Complex variability of *Kepler* AGN revealed by recurrence analysis

R. A. Phillipson [ORCID],[1]★ P. T. Boyd,[2] A. P. Smale[2] and M. S. Vogeley[1]

[1]*Department of Physics, Drexel University, 3141 Chestnut St, Philadelphia, PA 19104, USA*
[2]*Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA*

ABSTRACT

The advent of new time domain surveys and the imminent increase in astronomical data expose the shortcomings of traditional time series analysis (such as power spectra analysis) in characterizing the abundantly varied, complex, and stochastic light curves of Active Galactic Nuclei (AGNs). Recent applications of novel methods from non-linear dynamics have shown promise in characterizing higher modes of variability and time-scales in AGN. Recurrence analysis in particular can provide complementary information about characteristic time-scales revealed by other methods, as well as probe the nature of the underlying physics in these objects. Recurrence analysis was developed to study dynamical trajectories in phase space, which can be constructed from 1D time series such as light curves. We apply the methods of recurrence analysis to two optical light curves of *Kepler*-monitored AGN. We confirm the detection and period of an optical quasi-periodic oscillation in one AGN, and confirm multiple other time-scales recovered from other methods ranging from 5 to 60 d in both objects. We detect regions in the light curves that deviate from regularity, provide evidence of determinism and non-linearity in the mechanisms underlying one light curve (KIC 9650712), and determine realizations of a linear stochastic process describe the dominant variability in the other light curve (Zwicky 229–015). We discuss possible underlying processes driving the dynamics of the light curves and their diverse classes of variability.

Key words: accretion, accretion discs – methods: data analysis – galaxies: active – galaxies: Seyfert.

1 INTRODUCTION

Most of the extreme radiation output of Active Galactic Nuclei (AGN) originates from the accretion discs surrounding the supermassive black holes at their centres, where gravitational potential energy is converted into heat and viscous dissipation. The disc must transport angular momentum outwards, allowing matter to accrete inwards. The radiative flux emitted by the innermost regions of the accretion disc is highly variable across many decades of time, from hours up to months and years (Pica & Smith 1983; Krolik & Begelman 1988). The flux is often assumed to be thermal emission from a geometrically thin disc (Pringle 1981; Abramowicz & Fragile 2013), although a variety of possible accretion disc geometries have been proposed (e.g. slim discs, advective-dominated accretion flows, thick discs; Abramowicz & Fragile 2013). The non-periodic and stochastic variability of the radiation emitted from AGN promises to contain a wealth of information regarding the nature of the accretion flow, from the viscous mechanisms generating dissipation (Balbus & Hawley 1991) to the global geometry of the disc (Shakura & Sunyaev 1973).

Changes in the structure of the accretion flow can result in changes in the bulk variability properties that can be observed with photometric monitoring. A typical AGN accretion disc extends to roughly a few tenths of a parsec from the central supermassive black hole (Goodman 2003). Especially at extragalactic distances,

direct imaging of such objects is a highly arduous and difficult task (Akiyama et al. 2019), infeasible for the study of temporal changes. With the onset of upcoming sophisticated transient-hunting surveys such as the Large Synoptic Sky Telescope (LSST), the imaging of many thousands of objects observed every night, averaging an estimated 20 TB per 24 h period,[1] will surpass a combined decade of imaging data achieved by the Sloan Digital Sky Survey (SDSS)[2] (Ivezić et al. 2019). Acquiring corresponding spectroscopic data on the same scale therefore becomes a near impossibility. Astronomers have therefore invested in time series analysis of light curves as the leading probe of dynamical information across the electromagnetic spectrum of accreting, time-varying objects.

There are multiple theoretical processes that have been proposed that potentially explain the rapid and long-term optical variability of AGN. For example, it is theorized that reprocessing of the central X-ray radiation closest to the black hole (Krolik et al. 1991; Collier et al. 1998; Collier & Peterson 2001), and turbulent or limit cycle thermal processes (Shakura & Sunyaev 1973; Kato 1998) can manifest in oscillations and variability on the short and long term.

There are several methods to empirically translate the theoretical models of the variability into measurable, time-based quantities. For example, the propagating fluctuations model – where the fluctuations

---

★ E-mail: rebecca.a.phillipson@drexel.edu

[1]https://www.lsst.org/scientists/keynumbers
[2]The volume of all imaging data collected over a decade by the SDSS-I/II projects published in SDSS DR 7 (Abazajian et al. 2009) is approximately 16 TB.

in local viscosity have been shown to be driven by the magnetorotational instability (Hogg & Reynolds 2016) – predicts a lognormal distribution of the flux in a light curve and a power spectral density of fluctuations characterized by flicker noise (Lyubarskii 1997).

There are also statistical explanations for the variability. For example, the observed rms–flux relationship (McHardy et al. 2004), which correlates an increase in luminosity with an increase in variability, led to the relationship between the characteristic time-scale of X-ray variability and black hole mass across many decades of mass (Scaringi et al. 2015). The rms–flux relationship is proposed as a consequence of the multiplicative nature of the propagating fluctuations model (Balbus & Hawley 1998), however a clear physical interpretation of the rms–flux relationship is still needed. The statistical damped random walk model (DRW; Kelly, Bechtold & Siemiginowska 2009) predicts a power spectral distribution (PSD) with a power-law slope of −2 (MacLeod et al. 2010) in which some (unknown) mechanism drives random fluctuations and thus injects stochasticity about the mean flux value into the light curve. The physical source of the random fluctuations and stochasticity is also not well understood.

Lognormal flux distributions, rms–flux relationships, and high-frequency PSD slopes of −2 have been consistently recovered in the X-ray for AGN and X-ray Binaries (XRBs; galactic, stellar-mass black holes) alike (e.g. Uttley & McHardy 2001; Edelson et al. 2013). Similarly, in the X-ray bandwidth, a broken power-law model best fits the power spectrum of many XRBs and some AGN, in which the break frequency between the two power-law slopes scales with mass of the object (e.g. Uttley, M C Hardy & Papadakis 2002; Markowitz et al. 2003; McHardy et al. 2006). However, analyses of optical AGN light curves reveal more complex behaviour. Indeed, the rms–flux relationship found in the X-ray and the DRW model does not hold for many of the AGN observed in the optical by *Kepler* (Kasliwal, Vogeley & Richards 2015a; Smith et al. 2018a). Similarly, Moreno, Vogeley, and Richards (in preparation) finds an array of luminosity and rms–variability relationships for AGN observed by both the Sloan Digital Sky Survey (SDSS) and Catalina Real-Time Transient Survey (CRTS) in addition to a variety of PSD slopes (also uncovered by Smith et al. 2018a), which in general are more complex in the optical than in the X-ray. The relationship between the optical and X-ray variability thus remains an open question but could potentially constrain models for the accretion flow.

Commonly used methods for statistical characterization of light curves, such as the autocorrelation function or PSD, measure only the second-order moments of a distribution. By definition, such methods do not capture the higher-order moments, or traces of non-linearity, non-stationarity and direct probes of the nature of the underlying dynamics (Zbilut & Marwan 2008; Moreno et al. 2019). Although, for example, Fourier-based techniques have been the bread and butter of time-domain astronomers and remain some of the most powerful and sophisticated means of characterizing dynamical information from light curves, the abundance of discrepancies in empirical PSD-based measures across bandwidths of light (e.g. the rms–flux relationship) and decades of mass (e.g. the presence of a break frequency, and slopes of the PSD) for accreting systems prompts the pursuit of alternative and complementary analyses. Indeed, there has been success in extracting other types of information about the accretion flow by studying the variability of XRBs and AGN using other methods. For example, recurrence analysis (the study of recurrent, non-periodic information in non-linear dynamics) has been used to distinguish between stochastic, periodic, and chaotic structures underlying the light curves of six microquasars (XRBs exhibiting some of the properties of a quasar) observed in the X-ray

(Suková, Grzedzielski & Janiuk 2016). Topological methods derived from group theory (Gilmore 1998) related to recurrence analysis were used to positively correlate the non-linear light curve of an X-ray binary with the chaotic Duffing oscillator (Phillipson, Boyd & Smale 2018). Statistical analyses using CARMA (Continuous-time Auto-Regressive Moving-Average) applied to large AGN surveys have extracted multiple characteristic time-scales in the optical light curves describing the rate at which flux perturbations grow and decay (Kasliwal et al. 2015b; Kasliwal, Vogeley & Richards 2017; Moreno et al. 2019).

The importance of applying alternative methods, which are well established in other non-astrophysical fields such as statistics, economics, or geology, is two-fold. First, we desire a means to more directly probe the source of the time-scales over which various variability properties dominate and identify the mathematical structure of the equations that describe the underlying physics of the variability. For example, we would expect that random flaring in the accretion disc, local fluctuations in the viscosity or accretion rate, or other inherently random-driven processes that do not result in global, coherent structural changes in the accretion disc would lead to a light curve that is well-modelled by a linear, stochastic system. In contrast, the presence of non-linearity in a light curve identified by techniques from non-linear dynamics would provide evidence for a process that is a global instability (e.g. thermal-viscous instabilities or spiral wave modes; Shakura & Sunyaev 1973; Lightman & Eardley 1974; Wiita 1996), rather than due solely to local fluctuations (e.g. Mchardy 1988; Abramowicz et al. 1992; Edelson & Goddard 1999; Poutanen & Fabian 1999). Although time-scales are important for identifying the possible mechanisms that can exist, correlating specific variability features, such as quasi-periodicity, to a narrower mathematical model (such as non-linearity, or mere determinism), constrains the physical models we construct for accretion disc systems.

Our secondary motive for employing novel time series analysis techniques is to better prepare for the onset of large data sets from upcoming missions such as LSST. Many efforts are already underway classifying variability features (e.g. based on variability statistics or energy spectra) using automated and fast machine learning methods, such as principal component analysis (PCA) or self-organizing maps (Boroson & Green 1992; Francis et al. 1992; Faisst et al. 2019). We aim to add recurrence properties to the list of variability features that can be classified using machine learning techniques. Additionally, an alternative method to PSD modelling for probing universal variability characteristics will provide complementary constraints on current classification models and enable improved predictions for time sampling and baseline requirements for future surveys of active galaxies and other transients.

In this study we will combine the methods of recurrence analysis and topological non-linear analysis to establish a data-driven approach independent of an assumed model or PSD to extract multiple time-scales of interest and evidence for their underlying mechanisms from the well-sampled, multi-year optical light curves of two canonical AGN monitored by *Kepler*. The recurrence plot is the graphical representation of a 2D matrix which contains information about the recurrences (repetitive but not periodic features) present throughout the light curve, after an embedding into a higher-order space representative of the underlying dynamics. The embedding into a higher-order space, called the phase space, is akin to the transformation performed during singular value decomposition (SVD) or PCA, where the flux information from the light curve is recast into a mathematically convenient unit-less matrix, while simultaneously maintaining the same topological information that generated the original light curve. The transformation into phase
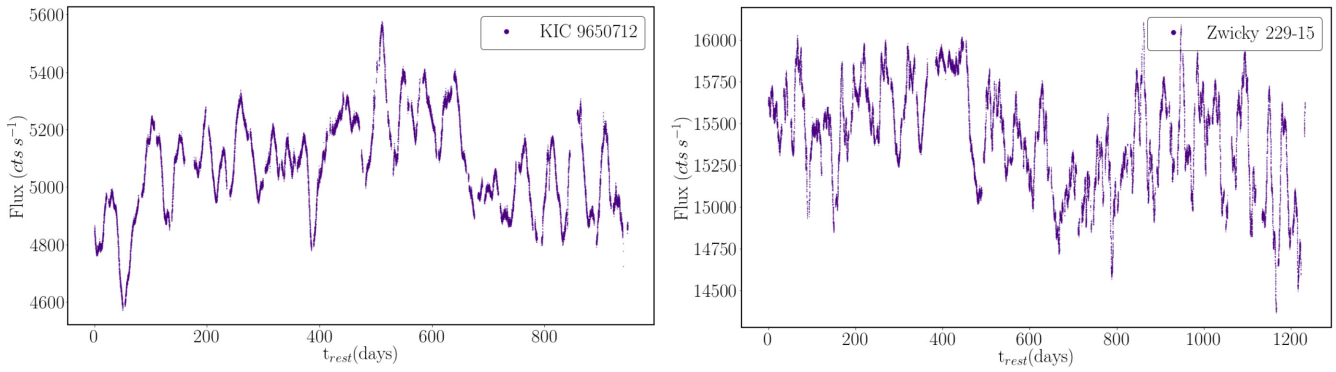
Figure 1. The *Kepler* light curves of (a) KIC 9650712 and (b) Zwicky 229–015, processed by the specialized AGN pipeline by Smith et al. (2018a). The 30 min cadence data points are shown in each light curve, with the typical error on each data point approximately 4 flux counts.

space and the resulting recurrences populate the entries in a matrix, which can be easily plotted into an image called the recurrence plot. The structures in the recurrence plot provide us with topological (dynamical) information about the physical processes that produce the light curve, rather than a merely statistical description.

In Section 2 we describe the *Kepler* satellite and the resulting data it obtained to construct the two AGN light curves. In Section 3 we introduce recurrence analysis and use it to identify three characteristic time-scales, all of which were recovered by multiple authors utilizing different methods for Zwicky 229–015 (Edelson et al. 2014; Kasliwal et al. 2015a,b; Smith et al. 2018a) and one of which was recovered as a low-frequency quasi-periodic oscillation (QPO) for KIC 9650712 (Smith et al. 2018b). We conclude our results with computing a dynamical invariant, the K_2 entropy related to the correlation dimension, as it compares to a series of statistical surrogates using the surrogate data method. The surrogate comparison enables us to possibly distinguish the presence of stochastic and deterministic underlying dynamics in the light curve, which we determine both exist at different horizons and give rise to different time-scales. In Section 4 we explore possible physical mechanisms responsible for the different driving mechanisms that exist in the light curves. We include an appendix with a more detailed overview of recurrence plots and their quantification, collectively called 'recurrence analysis,' as well as the Surrogate Data method used for establishing significance for our results.

2 THE *KEPLER* AGN SAMPLE

Kepler was launched in 2009 and operated for nine years with the scientific objective of exploring the structure and diversity of planetary systems (Borucki et al. 2010). This objective was achieved by searching for repetitive transits in the light curves of extrasolar planetary systems. All stars in the *Kepler* field of view (FOV) were monitored continuously in order to accumulate enough observation time of the transits which only last a fraction of a day. *Kepler* observed ~160 000 exoplanet search target stars with 30 min sampling for approximately 4 yr in the dense target FOV in the region of the sky in the constellations Cygnus and Lyra. Several dozen AGNs were discovered within the *Kepler* FOV (e.g. Mushotzky et al. 2011; Carini & Ryle 2012). The resulting *Kepler* AGN light curves remain the most well-sampled in the optical bandwidth to date.

A specialized pipeline to construct light curves of a sample of *Kepler*-monitored AGN, selected using infrared photometric selection (Edelson & Malkan 2012), X-ray selection from KSWAGS (*Kepler*-Swift Active Galaxies and Stars survey; Smith et al. 2015),

and optical spectroscopy, was developed by Smith et al. (2018a). The Smith et al. sample contains 21 confirmed AGN light curves with a wide range of accretion rates and black hole masses. We have chosen two long-baseline examples (Fig. 1) for which to perform recurrence analysis: the canonical Zwicky 229-015, the longest light curve of an AGN obtained from *Kepler*, and the optical QPO candidate, KIC 9650712. A summary of the physical properties of these two AGN, including masses, luminosities, and accretion rates, are listed in Table 1. We note that the statistical analyses of *Kepler* AGN light curves may be affected by systematics in the *Kepler* data (see e.g. Kasliwal et al. 2015b), which affect the slope of the PSD. Special treatment of the *Kepler* light curves is therefore required (Smith et al. 2018a provides an extended discussion on the various systematics and solutions), particularly when considering periodic or quasi-periodic behaviour. When comparing a re-processed light curve of Zw 229–015, calibrated against ground-based observatories to remove systematics introduced particularly at quarterly intervals, to a systematically contaminated light curve, the time-scale associated with a DRW at approximately 26 d remains unchanged (Kasliwal et al. 2015b). This indicates the presence of intrinsic, non-systematic, variability.[3]

2.1 Zwicky 229–015

There were seven AGNs known to lie in the *Kepler* FOV prior to the satellite's launch in 2009 (Barth et al. 2011; Mushotzky et al. 2011), one of which has been well-studied: KIC 6932990. Otherwise known as Zw 229–015 (and identified this way throughout the remainder of this paper), this AGN is a radio-quiet Type 1 Seyfert at a redshift of 0.0275 (Véron-Cetty & Véron 2003, or VCV catalogue). Edelson et al. (2014) initially found a 5 d characteristic period via power spectrum analysis using the *Kepler* light curve. The Kasliwal et al. (2015a) study recovered a de-correlation time-scale of ~27.5 d extracted from structure function analysis of the *Kepler* light curve and comparison to a DRW model. Using CARMA analysis, the same group (Kasliwal et al. 2017) extracted the previously noted time-scale of 5.6 d (Edelson et al. 2014) and an additional long-term 67 d time-scale; the CARMA model used was the higher-order damped-harmonic oscillator (DHO) perturbed by a coloured noise

---

[3]The analyses in this paper were performed on both the Kasliwal et al. (2015b) and Smith et al. (2018a) Zw 229–015 *Kepler* light curve and the time-scales recovered using recurrence analysis was the same for both, indicating a non-systematic origin to the variability that is characterized in this study.

Table 1. Table adapted from Smith et al. (2018a).

				Physical properties of the *Kepler* AGN				
Object	R.A.	Decl.	z^a	Kep. Mag.^{b}	V Mag.^{a}	$\log M_{BH}$ $(M_\odot)$	$\log L_{Bol}$ ^{c} $(erg\ s^{-1})$	L/L_{Edd} ^{d}
KIC 9650712	19 29 50.490	+46 22 23.59	0.128	16.64	−21.8	8.17^d	45.62	0.226
KIC 6932990	19 05 25.969	+42 27 40.07	0.025	11.13	−19.9	6.91^d, $7.0^{+0.19e}_{-0.24}$	44.11	0.125
(Zw229–015)	–	–	–	–	–	–	–	–

Notes. ^{a}Redshift and *V*-band absolute magnitudes from the VCV Catalogue (Véron-Cetty & Véron 2003)
^{b}Generic optical '*Kepler* magnitude' used in the KIC and calculated in Brown et al. (2011).
^{c}Bolometric luminosities calculated by Runnoe, Brotherton & Shang (2012).
^{d}Mass and Eddington ratio calculations from Smith et al. (2018a).
^{e}Mass based on H β reverberation mapping by Barth et al. (2011).

process. Finally, Smith et al. (2018a) fit a broken power law to the power spectrum of Zw 229–015 and extracted a characteristic 16.0 d break time-scale which, if we take a note from X-ray studies, should theoretically correlate with the black hole mass. In the same study, the best-fitting broken power law extracted a high-frequency PSD slope of −3.4, inconsistent with a DRW model (also found by Kasliwal et al. 2015a), affirming the need for higher-order statistical models such as the CARMA DHO model (which can accommodate a variety of fixed or bending PSD slopes; e.g. Moreno et al. 2019).

The multiple studies of Zw 229–015 and resulting characteristic time-scales, each by contrasting methods and differing calibration techniques, have consequently confused any singular explanation for the physical phenomena driving the intrinsic variability in this and, by extrapolation, other similar systems. We therefore seek to apply recurrence analysis to the *Kepler* light curve in order to determine whether we extract the same time-scales as previous studies using different methods, facilitating a more cohesive picture for the accretion process. Furthermore, we seek to develop a method that can add supporting evidence to the relationship between characteristic time-scales and variability features and the intrinsic physical characteristics of the systems (e.g. mass, luminosity, or accretion rate).

2.2 KIC 9650712

The other object in this study, KIC 9650712, was chosen not for its extensive research history but for the recent discovery of an optical QPO at approximately 44 d (Smith et al. 2018b) and its 'interesting' flux distribution, which is neither lognormal nor singularly Gaussian. The origin of QPOs in X-ray binaries remains largely unknown, but it is believed they are associated with the X-rays emitted near the inner edge of the accretion disc. Low-frequency QPOs (on the order of 0.1 to 10 Hz frequencies) have been associated with different spectral states in black hole XRBs (Stiele et al. 2013), an indication that changes in the accretion flow are connected to the manifestation of QPOs. Possible models that could lead to low-frequency QPOs and changing spectral states in XRBs include Lense-Thirring (Stella & Vietri 1998; Ingram & Done 2010) or orbital (Stella, Vietri & Morsink 1999; Ingram & Done 2012) precession of the accretion disc, spiral structures in the accretion disc (Tagger & Pellat 1999), radiation pressure instability (Janiuk & Czerny 2011), or viscous magneto-acoustic oscillations (Titarchuk & Fiorito 2004). QPOs have primarily been detected in X-rays, with the first X-ray QPO detected in an AGN in 2008 (Gierliński et al. 2008). While low-frequency QPOs are observed in nearly all transient black hole XRBs, there is only one low-frequency QPO AGN detection in the X-ray (Lin et al. 2013). KIC 9650712 marks the first detection of

a low-frequency QPO in an AGN in the optical bandwidth (Smith et al. 2018b). Non-linearity has been confirmed in the X-ray light curves of GRS 1915+105 (Misra et al. 2004) and GX 339–4 (Arur & Maccarone 2019), both XRBs, when a QPO was present. We seek to determine whether the same is true of the low-frequency QPO present in the optical light curve of KIC 9650712.

KIC 9650712 is more massive than Zw 229–015 and, notably, intrinsically brighter with a higher Eddington ratio, as detailed in Table 1. The Kasliwal et al. (2015a) study did perform the same structure function analysis and comparison to a DRW model as with Zw 229–015 and extracted a de-correlation time-scale of ∼48 d. Though the *Kepler* light curve used was only 4 quarters of data (compared to the 14 quarters of Zw 229–015 and the 12 used in this study), Kasliwal et al. (2015a) still noted weak indications of oscillatory behaviour of the shortened light curve. As with Zw 229–015, Smith et al. (2018a) found a broken power-law as the best-fitting model for the PSD of KIC 9650712 and extracted a 53 d break time-scale and a high-frequency PSD slope of −2.9. The method of close returns (Lathrop & Kostelich 1989; Mindlin & Gilmore 1992; Gilmore 1993), a subset of recurrence analysis, is capable of extracting unstable periodic orbits in addition to general recurrent behaviour, appropriate for the detection of quasi-periodic signals, which we will utilize to confirm a QPO in KIC 9650712 in Section 3.2.

3 RESULTS: CHARACTERISTIC TIME-SCALES AND DYNAMICAL INVARIANTS

Many traditional time series analysis methods deal with 1D time series and a statistical description of the variability features. In contrast, methods from non-linear dynamics are critically based on the phase space embedding of a time series – a higher-order space containing the topological information of the time series. The topological structures present in phase space, which most explicitly manifest as recurrences, contain information with direct relationships to the mathematical underpinnings and thus dynamics that generate the time series.

Recurrences that appear in phase space contain all the information about the dynamics of a system and constitute an alternative, and complete, description of a dynamical system (Robinson & Thiel 2009). Recurrence Plots (hereafter RPs) were introduced by Eckmann, Oliffson Kamphorst & Ruelle (1987) as a more general means to visualize the recurrences of trajectories embedded in phase space within dynamical systems. RPs are dynamic graphs which provide qualitative and quantitative information about the behaviour of the system of study, particularly indications of stochastic, periodic, or chaotic behaviour. Several of the quantitative measures, particularly

those based on diagonal-line features, are mathematically equivalent to a variety of dynamical invariants underlying the observed time series (Webber & Zbilut 1994). For an extensive overview of the history of RPs and their applications, see the seminal review by Marwan et al. (2007), from which we also draw many of the definitions in this study. Appendix A contains a more detailed discussion of recurrence plots, phase space, and their quantification specific to our analysis.

Following the notation of Marwan et al. (2007), suppose we have a dynamical system represented by the trajectory x_i for $i = 1, ..., N$ in a d-dimensional phase space. Then the recurrence matrix is defined as

$$\mathsf{R}_{i,j}(\epsilon) = \Theta(\epsilon - ||\vec{\mathbf{x}}_i - \vec{\mathbf{x}}_j||) \text{ for } i, j = 1, ..., N, \qquad (1)$$

where N is the number of time-ordered, measured points $(\vec{x}_i)$, ϵ is a threshold distance, and $\Theta(\cdot)$ is the Heaviside function. For states that persist in an ϵ-neighbourhood, i.e. return to within a threshold distance of a previous state in phase space, the following condition holds:

$$\vec{\mathbf{x}}_i \approx \vec{\mathbf{x}}_j \Leftrightarrow \mathsf{R}_{i,j} = 1. \qquad (2)$$

RPs are thus the graphical representation of the binary recurrence matrix, equation (1), where a colour represents each entry of the matrix (e.g. a black dot for unity and empty or white for zero). By convention, the axes increase in time rightwards and upwards. The RP is also symmetric about the main diagonal, called the 'line of identity' (LOI).

The distances between two positions in a time series are computed after the time series is embedded in phase space. The phase space is a higher-order space that properly resembles the same topological information as the mechanisms generating the time series (rather than merely the statistical information like the mean or standard deviation of the flux). For theoretical systems in which we know the equations of motion, we can directly embed the time series into a phase space constructed from the derivatives of the system. For experimental data in which we perhaps only have a single observable and no immediate knowledge of the equations of motion, we must construct the phase space. Construction of phase space from experimental data is similar to transforming a data set via PCA whereby each component of the new vector retains special information about the time series distinct from its scalar form (but does not necessarily retain the units, for example, of the original time series).

Given that we are dealing with a single observable, the flux, we must construct the multidimensional phase space from the 1D light curve. A commonly used method for construction is via the time delay method (Whitney 1935; Takens 1981), which is an embedding that is a one-to-one mapping to the original attractor that generates the 1D time series without loss of dynamical information (Sauer, Yorke & Casdagli 1991). Other approaches for constructing the higher-order phase space include independent component analysis (Hyvärinen, Hoyer & Inki 2001), singular value decomposition (Broomhead & King 1986), generalized mutual information (Fraser 1989), or numerical derivatives in a differential embedding (if the system is known to be low-dimensional; Gilmore 1998). We use the time delay method for phase space reconstruction in our analysis of KIC 9650712 and Zw 229–015, which uses flux values drawn from the light curve separated by the 'time delay' to construct the higher-order vector. The optimal parameters for the time delay embedding are an embedding dimension of 5 for both objects, and a time delay of 27 and 22 d for KIC 9650712 and Zw 229–015, respectively. For further details of our approach for constructing a phase space embedding based on the time delay method, see Appendix A1.

3.1 The recurrence plot: visualizing structure in the light curves

An example of the recurrence plots of KIC 9650712 and Zw 229–015 for a threshold corresponding to a recurrence rate of 30 per cent is shown in Fig. 2. This means that the threshold (ϵ) is set so that 30 per cent of the state vectors are within ϵ of another, thus 30 per cent of the RP is coloured black. This choice of threshold falls within a range of appropriate values for which the recurrence analysis will probe the same dynamical behaviour (see Appendix A1). The light curve of each object is embedded in a higher-order phase space using the time delay method (see Appendix A1 for details). The distances between every pair of time positions in phase space are computed and if the distance is less than the threshold, then a black dot is plotted (i.e. an entry of one is entered at that matrix position, and zero otherwise). The Euclidean distance metric is used to compute the distance, though any similarity metric could be appropriate. For an in depth discussion of the variety of qualitative features seen for a specific threshold recurrence plot, see Appendix A1. For observational data, and for the computation of invariant measures, it is useful to consider the recurrence plot as a function of threshold. We therefore introduce a colour bar representing a range of thresholds for both KIC 950712 and Zw 229–015 displayed in the upper left-hand panel of Figs 3 and 4, respectively (Zbilut & Webber 1992; Webber & Zbilut 1994). The colour bar in these figures indicates a range in threshold corresponding to a recurrence point density of ~ 1 per cent (purple) up to ~ 99 per cent (orange), which allows for the inspection of the texture of the less recurrent regions of the RP.

There are multiple features of interest that can provide some insight into the types of behaviour present in the light curves of both objects through a qualitative, visual inspection of the RPs. Both objects have RPs (Fig. 2) that display repetitive features vertically (or horizontally, above the LOI) and diagonally (parallel to the LOI) and large white bands and patches (represented by orange in the colour bars of Figs 3 and 4). Marwan et al. (2007) notes that periodic and quasi-periodic systems have RPs with diagonally oriented, periodic, or quasi-periodic recurrent structures, e.g. diagonal lines and 'checkerboard structures', the latter of which is most obvious in the KIC 9650712 recurrence plot. In contrast, vertical structures mark time intervals in which the state of the system evolves slowly (or not at all) and is consequently 'trapped' (Marwan, Thiel & Nowaczyk 2002a); these features are more obvious in the Zw 229–015 light curve.

Single, isolated recurrence points reflect both the observational noise and randomness in the light curve while features that fade with increasing distance from the LOI indicate non-stationarity, and large white (orange) bands or patches indicate abrupt changes in the dynamics of the system (e.g. state changes; Eckmann et al. 1987). By 'non-stationarity' we mean that the underlying dynamics that produces the light curve are experiencing fluctuations or time-invariance in the state parameters of the equations of motion, detectable over the length of the light curve. In Fig. 2, we observe large 'white' patches (regions that are only considered close for a large threshold; or orange in the colourbars of Fig. 3 or Fig. 4), indicating possible non-stationarity or dynamical changes in both light curves. We also note, by comparison, the size of recurrent structures (also called 'texture') in the RPs are different for these two systems, which may indicate stronger higher frequency recurrences or fluctuations in the Zw 229–015 light curve versus KIC 9650712.

The regions that are globally less recurrent in the un-thresholded RP still contain similar local texture to other regions of the RP (Figs 3 and 4), though less distinct. This may indicate that though the light curve is experiencing a variation in the parameters of the underlying system, the nature of the dynamics driving the variability in the light
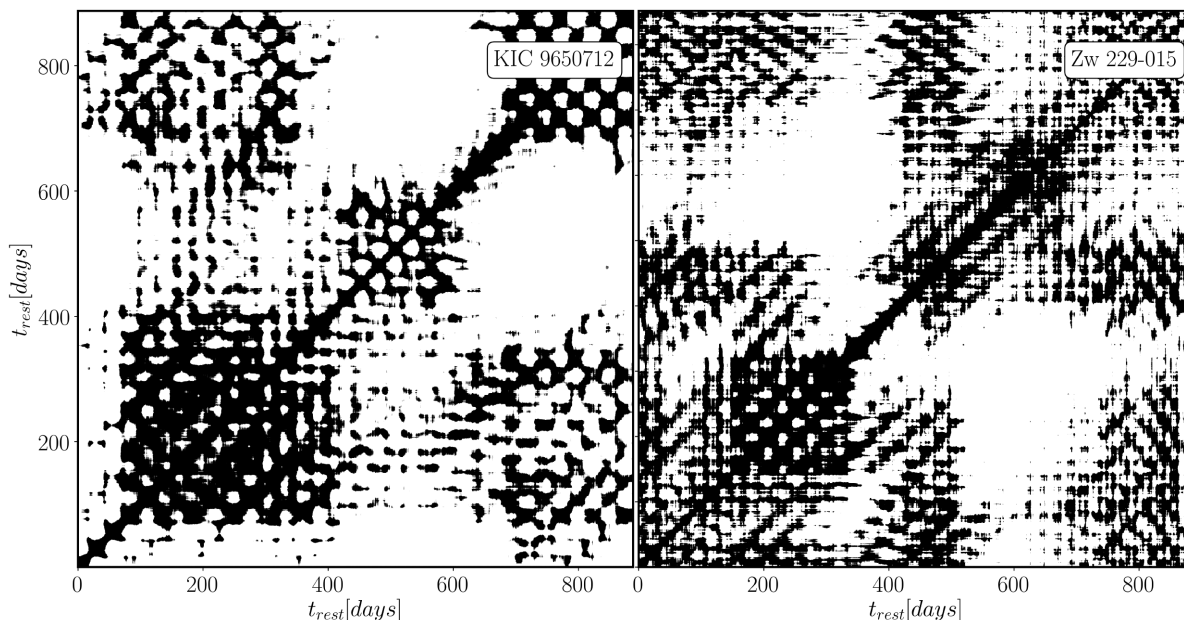
Figure 2. Recurrence plots of KIC9650712 (left-hand panel) and Zwicky 229-015 (right-hand panel) at a recurrence rate (see definition in Section A2) of 30 per cent; axes are on the same time-scale for comparison (though note that the full length of the Zw 229–015 is 200 d longer). A black dot is plotted where the difference in flux at the x and y time positions in the light curve (after embedding in phase space) is less than ϵ. This choice of threshold falls within a range of appropriate values for which the recurrence analysis will probe the same dynamical behaviour (see Appendix A1).

curve do not cease. We will explore the change in recurrence statistics in the KIC 9650712 recurrence plot in Section 3.3.3, where we find the light curve becomes more stochastically driven in the middle of the light curve.

Given the texture of the RPs of both objects, we expect that the light curves contain simultaneous stochastic (or chaotic) and quasi-periodic mechanisms, with the possibility of a dynamics transition particularly notable in the KIC 9650712 RP. The existence of simultaneous periodic and random components has been noted in X-ray binaries (Voges, Atmanspacher & Scheingraber 1987; Boyd & Smale 2004) and the preponderance of one or the other could correlate with intrinsic black hole properties, or specific dominant mechanisms such as a magnetic field (e.g. Suková et al. 2016; Ross, Latter & Tehranchi 2017).

3.2 Line features: quantifying structure in the light curves

3.2.1 Diagonal lines and close returns: recovering an optical quasi-periodic oscillation

The structures in the recurrence plot can be quantified using methods collectively referred to as recurrence quantification analysis, or RQA (see Appendix A2 for a discussion of the variety of RQA measures in more detail; Webber & Zbilut 1994). A variety of RQA measures correlate with specific dynamical invariants, such as Lyapunov exponents (which describe the topological structure of an attractor), dimension, determinism (regions of the time series with high predictability), and laminarity (tendency for regions of a time series to be time-invariant).

RQA measures can also be computed for each diagonal parallel to the LOI of an RP, thus describing recurrences as a function of time lag in the time series (note that this 'time lag' is not the same quantity as the 'time delay' used in creating an embedding). We define the 'recurrence rate' of an RP as the percentage of recurrent points with respect to the total size of the recurrence matrix, which is

of particular interest when studied as a function of time delay in the time series. That is, for those diagonal lines with distance τ (number of time-steps or observations) from the LOI, the τ-recurrence rate is defined as

$$RR_\tau = \frac{1}{N-\tau} \sum_{i=1}^{N-\tau} \mathsf{R}_{i,i+\tau} = \frac{1}{N-\tau} \sum_{l=1}^{N-\tau} l P_\tau(l), \tag{3}$$

where $P_\tau(l)$ is the number of diagonal lines of length l (in time-steps or observations) on each diagonal parallel to the LOI, offset by a distance τ (which, when multiplied by the cadence of the light curve, Δt can be represented in units of time). Here, we do not use a minimum line length to compute RR_τ; all recurrence points and lines are included in the τ-based recurrence rate.

The RR_τ measure can be represented by the so-called 'close returns' histogram, $H(RR_\tau)$, which we abbreviate to $H(\tau)$ for simplicity. The close returns histogram was introduced in non-linear dynamics outside the context of recurrence plots (Lathrop & Kostelich 1989; Mindlin & Gilmore 1992; Gilmore 1993). The name comes from the concept of trajectories in phase space returning 'closely' to a previously visited region in phase space, thus representing unstable periodic orbits (UPOs). These UPOs typically characterize the skeleton of a chaotic attractor and are important for understanding the predictability and periodic phases of a non-linear system (e.g. Boyd et al. 1994; Phillipson et al. 2018). Traditionally, close returns plots and histograms do not necessarily use an embedding, or they use a differential embedding (derived from the numerical derivatives of the time series). The RR_τ measure refers to the recurrence rate at a specific time delay, whereas $H(RR_\tau)$ refers to the full histogram of all recurrence rates as a function of time delay. Thus, the close returns histogram, $H(\tau)$, is identical to a histogram of RR_τ for every time delay, but in this context we use a time delay embedding method for constructing the phase space, rather than the differential embedding. In the view of the close returns histogram, we can observe modalities of specific periods.
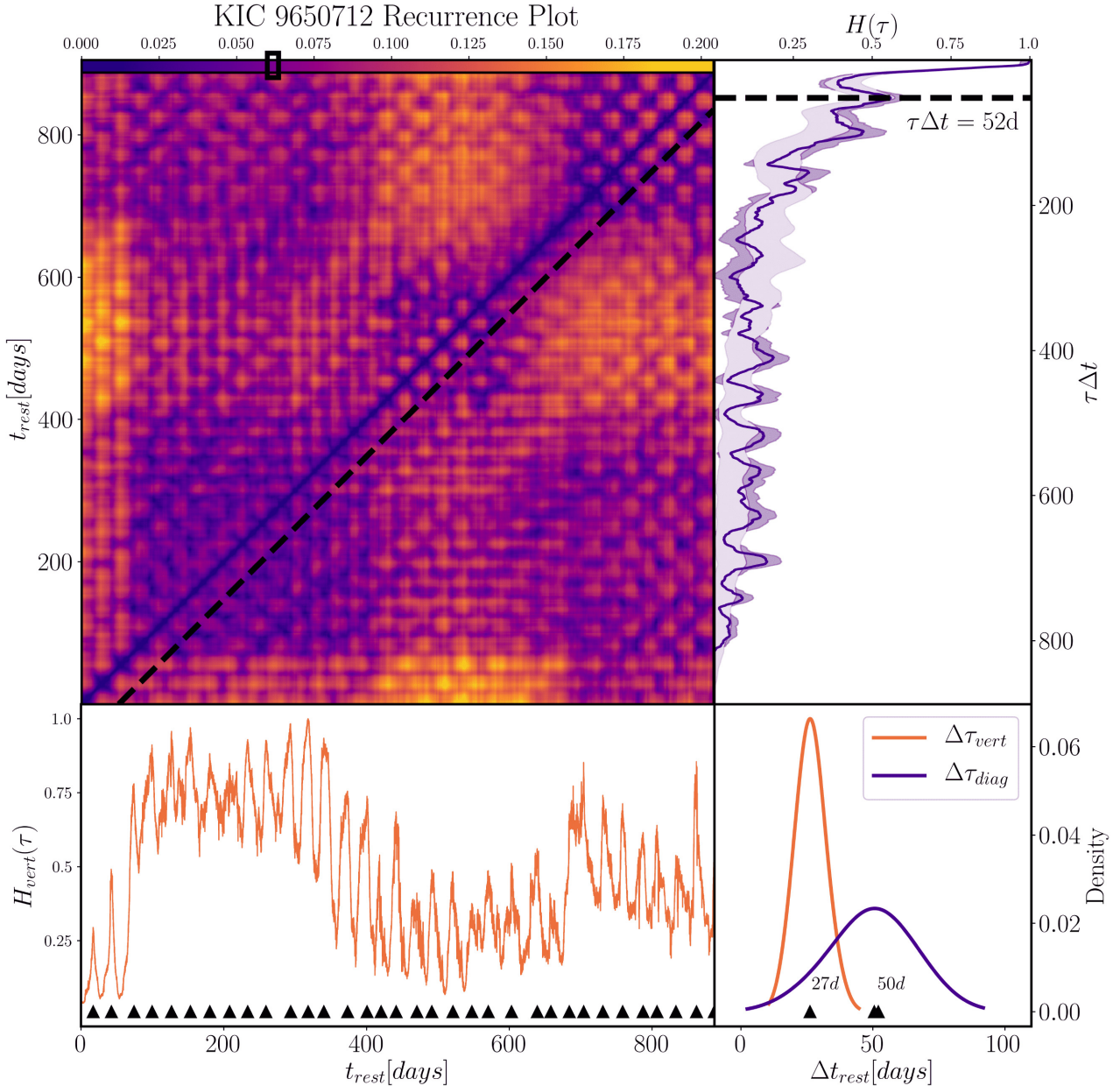
Figure 3. Top left-hand panel: The un-thresholded recurrence plot of KIC 9650712. The colour bar represents the threshold, ϵ, ranging from a corresponding recurrence rate of 1 per cent to 99 per cent, represented in colour by purple (dark) to orange (light). Top right-hand panel: The close returns histogram, $H(\tau)$ of the RP of KIC 9650712 for a threshold corresponding to RR = 30 per cent (marked by the open black rectangle in the colourbar, corresponding to dark purple). This threshold falls within a range of appropriate values as determined by a scaling region that exists between the recurrence rate and threshold when cast in logarithmic space (see Appendix A1). The dashed diagonal line from the RP (left) indicates, as an example, the diagonal that was used to compute the first peak in the close returns histogram aligned with the horizontal dashed line (right), at a time delay of $\tau \Delta t = 52 \pm 2$ d; this period persists with a standard deviation of 9 d for the full histogram. The solid purple (dark) line represents the raw close returns histogram. The regions marked by the purple patches represent the spread in close returns of 100 surrogates with an identical autocorrelation function (ACF) to the data – the dark (light) purple indicates more (less) than two standard deviations away from the ensemble of ACF surrogates, identifying regions of significance in the data's close returns. Bottom left-hand panel: The histogram of vertical lines (each column of the RP) at a threshold corresponding to RR = 30 per cent. The peaks in the vertical lines histogram are spaced on average by 26 d, with a standard deviation of 4 d, which is approximately the average length of a white patch or line in the standard thresholded RP (represented by the orange patches in the un-thresholded recurrence plot above). Bottom right-hand panel: The kernel-density estimation (KDE) of the intervals between successive peaks in the vertical lines histogram (orange), with a peak at 26 d marked by an upward triangle, and close returns histogram (purple), with a peak at 50 d marked by an upward triangle, calculated with the PANDAS package in Python. The first and most significant time delay of $\tau \Delta t = 52$ d found in the close returns histogram is also marked by an upward triangle, which we note is well aligned with the KDE.

The close returns histogram is conceptually similar to a generalized autocorrelation function (ACF) but describes higher-order correlations between the points of the trajectory in phase space as a function of time delay, τ (Marwan et al. 2007). A critical difference between $H(\tau)$ and the ACF is the fact that the close returns is drawn from the recurrence plot after embedding in a higher-order space has occurred – that is, the recurrence peaks in $H(\tau)$ trace the recurrences in the topology of the underlying system, between pairs of values of
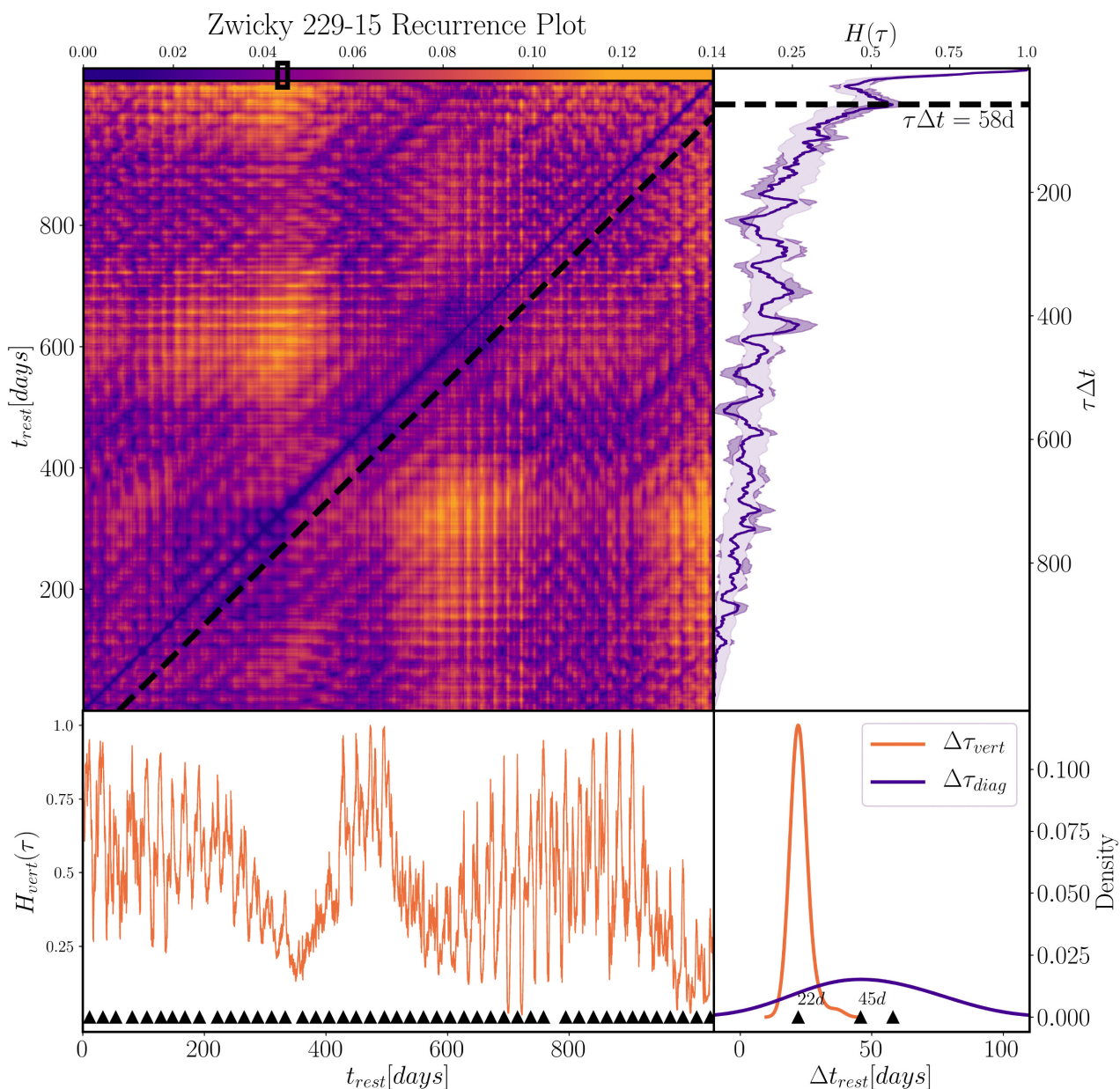
Figure 4. The same as Fig. 3 but for Zw 229–015. The close returns and vertical line histograms are both computed for a threshold corresponding to 30 per cent, marked by the open black rectangle on the colourbar of the un-thresholded recurrence plot. The average interval between peaks in the vertical lines histogram (bottom left) is 22 d (standard deviation: 3 d). The KDE of the intervals between successive peaks in the vertical lines histogram (orange, bottom right) also has a peak at 22 d marked by an upward triangle, and close returns histogram (purple), with a peak at 45 d marked by an upward triangle, calculated with the PANDAS package in Python. The first and most significant time delay of $\tau \Delta t = 58 \pm 2$ d found in the close returns histogram (upper right) is also marked by an upward triangle; the peak separations in the close returns histogram produce a standard deviation of 17 d, indicating the instability of the long-term periods, also evident by the deviation in the KDE.

the time series as a function of time delay. A further advantage of a close returns representation of the data over the (linear) ACF is that it is not an average over an entire sample or single observable but is instead constructed to identify specific, highly correlated segments of data within the time series (Gilmore 1993). This can also be interpreted as the probability that a state recurs to its ϵ-neighbourhood after τ steps. Indeed, any τ-based RQA measure is capable of finding non-linear similarities in short and non-stationary time series with high noise levels, appropriate for astronomical time series, which surpasses the capabilities of standard ACF techniques (Webber et al. 2009).

We plot the close returns histogram against the un-thresholded RP for KIC 9650712 and Zw 229–015 in Figs 3 and 4, respectively. We also construct the close returns of 100 stochastic surrogates (generated as phase-randomized samples from the light curves themselves) that have an identical standard ACF to the original light curve. The stochastic-generated close returns histograms are represented by the spread in light purple. The dark purple patches represent more than two standard deviations away from the ensemble of the ACF surrogates, which we can interpret as regions of significant structure present in the light curves which is not statistically recovered in the ACF. In other words, the fact that there are significant peaks of

the close returns histogram for the data with respect to stochastic surrogates which have an identical ACF demonstrates the additional structure that the close returns histogram uncovers versus a standard autocorrelation function.

The first peak in the close returns histogram represents the strongest recurrent period in the light curve, derived from the diagonal lines in the RP. The first peak is also significant, as it rises above the ensemble of surrogate close returns above the 95 per cent confidence level. We note that the diagonal lines are responsible for periodic and deterministic structure in the time series. For KIC 9650712 (Fig. 3), the first peak in $H(\tau)$ corresponds to a period of $\tau \Delta t = 52 \pm 2$ d, apparently consistent with the QPO detected by Smith et al. (2018b). If we then compute the distance between each successive peak in the close returns histogram, we find the 52 d period persists throughout the entirety of the light curve up to many times this fundamental time delay, with a standard deviation of 9 d. We therefore confirm the Smith et al. (2018b) finding that KIC 9650712 contains a QPO, persisting for long-memory times in the light curve.

In contrast, the Zw 229–015 recurrence plot also contains a long-term period of 58 ± 2 d, extracted from the first peak of the close returns histogram (Fig. 4), but which varies broadly throughout the light curve for large time delays, as can be seen by the flat and wide spread and deviating peak in the kernel density estimation, or KDE, in purple (Fig. 4, bottom right-hand panel) of the close returns peak separations. In contrast, the KDE of KIC 9650712 close returns peak separations in Fig. 3 is aligned with the QPO detection. We therefore conclude that though a long-term period exists in the Zw 229–015 light curve, its period does not remain stable for multiples of this fundamental period (i.e. for long time delays, the memory in the time series decays rapidly, varying with a standard deviation of 17 d) and the underlying mechanism driving the long-term quasi-periodic fluctuations likely does not dominate the light curve or may not be associated with a deterministic mechanism.

3.2.2 Vertical lines and recurrence periods

The vertical line structures within the RP result from the intermittent and laminar states of the time series (Marwan et al. 2002b; Marwan et al. 2007). The average length of a vertical line segment in an RP quantifies the amount of time that the trajectory in a particular state in the underlying system persists, called the 'trapping time,' *TT* (Marwan et al. 2002a). We can also interpret the trapping time as the length of time that fluctuations in an impulse-response system on average persist. For an accretion disc, these fluctuations originate in the accretion flow on a local scale. Similarly, the time that the trajectory needs to recur to the neighbourhood of a previously visited state, or the time between successive fluctuations, corresponds to a white vertical line in an RP (e.g. the gap between successive states; Zou et al. 2007). For example, for periodic motion of period T (perhaps embedded in a noisy signal), we expect a series of uninterrupted diagonal lines separated by a distance T. The vertical distance between successive line segments in the RP, called a 'white' vertical line, will have a length corresponding to T. The period T is often referred to as the recurrence period, $T_{\rm rec}$ (Gao 1999; Gao & Cai 2000), and is distinct from the dominant phase period, $T_{\rm ph}$, which corresponds to the dominant frequency in the power spectrum of a (possibly noisy, observational) time series (Thiel, Romano & Kurths 2003; Marwan et al. 2007).

We estimate $T_{\rm rec}$ through the average white patch length of an RP. The lower left-hand panel in Figs 3 and 4 is the sum of all vertical line segments in the RP in each column (for a specific threshold), including isolated recurrence points (e.g. $v_{\rm min} = 1$). Conceptually, this is similar to the close returns histogram, $H(\tau)$, but the successive peaks in recurrences correspond to trapped, time-invariant states represented by the vertical structure rather than the periodic recurrences represented by diagonal lines. That is, the peaks in the vertical lines histogram directly pinpoint regions in which we have a high frequency of vertical structure, whereas the distance between successive peaks corresponds to the time delays between the laminar states of the system (the average vertical distance between recurrent patches). The kernel density estimation (KDE) of the peak separations in both histograms is displayed in the bottom right-hand panels of Figs 3 and 4, which is much more narrowly isolated for the vertical structure compared to the periods extracted from the close returns. For KIC 9650712 we find the recurrence period to be 26 ± 4 d and for Zw 229–015 we find it to be 22 ± 3 d. We discuss how this time-scale relates to a de-correlation time-scale (e.g. the amount of time for two tangential segments to no longer be correlated), as computed by structure function analysis by Kasliwal et al. (2015a) and extracted from dynamical invariant calculations from the RP (Thiel et al. 2003; Marwan et al. 2007) in Appendix C.

3.3 Distinguishing deterministic versus stochastic mechanisms

3.3.1 K2 entropy: a dynamical invariant measuring complexity

The most important tracer of regular behaviour, including periodic, quasi-periodic, or deterministic behaviour, results from the existence of long diagonal lines in the RP. The longest diagonal line length in an RP is related to the largest Lyapunov exponent (Eckmann et al. 1987); $L_{\rm max}$ is in fact a good indicator of the presence of determinism (Marwan et al. 2007). However, it is the distribution of diagonal line lengths that is directly related to the correlation entropy (also called the Rényi entropy of second order, K_2; Faure & Korn 1998; Thiel et al. 2003), which is defined as the lower limit of the sum of the positive Lyapunov exponents (Ruelle 1978). The positive Lyapunov exponents dictate the rate at which trajectories on an attractor diverge for nearby initial conditions, and the negative Lyapunov exponents determine the boundedness of the attractor (Ott 2002). Chaotic systems contain a positive and finite maximal Lyapunov exponent, λ, resulting in a K_2 entropy that is finite and positive. Perfectly periodic systems have $\lambda = 0$ (and thus entropy is also zero), stable fixed points have $\lambda < 0$ (and thus entropy is negative or undefined), and noise has $\lambda = \infty$, resulting in an infinite entropy (Kantz & Schreiber 2004). Thus, the correlation entropy can be used as a discriminating statistic for probing determinism, periodicity, stochasticity, and chaos.

For complex systems with possible quasi-periodic signals, we would expect a small, finite value for the entropy and for deterministic systems, we expect the entropy to be smaller than its dynamics-free surrogates (e.g. statistically generated time series with the same first- and second-order variability features). The more non-linearity, chaos, or stochasticity present in a system, the larger the value of the entropy. When we compute the correlation entropy of observational data and compare against the entropy calculated from the data's surrogates, we can identify whether dynamical behaviour exists in the data and not in its surrogates and, if so, the nature of the dynamics (e.g. whether it is non-linear or deterministic, or not). In the context of light curves from AGN, the detection of dynamical behaviour that is periodic, deterministic, or non-linear present in the light curve but not in its surrogates, would narrow down the types of mechanisms that generate such behaviour. For example, detection of non-linearity underlying the light curves would rule out models that describe the

source of variability as due to, for example, superimposed linear processes (plus uncorrelated noise) of independent active regions in the accretion disc (e.g. Terrell N. James [1972](#); Gliozzi et al. [2010](#)).

It has been shown that an estimator for the Rényi entropy of the second order, K_2, can be obtained directly from an RP (Thiel et al. [2004](#)) by

$$K_2(\epsilon, l) = \frac{1}{l\Delta t} \log\left(\frac{1}{N^2} \sum_{t,s=1}^{N} \prod_{m=0}^{l-1} \mathsf{R}_{t+m,s+m}\right), \qquad (4)$$

where the quantity within the natural logarithm is the cumulative distribution of diagonal line lengths, $P_\epsilon^c(l)$, l is the length of a diagonal line (in number of data points) and Δt is the time sampling of the time series. In our case, hourly binning of the light curves was used. When we regard $P_\epsilon^c(l)$ as the probability of finding a diagonal line of at least length l in the RP, then the K_2 entropy is related by the approximation

$$P_\epsilon^c(l) \sim \epsilon^{D_2} \exp^{-K_2(\epsilon)l\Delta t}, \qquad (5)$$

where D_2 is the correlation dimension of the system (Thiel et al. [2003](#)). Thus, when we represent $P_\epsilon^c(l)$ in a (natural) logarithmic scale versus line length l we obtain a straight line with slope $-K_2(\epsilon)\Delta t$ for large l's, which is an estimate for the correlation entropy. The K_2 entropy as a function of thresholds, determined by a *RR* ranging from 1 per cent to 99 per cent (see e.g. Asghari et al. [2004](#)) should be monotonically decreasing and result in a scaling region. The scaling region over a range of thresholds provides a more rigorous estimate of the entropy compared to other methods (e.g. Grassberger-Procaccia method; Grassberger [1983](#)) and also accounts for the dynamical and observational noise in the light curves of these physical systems (Thiel et al. [2002](#), [2004](#)). The plateau (scaling region) in the slope of the curves for large l in dependence on ϵ can be found particularly for chaotic and deterministic systems (Thiel et al. [2003](#); Marwan et al. [2007](#)), and is not defined for purely stochastic systems (Thiel et al. [2004](#)). Thus the presence of a scaling region of the entropy with respect to viewing size (threshold) is as important as the value of the entropy for distinguishing between types of dynamical systems (e.g. by the surrogate data method).

To summarize, the correlation entropy describes the number of possible trajectories that the system can follow within l time-steps into the future. That is, the entropy is a proxy for the 'forecasting' time or horizon of the time series, or how well we can reasonably predict the future for $l\Delta t$ amount of time. From this perspective, for periodic systems, where the largest Lyapunov exponent is zero, the entropy is thus also zero, indicating only one possibility for a future trajectory of the system. For increasing entropy, the possible paths that can be taken into the future increases until, for pure white noise, there are infinite possibilities due to the inherent randomness.

We note that for well-sampled data, the lines directly above and below the LOI actually represent tangential motion about the LOI rather than distinct orbits. It is thus best practice to exclude this corridor entirely for the determination of dynamical invariants including the entropy (Gao & Zheng [1994](#)), choosing a width the size of the Theiler window (Theiler [1986](#)), generally comparable to the autocorrelation time. In other words, the entropy will be computed from line lengths that correspond to time-scales longer than approximately 20 d (beyond the de-correlation time-scale, as that found from the vertical lines in Figs [3](#) and [4](#)).

3.3.2 K_2 entropy: a comparison to stochastic surrogates

We will use the method of surrogate data (further discussed in Appendix B; Theiler et al. [1992](#); Small & Tse [2003](#)), to compare the computation of the K_2 entropy for KIC 9650712 and Zw 229–015 against three types of surrogate data sets. Each type of surrogate corresponds to a specific null hypothesis which we compare against using the computation of the entropy. The rejection of a null hypothesis indicates that the light curve is not described by that type of noise process. The three types of surrogates are:

(i) 'Shuffled' surrogates, which represent temporally independent Gaussian noise (e.g. random drawings from the flux distribution). These surrogates preserve the flux distribution of the original data but destroy the time ordering information (Theiler et al. [1992](#)), and thus represent random observations drawn from the same probability distribution of the data.

(ii) 'Phase' surrogates, which represent linearly correlated Gaussian noise, thereby preserving the autocorrelations (and by extension, the PSD) of the original data, but do not maintain the same flux distribution (Theiler & Prichard [1996](#)), and thus contain no non-linear determinism. These can be produced by randomizing the Fourier phases of the light curve.

(iii) The final surrogates are generated using the 'IAAFT' (iterative amplitude adjusted Fourier transform) algorithm, which preserve both the PSD and the flux distribution of the original data (Schreiber & Schmitz [1996](#)) and represent static monotonic non-linear transformations of linear noise.

We will follow the same general approach as introduced by Small & Tse ([2003](#)), and brought to astronomical time series by Suková et al. ([2016](#)) and Asghari et al. ([2004](#)), to compute the K_2 entropy of the data and their surrogates. Comparison to surrogates that are specifically generated by the light curves themselves means that systematics and noise in the light curves will also be imposed on the surrogates. Thus, the stochastic noise in addition to characteristic features present in the light curves will be replicated in the surrogates that are generated. Distinguishing between the noise- and feature-imprinted surrogates and the original light curves will enable us to determine whether deterministic processes are significant above the noise.

We utilize three different software packages for a variety of steps in the analysis. These include the publicly available software package TISEAN[4] (Hegger, Kantz & Schreiber [1999](#); Schreiber & Schmitz [2000](#)) for the production of surrogates; PYRQA[5] (Rawald, Sips & Marwan [2017](#)) for the production of RPs, cumulative diagonal line histograms, and other RQA measures; and finally the Python package PWLF[6] (continuous PieceWise Linear Fit) for the linear fitting of the cumulative histograms. In summary, this approach is as follows:

(i) We use the procedures `mutual` and `false_nearest` from TISEAN to determine the proper time delay embedding parameters for the construction of the recurrence plots of the light curves (see the Appendix A1 for a discussion on embedding parameter selection; also used for generating Figs [3](#)–[4](#)). We note that though the embedding parameters (time delay and embedding dimension) are required for the production of an RP using PYRQA, the results of the computation of the Rényi entropy are independent of these parameters (Thiel et al. [2004](#)).

---

[4]https://www.pks.mpg.de/~tisean/
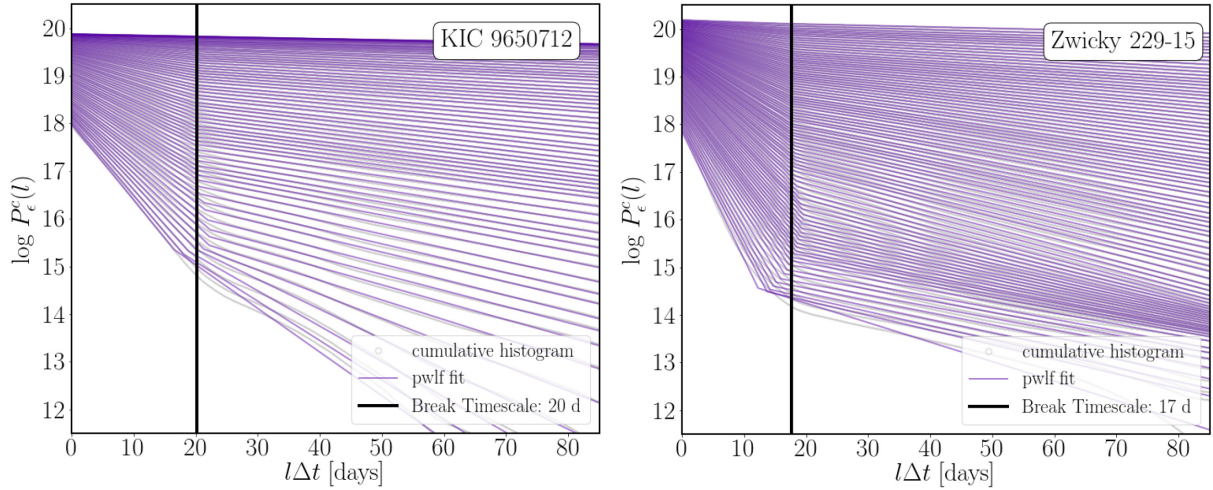[5]https://pypi.org/project/PyRQA/
[6]https://pypi.org/project/pwlf/

Figure 5. The natural logarithm of the cumulative diagonal line lengths histograms as a function of line length for KIC 9650712 (left-hand panel) and Zw 229–015 (right-hand panel). A histogram is plotted for each of the 100 thresholds with corresponding recurrence rates ranging between 1 per cent and 99 per cent. The raw cumulative histogram is plotted in light grey data points, and the purple lines are the piece-wise linear fits of the histograms using the Python package, PWLF. The solid vertical black line is the average position of the break between the two linear scaling regions, at 20 d for KIC 9650712 and 17 d for Zw 229–015. The slope of each purple line above the break time-scale is used in the computation of the K_2 entropy. The histograms are fit out to line lengths corresponding to approximately 10 per cent the length of the time series (Asghari et al. 2004).

(ii) To do the calculation of the entropy, the phase space reconstruction needs to be densely sampled in order to recover a scaling region in the final entropy estimate. This requires a total number of data points ranging from 10 000 to 30 000 (Eckmann & Ruelle 1992). As such, we use the hourly sampling rate for both light curves, which result in light curves of length 23 890 and 27 028 for KIC 9650712 and Zw 229–015, respectively. Though it is possible to compute the entropy using less data points, the confidence of the results would decline.

(iii) Using the TISEAN package, we produce 100 surrogates each of the shuffled, phase and IAAFT types. The surrogate generation algorithms are summarized in detail by Schreiber & Schmitz (2000). Shuffled surrogates are generated through a random shuffling of the original data. The phase surrogates are generated by randomizing the Fourier phases of the original data, but maintaining the amplitude of the complex conjugate pairs, and then performing an inverse Fourier transform. The IAAFT surrogates are generated by iteratively filtering towards the correct Fourier amplitudes and rank-ordering to the correct distribution, in an alternating fashion (i.e. an iterated combination of the first two algorithms).

(iv) For each of the original data of KIC 9650712 and Zw 229–015 and all of their surrogates we produce a recurrence plot for 100 thresholds ranging from ϵ_{min} to ϵ_{max}, corresponding to RR = 1 per cent to 99 per cent, using the PYRQA package. The colourbars in Figs 3 and 4 cover the range of all thresholds used.

(v) For each of the three types of surrogates and the original light curve of each object, we produce a cumulative distribution of diagonal line lengths, $P_\epsilon^c(l)$, for every threshold and use the PWLF package to fit the linear regions in the log $P_\epsilon^c(l)$ versus diagonal line length l plot. Fig. 5 shows the logarithmic plot of $P_\epsilon^c(l)$ and associated line fits for both objects, KIC 9650712 and Zw 229–015, for all thresholds.

(vi) With the line fits and resulting slopes of the cumulative histograms in hand, we compute the K_2 entropy as a function of threshold, ϵ, for all time series. Asghari et al. (2004) determined that the K_2 entropy should be fit by three 'clusters', where the region with the flattest slope represents the optimal estimate for the entropy

(recall, for deterministic and non-linear systems, there is a plateau in the entropy for a range of neighbourhood sizes, but a plateau may not exist if the system is linear stochastic, for example). We use the PWLF package again to fit three regions of the K_2 entropy and choose the smallest slope region as our best estimate for the entropy for every object and each of its surrogates. That is, we use the same threshold range to compute K_2 for all surrogates as we do for the original data, for consistency.

(vii) We compute the significance of the K_2 entropy against each of the surrogate types as a function of ϵ in the following two ways.

(viii) First, we use the standard rank-order test used for most statistical tests in the surrogate data method (see Appendix B; Theiler et al. 1992) to compute the significance of the deviation of the entropy of the data from each of its surrogates as a function of threshold. We select a residual probability α of a false rejection, corresponding to a significance $(1 - \alpha) \times 100$ per cent for a generated $M = K/\alpha - 1$ surrogates. The probability that the data by coincidence has one of the smallest values is exactly α. For our given 100 surrogates, a 95 per cent confidence level that the null hypothesis is rejected would correspond to our data representing one of the five smallest values of the entropy for a given threshold, as we expect purely stochastic systems to have higher entropy.

(ix) Secondly, the distribution of the K_2 entropy as a function of threshold is important for distinguishing non-linearity or determinism – that is, we expect the entropy to be invariant for a range of thresholds resulting in the existence of a plateau in Fig. 6. We therefore consider the cumulative distribution function (CDF) of the entropy for all thresholds for the data and each of the surrogates. The more localized the CDF and offset from the surrogates, the more likely there is an existence of a plateau and a distinction from surrogates. Specifically, a plateau in the entropy as a function of threshold will directly translate to a sharp jump in the CDF, whereas a monotonically decreasing entropy as a function of threshold will translate to a smooth increase in the CDF. Thus, the CDF is an alternative view of the distribution of entropy values for all thresholds. We use the two-sample Kolmogorov–Smirnov (KS) test (Smirnov 1939) using SCIPY to compare the CDF of the entropy
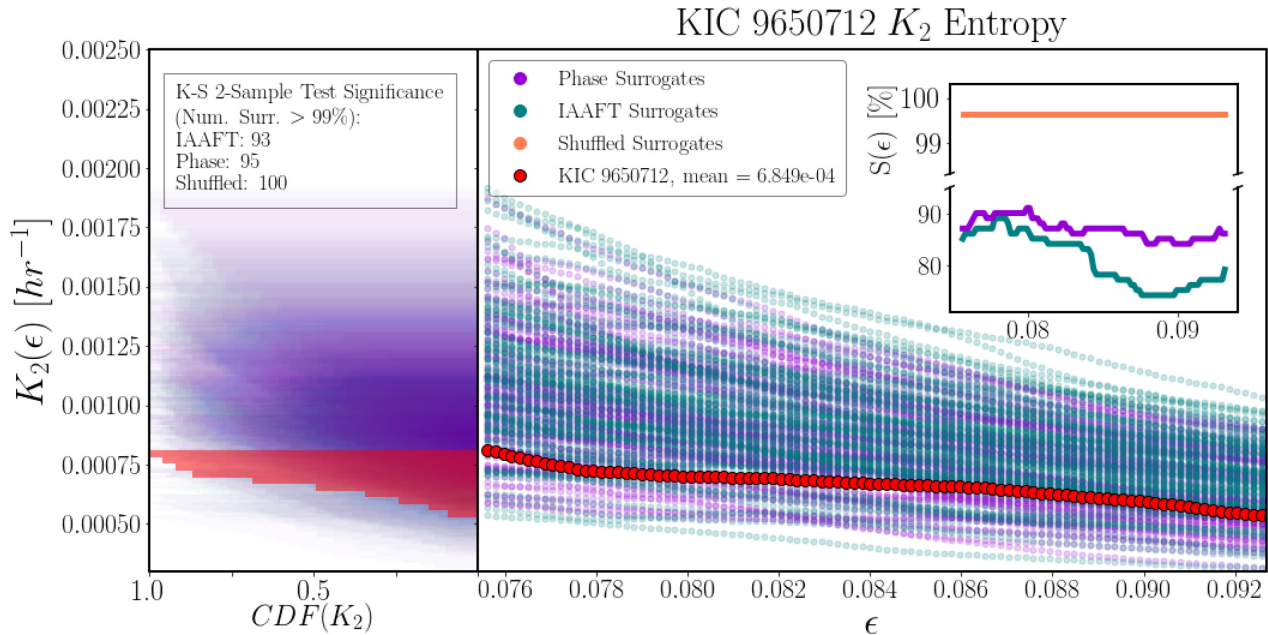
KIC 9650712 K_2 Entropy

Figure 6. The K_2 entropy of KIC 9650712 as a function of threshold, ϵ, for the original data (larger red-filled circles) and the three surrogate types: shuffled in orange, which is so different from the data and other surrogates that is out of the field of view of these plots; phase surrogates are in purple; and IAAFT surrogates are in teal. The mean entropy for KIC 9650712 is 6.849×10^{-4} h^{-1}. The inset in the plot shows the significance of the entropy calculation for the data as a function of threshold for each of the three surrogate types: for each column (threshold), a rank-order test is computed for the data against the surrogates, e.g. if the data (red) is one of the five smallest values including surrogates, it is significant against the surrogates at a 95 per cent confidence level. The shuffled surrogates are highly significant, the phase surrogates peak above a 90 per cent significance for low thresholds, and the IAAFT surrogates peak at 90 per cent significance. The left-hand panel is the cumulative distribution function (CDF) of the entropy for all thresholds for the data (red) and each of the surrogates (orange, purple, and teal as in the right-hand panel). The CDF is the cumulative distribution of the K_2 entropy for all thresholds, a view which incorporates both the median entropy value and how localized this value is for a range of thresholds (e.g. a quantifiable indication of a plateau in the entropy, which will translate to a sharp increase in the CDF, versus a slow increase in the CDF for the absence of a plateau). From the CDF we can determine whether the distribution of K_2 entropy for the data is similar or distinct from those of the surrogates using another statistical test. For each surrogate, a two-sample Kolmogorov–Smirnov test is performed against the data CDF: 12 surrogates total are indistinguishable from the data by the KS test at a 99.9 per cent confidence level – 7 from IAAFT surrogates, 5 from phase surrogates, and none from the shuffled surrogates.

of the data versus each of its surrogates. The two-sample KS test is a two-sided test for the null hypothesis that two independent samples are drawn from the same continuous distribution. The approximate critical value D_α to determine significance of the two-sample KS test is given by the equation

$$D_\alpha = c(\alpha) \sqrt{\frac{n_1 + n_2}{n_1 n_2}}, \qquad (6)$$

where the coefficient $c(\alpha)$, the inverse of the Kolmogorov distribution (Kolmogorov–Smirnov; Kolmogorov & Kolmogorov 1933), for 99.9 per cent confidence is $c(\alpha) = 1.95$ and n_1 and n_2 are the respective number of data points (the slopes as a function of threshold) for the data and each surrogate, which are equal (Smirnov 1939). We expect our data to have a significantly different distribution from all of its surrogates – i.e. the output of the two-sample KS test is greater than D_α – if it contains determinism or non-linearity.

The mean K_2 entropy in the full threshold range is 6.849×10^{-4} h^{-1} for KIC 9650712 and 1.035×10^{-3} h^{-1} for Zw 229–015. We reiterate that the absolute value of the entropy for each object in and of itself has little meaning without comparison to surrogates, since we are dealing with observational systems with inherent noise and systematics (versus theoretical dynamical systems with well-known dimension). Thus, we can only draw conclusions from a comparison of the entropy to different types of surrogates. The significance against the three types of surrogates is higher for

KIC 9650712 than is for Zw 229–015. For KIC 9650712, we see the rank-order test of the entropy as a function of threshold reveals above 90 per cent confidence level of significance against the phase and IAAFT surrogate types for small thresholds and above a 99 per cent confidence level against the shuffled surrogates. When performing the two-sample KS test of the distribution of entropy against all of the surrogate types, 12 total surrogates (none from the shuffled surrogates) were coincidentally similar to the data out of all 300 surrogates – i.e. the difference in distributions constituted a less than 95 per cent level of significance that the null hypothesis is false for only 12 surrogates. We conclude that the entropy for KIC 9650712 is modestly systematically lower than the surrogates, but strongly indicates the presence of determinism. For Zw 229–015, only the shuffled surrogates have a higher than 95 per cent confidence level of significance for both the rank-order test and the two-sample KS test; the phase and IAAFT surrogates never reach a 90 per cent confidence level for low thresholds in the rank-order test, and their distributions in entropy are not significantly different from the data when compared via the two-sample KS test.

The results of the K_2 entropy analysis include:

(i) KIC 9650712 as compared to Zw 229–015 contains more regular (or deterministic) behaviour. This is evident in the fact that there appears to be a plateau in the $K_2(\epsilon)$ plot of KIC 9650712 (Fig. 6) but not in that for Zw 229–015 (Fig. 7), which is what we
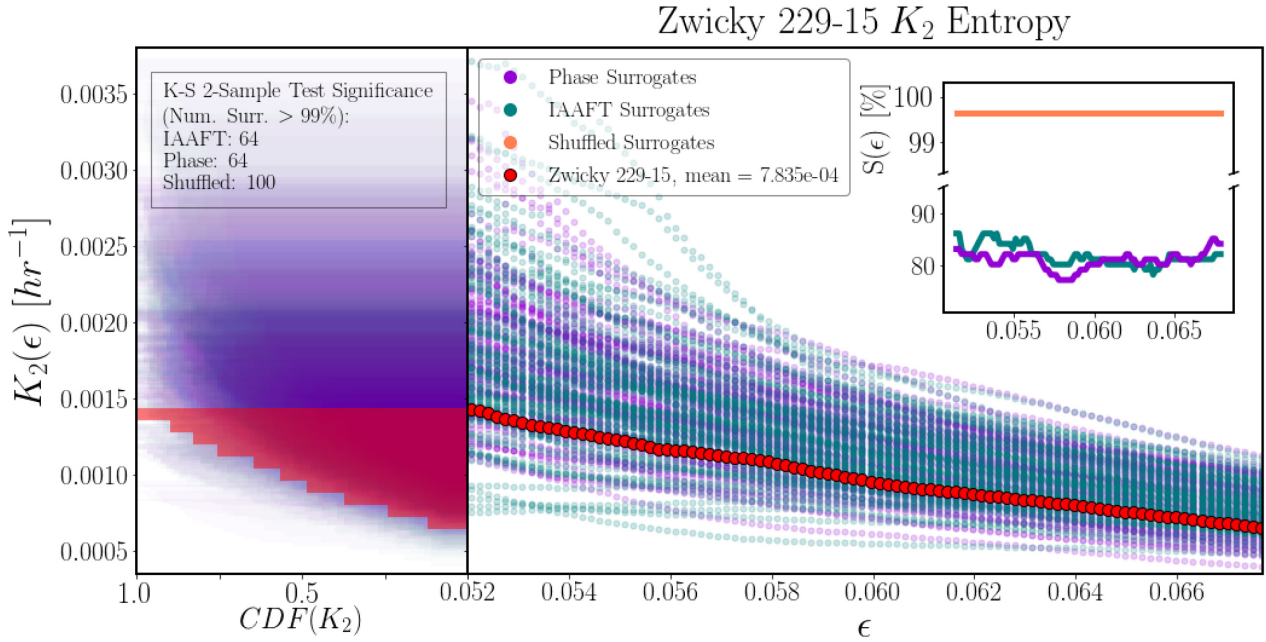
Figure 7. The same as Fig. 6. The K_2 entropy of Zw 229–015 as a function of threshold, ϵ, for the original data (larger red-filled circles) and the three surrogate types: shuffled in orange, which is so different from the data and other surrogates that it is out of the field of view of these plots; phase surrogates in purple; and IAAFT surrogates in teal. The mean entropy for Zw 229–015 is $1.035 \times 10^{-3}\ h^{-1}$, systematically higher than KIC 9650712. The significance of the entropy calculation for the data as a function of threshold for each of the three surrogate types is highly significant for the shuffled surrogates, but well below 90 per cent for both the phase and IAAFT surrogate types. There is no evidence for a plateau in the K_2 entropy. When computing the two-sample KS test of the cumulative histogram of the entropy for the data against each of the surrogates, many are indistinguishable from the data at the 99.9 per cent confidence level.

would expect from deterministic or chaotic systems, but not of linear stochastic ones.

(ii) The rejection of the null hypothesis from the shuffled surrogates for both objects is highly significant. This means we can, possibly unsurprisingly, rule out a temporally independent Gaussian process as a major contributor to the observed variability in both systems. At a minimum, the variability has significant correlations in both objects.

(iii) The null hypothesis of a linear correlated stochastic process (from the phase surrogates) can be likely ruled out for KIC 9650712 – a Gaussian process does not give rise to the variability – but is not significant enough for Zw 229–015. The same is true for a possibly non-linearly rescaled linear stochastic process (from the IAAFT surrogates, where the flux distribution is preserved in addition to the PSD) – non-linearity in the noise response of the KIC 9650712 light curve also does not dominantly contribute to the variability. This means KIC 9650712 either contains non-linear structure in one of the underlying physical mechanisms, or there is significant variation of the state parameters of the underlying system over the length of the light curve (e.g. a possible dynamical state change). Our analysis does not distinguish between non-linearity and non-stationarity of this kind.

The main results of this analysis are that there appears to be an underlying deterministic mechanism in the KIC 9650712 light curve driving variability on time-scales beyond 20 d (based on the comparison of the entropy to surrogate data; Fig. 6), including the QPO (Fig. 3), with the presence of possible non-stationarity or non-linearity in the underlying mechanism. Meanwhile, the variability of Zw 229–015 is indistinguishable from a linear stochastic process, when the entropy is compared to surrogate data (Fig. 7). In the case of Zw 229–015, this means that the light curve can be well

modelled by a typical stochastic process, such as the ARMA or CARMA(2,1) damped harmonic oscillator (Moreno et al. 2019) or similar, in which the linear autocorrelations recover a majority of the observed variability. In contrast, the KIC 9650712 light curve contains simultaneous stochastic (e.g. on time-scales less than the de-correlation time) and deterministic periodic components (e.g. possibly associated with the QPO) and thus we should look towards alternative models, such as non-linear oscillators, to characterize its variability.

It must be pointed out that we chose three specific null hypotheses for which to test the light curves of KIC 9650712 and Zw 229–015. When we reject a null hypothesis, as we did for all surrogate types for KIC 9650712 (though, modestly) and for just the shuffled surrogates of Zw 229–015, we can only state that these particular hypotheses do not represent the data for the specific discriminating statistic that we used (the correlation entropy). Similarly, for Zw 229–015, failure to reject a null hypothesis (e.g. the phase and IAAFT surrogates were not rejected) does not necessarily mean that the null hypothesis is true, only that the correlation entropy failed to be a probe of the differences between the null hypotheses and the data. We also point out that the usefulness of computing the entropy is reserved for well-sampled light curves that capture many cycles of a fundamental period (e.g. about 3 yr light curves for fundamental periods on the order of 50 d) and densely sample the phase space. We address these restrictions in the following section.

3.3.3 RQA: Determinism, time-scales, and transitions

We can confirm the presence of determinism by studying the maximum length of a diagonal line in the RP of each object with respect to each of its surrogates as a function of threshold. As referenced in Section 3.3.1, the longest diagonal line that is present
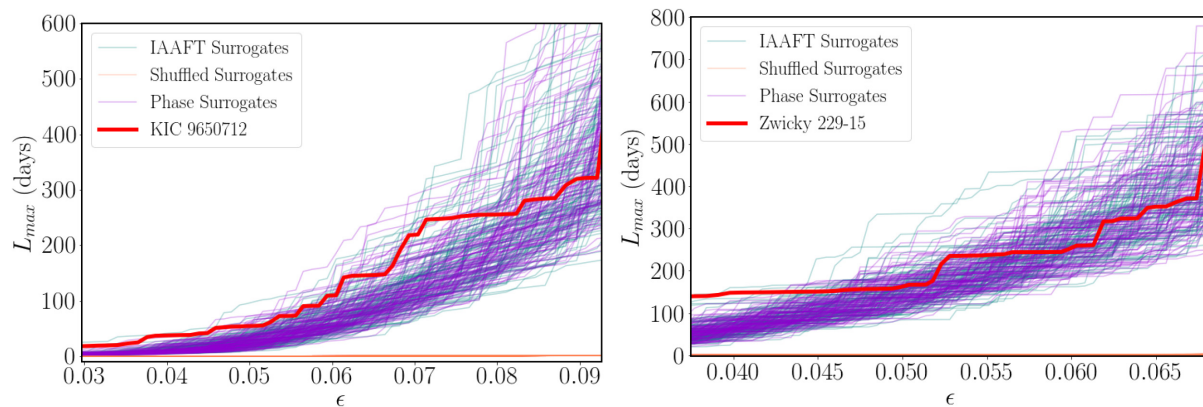
Figure 8. The length of the longest diagonal line found in the RP of KIC 9650712 (left-hand panel) and Zw 229–015 (right-hand panel) as a function of threshold, in a solid red line, against the three surrogate types (shuffled surrogates in orange, phase surrogates in purple, and IAAFT surrogates in teal). The longest diagonal line remains significant against all surrogates for a wide range of thresholds for KIC 9650712, and for only the smallest of thresholds for Zw 229–015. If the longest diagonal line length is significant against various surrogate types, then there is a strong likelihood that the underlying dynamics are deterministic.

in an RP is an indicator for deterministic structure, a calculation from the RP that is computationally much faster than the K_2 entropy and thus can be used for larger samples of objects. The entropy calculation is the most rigorous comparison to surrogates, but does require well-sampled data (on the order of 10 000 to 30 000 data points, such as from *Kepler*) in order to recover a scaling region as a function of threshold (Eckmann & Ruelle 1992) and can be computationally expensive.[7] The usefulness of computing the entropy is thus reserved for well-sampled and long light curves. Shorter and noisier light curves (e.g. less than approximately 1000 observations) would not be suitable with an entropy calculation. However, the standard recurrence quantities derived from the recurrence plot, such as the longest diagonal line, are proxies for the invariants of the light curve such as the entropy. The recurrence quantities can be computed with fewer data points, on the order of ~1000 observations or less (Marwan et al. 2007). In Fig. 8 we see that L_{max} is significant against all surrogates for a wide range of thresholds for KIC 9650712 (L_{max} is systematically longer), but not for Zw 229–015, when considering the well-sampled > 20 000 data point light curves. We therefore confirm the presence of determinism in the KIC 9650712 light curve using a different discriminating statistic from the entropy, and do not conclude determinism is evidenced in the Zw 229–015 light curve. An identical calculation was performed for multiple variability states of six XRBs by Suková et al. (2016), in which L_{max} also distinguished deterministic structure in certain variability states, and did not for stochastic variability states.

If we want an idea of how well-sampled a light curve we need, we can sub-sample the KIC 9650712 light curve at a daily rate, generate the three types of surrogates, and determine the L_{max} measure again for a range of thresholds. In Fig. 9, we see that the longest diagonal line in KIC 9650712 remains significant against the shuffled surrogates, and only marginally significant (just outside one standard deviation on average, and outside two standard deviations only for the smallest thresholds) for the phase and IAAFT surrogates, for a daily sampled light curve (totaling approximately 900 data points). Comparing against different sampling rates of light curves
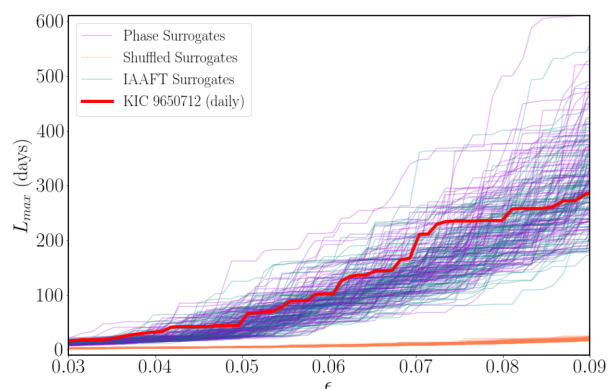

Figure 9. The same as Fig. 8 but for daily observations of KIC 9650712 (sub-sampled from the original light curve, resulting in less than 1000 data points). The L_{max} measure remains significant against both the shuffled and phase surrogates, but less significant against the IAAFT surrogates. We can conclude that sampling rate of a light curve will have an impact on the significance of RQA measures.

allows us to determine what sampling rate and signal to noise is required in order to distinguish the dynamics in a system with recurrence analysis. Using the surrogate data method, though noise present in the original light curve will be imprinted on the surrogates themselves, we may require a high enough sampling rate and time series length in order for the recurrence properties to uncover the dynamics (see e.g. Suková & Janiuk 2016, who compare recurrence properties of various simulated systems with added synthetic noise). Comparing recurrence properties such as L_{max} for an ensemble of daily-monitored AGN is the subject of a subsequent paper.

We can explore other recurrence statistics from the RPs of KIC 9650712 and Zw 229–015 related to characteristic time-scales and indications of transitions in the dynamics. We have recovered two characteristic time-scales from the recurrence plots of KIC 9650712 and Zw 229–015: a quasi-periodic long-term time-scale on the order of 50 d or more from the close returns (Figs 3 and 4), and a de-correlation time-scale from the frequency of vertical lines (also derived from the cumulative distribution of diagonal lines, Fig. 5). Both these time-scales are related to how often a state will recur and the probability of the occurrence of a particular state as a function

---

[7]The full process to calculate the K_2 entropy and significance outlined in Section 3.3.2 takes approximately 25 h on an Intel i7-6700K, 4.00 GHz processor running in serial for a light curve with approximately 24 000 data points and comparing against three types of surrogates, with 100 each.

of time lag. The third, and shortest, time-scale that can be recovered from recurrence analysis relates to how long a state will persist, which can be estimated by the average length of a diagonal line, L_{avg}, and the average length of a vertical line, TT, called the trapping time.

KIC 9650712 has a $TT = 4.9\,\mathrm{d}$ and an average diagonal line length of $L_{\mathrm{avg}} = 2.98\,\mathrm{d}$. Zw 229–015 has a $TT = 4.8\,\mathrm{d}$ and an average diagonal line length of $L_{\mathrm{avg}} = 2.7\,\mathrm{d}$. These values are for the recurrence plot in Fig. 2. We must be careful to exclude those points which are sometimes referred to as 'sojourn points' when calculating measures based on the vertical lines. Sojourn points are those which form vertical lines in the RP but do not represent tangential motion of the phase space trajectory and are thus false recurrences. We estimate the time below which we might have false recurrences in the vertical lines from the average white vertical line length (Gao 1999), also called the recurrence time. For both light curves, this recurrence time is 2 d (corresponding to the recurrence plots in Fig. 2). We can exclude spurious recurrences by setting a minimum line length (i.e. set $v_{\mathrm{min}} > 2\,\mathrm{d}$) when computing TT (or any other vertical line based measure). Similarly, when computing any recurrence properties based on the diagonal lines of the recurrence plot, we must exclude recurrence points that are correlated to one another. Just as we exclude the corridor immediately about the LOI in the RP for computing the K_2 entropy out to a delay corresponding to the de-correlation time, we must also exclude this same corridor when computing the longest diagonal line length and other diagonal-line based quantities (L_{avg}).

We note that a $\sim 5\,\mathrm{d}$ characteristic time-scale was recovered from the Zw 229–015 *Kepler* light curve via power spectrum analysis (Edelson et al. 2014) and from structure function analysis (Kasliwal et al. 2015a), the latter of which indicates that the time-scales at this length may be related to the average persistence time of an impulse fluctuation in an *AR*-type process. Indeed, if we consider how the average line lengths evolve over time, where we can compute L_{avg} and TT in a sliding window across the entire light curve, both L_{avg} and TT vary between 2 and 20 d, with shorter lines in the middle of the light curve. Similarly, if we divide the KIC 9650712 into three segments, compute the recurrence plot and from it the K_2 entropy for each segment separately, we find the entropy is highest (e.g. more noise-like) in the middle of the light curve ($K_2 = 2.78 \times 10^{-3}\,\mathrm{h^{-1}}$) compared to the beginning and end of the light curve ($K_2 = 1.56 \times 10^{-3}\,\mathrm{h^{-1}}$ and $K_2 = 1.72 \times 10^{-3}\,\mathrm{h^{-1}}$, respectively). The change in length of the average lines and longest diagonal line may therefore be quantitative measures for the change in texture in the recurrence plot and thus an analogue for the more computationally intensive entropy. An investigation into windowed recurrence analysis of a set of known state-transitioning X-ray binaries with a comparison to spectra is the subject of a subsequent paper.

4 CONCLUSIONS

The qualitative information that a recurrence plot can provide is in itself useful for distinguishing a variety of time series. However, the structure of recurrences, when quantified, not only indicates the underlying dynamical system but it has been shown that recurrences also contain all the information about the dynamics of a system and constitute an alternative, and complete, description of a dynamical system (Robinson & Thiel 2009). We determine the structure of recurrences using the Recurrence Plot for two Active Galactic Nuclei monitored by the *Kepler* satellite. We first confirm characteristic time-scales of interest identified by other methods, which verifies the validity of using RPs for AGN analysis. We secondly find evidence for low-dimensional determinism in one object (KIC 9650712), and

primarily stochastic realizations of underlying processes in the other (Zw 229–015).

In summary, we find three characteristic time-scales derived from the recurrence plot, which we correlate to three different processes:

(i) Both objects contain a long-term time-scale of $52 \pm 2\,\mathrm{d}$ for KIC 9650712 and $58 \pm 2\,\mathrm{d}$ for Zwicky 229–015. In the KIC 9650712 light curve, this period persists for many multiples of the fundamental period and is consistent with the previously detected optical low-frequency QPO (Smith et al. 2018b). Furthermore, the organization of the diagonal lines in the recurrence plot, which give rise to the $\sim 50\,\mathrm{d}$ time-scale, can be quantified by the correlation entropy. When the entropy is compared to a series of surrogate data, we see evidence that the long-term behaviour in the KIC 9650712 light curve is likely driven by a low-dimensional deterministic, possibly non-linear and/or non-stationary, process. In contrast, in the Zw 229–015 light curve, the period does not persist, but instead decays rapidly with time, and the mechanism determining the long-term variability is likely indistinguishable from a stochastic process.

(ii) A de-correlation time-scale of $26 \pm 4\,\mathrm{d}$ for KIC 9650712 and $22 \pm 3\,\mathrm{d}$ for Zw 229–015 was detected from the frequency of vertical line structures in the RP, which corresponds to the average amount of time between successive variability states in the light curve and indicates the amount of time that must pass before two points in the light curve are no longer correlated (Kasliwal et al. 2015a).

(iii) We determine the average length of a diagonal or vertical line in the recurrence plot is ~ 2–$5\,\mathrm{d}$ for both KIC 9650712 and Zw 229–015 and corresponds to the average length of time that a specific, localized variability state will persist. This can be interpreted as the average amount of time that a localized fluctuation persists in a statistical impulse-response model, as described by autoregressive processes (e.g. Moreno et al. 2019). Furthermore, the lengths of the average diagonal and vertical lines in the RP as a function of time are probes of periodic–chaos/chaos–periodic transitions, chaos–chaos transitions, and changes in laminar states (Marwan et al. 2007).

We conclude that recurrence analysis is capable of recovering time-scales probed by other methods, such as from the power spectrum, autocorrelation function, structure function, or stochastic modelling (Edelson et al. 2014; Kasliwal et al. 2015a; Smith et al. 2018a,b; Moreno et al. 2019).

Furthermore, recurrence analysis is capable of providing evidence for the nature of the underlying processes that produce the light curve related to these characteristic time-scales. We compute an estimate for the dynamical invariant of the Rényi entropy of second order, K_2 (also called the correlation entropy), directly from the recurrence plot of both KIC 9650712 and Zw 229–015 and compare the results to three types of surrogate data, each representing a different stochastic null hypothesis, using the surrogate data method of Theiler et al. (1992). Because the surrogate data are generated from the light curves themselves, the inherent noise processes are also imprinted on the surrogates. Similarly, if there are systematics in the light curve due to the pipeline that produces it or some instrument artefact, these features will also be imprinted on to the surrogates themselves. In fact, the surrogate data method is specifically designed for detecting processes distinct from noise present in the original time series. We determine that the KIC 9650712 light curve is likely driven by a deterministic process, with possible non-linearity or non-stationarity, on the order of many tens of days, while the Zw 229–015 light curve may be well-modelled by a linear, stochastic process in which the linear autocorrelations recover the majority of the observed variability. Though this is a case study of only two objects, we hypothesize that the determinism in the KIC 9650712 light curve is

related to the presence of the QPO, previously detected by Smith et al. (2018b).

Since the development of the surrogate data method, there have been advancements in more rigorous and sophisticated null hypotheses and testing procedures (e.g. as summarized by Lancaster et al. 2018), which may be more suitable for analysing the full 21 *Kepler* AGN sample from Smith et al. (2018a). For example, Moreno et al. (in preparation) finds that AGN observed by SDSS and the CRTS could be well-modelled by two classes of CARMA processes: one is the damped random walk, and the other is a stochastic damped harmonic oscillator. Rejection (or failure to reject) of a null hypothesis based on autoregressive moving average processes would corroborate the results from Moreno et al. and possibly further suggest two classes of the underlying physical process of the light curve. We also point out here that when we say 'non-stationarity', we refer to the time variance of the parameters of the underlying system or of dynamical transitions over the course of the light curve, which can be significant due to size effects. The methods we have utilized in this paper provide evidence for non-linearity, but we have not determined whether the source of the non-linearity is due to non-stationarity of this kind and thus it is possible that a state transition in the AGN light curve was captured over this time period. A windowed recurrence plot analysis would help illuminate whether state transitions have occurred (e.g. as discussed in Section 3.3.3).

QPOs have been uncovered in the X-ray light curves of both XRBs and AGN; the QPO signal in KIC 9650712 represents the first optical detection in an AGN, and its connection to X-ray variability remains unclear. The lack of a confirmation of the rms–flux relationship in the *Kepler* AGN light curves (Smith et al. 2018a), an empirical phenomenon previously detected in the X-ray of AGN, suggests that the propagating fluctuations model for the accretion disc may not be a consistent model for these observations in the optical and similarly the optical variability may not solely be due to reprocessing of the X-ray light from the innermost accretion disc or hot corona (e.g. there may be instabilities arising directly in the optical regions of the accretion disc). Given the deterministic nature of KIC 9650712 and the presence of a QPO, random flaring in the accretion disc or localized fluctuations in the accretion rate are unlikely to be the dominant source of the variability on the order of many days in the KIC 9650712 light curve. Instead, mechanisms capable of producing limit-cycle behaviour and entering a non-linear regime on a global scale must be the primary source of variability at these time-scales for KIC 9650712. Furthermore, that the correlation entropy with respect to stochastic surrogates is more significant for a QPO source is consistent with the results found for six microquasars using the same method (Suková et al. 2016). This suggests that there may be a common accretion mechanism in both XRBs and AGN that leads to QPO behaviour.

The detection of non-linearity alongside QPO signals has occurred in XRBs and microquasars at both short-term time-scales (e.g. seconds and sub-seconds; Suková et al. 2016) and long-term time-scales (e.g. many days; Phillipson et al. 2018). The apparent self-similarity across many decades of time, a hallmark of non-linear and chaotic systems, adds support to the prospect of a non-linear physical mechanism driving variability associated with quasi-periodic behaviour. However, some of the processes proposed for QPOs in XRBs would likely not be detected in AGN on the order of many days as in this study. For example, the radiation pressure instability would occur on the order of thousands of years or more (Janiuk & Czerny 2011), as would precession of the accretion disc connected to jet precession (Lu 1990) or the radiation-driven instability (Petterson 1977; Pringle 1996; Armitage & Pringle 1997).

However, the disc precession models are typically based on the assumption that information is transported through diffusion (Pringle 1997); if instead the propagation of a warp in the accretion disc was transported via wave-like processes, then the speed at which they propagate would be closer to the sound speed, corresponding to variability times much shorter than the viscous diffusion time, a feasible option for ∼50 d time-scales.

Another possible disc precession model is the magnetically driven instability (Aly 1980; Lai 2003), originating from the Bardeen-Petterson effect due to frame-dragging at the innermost edge of the accretion disc linking the spin of the central black hole to the magnetic field of the accretion disc (i.e. Lense-Thirring precession, Bardeen & Petterson 1975). In this case, the optical variability and QPOs would be intrinsically tied to, and possibly phase-shifted from, the X-ray QPOs (Veledina, Poutanen & Ingram 2013). A multi-wavelength study of AGN containing QPOs would be required to confirm this scenario, especially given the unclear manifestations of the rms–flux relationship in the optical.

Assuming the source of the QPO signals that arise in XRBs is the same as that detected in KIC 9650712, we are still left with the puzzle of why a QPO signal would be detected in the optical regime. Low-frequency QPOs on the order of 0.1 to 10 Hz translate to 1–1000 d time-scales for supermassive black holes if the process that produces these signals scales linearly with mass of the black hole and the emitting region corresponding the QPOs is the same. However, these are frequencies detected in the X-ray well within 10 gravitational radii for XRBs whereas the optical radiation from a thin disc about an AGN is expected to arise from at least 100 gravitational radii. Furthermore, the mode by which information is transported from the X-ray to the optical is unknown, but would likely increase those physical time-scales originating within 10 gravitational radii as a function of radius. An outstanding question is therefore whether the process that produces QPOs in different emitting regions of XRB and AGN discs is the same and, if so, how such information is translated between emitting regions.

We reiterate that the two-object sample in this study is clearly not sufficient to make any claims about which of these processes gives rise to the QPO signal in KIC 9650712 and further study of a large sample of AGN light curves, ideally multi-wavelength, is required. We merely hypothesize that the time-scale of the QPO and its moderate significance as a deterministic process with possibly non-linear origin indicates that the mechanisms producing the optical quasi-periodicity may either be due to an inner accretion disc process that propagates outwards (possibly distinct from the rms–flux relationship and propagating fluctuations model), or one that originates in the optical region of the accretion disc and is transported through wave-like processes (in order to occur on days-months periods). In either case the driving mechanism must be capable of operating in a non-linear regime. A sampling of the parameter space of various recurrence quantities with respect to physical characteristics of an ensemble of AGN and XRBs may help illuminate dependencies on physical characteristics of the systems, such as accretion rate and luminosity, and is the subject of a subsequent paper.

DATA AVAILABILITY

This research made use of data reduced and provided by Krista Lynne Smith while at the University of Maryland, originally collected by the *Kepler* mission and made publicly available through the Mikulski Archive for Space Telescopes (MAST). The code used to analyse the data are the publicly available packages TISEAN (Hegger et al. 1999; Schreiber & Schmitz 2000; available at https://www.pks.mpg.de/~tisean) for the production of surrogates; PyRQA (Rawald et al. 2017; available at https://pypi.org/project/PyRQA) for the production of RPs, cumulative diagonal line histograms, and other RQA measures; and finally the Python package PWLF (continuous PieceWise Linear Fit, available at https://pypi.org/project/pwlf). The combination of these packages for use in the analyses in this paper was facilitated by scripts written in Python and will be shared on reasonable request to the corresponding author.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Abramowicz M. A., Fragile P. C., 2013, Foundations of Black Hole Accretion Disk Theory. Springer, Berlin
Abramowicz M. A., Lanza A., Spiegel E. A., Szuszkiewicz E., 1992, Lett. Nature, 356, 41
Akiyama K. et al., 2019, ApJ, 875, L1
Aly J., 1980, A&A, 86, 192
Anishchenko V. S., Vadivasova T. E., Okrokvertskhov G. A., Strelkova G. I., 2003, Phys. A, 325, 199
Armitage P. J., Pringle J. E., 1997, ApJ, 488, L47
Arur K., Maccarone T. J., 2019, MNRAS, 486, 3451
Asghari N. et al., 2004, A&A, 426, 353
Babaei B., Zarghami R., Sedighikamal H., Sotudeh-Gharebagh R., Mostoufi N., 2014, Phys. A, 395, 112
Balbus S. A., Hawley J. F., 1991, ApJ, 376, 214
Balbus S. A., Hawley J. F., 1998, Rev. Mod. Phys., 70, 1
Bardeen J. M., Petterson J. A., 1975, ApJ, 195, L65
Barth A. J. et al., 2011, ApJ, 732, 121
Boroson T. A., Green R. F., 1992, ApJSS, 80, 109
Borucki W. J. et al., 2010, Science, 327, 977
Boyd P. T., Smale A. P., 2004, ApJ, 612, 1006
Boyd P. T., Mindlin G. B., Gilmore R., Solari H. G., 1994, ApJ, 431, 425
Broomhead D. S., King G. P., 1986, Phys. D, 20, 217
Brown T. M., Latham D. W., Everett M. E., Esquerdo G. A., 2011, AJ, 142, 112
Carini M. T., Ryle W. T., 2012, ApJ, 749, 70
Collier S., Peterson B. M., 2001, ApJ, 555, 775
Collier S. J. et al., 1998, ApJ, 500, 162
Eckmann J.-P., Ruelle D., 1992, Phys. D, 56, 185
Eckmann J.-P., Oliffson Kamphorst S., Ruelle D., 1987, Europhys. Lett., 4, 973
Edelson R., Goddard N., 1999, ApJ, 514, 682
Edelson R., Malkan M., 2012, ApJ, 751, 52
Edelson R., Mushotzky R., Vaughan S., Scargle J., Gandhi P., Malkan M., Baumgartner W., 2013, ApJ, 766, 16
Edelson R., Vaughan S., Malkan M., Kelly B. C., Smith K. L., Boyd P. T., Mushotzky R., 2014, ApJ, 795, 2
Faisst A. L., Prakash A., Capak P. L., Lee B., 2019, ApJ, 881, L9
Faure P., Korn H., 1998, Phys. D, 122, 265
Francis P. J., Hewett P. C., Foltz C. B., Chaffee F. H., 1992, ApJ, 398, 476
Fraser A. M., 1989, in Abraham N., ed., Measures of Complexity and Chaos. Plenum Press, New York, p. 117
Fraser A. M., Swinney H. L., 1986, Phys. Rev. A, 33, 1134
Gao J. B., 1999, Phys. Rev. Lett., 83, 3178
Gao J., Cai H., 2000, Phys. Lett. A, 270, 75
Gao J., Zheng Z., 1994, Phys. Rev. E, 49, 3807
Gierliński M., Middleton M., Ward M., Done C., 2008, Nature, 455, 369
Gilmore C. G., 1993, J. Econ. Behavior Org., 22, 209
Gilmore R., 1998, Rev. Mod. Phys., 70, 1455
Gliozzi M., Räth C., Papadakis I. E., Reig P., 2010, A&A, 512, A21
Goodman J., 2003, MNRAS, 339, 937
Grassberger P., 1983, Phys. Lett. A, 97, 227
Hegger R., Kantz H., Schreiber T., 1999, Chaos, 9, 413
Hilborn R., 2001, Chaos and Nonlinear Dynamics: An Introduction for Scientists and Engineers. Oxford Univ. Press, Oxford
Hogg J. D., Reynolds C. S., 2016, ApJ, 826, 40
Hyvärinen A., Hoyer P. O., Inki M., 2001, Neural Comput., 13, 1527
Ingram A., Done C., 2010, MNRAS, 405, 2447
Ingram A., Done C., 2012, MNRAS, 427, 934
Ivezić Ž. et al., 2019, ApJ, 873, 111
Janiuk A., Czerny B., 2011, MNRAS, 414, 2186
Kantz H., Schreiber T., 2004, Nonlinear Time Series Analysis. Cambridge Univ. Press, Cambridge
Kasliwal V. P., Vogeley M. S., Richards G. T., 2015a, MNRAS, 451, 4328
Kasliwal V. P., Vogeley M. S., Richards G. T., Williams J., Carini M. T., 2015b, MNRAS, 453, 2075
Kasliwal V. P., Vogeley M. S., Richards G. T., 2017, MNRAS, 470, 3027
Kato S., Fukue J., Mineshige S., ed., 1998, Black-Hole Accretion Disks. Kyoto Univ. Press, Japan
Kelly B. C., Bechtold J., Siemiginowska A., 2009, ApJ, 698, 895
Kennel M. B., Brown R., Abarbanel H. D. I., 1992, Phys. Rev. A, 45, 3403
Kolmogorov A. N., 1933, Sulla determinazione empirica di una legge di distribuzione. Gornale dell' Istituto Italiano degli Attuari, 4, 83
Krolik J. H., Begelman M. C., 1988, ApJ, 329, 702
Krolik J. H., Horne K., Kallman T. R., Malkan M. A., Edelson R. A., Kriss G. A., 1991, ApJ, 371, 541
Lai D., 2003, ApJL, 591, L119
Lancaster G., Iatsenko D., Pidde A., Ticcinelli V., Stefanovska A., 2018, Phys. Rep., 748, 1
Lathrop D. P., Kostelich E. J., 1989, Phys. Rev. A, 40, 4028
Lightman A. P., Eardley D. M., 1974, ApJ, 187, 1
Lin D., Irwin J. A., Godet O., Webb N. A., Barret D., 2013, ApJ, 776, L6
Lu J., 1990, A&A, 229, 424
Lyubarskii Y. E., 1997, MNRAS, 292, 679
MacLeod C. L. et al., 2010, ApJ, 721, 1014
Markowitz A. et al., 2003, ApJ, 593, 96
Marwan N., Thiel M., Nowaczyk N. R., 2002a, Nonlinear Process. Geophys., 9, 325
Marwan N., Wessel N., Meyerfeldt U., Schirdewan A., Kurths J., 2002b, Phys. Rev. E, 66, 026702
Marwan N., Carmen Romano M., Thiel M., Kurths J., 2007, Phys. Rep., 438, 237
McHardy I., 1988, Mem. Soc. Astron. Ital., 59, 239
McHardy I. M., Papadakis I. E., Uttley P., Page M. J., Mason K. O., 2004, MNRAS, 348, 783
McHardy I. M., Koerding E., Knigge C., Uttley P., Fender R. P., 2006, Nature, 444, 730
Mindlin G. B., Gilmore R., 1992, Phys. D, 58, 229
Misra R., Harikrishnan K. P., Mukhopadhyay B., Ambika G., Kembhavi A. K., 2004, ApJ, 609, 313
Moreno J., Vogeley M. S., Richards G. T., Yu W., 2019, PASP, 131, 63001
Mushotzky R. F., Edelson R., Baumgartner W., Gandhi P., 2011, ApJ, 743, L6

Ott E., 2002, Chaos in Dynamical Systems, 2nd edn. Cambridge Univ. Press, Cambridge

Petterson J. A., 1977, ApJ, 216, 827

Phillipson R. A., Boyd P. T., Smale A. P., 2018, MNRAS, 477, 5220

Pica A. J., Smith A. G., 1983, ApJ, 272, 11

Poincaré H., 1890, Acta Math., 13, 1

Pompe B., 1993, J. Stat. Phys., 73, 587

Poutanen J., Fabian A. C., 1999, MNRAS, 306, L31

Pringle J. E., 1981, ARA&A, 19, 137

Pringle J. E., 1996, MNRAS, 281, 357

Pringle J. E., 1997, MNRAS, 292, 136

Rawald T., Sips M., Marwan N., 2017, Comput. Geosci., 104, 101

Robinson G., Thiel M., 2009, Chaos, 19, 23104

Ross J., Latter H. N., Tehranchi M., 2017, MNRAS, 468, 2401

Ruelle D., 1978, Boletim da Sociedade Brasileira de Matemática, 9, 83

Runnoe J. C., Brotherton M. S., Shang Z., 2012, MNRAS, 422, 478

Sauer T., Yorke J. A., Casdagli M., 1991, J. Stat. Phys., 65, 579

Scaringi S. et al., 2015, Sci. Adv., 1, e1500686

Schreiber T., Schmitz A., 1996, Phys. Rev. Lett., 77, 635

Schreiber T., Schmitz A., 2000, Phys. D, 142, 346

Shakura N., Sunyaev R. A., 1973, A&A, 24, 337

Small M., Judd K., 1998, Phys. D, 120, 386

Small M., Judd K., 1999, Phys. Rev. E, 59, 1379

Small M., Tse C. K., 2002, Phys. D, 164, 187

Small M., Tse C. K., 2003, IEEE Trans. Circuit. Syst. I, 50, 663

Smirnov N. V., 1939, Byull. Mosk. Gos. Univ., i, 2

Smith K. L. et al., 2015, AJ, 150, 126

Smith K. L., Mushotzky R. F., Boyd P. T., Malkan M., Howell S. B., Gelino D. M., 2018a, ApJ, 857, 141

Smith K. L., Mushotzky R. F., Boyd P. T., Wagoner R. V., 2018b, ApJ, 860, L10

Stella L., Vietri M., 1998, ApJ, 492, 59

Stella L., Vietri M., Morsink S. M., 1999, ApJ, 524, 63

Stiele H., Belloni T. M., Kalemci E., Motta S., 2013, MNRAS, 429, 2655

Suková P., Janiuk A., 2016, preprint (arXiv:1604.02670)

Suková P., Grzedzielski M., Janiuk A., 2016, A&A, 586, A143

Tagger M., Pellat R., 1999, A&A, 349, 1003

Takens F., 1981, in Rand D., Young L. S., eds, Lecture Notes in Mathematics. Springer Verlag, Berlin, Heildelberg, p. 366

Terrell N. James J., 1972, ApJ, 174, L35

Theiler J., 1986, Phys. Rev. A, 34, 2427

Theiler J., Prichard D., 1996, Phys. D, 94, 221

Theiler J., Eubank S., Longtin A., Galdrikian B., Farmer J. D., 1992, Phys. D, 58, 77

Thiel M., Romano M. C., Kurths J., Meucci R., Allaria E., Arecchi F. T., 2002, Phys. D, 171, 138

Thiel M., Romano M. C., Kurths J., 2003, Applied Nonlinear Dynamics, 11, 20

Thiel M., Romano M. C., Read P. L., Kurths J., 2004, Chaos, 14, 234

Titarchuk L., Fiorito R., 2004, ApJ, 612, 988

Uttley P., McHardy I. M., 2001, MNRAS, 323, L26

Uttley P., M C Hardy I. M., Papadakis I. E., 2002, MNRAS, 332, 231

Veledina A., Poutanen J., Ingram A., 2013, ApJ, 778, 165

Véron-Cetty M. P., Véron P., 2003, A&A, 412, 399

Voges W., Atmanspacher H., Scheingraber H., 1987, ApJ, 320, 794

Webber C. L., Zbilut J. P., 2007, Int. J. Bifurcation Chaos, 17, 3467

Webber C. L., Marwan N., Facchini A., Giuliani A., 2009, Phys. Lett. A, 373, 3753

Whitney H., 1935, Proc. NAS, 21, 462

Wiita P. J., 1996, in Miller H. R., Webb J. R., Noble J. C., eds, ASP Conf. Ser. Vol. 110, Blazar Continuum Variability. Astron. Soc. Pac., San Francisco

Zbilut J. P., Marwan N., 2008, Phys. Lett. A, 372, 6622

Zbilut J., Webber C., 1992, Phys. Lett. A, 171, 199

Zbilut J. P., Zaldivar-Comenges J.-M., Strozzi F., 2002, Phys. Lett. A, 297, 173

Zou Y., Pazó D., Romano M. C., Thiel M., Kurths J., 2007, Phys. Rev. E, 76, 16210

APPENDIX A: RECURRENCE ANALYSIS: AN OVERVIEW

The concept of recurrent behaviour in time series was first introduced by Poincaré (1890), with visualization of recurrences in the form of Poincaré plots, and with return maps (Hilborn 2001). Recurrence Plots (RPs) were introduced by Eckmann et al. (1987) as a more general means to visualize the recurrences of trajectories embedded in phase space within dynamical systems. RPs provide qualitative information about the behaviour of the system of study, particularly indications of stochastic, periodic, or chaotic behaviour. Measures that quantify the structures present in RPs were introduced by Webber & Zbilut (1994) and subsequently applied to various fields including Mathematics, Geology, and Physiology (Gao & Cai 2000; Marwan et al. 2002a; Zbilut, Zaldivar-Comenges & Strozzi 2002, respectively) and others, where several of the quantitative measures, particularly those based on diagonal-line features, were mathematically equated to a variety of dynamical invariants underlying the observed time series. It is also possible to reconstruct a phase space from 1D observations without loss of dynamical information (Sauer et al. 1991). Recurrences that appear in phase space also contain all the information about the dynamics of a system and constitute an alternative, and complete, description of a dynamical system (Robinson & Thiel 2009).

Recurrence analysis has recently been brought to astrophysics where RPs were utilized in the study of the stability of terrestrial planets (Asghari et al. 2004), in the distinction of chaotic and stochastic behaviour in the X-ray variability of microquasars (Suková et al. 2016), and in the identification of non-linearity in the 14-yr X-ray light curve of an XRB (Phillipson et al. 2018). For an extensive overview of the history of RPs, RQA, and their applications, see the seminal review by Marwan et al. (2007). See also the appendices of Suková et al. (2016) for a very similar and detailed approach to computing the entropy and determining embedding parameters.

RPs are the graphical representation of the binary recurrence matrix, equation (1), where a colour represents each entry of the matrix (e.g. a black dot for unity and empty for zero). By convention, the axes increase in time rightwards and upwards. The RP is also symmetric about the main diagonal, called the 'line of identity' (LOI). The recurrence matrix is computed after the time series is embedded in phase space. In the following sections, we describe phase space and approaches for determining the embedding parameters, followed by the quantification of structure that is seen in the recurrence plot.

A1 Phase space

In order to compute the distances (often determined by the Euclidean metric) between two positions in a time series, which make up the entries of the recurrence matrix, we must first embed the time series in phase space. The most commonly employed embedding is a differential phase space, where the vectors in phase space represent successive derivates of an observable (e.g. position versus velocity). This type of embedding is also the most intuitive, because each component has an obvious relationship to the differential equation that describes the system in question, while also often having a physical interpretation. For example, the differential phase space embedding for a simple pendulum is the angular position versus angular velocity, which traces out a closed curve (e.g. a circle).

For a damped or driven oscillator, the structure in differential phase space becomes more complicated. As we increase the damping and driving parameters, we might see ellipses in phase space that do not close upon themselves, resulting in recurrent but not periodic

behaviour, and possibly entering into non-linear regimes with more complex trajectories. The relative structure of the nearly closed trajectories in phase space are unique to specific equations of motion and constitute a 'topological' perspective of the dynamics. Topological information thus describes how a neighbourhood of points in phase space evolves in time, and how positions in phase space are globally organized, relative to each other (Gilmore 1998). In fact, the way in which a neighbourhood of points is organized topologically is invariant under transformations – topological information such as how various trajectories are organized relative to each other do not change through a variety of different types of embeddings (differential, or otherwise).

Given that we are dealing with a single observable, the flux, we must construct the phase space from the 1D light curve. Furthermore, we do not know, or have an indication of, the dimension of the underlying attractor that generates the light curve and so a differential embedding is not possible. A commonly used method for reconstruction in this scenario is via the time delay method (Whitney 1935; Takens 1981), which reproduces the topological structure of the attractor from a single observable. The point here is that the differential information is invariant under transformations into higher-order spaces, even ones involving a single generic observable (Sauer et al. 1991).

Following notation from Gilmore (1998) (and Phillippson et al. 2018), for the scalar and discrete time series, $x(t)$, like a light curve, we construct vectors $y(t)$ with n components. This involves creating an n vector by the map

$$\mathbf{x}(t) \rightarrow \mathbf{y}(t) = (y_1(t), y_2(t), \ldots, y_n(t))$$
$$y_j(t) = x(t - \kappa_j), \quad j = 1, 2, \ldots, n, \qquad (A1)$$

where κ_j are called the time delays. The time delays are typically evenly spaced multiples of *the* time delay: $\kappa_j = (j - 1)\kappa_{\text{delay}}$. Note that this time delay, κ_{delay}, is a distinct quantity from the time lag introduced in the Close Returns analysis of Section 3.2.

Note that the embedding dimension using the time delay method does not carry physical units (much like the new vectors under PCA do not carry the same literal meaning as the original data). We are no longer dealing with direct differentials of the original time series, even though the newly constructed vectors, $y(t)$, are constructed from flux values in the original light curve. The term 'embedding' by definition means that the map from the space that contains the attractor into a reconstructed phase space is one-to-one and preserves differential information. In particular, Sauer et al. (1991) showed that an attractor with box-counting dimension d can be reconstructed into a new space, R^m, via m time-delayed versions of one generic observation, where $m \geq 2d + 1$ is a requirement. For example, for an attractor that exists in a space with box-counting dimension of 1.4, we would require an embedding dimension of at least 4 in order to unambiguously reconstruct the topological information in a new space, where each component of the new embedded vector consists of values from the original time series spaced by the time delay. We also point out that the requirement on the dimension is an inequality, and so an even higher embedding dimension (e.g. 5 or 6) would also be appropriate and recover the same topological information – indeed, many dynamical invariants are indifferent towards the embedding parameters (Thiel et al. 2004).

Recurrence plots, and other phase space based methods, reveal the dynamical information that generates the time series, but from a topological perspective. And since a time delay construction of phase space is an embedding, each of the m components of the embedded system reflect the global organization in phase space. This concept is readily evident for a simple pendulum embedded in a differential phase space: the points are organized in a circle, reflecting the periodic nature of the system. Recurrences probe this cyclical kind of global organization. An embedding of a pendulum constructed from only the positional information will carry the same cyclical information of the attractor, even with an embedding dimension greater than 2. The usefulness of recurrence plots comes from this critical concept, as topological information is a powerful and direct discriminant of different dynamical systems.

There are a variety of approaches for determining the optimal embedding parameters. For determining the time delay, one can use the first minimum in the mutual information function (Fraser & Swinney 1986; Pompe 1993). Mutual information operates similarly to an autocorrelation function, where the extrema provide us with information about correlations. However, mutual information also contains non-linear correlations, and is thus ideal for determining appropriate time delays when non-linearity might be a factor. Alternatively, one can use the first zero-crossing of the autocorrelation function – the point is to ensure the values we use to construct a new m-dimensional embedded vector are not correlated to one another.

For choosing an appropriate embedding dimension, we seek to reconstruct the attractor such that no two trajectories cross each other, as self-intersections would violate the rule that there are unique solutions to the equations of motion. The false nearest-neighbours algorithm (Kennel, Brown & Abarbanel 1992) is commonly used to determine the embedding dimension appropriate to avoid self-intersections. The false nearest-neighbours algorithm determines the minimal sufficient embedding dimension, m, such that neighbouring points in an embedded phase space represent neighbours in the fully reconstructed attractor. In practice, when we increase the embedding dimension, neighbouring points that then diverge from each other when you increase the dimension are called 'false neighbours.' The algorithm consists of systematically increasing the embedding dimension until we reach a minimal number of false neighbours (theoretically when this value becomes zero; conventionally when less than 10 per cent). The resulting dimension is the minimum embedding dimension required to recover unambiguous topological information according to Takens' theorem (Takens 1981).

Using the mutual information function and the false nearest-neighbours algorithm from the non-linear dynamics software TISEAN, we determine a minimum embedding dimension of $m = 5$ is appropriate for both objects (resulting in less than 10 per cent false neighbours, though we could have used an even higher dimension), and a time delay of 27 and 22 d for KIC 9650712 and Zw 229–015, respectively (note these time delays are comparable to the decorrelation times). Both these and other methods of determining the embedding parameters are reviewed by Marwan et al. (2007). The Takens' approach for embedding a scalar time series via the time delay method is only one approach for embedding. Other embedding methods include performing singular value decomposition (Broomhead & King 1986; Sauer et al. 1991), or independent component analysis (Hyvärinen et al. 2001), which similarly transform time series into other-dimensional spaces.

Once the embedding parameters are determined, one of the required parameters for generating RPs is the threshold, ϵ, which as a rule of thumb should not exceed the maximum phase space diameter of the time series embedded in phase space (Zbilut & Webber 1992), i.e. the threshold should not exceed the maximum size of the reconstructed space.

In practice, given that the observed light curve of an astronomical system is a superposition of a real signal and some observational noise, a method for extracting an optimal threshold to produce a
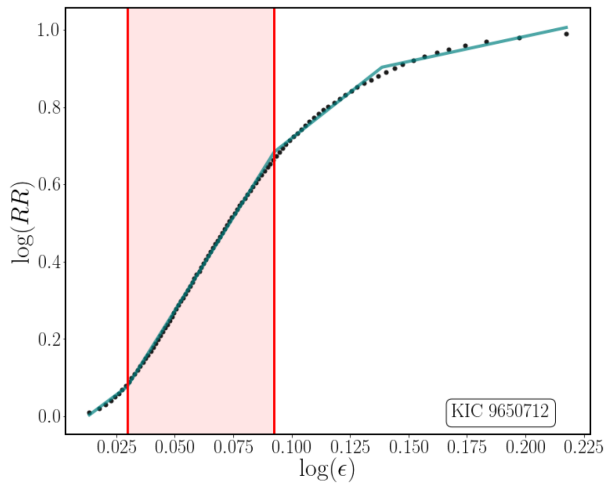
Figure A1. The natural logarithm of 100 thresholds corresponding to evenly spaced recurrence rates ranging from 1 per cent to 99 per cent for the KIC 9650712 light curve. The dotted grey line represents the thresholds. The blue line represents a piecewise linear fit of the thresholds. The red span highlights the linear relationship that exists for thresholds corresponding to recurrence rates between approximately 10 per cent and 70 per cent.

recurrence plot is to exploit the linear scaling relationship that exists between the threshold value and the corresponding recurrence rate. That is, if we produce a recurrence plot for every threshold that corresponds to a range of recurrence rates (called an 'un-thresholded' recurrence plot, as in Figs 3 and 4), there will be a range of values for which there exists a linear relationship between threshold and recurrence rate, when cast in logarithmic space. Any threshold that exists in the linear region of a log–log plot of threshold value versus recurrence rate will equivalently probe the dynamics of the system (Zbilut et al. 2002) without being dominated by noise (too small a threshold) or false-positive recurrences (too large a threshold). The log–log plot of threshold versus recurrence rate for KIC 9650712 is shown in Fig. A1. A recurrence rate of 30 per cent falls in the middle of the linear region of the log–log plot and is used for the recurrence plots of KIC 9650712 and Zw 229–015 shown in Fig. 2. Though small thresholds (e.g. corresponding to less than about 20 per cent recurrence rate) are preferred for determining quantitative measures in the RP, we use the higher recurrence rate in the figures throughout this paper primarily to aid in pointing out the structure present in the RP to the viewer. The significance of the Close Returns analysis in Section 3.2 for both KIC 9650712 and Zwicky 229–015 remains the same for thresholds sampled within the range of approximately 5 per cent recurrence rate up to 70 per cent recurrence rate. A range of thresholds is also required for the determination of dynamical invariants (see entropy calculation in Section 3.3).

A2 Recurrence quantification analysis and dynamical invariants

The structures in RPs have been classified via several measures of complexity known collectively as 'recurrence quantification analysis', or RQA. These measures are based on the recurrence point density, or recurrence rate, defined as

$$RR(\epsilon) = \frac{1}{N^2} \sum_{i,j=1}^{N} \mathsf{R}_{i,j}(\epsilon), \qquad (A2)$$

and the diagonal and vertical line structures and their distributions present in the RP. Here N is the length of the time series and, in

the limit $N \to \infty$, $RR(\epsilon)$ is the probability that a state recurs to its ϵ-neighbourhood in phase space (e.g. a time series that is a straight line, for example, would produce a 100 per cent recurrence rate).

RQA measures link the variety of large-scale patterns to specific behaviours of a system. For example, a Gaussian noise process produces an RP that is dominated by uniformly distributed random, isolated points. As a consequence, there are few uninterrupted lines throughout the RP and we therefore expect that a distribution based on diagonal line lengths, for example, would contain almost exclusively only short or unit-length lines. We can quantify the extent to which a recurrence plot is dominated by uncorrelated or weakly correlated behaviour (e.g. none or very short diagonal lines) through the ratio of recurrence points that form diagonal structures to all recurrence points, called the 'determinism' (DET) of the system (or predictability):

$$\mathrm{DET} = \frac{\sum_{l=l_{\min}}^{N} l P(l)}{\sum_{l=1}^{N} l P(1)}, \qquad (A3)$$

where the lower limit $l_{\min}$ is the minimum length line to consider, typically set to 2 (Babaei et al. 2014), and $P(l)$ is the histogram of diagonal line lengths (for a given threshold).

Another important quantity relates to the lengths of the diagonal lines in the RP, whereby the longest line found, $L_{\max}$, specifically relates to the exponential divergence of the phase space trajectory and, by consequence, the largest positive Lyapunov exponent (Eckmann et al. 1987), a quantity that characterizes an attractor in studies of non-linear dynamics. In fact, the cumulative distribution of diagonal line lengths present in an RP is directly related to the correlation entropy (also known as the Rényi entropy of second order, notated as K_2; Grassberger 1983; Thiel et al. 2003), which is a measure of the complexity and predictability of the system and can be used to distinguish determinism from randomness (Faure & Korn 1998) and estimate various dimensions (Grassberger 1983). A computation of the K_2 entropy is therefore useful for identifying non-linearity from linearity and stochasticity from determinism in a light curve and provides an estimate for the prediction horizon of the time series.

We have mentioned only a few of the RQA measures – the recurrence rate, RR, the determinism, DET, and the longest diagonal line, $L_{\max}$ – that are most often used in the recurrence analysis literature (see the seminal review by Marwan et al. 2007 for detailed definitions and applications). Diagonal lines trace recurrent behaviour in the time series whereby a state revisits itself at a later time. There are also measures based on the vertical lines and structures in an RP, which we discuss in Sections 3.2.2 and 3.3.3 and apply to the KIC 9650712 and Zw 229–015 light curves. For purely periodic dynamics we expect the RQA measures based on vertical structures (such as the longest vertical line, $V_{\max}$) to result in values of zero. In contrast, time series that instead contain regions with slowly changing, or unchanging, states will result in measurable vertical line structures in the RP. A distribution of the vertical line lengths can give indications about the length of time a particular state in a system will persist, or how long it will take to revisit a previous state. The vertical lines can provide information about fluctuations in the light curve, the rate at which they occur and how long they persist, in addition to state transitions (e.g. from chaos to periodicity or laminarity). For context, fluctuations are often related to turbulent mechanisms in the accretion disc.

Finally, the time dependence of RQA measures can be determined by computing these measures in small windows of the RP moving along the LOI. Windowed recurrence plots are particularly useful for detecting non-stationarity (i.e. fluctuations in the state parameters

of the underlying system) and state transitions in the time series (Marwan et al. 2002a). Such transitions could provide hints as to the dominating physical driving mechanisms underlying the light curve, or the times of their onset. For example, when *RR* remains unchanged whereas DET increases, a reorganization of the recurrent points from isolated positions to an assembly of diagonal lines occurs, indicating a transition from noise to periodicity or quasi-periodicity; similarly, the ratio of laminarity (*LAM*, ratio of recurrence points that form vertical structures to all recurrence points) to DET can indicate transitions between dominant localized fluctuations to global, periodic, or otherwise regular behaviour.

Other dynamical invariants, such as the K_2 entropy, based on either the diagonal or vertical line structures in an RP include generalized mutual information (Thiel et al. 2003), the correlation dimension, D_2 (Grassberger 1983), and the point-wise dimension (Gao 1999). The abstract concepts of dimension and entropy (relating to predictability and complexity of a time series) provide information about the family of differential equations that govern a particular system and give rise to the scalar time series of the observables. Translating to astronomy: entropy, dimension, and other invariant measures provide evidence for the types of physical mechanisms that produce the different modes of variability in a light curve beyond merely the time-scales at which they occur.

APPENDIX B: NUMERICAL METHOD USING SURROGATE DATA

Recurrence analysis becomes a powerful distinguishing probe of various stochastic, deterministic, and non-linear features in a time series when combined with the surrogate data method, introduced by Theiler et al. (1992) as an alternative, data-driven hypothesis test. In summary, a set of 'surrogate' data are generated that resemble the original data set. One then tests whether the original data set is a member of the class of dynamical systems that generate the surrogates. The mode by which we generate the surrogate data represents a null hypothesis for the origin of the observed structures in the time series of interest. Once we have an ensemble of surrogate data sets, we can then look for additional structure that is present in the real data and not in the surrogates via a variety of statistical tests (in our case, using recurrence properties) which will either fail to rule out our null hypothesis as a good model for our data, or instead indicate that higher order modes or non-linear mechanisms are responsible for the features we observe. The number of surrogates that we generate dictates the level of confidence in our results (Schreiber & Schmitz 2000).

To employ the surrogate data method, one must choose an appropriate null hypothesis for surrogate generation and an appropriate test statistic. The three algorithms introduced by Theiler et al. (1992) for surrogate generation still prominently used in the literature are based on Monte Carlo Fourier-based re-sampling techniques and generate surrogates that represent (i) independent and identically distributed noise (preserving the probability distribution), (ii) linearly filtered noise (preserving the power spectrum), and (iii) non-linear transformations of linearly filtered noise (preserving both the probability distribution and power spectrum). The methods of surrogate testing were initially introduced as a 'sanity' check for correlation dimension estimation (Small & Judd 1999; Small & Tse 2002; Small & Tse 2003); estimations for correlation entropy and dimension remain a leading choice for the test statistic to distinguish noise and determinism with the surrogate data method (Small & Judd 1998; Asghari et al. 2004; Suková et al. 2016). The test statistic must be independent of the surrogate generation method (thus, computing

the autocorrelation function or its derivatives would not be a good choice for the test statistic for surrogates generated via Fourier re-sampling techniques).

In this study, we follow the same procedure introduced by Small & Tse (2003), where we estimate the correlation entropy (K_2 entropy) directly from the recurrence plot, a computation that is a function of viewing scale, hence $K_2(\epsilon)$. The result is an estimate for the entropy which is a curve, and not a single number, dependent on threshold. The test statistic for the comparison to the surrogate data generated will therefore be the deviation of the entropy computed by the data with respect to the surrogates, which can be interpreted graphically as a function of ϵ. We note here that the full deterministic structure underlying the data can be seen only on length scales smaller than $\epsilon \sim e^{-h}$, where h is the theoretical correlation entropy; above a critical length scale (i.e. for larger thresholds), the data and surrogates look equivalent, and if the critical ϵ is smaller than the noise level (e.g. approximately 5σ of the photometric noise; Thiel et al. 2002), there is no way to distinguish signal from noise. Thus, our estimation of the K_2 entropy will be the most valid for the smallest ϵ values that rise above the noise (Kantz & Schreiber 2004).

The important distinction between the method of surrogate data and other statistical approaches is that it follows a 'constrained realizations' approach (Theiler & Prichard 1996) where we produce simulated time series generated by the original data set itself, imposing the features of interest in addition to observational and dynamical noise on to the surrogates, rather than by fitting a statistical model to the data. Furthermore, if we compare our real data to multiple types of surrogate data representing an array of null hypotheses (e.g. Small & Tse 2003; Suková et al. 2016), we can narrow down the nature of the underlying dynamics producing the time series of our data.

APPENDIX C: RECOVERING τ_{corr} FROM THE DIAGONAL LINE LENGTHS

A characteristic time-scale can be recovered from the frequency of vertical lines in the recurrence plot (Figs 3 and 4) that represents the 'recurrence period,' or the average amount of time between successive states in a time series. It has been shown that the same time-scale can be recovered from the cumulative diagonal line histogram in (natural) logarithmic space (e.g. Fig. 5) or from the autocorrelation function of the time series (Anishchenko et al. 2003; Thiel et al. 2003). For short line lengths, we observe a turnover of the slope in the log $P_\epsilon^c(l)$ plot (Fig. 5), used for the computation of the K_2 entropy. The cross-over time-scale separating the two scaling regions within the cumulative diagonal line distributions traces the de-correlation time-scale extracted from the autocorrelation function (ACF) (Anishchenko et al. 2003). The rate of the ACF decay in differential systems depends essentially on the structure of an attractor (containing the dynamics of the system) and on the influence of noise (Anishchenko et al. 2003). For certain non-hyperbolic attractor types, the autocorrelations decay exponentially. From the ACF, two time-scales can be distinguished, i.e. $\tau \leq \tau_{\rm cor}$ and $\tau > \tau_{\rm cor}$. In the first case the exponential decay is defined by fluctuations of the instantaneous amplitude, and in the second case it depends on the effective diffusion coefficient of the differential system (e.g. or for a Wiener process, the diffusion coefficient).

In the cumulative distribution of diagonal line lengths, the steeper the slope in log $P_\epsilon^c(l)$, the shorter the forecasting time (how far into the future we can reasonably predict the time series), which we would expect for the short-term fluctuations that occur for times $\tau \leq \tau_{\rm cor}$, as it is much more difficult to predict behaviour for times less than

the recurrence period. Meanwhile, for $\tau > \tau_{cor}$ we obtain time-scales that trace the long-term, deterministic dynamics (as also represented by mechanisms driving unstable periodic orbits as probed by the close returns plot) and result in the flatter slope for long line lengths in $\log P_\epsilon^c(l)$ corresponding to the Rényi entropy of the second order, K_2 – in this region, fluctuations in the light curve are no longer correlated to one another. Though the K_2 entropy is specifically defined in the limit of long line lengths, computing the slope for the line lengths in the region $\tau \leq \tau_{cor}$ can give an estimate for the forecasting time, or predictability, at smaller horizons (though it is

not clear this region is related to any kind of dynamical invariant). The de-correlation time separating these two regions can also be related to the turnover time as described by an AR(1) (or DRW) process (Kasliwal et al. 2015a). Indeed, Kasliwal et al. (2017) found a de-correlation time-scale of 27.5 d for Zw 229–015 (the transition in the cumulative distribution of diagonal line lengths for Zw 229–015 occurs at 22 ± 3 d).

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