJ*, 755:51, August 2012.

[Ofek *et al.*, 2012a] E. O. Ofek, R. Laher, N. Law, J. Surace, D. Levitan, B. Sesar, A. Horesh, D. Poznanski, J. C. van Eyken, S. R. Kulkarni, P. Nugent, J. Zolkower, R. Walters, M. Sullivan, M. Agüeros, L. Bildsten, J. Bloom, S. B. Cenko, A. Gal-Yam, C. Grillmair, G. Helou, M. M. Kasliwal, and R. Quimby. The Palomar Transient Factory Photometric Calibration. *PASP*, 124:62–73, January 2012.

[Ofek *et al.*, 2012b] E. O. Ofek, R. Laher, J. Surace, D. Levitan, B. Sesar, A. Horesh, N. Law, J. C. van Eyken, S. R. Kulkarni, T. A. Prince, P. Nugent, M. Sullivan, O. Yaron, A. Pickles, M. Agüeros, I. Arcavi, L. Bildsten, J. Bloom, S. B. Cenko, A. Gal-Yam, C. Grillmair, G. Helou, M. M. Kasliwal, D. Poznanski, and R. Quimby. The Palomar Transient Factory photometric catalog 1.0. *PASP*, 124:854–860, August 2012.

[Owen *et al.*, 1985] F. N. Owen, C. P. O'Dea, M. Inoue, and J. A. Eilek. VLA observations of the multiple jet galaxy 3C 75. *ApJL*, 294:L85–L88, July 1985.

[Palanque-Delabrouille *et al.*, 2011] N. Palanque-Delabrouille, C. Yeche, A. D. Myers, P. Petitjean, N. P. Ross, E. Sheldon, E. Aubourg, T. Delubac, J.-M. Le Goff, I. Pâris, J. Rich, K. S. Dawson, D. P. Schneider, and B. A. Weaver. Variability selected high-redshift quasars on SDSS Stripe 82. *A& A*, 530:A122, June 2011.

[Peebles, 1980] P. J. E. Peebles. *The large-scale structure of the universe*. 1980.

[Peeples *et al.*, 2016] M. S. Peeples, J. Tumlinson, A. Fox, A. Aloisi, T. R. Ayres, C. Danforth, S. W. Fleming, E. B. Jenkins, R. I. Jedrzejewski, B. A. Keeney, and C. M. Oliveira. The Hubble Spectroscopic Legacy Archive. In *American Astronomical Society Meeting Abstracts*, volume 227 of *American Astronomical Society Meeting Abstracts*, page 444.01, January 2016.

[Peters, 1964] P. C. Peters. Gravitational Radiation and the Motion of Two Point Masses. *Physical Review*, 136:1224–1232, November 1964.

[Pojmanski, 1997] G. Pojmanski. The All Sky Automated Survey. *AcA*, 47:467–481, October 1997.

[Popović, 2012] L. Č. Popović. Super-massive binary black holes and emission lines in active galactic nuclei. *NAR*, 56:74–91, February 2012.

[Preto *et al.*, 2011] M. Preto, I. Berentzen, P. Berczik, and R. Spurzem. Fast Coalescence of Massive Black Hole Binaries from Mergers of Galactic Nuclei: Implications for Low-frequency Gravitational-wave Astrophysics. *ApJL*, 732:L26, May 2011.

[Price-Whelan *et al.*, 2014] A. M. Price-Whelan, M. A. Agüeros, A. P. Fournier, R. Street, E. O. Ofek, K. R. Covey, D. Levitan, R. R. Laher, B. Sesar, and J. Surace. Statistical Searches for Microlensing Events in Large, Non-uniformly Sampled Time-Domain Surveys: A Test Using Palomar Transient Factory Data. *ApJ*, 781:35, January 2014.

[Pursimo *et al.*, 2000] T. Pursimo, L. O. Takalo, A. Sillanpää, M. Kidger, H. J. Lehto, J. Heidt, P. A. Charles, H. Aller, M. Aller, V. Beckmann, E. Benítez, H. Bock, P. Boltwood, U. Borgeest, J. A. de Diego, G. De Francesco, M. Dietrich, D. Dultzin-Hacyan, Y. Efimov, M. Fiorucci, G. Ghisellini, N. González-Pérez, M. Hanski, P. Heinämäki, R. K. Honeycutt, P. Hughes, K. Karlamaa, S. Katajainen, L. B. G. Knee, O. M. Kurtanidze, M. Kümmel, D. Kühl, M. Lainela, L. Lanteri, J. V. Linde, A. Lähteenmäki, M. Maesano, T. Mahoney, S. Marchenko, A. Marscher, E. Massaro, F. Montagni, R. Nesci, M. Nikolashvili, K. Nilsson, P. Nurmi, H. Pietilä, G. Poyner, C. M. Raiteri, R. Rekola,

G. M. Richter, A. Riehokainen, J. W. Robertson, J.-M. Rodríguez-Espinoza, A. Sadun, N. Shakhovskoy, K. J. Schramm, T. Schramm, G. Sobrito, P. Teerikorpi, H. Teräsranta, M. Tornikoski, G. Tosti, G. W. Turner, E. Valtaoja, M. Valtonen, M. Villata, S. J. Wagner, J. Webb, W. Weneit, and S. Wiren. Intensive monitoring of OJ 287. *A& AS*, 146:141–155, October 2000.

[Rafikov, 2013] R. R. Rafikov. Structure and Evolution of Circumbinary Disks around Supermassive Black Hole Binaries. *ApJ*, 774:144, September 2013.

[Rafikov, 2016] R. R. Rafikov. Accretion and Orbital Inspiral in Gas-assisted Supermassive Black Hole Binary Mergers. *ApJ*, 827:111, August 2016.

[Rahmer *et al.*, 2008] G. Rahmer, R. Smith, V. Velur, D. Hale, N. Law, K. Bui, H. Petrie, and R. Dekany. The 12K x 8K CCD mosaic camera for the Palomar Transient Factory. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7014 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 4, July 2008.

[Raiteri *et al.*, 2001] C. M. Raiteri, M. Villata, H. D. Aller, M. F. Aller, J. Heidt, O. M. Kurtanidze, L. Lanteri, M. Maesano, E. Massaro, F. Montagni, R. Nesci, K. Nilsson, M. G. Nikolashvili, P. Nurmi, L. Ostorero, T. Pursimo, R. Rekola, A. Sillanpää, L. O. Takalo, H. Teräsranta, G. Tosti, T. J. Balonek, M. Feldt, A. Heines, C. Heisler, J. Hu, M. Kidger, J. R. Mattox, E. J. McGrath, A. Pati, R. Robb, A. C. Sadun, P. Shastri, S. J. Wagner, J. Wei, and X. Wu. Optical and radio variability of the BL Lacertae object ¡ASTROBJ¿AO 0235+16¡/ASTROBJ¿: A possible 5-6 year periodicity. *A& A*, 377:396–412, October 2001.

[Rau *et al.*, 2009] A. Rau, S. R. Kulkarni, N. M. Law, J. S. Bloom, D. Ciardi, G. S. Djorgovski, D. B. Fox, A. Gal-Yam, C. C. Grillmair, M. M. Kasliwal, P. E. Nugent, E. O. Ofek, R. M. Quimby, W. T. Reach, M. Shara, L. Bildsten, S. B. Cenko, A. J. Drake, A. V. Filippenko, D. J. Helfand, G. Helou, D. A. Howell, D. Poznanski, and M. Sullivan. Exploring the Optical Transient Sky with the Palomar Transient Factory. *PASP*, 121:1334–1351, December 2009.

[Rieger, 2004] F. M. Rieger. On the Geometrical Origin of Periodicity in Blazar-type Sources. *ApJL*, 615:L5–L8, November 2004.

[Roberts *et al.*, 1987] D. H. Roberts, J. Lehar, and J. W. Dreher. Time Series Analysis with Clean - Part One - Derivation of a Spectrum. *AJ*, 93:968, April 1987.

[Robertson *et al.*, 2006] B. Robertson, J. S. Bullock, T. J. Cox, T. Di Matteo, L. Hernquist, V. Springel, and N. Yoshida. A Merger-driven Scenario for Cosmological Disk Galaxy Formation. *ApJ*, 645:986–1000, July 2006.

[Rodriguez *et al.*, 2006] C. Rodriguez, G. B. Taylor, R. T. Zavala, A. B. Peck, L. K. Pollack, and R. W. Romani. A Compact Supermassive Binary Black Hole System. *ApJ*, 646:49–60, July 2006.

[Roedig *et al.*, 2012] C. Roedig, A. Sesana, M. Dotti, J. Cuadra, P. Amaro-Seoane, and F. Haardt. Evolution of binary black holes in self gravitating discs. Dissecting the torques. *A& A*, 545:A127, September 2012.

[Romero *et al.*, 2000] G. E. Romero, L. Chajet, Z. Abraham, and J. H. Fan. Beaming and precession in the inner jet of 3C 273 — II. The central engine. *A& A*, 360:57–64, August 2000.

[Romero *et al.*, 2003] G. E. Romero, J.-H. Fan, and S. E. Nuza. The Binary Black Hole Scenario for the BL Lacertae Object AO˜0235+16. *ChJAA*, 3:513–525, December 2003.

[Roos *et al.*, 1993] N. Roos, J. S. Kaastra, and C. A. Hummel. A massive binary black hole in 1928 + 738? *ApJ*, 409:130–133, May 1993.

[Runnoe *et al.*, 2017] J. C. Runnoe, M. Eracleous, A. Pennell, G. Mathes, T. Boroson, S. Sigursson, T. Bogdanović, J. P. Halpern, J. Liu, and S. Brown. A large systematic search for close supermassive binary and rapidly recoiling black holes - III. Radial velocity variations. *MNRAS*, 468:1683–1702, June 2017.

[Ryan and MacFadyen, 2017] G. Ryan and A. MacFadyen. Minidisks in Binary Black Hole Accretion. *ApJ*, 835:199, February 2017.

[Salpeter, 1964] E. E. Salpeter. Accretion of Interstellar Matter by Massive Objects. *ApJ*, 140:796–800, August 1964.

[Sanders and Mirabel, 1996] D. B. Sanders and I. F. Mirabel. Luminous Infrared Galaxies. *ARA&A*, 34:749, 1996.

[Scargle, 1982] J. D. Scargle. Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *ApJ*, 263:835–853, December 1982.

[Schawinski *et al.*, 2011] K. Schawinski, M. Urry, E. Treister, B. Simmons, P. Natarajan, and E. Glikman. Evidence for Three Accreting Black Holes in a Galaxy at z ˜ 1.35: A Snapshot of Recently Formed Black Hole Seeds? *ApJL*, 743:L37, December 2011.

[Schmidt, 1963] M. Schmidt. 3C 273 : A Star-Like Object with Large Red-Shift. *Nature*, 197:1040, March 1963.

[Sesana *et al.*, 2017] A. Sesana, Z. Haiman, B. Kocsis, and L. Z. Kelley. Testing the binary hypothesis: pulsar timing constraints on supermassive black hole binary candidates. *ArXiv e-prints*, March 2017.

[Sesana, 2015] A. Sesana. Pulsar Timing Arrays and the Challenge of Massive Black Hole Binary Astrophysics. *Astrophysics and Space Science Proceedings*, 40:147, 2015.

[Sesar *et al.*, 2011] B. Sesar, J. S. Stuart, Ž. Ivezić, D. P. Morgan, A. C. Becker, and P. Woźniak. Exploring the Variable Sky with LINEAR. I. Photometric Recalibration with the Sloan Digital Sky Survey. *AJ*, 142:190, December 2011.

[Shannon *et al.*, 2015] R. M. Shannon, V. Ravi, L. T. Lentati, P. D. Lasky, G. Hobbs, M. Kerr, R. N. Manchester, W. A. Coles, Y. Levin, M. Bailes, N. D. R. Bhat, S. Burke-Spolaor, S. Dai, M. J. Keith, S. Osłowski, D. J. Reardon, W. van Straten, L. Toomey, J.-B. Wang, L. Wen, J. S. B. Wyithe, and X.-J. Zhu. Gravitational waves from binary supermassive black holes missing in pulsar observations. *Science*, 349:1522–1525, September 2015.

[Shen and Loeb, 2010] Y. Shen and A. Loeb. Identifying Supermassive Black Hole Binaries with Broad Emission Line Diagnosis. *ApJ*, 725:249–260, December 2010.

[Shen *et al.*, 2008] Y. Shen, J. E. Greene, M. A. Strauss, G. T. Richards, and D. P. Schneider. Biases in Virial Black Hole Masses: An SDSS Perspective. *ApJ*, 680:169–190, June 2008.

[Shen *et al.*, 2013] Y. Shen, X. Liu, A. Loeb, and S. Tremaine. Constraining Sub-parsec Binary Supermassive Black Holes in Quasars with Multi-epoch Spectroscopy. I. The General Quasar Population. *ApJ*, 775:49, September 2013.

[Shi and Krolik, 2015] J.-M. Shi and J. H. Krolik. Three-dimensional MHD Simulation of Circumbinary Accretion Disks. II. Net Accretion Rate. *ApJ*, 807:131, July 2015.

[Shi *et al.*, 2012] J.-M. Shi, J. H. Krolik, S. H. Lubow, and J. F. Hawley. Three-dimensional Magnetohydrodynamic Simulations of Circumbinary Accretion Disks: Disk Structures and Angular Momentum Transport. *ApJ*, 749:118, April 2012.

[Shields *et al.*, 2012] G. A. Shields, D. J. Rosario, V. Junkkarinen, S. C. Chapman, E. W. Bonning, and T. Chiba. LBQS 0103-2753: A Binary Quasar in a Major Merger. *ApJ*, 744:151, January 2012.

[Simm *et al.*, 2016] T. Simm, M. Salvato, R. Saglia, G. Ponti, G. Lanzuisi, B. Trakhtenbrot, K. Nandra, and R. Bender. Pan-STARRS1 variability of XMM-COSMOS AGN. II. Physical correlations and power spectrum analysis. *A& A*, 585:A129, January 2016.

[Smith *et al.*, 2010] K. L. Smith, G. A. Shields, E. W. Bonning, C. C. McMullen, D. J. Rosario, and S. Salviander. A Search for Binary Active Galactic Nuclei: Double-peaked [O III] AGNs in the Sloan Digital Sky Survey. *ApJ*, 716:866–877, June 2010.

[Springel *et al.*, 2005] V. Springel, T. Di Matteo, and L. Hernquist. Modelling feedback from stars and black holes in galaxy mergers. *MNRAS*, 361:776–794, August 2005.

[Tang *et al.*, 2017] Y. Tang, A. MacFadyen, and Z. Haiman. On the orbital evolution of supermassive black hole binaries with circumbinary accretion discs. *ArXiv e-prints*, March 2017.

[Taniguchi and Wada, 1996] Y. Taniguchi and K. Wada. The Nuclear Starburst Driven by a Supermassive Black Hole Binary. *ApJ*, 469:581, October 1996.

[Timmer and Koenig, 1995] J. Timmer and M. Koenig. On generating power law noise. *A& A*, 300:707, August 1995.

[Tsalmantza *et al.*, 2011] P. Tsalmantza, R. Decarli, M. Dotti, and D. W. Hogg. A Systematic Search for Massive Black Hole Binaries in the Sloan Digital Sky Survey Spectroscopic Sample. *ApJ*, 738:20, September 2011.

[Uttley *et al.*, 2002] P. Uttley, I. M. McHardy, and I. E. Papadakis. Measuring the broadband power spectra of active galactic nuclei with RXTE. *MNRAS*, 332:231–250, May 2002.

[Valtonen and Wiik, 2012] M. J. Valtonen and K. Wiik. Optical polarization angle and VLBI jet direction in the binary black hole model of OJ287. *MNRAS*, 421:1861–1867, April 2012.

[Valtonen *et al.*, 2012] M. J. Valtonen, S. Ciprini, and H. J. Lehto. On the masses of OJ287 black holes. *MNRAS*, 427:77–83, November 2012.

[Vanderplas *et al.*, 2012] J.T. Vanderplas, A.J. Connolly, Ž. Ivezić, and A. Gray. Introduction to astroml: Machine learning for astrophysics. In *Conference on Intelligent Data Understanding (CIDU)*, pages 47 –54, oct. 2012.

[Vasiliev *et al.*, 2015] E. Vasiliev, F. Antonini, and D. Merritt. The Final-parsec Problem in the Collisionless Limit. *ApJ*, 810:49, September 2015.

[Vaughan *et al.*, 2016] S. Vaughan, P. Uttley, A. G. Markowitz, D. Huppenkothen, M. J. Middleton, W. N. Alston, J. D. Scargle, and W. M. Farr. False periodicities in quasar time-domain surveys. *MNRAS*, 461:3145–3152, September 2016.

[Veilleux *et al.*, 2016] S. Veilleux, M. Meléndez, T. M. Tripp, F. Hamann, and D. S. N. Rupke. The Complete Ultraviolet Spectrum of the Archetypal "Wind-dominated" Quasar Mrk 231: Absorption and Emission from a High-speed Dusty Nuclear Outflow. *ApJ*, 825:42, July 2016.

[Villata and Raiteri, 1999] M. Villata and C. M. Raiteri. Helical jets in blazars. I. The case of MKN 501. *A& A*, 347:30–36, July 1999.

[Volonteri *et al.*, 2003] M. Volonteri, F. Haardt, and P. Madau. The Assembly and Merging History of Supermassive Black Holes in Hierarchical Models of Galaxy Formation. *ApJ*, 582:559–573, January 2003.

[Wang *et al.*, 2009] J.-M. Wang, Y.-M. Chen, C. Hu, W.-M. Mao, S. Zhang, and W.-H. Bian. Active Galactic Nuclei with Double-Peaked Narrow Lines: Are they Dual Active Galactic Nuclei? *ApJL*, 705:L76–L80, November 2009.

[Wang *et al.*, 2014] L. Wang, P. Berczik, R. Spurzem, and M. B. N. Kouwenhoven. The Link between Ejected Stars, Hardening and Eccentricity Growth of Super Massive Black Holes in Galactic Nuclei. *ApJ*, 780:164, January 2014.

[Wang *et al.*, 2017] L. Wang, J. E. Greene, W. Ju, R. R. Rafikov, J. J. Ruan, and D. P. Schneider. Searching for Binary Supermassive Black Holes via Variable Broad Emission Line Shifts: Low Binary Fraction. *ApJ*, 834:129, January 2017.

[Xie *et al.*, 2008] G. Z. Xie, T. F. Yi, H. Z. Li, S. B. Zhou, and L. E. Chen. Periodicity Analysis of the Radio Curve of PKS 1510-089 and Implications for its Central Structure. *AJ*, 135:2212–2216, June 2008.

[Yan *et al.*, 2015] C.-S. Yan, Y. Lu, X. Dai, and Q. Yu. A Probable Milli-parsec Supermassive Binary Black Hole in the Nearest Quasar Mrk 231. *ApJ*, 809:117, August 2015.

[Yunes *et al.*, 2011] N. Yunes, B. Kocsis, A. Loeb, and Z. Haiman. Imprint of Accretion Disk-Induced Migration on Gravitational Waves from Extreme Mass Ratio Inspirals. *Physical Review Letters*, 107(17):171103, October 2011.

[Zechmeister and Kürster, 2009] M. Zechmeister and M. Kürster. The generalised Lomb-Scargle periodogram. A new formalism for the floating-mean and Keplerian periodograms. *A& A*, 496:577–584, March 2009.

[Zheng *et al.*, 2016] Z.-Y. Zheng, N. R. Butler, Y. Shen, L. Jiang, J.-X. Wang, X. Chen, and J. Cuadra. SDSS J0159+0105: A Radio-Quiet Quasar with a Centi-Parsec Supermassive Black Hole Binary Candidate. *ApJ*, 827:56, August 2016.

[Zu *et al.*, 2013] Y. Zu, C. S. Kochanek, S. Kozłowski, and A. Udalski. Is Quasar Optical Variability a Damped Random Walk? *ApJ*, 765:106, March 2013.