Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **513,** 2841–2849 (2022)
Advance Access publication 2022 March 7

https://doi.org/10.1093/mnras/stac596

Detecting the periodicity of highly irregularly sampled light curves with Gaussian processes: the case of SDSS J025214.67−002813.7

Stefano Covino [1]★ Felipe Tobar[2] and Aldo Treves[1,3]

[1] *INAF/Brera Astronomical Observatory, via Bianchi 46, I-23807 Merate (LC), Italy*
[2] *Initiative for Data & Artificial Intelligence, Universidad de Chile, Santiago de Chile, 8370456, Chile*
[3] *Università degli Studi dell'Insubria, Via Valleggio 11, I-22100 Como, Italy*

Accepted 2022 March 2. Received 2022 January 20; in original form 2021 December 3

ABSTRACT

Based on a 20-yr-long multiband observation of its light curve, it was conjectured that the quasar SDSS J025214.67−002813.7 has a periodicity of ∼4.4 yr. These observations were acquired at a highly irregular sampling rate and feature long intervals of missing data. In this setting, the inference over the light curve's spectral content requires, in addition to classic Fourier methods, a proper model of the probability distribution of the missing observations. In this article, we address the detection of the periodicity of a light curve from partial and irregularly sampled observations using Gaussian processes, a Bayesian non-parametric model for time series. This methodology allows us to evaluate the veracity of the claimed periodicity of the above-mentioned quasar and also to estimate its power spectral density. Our main contribution is the confirmation that considering periodic component definitely improves the modelling of the data, although being the source originally selected by a large sample of objects, the possibility that this is a chance result cannot be ruled out.

Key words: methods: statistical – quasars: individual: SDSS J025214.67−002813.7.

1 INTRODUCTION

Since the seminal paper by Begelman, Blandford & Rees (1980), the search of binary supermassive black holes (BSBHs) has attracted the interest of a vast group of astronomers. This is because BSBHs are regarded as the parent population of the strongest gravitational wave (GW) signals due to the final merging of the two black holes. Though such events have not been detected yet, current developments associated with, e.g. (i) the discovery of GW bursts associated with collapsed star merging detected by the LIGO/Virgo Collaboration (Abbott 2020), (ii) the advancement of detection techniques through the Pulsar Timing Array (PTA) experiment (Verbiest, Oslowski & Burke-Spolaor 2021), and (iii) future space borne instrumentation like LISA (Vitale 2014; Sesana 2021), have strongly increased the development of the field in recent years.

Among the various approaches to finding BSBH, a promising possibility has been that of detecting periodicities in the light curves of Active Galactic Nuclei (AGN). In particular, it is known that if the mass of the holes is of the order of $10^8\,M_\odot$ and the period is somehow related to the orbital one, with ∼0.01 pc separations, the periodicity should be of the order of years. Two main types of AGN have been considered thus far: Blazars (i.e. AGN dominated by a relativistic jet) and Quasars. Blazars largely populate the gamma-ray sky, and most efforts related to the detection of their periodicities have been performed by examining data from the Fermi Gamma-Ray Space Telescope (Atwood et al. 2009), which covers the entire sky every three hours and has been collecting data for ∼11 yr. For Quasars, main observational data are in the form of large collections of optical frames partly obtained from blind searches of transients. Data sources that are worth noticing in this regard are the Catalina Real-Time Transient Survey (Djorgovski et al. 2011) and the Palomar Transient Factory (Law et al. 2009).

Though there are thousands of Fermi gamma-ray blazars, those which are bright enough for constructing a light curve suitable for a temporal analysis are only a dozen. Careful examination of these blazars, performed by several researchers, has been possible due to the free availability of the data (e.g. Covino, Sandrinelli & Treves 2019, and references therein). In general, the search of a period of about 1 yr is a challenging endeavour because: (1) due to uneven sampling, the required observation time needs to be about 10 times larger than the target period; (2) blazars are strongly variable sources, and the signal should be searched above a frequency-dependent (red) noise; (3) the appearance of periodicity could be *episodical*, that is, the signal may be identifiable on a time scale of a few years, but not of a decade or longer. However, one advantage of the Fermi light curves is that the sampling is homogeneous, without seasonal gaps.

Various methods have been used for the periodicity search. Some are more apt for sinusoidal signals, whereas others are essentially independent of the shape of signal, and often, unsurprisingly, results are widely distinct under the consideration of different methods (Rieger 2019). The resulting picture is therefore still controversial. Arguably, the only source where the hypothesis of periodicity is supported by sufficient evidence is PG 1553 + 113. The period $T \sim 2.2$ yr in the Fermi data was first proposed by Ackermann et al. (2015), and then confirmed by Tavani et al. (2018). Note, however, that its significance was challenged by Ait Benkhali et al. (2019) and Covino et al. (2019), essentially due to the short number of epochs covered by the Fermi observations.

★ E-mail: stefano.covino@inaf.it

As for quasars, more than a hundred periodic candidates have been proposed in the last decade (e.g. Graham et al. 2015; Charisi et al. 2016), but the significance of some of these periods has been challenged because of the incorrect treatment of the multiple-trial effect due to the large number of considered sources (Vaughan et al. 2016). Finally, we have to quote the case of OJ 287, which is a Blazar, but its periodicity $T \sim 12$ yr was identified from optical data using a century of archived plates (Carrasco, Dultzin-Hacyan & Cruz-Gonzalez 1985; Dey et al. 2019). In this case, some concerns on the significance of the periodicity have also been raised (e.g. Stothers & Sillanpää 1997; Butuzova & Pushkarev 2021).

The methods for detecting the periodicity in blazars and quasars have proceeded practically independently, with little interactions between the communities cultivating the two fields. See, however, Holgado et al. (2018), where the contributions of the two populations to the GW background, constrained by PTA observations, are compared and Charisi et al. (2021) for future perspective of the field.

Recently, Chen et al. (2020) presented a systematic search of the periodicity of 625 quasars with a median redshift ∼1.8 and with multiband photometry available for a time scale of ∼20 yr. They found five candidates with a periodicity of 3–5 yr. Liao et al. (2021), using the same data set, considered very carefully the most interesting case, SDSS J025214.67−002813.7 ($z = 1.53$), discussing the significance of the proposed ∼4.4 yr (∼1607 d). Chen et al. (2021) focused on the possible physical interpretation for the identified periodicity.

In this paper, we re-examine the available optical data for SDSS J025214.67−002813.7 (hereinafter J0252) using the same data set as in Liao et al. (2021). Our analysis is carried out by modelling the J0252 data by means of Gaussian processes (GPs), extending the procedure proposed in Covino et al. (2020) for the case of blazars. The main goal of our work is to assess the statistical significance of the proposed periodicity and to determine the Power Spectral Density (PDS) of the target by means of a Bayesian non-parametric analysis.

Section 2 describes the data for J0252, while Section 3 describes the GP-based analysis methods in detail. Results are reported in Section 4 and a discussion is developed in Section 5. Finally, conclusions are given in Section 6. Software libraries used for the computations carried out in this paper are listed in Appendix A.

2 SDSS J025214.67−002813.7

J0252 was selected from a set of 625 quasars with spectroscopic identification (Liao et al. 2021). For these objects, multiband *griz* photometry was available from the Dark Energy Survey (DES, Dark Energy Survey Collaboration 2016) and the Sloan Digital Sky Survey (SDSS, Ivezić et al. 2007), covering a range of about 20 yr of observations, with irregular samplings, gaps, etc. We refer the reader to Liao et al. (2021) for any details about data analysis. The light curves in the *griz* are shown in Fig. 1. The median spacing is about 5–6 d, although the longest interruptions lasted more than about 1000 d for all the light curves. In total, the data set consists of about 230–280 epochs.

Liao et al. (2021) carried out a detailed analysis of the optical light curves of J0252 and reported an interesting periodicity at $P = 1607 \pm 7$ d with a significance of 99.95 per cent in the g band. Moreover, the significance is larger than 99.43 per cent in any band. Given the length of the monitoring, this implies that less than 5 complete cycles of variations have been covered. The short duration light curve, in terms of the time scale under analysis, makes any periodicity claim intrinsically delicate. Liao et al. (2021) identified the periodicity by applying different analysis tools. A spectral analysis was carried out by the Lomb–Scargle algorithm (Lomb 1976; Scargle 1982; Zechmeister & Kürster 2009; VanderPlas 2018), and the significance of the maximum power in the derived periodogram compared to the null hypothesis that it is due to noise was assessed by simulating a large number of artificial light curves with the same sampling of the original one. The analyses were also carried out in the temporal domain, fitting a sine curve or physically motivated models to the data, and modelling the autocorrelation function (ACF) of the data as a damped periodic oscillation (Alexander 1997; Graham et al. 2015). J0252 turned out to be the most interesting case (i.e. with the lowest false alert probability) in a small set of five objects showing periodicities larger than 99.74 per cent in at least one band. The significance of the periodicity appears substantially stronger if all the bands are analysed together.

In Fig. 2, we show the Lomb–Scargle periodograms for the *griz* bands, i.e. the convolution of the periodogram of the observed window together with that of the data (e.g. van der Klis 1989). The maxima of the periodograms are, of course, consistent with those reported in Liao et al. (2021) and very similar in each band. However, the periodograms also show the typical signatures of correlated noise (e.g. Vaughan 2010), with a flat region at shorter periods, where the periodograms are probably dominated by Poissonian noise, and a rising trend going to the longer (shorter) periods (frequencies). The presence of correlated noise and the short duration of the monitoring, compared to the proposed periodicities, make any claim about their significance somewhat ambiguous (see also, e.g. Krishnan et al. 2021).

The Lomb-Scargle algorithm can also be obtained by a Bayesian analysis under the prior assumption that data periodicity are modelled by a sinusoidal functional form (Bretthorst 2003; Jaynes & Bretthorst 2003; Mortier et al. 2015). This scenario allows one to obtain the probability density function (PDF) for the frequency given the stated assumptions and the uncertainties associated with the maxima in the periodograms. In our case, the PDF maxima and their 1σ uncertainties are: 1648 ± 69, 1657 ± 97, 1663 ± 87, and 1650 ± 145 d, for the *griz* bands, respectively.

3 GAUSSIAN PROCESSES

Gaussian processes [see Rasmussen & Williams (2006), for a comprehensive introduction] are probabilistic generative models for time series (Roberts et al. 2012), which can be considered as the infinite-dimensional extension of the multivariate Normal distribution. As such, they provide a Bayesian non-parametric approach to smoothing, interpolation, and forecasting (Hogg & Villar 2021). Being defined on the real line rather than a finite, user-defined, discrete grid, GPs are particularly suited for (Bayesian) regression of irregularly sampled data, possibly with large gaps, while providing sound error bars.

The hyperparameters of the GP model are the mean and covariance functions. The former is usually set to zero or to linear/quadratic functions, while the latter is a positive semidefinite function. The covariance *kernel* is chosen based on a priori knowledge of the dynamic behaviour of the data, for instance, the existence of periodicities or the requirement for differentiability. In addition to the well-known collection of off-the-shelf covariance functions (Rasmussen & Williams 2006), the choice of a covariance function can rest upon the Wiener–Khinchin theorem (Brockwell & Davis 2016), which states that – in the stationary case – the covariance kernel and the PSD of the data are Fourier pairs.

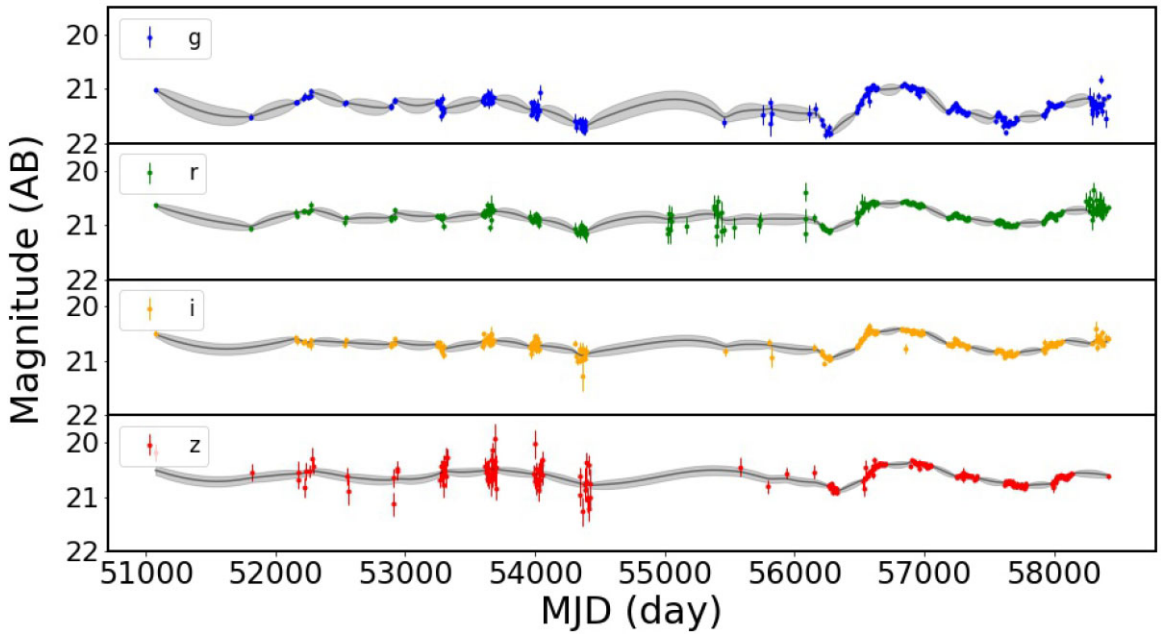

Figure 1. *griz* photometry for J0252 from Liao et al. (2021). For clarity, only points with 1σ magnitude error lower than 0.3 mag are plotted. The photometry covers about 20 yr although with seasonal gaps and interruptions and with an irregular sampling. Superposed to the four bands, we also plot the best-fitting multiband GP model and the 1σ uncertainty, as discussed in Section 4.

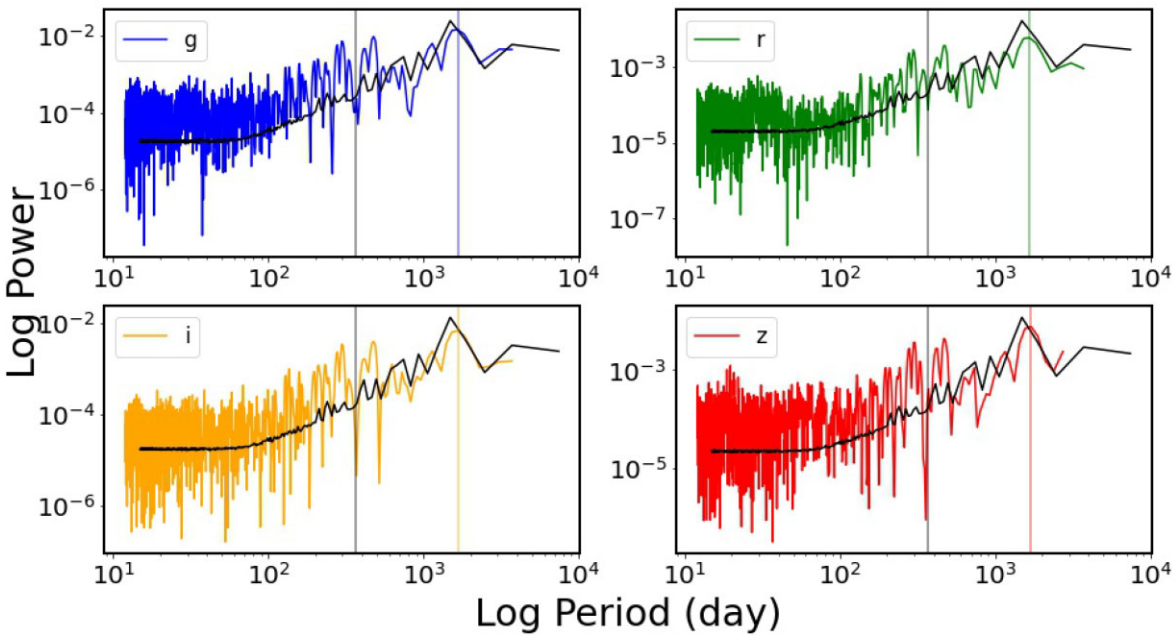

Figure 2. Lomb–Scargle periodograms for the *griz* bands of J0252. The coloured lines are the actual periodograms, while the black line is the PSD computed sampling at the maximum of the posterior distribution of the best-fitting GP model (see Section 4). The vertical lines indicate the maximum of the periodograms and the 1- yr period.

3.1 Periodicity assessment

GP techniques have been applied to identify possible periodicities in time series (Roberts et al. 2012; Tobar, Bui & Turner 2015; Durrande et al. 2016; Littlefair, Burningham & Helling 2017; Angus et al. 2018; Tobar 2018; Corani, Benavoli & Zaffalon 2020; Covino et al. 2020; Zhang et al. 2021). Here, we implement a procedure originally suggested by Littlefair et al. (2017) and then further developed by Covino et al. (2020). The idea is essentially to model the light curve of interest either with a stationary covariance function or with the same kernel but multiplied by a periodic (e.g. cosine) covariance function. Then, the problem of determining the presence of a periodic component hidden in the (often correlated) noise

affecting an astronomical time series can be addressed by means of Bayesian model selection (Kass & Raftery 1995; Jenkins 2014; Andreon & Weaver 2015; Trotta 2017).

Wilkins (2019) and Griffiths et al. (2021) carried out extensive simulations verifying the capability of some of the most common kernel functions (e.g. the squared exponential or radial basis, Matérn and rational quadratic kernels, Rasmussen & Williams 2006) to accurately modelling the data PSD. Results show that, depending on the specific sampling pattern and gap duration, either the Matérn or rational quadratic covariance functions typically outperform the squared exponential kernel. The Matérn kernel family is characterized by a parameter, ν, that drives the degree of 'smoothness' of the kernel. Time-series models corresponding to autoregressive processes of the first order are discrete-time equivalents of GP models with Matérn covariance functions with $\nu = 1/2$, which correspond, in one-dimension (1D), to the Ornstein–Uhlenbeck process (e.g. Roberts et al. 2012). The Ornstein–Uhlenbeck process is frequently invoked in modelling the statistical behaviour of AGN and blazars (e.g. Takata, Mukuta & Mizumoto 2018; Burd et al. 2021).

Given that the product and the sum of covariance kernels are again legitimate (i.e. symmetric and positive definite) covariance kernels, we can build a periodic covariance function by multiplying a stationary kernel with the so-called 'cosine' kernel (Covino et al. 2020). Since GP models are naturally suited for a Bayesian treatment, we can compute the probability favouring the more complex (i.e. periodic) model compared to a simpler one (non-periodic) by the Bayes factor. A remarkable by-product of this kind of analysis is that the PSD for the period can be easily obtained by marginalizing out (i.e integrating) the posterior probability distribution of the parameters obtained from the regression.

Spectral estimation by means of GP regression is an active research subject, which had a considerable boost after the seminal paper by Wilson & Adams (2013) where the PSD is modelled by a Gaussian mixture in the Fourier domain. This way, the resulting covariance function (via the Wiener–Khinchin theorem) is a *spectral mixture* of sub-kernels that allows one to approximate any stationary covariance kernel to arbitrary precision, given enough mixture components. The multi-output extension of the spectral mixture kernel was developed by Parra & Tobar (2017) and de Wolff, Cuevas & Tobar (2021). Other approaches have also been proposed. In the recent astrophysical literature, a mixture of the *celerite* covariance functions (Foreman-Mackey et al. 2017), a family of physically motivated kernels developed also for allowing a very fast computation, was applied (Yang et al. 2021) to derive the PSD from *Fermi*-LAT observations for a sample of blazars with published claims of quasi-periodicities. Another method termed Bayesian non-parametric spectral estimation (BNSE), was proposed by Tobar (2018); BNSE is the limit of the well-known Lomb–Scargle algorithm when an infinite number of components is considered with a Gaussian prior over the weights.

4 RESULTS

4.1 Periodic versus non-periodic kernels

We first modelled the light curves with the standard stationary kernels mentioned in Section3, i.e. the Square Exponential (SE), the Matérn with $\nu = \frac{1}{2}$ aka absolute exponential (AE) and the rational quadratic (RQ) (Rasmussen & Williams 2006). For all kernels, we denote $r = (t_i - t_j)$ the temporal difference between data points.

Table 1. Prior information adopted for analyses described in Section 4. Priors are properly normalized for the computation of the Bayes factors.

Hyper-parameter	Prior
$\ln A$	*Uniform* $[-20, 20]$
$\ln L$	*Uniform* $[-20, 20]$
$\ln \alpha$	*Uniform* $[-10, 10]$
$\ln P$	*Uniform* [ln (100), ln (3000)]

The SE kernel is given by

$$k_r = A \exp\left(-\frac{r^2}{2L^2}\right), \tag{1}$$

where $A > 0$ is the amplitude and $L > 0$ is the length-scale of the exponential decay. The AE kernel (or, in 1D, the kernel of the Ornstein–Uhlenbeck process) is given by:

$$k_r = A \exp\left(-\frac{r}{L}\right). \tag{2}$$

Lastly, the RQ kernel[1] is:

$$k_r = A \left[1 + \left(\frac{r^2}{2\alpha L^2}\right)\right]^{-\alpha}, \tag{3}$$

with $\alpha > 0$. This kernel can be obtained by a scale mixture (i.e. an infinite sum) of SE covariance functions with different characteristic length-scales drawn from a `gamma` distribution. The limit of the RQ covariance for $\alpha \to +\infty$ is, again, the SE covariance function.

In our regression experiments, the mean function of the GPs was set to the (constant) empirical mean of the observations. We also adopted flat priors for the kernel hyper-parameters as reported in Table 1. We considered both maximum likelihood implemented via L-BFGS-B (Byrd et al. 1995) and also integrated out the hyperparameters from the posterior density by a Markov Chain Monte Carlo (MCMC, Hogg & Foreman-Mackey 2018) based on the 'parallel-tempering ensemble' algorithm (Goggans & Chi 2004; Foreman-Mackey et al. 2013; Vousden, Farr & Mandel 2016). We started the Markov chains from small Gaussian balls centred on the best-fitting values. Then, we *thinned* the chain by discarding a fraction of the samples corresponding to a few times (typically four or five) the autocorrelation length of the computed chains, and we checked that a stationary distribution was reached (Sharma 2017). In most cases, the posterior distribution of the parameters (hereinafter just 'posterior') is single-peaked (unimodal) and model comparison could be quickly implemented by means of, e.g. the Bayesian Information Criterion (BIC, Schwarz 1978). However, in the general case of periodic behaviour assessment, we are interested in deriving the posterior PDF of the periods and therefore we carried out model comparison by a full computation of the Bayes factors (Kass & Raftery 1995; Trotta 2017).

Unsurprisingly, GP regression is very effective in modelling (i.e. interpolating) our data sets (e.g. Covino et al. 2020), yet Bayes factors show very different results for the adopted kernels (Table 2), reflecting their different abilities in describing the data covariance (Wilkins 2019; Griffiths et al. 2021).

Bayes factors computed from noisy data are themselves noisy (Jenkins 2014; Joachimi et al. 2021) therefore only very large values unambiguously identify the most suitable hypothesis, given the data

[1]Note that the analogous formula reported in Covino et al. (2020) is shown with a typo. We thank the anonymous referee for having pointed it out.

Table 2. Bayes factors (expressed as dB) computed for the light curves in all the four available bands for a stationary kernel (e.g. RQ or AE) over another (SE or AE).

Band	BF_{SE-RQ} dB	BF_{SE-AE} dB	BF_{AE-RQ} dB
g	53	44	09
r	55	68	−13
i	66	117	−51
z	5	−32	36

Table 3. Bayes factors (expressed as dB) computed for the light curves in all the four available bands for a periodic kernel over the best-fitting stationary one.

Band	BF_{RQ-CRQ} dB	BF_{AE-CAE} db
g	21	–
r	–	23
i	–	16
z	11	–

and the prior assumptions. Quite interestingly, results are different for each band, likely due to the different sampling and long-term coverage. The *g* and *z* band data are better described by the RQ kernel, while the *r* and *i* band data by the 'AE' kernel. The SE kernel is generally disfavoured although for the *z* band curve it behaves almost comparably to the RQ kernel. A possible interpretation of this behaviour depends on the presence of meaningful correlation for very long lags that, in case, can be more effectively modelled by the RQ kernel (Rasmussen & Williams 2006). The different results in different bands are probably due to the more or less effective coverage for long lags and the quality (i.e. the associated uncertainties) of the data.

Perhaps the simplest periodic covariance function is the cosine kernel:

$$k_r = A\cos(2\pi r/P), \tag{4}$$

where $A > 0$ is the amplitude, P is the period, and r is the temporal difference as defined above. Given that the product of two kernels is still a valid positive semidefinite kernel, we modelled our data with a kernel constructed by multiplying the best performing kernel in the previous experiment (Table 2) with the cosine kernel. We also adopted a flat prior distribution for the logarithm of the period, as reported in Table 1.

With this new kernel, we witnessed a relevant improvement as quantified by the Bayes factors shown in Table 3. Bayes factors, for ease of visualization, can be directly converted into probabilities conditioned on the data (e.g. Trotta 2007; Covino et al. 2020), obtaining factors of 99.2, 99.4, 97.6, and 92.1 per cent for the *g*, *r*, *i*, and *z* bands, respectively. In addition, the square root of the variances associated with each band, the *A* parameter in equations (1)–(4), can be interpreted as modulation amplitudes giving: ∼0.23, ∼0.14, ∼0.14, and ∼0.13 mag, respectively. These values are close to those reported in Liao et al. (2021; table 2), although derived following a different approach.

A direct byproduct of performing GP regression is the conditional covariance of the process given the observations, which can be compared to the ACF of the data. Computing the ACF for data affected by a very irregular sampling and long gaps is not an easy task. We followed here the discrete correlation function (DCF) algorithm proposed by Edelson & Krolik (1988), although we stress that the DCF computed in this paper is shown only for illustration purposes. Fig. 3 shows the (rather) noisy DCF compared to the GP regression best-fitting parameters. As mentioned already, the preference for the RQ or the AE kernel depends on the presence of correlation for long lags. The data, however, clearly show an oscillatory behaviour, yet the highly irregular sampling makes it difficult to assess whether we have a true periodicity or simply a recurrent behaviour with changing time scale.

Within the limits of our models, we can expect that the periods are better constrained when the RQ kernel is preferred since the AE kernel cannot support correlation on long lags. This is indeed the case, as we can see in Fig. 4. As expected, the PDF for the *g* and partly for the *z* bands are more peaked, while the range of possible periods for the *r* and *i* bands are larger. Periods maximising the GP regression posterior PDF tend to be shorter than those maximising the LS periodograms. A similar behaviour was also obtained by Liao et al. (2021) (their table 2), comparing results from LS and time-domain analyses.

Single-band study does not actually make the best use of all the available information. Having data collected in four different bands allows us to obtain a more solid inference by modelling the *g*, *r*, *i*, and *z* band data together and asking that the period be the same at all frequencies. This is an important step that is getting more and more significant given the common availability of multiband information from modern and future monitoring facilities (see, e.g. Huijse et al. 2012; Mondrik, Long & Marshall 2015; VanderPlas & Ivezić 2015; Saha & Vivas 2017; Huijse et al. 2018; Hu & Tak 2020; Elorrieta et al. 2021).

However, analysing multiband time-series requires some care since we may want to consider the, typically large, degree of correlation among the different bands. This is a basic topic in the GP literature (Pinheiro & Bates 1996; Osborne et al. 2012; Roberts et al. 2012; Parra & Tobar 2017) due to the clear interest in deriving information from multiple inputs in many sectors from research to industry (Roberts et al. 2012).

In this work, we have applied a simple technique known as '*intrinsic model of coregionalization*' (ICM, Bonilla, Chai & Williams 2008; Álvarez, Rosasco & Lawrence 2012) although many other, often more sophisticated, approaches are possible (e.g. van der Wilk et al. 2020; de Wolff et al. 2021). Assuming we have *P* different input data sets (or bands in our case), the kernel for our problem can be defined as:

$$\mathbf{K}_{\text{coreg}} = k(x, x') \cdot \mathbf{B}[i, j], \tag{5}$$

where x and x' are observations from any input data sets, $k(x, x')$ is a valid covariance function, as those introduced before, and $\mathbf{B}$ is a $P \times P$ positive-definite matrix. $\mathbf{K}_{\text{coreg}}$ therefore turns out to be

$$\mathbf{K}_{\text{coreg}} = \begin{bmatrix} B_{11}k(X_1, X_1) & \dots & B_{1P}k(X_1, X_P) \\ \vdots & \ddots & \vdots \\ B_{P1}k(X_P, X_1) & \dots & B_{PP}k(X_P, X_P) \end{bmatrix}. \tag{6}$$

Each output (band, in our case) can be seen as a linear combination of functions defined by the same covariance $k(x, x')$. The rank of the $\mathbf{B}$ matrix defines the number of components in the linear combination. Therefore, an ICM is generated by the product of two covariance functions: one that models the dependence between the outputs and one that models the input dependence (Álvarez et al. 2012). The restriction $B_{ij} = 0$ for $i \neq j$ describes the case with independent data for bands *i* and *j*.

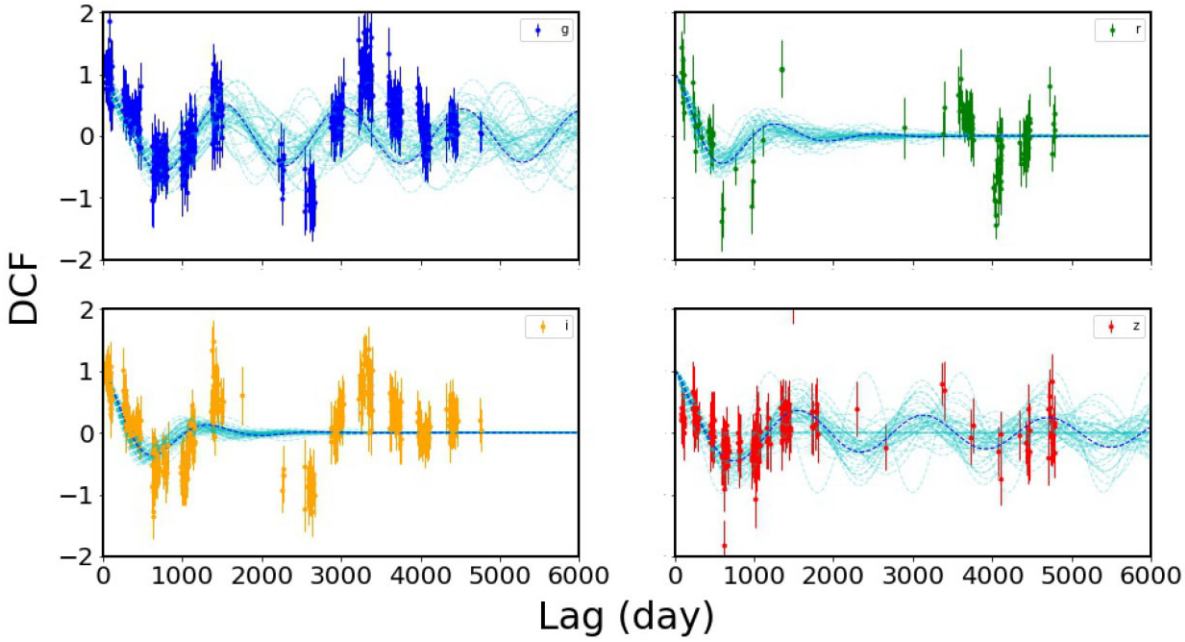

Figure 3. DCF for the *griz* bands of J0252. For clarity, only points with DCF errors lower than 0.4 are shown. The blue dashed line is the best-fitting periodic covariance function for each band, and the cyan lines are 50 randomly chosen samples extracted from the posterior distribution of parameters.

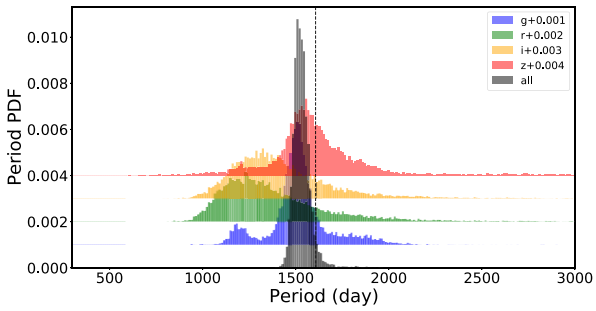

Figure 4. PDF for the periods computed marginalizing the posterior density distributions for the *griz* bands of J0252 and for the analysis with all the available bands together. The dashed vertical line identifies the period proposed in Liao et al. (2021). PDF for each band is artificially shifted for clarity by the amount reported in the legend. The most probable periods for the *griz* bands are 1511^{+103}_{-109}, 1317^{+323}_{-183}, 1570^{+220}_{-133}, and 1373^{+243}_{-215} d, respectively. The most probable period for the analysis including all the available bands is $P = 1525^{+42}_{-33}$ d.

ICM models can be, however, computationally demanding. In the high-rank case, the number of hyper-parameters to be learned is (in our case) larger than 20, making the computation of the likelihood computationally expensive. Therefore, for computational purposes, and also because the rather high degree of correlation and superposition among the four available bands makes high-rank coregionalized matrices unnecessary, we applied the simplest model with the lowest rank, thus modelling our four bands simultaneously with either the AE or the RQ covariance function.

As expected, once the kernel hyper-parameters are learned using the whole data set, the RQ kernel outperforms the AE, essentially because the former is better suited to model correlation for long lags. The introduction of a periodic component, now with the same period for all bands, is highly favoured by the computed Bayes factors

Table 4. Bayes factors (expressed as dB) computed for the light curves in all the four available bands considered together for a periodic kernel over the best-fitting stationary one.

BF_{AE-RQ} dB	BF_{RQ-CRQ} db
64	45

(Table 4), the probability in favour of the periodic model is about 4.2σ. The PDF for the period exhibits a strong peak (Fig. 4) and yields a period of $P = 1525^{+42}_{-33}$ d.

As mentioned before, adopting a higher-rank formulation quickly makes the (Monte Carlo) computation of the likelihood (and the Evidence) impractical due to the large number of parameters. Popular and, for simple inferences, effective alternatives are available as the already introduced BIC, which is a proxy of the actual Bayes factor, i.e. $\log \text{Evidence} \approx (-BIC/2)$ (see, e.g. Kass & Raftery 1995; Raftery 1999, for pros and cons of the BIC for model inference). Else, more advanced techniques for approximate Evidence computation that scale efficiently with the number of parameters have also been developed (e.g. variational inference, Opper & Archambeau 2009; Blei, Kucukelbir & McAuliffe 2017; Cherief-Abdellatif 2019).

Finally, as discussed in Section 3, another by-product of our analysis is an estimate of the PSD of the data. Applying again here a rank one ICM model with the best-fitting periodic covariance function, we can derive a PDF directly from the GP regression (Fig. 2). Given that the PSD can be computed for any needed sampling rate, it can be used for complex inferences or even fit as, e.g. in Vaughan (2010) with any physically driven model. More advanced GP based algorithms could also be applied, if needed, to derive the PSD with more general assumptions about the covariance of the data (e.g. de Wolff et al. 2021).

5 DISCUSSION

The analyses carried out in Section 4 show that a covariance function including a periodic component is supported by the data with a probability higher than 99 per cent for the g and r bands. These improvements are of interest, although somehow less significant compared to the results reported by Liao et al. (2021). Results obtained through different techniques should be carefully analysed since their direct comparison can be occasionally misleading due to their different assumptions. In addition, we also stress that here we report probabilities supporting a periodic kernel with 'any period' within the adopted prior, since we are considering the period as a nuisance parameter and thus integrating it out from our posterior. With essentially single-peaked unimodal distributions, it makes little difference, but this does not need to be generally the case (see also, e.g. Gregory & Loredo 1992, for a similar discussion).

The availability of multiple bands allows us to carry out the analysis making use of all the information provided by the data set, and the probability supporting a periodic kernel given the data becomes higher, corresponding to an improvement with respect to the purely stationary channel-independent solution of about 4.2σ. The best-fitting period, $P = 1525^{+42}_{-33}$ d, is shorter than those reported in Liao et al. (2021) and those obtained by our LS analysis (Section 2), although still roughly consistent given the uncertainties. This difference might indicate that the period is not stable during the monitoring, since LS analysis is essentially a change of basis, and data are simply described by the frequency content. A GP regression models a covariance function and the level of correlation for short and long lags affects how the best period describing the data is obtained. Shorter periods than those identified by an LS analysis were reported in Liao et al. (2021) too (see their table 2).

In any case, the source considered in Liao et al. (2021) and in this study was singled out by a sample of 625 analysed objects (Chen et al. 2020). This implies the reported probabilities have to be corrected for the so-called 'look elsewhere effect' (Bayer & Seljak 2020), i.e. the probability to find a positive result across multiple (independent) trials. In our case, the $\sim 4.2\sigma$ probability supporting the periodic kernel becomes, after the multiple trial correction (van der Klis 1989), $\sim 2.4\sigma$. This is still an interesting figure, yet far from being conclusive, implying that the periodicity proposed for J0252 can be just a chance result is still a plausible interpretation. We mention in passing that the multiple trial correction (see a discussion in Vaughan 2013) is often neglected since it is occasionally left unrecorded or badly described how large is the original sample of studied sources before devoting specific attention to the most interesting cases. Unfortunately, it is a factor that can easily dominate the final false alert probability for a proposed period.

6 CONCLUSIONS

In this paper, we have collected data for an AGN with a long multiband optical monitoring, SDSS J025214.67−002813.7. This source was singled out by Liao et al. (2021) from a larger sample of objects and, after an extensive analysis, they proposed a possible periodicity of ∼4.4 yr. The available data cover a long period, a couple of decades, but are affected by a very irregular sampling and long gaps. We re-analysed these data modelling the light curves by Gaussian processes, i.e. Bayesian non-parametric models for regression of time series. We analysed the single band data and by means of multi-output GP models the whole data set at once. We can confirm that a periodic component slightly longer than ∼4 yr improves the description of the data in all bands, and is also supported by the full analysis of the whole data set. Nevertheless, given that the considered source was originally identified in a large sample of several hundreds sources, after a correction due to the sample size the false alert probability of the periodicity does not appear to be low enough to rule out the possibility that this is just a chance result.

ACKNOWLEDGEMENTS

We thank the anonymous referee for her/his valuable and accurate comments. We acknowledge partial funding from the Agenzia Spaziale Italiana-Istituto Nazionale di Astrofisica grant I/004/11/5. FT acknowledges ANID grants Fondecyt-1210606, ACE210010, FB210005 & FB0008.

DATA AVAILABILITY

The data considered in this paper were published in Liao et al. (2021).

REFERENCES

Abbott B. P. E. A., 2020, Living Rev. Relativ., 23, 3
Ackermann M. et al., 2015, ApJ, 813, L41
Ait Benkhali F., Hofmann W., Rieger F. M., Chakraborty N., 2020, A&A, 634, 120
Alexander T., 1997, Is AGN Variability Correlated with Other AGN Properties? ZDCF Analysis of Small Samples of Sparse Light Curves. Springer, Dordrecht
Álvarez M. A., Rosasco L., Lawrence N. D., 2012, Found. Trends®i Mach. Learn., 4, 195
Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O'Neil M., 2016, IEEE Trans. Pattern Anal. Mach. Intell., 38, 252
Andreon S., Weaver B., 2015, Bayesian Methods for the Physical Sciences. Learning from Examples in Astronomy and Physics. Springer Int. Publ., Switzerland
Angus R., Morton T., Aigrain S., Foreman-Mackey D., Rajpaul V., 2018, MNRAS, 474, 2094
Astropy Collaboration, 2013, A&A, 558, A33
Astropy Collaboration, 2018, AJ, 156, 123
Atwood W. B. et al., 2009, ApJ, 697, 1071
Bayer A. E., Seljak U., 2020, J. Cosmology Astropart. Phys., 2020, 009
Begelman M. C., Blandford R. D., Rees M. J., 1980, Nature, 287, 307
Bisong E., 2019, Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners. Apress, New York
Blei D. M., Kucukelbir A., McAuliffe J. D., 2017, J. Am. Stat. Assoc., 112, 859
Bonilla E. V., Chai K., Williams C., 2008, in Platt J., Koller D., Singer Y., Roweis S., eds, Advances in Neural Information Processing Systems. Curran Associates, Inc., https://proceedings.neurips.cc/paper/2007/file/66368270ffd51418ec58bd793f2d9b1b-Paper.pdf
Bretthorst G. L., 2003, Frequency Estimation and Generalized Lomb-Scargle Periodograms. Springer, p. 309
Brockwell P. J., Davis R. A., 2016, Introduction to Time Series and Forecasting; 3rd ed. Springer Texts in Statistics. Springer, Cham
Burd P. R., Kohlhepp L., Wagner S. M., Mannheim K., Buson S., Scargle J. D., 2021, A&A, 645, A62
Butuzova M. S., Pushkarev A. B., 2021, Universe, 6, 191
Byrd R., Lu P., Nocedal J., Zhu C., 1995, SIAM J. Sci. Comput., 16, 1190
Carrasco L., Dultzin-Hacyan D., Cruz-Gonzalez I., 1985, Nature, 314, 146
Charisi M., Bartos I., Haiman Z., Price-Whelan A. M., Graham M. J., Bellm E. C., Laher R. R., Márka S., 2016, MNRAS, 463, 2145
Charisi M., Taylor S. R., Runnoe J., Bogdanovic T., Trump J. R., 2021, MNRAS, 510, 5929

Chen Y.-C. et al., 2020, MNRAS, 499, 2245
Chen Y.-C., Liu X., Liao W.-T., Guo H., 2021, MNRAS, 507, 4638
Cherief-Abdellatif B.-E., 2019, in Ruiz F., Zhang C., Liang D., Bui T., eds, Proceedings of The 1st Symposium on Advances in Approximate Bayesian Inference. PMLR, p. 11
Corani G., Benavoli A., Zaffalon M., 2020, preprint (arXiv:2009.08102)
Covino S., Sandrinelli A., Treves A., 2019, MNRAS, 482, 1270
Covino S., Landoni M., Sandrinelli A., Treves A., 2020, ApJ, 895, 122
Dark Energy Survey Collaboration, 2016, MNRAS, 460, 1270
de Wolff T., Cuevas A., Tobar F., 2021, Neurocomputing, 424, 49
Dey L. et al., 2019, Universe, 5, 108
Djorgovski S. G. et al., 2011, preprint (arXiv:1102.5004)
Durrande N., Hensman J., Rattray M., N. D. Lawrence, 2016, PeerJ Comput. Sci., 2, e50
Edelson R. A., Krolik J. H., 1988, ApJ, 333, 646
Elorrieta F., Eyheramendy S., Palma W., Ojeda C., 2021, MNRAS, 505, 1105
Foreman-Mackey D., 2016, J. Open Source Softw., 24, 1
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Foreman-Mackey D., Agol E., Ambikasaran S., Angus R., 2017, AJ, 154, 220
Goggans P. M., Chi Y., 2004, in Erickson G. J., Zhai Y., eds, AIPC Ser. Vol. 707, Bayesian Inference and Maximum Entropy Methods in Science and Engineering. Springer, p. 59
Graham M. J. et al., 2015, MNRAS, 453, 1562
Gregory P. C., Loredo T. J., 1992, ApJ, 398, 146
Griffiths R.-R. et al., 2021, ApJ, 914, 144
Hogg D. W., Foreman-Mackey D., 2018, ApJS, 236, 11
Hogg D. W., Villar S., 2021, PASP, 133, 093001
Holgado A. M., Sesana A., Sandrinelli A., Covino S., Treves A., Liu X., Ricker P., 2018, MNRAS, 481, L74
Hu Z., Tak H., 2020, AJ, 160, 265
Huijse P., Estevez P. A., Protopapas P., Zegers P., Principe J. C., 2012, IEEE Trans. Signal Process., 60, 5135
Huijse P., Estévez P. A., Förster F., Daniel S. F., Connolly A. J., Protopapas P., Carrasco R., Príncipe J. C., 2018, ApJS, 236, 12
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Ivezić Ž. et al., 2007, AJ, 134, 973
Ivezić Ž., Connelly A. J., VanderPlas J. T., Gray A., 2014, Statistics, Data Mining, and Machine Learningin Astronomy. Princeton University Press, Princeton and Oxford
Jaynes E. T., Bretthorst G. L., 2003, Probability Theory. Cambridge University Press, New York
Jenkins C., 2014, in Niven R. K., Brewer B., Paull D., Shafi K., Stokes B., eds, AIP Conf. Proc. 1636, 33rd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt 2013). Am. Inst. Phys., New York, p. 106
Joachimi B., Köhlinger F., Handley W., Lemos P., 2021, A&A, 647, L5
Kass R. E., Raftery A. E., 1995, J. Am. Soc. Mass Spectrom., 90, 773
Krishnan S., Markowitz A. G., Schwarzenberg-Czerny A., Middleton M. J., 2021, MNRAS, 508, 3975
Law N. M. et al., 2009, PASP, 121, 1395
Liao W.-T. et al., 2021, MNRAS, 500, 4025
Littlefair S. P., Burningham B., Helling C., 2017, MNRAS, 466, 4250
Lomb N. R., 1976, Ap&SS, 39, 447
Matthews A. G. d. G., van der Wilk M., Nickson T., Fujii K., Boukouvalas A., León-Villagrá P., Ghahramani Z., Hensman J., 2017, J. Mach. Learn. Res., 18, 1
Mondrik N., Long J. P., Marshall J. L., 2015, ApJ, 811, L34
Mortier A., Faria J. P., Correia C. M., Santerne A., Santos N. C., 2015, A&A, 573, A101
Oliphant T., 2006, NumPy: A guide to NumPy. Trelgol Publishing, USA, http://www.numpy.org/
Opper M., Archambeau C., 2009, Neural Comput., 21, 786
Osborne M. A., Roberts S. J., Rogers A., Jennings N. R., 2012, ACM Trans. Sensor Netw., 9, 1
Parra G., Tobar F., 2017, in Guyon I., Luxburg U. V., Bengio S., Wallach H., Fergus R., Vishwanathan S., Garnett R., eds, Advances in Neural Information Processing Systems 30. Curran Associates, Inc., https://proceedings.neurips.cc/paper/2017/file/333cb763facc6ce398ff83845f224d62-Paper.pdf
Pinheiro J. C., Bates D. M., 1996, Stat. Computi., 6, 289
Raftery A. E., 1999, Sociol. Methods Res., 27, 411
Rasmussen C. E., Williams C. K. I., 2006, Gaussian Processes for Machine Learning. MIT Press, Cambridge
Rieger F., 2019, Galaxies, 7, 28
Roberts S., Osborne M., Ebden M., Reece S., Gibson N., Aigrain S., 2012, Phil. Trans. R. Soc. A, 371, 20110550
Saha A., Vivas A. K., 2017, AJ, 154, 231
Scargle J. D., 1982, ApJ, 263, 835
Schwarz G., 1978, Ann. Statist., 6, 461
Sesana A., 2021, Frontiers Astron. Space Sci., 8, 7
Sharma S., 2017, ARA&A, 55, 213
Stothers R. B., Sillanpää A., 1997, ApJ, 475, L13
Takata T., Mukuta Y., Mizumoto Y., 2018, ApJ, 869, 178
Tavani M., Cavaliere A., Munar-Adrover P., Argan A., 2018, ApJ, 854, 11
Tobar F., 2018, in Bengio S., Wallach H., Larochelle H., Grauman K., Cesa-Bianchi N., Garnett R., eds, Advances in Neural Information Processing Systems 31. Curran Associates, Inc., p. 10127, http://papers.nips.cc/paper/8216-bayesian-nonparametric-spectral-estimation.pdf
Tobar F., Bui T. D., Turner R. E., 2015, in Cortes C., Lawrence N. D., Lee D. D., Sugiyama M., Garnett R., eds, Advances in Neural Information Processing Systems 28. Curran Associates Inc., p. 3501, http://papers.nips.cc/paper/5772-learning-stationary-time-series-using-Gaussian-processes-with-nonparametric-kernels.pdf
Trotta R., 2007, MNRAS, 378, 72
Trotta R., 2017, preprint (arXiv:1701.01467)
van der Klis M., 1989, ARA&A, 27, 517
van der Wilk M., Dutordoir V., John S., Artemev A., Adam V., Hensman J., 2020, preprint (arXiv:2003.01115)
van Rossum G., 1995, Technical Report CS-R9526, Python tutorial. Centrum voor Wiskunde en Informatica (CWI), Amsterdam
VanderPlas J. T., 2018, ApJS, 236, 16
VanderPlas J. T., Ivezić Ž., 2015, ApJ, 812, 18
Vanderplas J., Connolly A., Ivezić Ž., Gray A., 2012, in Conference on Intelligent Data Understanding. IEEE, p. 47
Vaughan S., 2010, MNRAS, 402, 307
Vaughan S., 2013, preprint (arXiv:1309.6435)
Vaughan S., Uttley P., Markowitz A. G., Huppenkothen D., Middleton M. J., Alston W. N., Scargle J. D., Farr W. M., 2016, MNRAS, 461, 3145
Verbiest J. P. W., Oslowski S., Burke-Spolaor S., 2021, preprint (arXiv:2101.10081)
Virtanen P. et al., 2020, Nature Methods, 17, 261
Vitale S., 2014, Gen. Relativ. Gravit., 46, 1730
Vousden W. D., Farr W. M., Mandel I., 2016, MNRAS, 455, 1919
Wilkins D. R., 2019, MNRAS, 489, 1957
Wilson A. G., Adams R. P., 2013, PMLR, 28, 1067
Yang S., Yan D., Zhang P., Dai B., Zhang L., 2021, ApJ, 907, 105
Zechmeister M., Kürster M., 2009, A&A, 496, 577
Zhang H., Yan D., Zhang P., Yang S., Zhang L., 2021, ApJ, 919, 58

APPENDIX A: SOFTWARE PACKAGES

We have developed software tools and used third-party libraries all developed with the PYTHON language (van Rossum 1995) (v. 3.7-8-9)[2] with the usual set of scientific libraries (NUMPY (Oliphant 06) (v. 1.15.4)[3] and SCIPY (Virtanen et al. 2020) (1.10).[4] The

[2]http://www.python.org
[3]http://www.numpy.org
[4]https://www.scipy.org

generalized LS algorithm we applied is part of the ASTROPY (v. 3.1.2)[5] suite (Astropy Collaboration 2013, 2018). Non-linear optimization algorithms and numerical integration tools are provided by the MINIMIZE and INTEGRATE subpackages of SCIPY library. MCMC algorithms are provided by the PTEMCEE[6] (v. 1.0.0) library (Foreman-Mackey et al. 2013; Vousden et al. 2016). GP analysis is carried out by the GEORGE package (v. 0.3.1)[7] (Ambikasaran et al. 2016) and by the GPFLOW package (V. 2.2.1)[8] (Matthews et al. 2017). Multi-output GP models have been implemented with the MOGPTK package (v. 0.2.5)[9] (de Wolff et al. 2021). The DCF is computed by the tools in the ASTROML library[10] (v. 0.4.1) (Vanderplas et al. 2012; Ivezić et al. 2014). Plots are produced within the MATPLOTLIB (Hunter 2007) (v. 3.0.2)[11] framework. Multidimensional projection plots were obtained with the CORNER (Foreman-Mackey 2016) (v. 2.0.2)[12] library. Some of the developed code was executed by means of the GOOGLE COLABtoolkit (Bisong 2019).

[5]http://www.astropy.org
[6]https://github.com/willvousden/ptemcee
[7]https://george.readthedocs.io/en/latest/
[8]https://gpflow.readthedocs.io/en/master/index.html
[9]https://github.com/GAMES-UChile/mogptk
[10]https://www.astroml.org/
[11]https://www.matplotlib.org
[12]https://corner.readthedocs.io/en/latest/

This paper has been typeset from a TEX/LATEX file prepared by the author.