Recurrence complexity analysis of oscillatory signals with application to general anesthesia EEG signals

Mariia Fedotenkova, Peter Beim Graben, Tamara Tošić, Jamie Sleigh, Axel Hutt

Recurrence complexity analysis of oscillatory signals with application to general anesthesia EEG signals

Mariia Fedotenkova*[a,b,c], Peter beim Graben[d], Tamara Tošić[a,b,c], Jamie Sleigh[e], and Axel Hutt[f]

[a]NEUROSYS team, INRIA, Villers-lès-Nancy, F-54600, France
[b]UMR n° 7503, CNRS, Loria, Vanduvre-lès-Nancy, F-54500, France
[c]Université de Lorraine, Villers-lès-Nancy, F-54600, France
[d]Bernstein Center for Computational Neuroscience, Berlin, Germany
[e]Waikato Clinical School of the University of Auckland, New Zealand
[f]Deutscher Wetterdienst, Offenbach am Main, Germany
*maria.fedotenkova@gmail.com

May 4, 2016

Abstract

Recurrence structures in univariate time series are challenging to detect. We propose a combination of recurrence and symbolic analysis in order to identify such structures in a univariate signal. This method allows to obtain symbolic representation of the signal and quantify it by calculating its complexity measure. To this end, we propose a novel method of phase space reconstruction based on the signal's time-frequency representation and show that the proposed method outperforms conventional phase space reconstruction by delay embedding techniques. We evaluate our method on synthetic data and show its application to experimental EEG signals.

Keywords: recurrence analysis, recurrence complexity, symbolic dynamics, time-frequency representation, EEG, general anesthesia.

1 Introduction

Recurrent temporal dynamics is a phenomenon frequently observed in time series measured in complex nonlinear systems. Examples range from intermittent behavior in atmospheric aerosols [3] over bird songs exhibiting recurring temporal structures [46] to returning epileptic seizures [2] and recurrent cognitive states in neural language processing [5] and in early auditory neural processing [19]. To detect such temporal recurrence structures, typically one applies recurrence analysis based on recurrence plots introduced by Eckmann [9, 31], which take their origin in Poincaré's theorem [35]. This approach allows to detect recurrence structures in multi-dimensional time series.

However, in some cases a single univariate time series is available only. For instance, electroencephalographic activity (EEG) on the scalp may be measured by a single electrode only, as typically applied in surgery under general anesthesia. To retrieve recurrence structures from such univariate time series, several methods have been suggested, such as delay embedding techniques [41]. This method estimates the embedding dimension and the time delay from the data, but does not take into account the data's temporal structure. This latter disadvantage is a

*Corresponding author

strong limitation as will be seen in the subsequent sections since experimental signals obtained in many complex systems exhibit oscillatory behavior. However, most previous methods do not take into account specifically the oscillatory nature of the signals. To improve recurrence analysis, we consider the oscillatory nature of time series in recurrence analysis and propose a technique to embed the univariate time series in a multi-dimensional space reflecting the time series oscillatory activity. The approach is based on the signals time-frequency representation.

In a previous work we have sketched this approach in a different context [44] but without discussing its performance subject to different time-frequency representations. Such time-frequency representations are the results of spectral analysis techniques that resolve the spectral power in time. Examples for such techniques are spectrogram, multi-taper or wavelet analysis. In contrast to the previous work, where a single specific time-frequency representation is applied, the present work compares results utilizing different time-frequency representations from spectrogram and two variants of wavelet analysis. Moreover, we compare the obtained results with results utilizing a conventional embedding technique. As a new application of the spectral recurrence analysis introduced in this paper, we show that the method may provide a novel approach to distinguish systems based on the differences in their recurrent activity by considering the signals recurrence complexity. Applications to synthetic data allows to assess the method in a comparison to results obtained from a conventional technique. This evaluation permits to reveal the temporal complexity structures in experimental EEG obtained under general anesthesia and to distinguish different anesthetics administered on the basis of the measured data.

2 Materials and Methods

2.1 Symbolic Recurrence Structure Analysis

The core method of this paper is the symbolic recurrence structure analysis (SRSA). Its aim is to identify transient and metastable states in the system's dynamics based on its recurrence. The first part of the method originates form Poincaré's recurrence theorem [35] which can be visualized by the means of recurrence plots (RPs), first introduced by Eckmann et al. in [9]. The RP is a visual representation of the recurrence matrix given by

$$R_{ij} = \Theta\left(\varepsilon - ||\mathbf{x}_i - \mathbf{x}_j||\right), \qquad i, j = 1, 2, \ldots, N \ . \tag{1}$$

Here $\mathbf{x}_i \in \mathbb{R}^d$ is the state of the complex system in the phase space of dimension d at a time instance i, $||\cdot||$ denotes a metric, Θ is the Heaviside step function, and ε is a threshold distance.

Recurrence plots allow to partition a system's trajectory into three groups of activity: (i) stable states, meaning that the system evolves close to a single state within a long time interval, (ii) recurrent states, which are the states to which trajectory returns after some time and stays there for a finite amount of time, and (iii) transient states, which evolve fast, are unique for the system and do not repeat within the time of observation. In RPs, a black-colored dot denotes $R_{ij} = 1$ and white-colored space represents $R_{ij} = 0$. Hence, transient states will be shown in white and recurrent and stable domains will be shown in black.

Previous studies [6, 5, 7] have shown how to detect in a simple way each of the three types of states by means of symbolic dynamics. In detail, the symbols in the starting symbolic sequence are their corresponding time index $s_i = i$. If two states at times i and j are recurrent, i.e., $R_{ij} = 1$, and $i > j$, we can rewrite the symbol at time i with the symbol s_j. We can now

see that the recurrence matrix can be interpreted as a rewriting grammar and higher indices are replaced by the lower ones if the states at these time indices are recurrent. In addition, this method resolves cases when three states are recurrent, i.e., $R_{ij} = 1$ and $R_{ik} = 1$. The aforementioned rewriting rules can be formulated as follows:

$$\begin{array}{ll} i \to j & \text{if } i > j \text{ and } R_{ij} = 1 \\ \left.\begin{array}{l} i \to k \\ j \to k \end{array}\right\} & \text{if } i > j > k \text{ and } R_{ij} = 1,\ R_{ik} = 1 \ . \end{array} \tag{2}$$

The process of rewriting the initial symbolic partition accordingly should be performed twice in order to resolve both rules in Eq. (2). The subsequent step is to identify the transient states. Since we are not interested in each of these states in particular, we map sequentially increasing indices to the symbol $s_0 = 0$ to mark them as transitions from one quasi-stable state to another [6, 5].

The method described allows the visual representation of a system's dynamics and is rather simple with only one parameter to estimate. This parameter is the threshold distance ε. As we can see from Eq. (1), this parameter is crucial to the method. If ε is too small no recurrence can be formed and each state will be detected as a distinct state. Conversely, if the value of ε is too large, then all the states will be observed as one stable state.

Several estimates of the optimal threshold distance exist [30], however most of them are recommendations or rules of thumb without a grounding in the data. The present work utilizes the estimates of ε based on the assumption of the system's dynamics as it was done in [6]. The main idea of the method is to choose the optimal ε such that it maximizes a utility function which represents our assumption. This utility function reflects a given model of the temporal recurrence structure. In the present work, we assume that the system spends an equal amount of time in each of the recurrence domains and in the transient states. This hypothesis can be summarized by the principle of maximum entropy. In detail, to obtain the optimal threshold distance, first one varies values of ε and for each value one constructs a symbolic sequence. Then, the optimal threshold distance ε^* maximizes the utility function

$$\varepsilon^* = \arg\max_{\varepsilon} h(\varepsilon) \ , \tag{3}$$

where $h(\varepsilon)$ is the normalized entropy

$$h(\varepsilon) = -\frac{\sum_{k=1}^{M(\varepsilon)} p_k \log p_k}{M(\varepsilon)} \ . \tag{4}$$

Here p_k is the relative frequency of symbol $s_i = k$ and $M(\varepsilon)$ is the cardinality of the symbolic sequence. The denominator in Eq. (4) ensures that a too small or too large alphabet does not bias the estimate.

2.2 Phase Space Reconstruction

To analyze complex temporal properties of single observables, one may analyze time series of corresponding measurements. Such properties under study may be their fractal dimension [29, 16], their degree of nonlinearity [38, 13] or their recurrence structure [32, 47]. To this end, it has been shown that it is advantageous to reconstruct the system's supposed high-dimensional phase space to better retrieve underlying dynamical features. Then, the system's trajectory

in the new reconstructed phase space serves as the basis for the corresponding analysis. The most common reconstruction method is the delay embedding technique [41] (for other existing methods see, for example, [34] or [24]). In this paper, we propose a different approach, based on time-frequency representation of the univariate signal and compare it to the delay embedding technique.

2.2.1 Time-Frequency Embedding

A time-frequency representation is a distribution of the signal's power over time and frequency. It provides us with several time series, one for each frequency, which we then use to reconstruct phase space, where each frequency band contributes to the multi-dimensional phase trajectory. The main advantage of this approach over existing ones is that it takes into account the oscillatory nature of the signal and is more robust than analyzing RPs constructed from individual amplitude values. In the present work, we study recurrence symbolic sequences in two different time-frequency representations of the signal under study: the spectrogram and two variants of a continuous wavelet transform.

Spectrogram

The spectrogram $S(t, f)$ of a signal $x(t)$ in an infinite time interval is the squared magnitude of its short-time Fourier transform (STFT)

$$X(t, f) = \int_{-\infty}^{+\infty} x(\tau) h^*(t - \tau) \mathrm{e}^{-2\pi f i \tau} \, \mathrm{d}\tau \ , \tag{5}$$

where $h(t)$ is a smoothing window and $*$ denotes the complex conjugate, i.e., $S(t, f) = |X(t, f)|^2$. The spectrogram is computed in three steps: dividing the signal into several possibly overlapping segments, applying the smoothing window $h(t)$ to the segments and computing the power spectrum for each of them. This approach makes the spectrogram a very good analysis tool for non-stationary data.

In the present work the spectrograms were computed with a Gaussian window [15]

$$w(t) = \exp\left[-\frac{1}{2}\left(\alpha \frac{t}{(T-1)/2}\right)^2\right] \ , \tag{6}$$

where $-(T-1)/2 \leqslant t \leqslant (T-1)/2$, T is the length of the window, and α is a parameter controlling the width of the window. These parameters are inversely related: larger α lead to narrower windows. Here, $\alpha = 2.5$ and overlap between consecutive segments is 80%. The length of the segments was chosen such as to provide a resolution of 0.3 Hz for synthetic data and 1 Hz for EEG data. Afterwards the segments were zero-padded to the closest power of 2 for computational efficiency.

Moreover, it is important to mention edge effects at the borders of the time interval under study, cf. hatched areas in Fig. 3 (A) and 4 (A). These result from the fact that the time interval is finite. For the spectrogram, the time window length is independent from the frequency leading to an edge effect of constant width. In order to make the boundary effect less prominent, we padded the signals with a mirrored version of itself, where length of the padding is half of the window length [42].

Scalogram

By analogy with the spectrogram, the scalogram is the energy distribution of a signal calculated as the squared magnitude of the wavelet transform. The continuous wavelet transform (CWT) is obtained by convolving the signal with a set of functions $\psi_{a,b}(t)$ obtained by translation b and dilation a of a mother wavelet function $\psi_0(t)$ [1]

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi_0\left(\frac{t-b}{a}\right) \quad . \tag{7}$$

The mother wavelet function should be admissible [10], i.e., it should be square integrable, centered around zero with compact support and a zero mean. With the definition of the CWT for an infinite time domain

$$T_\psi(b,a) = \int_{-\infty}^{+\infty} x(t)\psi^*_{a,b}(t)\,\mathrm{d}t \quad , \tag{8}$$

the scalogram is given by

$$W_\psi(b,a) = |T_\psi(b,a)|^2 \quad . \tag{9}$$

Similarly to the spectrogram, the continuous wavelet transform is well suited for analyzing non-stationary data. However, unlike the spectrogram, where one faces a trade-off between the resolution in time and in frequency (called the Heisenberg relation), the wavelet transform has an adaptive resolution. In spectrograms, the resolution is determined by the length of the segments: short segments provide good time resolution but have wide spread in frequency domain and vice versa. In contrast, wavelet functions have larger temporal and smaller frequency spreading for lower-frequencies and the opposite for higher frequencies. Due to this property of the CWT the choice of the scales at which the transform is computed plays a very important role.

Scales in wavelet transform are inversely proportional to frequencies in a good approximation, i.e., $a \sim 1/f$ and the pseudo-frequencies f can be computed from scales by [43]:

$$f = \frac{1}{\lambda a} \quad . \tag{10}$$

The parameter λ is the equivalent to the primary Fourier period of the mother wavelet and hence is different for each wavelet type, i.e., each choice of mother wavelet function [33].

In this paper, we perform the analysis using the analytical Morlet wavelet [43]

$$\psi_0(t) = \pi^{-1/4}\mathrm{e}^{i\omega_0 t}\mathrm{e}^{-t^2/2} \quad , \tag{11}$$

where ω_0 is the center frequency of the wavelet. Here, $\omega_0 = 6$ in order to satisfy admissibility condition. The primary Fourier period in Eq. (10) for this mother wavelet is $\lambda = 1.03$ s.

Despite the fact that scales can be chosen arbitrarily several general practices exist. For a given range of frequencies, one has to choose the corresponding smallest and the largest scales a_0 and $a_{\max}$, respectively. The smallest scale is defined through the largest frequency which is limited by the Nyquist frequency, i.e., half of the sampling frequency. In the present work we do not choose this upper limit since the major activity of interest evolves at lower frequencies. For that reason we choose the maximum frequency to $f_{\max} = 10$ Hz for the transient oscillations, $f_{\max} = 5$ Hz for the Lorenz system and $f_{\max} = 40$ Hz for EEG data.

Similar to the spectrogram, we can observe edge effects occurring at the borders, cf. hatched areas in Fig. 3 (B,C) and 4 (B,C). In wavelet analysis, the sliding time window changes with

scale, i.e., with frequency, and the edge effect is larger for low frequencies than high frequencies due to the corresponding longer sliding time windows for low frequencies. Such nonlinear nature of boundary effect gives us an idea for estimating the largest scale, and thus the lowest frequency. Its value is chosen such that 50% of the CWT coefficients will be affected by the edge effect of the CWT [20] only. The next step is to sample the scale values within the above limits. In this work, we study the influence of the scales vector on the SRSA and compare two different scale sampling schemes.

In a first method variant, the set of scales was constructed similarly to the spectrogram. We sampled the frequency range under study with a fixed frequency step Δf and converted the frequencies of number N_a into scales by inverting Eq. (10):

$$a_{n+1} = a_n \frac{f_{\max} - n\Delta f}{f_{\max} + (n+1)\Delta f} \quad , \; n = 0, \ldots, N_a - 1 \tag{12}$$

with the maximum frequency $f_{\max}$. This method corresponds to the conventional wavelet analysis.

In another variant, to populate the scale vector, subsequent scales are proportional to previous scales [20]

$$a_{n+1} = K a_n \quad , \; n = 0, \ldots, N_a - 1 \; . \tag{13}$$

The factor K ensures that scale values cover sufficiently the frequency domain and is defined by

$$K = \exp\left[\frac{\ln a_{\max} - \ln a_0}{N_a - 1}\right] \quad , \tag{14}$$

Where $a_{\max}$ and a_0 are the largest and the smallest scales, respectively, and N_a is the total number of scales. In the following, N_a is chosen as the number of frequency bins used in the spectrogram for comparison reasons.

2.2.2 Delay Embedding

To compare the time-frequency embedding technique shown in the previous subsection to a conventional embedding technique, we employ the famous delay embedding technique first published by Takens [41]. Let us assume to have a time series of scalar measurements of a system's observable in discrete time

$$x_n = x(n\Delta t), \quad n = 1, \ldots, N \; , \tag{15}$$

where Δt is the measurement sampling time. Then the reconstructed phase space is given by

$$\mathbf{y}_n = \left[x_n, \, x_{n+\tau}, \, x_{n+2\tau}, \, \ldots, \, x_{n+(m-1)\tau}\right], \quad n = 1, \ldots, N - (m-1)\tau \; . \tag{16}$$

The parameters m and τ is the embedding dimension and the time delay, respectively. These parameters play an important role to reconstruct correctly phase space dynamics.

The optimal time delay τ should be chosen such that delay vectors from Eq. (16) are sufficiently independent. Several approaches to estimate this parameter exist. Among them, one well-established technique is based on the average mutual information [11, 27], which we have chosen to estimate the proper value of the time delay τ. Moreover, the main attribute of the appropriately chosen dimension m is that the original d-dimensional manifold will be embedded into m-dimensional space without ambiguity, i.e., without self-crossing and intersections. We apply the method of false nearest neighbors [22, 21] that permits to estimate the minimal embedding dimension.

2.3 Complexity Measure

To quantitatively compare recurrence structures gained from the recurrence analysis, we suggest to compare the complexity of their corresponding resulting symbolic sequences. Besides several complexity measures for symbolic sequences, such as the Lempel-Ziv complexity [26] or the Kolmogorov complexity [23], another good candidate for this measure is the number of unique words extracted from a sequence [17]. We define a word as a consecutive sequence of the same symbol. As an example let us look at the sequence $s = 0011110002220033000$. The extracted words are $w = \{00 \cdot 1111 \cdot 000 \cdot 222 \cdot 00 \cdot 33 \cdot 000\}$. Here, the total number of words is 7 and the number of unique words is 5. Hence the complexity value is $c = 5$.

2.4 Data

2.4.1 Synthetic Data

Synthetic data under investigation includes transient oscillations and the solution of the chaotic Lorenz system. Examples of such signals are shown on the Fig. 1.

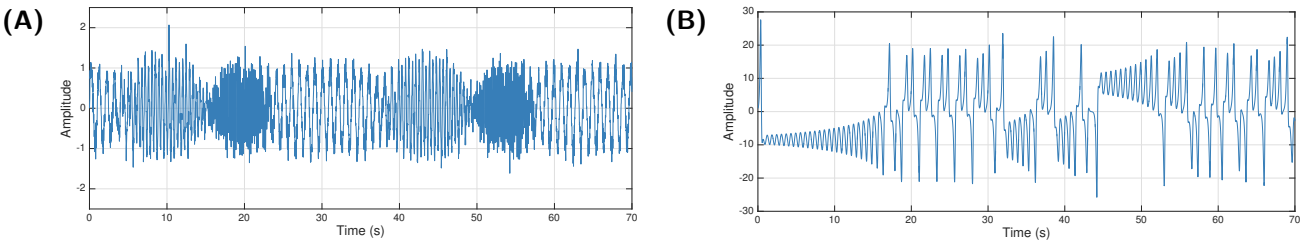

Figure 1: Example signals of the synthetic data. (A) Transient oscillations with frequencies $f_1 = 1$ Hz, $f_2 = 2.25$ Hz and $f_3 = 6.3$ Hz and the sequence $f_1 \cdot f_2 \cdot f_3 \cdot f_1 \cdot f_2 \cdot f_3 \cdot f_1$; (B) Solution of the Lorenz system along a single dimension.

Transient Oscillations. In detail, a modified Lotka-Volterra model with $N = 3$ interactive elements [37, 36]

$$\dot{x}_i(t) = x_i(t)\left(\sigma_i - \sum_{j=1}^{N} \rho_{ij} x_j(t)\right), \quad i = 1, 2, 3 \tag{17}$$

serves as an abstract model of a heteroclinic sequence. It has been shown to serve as a good model for the temporal structure of event-related brain potentials [5] and bird songs [46]. Here $x_i(t) \geqslant 0$ is the activity rate of the element i, σ_i is the growth rate of the i-th element and ρ_{ij} is an entry of the $N \times N$ interaction matrix. In our setup, $\sigma_1 = 1$, $\sigma_2 = 1.2$ and $\sigma_3 = 1.6$, $\rho_{ii} = 1$, $\rho_{12} = 1.33$, $\rho_{13} = 1.125$, $\rho_{21} = 0.7$, $\rho_{23} = 1.25$, $\rho_{31} = 2.1$, and $\rho_{32} = 0.83$. The output signal $s(t)$ to be investigated is a linear superposition of transient oscillations with frequencies $f_1 = 1$ Hz, $f_2 = 2.25$ Hz, $f_3 = 6.3$ Hz, where at one time instance, only one of these three components is dominant, see more details below. We point out that these frequencies are chosen rather arbitrarily for an optimal illustration.

The activity rate x_i defines the amplitude of the signal component y_i with frequency f_i and the output signal is given by

$$u(t) = \sum_{i=1}^{N} y_i(t) \sin 2\pi f_i t + \xi(t) \ , \qquad y_i(t) = \exp\left[-\frac{(x_i - \sigma_i)^2}{2\eta_i^2}\right] \ , \tag{18}$$

with $\eta_1 = 0.5, \eta_2 = 0.33, \eta_3 = 0.4$ and $\xi(t) \sim \mathcal{N}(0, 0.01)$ is additive white Gaussian noise. By this construction, the signal components y_i oscillates in a certain time window outside of which they almost vanish. These windows of the three oscillation modes do not overlap and the transitions between them are rapid. The sampling rate of the dataset is 50 Hz.

Lorenz Dataset. The Lorenz system [28] is a well-studied three-dimensional differential equation system

$$\dot{x} = -\sigma x + \sigma y \ , \quad \dot{y} = \rho x - y - xz \ , \quad \dot{z} = -\beta z + xy \tag{19}$$

with $\sigma = 10$, $\rho = 28$, $\beta = 8/3$. The system describes, inter alia, weather phenomena [28] and some dynamics in lasers [14] and electric circuits [8]. Its solutions show non-trivial transient dynamics and their wings represent metastable states, cf. [6]. The projection of the system solution onto a single coordinate represents a time series that may serve as a model of a macroscopic measured signal, such as an chaotic EEG recording [39, 4, 12]. In this paper, we study the univariate time series $u(t) = y(t)$ with sampling rate equal to 50 Hz.

2.4.2 Experimental Data

General anesthesia is an everyday practice in hospital. To improve the monitoring of brain activity and hence the supervision of patients during surgery, more and more hospitals register EEG on the scalp of the patient. Typically, a single EEG electrode is fixed on the forehead of the patient. The data investigated in the present work has been registered during surgery in a large set of patients with an effective sampling rate of 128 Hz. The dataset [40] comprises time series of 115 patients who have been under surgery applying the anesthetics *desflurane* (number of patients: 55) and *propofol* (number of patients: 60). The EEG data investigated represent 30-second signals recorded during surgical operations before and after skin incision in the beginning of surgery. Patients are divided into two groups based on the anesthetic drug they received: desflurane or propofol. For illustration, Fig. 2 shows examples of signals in each of the groups taken from pre- and post-incision periods. We investigate the recurrence structure of such EEG time series and whether it is possible to distinguish the anesthetic agents administered in the EEG data based on the recurrence structure.

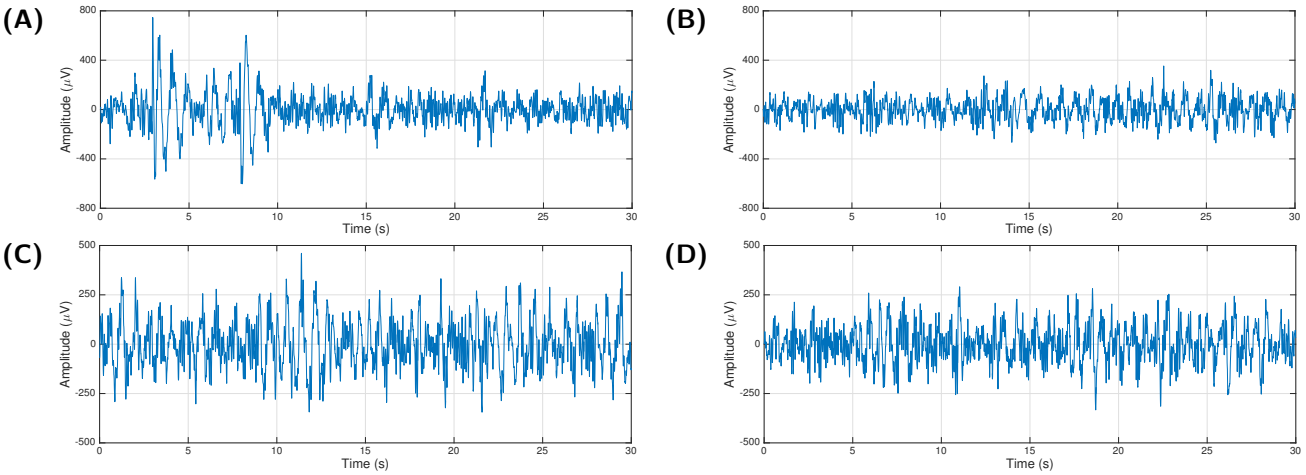

Figure 2: Example signals of the EEG data. (A) Desflurane pre-incision; (B) Desflurane post-incision; (C) Propofol pre-incision; (D) Propofol post-incision.

3 Recurrence Analysis

In this section we evaluate the method proposed by applying it to the synthetic data and experimental EEG data. In the analysis of synthetic and experimental data, we compute recurrence symbolic dynamics employing time-frequency embedding and delay embedding.

3.1 Synthetic Data

3.1.1 Time-Frequency Embedding

Figure 3 shows the recurrence analysis results (right hand side) obtained from transient oscillations with different time-frequency representations (left hand side) for the signal with initial conditions $\mathbf{x}_0 = [1, 0.17, 0.01]$.

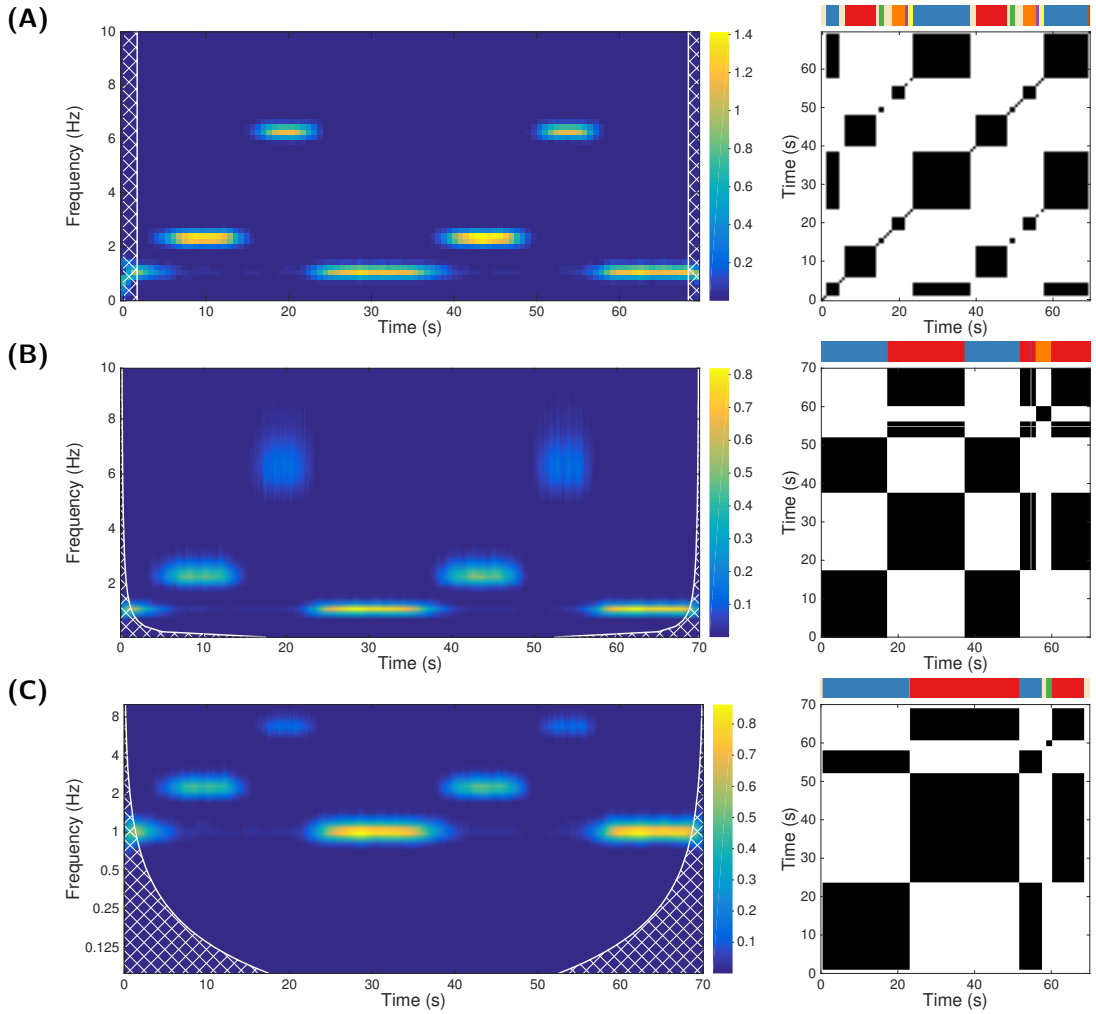

Figure 3: Results for transient oscillations signal. (A) spectrogram; (B) scalogram (equidistant frequencies); (C) scalogram (non-equidistant frequencies). Left: time-frequency representation. Right: RPs with corresponding symbolic sequences above them. The colors in the symbolic sequences reflect a metastable state and transitions between the states, i.e., $s_i = 0$, are shown in beige color. In the time-frequency representations hatched regions mark the cone of influence reflecting edge effects.

We observe that the recurrence plot and the corresponding symbolic sequence obtained from the spectrogram are in good accordance with the underlying dynamics of the signal. Namely, we see three oscillatory components shown in blue, red and orange in the correct time windows. In addition, it appears that the end of one component and the beginning of the next one creates additional states of brief duration, shown in green, purple and yellow on Fig. 3 (A).

The recurrence analysis based on the time-frequency representations obtained by the wavelet transform techniques perform worse and tend to neglect higher-frequency oscillatory components. They detect two main states at lower frequencies only. Equally spaced wavelet transform (cf. Fig. 3 (B)) creates two symbols representing 1 Hz and 2.25 Hz oscillations while neglecting 6.3 Hz component. The wavelet transform with non-equidistant frequencies separates two low-frequency oscillations from the 6.3 Hz oscillations. The primary reason of the neglect of components at larger frequencies is the well-known $1/f$ scaling of the power intrinsic in the wavelet spectrum, i.e., wavelet power decreases with higher frequencies due to the spread of the wavelet's frequency support while conserving total power. In phase space, this decrease of power moves the high-frequency component closer to zero power and the corresponding data points are hard to distinguish from the background activity. Moreover, in wavelet analysis high-frequency components are resolved worse in frequency (due to the Heisenberg principle) and hence the power peak is spread over several frequencies. This prevents phase space reconstruction from properly capturing the dynamics, increases distances between data points and hence diminishes the performance of the recurrence analysis.

The results for the Lorenz system with initial conditions $[x_0, y_0, z_0] = [0, 1, 0]$ are shown in Fig 4. As it can be seen, the spectrogram exhibits good correspondence with the underlying dynamics. We can clearly see two states ("wings") interchanging each other, whereas results from wavelet are much worse. One of the reasons for that is the aforementioned $1/f$ scaling of wavelets' power and the increased frequency spread at higher frequencies. Another reason can be found by examining the low-frequency parts of the corresponding time-frequency representations. It is known that the wavelet transform can not estimate power around $f = 0$ Hz, since this would correspond to an infinite scale due to $f \sim 1/a$. Taking a close look at the Lorenz time series in Fig. 1 (B), one observes that the "wings" of Lorenz system are not only defined by the frequency of the oscillations but also by the bias of these oscillations, despite the fact that the whole signal is zero-mean. Since wavelet techniques can not resolve this constant shift, they miss an important feature and hence can not properly handle the signal's dynamics.

3.1.2 Delay Embedding

To compare the suggested method to the conventional delay embedding technique, we compute the optimal embedding parameters m and τ for both synthetic datasets, embed the signals in the corresponding high-dimensional space and apply the SRSA to obtain the recurrence plot and corresponding symbolic sequences.

Figure 5 shows that the recurrence analysis does capture much worse the underlying recurrence structure using delay embeddings. One can poorly identify main recurrence domains for the transient oscillations (Fig. 5 (A)). For the Lorenz signal (Fig. 5 (B)) we observe two interchanging quasi-stable states, however the detection is worse than with the spectrogram-based method (Fig. 4 (A)). We notice that the method based on the delay embedding tends to extract stable states based on the amplitude values of the signal. This is especially noticeable in case of the transient oscillations, where symbolic sequence consist of almost 30 unique symbols.

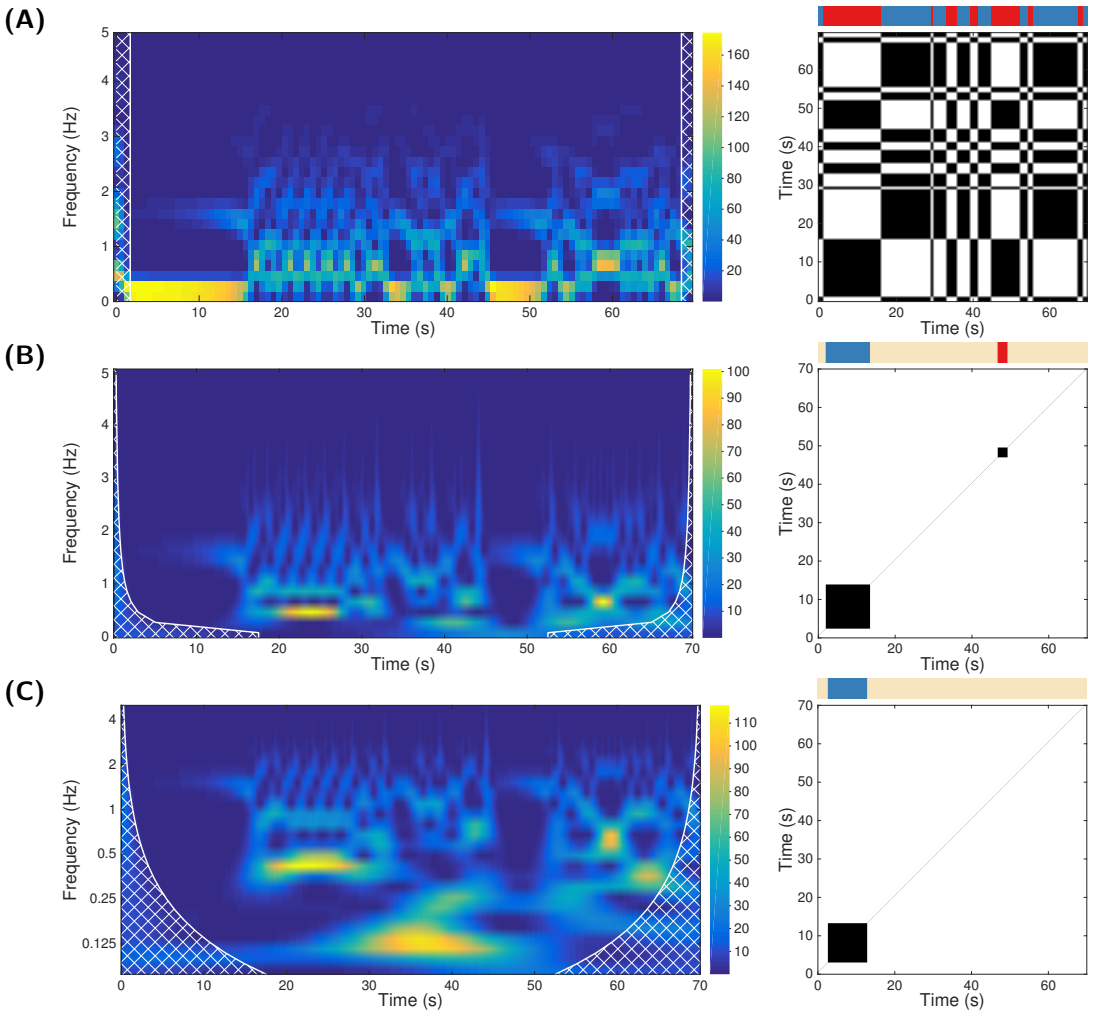

Figure 4: Results for Lorenz system. (A) spectrogram; (B) scalogram (equidistant frequencies); (C) scalogram (non-equidistant frequencies). Left: time-frequency representation. Right: RP with corresponding symbolic sequence above them. In the time-frequency representations hatched regions mark the cone of influence reflecting edge effects. Color-coding in symbolic sequences as in Fig. 3.

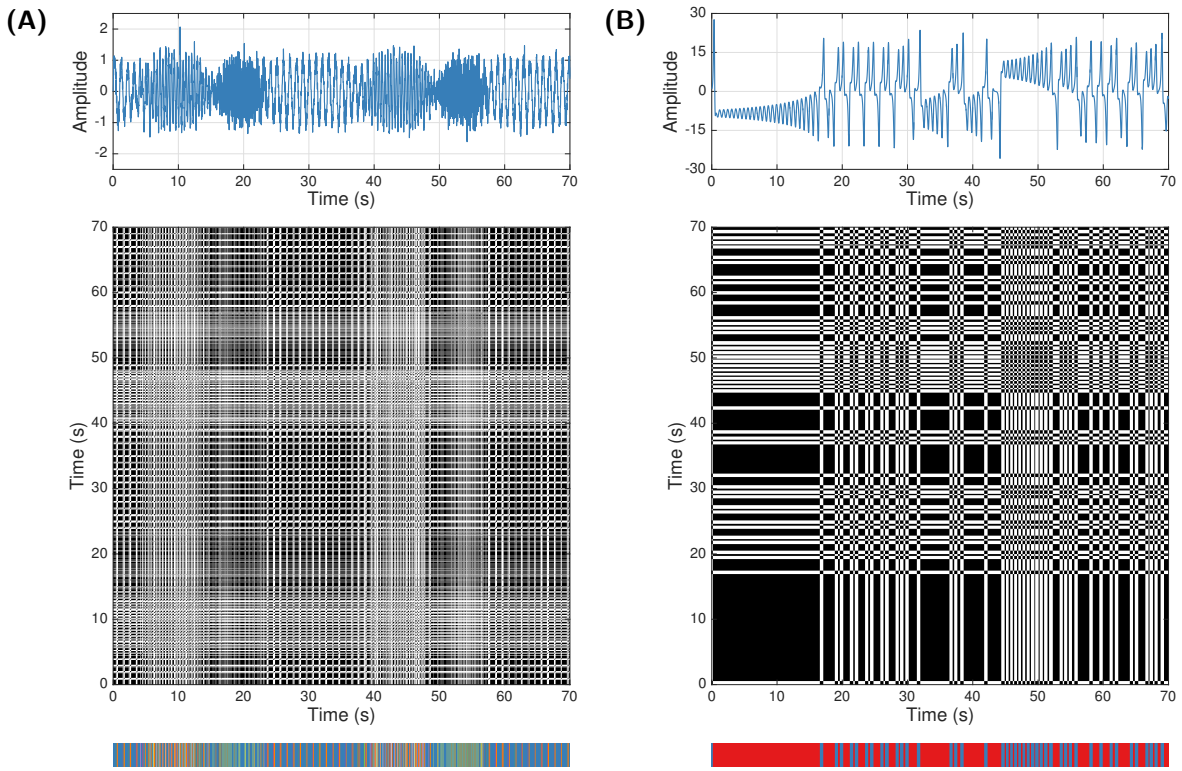

Figure 5: Results obtained with delay embeddings. (A) Transient oscillations, reconstruction parameters: $m = 5$, $\tau = 0.1$ s; (B) Solutions of the Lorenz system, reconstruction parameters: $m = 3$, $\tau = 0.18$ s.

3.2 Experimental data

There is a strong evidence that the brain passes a sequence of metastable states under certain conditions, e.g., in bird songs [46] or during the emergence from anesthesia [18]. Since the previous results from synthetic data clearly reveal that the spectrogram and the conventional scalogram (with equidistant frequencies) yield recurrence structures that resemble the underlying dynamics in contrast to the wavelet analysis with non-equidistant frequencies, we pick these two former methods for the following analysis.

Figures 6 and 7 shows recurrence analysis results for time-frequency representations gained from spectrogram and conventional wavelet analysis, respectively. Firstly, one observes that the signals evolve with a maximum range in the $\alpha-$range, i.e., in the frequency interval [8 Hz, 12 Hz] with a strong activity in this band. In addition, pronounced activity can be observed in the $\delta-$band with frequencies in the interval [1 Hz, 4 Hz]. This enhanced power in the $\delta-$ and $\alpha-$bands is in accordance with the previous findings in [40] and is a typical spectral feature in frontal EEG electrodes during deep anesthesia. A large number of previous studies have extracted stationary neural activity, but very few have investigated fluctuations of spectral power [25, 45].

Moreover, Fig. 6 shows well-located spectral power in the $\alpha-$ and $\delta-$band whereas spectral peaks are smeared out in Fig. 7, especially at larger frequencies. As mentioned above, this smearing is an intrinsic property of conventional wavelet analysis. Hence, one expects that the recurrence analysis based on the wavelet time frequency representations captures primarily the temporal structure at lower frequencies, i.e., in the $\delta-$band. Consequently, this leads to a coarser recurrence structure than for the spectrogram and hence sequences with less symbols. In fact, this is seen in Figs. 6 and 7.

Comparing recurrence structures between desflurane and propofol or between pre- and post-incision, it is difficult to identify fundamental differences by visual inspection. It seems that the recurrence structure computed from scalogram exhibits longer recurrent periods in propofol data than in desflurane data. However, such a visual impression has to be validated statistically. This is attempted in the subsequent section.

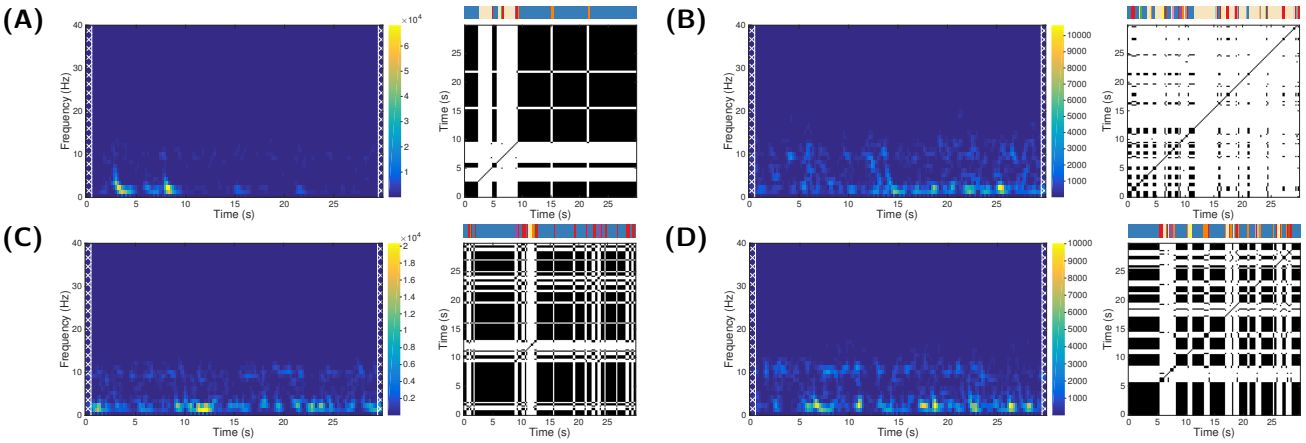

Figure 6: Spectrograms and corresponding recurrence dynamics. (A) Desflurane pre-incision; (B) Desflurane post-incision; (C) Propofol pre-incision; (D) Propofol post-incision. In the time-frequency representations hatched regions mark the cone of influence reflecting edge effects. The frequency range is [0 Hz; 40 Hz].

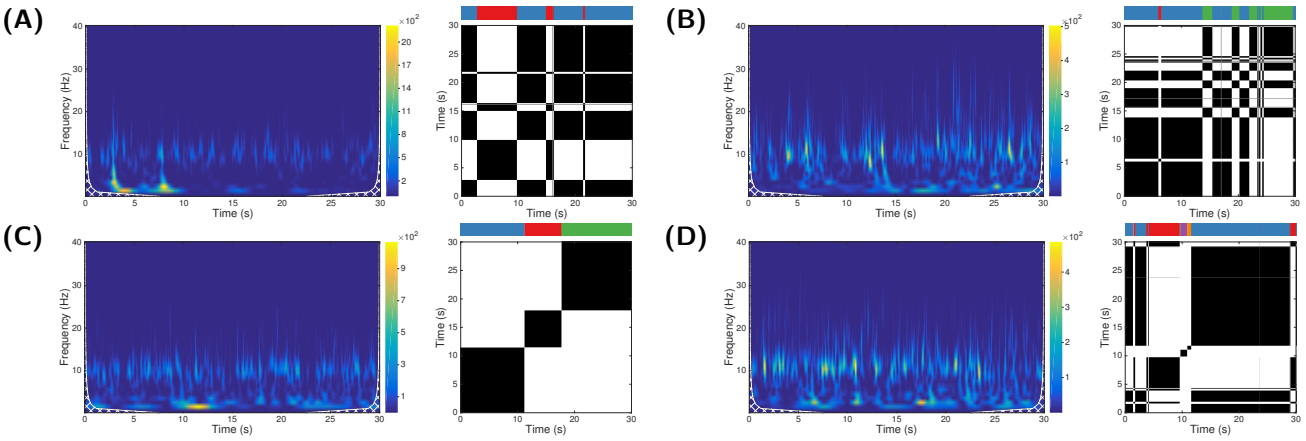

Figure 7: Scalograms (equidistant frequency) and corresponding recurrence dynamics. (A) Desflurane pre-incision; (B) Desflurane post-incision; (C) Propofol pre-incision; (D) Propofol post-incision. In the time-frequency representations hatched regions mark the cone of influence reflecting edge effects. The frequency range is [0.18 Hz; 40.18 Hz].

4 Recurrence complexity

The symbolic sequence extracted for each data set reflects the recurrence structure in the signal and may serve as a unique statistical feature. Consequently, the statistical comparison of symbolic sequences of different signals may represent a methodology to compare intrinsic recurrence structures and hence compare the corresponding system's underlying mechanism. As quantitative feature of a symbolic sequence, we consider its complexity (cf. Section 2.3).

4.1 Synthetic data

To demonstrate the method we generated a set of 100 trials of each type of synthetic signals with random initial conditions and computed distributions of the complexity measure. The initial condition are drawn from normal distribution, i.e., $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}, \sigma^2)$, the parameters of the distribution are: $\boldsymbol{\mu} = [1, 0, 0]$ and $\sigma = 0.2$ for transient oscillations and $\boldsymbol{\mu} = [20, 5, -5]$ and $\sigma = 15$ for the Lorenz system. Figure 8 shows the distributions of complexity measures for the different time-frequency representations. The two types of scalograms result in very similar distributions for both transient oscillations and the Lorenz system, whereas the complexity measure distributions of both signals are distinct applying the spectrogram. Kolmogorov-Smirnov tests show that all the distributions are significantly different with $p < 0.001$.

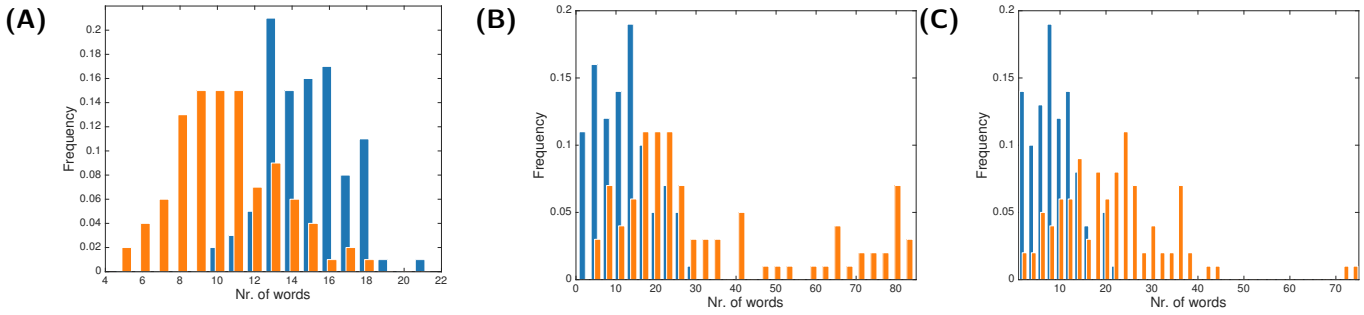

Figure 8: Complexity distributions for the transient oscillations (orange) and the Lorenz system (blue) obtained with different time-frequency representations. (A) spectrogram; (B) conventional scalogram (equidistant frequencies); (C) scalogram with non-equidistant frequencies. For all time frequency distributions, both distributions are significantly different (Kolmogorov-Smirnov test with $p < 0.001$).

4.2 Experimental data

Finally, we study the large experimental EEG datasets. As described in Sect. 2.4.2, EEG dataset consist of four classes: desflurane, propofol, pre-incision and post-incision. We focus on distinguishing between desflurane and propofol, separately in pre-incision and post-incision signals. Similarly to synthetic dataset we computed the complexity distributions for all the time-frequency representations, corresponding p-values of Kolmogorov-Smirnov test are summarized in Table 1.

Table 1: p-values for Kolmogorov-Smirnov test for complexity distributions obtained with different embedding types. Significant differences between propofol and desflurane are marked by bold values with $p < 0.05$.

Embedding	Pre-incision	Post-incision
Delay embedding	0.2207	0.0744
Spectrogram	0.3707	0.6839
Conventional scalogram	**0.0148**	**0.0417**
Scalogram with non-equidistant frequencies	**0.0182**	**0.0120**

We observe that symbolic sequences obtained with delay embeddings or spectrogram embedding do not allow to distinguish desflurane and propofol anesthesia. Conversely both wavelet techniques distinguish time series of both anesthetics. For illustration reasons, Fig. 9 presents

the corresponding complexity distributions for the conventional scalogram. We observe that signals from patients under propofol anesthesia exhibit a broad distribution of complexities before the incision (A) and a rather compact distribution at lower complexities in signals after incision (B). Such a change of distribution is less visible in signals under desflurane anesthesia.

The fact that recurrence structures in these data based on the wavelet analysis are significantly different compared to spectrogram and the delay embedding data indicates that activity at low frequencies in the $\delta-$band exhibit unique recurrence structures. These unique recurrence patterns exist in spectrogram data as well, however the additional $\alpha-$activity may disturb the detection of the recurrence structure in the $\delta-$band and hence the recurrence structures involving both bands are blurred.

An additional statistical inference test on differences between pre- and post-incision data separated for each anesthetic does not show any significant differences between their recurrence structures (not presented here for brevity).

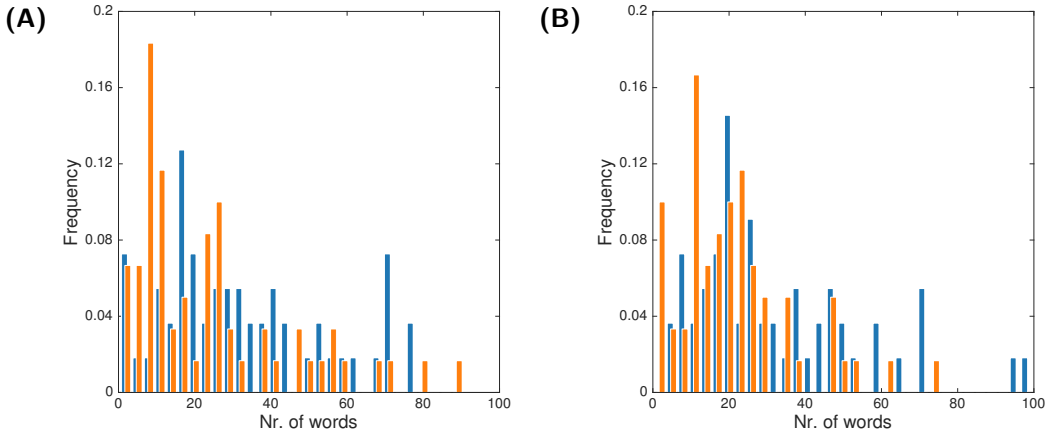

Figure 9: Desflurane/propofol distributions of complexity measures obtained with scalograms (linear frequency). (A) pre-incision signals; (B) post-incision signals. Desflurane is shown in blue and propofol in orange.

5 Conclusions

The present work shows a detailed investigation how to compute recurrence structures in univariate time series taking into account their oscillatory activity. The major idea is to compute the time series time-frequency representation, considering the instantaneous power at certain frequencies as a multivariate time series and feeding this high-dimensional signal into a recurrence analysis. For synthetic data, we observe that the spectrogram provides a good time-frequency representation and allows to extract the underlying recurrence structure of a signal. Wavelet analysis yields worse results. A comparison to the conventional delay embedding technique demonstrates clearly that the suggested method outperforms results gained from delay embedding.

In order to classify observed time series based on their recurrence structure, we suggest to compute their recurrence complexity. A statistical inference test on synthetic data demonstrates that spectral embedding techniques may allow to distinguish the time series by their different recurrence complexity. This classification method also allows to distinguish experimental EEG data obtained during general anesthesia with two different anesthesia agents. In contrast to

the synthetic data, in this classification scalogram works well whereas the spectrogram and the conventional delay embedding technique fail. We provide possible explanations for this finding based on the power distribution over frequencies in the data.

Future work will focus on the importance of certain frequency bands in the time-frequency representations in the context of complexity measures. Moreover, novel criteria for the choice of the threshold ε may yield different results and may elucidate how to introduce a model for the recurrence structure into recurrence analysis.

References

[1] P. S. Addison. *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine, and finance*. Institute of Physics Pub., 2002.

[2] C. Allefeld, H. Atmanspacher, and J. Wackermann. Mental states as macrostates emerging from EEG dynamics. *Chaos*, 19:015102, 2009.

[3] E. Balkovsky, G. Falkovich, and A. Fouxon. Intermittent distribution of inertial particles in turbulent flows. *Phys. Rev. Lett.*, 86:2790, 2001.

[4] E. Basar. *Brain Function and Oscillations*, chapter Chaos in Brain Function, pages 169–198. Springer Heidelberg, 2006.

[5] P. beim Graben and A. Hutt. Detecting event-related recurrences by symbolic analysis: applications to human language processing. *Phil. Trans. A Math. Phys. Eng. Sci.*, 373(2034):20140089.

[6] P. beim Graben and A. Hutt. Detecting recurrence domains of dynamical systems by symbolic dynamics. *Phys. Rev. Lett.*, 110(15).

[7] P. beim Graben, K. K. Sellers, F. Fröhlich, and A. Hutt. Optimal Estimation of Recurrence Structures from Time Series. *submitted*, 2016.

[8] K. M. Cuomo and A. V. Oppenheim. Circuit implementation of synchronized chaos with applications to communications. *Phys. Rev. Lett.*, 71(1):65–68, 1993.

[9] J.-P. Eckmann, S. Oliffson Kamphorst, and D. Ruelle. Recurrence Plots of Dynamical Systems. *Europhysics Letters (EPL)*, 4(9):973–977, 1987.

[10] M. Farge. Wavelet Transforms and their Applications to Turbulence. *Annu. Rev. Fluid Mech.*, 24(1):395–458, 1992.

[11] A. M. Fraser and H. L. Swinney. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A*, 33(2):1134–1140, 1986.

[12] R. Friedrich and C. Uhl. Spatio-temporal analysis of human electroencephalograms: Petit-mal epilepsy. *Physica D*, 98:171–182, 1996.

[13] A. Gunduz and J. C. Principe. Correntropy as a novel measure for nonlinearity tests. *Signal Processing*, 89(1):14–23, 2009.

[14] H. Haken. Analogy between higher instabilities in fluids and lasers. *Phys. Lett. A*, 53(1):77–78, 1975.

[15] F. J. Harris. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proceedings of the IEEE*, 66(1):51–83, 1978.

[16] H. G. E. Hentschel and I. Procaccia. The infinite number of generalized dimensions of fractals and strange attractors. *Physica D*, 8(3):435–444, 1983.

[17] J. Hu, J. Gao, and J. Principe. Analysis of Biomedical Signals by the Lempel-Ziv Complexity: the Effect of Finite Data Size. *IEEE Transactions on Biomedical Engineering*, 53(12):2606–2609, 2006.

[18] A. E. Hudson, D. P. Calderon, D. W. Pfaff, and A. Proekt. Recovery of consciousness is mediated by a network of discrete metastable activity states. *Proc. Natl. Acad. Sci. USA*, 111:9283–9288, 2014.

[19] A. Hutt and H. Riedel. Analysis and modeling of quasi-stationary multivariate time series and their application to middle latency auditory evoked potentials. *Physica D*, 177:203–232, 2003.

[20] D. Jordan, R. W. Miksad, and E. J. Powers. Implementation of the continuous wavelet transform for digital time series analysis. *Review of Scientific Instruments*, 68(3):1484, 1997.

[21] M. B. Kennel and H. D. I. Abarbanel. False neighbors and false strands: A reliable minimum embedding dimension algorithm. *Phys. Rev. E*, 66(2), 2002.

[22] M. B. Kennel, R. Brown, and H. D. I. Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A*, 45(6):3403–3411, 1992.

[23] A. N. Kolmogorov. Three approaches to the quantitative definition of information. *Problemy Peredachti Informatsii*, 1(1):3–11, 1965.

[24] D. Kugiumtzis and N. D. Christophersen. State space reconstruction: method of delays vs singular spectrum approach. *Res. Rep. Httpurn Nb NoURN NBN No-35645*, 1997.

[25] J. Lefebvre, A. Hutt, J. F. Knebel, K. Whittingstall, and M. Murray. Stimulus statistics shape oscillations in nonlinear recurrent neural networks. *J. Neurosci.*, 35(7):2895–2903, 2015.

[26] A. Lempel and J. Ziv. On the complexity of finite sequences. *IEEE Transactions on Information Theory*, 222(1):75–81, 1976.

[27] W. Liebert and H. G. Schuster. Proper choice of the time delay for the analysis of chaotic time series. *Phys. Lett. A*, 142(2):107–111, 1989.

[28] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmosph. Sci.*, 20(2):130–141, 2016.

[29] B. B. Mandelbrot. Self-Affine Fractals and Fractal Dimension. *Physica Scripta*, 32(4):257, 1985.

[30] N. Marwan, M. Carmenromano, M. Thiel, and J. Kurths. Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5-6):237–329, 2007.

[31] N. Marwan and J. Kurths. Line structures in recurrence plots. *Phys. Lett. A*, 336:349–357, 2005.

[32] N. Marwan and C. L. Webber, Jr. Mathematical and Computational Foundations of Recurrence Quantifications. In C. L. Webber, Jr. and N. Marwan, editors, *Recurrence Quantification Analysis*, Understanding Complex Systems, pages 3–43. Springer International Publishing, 2015.

[33] S. D. Meyers, B. G. Kelly, and J. J. O'Brien. An Introduction to Wavelet Analysis in Oceanography and Meteorology: With Application to the Dispersion of Yanai Waves. *Monthly Weather Review*, 121(10):2858–2866, 1993.

[34] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. Geometry from a time series. *Phys. Rev. Lett.*, 45(9):712, 1980.

[35] H. Poincaré. Sur le problème des trois corps et les équations de la dynamique. *Acta mathematica*, 13(1):3–270, 1890.

[36] M. I. Rabinovich, R. Huerta, and G. Laurent. Transient dynamics for neural processing. *Science*, 321(5885):48 – 50, 2008.

[37] M. I. Rabinovich, R. Huerta, P. Varona, and V. Afraimovich. Transient cognitive dynamics, metastability and decision making. *PLOS Comput. Biol.*, 4(5), 2008.

[38] M. Sano and Y. Sawada. Measurement of the Lyapunov Spectrum from a Chaotic Time Series. *Phys. Rev. Lett.*, 55(10):1082–1085, 1985.

[39] C. A. Skarda and W. J. Freeman. How brains make chaos in order to make sense of the world. *Behav. Brain Sci.*, 10(2):161, 1987.

[40] J. W. Sleigh, K. Leslie, and L. Voss. The effect of skin incision on the electroencephalogram during general anesthesia maintained with propofol or desflurane. *J. Clin. Mon. Comput.*, 24:307–318, 2010.

[41] F. Takens. *Detecting strange attractors in turbulence*. Springer, 1981.

[42] G. Thakur, E. Brevdo, N. S. Fučkar, and H.-T. Wu. The Synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications. *Signal Processing*, 93(5):1079–1094, 2013.

[43] C. Torrence and G. P. Compo. A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79:61–78, 1998.

[44] T. Tošić, K. Sellers, F. Fröhlich, M. Fedotenkova, P. beim Graben, and A. Hutt. Statistical frequency-dependent analysis of trial-to-trial variability in single time series by recurrence plots. *Fronti. Syst. Neurosci.*, 9(184), 2016.

[45] M. A. Whittington, R. D. Traub, and J. G. Jefferys. Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. *Nature*, 373:612–615, 1995.

[46] I. B. Yildiz and S. J. Kiebel. A hierarchical neuronal model for generation and online recognition of birdsongs. *PLoS Comput. Biol.*, 7:e1002303, 2011.

[47] J. P. Zbilut, N. Thomasson, and C. L. Webber. Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals. *Medical Engineering & Physics*, 24(1):53–60, 2002.