

Searching for unresolved massive black hole pairs through active galactic nucleus photometric variability

Lorenzo Bertassi^{1,2,3,*}, Maria Charisi^{5,6}, Fabio Rigamonti^{2,3,4}, Stefano Covino^{2,4}, and Massimo Dotti^{1,2,3}

¹ Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy

² INAF – Osservatorio Astronomico di Brera, Via Brera 20, I-20121 Milano, Italy

³ INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy

⁴ Como Lake centre for AstroPhysics (CLAP), DiSAT, Università dell’Insubria, Via Valleggio 11, 22100 Como, Italy

⁵ Department of Physics and Astronomy, Washington State University, Pullman, WA 99163, USA

⁶ Institute of Astrophysics, FORTH, GR-71110 Heraklion, Greece

Received 25 August 2025 / Accepted 20 March 2026

ABSTRACT

Since the discovery of active galactic nuclei (AGN), their light curves have proved to be intrinsically variable. In the optical/UV band, this variability is consistent with correlated or red noise and is particularly well described by the damped random walk (DRW) model. In this work, we evaluate the feasibility of a new method for identifying spatially unresolved couples of active galactic nuclei through a fully Bayesian time-domain analysis of the observed light curves. More specifically, we check whether observed light curves are better described by a single DRW, which we interpret as emitted by a single massive black hole, or a pair of independent damped random walks, generated by a pair of massive black holes. We test the method on mock light curves associated with single massive black holes and pairs generated with different cadences and lengths of observational campaigns. We constrained the occurrence of false positives, that is, the percentage of single MBH light curves that show substantial evidence in favour of the unresolved MBH pair scenario, finding a fraction of 0.2% and 0.59% in the even and uneven sampling scenarios. We discuss how well the method recovers the model parameters, showing that about 51% and 7% of the simulated light curves have all the recovered parameters within 20% of their true values in our best scenario of evenly sampled light curves for the single massive black hole and massive black hole pair scenarios, respectively. We finally study the region of the parameter space in which the detection of a massive black hole pair is possible, finding that such objects can be correctly identified if the timescales of the process describing the noise are very different, with a ratio smaller than ~ 0.2 , and the variability amplitudes are similar, with their ratio bigger than ~ 0.2 . When limiting to such a region of the parameter space, the fraction of pairs with all the recovered parameters within 20% of the injected values increases up to about 14% and 8% for evenly and unevenly sampled light curves, respectively.

Key words. methods: data analysis – methods: statistical – techniques: photometric – galaxies: active – galaxies: nuclei – quasars: supermassive black holes

1. Introduction

In the hierarchical model of galaxy formation, smaller galaxies merge to form the larger galaxies (White & Rees 1978). As nearly all massive galaxies are expected to host a massive black hole (MBH) at their centre (Kormendy & Gebhardt 2001), massive black hole pairs (MBHPs) and, at even smaller separations, massive black hole binaries (MBHBs) are predicted to form in the aftermath of the merger (Begelman et al. 1980).

In our current view, two MBHs, still unbound from one another and residing at the centres of the merging galaxies, are expected to drift towards the centre of the newly formed galaxy because of dynamical friction. The dynamical friction efficiency at large MBH separations (0.001–10 kpc) is still poorly constrained (see Dotti et al. 2012), due to both the possible interaction between the MBHs and the galaxy substructures such as spiral arms, bars, and stellar or gaseous clumps (e.g. Fiacconi et al. 2013; del Valle et al. 2015; Souza Lima et al. 2020; Bortolas et al. 2020, 2022), and tidal effects. More specifically, tides can significantly increase the pairing timescale

by decreasing the mass of the progenitor galaxy cores¹, (e.g. Governato et al. 1994; Pfister et al. 2019; Varisco et al. 2024).

Past this stage, the evolution is expected to be driven by three-body scatterings with stars (Begelman et al. 1980; Mikkola & Valtonen 1992) and/or the interaction with the gas surrounding the two MBHs (Armitage & Natarajan 2002; Haiman et al. 2009). If the above-mentioned processes manage to efficiently remove orbital energy and angular momentum from the two MBHs, the resulting binary, i.e. the system where the two MBHs are gravitationally bound to each other, will coalesce due to gravitational wave emission (see Thorne & Braginskii 1976), thus becoming one of the most promising source populations to be detected by space-borne interferometers, for example the Laser Interferometer Space Antenna (LISA; Amaro-Seoane et al. 2023), and pulsar timing arrays (see Verbiest et al. 2016). Recently, evidence for a gravitational wave background possibly generated by inspiralling MBHBs has been found in pulsar timing array data (see Agazie et al. 2023; Reardon et al. 2023; Xu et al. 2023; EPTA Collaboration and InPTA Collaboration 2024; Miles et al. 2025). Even though the background is likely generated by

¹ The efficiency of tidal stripping depends on the possible deepening of the potential well of the core, as a consequence of merger driven gas inflows (e.g. Callegari et al. 2009).

* Corresponding author: l.bertassi@campus.unimib.it

MBHBs, the uncertainties in stellar- and gas-driven evolution result in a significant scatter in the predicted gravitational wave signal (e.g. Varisco et al. 2021; Franchini et al. 2023, and references therein). Stringent EM constraints on MBHPs and MBHBs are needed for a comparison between theory and observation.

To date, there are only ~ 100 confirmed observations of MBHPs resolved as dual AGNs (see De Rosa et al. 2019 for a review, Zhang et al. 2021 for a more recent collection of ~ 10 kpc MBHPs, Hwang et al. 2020; Mannucci et al. 2022; Scialpi et al. 2024; Chen et al. 2023; Wu et al. 2024; Chen et al. 2025; Schwartzman et al. 2024 for recent searches of kiloparsec-scale MBHPs and Trindade Falcão et al. 2024 for a candidate with a ~ 100 pc separation). A secure identification of spatially unresolved MBHPs and binaries is challenging. Current searches for unresolved signatures (see D’Orazio & Charisi 2023, for a review) have focused on close MBHBs where quasi-periodic modulations of the continuum (Valtonen et al. 2008; Ackermann et al. 2015; Graham et al. 2015; D’Orazio et al. 2015; Li et al. 2016; Charisi et al. 2016; Sandrinelli et al. 2016, 2018; Severgnini et al. 2018; Li et al. 2019; Liu et al. 2019; Hu et al. 2020; Chen et al. 2020; Zhu & Thrane 2020; Cocchiararo et al. 2024; Rigamonti et al. 2025a; Bertassi et al. 2025a) and of the broad emission line (BEL) profile (Shen & Loeb 2010; Tsalmantza et al. 2011; Eracleous et al. 2012; Decarli et al. 2013; Runnoe et al. 2017, 2025; Sottocorno et al. 2026; Rigamonti et al. 2025b; Bertassi et al. 2025b) are expected to occur on timescales of months to centuries.

At large enough separations, the search for photometric and spectroscopic features becomes either impossible (as the underlying assumptions of the models break) or unfeasible (as the required length of the observational campaigns to observe such modulations becomes too long). One exception to this is a test originally proposed by Gaskell (1988) and recently quantitatively detailed by Dotti et al. (2023), which assumes that the two MBHs are far enough from each other for their emission (both in the continuum and in the BELs) to be uncorrelated on observational timescales. In this case, different spectral regions of BELs react at two independent continua. This test identifies pairs and binaries with orbital periods of decades to centuries using reverberation-mapping campaigns lasting $\lesssim 1$ yr (Dotti et al. 2023), but fails when the separation is so large (e.g. ~ 0.5 pc for a binary of $\sim 10^6 M_{\odot}$) that the two independent broad-line regions have a too-small relative shift, due to the small orbital velocity of the binary ($\lesssim 100$ km/s, see Dotti et al. 2023).

In this work, we present a new photometric test for couples of active MBHs (either MBHBs or MBHPs) too widely separated to be identified through continuum periodicity (e.g., with separations ≥ 0.001 pc for a binary of $\sim 10^6 M_{\odot}$), but too close to be spatially resolved. For example, assuming an angular resolution of ≈ 0.65 arcsec in the r band for the *Vera Rubin* observatory (Ivezic et al. 2019), the maximum projected separation between the two MBHs for the test to succeed is ≈ 5 kpc at $z \approx 1$ and ≈ 1.2 kpc at $z \approx 0.1$. In the following, we refer to such systems as unresolved black hole duos (UBHDs).

The test leverages the intrinsic variability of AGN light curves (see Paolillo & Papadakis 2025, for a review) well modelled with a damped random walk (DRW, see Kelly et al. 2009; Kozłowski et al. 2010; MacLeod et al. 2010), a stochastic process with red-noise², i.e. a frequency-dependent noise with

power increasing towards the lower frequencies and becoming white noise, i.e. frequency independent at the lowest frequencies (see Section 2).

At the separations considered in this study, the short-timescale variability properties of the two MBHs were uncorrelated (as in the spectroscopic test proposed by Gaskell 1988 and Dotti et al. 2023). With simulations of UBHDs consisting of two DRW components, we tested whether these systems can be distinguished from the typical AGN behaviour described by a single DRW. For this, we employed a Bayesian model comparison.

By allowing the identification of unresolved UBHDs (associated with long residence timescales, which are expected to be abundant, see e.g. Haiman et al. 2009), the test has the potential to reveal large samples of UBHDs compared to the close MBHBs already searched for in time-domain surveys. These surveys are current, for example the *Zwicky* Transient Facility (ZTF; Bellm et al. 2018) and the Catalina Real-Time Transient Survey (CTRS; Djorgovski et al. 2011), and upcoming, for example the *Vera Rubin* Observatory’s ten-year Legacy Survey of Space and Time (LSST; Ivezić et al. 2019) and the *Roman* Space Telescope’s High Latitude Time Domain Survey (HLTDS; Haiman et al. 2023).

In Section 2, we describe the properties of the DRW model. We also describe how mock light curves are generated, and discuss the methods with which the parameters are constrained and the model comparison is performed. In Section 3, we discuss the fraction of false positives expected, we quantify how well the injected parameters are retrieved, and identify the regions in the parameter space in which the test correctly identifies an UBHD. Our main conclusions are presented in Section 4. Appendix A discusses the properties of light-curve periodograms, and highlights the limitations that prompted the design of the analysis discussed in the main text. A less statistically solid test searching for UBHDs by fitting the power spectral density (PSD) obtained through the Lomb-Scargle periodogram (see VanderPlas 2018, and references therein) shape directly is discussed in Appendix B.

2. Methodology

For this work, we focused on UBHDs, i.e. MBHs separated enough to have independent emission variability, so that the PSD of the resulting time series is given by the sum of the power spectra of the two processes. Furthermore, we assumed that the only source of variability for each MBH is the intrinsic one produced by a DRW. We generated mock DRW light curves for each component MBH using *celerite*³ (see Foreman-Mackey et al. 2017) and created mock UBHD light curves by summing the realisations of the flux of two different MBHs (as detailed in Section 2.2). First, we considered evenly sampled time series, and then generalised the work to unevenly sampled time series, simulating the expected cadence of LSST. Next, we estimated the parameters and the evidence of a given model, either single MBH or UBHD, sampling the likelihood with a nested sampling algorithm (see Skilling 2006) implemented in *RayNest*⁴ (Veitch et al. 2024). Finally, we selected the best model by comparing the statistical evidence for each model.

³ The documentation about the package can be found at the following link: <https://celerite.readthedocs.io/en/stable/python/install/>

⁴ The documentation about this package can be found at this link: <https://github.com/wdpozso/raynest>

² Notably, this red noise presents a major limitation in the photometric periodicity searches for close MBHBs, leading to false positives (Vaughan et al. 2016).

2.1. Noise models

The behaviour of a quantity evolving under a DRW is described by the solution to the following stochastic differential equation (see Kelly et al. 2009):

$$dX(t) = -\frac{1}{\tau}X(t) dt + \hat{\sigma} \sqrt{dt} \epsilon(t) + b dt \text{ with } \tau, \hat{\sigma}, b > 0. \quad (1)$$

Here $X(t)$ is the process that is being observed as a function of time, τ is the damping timescale of the process (i.e. the time that is needed for the time series to become roughly uncorrelated), $\epsilon(t)$ is a white noise with zero mean and unit variance, and $\hat{\sigma}$ describes the light curve variations at small timescales and is related to the variance of the process $\sigma = \hat{\sigma}\tau/2$. Finally, b is related to the mean of the process given by $b\tau$. This noise model has a PSD of the form:

$$P_{\text{sing}}(f) = \frac{2\hat{\sigma}^2\tau^2}{1 + (2\pi\tau f)^2}, \quad (2)$$

while its covariance matrix can be written as:

$$C_{i,j}^{\text{sing}} = C_{i,j}^{\text{DRW}} = \frac{1}{2}\hat{\sigma}^2\tau \exp\left(-\frac{\Delta t_{i,j}}{\tau}\right) \quad (3)$$

where $\Delta t_{i,j} = |t_i - t_j|$ and $i, j = 1 \dots N$, represent two arbitrary observations out of the total number of N pointings.

When considering an UBHD, assuming the light curve from one MBH to be independent of the other, the PSD can be written as the sum of the two PSDs:

$$P_{\text{UBHD}}(f) = \frac{2\hat{\sigma}_1^2\tau_1^2}{1 + (2\pi\tau_1 f)^2} + \frac{2\hat{\sigma}_2^2\tau_2^2}{1 + (2\pi\tau_2 f)^2} \quad (4)$$

where $\tau_1, \hat{\sigma}_1$ and $\tau_2, \hat{\sigma}_2$ are the damping timescales and short-term variability magnitudes of the two processes. The covariance function for this model can be written as the sum of two covariance functions:

$$C_{i,j}^{\text{UBHD}} = C_{i,j}^{\text{DRW},1} + C_{i,j}^{\text{DRW},2} \\ = \frac{1}{2}\hat{\sigma}_1^2\tau_1 \exp\left(-\frac{\Delta t_{i,j}}{\tau_1}\right) + \frac{1}{2}\hat{\sigma}_2^2\tau_2 \exp\left(-\frac{\Delta t_{i,j}}{\tau_2}\right). \quad (5)$$

This is consistent with the assumption of two independent, stationary, zero-mean processes being added, as the absence of cross-correlations between the two processes allows both the covariance functions and PSDs to be additive.

Several studies have been carried out to find a relation between the parameters of the stochastic equation and the physical parameters of the MBH (see Kelly et al. 2009; MacLeod et al. 2010; Burke et al. 2021; Arévalo et al. 2023, 2024; Su et al. 2024; Benati Gonçalves et al. 2025 for some examples), but there is currently no consensus regarding these relationships. The methodology followed in this work does not require any strong assumptions about the UBHD physical parameters. Since our objective is to explore the parameter space as broadly and agnostically as possible, we do not assume any relation between the parameters of the noise model and those of the MBHs.

2.2. Time series generation

Each light curve was generated by sampling the flux of the MBH using `celerite`. Once the kernel of the Gaussian process (GP) was chosen, the light curve could be sampled with an arbitrary

sampling pattern and baseline. For this work, the injected parameters were drawn from log-uniform distributions for the damping timescales and uniform distributions for the variances of the two processes, as follows:

– Single MBH case:

$$10 \text{ d} < \tau < 500 \text{ d},$$

$$0 < \hat{\sigma}^2 < 1$$

– UBHD case:

$$10 \text{ d} < \tau_1 < 500 \text{ d},$$

$$10 \text{ d} < \tau_2 < \tau_1,$$

$$0 < \hat{\sigma}_1^2 < 1,$$

$$0 < \hat{\sigma}_2^2 < 1.$$

For our work, we sampled the time series for 10 years to resemble the LSST nominal survey. As noted in Kozłowski (2017), to obtain good constraints on the flat part of the PSD, the process should be observed for at least 10τ . With the assumed observational baseline, the flat part of the PSD can be observed particularly well for MBHs with a mass up to $\sim 10^8 M_\odot$ since they are expected to have a damping time of the order of 200 days (Burke et al. 2021). We initially sampled the light curves evenly in time every 3 days, and then generalised the test to more realistic light curves by omitting three to four months of observations each year to incorporate the inevitable gaps of ground-based observations. We assumed a relative flux uncertainty of 0.01. Such an uncertainty corresponds to a magnitude of ~ 21 in the r band of a single LSST visit, based on the expected photometric precision Ivezić et al. (2019). In the case of an evenly sampled light curve, we assumed a time step of $\Delta t = 3$ day, roughly corresponding to the cadence of the Wide Fast Deep survey mode of LSST, which will cover about 95% of the survey time. Examples of light curves generated in such a way, along with their respective PSDs, are shown in Figure 1. Here it is possible to see how a change in the parameters changes the observed light curve and its power spectrum. More specifically, a change in the parameter $\hat{\sigma}$ results in a uniform vertical shift of the PSD, with higher values of $\hat{\sigma}$ leading to increased power at all frequencies, as shown in (b) panel of Figure 1. A change in the damping timescale τ causes a change in the break frequencies with a higher τ moving the flat part of the spectrum towards lower frequencies, as shown in panel (d) of Figure 1. The light curves for the UBHD case are constructed by summing two light curves generated as in the single MBH scenario with equal weight.

2.3. Parameter estimation and model selection

For this work, we combined nested sampling and GPs to estimate the posterior distributions of the parameters and to compare the two competing models (for works that use a similar approach, see Covino et al. 2020; Zhu & Thrane 2020; Covino et al. 2022; Witt et al. 2022). While GPs often employ a maximum likelihood approach to optimise kernel hyperparameters for computational convenience, we took a fully Bayesian approach by using nested sampling to marginalise over these parameters. More specifically, nested sampling transforms the multidimensional integral for the evidence into a one-dimensional integral over the prior volume and explores nested contours of increasing likelihood by iteratively replacing low-likelihood samples with higher-likelihood ones. Unlike Markov chain Monte Carlo (MCMC) algorithms, nested sampling allows us to capture

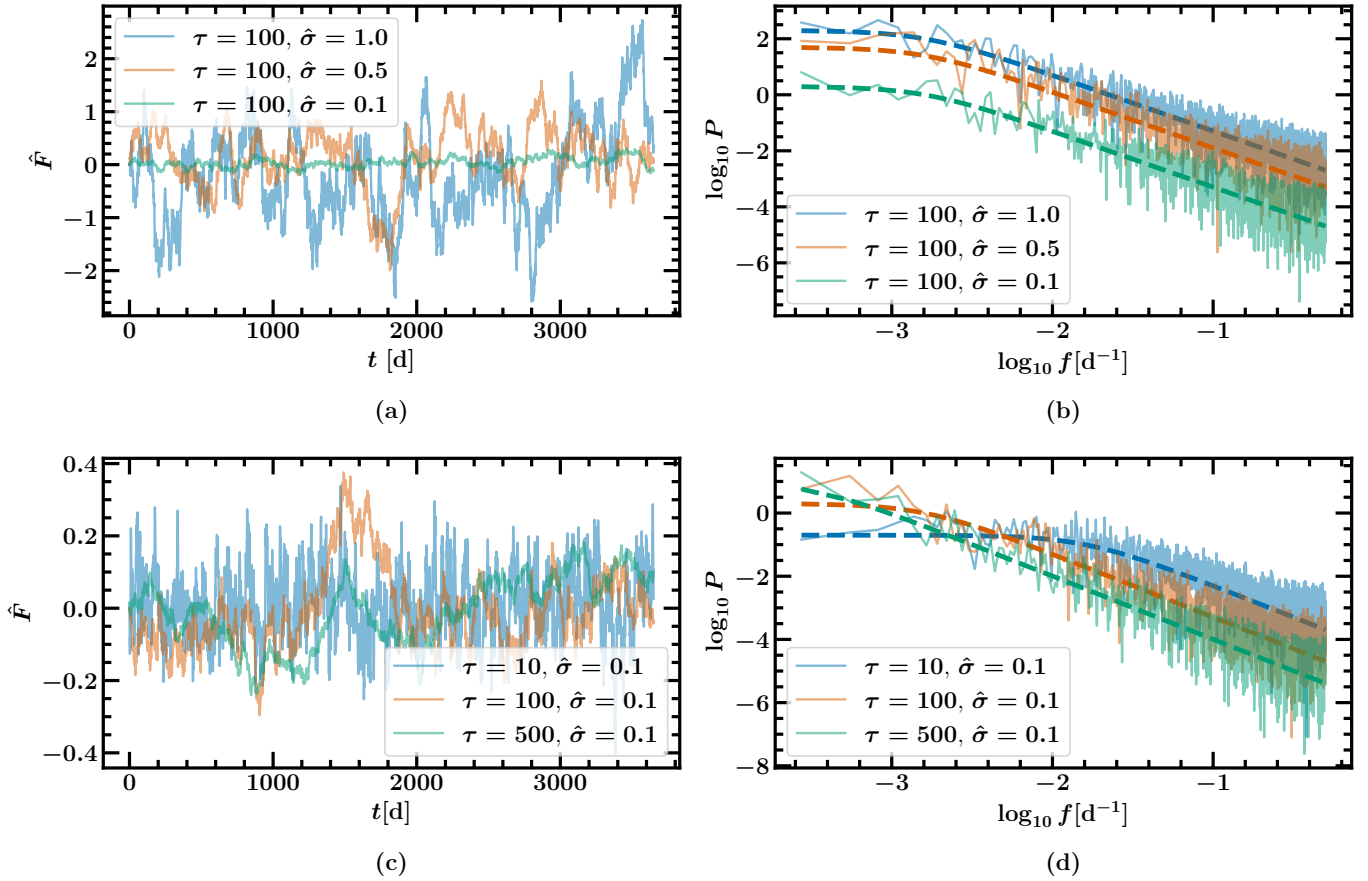


Fig. 1. Examples of light curves generated using `celerite` varying the process parameters: damping timescale τ (panel a), and the variability amplitude $\hat{\sigma}$ (panel c). In the right panels, the retrieved periodograms (computed as discussed in Appendix A) as well as the theoretical expected PSD (dashed lines) are shown. The time series (\hat{F}) are centred at zero and have arbitrary units. The parameter σ and power (P) should be considered in terms of the same dimensional scale as \hat{F} , with P having units of the square of the units of \hat{F} .

the full posterior distribution and compute the model evidence, enabling a statistically robust comparison between models (Skilling 2006; Ashton et al. 2022; Buchner 2023). We note that, due to the parallelism between Gaussian process kernels and PSD, UBHDs can be searched for through the fitting of the periodograms of evenly spaced light curves. However, for uneven data samplings, a consistent Bayesian analysis is not possible (as detailed in Appendix A). An analysis directly performed on the PSDs is presented in Appendix B together with its shortcomings.

The fitting of the light curve through GPs is performed using the GP regressor `celerite`, using a kernel in the form

$$k_{\text{single } i,j} = ae^{-c \Delta t_{i,j}}. \quad (6)$$

Equation (6) becomes equal to the kernel in Equation (3) by setting $a = 0.5\tau\sigma^2$ and $c = 1/\tau$. In the context of `celerite`, a kernel such as the one in Equation (6) is known as `RealTerm`. The UBHD model can then be written as the sum of two `RealTerm` kernels:

$$k_{\text{UBHD}}(\Delta t) = a_1 e^{-c_1 \Delta t} + a_2 e^{-c_2 \Delta t}. \quad (7)$$

We note that we chose to use `celerite` due to its computational efficiency in evaluating the likelihood for kernels such as those presented in Equation (6). Specifically, the computational cost of `celerite` scales linearly with the number of data points, in contrast to the cubic scaling of standard GP methods.

The nested sampling is performed using `RayNest`. The priors are a log-uniform distribution in τ in between 10 and 500 d, and log-uniform in $\hat{\sigma}^2$ in between 10^{-6} and 10^6 . We emphasise that the prior distribution on the variance $\hat{\sigma}^2$ is much wider than the distribution used to generate the light curves. This choice was made to ensure that the results and posteriors are not biased by an artificial truncation of the prior boundaries, which can affect the resulting posterior and evidence.

For model comparison, we computed the Bayes factor, defined as the ratio of the evidence of the two models:

$$B_{\text{pair,sing}} = \frac{\int p(D|M_{\text{pair}}, \theta_{\text{pair}}) p(\theta_{\text{pair}}|M_{\text{pair}}) d\theta_{\text{pair}}}{\int p(D|M_{\text{sing}}, \theta_{\text{sing}}) p(\theta_{\text{sing}}|M_{\text{sing}}) d\theta_{\text{sing}}}. \quad (8)$$

Here $p(D|M, \theta, I)$ is the likelihood function and $p(\theta|M, I)$ is the prior; D is the data, i.e. the observed light curve (t_i, F_i) , with $i, 1, 2, \dots, N$ the number of data points and F_i being the flux observed at time t_i ; M is the considered model; and θ are the parameters for this particular model. In this work, we assume that the two models have the same prior probability. In the particular case of GPs, the likelihood of the data is given by (see Rasmussen & Williams 2006):

$$p(D|M, \theta, I) = \frac{1}{(2\pi)^{\frac{N}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(D - \mu)^T \Sigma^{-1}(D - \mu)\right) \quad (9)$$

where μ is the mean function of the GP, N is the number of observations, and Σ and $|\Sigma|$ are the covariance matrix of the GP and

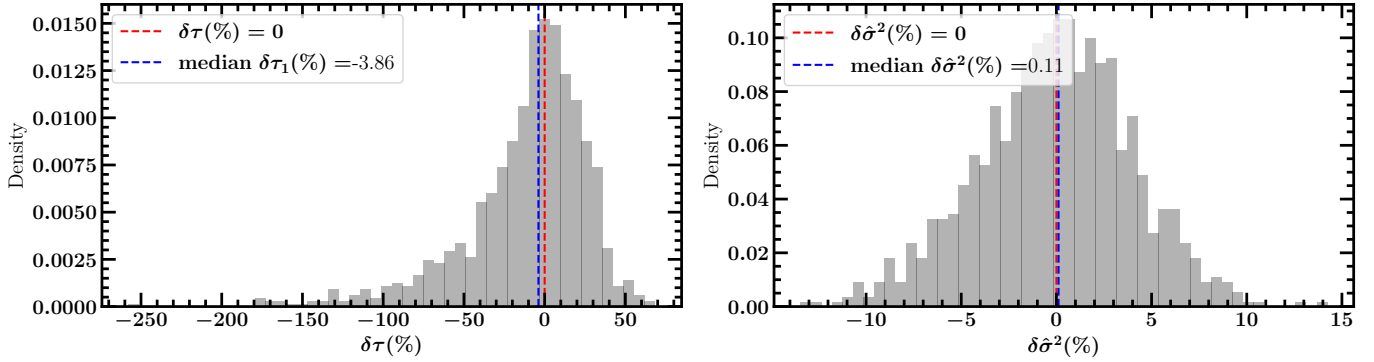


Fig. 2. Distribution of relative errors for the retrieved parameters in the single MBH evenly sampled case.

its determinant. Here we consider the evidence for UBHD to be significant when $B_{\text{pair,sing}} > 3$, i.e. if it shows at least substantial evidence according to the Jeffreys scale (Jeffreys 1998).

3. Results

The aim of the test proposed in this work is to assess the presence of an UBHD directly from the analysis of the observed light curve. To achieve this, we analysed realistic time series using GPs and nested sampling.

In this section, we present the main findings of our study. First, we assessed the false positive rate of the proposed signature by analysing single MBH light curves with both models and computing the corresponding Bayes factor (Section 3.1) to infer the percentage of single MBH light curves wrongly identified as UBHDs. Next, we evaluated how accurately the nested sampling procedure recovers the injected parameters (Section 3.2), considering both the single MBH and UBHD scenarios, and using both evenly and unevenly sampled light curves. Finally, we explored which regions of the model parameter space allow a reliable identification of UBHDs. To do this, we randomly sampled parameters and evaluated the Bayes factor to determine the conditions under which the test is most effective (Section 3.3).

3.1. False positives

To quantify the robustness of the proposed signature, it is important to quantify how unique it is versus how often it can be confused with the typical DRW variability of a single MBH. For this, we assessed the fraction of false positives produced by the test by simulating 1000 evenly sampled and 1000 unevenly sampled single MBH light curves and calculating the Bayes factor for the UBHD versus the single MBH model, as described in Section 2.3. We then computed the fraction of light curves that our test identifies as UBHDs, i.e. the fraction of light curves for which the Bayes factor is greater than 3. We find the 0.2% and 0.59% of false positives in the evenly and unevenly sampled scenarios, respectively. The percentage of false positives remains smaller than 1%, increasing when the sampling becomes uneven, even if the length of the observational campaign remains the same. We emphasise that the overall false detection rate is very low (regardless of light curve quality).

3.2. Parameter recovery

Before searching for the best parameter space in which the analysis of the photometric light curve can show evidence of the

Table 1. Parameter recovery for the single MBH model.

	Even sampling (%)	Uneven sampling (%)
$\delta\tau$	-3.86	-4.13
$\delta\hat{\sigma}$	0.11	0.03
σ_τ	82.76	84.18
$\sigma_{\hat{\sigma}}$	0.04	0.04

Notes. Median values of the relative error $\delta\theta$ and relative uncertainty range σ_θ for the parameters τ and $\hat{\sigma}$ of the single MBH model.

presence of an UBHD, we tested how well our algorithm can recover the parameters of the injected light curve for each model. This comparison was not a simple consistency check between model and data realisations, as the processes we looked at are intrinsically stochastic, i.e. individual realisations would differ even for the same set of parameters. To assess the goodness of the parameter estimation, we computed the relative error between the injected θ_{inj} and recovered θ_{rec} parameters as:

$$\delta\theta = \frac{\theta_{\text{rec}} - \theta_{\text{inj}}}{\theta_{\text{inj}}} \times 100\% \quad (10)$$

and the relative uncertainty range (σ_θ) defined as:

$$\sigma_\theta = \frac{q_{\theta,0.84} - q_{\theta,0.16}}{\theta_{\text{inj}}} \times 100\% \quad (11)$$

where θ_{inj} is one of the parameters with which the light curve has been generated, θ_{rec} is the median value of the posterior distribution for that particular parameter, $q_{\theta,0.16}$ and $q_{\theta,0.84}$ are respectively the 0.16 and 0.84 quantiles of the posterior distribution.

In Figure 2, we show the distribution of the relative errors $\delta\theta$ for 1000 evenly sampled light curves for the single MBH case. The dashed vertical red line indicates the optimal scenario $\delta\theta = 0$ while the dashed vertical blue line indicates the median value of $\delta\theta$. We see that both parameters are recovered well as these distributions are scattered around $\delta\theta = 0$ for both τ and $\hat{\sigma}$. The values we found for the median relative error $\delta\theta$ and relative uncertainty range σ_θ for both the even and uneven scenarios are reported in Table 1.

For the damping timescale, we see a tail at lower values, due to the fact that the length of the observational campaign is not long enough to efficiently sample the flat part of the PSD. In Figure 3, we do the same for the UBHD case. Here, the scatter from the true value is higher than that in the single DRW scenario, as can be seen from the values of the median relative error $\delta\theta$ and relative uncertainty range σ_θ reported in Table 2.

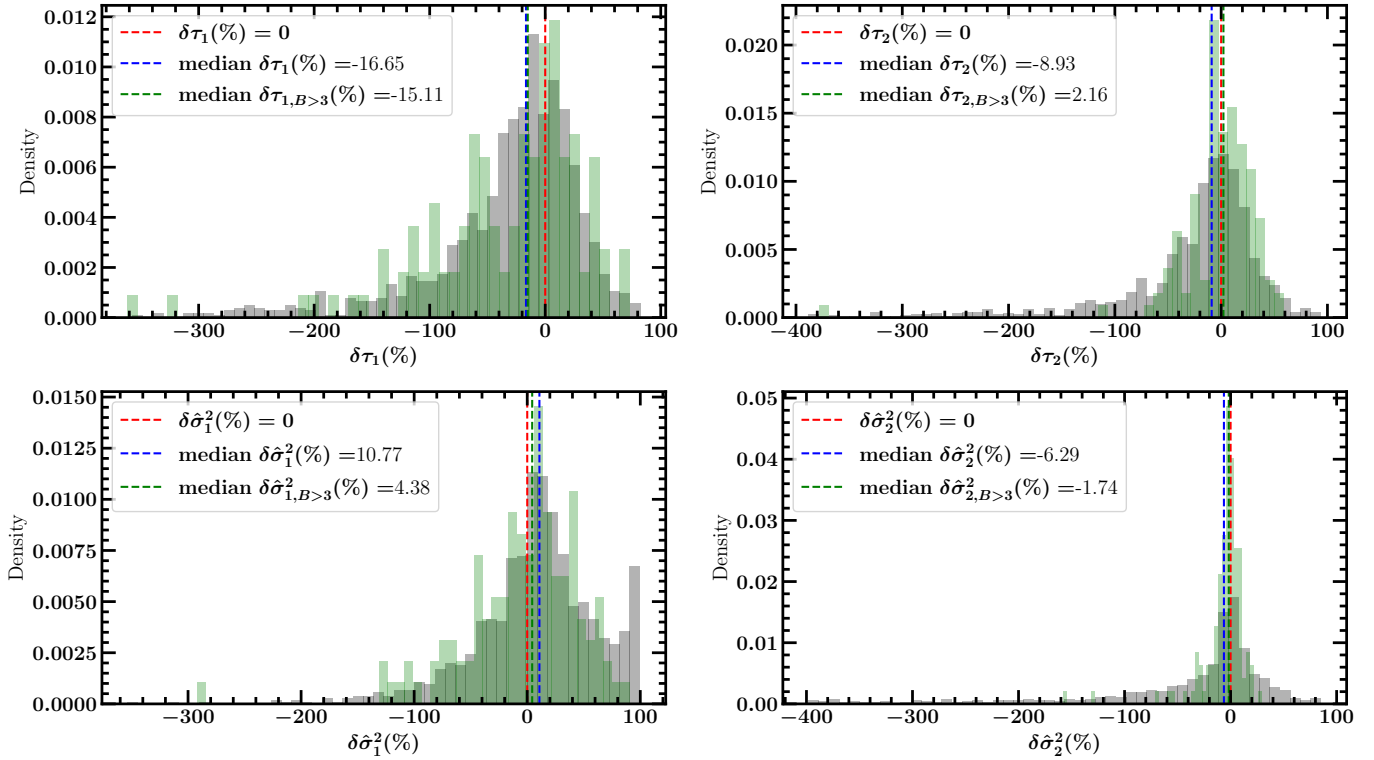


Fig. 3. Distribution of relative errors for the retrieved parameters in the UBHD evenly sampled case. Shown are the distributions for the cases in which the UBHD model is favoured (in green).

Table 2. Parameter recovery for the UBHD model.

	Even sampling (%)	Uneven sampling (%)
$\delta\tau_1$	-16.65	-13.31
$\delta\hat{\sigma}_1$	10.77	12.15
$\delta\tau_2$	-8.93	4.20
$\delta\hat{\sigma}_2$	-6.29	-0.73
σ_{τ_1}	260.56	266.95
$\sigma_{\hat{\sigma}_1}$	9.93	52.54
σ_{τ_2}	103.23	77.98
$\sigma_{\hat{\sigma}_2}$	6.08	18.01

Notes. Median values of the relative error $\delta\theta$ and relative uncertainty range σ_θ for the parameters τ_1 , τ_2 , $\hat{\sigma}_1$, $\hat{\sigma}_2$ of the UBHD model.

The green distributions in Figure 3 represent the distributions of $\delta\theta$ for the combination of parameters where the test correctly identifies an UBHD. The median values of the relative errors and the relative uncertainty range of such light curves are reported in Table 3. In this case, one can see that the scatter overall becomes smaller, but with τ_1 still showing wide tails in the distributions, similarly to the single DRW case.

In Table 4, we report the percentage of light curves for which all the parameters are retrieved with a relative error within the 10%, 20%, and 50% respectively for the single MBH, UBHD, and correctly identified (green distributions in Figure 3) cases.

3.3. True positives

Finally, we searched for the best region of the parameter space for which the test can efficiently identify UBHDs. More

Table 3. Parameter recovery for the correctly identified UBHD light curves.

	Even sampling (%)	Uneven sampling (%)
$\delta\tau_1$	-15.11	-19.09
$\delta\hat{\sigma}_1$	4.38	10.44
$\delta\tau_2$	2.16	2.89
$\delta\hat{\sigma}_2$	-1.74	-2.42
σ_{τ_1}	155.52	160.09
$\sigma_{\hat{\sigma}_1}$	3.84	4.08
σ_{τ_2}	45.10	50.75
$\sigma_{\hat{\sigma}_2}$	0.69	0.63

Notes. Median values of the relative error $\delta\theta$ and relative uncertainty range σ_θ for the parameters τ_1 , τ_2 , $\hat{\sigma}_1$, $\hat{\sigma}_2$ of the UBHD light curves correctly identified by the test (the green distributions in Figure 2).

specifically, we sampled 1500 light curves (for both evenly and unevenly sampled light curves) varying the ratio of the two damping timescales $q_\tau = \tau_2/\tau_1$ with $\tau_2 < \tau_1$ and the ratio of the amplitudes of the two processes $q_\sigma = \sigma_2/\sigma_1 = (\hat{\sigma}_2 \sqrt{\tau_2})/(\hat{\sigma}_1 \sqrt{\tau_1})$ with $\sigma_2 < \sigma_1$. For each pair of q_τ and q_σ , we computed the Bayes factor in order to highlight where the test works the best, i.e. where the Bayes factor is greater than 3.

In Figure 4, we show the result in the case of unevenly sampled light curves observed for 10 years with a cadence of 3 d. The points are colour-coded by the value of $\log_{10} B$, and the set of parameters for which the test yields a Bayes factor greater than 3 is represented by stars. The result does not significantly change when considering evenly sampled light curves. More specifically, in the evenly sampled scenario, 4.65% of all the

Table 4. Percentages of light curves for which all the parameters are retrieved with a relative error within the 10%, 20% and 50%.

	single MBH (E) [%]	single MBH (U) [%]	UBHD (E) [%]	UBHD (U) [%]	UBHD ★ (E) [%]	UBHD ★ (U) [%]
10%	29.55	27.69	1.10	0.67	2.38	1.43
20%	51.16	50.90	7.13	5.35	14.29	8.31
50%	84.92	86.15	35.14	20.74	48.41	39.54

Notes. From left to right: percentages of light curves with both retrieved parameters relative errors within the 10%, 20%, and 50% respectively in the single MBH scenario, UBHD scenario and for UBHD light curves that are correctly identified (green distribution in Figure 3, indicated in the Table with UBHD ★), respectively. E and U in the parentheses indicate the even and uneven sampling, respectively.

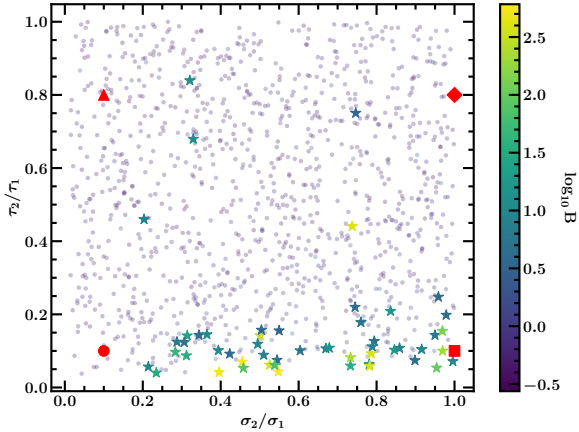


Fig. 4. Map of Bayes factors for 1500 unevenly sampled realisations of UBHD light curves with a ten-year baseline. The colour scale represents the base-10 logarithm of the Bayes factor, while the axes show the ratio of the model parameters between the two light curves. The red square, diamond, triangle and circle correspond to the light curves used to compute the periodograms shown in panels a,b,c, and d, respectively, in Figures 6 and B.1.

considered light curves are correctly identified as UBHDs, while in the unevenly sampled scenario, this fraction reduces to 3.77%.

Figure 4 shows clearly that the test we propose works only in a small region of the parameter space where $q_\tau \lesssim 0.2$ and $q_\sigma > 0.2$. In principle, setting a lower Bayes factor threshold could increase the fraction of correctly retrieved UBHD light curves and expand the region of the parameter space where the test works. For example, thresholds on B of 2, 1, and 0.5 would increase the fraction of correctly retrieved UBHD light curves from 3.77% (when the threshold is set to $B = 3$) to 5.3%, 11.3% and 73.9% in the case of unevenly sampled light curves and from 4.6% (for $B = 3$) to 5.4%, 12.5% and 87.6% in the case of evenly sampled light curves. The ratio of the damping timescales below which the test works moves from $q_\tau = 0.21$ for $B = 3$ to $q_\tau = 0.31$, $q_\tau \sim 0.65$ and $q_\tau \sim 0.87$ for $B = 2$, 1, and 0.5, respectively. At the same time, lowering the threshold would increase the false positive fraction. For the uneven sampling scenario, the false positive fraction would increase from 0.59% (for $B = 3$) to 1%, 6% and 82%, while in the even sampling scenario the false positives would increase from 0.59% (for $B = 3$) to 1.07%, 6%, and 82% for Bayes factor thresholds of 2, 1, and 0.5, respectively. Therefore, lower thresholds would make the statistical evidence in favour of the UBHD scenario less robust. For this reason, we chose a more conservative limit for the threshold.

In Figure 5, we show 500 new simulations in order to provide a zoomed-in image of this region of parameter space for UBHD light curves. In particular, UBHDs are identified when

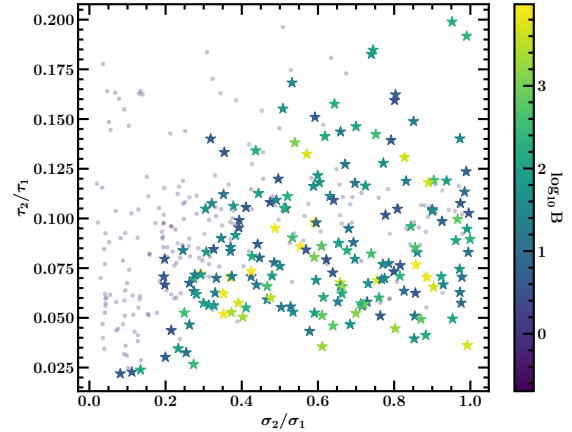


Fig. 5. Same kind of map as reported in Figure 4 but for 580 light curves in the region of the parameter space where the UBHDs can be correctly identified.

the two MBHs have very different damping timescales and similar variability amplitudes. This is not surprising if one considers the shape of the PSD. In Figure 6, we show the PSDs for four cases, one of which is identified (in panel a) as an UBHD. More specifically, the search identifies an UBHD if, after the first plateau of the PSD up to the break frequency related to the longer damping timescale, a second plateau starts appearing up to the second break frequency related to the shorter damping timescale.

Since the damping timescale positively correlates with the MBH luminosity and with the mass of the MBH, we can conclude that pairs where the two MBHs have very different luminosities or very different masses (see MacLeod et al. 2010), i.e. a low luminosity ratio or a low mass ratio, are expected to be more easily identified. The ability of this analysis to identify UBHDs large-scale photometric surveys (which provide unevenly sampled data) makes it relevant for currently available photometric surveys, most of which feature gaps and uneven pointings.

We repeated the same procedure, changing the length of the observational baseline. More specifically, we studied the cases of light curves observed for 6 and 3 years, keeping the uneven sampling discussed in Section 2.2. These baselines are relevant for current time-domain surveys (e.g. ZTF), and for the early years of LSST. As can be seen in Figure 7, as the length of the observation decreases, the region where the test is able to correctly retrieve the presence of an UBHD shrinks to a region of higher q_σ . Specifically, by repeating simulations in the region with $q_\tau < 0.2$, analogous to those used to generate Figure 5, we find that the value of q_σ above which 90% of the correctly retrieved light curves lie moves from $q_\sigma = 0.21$ for the 10 yr baseline to $q_\sigma = 0.36$ and $q_\sigma = 0.40$ for baselines of 6 and 3 yr respectively.

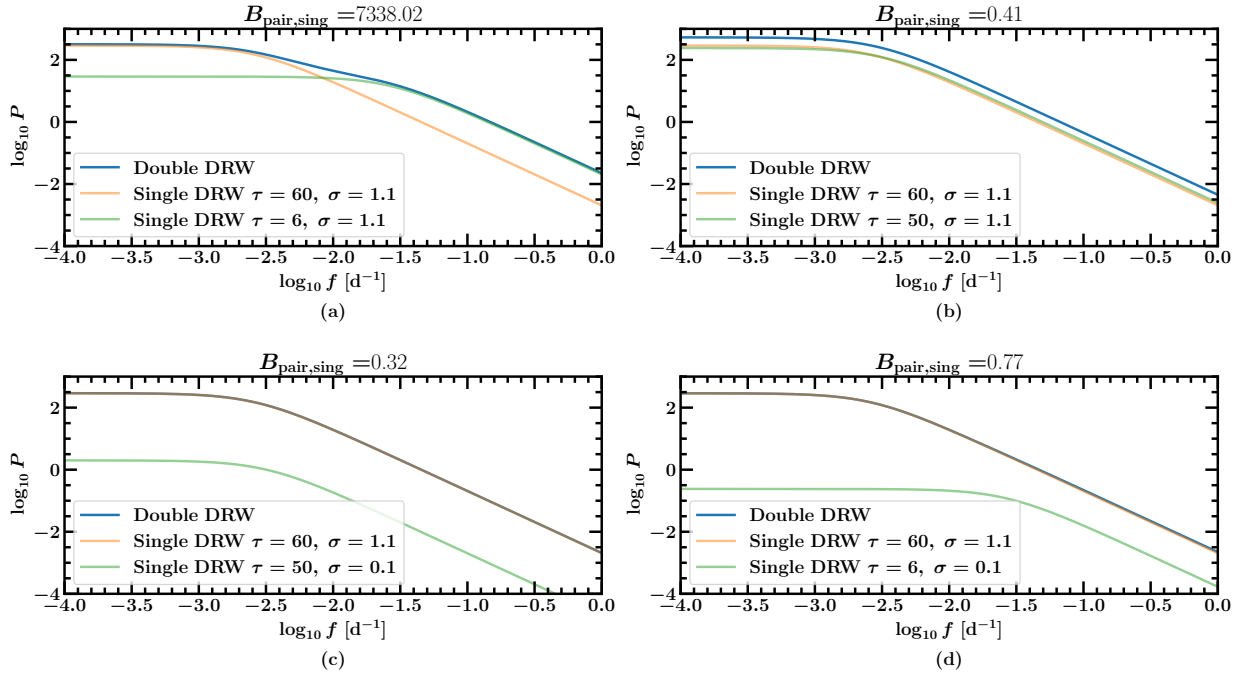


Fig. 6. PSD of double DRW kernels for different parameter ratios. The Bayes ratio resulting from the GP fitting is reported above each panel. In all the panels, the light curves are sampled every 3 d for 10 yr. The power is reported in arbitrary units. Case (a): Similar variability amplitudes but considerably different damping timescales. The test can correctly identify the UBHD (shown as the red square in Figure 4). Case (b): Similar amplitudes and similar damping timescales (shown as the red diamond in Figure 4). Case (c): Considerably different amplitudes but similar damping timescales (shown as the red triangle in Figure 4). Case (d): Considerably different amplitudes and damping timescales (shown as the red circle in Figure 4).

4. Discussion and conclusions

In this paper, we studied the possibility of detecting the presence of an UBHD using AGN light curves. In particular, we examined whether we can distinguish the presence of such a system from the typical variability of an AGN with a single MBH (typically described with a DRW model). For this signature, we focused on the optical variability, where we will soon have large samples of quasar light curves.

We generated both single MBH mock light curves and UBHD mock light curves by summing two single MBH time series with different cadences (either evenly sampled light curves with a three-day cadence or including observational gaps of a few months every year) and different lengths in terms of observation time (10, 6, or 3 yr). We then modelled the simulated time series employing GPs, using either a single DRW kernel or the sum of two uncorrelated DRW kernels. Finally, we checked whether the presence of a double DRW could be effectively observed through Bayesian model comparison. In particular, we used nested sampling to compute the evidence of the two models (single and double DRW) and calculated the Bayes factor, which allowed us to assess which model describes the simulated data better.

We studied how many false positives, i.e. light curves generated as single MBH, are better described by the double DRW model according to the Bayes factor (Section 3.1). To do so, we generated single MBH light curves with different cadences, and for each of the time series, we computed the Bayes ratio. We considered as a false positive a time series that presented a Bayes ratio higher than 3, i.e. the Bayes ratio indicates substantial evidence in favour of the double DRW model following the Jeffreys scale. More specifically, 1000 evenly sampled light curves with a

10 yr baseline show a false positive rate of 0.2%, while the same number of unevenly sampled light curves with the same baseline length result in 0.59%, demonstrating that the proposed signature has a low confusion rate with the typical AGN variability.

We assessed whether nested sampling successfully recovered the injected parameters, both for the single MBH scenario and for UBHDs (Section 3.2). To do so, we computed the relative per cent error $\delta\theta$ (i.e. the normalised difference between the injected and recovered parameter, see Equation (10)) and the relative per cent uncertainty range (i.e. the normalised uncertainty range between by the injected parameter, see Equation (11)). We showed that 51.16% of evenly sampled single MBH light curves, 50.90% of unevenly sampled single MBH light curves, 7.13% of evenly sampled UBHD light curves and 5.35% of unevenly sampled UBHD light curves with a 10 yr baseline show recovered parameters within 20% of the injected parameters.

We studied which region of the parameter space allows us to detect an UBHD (Section 3.3). In particular, we generated light curves with different ratios of the two damping timescales, and variances of the two DRW and constructed maps that show which combinations of these two ratios provide substantial evidence for the double DRW model. We found that our test is able to identify UBHDs that show very different damping timescales, $q_\tau < 0.2$, and similar variability amplitudes $q_\sigma > 0.2$ for both evenly and unevenly sampled light curves observed for 10 years. For the correctly identified UBHDs, the fraction of light curves with $\delta\theta < 20\%$ grows to 14.29% (8.31%) for the evenly (unevenly) sampled data. The low ratio in the damping timescales implies that only UBHDs with very different luminosities (and masses) will be correctly classified. Furthermore, we changed the length of the observation, maintaining an uneven sampling, and found that shortening the length of the

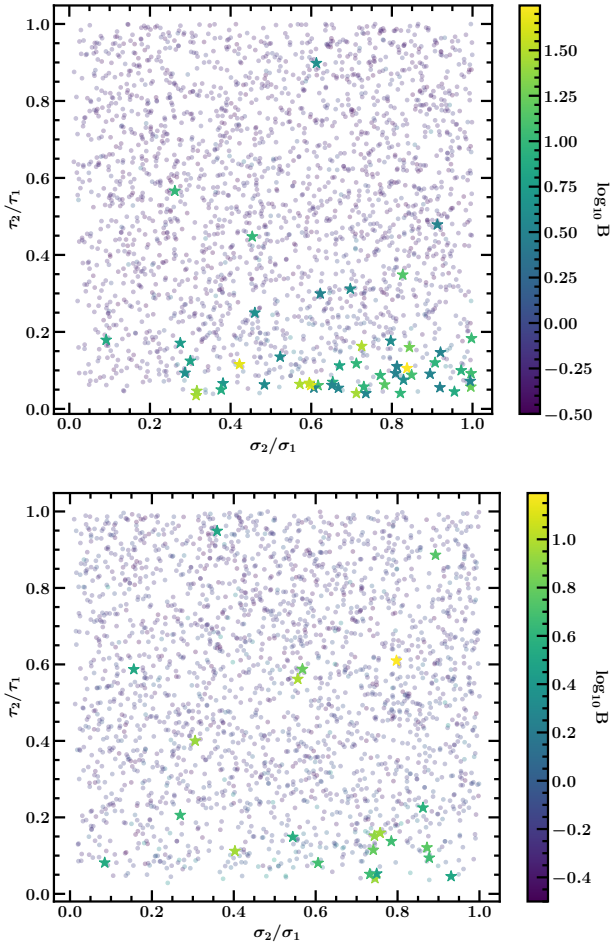


Fig. 7. Similar to the map in Figures 4 but with different observational baselines. In the upper panel, the light curves are observed for 6 years, while in the lower panel, the light curves are observed for 3 years.

observational baseline increases the q_σ needed to correctly identify an UBHD.

In the Appendices, we explored another possible method to detect the presence of an unresolved UBHD by studying the shape of the periodogram applicable only for evenly sampled data, which is not readily available. In Appendix A, we showed how the periodogram can be computed and discussed the statistical properties of the classical periodogram. In Appendix B, we showed that by using the Whittle likelihood (see Whittle 1951) to describe the distribution of the retrieved periodogram with respect to the true PSD, one can assess whether a single or double DRW better describes the observed periodogram shape. This method works only in the case of evenly sampled time series. In the case of unevenly sampled time series, this approach cannot be used as the statistical properties are different from the classical properties and depend on the particular noise considered.

Once candidates are identified, confirming the UBHD nature of objects with intermediate values of $B_{\text{pair,sing}}$ and discerning whether the two MBHs are bound in a binary or not will require follow-up observations. High-resolution imaging would have the chance to directly identify MBHPs (e.g. Hwang et al. 2020; Mannucci et al. 2022; Scialpi et al. 2024; Chen et al. 2023), while optical spectroscopy allows us to check for the presence of two sets of narrow emission lines (NELs), supporting the UBHD scenario (e.g. Comerford et al. 2009). If a single set of

NELs is observed, the identification of BELs that are asymmetric or offset with respect to the host rest frame in optical/UV spectra would be an indication of the presence of a close binary (e.g. Shen & Loeb 2010; Tsalmantza et al. 2011; Eracleous et al. 2012). In this case, reverberation mapping studies could be used to further support binary interpretation, searching for uncorrelated variability of different parts of single BELs (Gaskell 1988; Dotti et al. 2023).

It should be noted that our analysis was able to identify systems composed of two MBHs whose hosts are not in interaction, for example, chance superpositions of two AGNs at different redshifts or within a galaxy cluster (e.g. Dotti & Ruszkowski 2010). This particular scenario can be tested by looking at the frequency shift between two sets of NELs and by looking at the local density of galaxies in the UBHD-candidate neighbourhood (e.g. Decarli et al. 2014). Another scenario that could be similar to the one explored in this work is the case of a strongly lensed quasar with angularly unresolved multiple images. However, such a scenario cannot produce false UBHDs as the light curves of the different images generated by gravitational lensing will have the same damping timescale, generated by the same quasar.

The only differences between the lensed light curves would be the amplitude, due to the different magnification of the images, and a time delay. To test this scenario, we simulated 2000 unresolved lensed quasar light curves by simulating realisations of a single DRW and duplicating them. The two duplicated light curves are then multiplied by different magnifications computed assuming an un-truncated Singular Isothermal Sphere (SIS) as the matter distribution for the lens, and sampling from a uniform distribution between 0 and 1 for the ratio between the source-lens angular separation and the lens Einstein radius. A delay sampled from a log-uniform distribution between a minute and two weeks is then introduced between the two light curves. The Bayes factor for all the simulated light curves is smaller than $B = 1$, thus favouring the single DRW model with respect to the UBHD model, as expected.

Finally, in this work, we examined our proposed signature in single-band AGN light curves. However, the upcoming LSST will deliver multi-band light curves, which may be advantageous to combine. Therefore, we plan to extend the analysis presented in this paper to light curves obtained in different wavelength bands (as explored in Covino et al. 2022, for the identification of periodicities in AGN light curves).

All of the above discussion and conclusions rest on the assumption that AGN intrinsic variability is accurately described by the DRW model. However, deviations from the DRW have been reported (e.g. Zu et al. 2013; Kasliwal et al. 2015; Wang & Shi 2019), and alternative models have been proposed, such as the damped harmonic oscillator (DHO) model (Yu et al. 2022). If AGN light curves follow a DHO process rather than a DRW, our results may be affected. In the over-damped DHO case (see Yu et al. 2022), the covariance function reduces to a sum of two exponential kernels, resembling our UBHD model. Nevertheless, the two models predict different parameter correlations, which could, in principle, be used to distinguish between them. We plan to investigate the impact of including the DHO model on our results in future work.

Acknowledgements. LB acknowledges ISCR for awarding this project access to the LEONARDO supercomputer, owned by the EuroHPC Joint Undertaking, hosted by CINECA (Italy). LB wishes to thank the ‘‘Summer School for Astrostatistics in Crete’’ for providing training on the statistical methods adopted in this work. MC acknowledges support by the European Union (ERC,

MMMonsters, 101117624). FR acknowledges the support from the Next Generation EU funds within the National Recovery and Resilience Plan (PNRR), Mission 4 – Education and Research, Component 2 – From Research to Business (M4C2), Investment Line 3.1 – Strengthening and creation of Research Infrastructures, Project IR0000012 – “CTA+ – Cherenkov Telescope Array Plus. MD acknowledge funding from MIUR under the grant PRIN 2017-MB8AEZ, financial support from ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union – NextGenerationEU, and support by the Italian Ministry for Research and University (MUR) under Grant ‘Progetto Dipartimenti di Eccellenza 2023-2027’ (BiCoQ). We acknowledge a financial contribution from the Bando Ricerca Fondamentale INAF 2022 Large Grant, *Dual and binary supermassive black holes in the multi-messenger era: from galaxy mergers to gravitational waves*. FR acknowledges support from ELSA ELSA Euclid Legacy Science Advanced analysis tools” (Grant Agreement no. 101135203) is funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or Innovate UK. Neither the European Union nor the granting authority can be held responsible for them. UK participation is funded through the UK Horizon guarantee scheme under Innovate UK grant 10093177.

References

- Ackermann, M., Ajello, M., Albert, A., et al. 2015, *ApJ*, **813**, L41
 Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2023, *ApJ*, **951**, L8
 Amaro-Seoane, P., Andrews, J., Sedda, M. A., et al. 2023, *Liv. Rev. Relat.*, **26**, 2
 Arévalo, P., Lira, P., Sánchez-Sáez, P., et al. 2023, *MNRAS*, **526**, 6078
 Arévalo, P., Churazov, E., Lira, P., et al. 2024, *A&A*, **684**, A133
 Armitage, P. J., & Natarajan, P. 2002, *ApJ*, **567**, L9
 Ashton, G., Bernstein, N., Buchner, J., et al. 2022, *Nat. Rev. Methods Primers*, **2**, 39
 Barret, D., & Vaughan, S. 2012, *ApJ*, **746**, 131
 Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, *Nature*, **287**, 307
 Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2018, *PASP*, **131**, 018002
 Benati Gonçalves, H., Panda, S., Storchi Bergmann, T., Cackett, E. M., & Eracleous, M. 2025, *ApJ*, **988**, 27
 Bertassi, L., Charisi, M., Buscicchio, R., et al. 2025a, *A&A*, submitted [arXiv:2512.13688]
 Bertassi, L., Sottocorno, E., Rigamonti, F., et al. 2025b, *A&A*, **702**, A165
 Bloomfield, P. 2004, *Fourier Analysis of Time Series: An Introduction* (Wiley)
 Bortolas, E., Capelo, P. R., Zana, T., et al. 2020, *MNRAS*, **498**, 3601
 Bortolas, E., Bonetti, M., Dotti, M., et al. 2022, *MNRAS*, **512**, 3365
 Buchner, J. 2023, *Stat. Surv.*, **17**, 169
 Burke, C. J., Shen, Y., Blaes, O., et al. 2021, *Science*, **373**, 789
 Callegari, S., Mayer, L., Kazantzidis, S., et al. 2009, *ApJ*, **696**, L89
 Charisi, M., Bartos, I., Haiman, Z., et al. 2016, *MNRAS*, **463**, 2145
 Chen, Y.-C., Liu, X., Liao, W.-T., et al. 2020, *MNRAS*, **499**, 2245
 Chen, Y.-C., Liu, X., Lazio, J., et al. 2023, *ApJ*, **958**, 29
 Chen, Y.-C., Gross, A. C., Liu, X., et al. 2025, *ApJ*, **988**, 126
 Cocchiariaro, F., Franchini, A., Lupi, A., & Sesana, A. 2024, *A&A*, **691**, A250
 Comerford, J. M., Gerke, B. F., Newman, J. A., et al. 2009, *ApJ*, **698**, 956
 Covino, S., Landoni, M., Sandrinelli, A., & Treves, A. 2020, *ApJ*, **895**, 122
 Covino, S., Tobar, F., & Sandres, A. 2022, *MNRAS*, **513**, 2841
 De Rosa, A., Vignali, C., Bogdanović, T., et al. 2019, *New Astron. Rev.*, **86**, 101525
 Decarli, R., Dotti, M., Fumagalli, M., et al. 2013, *MNRAS*, **433**, 1492
 Decarli, R., Dotti, M., Mazzucchelli, C., Montuori, C., & Volonteri, M. 2014, *MNRAS*, **445**, 1558
 del Valle, L., Escala, A., Maureira-Fredes, C., et al. 2015, *ApJ*, **811**, 59
 Djorgovski, S. G., Drake, A. J., Mahabal, A. A., et al. 2011, ArXiv e-prints [arXiv:1102.5004]
 D’Orazio, D. J., & Charisi, M. 2023, ArXiv e-prints [arXiv:2310.16896]
 D’Orazio, D. J., Haiman, Z., & Schiminovich, D. 2015, *Nature*, **525**, 351
 Dotti, M., & Ruszkowski, M. 2010, *ApJ*, **713**, L37
 Dotti, M., Sesana, A., & Decarli, R. 2012, *Adv. Astron.*, **2012**, 940568
 Dotti, M., Rigamonti, F., Rinaldi, S., et al. 2023, *A&A*, **680**, A69
 EPTA Collaboration and InPTA Collaboration, (Antoniadis, J., et al.) 2024, *A&A*, **690**, A118
 Eracleous, M., Boroson, T. A., Halpern, J. P., & Liu, J. 2012, *ApJS*, **201**, 23
 Fiacconi, D., Mayer, L., Roškar, R., & Colpi, M. 2013, *ApJ*, **777**, L14
 Foreman-Mackey, D., Agol, E., Angus, R., & Ambikasaran, S. 2017, *AJ*, **154**, 220
 Franchini, A., Lupi, A., Sesana, A., & Haiman, Z. 2023, *MNRAS*, **522**, 1569
 Gaskell, C. M. 1988, in *Active Galactic Nuclei*, eds. H. R. Miller, & P. J. Wiita, 307, 61
 Governato, F., Colpi, M., & Maraschi, L. 1994, *MNRAS*, **271**, 317
 Graham, M. J., Djorgovski, S. G., Stern, D., et al. 2015, *Nature*, **518**, 74
 Haiman, Z., Kocsis, B., & Menou, K. 2009, *ApJ*, **700**, 1952
 Haiman, Z., Xin, C., Bogdanović, T., et al. 2023, ArXiv e-prints [arXiv:2306.14990]
 Hu, B. X., D’Orazio, D. J., Haiman, Z., et al. 2020, *MNRAS*, **495**, 4061
 Hwang, H.-C., Shen, Y., Zakamska, N., & Liu, X. 2020, *ApJ*, **888**, 73
 Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, **873**, 111
 Jeffreys, H. 1998, *The Theory of Probability* (Oxford: OUP), Oxford Classic Texts Phys. Sci.
 Jenkins, G., & Watts, D. 1968, *Spectral Analysis and Its Applications* (Holden-Day)
 Kasliwal, V. P., Vogeley, M. S., & Richards, G. T. 2015, *MNRAS*, **451**, 4328
 Kelly, B. C., Bechtold, J., & Siemiginowska, A. 2009, *ApJ*, **698**, 895
 Kormendy, J., & Gebhardt, K. 2001, *AIP Conf. Ser.*, **586**, 363
 Kozłowski, S. 2017, *A&A*, **597**, A128
 Kozłowski, S., Kochanek, C. S., Udalski, A., et al. 2010, *ApJ*, **708**, 927
 Li, Y.-R., Wang, J.-M., Ho, L. C., et al. 2016, *ApJ*, **822**, 4
 Li, Y.-R., Wang, J.-M., Zhang, Z.-X., et al. 2019, *ApJS*, **241**, 33
 Liu, T., Gezari, S., Ayers, M., et al. 2019, *ApJ*, **884**, 36
 Lomb, N. R. 1976, *Ap&SS*, **39**, 447
 MacLeod, C. L., Ivezić, Ž., Kochanek, C. S., et al. 2010, *ApJ*, **721**, 1014
 Mannucci, F., Pancino, E., Belfiore, F., et al. 2022, *Nat. Astron.*, **6**, 1185
 Mikkola, S., & Valtonen, M. J. 1992, *MNRAS*, **259**, 115
 Miles, M. T., Shannon, R. M., Reardon, D. J., et al. 2025, *MNRAS*, **536**, 1489
 Paolillo, M., & Papadakis, I. 2025, *Nuovo Cimento Riv. Ser.*, **48**, 537
 Papadakis, I. E., & McHardy, I. M. 1995, *MNRAS*, **273**, 923
 Pfister, H., Volonteri, M., Dubois, Y., Dotti, M., & Colpi, M. 2019, *MNRAS*, **486**, 101
 Priestley, M. 1981, *Spectral Analysis and Time Series* (Academic Press)
 Rasmussen, C. E., & Williams, C. K. I. 2006, *Gaussian Processes for Machine Learning*
 Reardon, D. J., Zic, A., Shannon, R. M., et al. 2023, *ApJ*, **951**, L6
 Rigamonti, F., Bertassi, L., Buscicchio, R., et al. 2025a, *A&A*, **702**, A242
 Rigamonti, F., Severgnini, P., Sottocorno, E., et al. 2025b, *A&A*, **693**, A117
 Runnoe, J. C., Eracleous, M., Pennell, A., et al. 2017, *MNRAS*, **468**, 1683
 Runnoe, J. C., Eracleous, M., Bogdanović, T., Halpern, J. P., & Sigurdsson, S. 2025, *ApJ*, **984**, 17
 Sandrinelli, A., Covino, S., Dotti, M., & Treves, A. 2016, *AJ*, **151**, 54
 Sandrinelli, A., Covino, S., Treves, A., et al. 2018, *A&A*, **615**, A118
 Scargle, J. D. 1982, *ApJ*, **263**, 835
 Schwartzman, E., Clarke, T. E., Nyland, K., et al. 2024, *ApJ*, **961**, 233
 Scialpi, M., Mannucci, F., Marconcin, C., et al. 2024, *A&A*, **690**, A57
 Severgnini, P., Ciccone, C., Della Cecca, R., et al. 2018, *MNRAS*, **479**, 3804
 Shen, Y., & Loeb, A. 2010, *ApJ*, **725**, 249
 Skilling, J. 2006, *Bayesian Anal.*, **1**, 833
 Sottocorno, E., Ogborn, M., Bertassi, L., et al. 2026, *A&A*, **708**, A153
 Souza Lima, R., Mayer, L., Capelo, P. R., Bortolas, E., & Quinn, T. R. 2020, *ApJ*, **899**, 126
 Su, Z.-B., Cai, Z.-Y., Sun, M., et al. 2024, *ApJ*, **969**, 78
 Thorne, K. S., & Braginskii, V. B. 1976, *ApJ*, **204**, L1
 Trindade Falcão, A., Turner, T. J., Kraemer, S. B., et al. 2024, *ApJ*, **972**, 185
 Tsalmantza, P., Decarli, R., Dotti, M., & Hogg, D. W. 2011, *ApJ*, **738**, 20
 Uttley, P., McHardy, I. M., & Papadakis, I. E. 2002, *MNRAS*, **332**, 231
 Valtonen, M. J., Lehto, H. J., Nilsson, K., et al. 2008, *Nature*, **452**, 851
 van der Klis, M. 1989, in *Fourier Techniques in X-Ray Timing*, eds. H. Ögelman, & E. P. J. van den Heuvel (Dordrecht, Netherlands: Springer), 27
 VanderPlas, J. T. 2018, *ApJS*, **236**, 16
 Varisco, L., Bortolas, E., Dotti, M., & Sesana, A. 2021, *MNRAS*, **508**, 1533
 Varisco, L., Dotti, M., Bonetti, M., Bortolas, E., & Lupi, A. 2024, *A&A*, **689**, A279
 Vaughan, S., Edelson, R., Warwick, R. S., & Uttley, P. 2003, *MNRAS*, **345**, 1271
 Vaughan, S., Uttley, P., Markowitz, A. G., et al. 2016, *MNRAS*, **461**, 3145
 Veitch, J., Del Pozzo, W., Lyttle, A., et al. 2024, <https://doi.org/10.5281/zenodo.12801702>
 Verbiest, J. P. W., Lentati, L., Hobbs, G., et al. 2016, *MNRAS*, **458**, 1267
 Vio, R., Andreani, P., & Biggs, A. 2010, *A&A*, **519**, A85
 Wang, H., & Shi, Y. 2019, *Ap&SS*, **364**, 27
 White, S. D. M., & Rees, M. J. 1978, *MNRAS*, **183**, 341
 Whittle, P. 1951, *Hypothesis Testing in Time Series Analysis* (Almqvist & Wiksells boktr.)
 Witt, C. A., Charisi, M., Taylor, S. R., & Burke-Spolaor, S. 2022, *ApJ*, **936**, 89
 Wu, Q., Scialpi, M., Liao, S., Mannucci, F., & Qi, Z. 2024, *A&A*, **692**, A154
 Xu, H., Chen, S., Guo, Y., et al. 2023, *Res. Astron. Astrophys.*, **23**, 075024
 Yu, W., Richards, G. T., Vogeley, M. S., Moreno, J., & Graham, M. J. 2022, *ApJ*, **936**, 132
 Zhang, Y.-W., Huang, Y., Bai, J.-M., et al. 2021, *AJ*, **162**, 276
 Zhu, X.-J., & Thrane, E. 2020, *ApJ*, **900**, 117
 Zu, Y., Kochanek, C. S., Kozłowski, S., & Udalski, A. 2013, *ApJ*, **765**, 106

Appendix A: Periodogram computation

The periodogram is a useful tool to estimate the true PSD of a process (see Priestley 1981; Bloomfield 2004). In its classical definition, the periodogram can be defined as the square of the discrete Fourier transform (see Scargle 1982). Panels (b) and (d) of Figure 1 show examples of the periodograms computed using the classical definition.

The sampled periodogram is not a smooth function of the frequencies but is scattered around the true PSD (see Vaughan et al. 2003). More specifically, the periodogram computed at a given frequency is distributed as a χ^2 distribution with two degrees of freedom, i.e. it is exponentially distributed as (see Scargle 1982; van der Klis et al. 1989):

$$p(I_j/S_j) = \frac{1}{S_j} e^{-\frac{I_j}{S_j}} \quad (\text{A.1})$$

where S_j and I_j are respectively the true PSD and the value of the periodogram computed at the frequency $f_j = j/N\Delta t$ with $j = 1 \dots N/2$. The shape of this distribution is due to the fact that the real and imaginary parts of the discrete Fourier transform are normally distributed for a stochastic process (see Jenkins & Watts 1968; Priestley 1981). As an example, in Figure A.1, we show an example of a white noise light curve, its periodogram with the true PSD and the distribution of the periodogram around the true PSD, along with the theoretical χ^2 distribution. The classical definition works particularly well for evenly sampled time series, but it is not without flaws. The periodogram is not a consistent estimator as the scatter does not decrease as the number of data points in the light curve

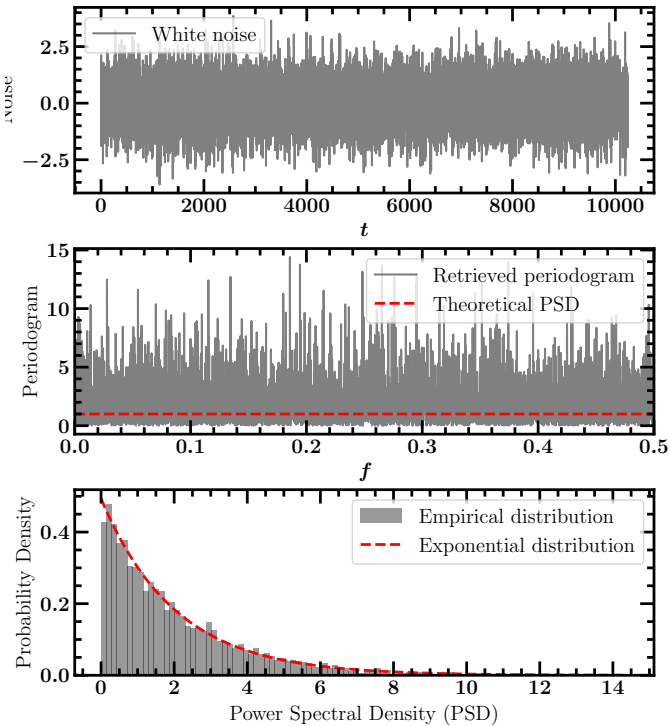


Fig. A.1. Example of a distribution for a white noise periodogram. Upper panel: noisy time series for a white noise with unity variance. Middle panel: comparison between the theoretical PSD (red dashed line) and the retrieved periodogram. Lower panel: distribution of the individual frequency bin estimates of the retrieved periodogram around the true PSD compared with the expected exponential distribution shown in Equation A.1 (red dashed lines).

increases (Jenkins & Watts 1968). Also, periodograms measured from finite data tend to be biased by windowing effects (see Vaughan et al. 2003, and references therein).

When the sampling is uneven, one can use the Lomb-Scargle periodogram (see Lomb 1976; Scargle 1982). The Lomb-Scargle periodogram is a generalisation of the classical periodogram, but its statistical properties are different. In particular, the observed periodogram is distorted from the underlying PSD by the sampling pattern of the underlying light curve. More specifically, in the context of red noise, effects such as aliasing and red noise leakage become important (see Papadakis & McHardy 1995; Uttley et al. 2002). Furthermore, the distribution of periodogram peaks is generally unknown, and correlations between the power at different frequency bins arise unless the light curve is evenly sampled and the periodogram is computed at the Fourier frequencies or another orthonormal frequency grid (see Vio et al. 2010; Covino et al. 2022).

Appendix B: Fitting directly the PSD

The best parameters of the stochastic process can be estimated directly from the periodogram. Also in this case, one can use nested sampling to do both evidence and best parameter estimation. In particular, the likelihood for the sampled periodogram points is given by the Whittle likelihood (see Barret & Vaughan 2012, for an application to X-ray power spectra). Given the probability (A.1), the Whittle log-likelihood can be written as:

$$\log L = \sum_{j=1}^{N-1} \frac{I_j}{S_j} + \log S_j. \quad (\text{B.1})$$

This approach can be used only in light curves where the time is evenly sampled. In the case of uneven sampling, the computation of the periodogram through the Lomb-Scargle procedure introduces correlations between different frequencies so that the Whittle likelihood does not approximate the true likelihood of the periodogram anymore. In particular, the specific powers at each frequency bin are not always independent. In principle, the presence of an UBHD can be searched for by fitting a Lomb-Scargle periodogram from an unevenly sampled light curve, but the statistical issues mentioned above will prevent a robust statistical estimate of the evidence of the UBHD hypothesis. In the following, we present for completeness a few examples of evenly sampled mock UBHD light curves fitted with two red-noise PSDs, indicating the Bayes ratios obtained under the two assumptions⁵. In Figure B.1, we show the result of the fitting procedure using the PSD and Whittle likelihood for a time series sampled every 3 d for 10 yr with the same parameter ratios used in Figure 6. The dashed green line represents the true PSD, while the solid red and blue lines show the best fit from the UBHD and single MBH models, respectively. The red and blue shaded areas represent the 68% and 95% confidence intervals of the UBHD and single MBH models, respectively. Also in this case, we reported the Bayes ratio resulting from the nested sampling.

Another approach to assess the presence of an UBHD without a fully Bayesian analysis, which we do not explore here, is to generate a large number of simulated curves assuming a large grid of values for the parameters of the assumed noise model. By defining a suitable test-statistic, such as the χ^2 , one can find the set of parameters that best describe the observed light curve, and

⁵ We emphasise again that the ratio obtained from fitting the Lomb-Scargle periodogram in unevenly sampled data should not be considered as a statistically robust indicator of the presence of an UBHD.

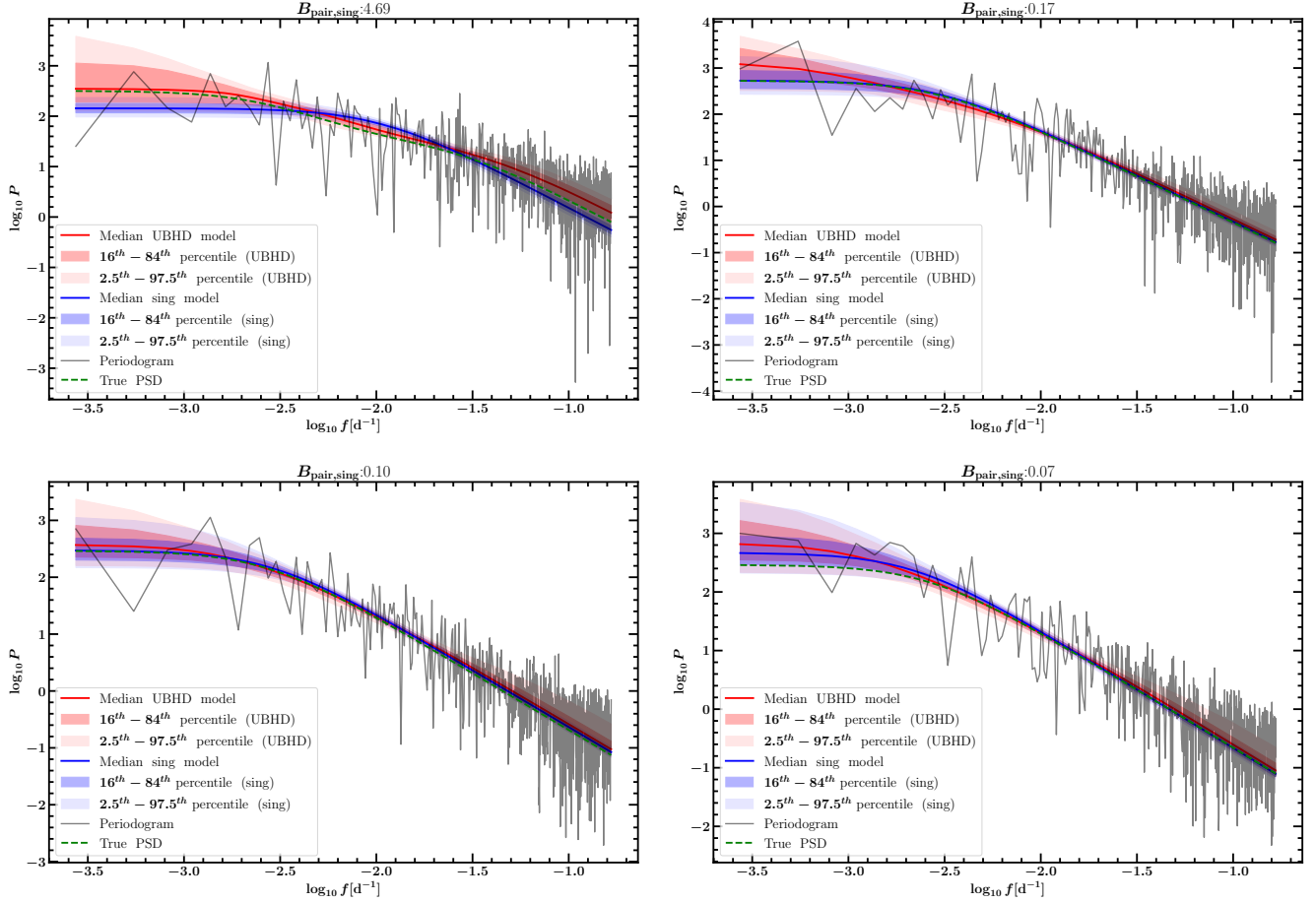


Fig. B.1. Results of the direct fit of the PSD for evenly sampled light curves with a 3 d cadence for 10 yr. The grey line indicates the light curve periodogram. The red and blue solid lines show the median model for the double and single DRW, respectively. The red and blue shaded areas show confidence intervals (16th – 84th and 2.5th – 97.5th) for the UBHD and single DRW, respectively.

given a goodness-of-fit statistic, one can compare the results for different noise models. We note, though, that while such a procedure may reduce the computational cost of the fitting procedure, it typically provides only point estimates of the parameters and approximations to their uncertainties, and it does not fully account for correlations or degeneracies between parameters, in contrast to the fully Bayesian approach presented in this work.