Deterministic Aspect of the $\gamma$-Ray Variability in Blazars

Gopal Bhatta$^1$, Radim Pánis$^2$, and Zdeněk Stuchlík$^2$

1 Astronomical Observatory of the Jagiellonian University ul. Orla 171 30-244 Kraków, Poland; gopal@oa.uj.edu.pl
2 Research Centre for Theoretical Physics and Astrophysics, Institute of Physics Silesian University in Opava Bezručovo nám.13, CZ-74601 Opava, Czech Republic

Received 2020 July 14; revised 2020 October 26; accepted 2020 October 28; published 2020 December 24

Abstract

Linear time series analysis, mainly the Fourier transform-based methods, has been quite successful in extracting information contained in the ever-modulating light curves of active galactic nuclei, and thereby contribute in characterizing the general features of supermassive black hole systems. In particular, the statistical properties of $\gamma$-ray variability of blazars are found to be fairly represented by flicker noise in the temporal frequency domain. However, these conventional methods have not been able to fully encapsulate the richness and the complexity displayed in the light curves of the sources. In this work, to complement our previous study on a similar topic, we perform nonlinear time series analysis of the decade-long Fermi/LAT observations of 20 $\gamma$-ray bright blazars. The study is motivated to address one of the most relevant queries: whether the dominant dynamical processes leading to the observed $\gamma$-ray variability are of deterministic or stochastic nature. For the purpose, we perform recurrence quantification analysis of the blazars and directly measure the quantities, which suggest that the dynamical processes in blazars could be a combination of deterministic and stochastic processes, while some of the source light curves revealed significant deterministic content. The result, with possible implication of strong disk-jet connection in blazars, could prove to be significantly useful in constructing models that can explain the rich and complex multiwavelength observational features in active galactic nuclei. In addition, we estimate the dynamical timescales, so-called trapping timescales, in the order of a few weeks.

Unified Astronomy Thesaurus concepts: Time series analysis (1916); Time domain astronomy (2109); Blazars (164); Relativistic jets (1390); Gamma-ray transient sources (1853); High energy astrophysics (739); Active galactic nuclei (16); Supermassive black holes (1663)

1. Introduction

Blazars are extragalactic, supermassive black hole systems that display relativistic jet closely pointed toward the Earth. The sources come mainly in two flavors: flat-spectrum radio quasars (FSRQ), the more luminous kind that shows emission lines over the continuum, and BL Lacertae (BL Lac) sources, the less powerful objects that show weak or no such lines. The current and widely accepted models paint a spectacular picture of the blazar systems: as plasma material swirls inward close to the supermassive black holes of the masses in the order $\sim10^9 M_\odot$, the magnetic field in conjunction with the fast rotation of the supermassive black hole contributes to the launching of the bipolar relativistic jets, which then travel up to Mpc scale distance (Blandford & Znajek 1977; Blandford et al. 2019). While the jets plow through the intergalactic medium, any small velocity gradient can lead to formation of shock waves and consequently create favorable condition for the violent episodes giving rise to the large amplitude flares, as observed in the light curves (see Marscher 2016). It is believed that the jet contents could be dominated by the Poynting flux such that the relativistic electrons give rise to synchrotron emission, and the accelerated charged particles upscatter either the population of co-spatial synchrotron photons (Synchrotron Self-Compton model; e.g., Maraschi et al. 1992; Mastichiadis & Kirk 2002) or the lower-energy photons coming from various parts, e.g., accretion disk (AD; Dermer & Schlickeiser 1993), broad-line region (Sikora 1994), and dusty torus (Bläzejowski et al. 2000) (External Compton model). As a result, blazars become dominant sources of high energy emission along with possible extragalactic sources of neutrinos (see IceCube Collaboration et al. 2018a, 2018b).

Flux variability in diverse temporal and spatial frequencies is one of the defining and fascinating properties of blazars (e.g., $\gamma$-ray; Rajput et al. 2020, X-ray; Bhatta et al. 2018, optical; Bhatta & Webb 2018). The power spectral density analysis reveals that the statistical nature of the blazar $\gamma$-ray variability can be well described by power-law-type noise with mostly single power-law index (Bhatta & Dhital 2020, and the references therein). A number of models attempt to explain the variability linking its origin to various mechanisms, e.g., magnetohydrodynamic instabilities in the jets (e.g., Bhatta et al. 2013; Marscher 2014), shocks traveling down the turbulent jets (e.g., Marscher & Gear 1985; Böttcher & Dermer 2010), magnetic reconnection in the turbulent jets (Sironi et al. 2015; Werner et al. 2016), and relativistic effects due to jet orientation (e.g., Camenzind & Krockerberger 1992; Raiteri et al. 2017). However, working out the exact details of the underlying processes has been part of ongoing research.

In general, the time series analysis, mostly power spectrum density analysis, are treated as one of the most powerful tools in characterizing the statistical nature of the observed variability. However, usage of such analyses are limited to the second-order moments of the flux distribution and static properties of the light curves. Consequently, the methods fail to incorporate the information about the inherent nonlinearity and non-stationarity, which are contained in the higher-order moments and which directly reflect into the dynamical nature of the black hole systems (see Green et al. 1999; Zblut & Marwan 2008; Shoji et al. 2020). Moreover, the attempts to constrain the observed variability in the blazar within the framework of linear stochastic systems probe into the randomly occurring flaring episodes, such as local fluctuations in the viscosity, accretion rate at the AD, and/or stochastic shock
events prevailing in the jet regions. Such linear stochastic changes are not likely to affect global perturbations, which ultimately materialize in the observed flux changes in the sources. On the other hand, the observational features such as rms-flux relation and log-normal flux distribution (e.g., Uttley et al. 2005; Bhattacharyya et al. 2020; Bhatta & Dhillon 2020) point to the nonlinear dynamics inherent in the disk-jet systems, and therefore explore the processes that lead to global perturbations giving rise to the instabilities that persist and remain coherent over the entire system (however, for a shot noise interpretation of such observations, see Scargle 2020). Studies of black hole systems taking the nonlinear time series approach to the AGN light curves can be found in several works (e.g., Leighly & O’Brien 1997; Bachev et al. 2015; Phillipson et al. 2020; Shoji et al. 2020). Besides, the nonlinear time series analysis can be used to distinguish sources that have similar sets of nonlinear properties, as well as to measure characteristic timescales, e.g., trapping timescales, which represents the average time a system spends on a particular state (see Marwan et al. 2002).

More important, the query of whether the basic nature of the variability should be treated as stochastic or deterministic stands out as one of the most relevant questions to be asked (see Kiehlmann et al. 2016; in the context of microquasars, see Suková & Janiuk 2016). The answer to such queries has far-reaching impact in our attempts to constrain that physical process that leads the multi-timescale variability, e.g., the physical conditions prevailing over the innermost regions of blazar jets, the nature of the dominant particle acceleration and energy dissipation mechanism, magnetic field geometry, jet content, etc. It is most likely that the roots of variability phenomenon can be related to the nonlinear magnetohydrodynamical flows at the accretion-jet systems, which are governed by the combined effect of the ambient magnetic field and the rotation of the innermost regions around the supermassive black holes. In such a scenario, nonlinear time series analysis estimating the changes in the dynamical states of the system can contribute to establishing a strong connection between the AD and the jet in radio-loud AGN systems (see Bhatta et al. 2018 for the observational signature of the disk-jet connection).

In this work, we carry out nonlinear time series analysis of 20 blazars utilizing decade-long Fermi/LAT light curves presented in our previous work (see Bhatta & Dhillon 2020). In Section 3, the details of the analysis, in particular, recurrence quantification analysis (RQA), which provides various measures including determinism, predictability, and entropy, are discussed. In addition, the results of the analyses on the γ-ray light curves are also presented. Discussion of the results, along with their possible implications for the nature of γ-ray emission from the sources, are presented in Section 4, and the conclusions of the study are summarized in Section 5.

2. Source Sample

The source sample consists of 20 γ-ray bright blazars, such that weekly binned light curves can be constructed.3 The sources are listed in Column 1 of Table 1 along with their positions in the sky, R.A. (Col. 2) and decl. (Col. 3); 3FGL catalog names (Col. 4); source class (Col. 5); and their redshifts as listed in NED.4

3. Analysis

To further explore the nature of variability in γ-ray light curves of the sample blazars, we adopted a number of approaches to the chaos study. The description of the methods and the corresponding results of the analyses are presented below.

3.1. Deterministic Study

Nonlinear time series analysis (NLTSA) serves as a powerful apparatus that can directly probe into the dynamical states of deterministic systems. It also provides a framework for the inverse problem complexity, which in some cases can help regain the equations of motion of the underlying system. In the current work, as a complementary study to Bhatta & Dhillon (2020) and in direct contrast to stochastic modeling of the observed variability study, NLTSA is carried out on the γ-ray light curves of 20 blazars with a deterministic approach such that the light curves are modeled as the output from the low/high order dynamical systems. As in the Fourier transform-based analyses, to characterize the deterministic properties of the astronomical observations could be a challenging task, because the methods, in principle, demand observations for an infinite length of time (see Takens 1981, for mathematical proof).

Moreover, in NLTSA, the appearance of chaotic properties of deterministic systems could be of a diverse nature and depend upon a number of factors, e.g., the number of degrees of freedom of the underlying systems, the measurement error, and the signal-to-noise ratio (see Bradley & Kantz 2015, for an overview on Chaos). Nevertheless, estimates of the relation between these quantities can be found, for example, in the well-known invariant method of estimation of fractal dimension; the relation between the number of observations in the original state space and the correlation dimension can be constrained as $N > 4D_2$ (Smith 1988), where $D_2$ is correlation dimension (an effective algorithm of correlation dimension can be found in Grassberger & Procaccia 1983). In particular, Chaos analysis of scalar time series can be approached following several methods, e.g., invariant methods such as fractal dimensions (box counting, correlation dimension) or numerically calculated Lyapunov exponents (see Kantz & Schreiber 2003). We have demonstrated the applicability of some of these methods while treating relations of the chaotic and regular motion around magnetized black holes (Páns et al. 2019). From the well-established NLTSA methods, we employ the recurrence quantification analysis (Zbilut & Webber 1992) and the practical method for determining the minimum embedding dimension of a scalar time series introduced by Cao (1997). These two approaches are preferable for the work, as the methods are less sensitive to the gaps in the data and work well for a reasonably finite number of observations. The number of observations for the 20 blazars presented in Bhatta & Dhillon (2020) vary around 430 and are reasonably evenly sampled, and therefore well suited for such an analysis.

Furthermore, we also realize the importance of the choice of the parameters in obtaining most reliable results. In the context

---

3 The Fermi/LAT data acquisition and processing are discussed in Bhatta & Dhillon (2020).

4 https://ned.ipac.caltech.edu/
of the RQA method, we adopt an approach where we consider presenting RQA as the function of thresholds instead of an RQA measure for a single value of a threshold. A similar approach was implemented in Suková et al. (2016) for the calculation of the significance of chaotic processes. The underlying assumption of this approach is that the recurrence measures are significant on different scales (thresholds). This aspect of the analysis based on recurrence plots (RP) can be observed in unthresholded recurrence plots (e.g., see RPs in Charles & Cornell IoanaNorbert Marwan 2015). Consequently, an RQA measure over a range of thresholds should be more accurate than just considering one fixed value. It exhibits more rigorous deterministic behavior in a system, along with its properties on different scales.

For the estimation of the embedding in relation to the degrees of freedom of the underlying system, we use the
method developed by L. Cao, which is particularly well known for not being sensitive to the number of observations; this embedding dimension is later used as the input for the RQA analysis of the light curves. There is emphasis given to obtain the most unbiased results, and for this purpose we present Tables 2, A1, and B3 of different configurations of the
algorithms applied on real data, where the emphasis is given to the task of distinguishing between less and more deterministic signals present in the observations. For this purpose, the tables of main results, 2, A1, and B3, are presented in the descending order of the fourth column value, the averaged Determinism measure as described in Section 3.3 by Equation (8). The computation of the relevant quantities, i.e., average mutual information, L. Cao algorithm, and RQA (see Sections 3.1.1, 3.1.2, 3.3, respectively) were performed using the “NonlinearTseries” package in R (Garcia 2019). Optimal parameters for the above functions have been set up by testing the performance on artificial light curves (ALC) produced with the RobPer (Thieler et al. 2016) library (see Section B). The application on real data in three ways has been done according the results of testing presented in Table B1. The ALC with different configurations had especially different noise-to-signal ratios, and the values of input parameters were tuned in consideration of the ordering of the signals according to their deterministic content.

3.1.1. Time Delay and the Average Mutual Information

The roots of nonlinear time series analysis are bounded with the state space reconstruction. One can reconstruct the dynamics of a multidimensional nonlinear system from a single time series using theoretical formulation based on mathematical theorems, for example (Takens 1981). However the term reconstruct is meant in the sense of topological properties, which can be very useful in exploring the behavior of the underlying systems. The standard approach for state space reconstruction is the delay coordinate embedding. The original scalar vector from the time series is simply mapped into a new space, which is defined by the number of delayed dimensions. The \( m \)-dimensional delayed vector \( \mathbf{X}(t) \) constructed from \( m \) samples of the \( y(t) \) with the delay \( \tau \) is defined as:

\[
\mathbf{X}(t) = [y(t), y(t - \tau), y(t - 2\tau), ..., y(t - (m - 1)\tau)].
\]

The embedding theorems require \( \tau \) to be any nonzero, not necessarily a multiple of any orbits period. However, this is true only in case of an infinite amount of noise-free data. When dealing with real observations, one works with finite data added, with noise and measurements errors. In practice, the \( \tau \) is a significant factor when reconstructing the phase space, and if \( \tau \) is too small, the \( m \) coordinates in each of these vectors are highly correlated, and the points from embedded dynamics are close to the main diagonal of the reconstruction space and may not show any interesting structure. If \( \tau \) is too large, the different coordinates may be not correlated and the reconstructed attractor may not be very similar to that of the underlying system.

The time delay vector defined by \( \tau \) has significant impact when comparing the results of chaoticity measures for different set of observations. This has been observed when running many simulations with artificially produced data; therefore, we tested in Section B three configurations of the \( \tau \) choice, namely a) the same value for every set of compared observations chosen as a maximum or mean of the set, or b) its own value for every observation. This \( \tau \) value, calculated by the average mutual information (AMI; see Section 3.1.1), was taken as the maximum for RQA input in main results in Tables 2 and B3 for the compared sets of light curves of not-interpolated observations, and as its own value for every observation in Table A1 of interpolated observations.

Mutual information, a measure of the information shared between two random variables, is also used by stochastic modeling (Jiao et al. 2017). In the framework of NLTSA of a observed time series \( x(t) \), AMI denotes the amount of knowledge mined into the neighborhood of \( x(t + \tau) \). The AMI algorithm as described in Kantz & Schreiber (2003) uses the interval explored by the data where it constructs a histogram of \( \epsilon \) resolution for the probability distribution of the data. If \( p_i \) is probability that the signal has a value in \( i \)-th bin of the histogram and \( p_{ij} \) is probability that \( x(t) \) is in \( i \)-th bin and \( x(t + \tau) \) is in \( j \)-th bin, then AMI for the given \( \tau \) is written as

\[
\text{AMI}_i(\tau) = \sum_{i,j} p_{ij}(\tau) \ln p_{ij}(\tau) - 2 \sum_i p_i \ln p_i. \tag{2}
\]

The \( \tau \) is then selected by first minimum approach, that is, \( \tau \) for which AMI function reaches its first minimum. Another method for calculating appropriate \( \tau \) is the graphical approach and the autocorrelation function (ACF). By graphical approach, one observes the structure of the reconstructed state space. However, the graphical approach is definitely not suitable when dealing with a large amount of data, as it could be computationally expensive. To manually tune \( \tau \) and then plot the reconstructed phase space and decide whether it is appropriate or not is also time consuming. There are arguments against the use of ACF in the context of nonlinear analysis, as the method is based on linear statistics and it could omit the nonlinear dynamical correlations (Kantz & Schreiber 2003). In such context, it is often stated that the product \( m \times \tau \) becomes a more relevant and meaningful measure rather than the exact values of \( m \) and \( \tau \) (Bradley & Kantz 2015).
3.1.2. L. Cao’s Practical Method for Determining the Minimum Embedding Dimension of a Scalar Time Series

This method bears several practical advantages in the estimation of the minimum embedding dimension of time series. One of them is the low number of subjective input parameters, and namely only the time delay parameter $\tau$. This is a big advantage; performing chaos analysis by nature is strongly sensitive not only to the initial conditions but also to the number of input parameters used in modeling the data. Moreover, the implementation of the method consumes less computational time compared with other methods, e.g., invariant methods.

Let the $x_1, x_2, \ldots, x_N$ be the observed data, and let the reconstructed time delay vector have the form $y_i(m) = (x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau})$, where $i = 1, 2, \ldots, N - (m - 1)\tau$, where $m$ denotes the embedding dimension and $\tau$ the time delay, then $y_i(m)$ denotes the $i$th reconstructed vector in state space with embedding dimension $m$. Next the variable $a(i, m)$ is defined similarly to the false nearest neighbors method:

$$a(i, m) = \frac{||y_i(m + 1) - y_{n(i,m)}(m + 1)||}{||y_i(m) - y_{n(i,m)}(m)||}$$

with $i = 1, 2, \ldots, N - m\tau$ (3)

where $n(i, m) \in (1 \leq n(i, m) \leq N - m\tau)$ is an integer for which $y_{n(i,m)}(m)$ is the nearest neighbor of $y_i(m)$ in the reconstructed $m$-dimensional state space; naturally $y_{n(i,m)}(m)$ is nearest to the $y_i(m)$ in some Euclidean norm when $y_{n,i,m}(m) = y_i(m)$; the nearest neighbor is the smallest $n(i, m)$ for which $y_{n(i,m)}(m) \neq y_i(m)$. When two points are close in the $m$-dimensional reconstructed space, and also $(m + 1)$-dimensional reconstructed space, they are called true neighbors; otherwise, they are false neighbors. The feature of true neighbors comes from the embedding theorems, such as Takens (1981), and for a perfect embedding, no false neighbors does exist. In order to omit defining some value for which

Figure 3. The distribution of the averaged DET, L, and ENTR as derived from the recurrence analysis and as presented in Table 2 are shown in the left, middle, and right panels, respectively. The sources FSRQs and BL Lacs are distinguished by the red and blue colors, respectively.

Figure 4. The characteristic dynamical timescales as represented by the average vertical line features in the recurrence plot of the sample sources, along with the legend, are shown here. The RQA parameter $V_{\min}$, denoting the minimum number of points considered as a vertical line, scales by the same criteria as the selection of $l_{\min}$ (see the caption of Table B1). The sources FSRQs and BL Lacs are distinguished by the red and blue colors, respectively.
\[ E(m) = \frac{1}{N - m^2} \sum_{i=1}^{N-m^2} a(i, m). \]  

To determine right \( m \), the \( E1(m) = E(m)/E(m+1) \) is introduced; \( E1(m) \) stops changing when \( m \) reaches some value of \( m_0 \) and then \( m_0 + 1 \) is the estimation of the embedding dimension. For an illustration, the analysis on the light curve of the blazar AO 0235+164 is presented in Figure 2 which shows \( E1 \) (black curve) and \( E2 \) (red curve) as a function of dimension, and the confidence interval is marked by the green dotted lines.

### 3.2. Recurrence Quantification Analysis

The recurrence quantification analysis (RQA) is an apparatus that measures the properties of the recurrence plot, a graphical tool introduced by Eckmann et al. (1987) that is used for investigating the state space trajectories. RQA was introduces in 1992 by Zbilut & Webber (1992) and later improved by Marwan (2008). In RQA, the basis for calculating RP is provided by the matrix defined as

\[ R_{ij} = H(\epsilon - \|x_i - x_j\|) \quad i, j = 1, \ldots, N, \]

where \( N \) is the number of measured points, \( x_i \), and \( \epsilon \) is a threshold distance, a crucial value that has a strong effect on the result. \( H(\cdot) \), the Heaviside function, is defined as

\[ H(\epsilon) = \begin{cases} 0, & \epsilon < 0 \\ 1, & \epsilon \geq 0. \end{cases} \]

It can be seen that Equation (5) gives rise to a symmetrical square matrix that consists of binary values, i.e., zeros and ones. The RP is obtained as a plot of this square matrix. As threshold value parameter largely determines density of the RP plot, there appears to be some ambiguity over a consistent choice of the \( \epsilon \) (Schinkel et al. 2008). Therefore instead of looking for one single set of correct values, a more rigorous result could be obtained by averaging over more thresholds.

In this work, we follow the approach in the implementation of the RQA method by averaging over the span of thresholds. In our case, the thresholds are calculated for a wide range of percentages of points in RP (ones in the binary matrix) and thereby consider the significance of RQA measures on divergent scales. This percentage is actually part of the RQA analysis defined as the recurrence rate (RR)

\[ RR = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{ij}, \]

which provides a measure for the density of the recurrence points in the RP. As an example case, the RP for the blazar CTA 102 are shown in the panels of Figure 1. Dictated by Equation (6), the panels shows that as the RR is increased from 1%-95% with a step of 5%, i.e., \([1, 5, 10, \ldots, 95]\%\), the area in the plot is gradually populated by larger numbers of blue symbols.

Determinism—Determinism is computed considering the RR of the points that align along the diagonal lines of the RP. The quantity tells how deterministic or well behaved a system is. The mean determinism, over the range of the RR\(%\) considered here, of the sample blazar \( \gamma \)-ray light curves, denoted as “mD,” is presented in the fourth column of the

\[ \gamma \] denotes as the recurrence rate of the points that align along the diagonal lines of the RP.

### Tables and A1

Tables 2 and A1, and note that all the other column values in the tables are sorted according to descending order of the mean determinism values.

\[ DET = \frac{\sum_{l=l_{\text{min}}}^{N} P(l)}{\sum_{l=l_{\text{min}}}^{N} l}, \]

where \( P(l) \) denotes the frequency distribution of the lengths \( l \) of the diagonal lines.

\[ L = \frac{\sum_{l=l_{\text{min}}}^{N} l P(l)}{\sum_{l=l_{\text{min}}}^{N} P(l)} \]

The mean line length, over the range of the RR\(%\) considered here, of the sample blazar \( \gamma \)-ray light curves, denoted as “mL,” are presented in the fifth column of Tables 2, A1, and B3.

ENTR—Entropy, computed as the probability distribution of the diagonal line of lengths \( p(l) \) of the RP, provides a measure of the complexity of the data. In other words, the quantity conveys how much information the observation does contain—or richness of the data.

\[ \text{ENTR} = - \sum_{l=l_{\text{min}}}^{N} p(l) \log p(l), \]

where \( p(l) \) is the probability that a diagonal line in the RP is exactly of the length \( l \)—it can be estimated from the frequency distribution \( P(l) \) with \( p(l) = \frac{P(l)}{\sum_{l=l_{\text{min}}}^{N} P(l)} \). The mean entropy, over the range of the RR\(%\) considered here, of the sample blazar \( \gamma \)-ray light curves, denoted as “mENTR,” are presented in the sixth column of the Tables 2, A1 and B3.

In the computational process, for every RR\(%\) \( \in [1, 2, 3] \) in Table 2, RR\(%\) \( \in [1, 2, 3] \) in Table A1, and RR\(%\) \( \in [5, 10, \ldots, 95] \) in Table B3, RQA measures (D, L, and EN) were calculated for the observations and averaged into mL, mD, and mENTR, as well as listed in descending order of D, L, EN, sequentially. Then integers from 1 to 20 (number of analyzed observations) were assigned to them as sD, sl, and sEN, respectively. One can visualize the scoring sD, sl, and sEN measures for some RR and all the observations as a column with values of \((20, 19, \ldots, 1)\) aligned in descending order of D, L, and EN measures for 20, highest value of D, L, and EN, 19 for second highest, etc.).

The additional statistics of averaged scoring measures mL, mD, and mENTR of sD, sl, and sEN provide additional information about times the observation scored at some place for considered percentage of RR. By comparing mL and mD columns, it can be seen that the ordering of averaged scoring can be different from the averaged RQA measures mL, mD, and mENTR. This means that a source that is most deterministic by mL measure might have slightly less deterministic scoring.

For the sample, the averaged RQA quantities DET, L, and ENTR for RR\(%\) \( \in [1, 2, 3] \) are presented in Table 2 (denoted by mL, mD, mENTR), and are also depicted in Figure 3. From the table and the figure, it can be seen that the averaged RQA measure DET is relatively large for some of the sources in the sample, especially for FSRQs. This implies the dominant role
of the deterministic processes that lead to the observed $\gamma$-ray variability in the sources. Moreover, Figure 3 shows an interesting feature: that the RQA measures on average as FSRQs have larger values of $\text{DET}$, $\text{L}$, and $\text{ENTR}$ compared with BL Lacs. This pattern can also be observed in Table B3 with $\text{RR}[] \% \in \{1, 5, 10, \ldots, 95\}$ and in A1 of linearly interpolated observations, where $\text{RR}[] \% \in \{1, 2, 3\}$ with different orders. We also note that the analysis results zero mean entropy for eight of the sources, mostly BL Lacs. This could have been caused by the relative “low information content” in the light curves of these sources.

Because nonlinear phenomena and chaotic systems are sensitive to initial conditions, the nonlinear methods and algorithms are also sensitive to the inputs. Therefore, it is not surprising that the results presented in Tables 2, A1, and B3 differ in the orders and magnitudes of the measures. However, the difference in Tables 2, A1, and B3 is more significant in the magnitude of the RQA measures than in the order of sources according to their deterministic content; therefore, rather than compare the magnitudes of the RQA measures among the tables, comparing the RQA measures of the sources within a single table makes more sense. Figure 3 shows the pattern of FSRQs with higher deterministic content, which is also observable from the tables. However, it is important to note that the position of blazars W Comae and AO 0235+164 can be questionable when taking into account their least number of observations—around half of the mean length (see second column in Table 2), while the length is an important factor when handling nonlinear phenomena from both theoretical and practical points.

In such context, the source blazar 3C 454.3 can be seen as one with the highest averaged deterministic value according to Table 2; this blazar is second in Table A1 behind 4C+21.35, and in Table B3, the FSRQ PKS 1502+106 has the most deterministic content.

In Figure 3 we observe that in the mEN measure, the FSRQs also correlate with the mD measure and show more information than BL Lacs. Consequently, the sources are less predictable according to the mL measure, following from the fact that the complex nonlinear and chaotic systems are likely to be less predictable.

In Table B3, where the RQA measures are averaged to a very high percentage of $\text{RR}$, the distinction between FSRQs and BL Lacs is also observed. In this case, where $\text{RR} \in \{1, 5, 10, \ldots, 95\} $, the order of sources would almost not change if the averaging were to be set just 50%, while the higher the limit of averaging is set up, the bigger the gain of mD and mEN measures. Overall, while the mD and mEN measures in this case gained some values, the mL column is lower in the sample sources using RR[()]. The RPs were constructed for the sample sources using RR[()] from 1 to 3, and the recurring features, e.g., lengths of the vertical lines and time delays between diagonal lines, were computed. However, comparing different configurations of the setup of RQA, the diagonal timescales are too dependent on the setup of RQA parameters, and therefore are not presented in this work. The average of the timescales corresponding to the range of the RR[()] was taken as the dominant timescale in the light curves. The resulting vertical timescales in weeks ($\gamma$-axis in the plot) for the sample sources are shown in Figure 4. It’s seen that trapping timescales, which reflect the stability/instability of a system, are in the order of 5–15 weeks.

4. Discussion

The nonlinear time series analysis performed on the Fermi/LAT light curves of a sample of 20 blazars has revealed interesting results. In this section, we present discussion on the possible interpretation of the results in the context of currently accepted blazar models.

As known, the measurements by sensors do not have smooth time bases. The embedding theorems require evenly spaced observations. This is definitely a less luxurious demand than the infinite amount of observations, and one can achieve this by the well-known technique of interpolation. So to obtain most unbiased results, we apply the nonlinear analysis also on the interpolated data. We compare the results obtained with the “raw” data in Table 2, so as the interpolated ones in Table A1, and try to judge the effect of added interpolated dynamics.

When providing RQA, the choice of the threshold is crucial; the recommendations for its choice have been given by many authors, and most of them are derived from the variance of the observed data (see Schinkel et al. 2008). As mentioned earlier, for a more robust result, instead of one single value of the threshold, the RQA measures averaged (see Section 3.3) over a range of thresholds are preferred. This approach reflects the behavior of the underlying system in diverse scales, providing more objectivity to the results. One of the important parameters of this approach is the interval of percentages one considers for this averaging. In this work, many setups of RQA were massively tested on artificial data in the sense of choice of minimal diagonal length (imin), time delay ($\tau$), embedding dimension ($m$), and the value of recurrence rate (RR), before the analysis with the most suitable parameters was performed on the real observations (see Appendix B).
To present the most unbiased results possible, three different approaches to the RQA are performed on the real data, which were based on the results from the extensive testing processes using similar artificial data as presented in Table B1. For every RR[%] in the first column, 90 different setups of RQA were performed (10 different ways for the lmin, 3 ways for the $\tau$, and 3 ways for $m$). With an aim to configure the most suitable parameters, i.e., lmin, $\tau$, and $m$, for every RR[%], the optimal setup was selected based on the ability to sort the signals in the ALCs according to their deterministic strengths. (For details on the testing process, refer to Appendix B). Every setup was tested on 5 sets of 10 ALC generated using 5 different configurations of the generator: in the first 3 ways, light curves were generated using the same configuration for all 10 different signal-to-noise ratios (S/N), and in the other two ways, all of the 10 light curves with different S/N had randomized parameters. On top of these, the ALCs were provided with high red noise content to account for the observed power-law shape of the power spectral density of the $\gamma$-ray observations. Next, to mimic our condition the most, gaps in the data were made. From a fixed length of the generated data set up to 513, the random amount of data up to 50% was replaced by Not a Number (NaN) values. The artificial data were then also linearly interpolated, making in common two huge sets of artificial data.

From Table B1, two approaches are chosen and presented on “raw” data: a) the approach where $\tau = 9$ and $m = 11$ are set up as maximum values from the set of 20 blazars, while it is averaged up to $RR = 2\%$ only by the step of 1% (see Table A1); and b) the setup of maximal $m$ and the corresponding $\tau$ for every source, where in order to include the effects of a wide range of threshold, large range of $\epsilon$ is included in terms of RR between 1%–95% with the step of 5% (see Table B3), and then the measures are averaged over the values of $\epsilon$ to obtain the final results. It is noted that this setup was found to produce quite stable results when considering different inputs of $\tau$ and $m$. Although the weekly binned decade-long $\gamma$-ray observations are fairly well sampled in terms of number of observations, they are not strictly evenly sampled. Therefore, to assess the possible effects of uneven sampling of the light curves on the analysis, the light curves were made evenly spaced by linear interpolation, and the analysis is performed in a similar way. To illustrate the data interpolation, the real and interpolated observations for the two blazars, namely 1ES 1959+65 and TON 0599, are shown Figure B2. The corresponding quantities resulting from the analysis are presented in Table A1 in the Appendix. We note that there are no significant changes in terms of the measure RQA quantities in comparison with Table 2, but now as more points due to interpolation are introduced, as the RR is averaged slightly higher by 1%, the mEN gained some values. The setup chosen for interpolated data has also a maximal value of $m$ and an own value of $\tau$ for RQA input, and it was averaged until 3% by the step of 1%. The precision was set up to one-hundredth for $RR \leq 5\%$ and one-tenth for $RR > 5\%$.

The processes in an AGN system could consist of both deterministic and stochastic natures. Nevertheless, both log-normal flux distribution and logarithmically increasing variability as reported in Bhatta & Dhital (2020) present evidence that the AD-related modulations still dominate over the large spatial and temporal extension, providing an overall nature of the variability as deterministic. Kiehlmann et al. (2016), in their study of electric vector position angle (EVPA) rotations in the blazar 3C 279, came to the conclusion that while the low-brightness states in the blazar could be a result of a stochastic process, most of the high-amplitude variability should be the result of the underlying deterministic processes. The results point to the strong coupling between the accretion and the jet in the sense that the disk and the jet interact in a coordinated manner, such that information about the disk process remains intact while they propagate into the jets. In blazar systems, this also could have an implication that although the observed flux is largely dominated by the non-thermal emission from the jets, variability probes employing suitable time series analysis could still reveal the origin of variability phenomenon to the AD.

The result is that the dominant physical processes in FSRQs of a more deterministic nature can be interpreted in the widely accepted scenario that jets are powered through the extraction of the rotational energy of a supermassive Kerr black hole surrounded by magnetically arrested AD. It is also possible that FSRQs are disk dominated, showing features of more powerful AD. Their jets possibly are less magnetized and consequently provide less favorable conditions to stochastic processes, e.g., rampant shock and/or magnetic reconnection events. In contrast, BL Lacs jet have been found to be abundant with streaming particles that can contribute to the enhanced stochasticity (Zhang et al. 2014).

Alternatively, the deterministic features of high energy emission can be linked to the shock compression events in ordered magnetic fields in axisymmetric linear jets such that the turbulent activities are less prevalent (e.g., see Zhang et al. 2015; Aller et al. 2020). The geometry of the magnetic field of the blazar jets has been routinely explored using multi-frequency polarimeters (e.g., optical band, Blinov et al. 2016; and radio band, Anderson et al. 2019). In particular, highly polarized jets are indicators of large ambient jet magnetic fields. In addition, the sudden EVPA rotations (e.g., Marscher et al. 2008) have been routinely observed. Such EVPA rotations could be indicative of the deterministic process (see the discussion in Kiehlmann et al. 2016), which can be related to the strongest $\gamma$-ray flares frequently observed in blazars (e.g., Abdo et al. 2010; Blinov et al. 2015; also see the $\gamma$-ray light curves of the blazars presented in Bhatta & Dhital 2020). However, the role of stochastic processes (e.g., Jones et al. 1985; Lehto 1989; Bhatta et al. 2013; Marscher 2014) cannot be completely ruled out.

Similarly, the presence of circum-nuclear material, as suggested by relatively stronger emission lines in FSRQs, that is believed to provide the low-energy photons for the inverse-Compton process (External Compton model), giving rise to dominant $\gamma$-ray emission, make such systems more complex, whereas the self-Compton origin of the high energy emission (Synchrotron Self-Compton model), involving a lower number of interacting components, e.g., electron density distribution, magnetic field, and Doppler factor, could imply a relatively simpler scenario.

As seen in Figure 4, the timescales derived from the vertical distribution of the points in the RP are in the range of ~5–15 weeks. The timescales in the former case are an indication of the average recurrent timescale of the dynamical processes signifying how frequently the system revisits a particular state. Moreover, it is interesting to note that the average predictability of timescales is comparable to the recurrent timescale. It should be noted that the flux distribution being log-normal, the light
curves are dominated by lower fluxes, and therefore it is natural to expect the trapping timescales to be in the order of a few weeks. This means the resulted average dynamical timescales represent low flux level fast variability, giving lower weight to large flares lasting several months. Nevertheless the resulting timescales could be relativistically diluted through the relation $R \gtrsim \beta t/(1 + z)$, reflecting the cosmic expansion. For an average redshift of $z = 0.6$, and for a moderate value of Doppler factor $\delta = 10$, the timescales can be translated into the size of the regions following a causality argument. For a typical black hole mass of $10^9 M_\odot$ with gravitational radius $R_g = GM/c^2$, the size corresponding to 15 weeks corresponds to 0.5 pc, which is comparable to the size of the inner AD; and 10 weeks represents a few thousands of gravitational radii, within which most of the gravitational potential energy is converted into the radiation energy. In such interpretation, the inner AD might be treated as the main component of a dynamical state of an AGN, and the modulations driven by various instabilities, e.g., radiation pressure, viscous instabilities (Janiuk et al. 2002; Janiuk & Czerny 2011), occurring within this region leading to the change in the states.

5. Conclusions

We probed a sample of 20 blazars by performing nonlinear time series analysis of their decade-long $\gamma$-ray light curves from the Fermi/LAT telescope. The results of the analysis suggest that the dynamical processes responsible for the $\gamma$-ray variability of the blazars are mostly a mixture of deterministic and stochastic in nature, although in some of the sources, e.g., blazar 3C 454.3, 4C+21.35, and CTA 102, displayed high deterministic content. The result could be significantly useful in formulating the model that explains the interplay between the disk and jet processes ubiquitous in black hole systems. In addition, the analysis reveals characteristic timescales of several weeks, which could be interpreted as so-called trapping timescales ($\sim 5$–15 weeks). The timescales in combination with the results from the multifrequency studies could provide further insights about the nature of origin of the $\gamma$-ray in radio-loud jets.

G.B. acknowledges the financial support by the Narodowe Centrum Nauki (NCN) grant UMO-2017/26/D/ST9/01178. R.P. thanks to the SGF program CZ.02.2.69/0.0/0.0/19_073/0016951 of the Silesian University in Opava. Z.S. thanks the support of the Institute of Physics and the Research Centre of Theoretical Physics and Astrophysics, at the Silesian University in Opava. We are very thankful to the anonymous referee for a thorough and careful reading of the paper and very useful comments and suggestions, which helped improve the presentation of the paper significantly. This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration.

### Appendix A

#### RQA of the Interpolated Blazar Light Curves

To assess the effect of the uneven sampling on the resulted RQA measures, the analysis was also performed on the source light curves that were made evenly spaced by the linear interpolation. The resulting measures, similar to Table 2, are tabulated in Table A1.

#### Table A1

The Nonlinear Time Series Analysis Applied on 20 Blazars with Linearly Interpolated Values, where the Columns Have the Same Meaning as in Table 2

| Source   | Len | $\tau$ | m | mD     | mL   | mEN   | msD   | msL   | msEN  |
|----------|-----|--------|---|--------|------|-------|-------|-------|-------|
| 4C+21.35 | 512 | 4      | 8 | 0.5993 | 15.6162 | 2.5014 | 19.00 | 8.67  | 17.33 |
| 3C 454.3 | 513 | 9      | 9 | 0.5619 | 7.6438  | 1.4513 | 18.00 | 1.00  | 11.00 |
| CTA 102  | 513 | 10     | 9 | 0.5109 | 10.9990 | 1.7812 | 18.67 | 3.67  | 13.67 |
| 3C 279   | 514 | 8      | 8 | 0.4778 | 10.9323 | 2.0061 | 17.67 | 4.67  | 16.00 |
| PKS 0454−234 | 515 | 3     | 9 | 0.3730 | 21.7236 | 2.5954 | 14.67 | 11.00 | 19.33 |
| AO 0235+164 | 511 | 17   | 6 | 0.3528 | 9.0909  | 1.0025 | 14.00 | 3.33  | 8.00  |
| PKS 1502+106 | 515 | 8     | 9 | 0.3413 | 10.9701 | 1.6412 | 14.33 | 5.67  | 13.00 |
| W Comae  | 509 | 4      | 10| 0.3254 | 15.3842 | 2.1845 | 13.33 | 9.67  | 17.00 |
| S 0716+714 | 515 | 4     | 8 | 0.3217 | 18.0873 | 2.0249 | 12.67 | 10.33 | 14.33 |
| PKS 1424−418 | 506 | 13    | 7 | 0.3037 | 12.5349 | 1.5297 | 12.33 | 6.67  | 10.33 |
| 3C 273   | 513 | 5      | 10| 0.2075 | 17.4184 | 1.5992 | 9.33  | 9.67  | 12.00 |
| 4C+38.41 | 513 | 7      | 10| 0.2044 | 13.8741 | 0.7189 | 9.33  | 8.00  | 4.33  |
| TON 0599 | 513 | 9      | 8 | 0.1729 | 32.5414 | 1.0351 | 6.00  | 14.67 | 6.67  |
| PKS 2155−304 | 515 | 6     | 10| 0.1699 | 18.1636 | 0.5620 | 6.33  | 9.67  | 5.00  |
| 3C 66A   | 515 | 3      | 8 | 0.1698 | 47.9726 | 1.8597 | 7.00  | 16.00 | 15.00 |
| Mk 501   | 513 | 3      | 9 | 0.1578 | 77.5414 | 1.6060 | 4.67  | 17.00 | 11.67 |
| Mk 421   | 515 | 6      | 8 | 0.1443 | 59.4289 | 0.7013 | 4.00  | 16.67 | 5.00  |
| BL Lac   | 514 | 5      | 10| 0.1440 | 184.2397| 0.1867 | 4.33  | 16.00 | 1.67  |
| 1ES 1959+65 | 514 | 4    | 11| 0.1361 | 102.9524| 0.9813 | 2.67  | 18.33 | 6.33  |
| ON+325   | 513 | 4      | 9 | 0.1310 | 368.5556| 0.2122 | 1.67  | 19.33 | 2.33  |

Note. It can be observed that in comparison with Table 2, the lengths in the second named column are increased, as the missing values (or “NaN”) were replaced by interpolation. In this analysis, every source has its own $\tau$ for the input of RQA, and $m$ is taken again as the maximum of all the $ms$ calculated, while this choice is made according to the results in Table B1. For this setup, the distinction between FSRQs and BL Lacs, in terms of deterministic content as represented by mD, is also significant.

Bhatta, Pánis, & Stuchlík
Table B1

| RR          | Part A | Part B |
|-------------|--------|--------|
| %           | lmin#  | emb#   | tau#  | measure1 | measure2 | lmin#  | emb#   | tau#  | measure1 | measure2 |
| 1           | 6      | 2      | 3     | 12.43    | 26.67    | 9      | 3      | 1     | 13.57    | 28.00    |
| 2           | 7      | 3      | 3     | 10.35    | 25.00    | 1      | 3      | 1     | 11.35    | 27.23    |
| 3           | 7      | 3      | 3     | 11.72    | 26.33    | 7      | 3      | 1     | 11.12    | 26.12    |
| 4           | 7      | 3      | 3     | 11.72    | 27.00    | 7      | 3      | 1     | 12.78    | 26.33    |
| 5           | 6      | 2      | 2     | 13.99    | 29.33    | 10     | 3      | 1     | 14.12    | 29.33    |
| 10          | 2      | 3      | 3     | 13.12    | 31.33    | 4      | 3      | 2     | 14.93    | 34.67    |
| 15          | 1      | 2      | 3     | 12.48    | 28.00    | 4      | 3      | 2     | 14.33    | 33.33    |
| 20          | 1      | 3      | 3     | 12.43    | 30.67    | 3      | 3      | 1     | 14.77    | 32.67    |
| 25          | 1      | 1      | 3     | 11.98    | 28.00    | 5      | 2      | 1     | 14.93    | 33.33    |
| 30          | 1      | 1      | 3     | 11.97    | 27.33    | 5      | 2      | 1     | 14.93    | 33.33    |
| 35          | 2      | 3      | 3     | 12.27    | 30.67    | 9      | 3      | 1     | 14.77    | 33.33    |
| 40          | 2      | 2      | 3     | 12.27    | 28.00    | 6      | 3      | 1     | 14.61    | 33.33    |
| 45          | 2      | 2      | 1     | 12.27    | 28.00    | 9      | 1      | 1     | 14.76    | 32.67    |
| 50          | 2      | 1      | 1     | 12.27    | 28.00    | 9      | 2      | 3     | 14.76    | 32.67    |
| 55          | 2      | 1      | 1     | 12.27    | 28.00    | 3      | 3      | 3     | 14.11    | 33.33    |
| 60          | 2      | 1      | 1     | 12.27    | 28.00    | 10     | 3      | 3     | 14.11    | 34.00    |
| 65          | 2      | 3      | 1     | 12.61    | 29.33    | 10     | 3      | 3     | 14.11    | 34.00    |
| 70          | 2      | 3      | 1     | 12.27    | 28.67    | 10     | 3      | 1     | 14.11    | 34.00    |
| 75          | 2      | 3      | 1     | 12.27    | 28.67    | 10     | 3      | 1     | 14.11    | 34.00    |
| 80          | 2      | 3      | 1     | 12.27    | 28.67    | 9      | 3      | 1     | 14.35    | 34.67    |
| 85          | 2      | 3      | 1     | 12.27    | 28.67    | 7      | 3      | 1     | 14.35    | 35.33    |
| 90          | 2      | 3      | 1     | 12.61    | 29.33    | 3      | 3      | 1     | 14.94    | 34.67    |
| 95          | 2      | 3      | 1     | 12.27    | 28.67    | 3      | 3      | 1     | 14.44    | 36.00    |

Note. Part A of the table corresponds to the artificial light curves with introduced gaps, and Part B shows the results of the data with linearly interpolated gaps. The lmin# column takes integer numbers from 1 to 10 and denotes setup of minimal line length, where lmin# 1–4 belongs to fixed lmin of values equal to 2, 3, 4, and 5, and lmin# 5–10 belongs to lmin adjusted with RR. Lmin6 is scaled according to the RR by the rule lmin5 = RR[%, lmin6 = RR[%, lmin7 = RR[%] + 1, and similarly up to lmin10. The columns emb# and tau# have been considered for testing of the set of ALCs as 1) individual values corresponding to each ALC; 2) the mean of the whole considered set; and 3) the maximal value from the considered set of 10 ALCs with S/N of values from [0.005, 0.01, 0.025, 0.5, 0.75, 1, 1.5, 2, 3, 5]. The column measure1 and measure2 correspond to the summed absolute value of difference between the vector of method ordered set of 10 ALCs and the real defined ascending order (the measures are described in Section B). The values in bold correspond to best—lowest scoring configurations and its measures, while measure1 ∈[0, 22.42] and measure2 ∈[0, 50].

Appendix B

Testing Process

To configure the optimal setup of the parameters to be employed in the analysis of the real observations, several tests on artificial data were performed prior to the application of averaged RQA analysis on the data set of decade-long γ-ray light curves of 20 blazars. The tests made use of several time series algorithms that deal with nonlinear phenomena. The choice of the methods used in the analysis are based on the two main criteria: First, the methods should involve a minimal amount of inputs; this is crucial because when dealing with nonlinear phenomena such as chaos, the results are very sensitive to the inputs of the algorithms. In addition, the usage of a low number of inputs has several benefits, e.g., the analysis can be performed for a number of combinations of the inputs within computational resources, and the interpretation of the results is easier in terms of physical theories—in contrast to the machine learning algorithms that can provide better fits, however, the interpretability, in some sense, is often in terms of a black box. These tests are carried out in order to select suitable parameters from bounded parameter space for estimating embedding dimension, time delay, and optimal recurrent rate, which subsequently can be fed to AMI (Section 3.1.1), the L. Cao algorithm, (Section 3.1.2), and RQA (Section 3.3) analysis. The RQA measure can mathematically be described in terms of a compound function such as:

\[
RQA(c) = RQA(Imin(RR(c))),
\]

where Imin is the parameter for calculation; the diagonal features are RP, which defines how minimally many diagonally connected points are considered as a line; RR is the recurrence rate defined by Equation (7); and \(c\) is the threshold value see Equation (5), which simply says what is the distance between points in order to denote it by 1 by Heaviside function and later denote it by color (not white in RP).

The approach of the testing, i.e., searching for the optimal parameter setup, is motivated to distinguish deterministic signals from stochastic noise. In the context of RQA, the strength of the deterministic part of the signal is represented by the DET measure (Section 8). The R package RobPer allows for generating ALC with different configurations, while one of the parameters is the strength of the signal, where the power-law noise is generated according to the prescription discussed in Timmer & Koenig (1995).

For training purposes, the parameters for 10 ALC were configured for several values of S/N (see S/N from tsgen function in Thieler et al. 2016), taking values from the vector \([0.005, 0.01, 0.025, 0.5, 0.75, 1, 1.5, 2, 3, 5]\). To produce most unbiased result on the data of 20 blazars, the setup is applied to both interpolated and not-interpolated data. The analyses were performed on artificial data, which mimicked the sampling of
the real observations (see Figure B1). From the generated ALCs with the total length of 513 weeks (~10 years), observations up to 50% were randomly deleted to include the effects of the shortest data length of the source W Comae. The real and interpolated observations for the two blazars, namely 1ES 1959+65 and TON 0599, are shown Figure B2.

To encapsulate the RQA behavior of the varied nature of the light curves, for each of 5 different configurations, 10 ALCs with different S/N and noise content were generated (see Section 4). In addition, the ALCs were provided 10 additional different parameters, including length of the observations (see tsen Thieler et al. 2016 for details). During the process, for 3 out of 5 configurations, these parameters were fixed for every S/N, whereas for the other 2 configurations, most of the parameters (for given boundaries) were randomized. In addition, the ALCs had high red noise ratio, up to 95%, in the white noise/red noise mixture. In the case of the first 3 configurations for generating ALCs, as for the selection of the main inputs to the RQA, namely time delay (τ), embedding dimension (m), and the choice of the minimum line length (lmin), approaches based on multiple criteria were adopted. The values of τ and m have been considered for testing for the set of ALCs as a) specific to each ALC, b) the mean of the whole considered set, and c) the maximal value from the considered set. So there are 9 ways to configure the τ/m setting for the RQA. Lmin plays one of the key roles in RQA, also in the sense of the magnitude of the RQA measures. In many computational libraries, e.g., “NonlinearTseries,” “RHRV,” and “crqa,” this value is predefined as 2 or 3. In our parameter space, lmin is configured in 10 ways; 4 of them are of the fixed value of 2, 3, 4, and 5, and the rest have a changing value of lmin according to the RR value that is used. Changing lmin value with higher RR seems like a natural approach that can avoid “tangential motion,” which appears when the RR-e value is higher (Marwan et al. 2007). In Theiler (1986), the author recommends that for calculation of correlation integral, the choice of lmin should be similar to the choice of the Theiler window. However, the configuration of lmin according to RR has not been explored much and could be an active field of enormous possibilities. We considered 6 approaches for choosing lmin according to RR in Table B1.

All the computation has been performed on Intel TM core i7 processor of 7th generation. The computation is mostly demanding in the sense of having to repeat (loop) the computation of RQA many times to find the desired threshold (ε) for given RR. The computation of RQA could not have been improved by using available R libraries, which use a graphics processing unit (GPU), while the efficiency of GPU comes with long data sets. In frames of the computational resources, the precision of finding threshold for RR has been within tolerance of 5 hundredths for RR lower than 5% and 5 tenths is equal or greater 5%.

To summarize, in order to set up and search the optimal parameters, an extensive test was performed using a large number of ALCs, i.e., on each group of 10 ALCs with 5 different settings (divided into 2 groups—with gaps and

Figure B1. An example of artificial light curve generated using the tsen function from the R package RobPer (Thieler et al. 2016), which was used to configure the most suitable RQA parameters and later applied to real data. The red points denote the linearly interpolated values, which were deleted in order to mimic the sampling of the blazar γ-ray light curves.

Figure B2. The weekly binned Fermi/LAT observations (black symbols) were made evenly spaced by the linear interpolation (red symbols). The upper and lower panels show the light curve of the blazars 1ES 1959+65 and TON 059, respectively. The uncertainty in the flux is not shown here for clarity.
interpolated ones). RQA algorithm ran on loop while desired 
thresholds for given RR with defined precision were found; this 
process was repeated 9 times for a combination of 3 \( \tau \)s and 
3 ms. After the desired threshold for RR was found, 10 different 
approaches of choosing \( l_{\text{min}} \) were computed. Eventually, on each of 100 ALCs, 90 ways of setting the 
RQA were tested.

The most time demanding operation in this approach was the 
search of thresholds belonging to considered RR, while the 10 
approaches of choosing \( l_{\text{min}} \) were computed very fast after-
wards. The search of thresholds belonging to considered RR, while the 10 
values of \( S \) were averaged to high values of RR.

when reaching the region close to desired RR, the step was 
divided in loop by 10 to not jump over the desired precision.

The total time taken for the entire computation/calculations was 
\( \sim 100 \) hours, including the process of saving the 
intermediate results, using the available computation facilities.

The measure of the ability to sort the ALCs according to the 
deterministic content (estimated by \( m_{\text{DET}} \) measure) has been 
defined in two ways: a) by putting more significance to the 
ordering of stronger signals—when the absolute value of the 
difference of computed order by an RQA setup and the order of 
intermediate results, using the available computation facilities.

The measure of the ability to sort the ALCs according to the 
deterministic content (estimated by \( m_{\text{DET}} \) measure) has been 
defined in two ways: a) by putting more significance to the 
ordering of stronger signals—when the absolute value of the 
difference of computed order by an RQA setup and the order of 
intermediate results, using the available computation facilities.

The measure of the ability to sort the ALCs according to the 
deterministic content (estimated by \( m_{\text{DET}} \) measure) has been 
defined in two ways: a) by putting more significance to the 
ordering of stronger signals—when the absolute value of the 
difference of computed order by an RQA setup and the order of 
intermediate results, using the available computation facilities.

The measure of the ability to sort the ALCs according to the 
deterministic content (estimated by \( m_{\text{DET}} \) measure) has been 
defined in two ways: a) by putting more significance to the 
ordering of stronger signals—when the absolute value of the 
difference of computed order by an RQA setup and the order of 
intermediate results, using the available computation facilities.

The measure of the ability to sort the ALCs according to the 
deterministic content (estimated by \( m_{\text{DET}} \) measure) has been 
defined in two ways: a) by putting more significance to the 
ordering of stronger signals—when the absolute value of the 
difference of computed order by an RQA setup and the order of 
intermediate results, using the available computation facilities.

The measure of the ability to sort the ALCs according to the 
deterministic content (estimated by \( m_{\text{DET}} \) measure) has been 
defined in two ways: a) by putting more significance to the 
ordering of stronger signals—when the absolute value of the 
difference of computed order by an RQA setup and the order of 
intermediate results, using the available computation facilities.

The measure of the ability to sort the ALCs according to the 
deterministic content (estimated by \( m_{\text{DET}} \) measure) has been 
defined in two ways: a) by putting more significance to the 
ordering of stronger signals—when the absolute value of the 
difference of computed order by an RQA setup and the order of 
intermediate results, using the available computation facilities.

The measure of the ability to sort the ALCs according to the 
deterministic content (estimated by \( m_{\text{DET}} \) measure) has been 
defined in two ways: a) by putting more significance to the 
ordering of stronger signals—when the absolute value of the 
difference of computed order by an RQA setup and the order of 
intermediate results, using the available computation facilities.
between the signals strengths, as in this case the stronger signals would not contribute more to the measure. The two measures, when taking into account 10 ALCs with different S/N, can mathematically be expressed as:

$$\text{measure} = \sum_{i=1}^{10} |x_i - y_i|, \quad (B2)$$

where $x_i$ is the $i$th element of the vector $x$ denoting the S/N/position of ALC according to its mDET measure computed by some setup of RQA, and $y_i$ is the element of the vector of $y$ of the S/N/positions as defined. Most naturally, the $x_i - y_i$ are subtracted when vector $y$ is ordered in either descending or ascending fashion.

One can observe how this measures works in Table B2, which shows that in this particular case, the measure1 = 8.51 and measure2 = 16. Consequently, we consider the 10 particular S/N, the measures for which are bounded in the intervals, where measure1 $\in [0, 22.42]$ and measure2 $\in [0, 50]$. The important results from above described testing are presented in Table B1, where columns of measure2 and measure1 are averaged for 5 different generators of ALCs. The best performances for considered RRs are provided in the table; it can be observed that the second table on the right has measures1 and 2 higher, which suggests that the ability to recognize the different S/N slightly worsens in the case of interpolated data.

The lmin adjusted to RR performs better than fixed ones, especially when averaged to lower RRs $\leq 5\%$ for both data sets. Naturally in this case the lmin adjusts itself with only a few RRs. For interpolated values, the adjusted lmin works better for most of the RRs. For data with gaps the fixed lmin shows better performance above RR $= 5\%$, and for both data sets, better performance when averaged to the highest values of RR $= 95\%$.

When differing between the performance of the ALCs generated by the same setup for 10 different S/N and the ALCs, where every one of the 10 ALCs has a different randomized generator, the ability to sort by S/N is naturally worse for the randomized ones, and interestingly, the methods with adjusted lmin perform better than the fixed ones. The randomized generators of ALCs are in the training sets represented by 2/5, with the assumption that the blazar variability might be governed by similar underlying processes. Table B1 shows the results applied on all 5 ALCs generators.

The configuration with lowest measure1 and 2 for data with gaps appears for RR $= 2\%$, where lmin=# = 7, emb#= 3, and tau# = 1, denoting the adjusted lmin scales with RR by the rule lmin = RR + 2, with $\tau$ and $m$ taken as the maximum from the calculated time lags and embeddings. This configuration applied to real data is presented in Table 2.

The configuration with lowest measure1 and 2 for data with interpolated values appears for RR $= 3\%$, where lmin=# = 7, emb#= 3, and tau# = 1, so the lmin scales the same as the data with gaps, with $\tau$ value as its own for every light curve and $m$ taken again as the maximum from the calculated embeddings. This configuration applied to real data is presented in Table A1.

The best configuration with fixed lmin averaged until RR $= 95\%$ for data with gaps appears for lmin=# = 2, emb#= 3, and tau# = 1, where the lmin = 3, with $\tau$ and $m$ chosen as in previous cases. This configuration applied to real data is presented in Table B3.
Suková, P., & Janiuk, A. 2016, A&A, 591, A77
Takens, F. 1981, Lecture Notes in Mathematics (Berlin: Springer), 366
Theiler, J. 1986, PhRvA, 34, 2427
Thieler, A., Fried, R., & Rathjens, J. 2016, Journal of Statistical Software, 69, 69
Timmer, J., & Koenig, M. 1995, A&A, 300, 707
Uttley, P., McHardy, I. M., & Vaughan, S. 2005, MNRAS, 359, 345
Werner, G. R., Uzdensky, D. A., Cerutti, B., et al. 2016, ApJL, 816, L8
Zbilut, J. P., & Marwan, N. 2008, PhLA, 372, 6622
Zbilut, J. P., & Webber, C. L. 1992, PhLA, 171, 199
Zhang, H., Chen, X., Böttcher, M., et al. 2015, ApJ, 804, 58
Zhang, J., Sun, X.-N., Liang, E.-W., et al. 2014, ApJ, 788, 104