Article

Spike Spectra for Recurrences

K. Hauke Kraemer [1,*], Frank Hellmann [1], Mehrnaz Anvari [1], Jürgen Kurths [1,2,4] and Norbert Marwan [1,2,3,*]

[1] Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, 14473 Potsdam, Germany
[2] Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany
[3] Institute of Geosciences, University of Potsdam, 14476 Potsdam, Germany
[4] Institute of Physics, Humboldt Universität zu Berlin, 12489 Berlin, Germany
* Correspondence: hauke_kraemer@hotmail.com; marwan@pik-potsdam.de

Abstract: A novel kind of power spectrum is constructed, the *inter-spike spectrum*, which transforms any signal into its spike-frequency domain. This method clearly shows the apparent cycles in the data and overcomes the problems for spike-train-like signals when using the obvious idea of Fourier-transforming it. We invent this instructive approach with the idea of transforming the τ-recurrence rate of a recurrence plot (RP), which often has a spiky appearance. The τ-recurrence rate is the density of recurrence points along diagonals of the RP, which are parallel to the main diagonal with a distance of τ. In this context the inter-spike spectrum can be interpreted as a nonlinear power spectrum of a potentially high dimensional system which constitutes the RP. The proposed measure is robust to noise and is able to detect and analyze bifurcations.

Keywords: Decomposition; Frequency Analysis; Recurrence Analysis; Bifurcations

PACS: 05.45.Tp, 05.90.+m, 89.90.+n, 02.70.Uu, 05.10.Ln, 05.45.-a, 05.45.Ac

1. Introduction

Recurrence Plots (RPs) provide a vivid representation of complex dynamics stemming from potentially high dimensional systems [1]

$$R_{i,j}(\varepsilon) = \Theta(\varepsilon - D_{i,j}) = \Theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|), \qquad \vec{x} \in \mathbb{R}^d, \quad i,j \in [1,\ldots,N]. \tag{1}$$

The crucial free parameter ε (recurrence threshold) determines the visible structures in the RP. It can be chosen such that the recurrence rate $RR(\varepsilon) = N^{-2}\sum_{i,j}^{N} R_{i,j}(\varepsilon)$ exhibits a certain value [2]. The simple idea to track recurring states of the d-dimensional trajectory $\vec{x}_i$ of the system under study not only allows for a beneficial visualization of the dynamics, but also for its quantification, using certain structures in the RP, such as diagonal or vertical lines [1].

Some of these recurrence quantification measures, the entropy of diagonal lines and the entropy of recurrence times, can be related to basic characteristics of complex systems, such as the Kolmogorov-Sinai entropy [3,4]. However, these quantifiers have a free parameter, the minimal considered line length, and are usually biased, due to the finite size of the RP and thickened diagonal lines, which needs to be corrected [5]. Moreover, the mentioned statistics cannot account for changing regular (non-chaotic) dynamics, such as period-doubling bifurcations. A rather simple idea is to look at the τ-recurrence rate of the RP (τ-RR, Eq. 2) [6,7]. This is the density of recurrence points along the diagonals of the recurrence matrix, as a function of the distance τ (sampling units) to the main diagonal:

$$\tau\text{-RR}(\varepsilon) = RR(\tau,\varepsilon) = \frac{1}{N-\tau}\sum_{i=1}^{N-\tau} R_{i,i+\tau}. \tag{2}$$

τ-RR serves as an estimator for the probability that the system recurs after time $\tau\Delta t$, with Δt being the sampling time of the trajectory $\vec{x}_i = \vec{x}(\Delta t \cdot i)$, $i = 1,\ldots,N$. Zbilut and Marwan [7] pointed out that τ-RR could be used as a plugin value for the auto-correlation function $C(\tau)$ and, hence, via

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

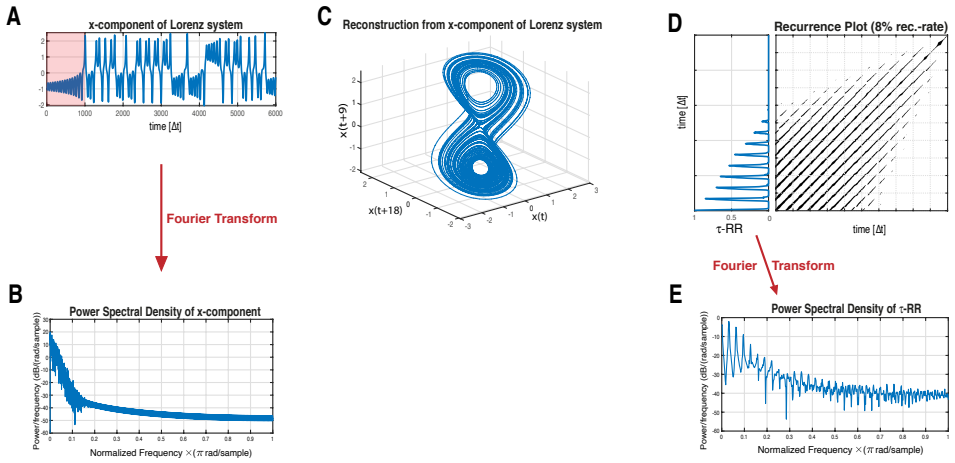

Figure 1. Schematic illustration of a τ-recurrence rate based spectrum. **A** x-component time series of the Lorenz63-System (Eq. (A1)) and **B** its corresponding Fourier power spectrum. **C** Reconstructed state space portrait from the time series shown in **A** using PECUZAL time-delay embedding [8]. **D** Subset of the recurrence plot and the corresponding τ-recurrence rate obtained from the state space trajectory in **C**. **E** Fourier Power spectrum obtained from the τ-recurrence rate (subset shown in panel **D**) [7]. **D** and **E** show the results of a part of the time series, which is highlighted in pink in **A**.

the Wiener-Khinchim theorem a "generalized" powerspectrum can be obtained. This is reasonable, since the average distances for a given lag τ

$$\overline{D}(\tau) = \frac{1}{N-\tau} \sum_{i=1}^{N-\tau} D_{i,i+\tau} \tag{3}$$

can be directly read from the distance matrix **D** and is also preserved in its thresholded version τ-RR. There are clear advantages for a recurrence-derived powerspectrum, i.e., Fourier transforming (FT) τ-RR (Fig. 1D,E), instead of $C(\tau)$: There are no assumptions for stationarity or sampling, when constructing a RP. Furthermore, the correlation structures of higher dimensional spaces can be resolved in the recurrence-derived Fourier-spectrum.

However, there is a drawback to this approach. Whenever τ-RR is a spike-train-like signal, which it is in most cases (see Fig. 1) especially for map-data (low-resolution data), a FT of such a signal leads to a spike-train-like image in the frequency domain (e.g., [9,10], see Fig. 1E). Thus, it is not intuitive how to extract meaningful information about dominant frequencies of the systems' state space trajectory.

For clarification, consider the signal we would like to analyze (e.g., the τ-RR of a system) to be a Dirac comb (DC) with inter-spike period T_{is}:

$$\text{DC}_{\text{is}}(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT_{\text{is}}), \tag{4}$$

i.e., a series of Dirac delta functions for a period T_{is}. There is only one single period, T_{is}, in this signal (Fig. 2A, D), so in principle we would strive for a single peak in the frequency domain of this signal at a frequency $f = 1/T_{\text{is}}$. Surprisingly the Fourier spectrum does not meet this expectation and instead of a single frequency, there are exceptionally many frequencies excited (Fig. 2B, E). This is because the Fourier components add constructively for every frequency $1/T_{\text{is}}$ and therefore $\text{DC}_{\text{is}}(t)$ coincides with its own Fourier transform up to a factor $1/T_{\text{is}}$.

In this article we propose a new way of transforming a spike-train-like signal into its frequency domain. This novel *inter-spike spectrum* does not show resonance behavior of the signal's inherent inter-spike frequencies (Fig. 2C, F). Section 2 explains the idea, which can be used to decompose any arbitrary signal and is not restricted to the τ-RR, which we exemplify here. However, the more spiky the signal is, the more outstanding our new approach is compared to a FT. In Section 3 we

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

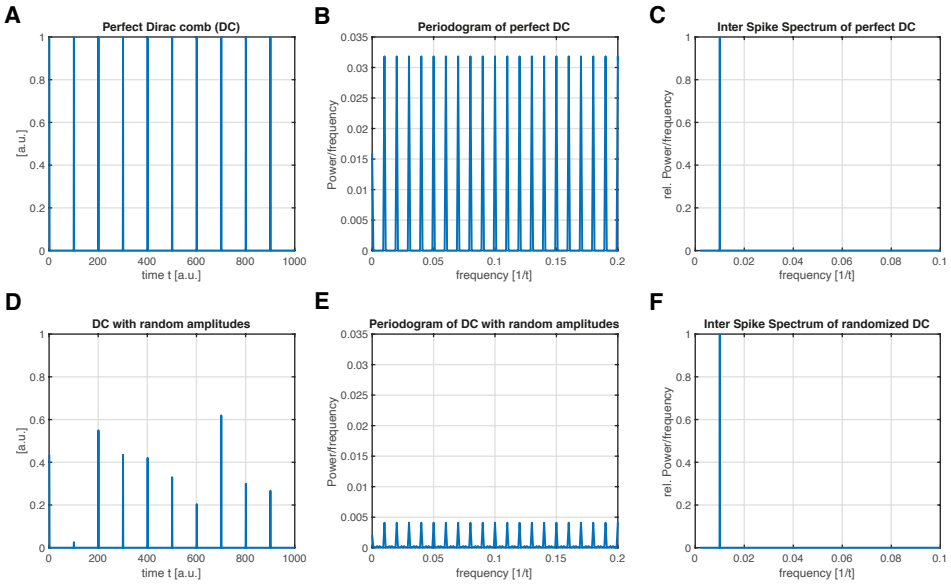

Figure 2. The transformation of a Dirac comb (series of Dirac delta functions) with a single inter-spike period $T_{\text{is}} = 100$ ($\widehat{=} f = 0.01$) into the frequency domain. **A** Dirac Comb (DC) with equal amplitudes and **B** its FFT-based powerspectrum. **C** Proposed inter-spike spectrum of the signal in **A** showing a single frequency, which corresponds to the inter-spike period T_{is} ($f = 0.01$). **D** DC with randomly chosen amplitudes and same T_{is} as in **A**, and **E** its FFT-based powerspectrum. **F** Proposed inter-spike spectrum of the signal in **D** showing a single frequency, which corresponds to the expected inter-spike period T_{is} ($f = 0.01$). Inter-spike spectra were obtained with a LASSO regression and a regularization threshold corresponding to $\rho = 0.9$ accordance of the signals in (**A**,**D**) and its re-composed signals (c.f. Section 2).

exemplify the usage of inter-spike spectrum when transforming the τ-RR of a system under study. In this case, the inter-spike spectrum can unravel characteristic time scales of high dimensional systems, which is not possible when using a FT. Finally, in Section 5 our results are summarized.

2. Method

The signal, which we would like to transform is decomposed into a set of appropriate basis functions.[1] Instead of using trigonometric functions, as it is the idea in the Fourier decomposition, we use Dirac combs (DC) with different inter-spike periods as basis functions, Eq. (4). Let $s(t_i)$ be the normalized signal we want to transform of length N and $t_i = i \cdot \Delta t$, $i = 1,\ldots,N$, where Δt denotes the sampling time and $s(t_i) \in [0,\ 1]\ \forall\ i$. In the following we label this time series as a $(1 \times N)$-dimensional vector $\vec{s}$. First, $\tilde{N}$ different DCs of length N are constructed with inter-spike periods $T_{\text{is}} \in [1,\ldots,\tilde{N}]$ and $\tilde{N} = \lceil N/2 \rceil + 1$. Second, in order to account for possible phase shifts of these basis functions occurring in $\vec{s}$, each of these $\tilde{N}$ different DCs also need to be shifted one step further $T_{\text{is}} - 1$ times. This leaves us with a total number of $M = \sum_{i=1}^{\tilde{N}} i$ basis functions which we can arrange as rows of a $(M \times N)$-sized matrix $\mathbf{X}$ (Fig. 3 illustrates the described procedure):

$$\mathbf{X}_{i,j} = \sum_{k=0}^{N} \delta\big(j - 1 - kT_{\text{is}}(i) - i + T_{\text{is}}(i)\big), \quad i = 1,\ldots,\lceil N/2 \rceil + 1, \quad j = 1,\ldots,N \tag{5}$$

$$T_{\text{is}}(i) = n, \quad s.t.\ n: \frac{n(n-1)}{2} + 1 \leq i < \frac{n(n+1)}{2} + 1, \quad n \in \mathbb{N}_+. \tag{6}$$

Note that due to the shifting of each of the basis functions of inter-spike period T_{is}, $\mathbf{X}$ is not linear independent anymore. Furthermore, there will be identical basis functions and also basis functions, which do not allow for an unambiguous inter-spike period, if we would include all N

[1] We focus on τ-RR in this article, Eq. (2), since such time series are often spiky, which favors the proposed decomposition technique. But the proposed method can be applied to any sort of signal.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

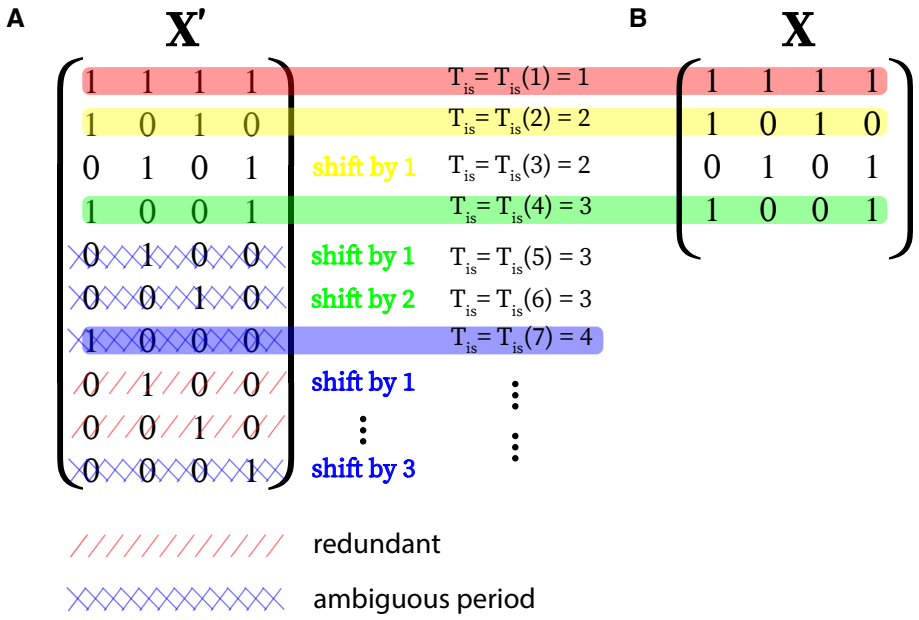

Figure 3. A Example for a full set of basis functions for an input signal of length $N = 4$, aligned in the matrix $\mathbf{X}'$. inter-spike periods T_{is} larger than $\lceil N/2 \rceil + 1$ lead to redundant basis functions (i.e., repeated lines in $\mathbf{X}$, red sheared) or basis functions, which cannot be uniquely assigned to a certain inter-spike period (blue sheared). T_{is} for each row can be obtained by Eq. (6). **B** The final set of unambiguous, but still linearly dependent, basis functions aligned in the matrix $\mathbf{X}$.

possible inter-spike periods for a signal of length N instead of $\lceil N/2 \rceil + 1$ (Fig. 3A). The reason is that in contrast to a trigonometric decomposition, where the Nyquist frequency marks a lower bound for the corresponding wave period, here the maximum considered inter-spike period is bounded by $T_{\text{is}}^{\max} = \lceil N/2 \rceil + 1$ (schematically illustrated in Fig. 3B).

Eventually, an under-determined linear system

$$\mathbf{X}^T \beta = \vec{s} \tag{7}$$

has to be solved for β, a $(M \times 1)$-sized vector carrying the loadings we are interested in. Along a variety of algorithms which can solve this problem, we are particularly interested in those solutions, which promote sparsity in β, since our goal is to decompose the signal $\vec{s}$ into a minimal number of basis functions (for an excellent overview of the topic we refer to Brunton and Kutz [11]). In this paper we either use the *least absolute shrinkage and selection operator* (LASSO) [12] or a *sequentially thresholded least squares* (STLS) regression [11,13] to obtain a solution $\hat{\beta}$. Finally, we group loadings which correspond to basis functions having the same period T_{is} into $\hat{\beta}_f$ and obtain the *inter-spike spectrum* by simply plotting $\hat{\beta}_f$ as a function of the frequency $f = T_{\text{is}}^{-1}$, with $T_{\text{is}} = \Delta t, 2\Delta t, \ldots, (\lceil N/2 \rceil + 1)\Delta t$ (Fig. 2C, F).

In addition to the computational complexity of solving Eq. (7), the solving algorithm must also deal with the free choice of the regression regularization parameter α, which determines the sparsity of the regression coefficient vector $\hat{\beta}$. Obviously, there will be numerous peaks in the inter-spike spectrum of a given signal, when α is small, since $\alpha = 0$ corresponds to the least squares solution and the corresponding $\hat{\beta}$ is an exact solution of Eq. (7) and not sparse. On the other hand, only a sparse number of non-zero loadings in $\hat{\beta}$, i.e., peaks in the inter-spike spectrum, are obtained when α is sufficiently large. The actual implementation of the regularization depends on the regression method chosen (in our case, LASSO or STLS), and consequently a given value of α will generally result in different sparse loading vectors $\hat{\beta}$. Therefore, in this paper we select α such that the re-composed signal $\tilde{\vec{s}} = \mathbf{X}^T \hat{\beta}$ matches a given (Pearson) correlation coefficient ρ between the original signal $\vec{s}$ and itself.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

3. Results

We exemplify the use of the inter-spike spectrum in combination with the τ-RR as outlined in Section 1 on several interesting research questions. The procedure is the following:

(1) Compute a RP of the trajectory of the system, Eq. (1). If only univariate data is available, perform a state space reconstruction for obtaining the trajectory first.
(2) Compute the τ-RR of that RP, Eq. (2).
(3) Transform the τ-RR into the proposed inter-spike spectrum, see Section 2.

3.1. Period estimation for different dynamics in the Rössler system

First, we consider the Rössler system (Eq. (A2) in Appendix A.2) in three different dynamical setups. We use the proposed inter-spike spectrum to identify the type of dynamics. We set the parameters $b = 2$, $c = 4$ and analyze period-2 limit cycle dynamics ($a = 0.36$ in Fig. 4A, D, G, J), period-3 limit cycle dynamics ($a = 0.41$ in Fig. 4B, E, H, K) and chaotic dynamics ($a = 0.428$ in Fig. 4C, F, I, L).

The inter-spike spectra unravel the specific dynamics, which are also apparent in the state space portraits (Fig. 4A, B, C) and in the τ-RRs (Fig. 4G, H, I). The proposed idea is also robust to noise (see Fig. A2 in the Appendix). This is because the peaks of the τ-RR are insensitive to noise. While the peak shape does change in the presence of noise, its position does not, and this is what the inter-spike spectrum encrypts after all.

3.2. Bifurcations in the Logistic map

We consider the Logistic map $x_{n+1} = r \cdot x_n(1 - x_n)$ for changing control parameter r. We vary r from $r = 3.4$ to $r = 4$ in steps of 0.001. For each setting of r

(1) a time series of length $N = 201$ is computed with a random initial condition $u_0 \in [0,\ 1]$, neglecting the first $1,000$ samples as transients,
(2) 100 iterative Amplitude Adjusted Fourier Transform (iAAFT) surrogates [14,15] are computed,
(3) the time series and its iAAFT surrogates are embedded in a 2-dimensional state space using a time delay of unity,
(4) from the 2-dimensional trajectories RPs, Eq. (1), are computed under a threshold $\varepsilon = 0.05$,
(5) τ-RR, Eq. (2), is computed from the RP of the signal and from the RPs of the surrogates,
(6) inter-spike spectra are obtained from τ-RR of the signal and from the τ-RRs of the surrogates, see Section 2, and finally,
(7) from the distribution of the surrogate inter-spike spectra the 95^{th} percentile is computed. The peaks of the inter-spike spectrum of the signal which exceed this percentile are counted.

In this example, the according null hypothesis for constructing the surrogate data is that the data stems from a process which yields the same auto-correlation, hence the same Fourier powerspectrum, and the same amplitude distribution. We consider the number of significant peaks in the inter-spike spectrum with respect to the control parameter in order to distinguish the corresponding dynamics (Fig. 5C). A correlation with the positive Lyapunov exponent (Fig. 5A) is discernible ($\rho_{\text{Pearson}}(\text{Lyapunov}) = 0.72$). This analysis can tackle period-doubling, since it "measures" the dominant cycles via the inter-spike spectrum. However, whenever the periods of the "new" cycles coincide with integer multiples of the periods of already existing cycles, this approach cannot detect period-doubling. A similar case can be observed in Fig. 4J,K where the number of peaks does not change, but their mutual distance does.

A less computationally intensive approach is to compute surrogates for the τ-RR analytically, rather than computing a RP and its τ-RR for each iAAFT surrogate of the time series. This translates into a null hypothesis that the τ-RR and its corresponding inter-spike spectrum stems from a RP of a random signal. In this case the probability of finding a black point in the RP can be obtained from a binomial distribution with probability parameter p set to the recurrence rate of the RP of the signal. This way 100 surrogate τ-RRs are computed in step (5). The results are even slightly better compared to the ones obtained from the iAAFT surrogates (Fig. 5B, $\rho_{\text{Pearson}}(\text{Lyapunov}) = 0.81$). The first period doubling at $r \approx 3.458$ cannot be detected by any of the surrogates.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

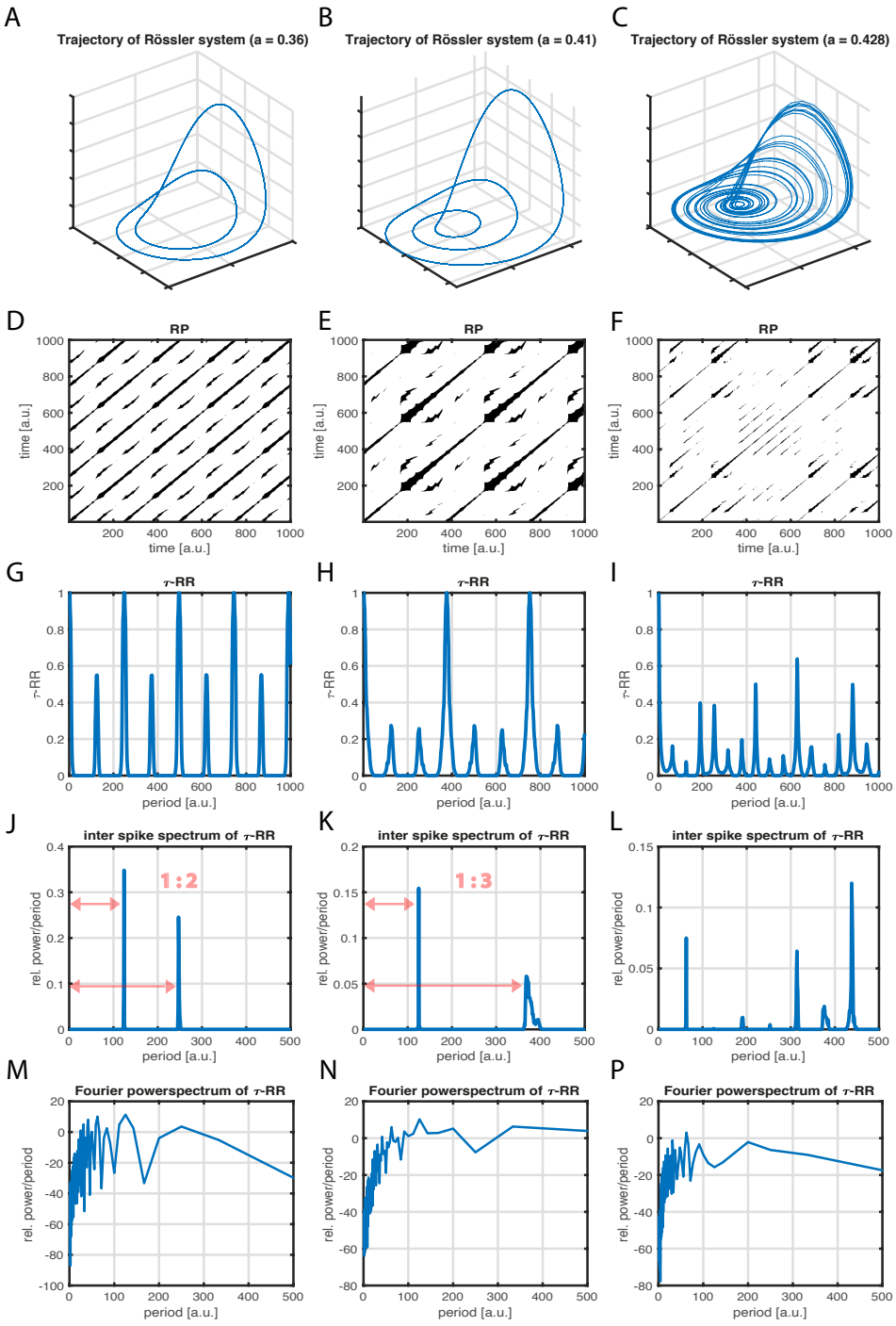

Figure 4. Inter-spike spectra of the τ-RR of the Rössler system in three different dynamical regimes with parameters $b = 2$, $c = 4$. **A** Trajectory of the system in a period-2 (parameter $a = 0.36$), **B** in a period-3 (parameter $a = 0.41$) and **C** in a chaotic regime (parameter $a = 0.428$). **D, E, F** The corresponding RPs, obtained by using a recurrence threshold corresponding to a 10% global recurrence rate for D & E and 5% for F. **G, H, I** τ-RRs of the shown RPs. **J, K, L** The proposed inter-spike spectra of the τ-RRs shown in panels G, H, I. Spectra were obtained with a LASSO regression and a regularization threshold corresponding to $\rho = 0.95$ accordance of τ-RRs and re-composed signals. The distance ratio of the peaks reflect the limit cycle dynamic. **M, N, P** Fourier power spectra of the τ-RRs shown in panels G, H, I.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

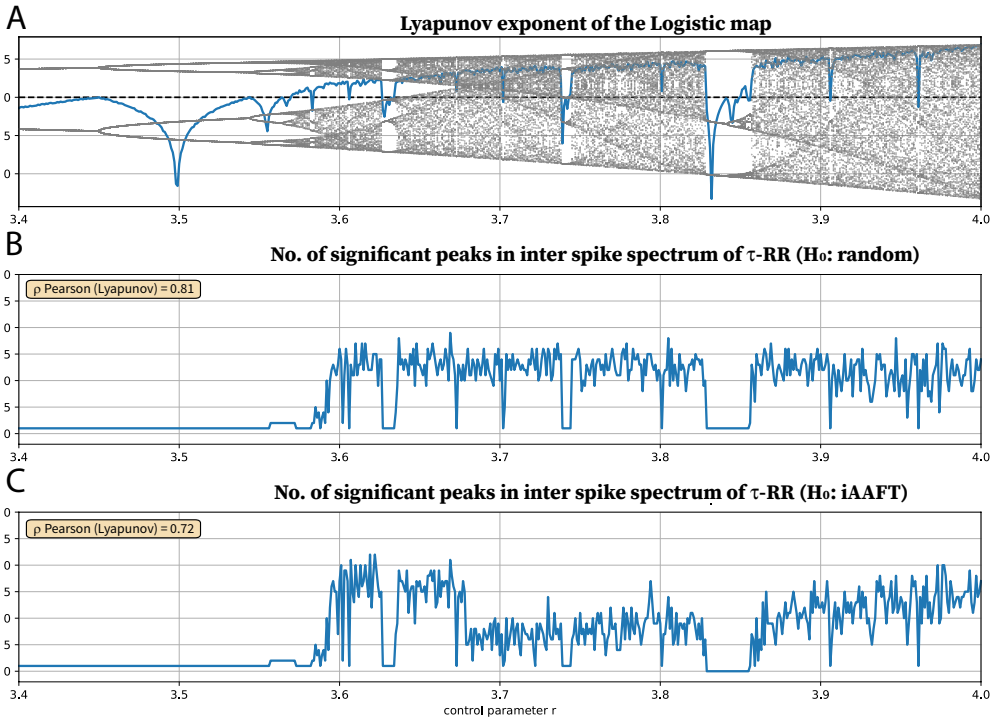

Figure 5. A Bifurcation diagram and Lyapunov exponent and of the Logistic map as a function of the control parameter r. **B** Number of significant peaks ($\alpha = 0.05$) in the inter-spike spectrum of the τ-RR and its Pearson correlation coefficient to the Lyapunov exponent shown in **A** (white noise surrogates). **C** Same as **B**, but for iterative Amplitude Adjusted Fourier Transform (iAAFT) surrogates [14,15]. For obtaining the inter-spike spectra we used a LASSO regression and a regularization threshold corresponding to $\rho = 0.95$ accordance of τ-RRs and re-composed signals.

The described procedure does work well for map data, because most often the τ-RR for those kind of data reveals a "spiky enough" nature. On the contrary, highly sampled (flow-) data often yield not as spiky τ-RRs and, thus, the number of significant peaks in the inter-spike spectrum may not be sensitive enough to detect period-doubling bifurcations. Moreover the sensitivity of the inter-spike spectrum to detect the regime shifts also depends on the critical regularization threshold. Nevertheless the according inter-spike spectra is still revealing important information (Fig. 4) and practitioners can design appropriate quantifying statistics based on these spectra, which suit the research task.

3.3. inter-spike spectra of power grid frequency data

Power grids are large, synchronized, complex networks whose stable functioning is indispensable for modern societies. To maintain the stability of a power grid, the balance between energy consumption and energy generation must be ensured. In an AC-power grid, the grid frequency is an observable variable that reflects how well this balance is satisfied. In this process, the grid frequency and its deviations from the nominal frequency are continuously recorded and monitored by the grid operators (in Europe and many parts of the world this is 50 Hz or 60 Hz in America and, for example, southern Japan). For example, if there is more (less) demand than supply, the network frequency decreases (increases) compared to the nominal frequency [16].
The frequency variations can include other information, such as the functionality of control systems [17], the effect of fluctuations in renewable energies (REs), demands on the grid [18] and, moreover, the effect of regular dispatches due to the trading market [19]. The latter induce periodic frequency jumps. Here we look at the frequency time series for the Great Britain (GB) and Continental Europe (CE) (Appendix B and Fig. A1A,C). Clear jumps every 30 and 60 minutes are discernible and quantitatively reflected in the corresponding autocorrelations (Fig. A1B,D). Furthermore, the autocorrelation of the CE frequency time series shows regular peaks every 15 min (see Fig. A1B).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

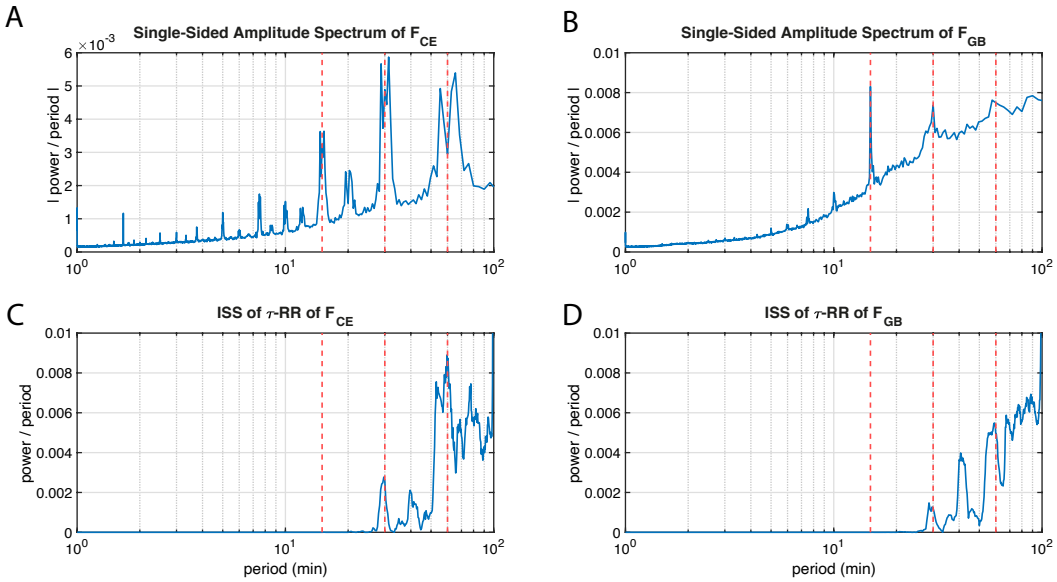

Figure 6. Averaged Fourier power spectra of recorded power grid frequency time series of **A** Central Europe (CE) and **B** Great Britain (GB) (see Appendix B). The corresponding averaged inter-spike spectra of the according τ-RRs, Eq. (2), are shown in **C** and **D**, respectively. Vertical red dashed lines correspond to 15, 30 and 60 min. For technical details on the calculation of the spectra shown, the reader is referred to the Appendix B.

These peaks are caused by a mismatch of power supply and demand [20] during dispatches. In most electricity grids the operation of dispatchable power plants is scheduled in 1-hour blocks, where additional (shorter) 30 and 15 min intervals might exist.

The Fourier frequency spectra for the Central Europe data in Figure 6A, however, does not display sharp peaks exactly at 15, 30 and 60 min (for technical details on the calculation of the spectra shown, the reader is referred to the Appendix B). Especially the 30 and 60 min peaks are split into two adjacent peaks and the local minimum in between these "double"-peaks correspond to the exact times. The Great Britain counterpart in Fig. 6B has sharp peaks at 15 and 30 min and also "local predecessor peaks" for 30 and 60 min at the same positions as in panel A. In contrast, the 15 min peak is completely absent from the inter-spike spectra of the τ-RRs of the frequency data (Fig. 6C,D), but these show sharp peaks at 30 and 60 min. In front of these peaks, smaller local peaks can be seen, which correspond to the local peaks of the Fourier power spectra at 28 and 55 min, respectively. Moreover, there is an additional peak at 40 min for both datasets, which is absent in the Fourier spectrum and which is not a multiple of the missing 15 min peak. The position and magnitude of the peaks in the shown inter-spike spectra are robust to the chosen recurrence threshold, the regression method and its regularization as well as the sampling time of the original signal.

We interpret the results presented as follows. The missing 15 min in the inter-spike spectra is due to the much stronger autocorrelation at 30 and 60 min (see Fig. A1B,D) and because these periods are integer multiples of 15 min the inter-spike spectra are not able to detect it and the sparse regression "drags" the 15 min periods into the mentioned 30 and 60 min peaks. Moreover, the inter-spike spectra, unlike the Fourier spectra demonstrates sharp peaks exactly at 30 and 60 min, i.e. during dispatches (not valid in case of the GB dataset). Eventually, here we have found a clear sharp peak at 40 min which can occur because of any regular controls in a power grid, and can be a hint to develop the existing stochastic processes to model precisely the power grid frequency [17].

3.4. Evolutionary inter-spike spectra of Earth's orbit data

When applying the proposed inter-spike spectrum to the τ-RR of a time series we expect additional frequency/period information, due to the fact that the recurrence plot (RP), Eq. (1), and its corresponding τ-RR, Eq. (2), visualize the trajectory of the embedded time series in an embedding

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

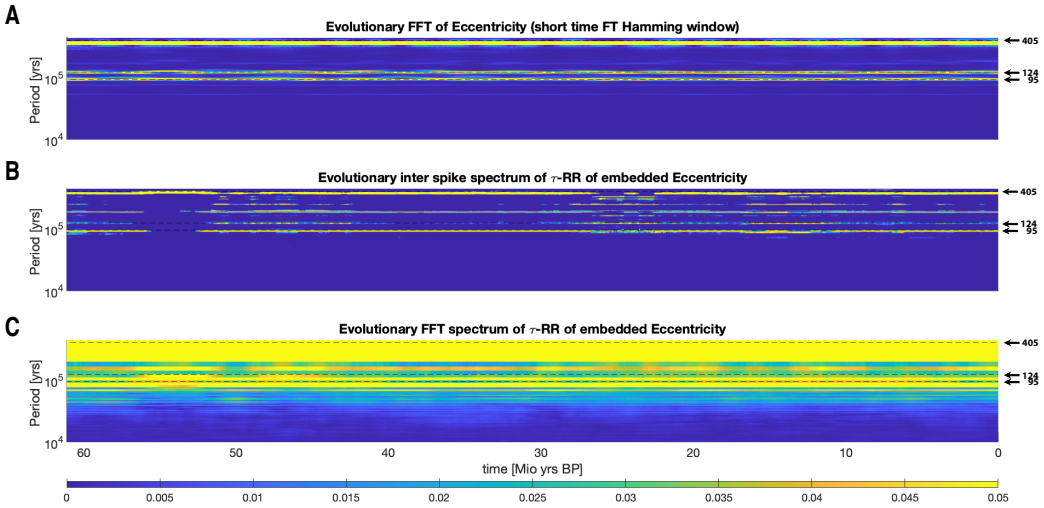

Figure 7. A Evolutionary Fourier power spectra of eccentricity time series. **B** Inter-spike spectrogram of the τ-recurrence rate of the eccentricity time series and *C* its Fourier spectrogram. Horizontal black dashed lines highlight the analytically expected orbital periods of 405, 124 and 95 kyrs. For comparability, in all cases the spectra aligned in the columns of the shown plots are normalized to probabilities (sum of unity for each power spectrum). For further computational details, please refer to the main text.

space of higher dimension. However, given a sufficient embedding of the time series, we would also expect that major frequencies/periods of the non-embedded time series are incorporated in the RP, its τ-RR, and eventually in the inter-spike spectrum of the τ-RR. In order to demonstrate this we apply the inter-spike spectrum to the freely available eccentricity time series of Laskar, J. *et al.* [21]. This astronomical computation of the orbital motion of the Earth (here we focus on the eccentricity only) has a clear expectation value of the incorporated frequencies/periods. The three leading eccentricity cycles of 405 kyr period, 95 kyr period and 124 kyr period are well known in palaeoclimate studies [22,23]. Our aim in this section is to show that the inter-spike spectrum of the τ-RR of the embedded eccentricity time series will reflect these cycles in a similar fashion as the Fourier powerspectrum of the non-embedded eccentricity time series. We will further show that Fourier-transforming the τ-RR instead of applying the proposed inter-spike spectrum will lead to non-satisfying results, since the spiky τ-RR excites a variety of harmonics in the corresponding Fourier spectrum (cf. Section 1). We use a time series which covers the past ~67 Mio. years (Myr), with a total length of $N = 13,421$ samples and a sampling period of $\Delta t = 5,000$.

First we compute an evolutionary short time Fourier transform (FT) using a windowsize of $ws = 1,000$ samplepoints ($\equiv 5$ Myr) shifted by unity and a Hamming window. The spectrogram reveals the expected periods mentioned, which are highlighted and clearly visible (Fig. 7A).

Then we construct the inter-spike spectrogram of the τ-RR by first determining an appropriate embedding. By using a recent Tree-Embedding-Ansatz [24] we minimize the *false nearest neighbor* statistic [25] and use the continuity statistic [26] for potential delays. Eventually we obtain an 8-dimensional embedding with delays $\tau s = [0, 18, 36, 48, 62, 101, 114]$ (in sampling units) for the entire time series. Similar to the preceding approach we embed the time series with these embedding parameters on windows of size $ws = 1,200$ and a unity shift.[2] The RPs are computed on these embedded trajectories with a fixed recurrence threshold corresponding to 10% global recurrence rate [2] and the inter-spike spectra of the corresponding τ-RRs is obtained by using STLS regression (see Section 2) and a regularization threshold corresponding to $\rho = 0.9$ accordance of τ-RRs and re-composed signals. We only use the first 200 data points of the τ-RR (covering a time span of 1 Myr). The spectrogram also highlights the 95, 124 and 405 kyr periods as expected (Fig. 7B). Additional power is distributed in its harmonics at 190 and 248 kyr periods. Finally, the standard FT with a Hamming window of the same τ-RRs used to obtain the inter-spike spectrogram in panel B,

[2] We use a slightly larger window here than in the short FT, because the embedding causes a "loss" of data points and we want to cover similar time spans.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

yields a smeared spectrogram which does not reflect the expected periods, but rather suffers from the spike train behavior, i.e., many excited harmonics, of the FT described in Section 1 and Fig. 2. The shown results are robust to a change of embedding parameters and windowsizes. However, a too low regularization threshold smears the clear spectrogram in panel B.

4. Discussion

We have successfully used the idea of transforming the τ-RR for the detection of bifurcations in the Logistic map. By constructing appropriate surrogates of the inter-spike spectra, and thus a null model, the number of significant peaks in the inter-spike spectrum correlated well with the positive Lyapunov exponent. This measure was also able to resolve period-doubling bifurcations. However, we have to admit this way the detection of a bifurcation is only possible when the additional period(s) is not an integer multiple of the former period(s). This behavior is described in the application to the Rössler system, where we explicitly showed the different inter-spike power spectra for period-2, period-3, and chaotic dynamics. Further development might potentially incorporate the mutual distance of peaks in the spectrum for a better correlation to the Lyapunov exponent. The inter-spike spectra of power grid frequency data illustrate that our proposed method may serve as a valuable source of information in addition to a standard Fourier analysis. And last but not least, we showed that our approach really incorporates frequencies, which are apparent in the Fourier spectrum of the signal, by applying it to analytically derived eccentricity data, where the dominant frequencies are well known.

We discuss some more technical details in the following, which will also affect any application of the proposed method. First of all, the number of required basis functions $M = \sum_{i=1}^{\lceil N/2 \rceil + 1} i$ for an input signal of length N is the crucial bottleneck of this approach, which is why it does not show good scaling behavior. The subsequent sparse regression, therefore, gets computationally intensive for $N > 1{,}000$. Depending on the memory of the computer used, input signals $N > 2{,}000$ usually do not work anymore. This means that signals often need to get downsampled as a preprocessing step (e.g., see Appendix B). Second, the regularization parameter α for the regression is a crucial free parameter. As described in Sect. 2, our idea in this paper is to select the α such that the re-composed signal $\tilde{\vec{s}} = \mathbf{X}^T \hat{\beta}$ matches a given (Pearson) correlation coefficient $\rho_{\vec{s},\tilde{\vec{s}}}$ between the original signal $\vec{s}$ and itself. This ensures that α adjusts itself to the data as well as to the used regression method. We found that this increases the comparability of different spectra, especially when performing a running window approach in order to obtain an evolutionary spectrogram (Fig. 7). However, the two different sparse regression algorithms we encountered in this article (LASSO and STLS) yield different results for the same desired $\rho_{\vec{s},\tilde{\vec{s}}}$. Even if the spectra obtained in this way look qualitatively similar, they are not always quantitatively similar. The reason for this is that $\rho_{\vec{s},\tilde{\vec{s}}}$ is not a smooth function of α in case of STLS, due to the hard-thresholding involved [13], which is shown in Fig. A3. Third, when adopting our idea of applying the inter-spike spectrum to the τ-recurrence rate of the signals state space trajectory, the embedding process induces additional free parameters. This is not a drawback of the proposed method, but rather a drawback of the particular application method that we have heavily used in this article and which was the main motivation for developing the proposed method. As a very last remark, we draw attention to the fact that sparse regression can be transformed into sparse logistic regression when the signal we would like to transform is binary.

5. Conclusion

A novel type of powerspectrum, the *inter-spike spectrum*, has been proposed. The method decomposes any arbitrary signal into basis functions which consist of (lagged) Dirac combs (DC) of different inter-spike period. The loading for each period is obtained by a regularized regression, which promotes sparsity in its solution. Wee chose LASSO or a sequentially thresholded least squares regression STLS in this letter. Since there are $M = \sum_{i=1}^{\lceil N/2 \rceil + 1} i$ basis functions for a signal of length N the regression can get computationally intensive for $N > 1{,}000$. When plotting the computed loadings as a function of the period (or frequency) the inter-spike spectrum is obtained. Although this novel power spectrum is superior to an ordinary FFT-based power spectrum when the signal has a spike-train-like appearance, the authors suggest that this method should be considered as an additional source of information but not as a substitute for ordinary Fourier analysis. Due to

 doi:10.20944/preprints202209.0054.v1

the sparse regression underlying the method, there is no unique inverse of the transformation and the regularization parameter plays a crucial role and determines the appearance of the obtained inter-spike spectrum. Moreover, similar to the Nyquist frequency barrier in the Fourier Transform which sets a lower bound for the corresponding wave period, here the maximum considered inter-spike period is bounded by $T_{\text{is}}^{\max} = \lceil N/2 \rceil + 1$.

The invention of the proposed method has been motivated by the idea of transforming τ-recurrence rate signals (τ-RRs) into their frequency domain. This general idea [7] allows for a frequency analysis of high dimensional systems, because the RP is a representation of the system's state space trajectory. The τ-RR of a recurrence plot (RP) usually has a spiky shape, especially for map-like data, and the inter-spike spectrum can reliably reveal the system's dominant frequencies, which is not possible when Fourier transforming the τ-RR or the underlying signal itself. Since the position of the peaks in the τ-RR are not sensitive to noise, the corresponding inter-spike spectrum also yields robust results in the presence of noise.

We could think of a broad range of applications of the proposed idea. The inter-spike spectrum itself can serve as a valuable tool for the analysis of any sort of spike-train-like data. On the other hand, the inter-spike spectrum of the τ-RR of a signal can serve as a generalized, nonlinear frequency analysis tool for complex systems. When there is only a subset of state variables available, the state space has to be reconstructed as a pre-processing step. Recent findings [8,24] show that this reconstruction process can be reliably automated and applied to multivariate data as well. This would allow for a "running window" approach, in order to detect transitions. Due to the mentioned computational constraints of our proposed method, a window size $w \leq 1,000$ would possibly suffice for most data, especially when it is map-like, i.e., not highly sampled.

Author Contributions: Conceptualization, KHK and FH; methodology, KHK, FH and NM; software, KHK and FH; validation, KHK, FH, MA and NM; formal analysis, KHK; data curation, KHK and MA; writing—original draft preparation, KHK; writing—review and editing, KHK, FH, MA and NM; visualization, KHK; supervision, NM; project administration, JK; All authors have read and agreed to the published version of the manuscript.

Funding: This work has been financially supported by the German Research Foundation (DFG projects MA4759/8 and MA4759/9).

Data Availability Statement: The study that we present here is available as a fully reproducible code base **Repository will be published and cited here, when accepted** and the method will be available in the Julia language **Package will be published and cited here, when accepted** and as a MATLAB® toolbox **Toolbox will be published and cited here, when accepted**.

Acknowledgments: All computations have been carried out in MATLAB® and the Julia language and made use of the packages *DynamicalSystems.jl* [27] and *DifferentialEquations.jl* [28].

Conflicts of Interest: The authors declare that they have no conflict of interest.

 doi:10.20944/preprints202209.0054.v1

Appendix A. Exemplary models

Appendix A.1. Lorenz system

The classical Lorenz-63 system [29] is defined as

$$\begin{aligned} \dot{x} &= \sigma(y-x) \\ \dot{y} &= x(r-z)-y \\ \dot{z} &= xy-\beta z. \end{aligned} \tag{A1}$$

For producing Fig. 1 we set the initial condition to $u_0 = [0.0, 10.0, 0.0]$, used a sampling time of $\Delta t = 0.01$ and discarded the first 2,000 points of the integration as transients. The parameters have been set to $\sigma = 10, \beta = 8/3, \rho = 28$ and we used a time series consisting of 6,000 samples.

Appendix A.2. Rössler system

The Rössler system [30] is defined as

$$\begin{aligned} \dot{x} &= -y-z \\ \dot{y} &= x+ay \\ \dot{z} &= b+z(x-c). \end{aligned} \tag{A2}$$

For producing Figs. A2 and 4, the initial condition for producing panels **A** & **B** was set to $u_0 = [0.7, -1, 0.4]$ with a sampling time of $dt = 0.05$ and in case of panel **C**, $u_0 = [-0.1242, -2.5415, 0.2772]$ with a sampling time of $dt = 0.1$. The first $5,000$ samples were discarded as transients and trajectories of length $N = 5,000$ were obtained from which we computed the RPs and the corresponding τ-RRs. For the inter-spike spectra only the first $1,000$ values of the τ-RRs were considered.

Appendix B. Power grid frequency time series

The raw frequency time series have length $\tilde{N}_{\mathrm{CE}} = 58,752,000$ and $\tilde{N}_{\mathrm{GB}} = 31,622,400$ with sampling times $\Delta t_{\mathrm{CE}} = 0.2$s and $\Delta t_{\mathrm{GB}} = 1$s, respectively. GB frequency data was measured during 2016, and has been gotten from [31]. CE frequency data belongs to 2017 and has been gotten from [32].

We downsampled these time series to a sampling time of $\Delta t_{\mathrm{CE}} = \Delta t_{\mathrm{GB}} = 20$s, which led to total time series lengths of $N_{\mathrm{CE}} = 587,520$ and $N_{\mathrm{GB}} = 1,581,120$ which we used for the further analysis.

In this analysis we divided the time series into non-overlapping blocks of length $N_{\mathrm{block}} = 4,320$, covering a time span of 24 hours. For each time series block we computed Fourier spectra and recurrence plots (RPs), Eq. (1), along with their τ-RRs, Eq. 2. The RPs were obtained from a uniform 5-dimensional time-delay embedding of each time series with the delay set to the first minimum of the mutual information [25,33] and a fixed recurrence threshold corresponding to 8% recurrence rate was used in order to ensure comparability [2]. The first 600 data points of the τ-RRs were used for finally obtaining inter-spike spectra with a LASSO regression and a regularization threshold corresponding to $\rho = 0.95$ accordance of τ-RRs and re-composed signals. The spectra shown in all panels of Figure 6 are averages over all blocks, following Meyer *et al.* [19].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

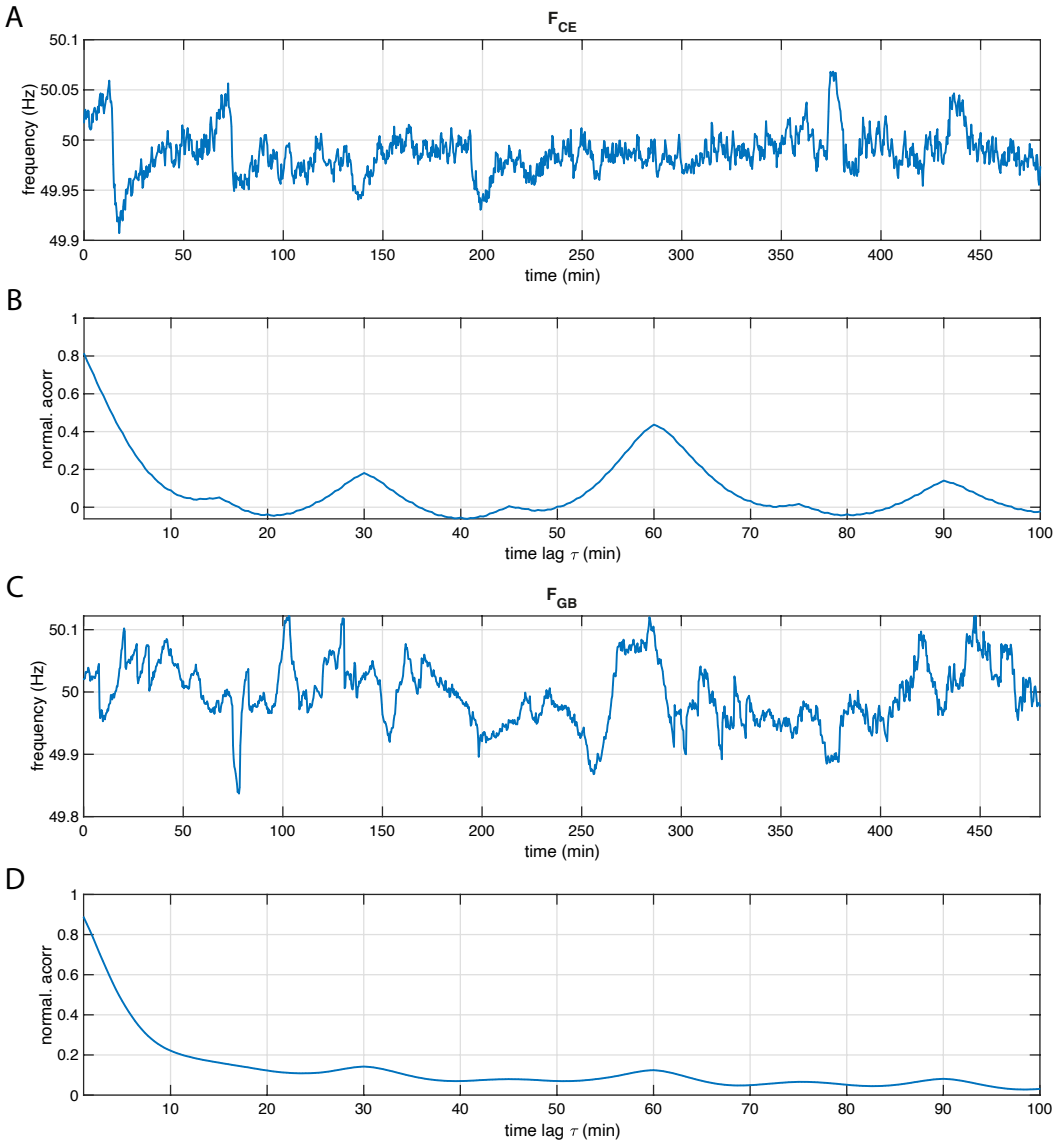

Figure A1. Subset of the frequency time series of **A** Central Europe (CE) [32] and **C** Great Britain (GB) [31] along with their autocorrelation functions in **B** and **D**, respectively. The normalized autocorrelations in **B**, **D** have been computed on the entire time series, wheres as the subsets shown in **A**, **C** contain only 1,441 samples.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

Appendix C. inter-spike spectra for noisy Rössler system

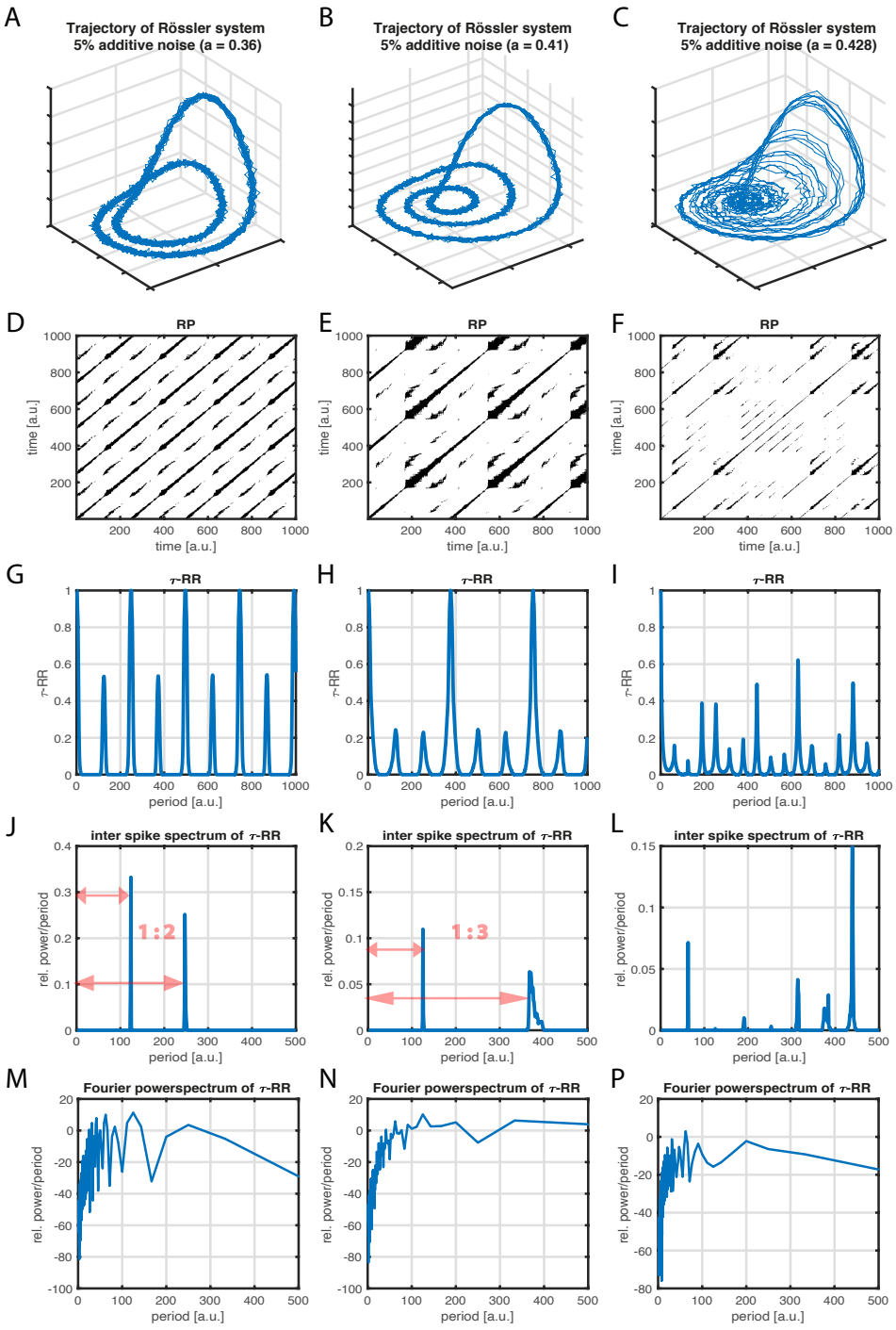

Figure A2. Same as in Fig. 4, but here with 5% additive Gaussian white noise on each component *x*, *y* and *z*. The appearance of the inter-spike spectra in **J, K, L**, and the Fourier spectra in **M, N, P** are unaffected by the additive noise.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 September 2022 doi:10.20944/preprints202209.0054.v1

Appendix D. Regularizationparameters for different sparse regression techniques

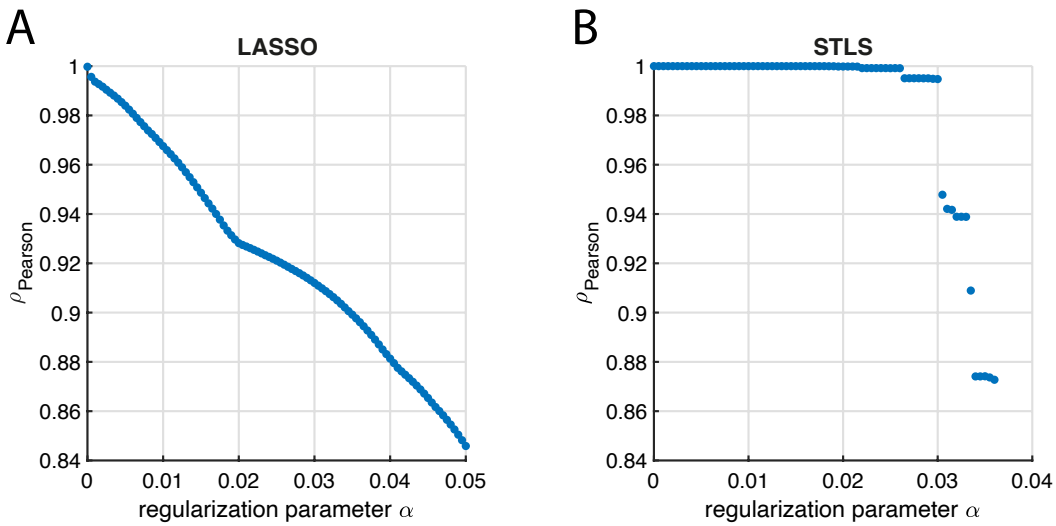

Figure A3. Pearson correlation coefficient $\rho_{\vec{s},\tilde{\vec{s}}}$ between the input signal signal $\vec{s}$ and the re-composed signal $\tilde{\vec{s}} = \mathbf{X}^T \hat{\beta}$ as a function of the regularization parameter α for **A** LASSO and **B** sequentially thresholded least squares (STLS) regression (see also Section 2). The input time series is of length $N = 200$ and stems from the τ-RR of Rössler system in regular dynamics.

References

1. Marwan, N.; Romano, M.C.; Thiel, M.; Kurths, J. Recurrence Plots for the Analysis of Complex Systems. *Physics Reports* **2007**, *438*, 237–329. https://doi.org/10.1016/j.physrep.2006.11.001.
2. Kraemer, K.H.; Donner, R.V.; Heitzig, J.; Marwan, N. Recurrence threshold selection for obtaining robust recurrence characteristics in different embedding dimensions. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2018**, *28*, 085720, [https://doi.org/10.1063/1.5024914]. https://doi.org/10.1063/1.5024914.
3. March, T.K.; Chapman, S.C.; Dendy, R.O. Recurrence plot statistics and the effect of embedding. *Physica D* **2005**, *200*, 171–184. https://doi.org/10.1016/j.physd.2004.11.002.
4. Baptista, M.S.; Ngamga, E.J.; Pinto, P.R.F.; Brito, M.; Kurths, J. Kolmogorov-Sinai entropy from recurrence times. *Physics Letters A* **2010**, *374*, 1135–1140. https://doi.org/10.1016/j.physleta.2009.12.057.
5. Kraemer, K.H.; Marwan, N. Border effect corrections for diagonal line based recurrence quantification analysis measures. *Physics Letters A* **2019**, *383*, 125977. https://doi.org/https://doi.org/10.1016/j.physleta.2019.125977.
6. Marwan, N.; Kurths, J. Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A* **2002**, *302*, 299–307. https://doi.org/10.1016/S0375-9601(02)01170-2.
7. Zbilut, J.P.; Marwan, N. The Wiener–Khinchin theorem and recurrence quantification. *Physics Letters A* **2008**, *372*, 6622–6626. https://doi.org/https://doi.org/10.1016/j.physleta.2008.09.027.
8. Kraemer, K.H.; Datseris, G.; Kurths, J.; Kiss, I.Z.; Ocampo-Espindola, J.L.; Marwan, N. A unified and automated approach to attractor reconstruction. *New Journal of Physics* **2021**, *23*, 033017. https://doi.org/10.1088/1367-2630/abe336.
9. Schild, D. An Efficient Method for the Fourier Transform of a Neuronal Spike Train. *International Journal of Neuroscience* **1982**, *17*, 179–182, [https://doi.org/10.3109/00207458208985921]. https://doi.org/10.3109/00207458208985921.
10. Córdoba, A. Dirac combs. *Letters in Mathematical Physics* **1989**, *17*, 191–196. https://doi.org/10.1007/BF00401584.
11. Brunton, S.L.; Kutz, J.N. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*; Cambridge University Press, 2019. https://doi.org/10.1017/9781108380690.
12. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* **1996**, *58*, 267–288, [https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.2517-6161.1996.tb02080.x]. https://doi.org/https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.
13. Brunton, S.L.; Proctor, J.L.; Kutz, J.N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* **2016**, *113*, 3932–3937, [https://www.pnas.org/doi/pdf/10.1073/pnas.1517384113]. https://doi.org/10.1073/pnas.1517384113.
14. Schreiber, T.; Schmitz, A. Improved Surrogate Data for Nonlinearity Tests. *Phys. Rev. Lett.* **1996**, *77*, 635–638. https://doi.org/10.1103/PhysRevLett.77.635.
15. Schreiber, T.; Schmitz, A. Surrogate time series. *Physica D: Nonlinear Phenomena* **2000**, *142*, 346–382. https://doi.org/https://doi.org/10.1016/S0167-2789(00)00043-9.

16. Kundur, P.; Balu, N.J.; Lauby, M.G. *Power system stability and control*; Vol. 7, McGraw-hill New York, 1994.
17. Gorjão, L.R.; Anvari, M.; Kantz, H.; Beck, C.; Witthaut, D.; Timme, M.; Schäfer, B. Data-driven model of the power-grid frequency dynamics. *IEEE access* **2020**, *8*, 43082–43097.
18. Anvari, M.; Gorjão, L.R.; Timme, M.; Witthaut, D.; Schäfer, B.; Kantz, H. Stochastic properties of the frequency dynamics in real and synthetic power grids. *Physical review research* **2020**, *2*, 013339.
19. Meyer, P.G.; Anvari, M.; Kantz, H. Identifying characteristic time scales in power grid frequency fluctuations with DFA. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2020**, *30*, 013130.
20. Weißbach, T.; Welfonder, E. High frequency deviations within the European power system: Origins and proposals for improvement. In Proceedings of the 2009 IEEE/PES Power Systems Conference and Exposition. IEEE, 2009, pp. 1–6.
21. Laskar, J..; Fienga, A..; Gastineau, M..; Manche, H.. La2010: a new orbital solution for the long-term motion of the Earth. *A&A* **2011**, *532*, A89. https://doi.org/10.1051/0004-6361/201116836.
22. Laskar, J..; Robutel, P..; Joutel, F..; Gastineau, M..; Correia, A. C. M..; Levrard, B.. A long-term numerical solution for the insolation quantities of the Earth. *A&A* **2004**, *428*, 261–285. https://doi.org/10.1051/0004-6361:20041335.
23. Westerhold, T.; Marwan, N.; Drury, A.J.; Liebrand, D.; Agnini, C.; Anagnostou, E.; Barnet, J.S.K.; Bohaty, S.M.; De Vleeschouwer, D.; Florindo, F.; et al. An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science* **2020**, *369*, 1383–1387, [https://science.sciencemag.org/content/369/6509/1383.full.pdf]. https://doi.org/10.1126/science.aba6853.
24. Kraemer, K.H.; Gelbrecht, M.; Pavithran, I.; Sujith, R.I.; Marwan, N. Optimal state space reconstruction via Monte Carlo decision tree search. *Nonlinear Dynamics* **2022**, *108*, 1525–1545. https://doi.org/10.1007/s11071-022-07280-2.
25. Hegger, R.; Kantz, H. Improved false nearest neighbor method to detect determinism in time series data. *Phys. Rev. E* **1999**, *60*, 4970–4973. https://doi.org/10.1103/PhysRevE.60.4970.
26. Pecora, L.M.; Moniz, L.; Nichols, J.; Carroll, T.L. A unified approach to attractor reconstruction. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2007**, *17*, 013110, [https://doi.org/10.1063/1.2430294]. https://doi.org/10.1063/1.2430294.
27. Datseris, G. DynamicalSystems.jl: A Julia software library for chaos and nonlinear dynamics. *Journal of Open Source Software* **2018**, *3*, 598. https://doi.org/10.21105/joss.00598.
28. Rackauckas, C.; Nie, Q. Differentialequations.jl–a performant and feature-rich ecosystem for solving differential equations in julia. *Journal of Open Research Software* **2017**, *5*.
29. Lorenz, E.N. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences* **1963**, *20*, 130–141. https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.
30. Rössler, O. An equation for continuous chaos. *Physics Letters A* **1976**, *57*, 397 – 398. https://doi.org/https://doi.org/10.1016/0375-9601(76)90101-8.
31. National Grid, Frequency data (2014–2018). Accessed: 2022-07-25.
32. Haehne, H.; Schottler, J.; Waechter, M.; Peinke, J.; Kamps, O. The footprint of atmospheric turbulence in power grid frequency measurements. *EPL (Europhysics Letters)* **2018**, *121*, 30001.
33. Fraser, A.M.; Swinney, H.L. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* **1986**, *33*, 1134–1140. https://doi.org/10.1103/PhysRevA.33.1134.