

## Highlights

### **Complexity from ordinal pattern positioned slopes (COPPS)**

Jean Sire Armand Eyebe Fouda, Wolfram Koepf, Norbert Marwan, Jürgen Kurths, Thomas Penzel

- New data encoding combining slopes and positions to enhance complexity evaluation
- New COPPS method measuring complexity from short measurement data sets
- Distinction between chaotic and stochastic data and robustness against noise
- Self-similarities and model assumptions are not considered
- Easy detection of phase transitions and application to ECG beat classification

# Complexity from ordinal pattern positioned slopes (COPPS)

Jean Sire Armand Eyebe Fouda<sup>a,\*</sup>, Wolfram Koepf<sup>b</sup>, Norbert Marwan<sup>c</sup>,  
Jürgen Kurths<sup>c</sup>, Thomas Penzel<sup>d</sup>

<sup>a</sup>*Department of Physics, Faculty of Science, University of Yaoundé I, P.O. Box  
812, Yaoundé, Cameroon*

<sup>b</sup>*Institute of Mathematics, University of Kassel, Kassel, Germany*

<sup>c</sup>*Potsdam Institute for Climate Impact Research, Member of the Leibniz  
Association, P.O. Box 601203, Potsdam, D-14412, Germany*

<sup>d</sup>*Department of Internal Medicine, Center of Sleep Medicine, Charité University of  
Medicine, Berlin, D-10117, Germany*

---

## Abstract

Measuring complexity allows to characterize complex systems. Existing techniques are limited to simultaneously measure complexity from short length data sets, detect transitions and periodic dynamics. This paper presents an approach based on ordinal pattern positioned slopes (OPPS). It considers exclusively OPPS group occurrences to compute the complexity from OPPS (COPPS) as the average number of patterns and applies to short data series. The COPPS measure was successfully applied to simulation data for measuring complexity, detecting transition phases and regular dynamics, distinguishing between chaotic and stochastic dynamics; and to real-world data for detecting arrhythmia ECG beats.

*Keywords:*

Ordinal patterns, Complexity, Time series

---

## 1. Introduction

Measuring the complexity of dynamical systems is as important as it is helpful for understanding their internal behaviour, hence it is useful for predicting important phenomena such as phase transitions, synchronization, tipping points or extreme events [1, 2]. Mostly used complexity measures are entropy, correlation dimension and Lyapunov exponents (LE). The basic conceptual problem with these measures is that they apply to presumably

---

\*Corresponding author

*Email address:* fouda@mathematik.uni-kassel.de, efoudajsa@gmail.com (Jean Sire Armand Eyebe Fouda)

*URL:* www.mathematik.uni-kassel.de/~fouda (Jean Sire Armand Eyebe Fouda)

ergodic dynamical systems and require infinite data series, or modelling equations [3, 4, 1, 2, 5]. Empirical measures like the renormalized entropy or the approximate entropy often lack a theoretical foundation or are not well interpretable [6, 5, 7]. There are other time series-based complexity measures including sample entropy, Renyi’s entropy, the effective measure complexity, the algorithmic complexity, the  $\varepsilon$ -complexity, or recurrence quantification measures that were introduced during the last four decades [6, 8, 9, 10, 11, 12]. Despite their efficiency, these methods rely on the definition of a threshold allowing to determine self-similarities in the underlying time series. Although several approaches have been proposed for some cases, the choice of the threshold depends on the specific question and does not necessarily assess the natural complexity of the dynamical system [13]. As a consequence, they work well for simulated time series from low-dimensional dynamical systems, but are limited when analysing noisy data, thereby real-world data [2].

New methods commonly referred to as symbolic encoding were proposed to approximate theoretical methods. Symbolic encoding emphasizes typical properties, although it may result in a loss of information on the detailed state of the system under investigation [14, 15, 16]. In the case of lossy techniques like the ordinal encoding (OE) [17],  $n!$  possible permutations resulting from rank order sequences of length  $n$  (embedding dimension) are enumerated. These sequences result in a symbolic encoding that allows to define statistical methods like permutation entropies and order recurrence plots [3, 4, 18]. The permutation entropy (PE) introduced by Bandt and Pompe does not require the definition of a threshold for determining self-similarities in the time series. The symbol sequence naturally comes from the time series, without further model assumptions. PE efficiently estimates complexity of time series and behaves similarly to the Lyapunov exponent over the chaotic regime, whereas it fails to detect periodic time series with a zero entropy. PE also has a practically important invariance property that implies its high sensitivity to noise in the case of constant and periodic time series. This sensitivity limits its application to noisy data, as it yields the highest entropy value for periodic dynamics, similar to purely random dynamics [4].

Recently, we proposed to consider the first order difference-filtering of the ordinal patterns [19]. The method consists of considering the largest slope of ordinal patterns and is referred to as ordinal pattern slope (OPS)-encoding. The resulting permutation largest slope entropy (PLSE) characterizes regular and nonregular dynamics respectively with zero and positive entropy values. It was successfully applied to the detection of strange nonchaotic attractors as regular dynamics [20], and the analysis of real-world data by clearly identifying regular dynamics and transitions in a noisy environment [21]. The problem with the PLSE is that it does not allow to compare different nonregular dynamics. This limitation of the OPS encoding results from

the further information loss as only largest slopes are considered.

In this paper, we suggest to combine the largest slope of the ordinal pattern with its position in the embedding vector, in order to compensate the information loss mentioned above. Such an encoding that we refer to as the *ordinal pattern positioned slope* (OPPS)-encoding also does not make any further model assumption. The complexity of a time series is naturally measured as the ratio of symbols it can generate over the total number of possible symbols. The method is applied to simulation data derived from the sine-circle and logistic maps in order to prove its efficiency to distinguish between different nonregular dynamics, detect transition phases and regular dynamics; and to electrocardiogram (ECG) beats as short length real-world data to detect arrhythmia, which could help early alerting for cardiac disorders. We also apply it to various colored noises in order to further confirm its ability to distinguish between deterministic and stochastic data.

## 2. The proposed COPPS measure

### 2.1. Basic definitions

Let us consider a time series  $\{x_t\}_{t=0,1,\dots,T-1}$  of length  $T$  where  $t$  is the time index; and an embedding dimension  $n$ . By comparing neighboring values sorted into the ascending order in embedding vectors  $\mathbf{x}_k = (x_k, x_{k+\tau}, \dots, x_{k+(n-1)\tau})$ , where  $0 \leq k \leq T - (n-1)\tau - 1$ ,  $n \in \mathbb{N}_{\geq 3}$ , permutations  $\pi_k$  of length  $n$  are obtained,  $\tau \in \mathbb{N}_{\geq 1}$  is the delay time of the samples or time lag [19, 3, 4]. Indeed, each embedding vector is associated with the ordered vector of positions  $\pi^o = (1, 2, \dots, n)$ ; sorting  $\mathbf{x}_k$  into ascending order yields to the sorted vector  $\mathbf{x}_k^o$ , and rearranging values of  $\pi^o$  according to their previous position in  $\mathbf{x}_k$  leads to the ordinal pattern  $\pi_k$ . By considering the permutation  $\pi_k$  as a piece-wise linear interpolation function, the slope of each interpolation function is defined as

$$\vartheta_k(i) = \pi_k(i+1) - \pi_k(i), 1 \leq i \leq n-1, \quad (1)$$

where  $\vartheta_k(i)$  is the difference between pairs of neighboring values in  $\pi_k$ . The OPS encoding was defined by considering the largest slope  $S_k$  of  $\vartheta_k = (\vartheta_k(i))$  as  $S_k = \max(\{\vartheta_k(i)\}_{1 \leq i \leq n-1})$  over  $i$ . It was shown in the case of periodic dynamics with period  $L$  that  $|S_k| = L$  for  $n > L$  [19, 22]. The maximum number of distinct largest slopes is  $n$ , and all of them belong to the set  $\mathcal{S} = \{-1, 1, 2, \dots, n-1\}$  of largest slopes. We now combine the largest slope with its position  $i_k$  in  $(\vartheta_k(i))$  to define the OPPS encoding. The corresponding symbols are therefore denoted as  $P_k = (i_k, S_k)$ . An example is given for  $(x_t) = (3, 2, 3, 1, 3, 5, 2, 7, 4, 5)$ ,  $n = 4$ , and  $\tau = 1$ . In the case of the first embedding vector, sorting  $\mathbf{x}_0 = (3, 2, 3, 1)$  into ascending order leads to  $\mathbf{x}_0^o = (1, 2, 3, 3)$ , while the corresponding vector of positions  $\pi^o = (1, 2, 3, 4)$  leads to the OP  $\pi_0 = (4, 2, 1, 3)$ , by rearranging its values as

stated above. The application of Eq. (1) to  $\pi_0$  gives  $\vartheta_0 = (-2, -1, 2)$ . The corresponding largest slope is then  $S_0 = 2$  and its position is  $i_0 = 3$ , thus leading to the OPPS  $P_0 = (3, 2)$ . The complete OPPS encoding process is shown in Fig. 1.

There are  $n - 1$  possible slope positions  $1 \leq i_k \leq n - 1$ , and  $n$  possible slopes. In the case of redundant values of largest slope in  $\vartheta_k$ , the smallest index  $i$  is used as slope position  $i_k$ . Such a case is observed for  $\pi_5 = (2, 4, 1, 3)$  in Fig. 1 where the largest slope  $S = 2$  appears at positions  $i = 1$  and  $i = 3$ . While considering this precaution, slope  $S = -1$  only appears at position  $i_k = 1$ . Therefore, the cardinality of the set of OPPS is  $\Lambda_0(n) = (n - 1)^2 + 1$ .

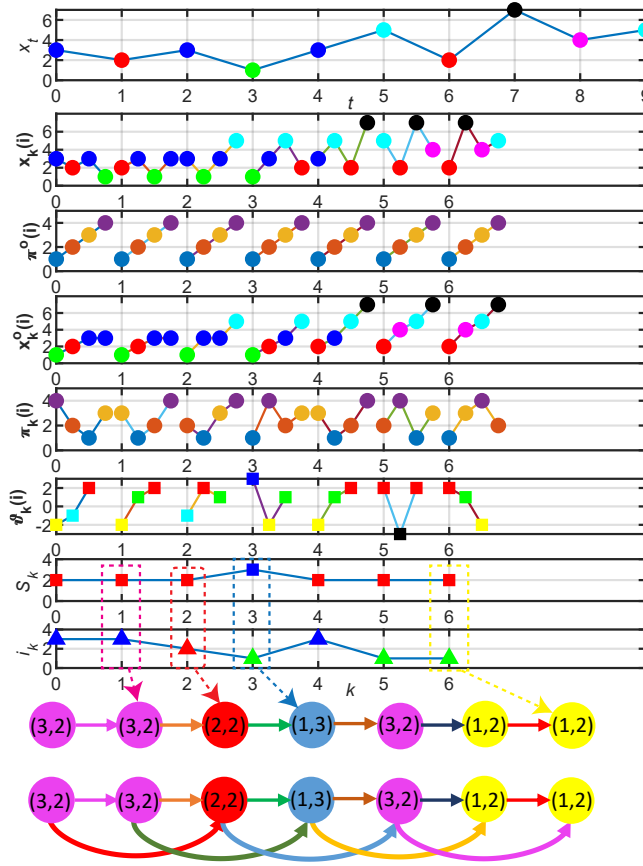


Figure 1: Example of OPPS encoding for  $n = 4$  and  $\tau = 1$ . From top to bottom:  $(x_t)$  time series;  $(\mathbf{x}_k)$  series;  $(\pi^o)$  series;  $(\mathbf{x}_k^o)$  series;  $(\pi_k)$  series;  $(\vartheta_k)$  series;  $(S_k)$  series;  $(i_k)$  series; 2-OPPS linking and 3-OPPS linking series.

## 2.2. The COPPS measure

We are not interested in the probability distribution [4] of OPPS, but only in their occurrence. Therefore, a system is complex as it produces

the whole set of possible OPPS. Such a definition gives the possibility to quantify the complexity of dynamical systems even from short observation times. In our assumption, the complexity of a dynamical system corresponds to its ability of generating new patterns as the time evolves. The approach is equivalent to considering the complementary of the number of forbidden/missing patterns as defined in the case of PE [23, 24, 25]. It has been established that forbidden patterns have the potential ability for distinguishing deterministic behavior (chaos) from randomness in finite time series contaminated with observational white noise [26, 27, 28]. Some systems may produce their complete set of patterns in a very short time, while the distribution of these patterns may continue to vary on time. Let  $R(n, \tau)$  be the number of distinct patterns generated by the system under investigation,  $R$  depends on the embedding dimension  $n$  and the time lag  $\tau$ . For a system with a monotonous dynamics (uniformly constant, increasing or decreasing), only a single pattern ( $R = 1$ ) is generated. For a purely random source, the full set of patterns is generated ( $R = (n - 1)^2 + 1$ ). While considering a period- $L$  dynamics, choosing  $\tau = L$  allows also to get  $R = 1$  as in the case of monotonous dynamics.

We further exploit the information in the time series by grouping consecutive OPPS to form new patterns. This approach is similar to constructing a transition network with OPPS as nodes. An example of 2-OPPS and 3-OPPS grouping/linking is shown in Fig. 1. Let  $s + 1$  be the number of OPPS grouped, the number of possible groups/patterns is  $\Lambda_s(n) = \left((n - 1)^2 + 1\right)^{s+1}$ . Therefore, we can define the  $s$ -link group multi-lag average  $r_s(n)$  as

$$r_s(n) = \left( \prod_{\tau=1}^{\tau_{max}} r_E(n, \tau) \right)^{\frac{1}{\tau_{max}}}, \quad (2)$$

where  $r_E(n, \tau) = \frac{R(n, \tau) - 1}{((n - 1)^2 + 1)^{s+1} - 1}$  is the normalized  $R(n, \tau)$ . By this definition, a regular dynamics is detected with  $r_s(n) = 0$  as there exists  $1 \leq \tau_d \leq \tau_{max}$  such that  $R(n, \tau_d) = 1$ , hence  $r_E(n, \tau_d) = 0$ . We define  $s$  as the complexity order or network depth. The case  $s = 0$  corresponds to the node average  $r_0(n)$ . The link average  $r_s(n)$ ,  $s > 0$ , is efficiently evaluated as  $T \gg T_{min}$ , with  $T_{min} = \Lambda_s(n)$ . The grouping approach gives us the possibility to increase the number of patterns and to further discriminate between complex systems.

By considering Eq. (2), we define the COPPS measure as the ratio of  $r_s$  when increasing the number of links

$$\lambda_s(n) = \frac{r_{s+1}(n)}{r_s(n)}, \quad s > 0. \quad (3)$$

Similar to the conditional probability,  $\lambda_s$  is the conditional  $r_s$  ensuring that two patterns that are similar for  $s$  points remain similar at the next point.

Increasing  $s$  allows to increase the number of possible patterns, therefore increasing their diversity in order to better assess the complexity of the time series. Regular dynamics are characterized by  $\lambda_s(n) = 0$ , while  $\lambda_s(n) = 1$  for purely random dynamics.  $\lambda_s(n) = 0$  in the case  $r_s(n) = r_{s+1}(n) = 0$ . The COPPS measure thus defined can also be considered with the OPS encoding where only largest slopes are taken into account, or the OP encoding where the entire permutations are considered.

### 3. Results and discussion

To demonstrate the potential of the new measure, we apply it on prototypical systems to measure complexity, detect transition phases and regular dynamics. We compare the results with the recurrence plot DET measure. An example of application to real-world systems is also considered with ECG data to demonstrate the ability of the method to perform with short length data and to detect arrhythmia.

#### 3.1. Application to chaotic time series

##### 3.1.1. Detection of the quasi-periodic route to chaos

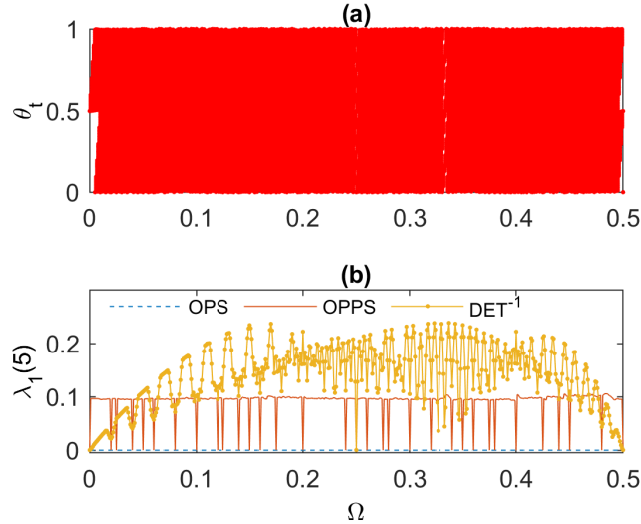


Figure 2: Detection of periodic dynamics with large periods in the rotation map. (a) Feigenbaum diagram; (b):  $DET^{-1}$  measure, COPPS diagram using OPS and OPPS encoding for  $T = 5 \times 10^3$ ,  $n = 5$ ,  $\tau_{max} = 30$ ,  $s = 1$ , and  $\gamma = 0$ .

We consider the sine-circle map defined as  $\theta_{t+1} = \Omega + \theta_t + \frac{\gamma}{2\pi} \sin(2\pi\theta_t) \bmod 1$ ,  $t \in \mathbb{N}$ .  $\Omega$  is the frequency ratio parameter and  $\gamma$  the nonlinearity parameter. For  $\gamma > 0$ , the map exhibits periodic, quasi-periodic and chaotic dynamics, depending on  $\Omega$ . When  $\Omega$  is a rational number and  $\gamma = 0$ , the map corresponds to the rotation map and exhibits periodic dynamics. Such

dynamics are difficult to detect as regular given their large period. We evaluated the COPPS measure of the rotation map for  $\theta_0 = 0.35$  and  $0 \leq \Omega \leq 0.5$ , varying by step size  $\Delta\Omega = 10^{-3}$ . We set  $n = 5$  and  $T = 10^4$ . The detection result is shown in Fig. 2. The COPPS measure is close to zero, and takes zero values for all dynamics at rational numbers  $\Omega$  whose denominators are smaller than or equal to  $n^{s+2}$ , instead of only  $n$  as compared to previous results in [29, 30, 31]. Thus, the COPPS measure is suited for the detection of regular dynamics with periods much larger than the embedding dimension. By increasing  $T$ , we observed that the value of  $\lambda_s$  remains unchanged, starting from  $T = 300$  for  $\tau = 30$ . While considering the OPS encoding, even those regular dynamics with large periods that failed to output  $\lambda_s(n) = 0$  with the OPPS encoding are well detected, as  $\lambda_s(n) = 0$  for  $0 \leq \Omega \leq 0.5$ .

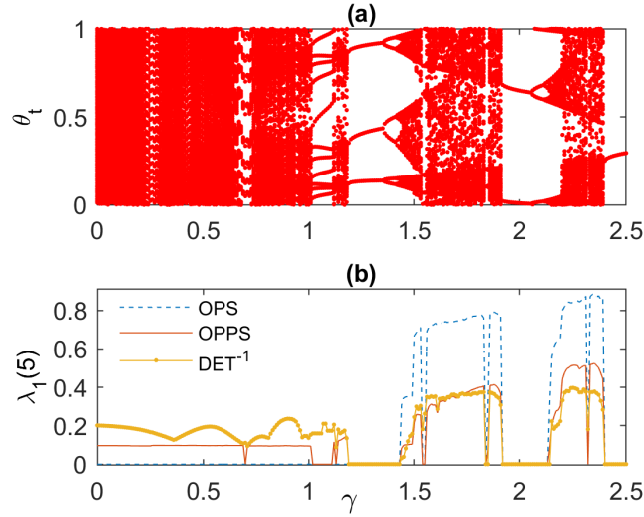


Figure 3: Detection of quasi-periodic route to chaos in the sine-circle map. (a) Feigenbaum diagram; (b):  $DET^{-1}$  measure, COPPS diagram using OPS and OPPS encoding for  $T = 5 \times 10^3$ ,  $n = 5$ ,  $\tau_{max} = 30$ ,  $s = 1$ , and  $\Omega = \frac{-1+\sqrt{5}}{2}$ .

For  $\Omega = \frac{-1+\sqrt{5}}{2}$ , the system exhibits quasi-periodic dynamics for  $0 \leq \gamma \leq 1$ , and periodic and chaotic dynamics as  $\gamma > 1$ . The detection results are depicted in Fig. 3. As in the case of the rotation map, quasi-periodic dynamics are detected with zero complexity while using the OPS encoding. In the case of the OPPS encoding, their complexity is close to zero and does not depend on  $T$  as it was the case for periodic dynamics with large periods. Thus, the COPPS measure efficiently detects quasi-periodic dynamics as regular dynamics. Indeed, the above results show that regular dynamics generate their complete set of patterns after a given observation time (approximately  $T = 300$  for  $\tau_{max} = 30$ ) as confirmed in Fig. 6, which justifies our definition for the COPPS measure to consider exclusively the occurrence of OPPS. Using the OPPS encoding gives more detail on the complexity of

the system. Such an observation is confirmed for  $\gamma = 1.61$  where a decrease on the  $\lambda_1(5)$  is observed on the OPPS plot, whereas the OPS plot remains smooth. However, both OPPS and OPS encodings shows large jumps in the COPPS measure when a band merging transition occurs (at  $\gamma = 1.49$  for example), which allows to easily detect it.

We further compare the result with the *DET* measure of the recurrence quantification analysis. The *DET* measure estimates the ratio of recurrence points that form diagonal structures. Indeed, processes with uncorrelated or weakly correlated, stochastic or chaotic behaviour are assumed to form none or very short diagonals, whereas deterministic processes cause longer diagonals and less single points [12]. *DET* is thus large (approximately half the data length) for periodic dynamics and close to zero for nonregular dynamics. We consider the inverse value of *DET* as a complexity measure. For periodic dynamics,  $DET^{-1}$  is supposed to be close to 0, depending on the data length; and for nonregular dynamics, it is supposed to be greater than zero. While looking at Figs. 2-3,  $DET^{-1}$  is close to 0 for period- $L$  dynamics as  $L$  is small. Both regular dynamics with large  $L$  and nonregular dynamics are detected with approximately similar large  $DET^{-1}$  values.

### 3.1.2. Detection of the period doubling route to chaos

The logistic map is defined as  $x_{t+1} = \mu x_t(1 - x_t)$ , where  $\mu$  is the control parameter and  $x_t$  the phase space amplitude at time index  $t$ . The system exhibits period doubling route to chaos for  $3.5 \leq \mu \leq 4$ , starting from period-4 dynamics. We evaluate the efficiency of the  $\lambda_s(n)$  measure by plotting the bifurcation diagram of the logistic map as depicted in Fig. 4. Fig. 4(b) shows that period- $L$  dynamics are detected as  $L < \tau_{max}$ . We verified that  $\lambda_s(n) = 0$  for period- $L$  dynamics as  $L \leq 2\tau_{max}$ . It is also observed that transition phases such as band mergings where the complexity of the system significantly increases, are evidently detected as  $\tau_{max}$  increases, although its computational cost also increases. Depending on the complexity order  $s$ , they are better detected with  $\lambda_1(n)$  than  $\lambda_2(n)$ , whose the transition detection efficiency is improved by increasing  $T$  and  $\tau_{max}$ . Fig. 4(c) presents the impact of the observation time on  $\lambda_s(n)$ . We observe that the sensitivity of  $\lambda_s(n)$  on  $T$  increases with the randomness of the dynamics under investigation, and  $T > \Lambda_{s+1}(n)$ . Indeed, for periodic and weakly chaotic dynamics,  $\lambda_s(n)$  does no longer significantly increase with  $T$  as  $T > \Lambda_{s+1}(n)$ . We verify that for a given  $\tau_{max}$ ,  $\lambda_s(n)$  is accurately evaluated as  $\Lambda_{s+1}(n) < T \leq 2\Lambda_{s+1}(n)$ . In the case  $T \gg 2\Lambda_{s+1}(n)$ , it is advisable to increase  $n$ . Furthermore, the complexity is better evaluated as  $n$  is large, but we recommend  $3 \leq n \leq 10$ . In the case  $T < \Lambda_{s+1}(n)$ , the length of the observation can be balanced by increasing  $\tau_{max}$ , but it is better to reduce  $n$  whenever  $n > 3$ . Fig. 4(d) shows that periodic dynamics are efficiently detected when comparing  $\lambda_s(5)$  with the Lyapunov exponent. The Lyapunov exponent is better approximated by  $\lambda_2(n)$  than  $\lambda_1(n)$  (see

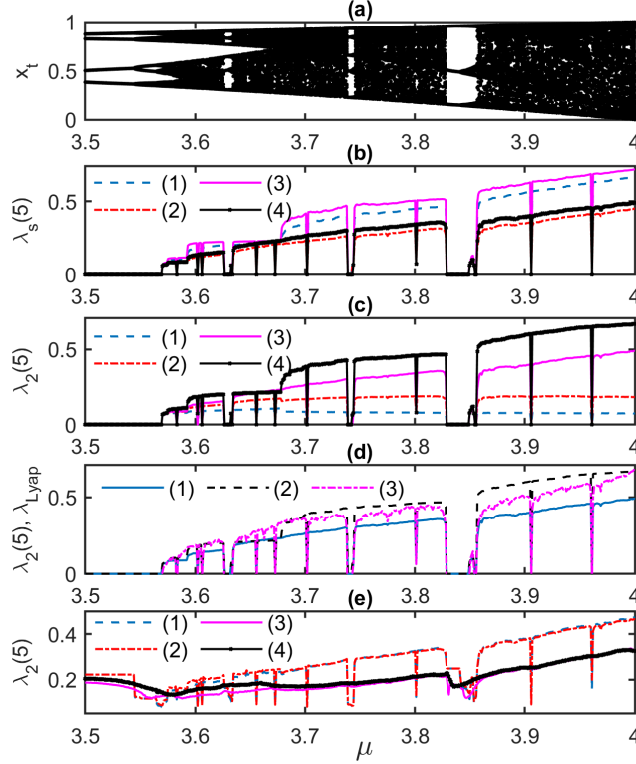


Figure 4: Logistic map for varying control parameter  $\mu$  (step size  $\Delta\mu = 10^{-3}$ ),  $x_0 = 0.35$ . (a): Bifurcation diagram; (b):  $\lambda_s(5)$  for (1)  $s = 1$  and  $\tau_{max} = 30$ , (2)  $s = 2$  and  $\tau_{max} = 30$ , (3)  $s = 1$  and  $\tau_{max} = 90$ , and (4)  $s = 2$  and  $\tau_{max} = 90$ ; (c):  $\lambda_2(5)$  for (1)  $T = 10^3$ , (2)  $T = 10^4$ , (3)  $T = 10^5$  and (4)  $T = 10^6$ ; (d): comparison of  $\lambda_2(5)$  (1)  $T = 10^5$  and  $\tau_{max} = 90$ , (2)  $T = 10^6$  and  $\tau_{max} = 30$ , and (3) Lyapunov exponent; (e): robustness of  $\lambda_s(n)$  against observational noise with (1)  $\sigma_b = 5 \cdot 10^{-4}$  and (2)  $\sigma_b = 10^{-3}$ , and dynamical noise with (3)  $\sigma_b = 5 \cdot 10^{-4}$  and (4)  $\sigma_b = 10^{-3}$ .

Fig. 4(b)). Indeed, the shape of  $\lambda_s(n)$  is close to the Lyapunov exponent as  $s$  is large and  $T \geq 2\Lambda_{s+1}(n)$ , which suggests that the Lyapunov exponent is better approximated by a high order complexity. Fig. 4(e) confirms the high robustness of  $\lambda_s(n)$  against both the observational and dynamical noise, as the shape of the  $\lambda_2(5)$  spectrum is preserved. This result also suggests that the  $\lambda_s(n)$  measure is suitable for distinguishing between chaotic and stochastic dynamics. We considered for this example the gaussian white noise (GWN) with standard deviation  $\sigma_b = 5 \cdot 10^{-4}$  and  $\sigma_b = 10^{-3}$ .

We also compare the ability of OP, OPS and OPPS encoding methods to discriminate time series. Fig. 5 shows the bifurcation diagrams of the logistic map obtained with the three approaches. It appears that OP and OPPS encoding behave similar in the case of chaotic dynamics, although the OP encoding requires large observations as  $n$  increases. Such a requirement justifies the weak amplitude of  $\lambda_1(5)$  for the OP encoding, as the observation

time  $T = 10^4$  is smallest than the minimal observation time required,  $T_{min} > (5!)^3$ . In addition, while increasing  $n$  is necessary for improving the analysis of the data sequence, the resulting large number of OP is memory and computationally costly, thus limiting the use of the OP encoding. The OPS approach better highlights transition phases (band merging transitions), but  $\lambda_s(n)$  values are large even for weakly chaotic dynamics, which reduces its discriminating power.

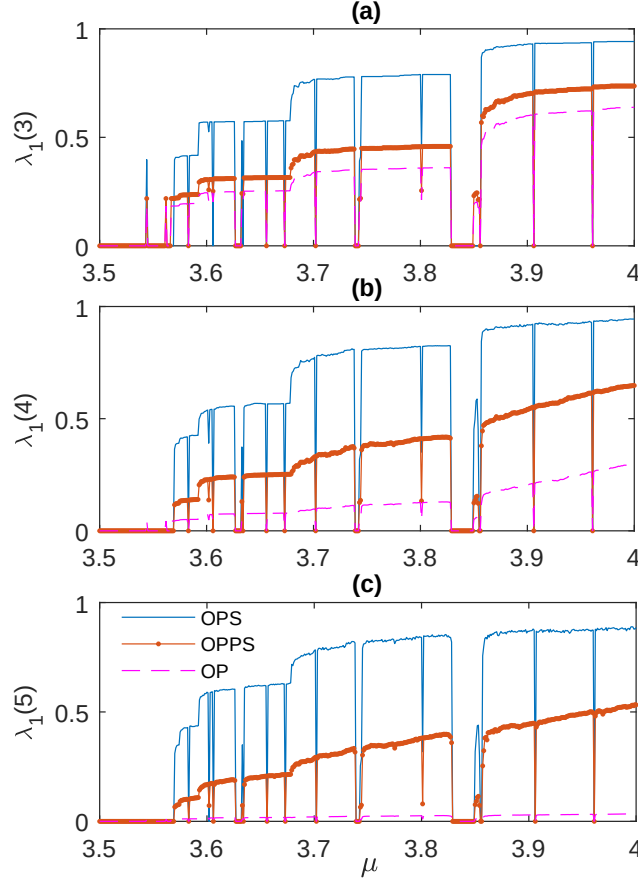


Figure 5: Comparison of OP, OPS and OPPS encoding methods for  $T = 10^4$ , and  $\tau_{max} = 30$ .

### 3.1.3. Distinguishing between chaotic and stochastic data

In this section, we compare the behavior of the COPPS measure for chaotic and stochastic time series. We considered colored noises as red (RN), pink (PN), blue (BN), and violet (VN), in addition to the GWN. While setting  $T = 10^6$  and  $\tau_{max} = 90$ , we find  $\lambda_1(5) = 0.7201$  and  $\lambda_2(5) = 0.6752$  for the logistic map with  $\mu = 4$ , and  $\lambda_1(5) = 0.9825$  and  $\lambda_2(5) = 0.8447$  for the GWN. This result shows that the complexity of the two dynamics

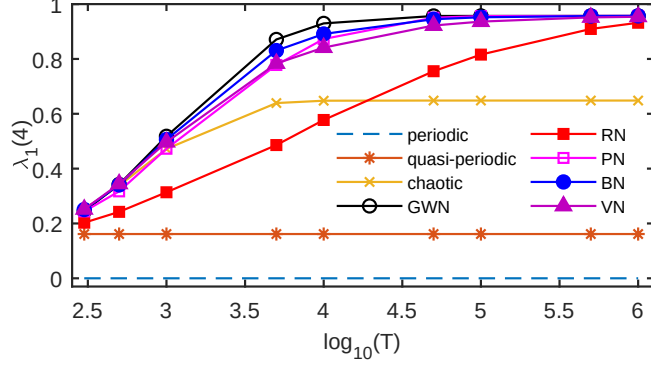


Figure 6: Behavior of COPPS measure with respect to the observation length  $T$ ,  $\tau_{max} = 30$ .

is different by a wide margin. Fig. 6 shows the behavior of the COPPS measure with respect to  $T$ , for period-4 and chaotic dynamics obtained from the logistic map with  $\mu = 3.5$  and  $\mu = 4$ , respectively; a quasi-periodic dynamics generated by the sine-circle map with  $\gamma = 0.2$ ; the RN, PN, BN, VN and GWN. We observe that  $\lambda_1(5)$  does not depend on  $T$  for regular dynamics, and converges to a limit value for each dynamics. The chaotic dynamics converges faster to its limit value that is smaller than those of the stochastic dynamics (RN, VN, and GWN). This observation suggests that chaotic dynamics generate faster their complete set of OPPS than stochastic dynamics.

Although the direct application of the COPPS measure to  $\{x_t\}$  allows to determine the effective complexity of the time series, it requires large values of  $T$  ( $T \geq 10^5$ ) to discriminate between chaotic and stochastic data, as shown in Fig. 6. Let us now consider  $\{x_t^{(m)}\}_{0 \leq t < T-1}$  as the series of OPPS computed from  $\{x_t\}$  with an embedding dimension  $m$ . We compute the complexity of  $\{x_t^{(m)}\}$  for  $3 \leq m \leq 25$  to distinguish between deterministic and stochastic data. Fig. 7 depicts the results for the chaotic dynamics generated from the logistic map with  $\mu = 4$ , and the stochastic dynamics. The above noises are power-law noise defined as signal with components at all frequencies  $f$ , with a power spectral density per unit of bandwidth proportional to  $f^\beta$ ,  $\beta \in \mathbb{R}$ . Each noise is then characterized by  $\beta$  and RN, PN, GWN, BN and VN are defined with  $\beta = -2$ ,  $\beta = -1$ ,  $\beta = 0$ ,  $\beta = 1$  and  $\beta = 2$ , respectively. The above transformation allows to make a clear difference between all the dynamics, ranging from red to violet noise. Indeed, the maximum value of  $\lambda_2(5)$  for the chaotic dynamics occurs at  $m = 2n$ . For  $m > 2n$ , values of  $\lambda_2(5)$  are increasingly ranged according to their respective values of  $\beta$ , starting from  $\beta = -2$ . For  $m \leq 2n$ , the complexity of all stochastic dynamics is greater than that of the chaotic dynamics. Given the large gap between the complexity of the two types of

data, it appears that the use of  $\{x_t^{(m)}\}$  improves the discriminating power of the COPPS measure as  $m < 2n$ .

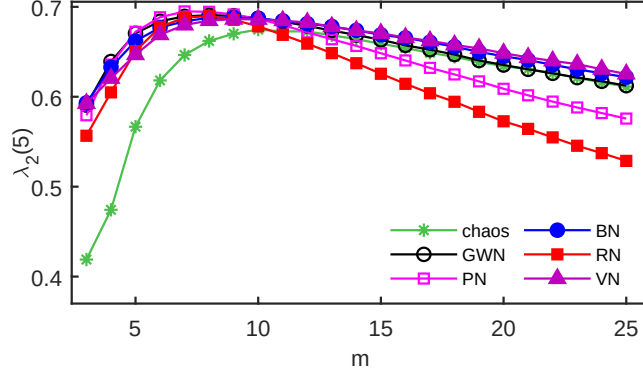


Figure 7: Comparison between chaotic and stochastic dynamics. We set  $T = 5 \times 10^5$  and  $\tau_{max} = 30$ .

Let us now consider  $\{x_t^{(m)}\}$  with  $m \leq n$  and investigate the smallest  $T$  required for discriminating between deterministic and stochastic data. The corresponding results for  $m = 3$  are depicted in Fig. 8, where it is observed that chaotic and stochastic data are discriminated, starting from  $T = 500$ , instead of  $T > 10^5$  as it is the case for the direct application. These results confirm that applying the COPPS measure to the OPPS series allows to improve the discriminating power of the COPPS method, hence to apply it to short length time series. Furthermore, these results improve the minimal time series length  $T_{min} = \Lambda_{s+1}$  required for direct application of the COPPS measure to  $\{x(t)\}$ .

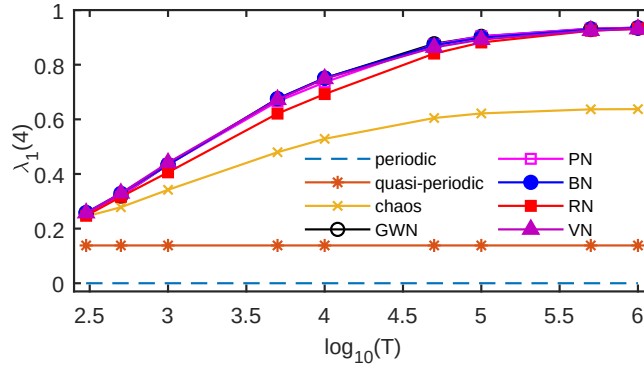


Figure 8: Discriminating between chaotic and stochastic data with the COPPS measure of the transformed data series  $\{x_t^{(3)}\}$  (OPPS series),  $\tau_{max} = 30$ .

We further consider the logistic map to evaluate the effectiveness of the COPPS measure to determine the complexity of short length sequences. We set  $n = 3$ ,  $s = 1$ ,  $\tau_{max} = 30$ , and  $T = 300$  and  $T = 1000$ . This choice is

motivated by the previous requirement for  $T > 2\Lambda_{s+1}$ , with  $\Lambda_{s+1} = 125$  for  $n = 3$  and  $s = 1$ . The corresponding results are depicted in Fig. 9. The results confirm the ability of the COPPS measure to characterize short length time series. Increasing  $T$  does not change the result for weakly chaotic dynamics, but increases the complexity value for strongly chaotic sequences. Transforming the original time series  $\{x(t)\}$  into a series of OPPS  $\{x^{(m)}(t)\}$  allows to improve the detection of periodic dynamics as shown in Fig. 9. Indeed, periodic dynamics ( $\mu = 3.602$  and  $\mu = 3.606$  for example) with large periods that are detected as chaotic ( $\lambda_1(3) > 0$ ) while applying the COPPS measure to  $\{x(t)\}$ , are clearly detected with  $\lambda_1(3) = 0$  while using  $\{x^{(m)}(t)\}$  with  $m = 3$  and  $m = 5$ . Such a result suggests that the transformation contributes to reduce the period of the original time series, thus making possible their detection even with short length sequences.

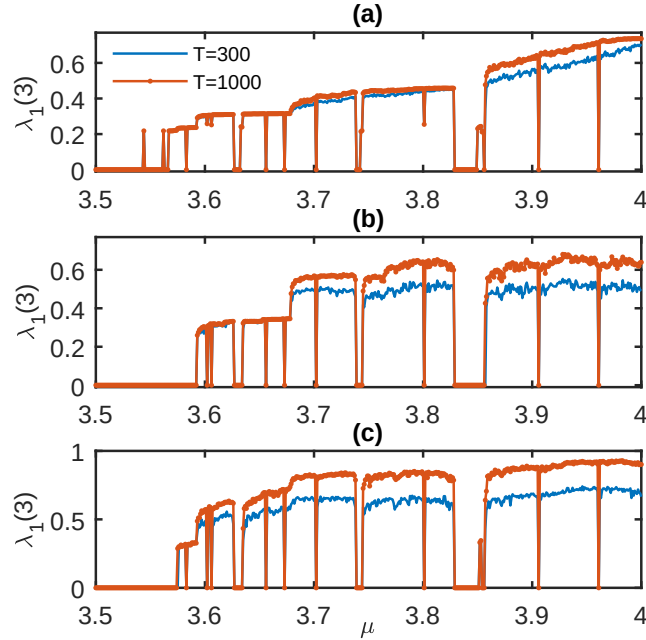


Figure 9: Chaos detection with application of COPPS measure to short length sequences of the logistic map,  $T = 300$ ,  $T = 1000$  and  $\tau_{max} = 30$ . We considered (a) original time series  $\{x(t)\}$ ; (b) OPPS series  $\{x^{(m)}(t)\}$  with  $m = 3$ ; and (c) OPPS series  $\{x^{(m)}(t)\}$  with  $m = 5$

### 3.2. Application to ECG signal analysis

An ECG allows physicians to diagnose cardiac function and cardiac disorders. Patient with cardiac arrhythmia can have a high risk to stroke, cardiac arrest, or sudden cardiac death. The heart rhythm abnormalities known to be responsible for these problems are principally detected as premature atrial contraction (PAC) and premature ventricular contraction (PVC) [32, 33].

The early detection of PAC and PVC allows to prevent and reduce the mortality due to cardiac disorders. We exploit the ability of the COPPS method

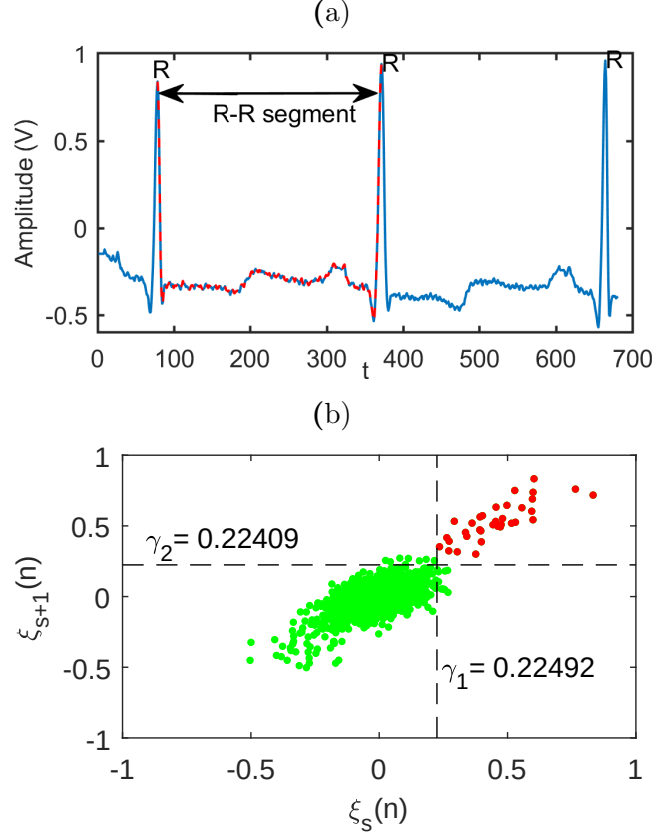


Figure 10: Example of arrhythmia detection in recording-100 with  $s = 1$ ,  $n = 3$ ,  $\alpha_1 = \alpha_2 = -0.25$ : (a) time representation of R-R signal, (b) classification result. 34 over 34 abnormal (red dots) and 2,237 normal (green dots) beats are well classified.

to detect transition phases by applying it to 30 minutes ECG recordings, with sampling frequency 360 Hz, of 20 patients from the MIT-BIH database [34, 35]. Indeed, we assume that each ECG beat in the whole recording is driven by a different unknown control parameter value, and significant changes in the control parameter, thereby the heart rhythm, may result in large variations in the beat COPPS measure. Raw data containing 41,054 beats (R-R segments) are analyzed, with 1,434 abnormal beats (type-”A” or PAC, and type-”V” or PVC beats). Assuming ECG beats are short length data with  $150 < T < 300$  samples for each beat, we set  $n = 3$ ,  $\tau_{max} = 30$  and  $s = 1$ . In order to detect abnormal beats, we consider the  $(\xi_s, \xi_{s+1})$  plane and two thresholds defined as  $\gamma_j = \left| \frac{M_{s+j-1} + m_{s+j-1}}{2} \right| + \alpha_j \sigma_j$ ,  $j \in \{1, 2\}$ , where

$m_{s+j-1} = \min(\xi_{s+j-1}(n))$ , and

$$\xi_{s+j-1}(n) = \frac{\Delta\lambda_{s+j-1}(n)}{1.2 \max(|\Delta\lambda_{s+j-1}(n)|)}.$$

The median and standard deviation of  $\xi_{s+j-1}(n)$  are respectively  $M_{s+j-1}$  and  $\sigma_{s+j-1}$ , and  $\xi_s$  is the normalized difference of  $\lambda_s$  measures for consecutive beats. It allows to detect transition phases with large gaps in the cardiac rhythm, thereby abnormal beats, as depicted in Fig. 10.  $\alpha_j \in \mathbb{R}$  are learning parameters such that  $-1 \leq \alpha_j \leq 1$ . Any beat with  $\xi_s(n) > \gamma_1 \wedge \xi_{s+1}(n) > \gamma_2$  is assumed to be abnormal. After analysis with  $\alpha_1 = \alpha_2 = -0.5$ , we obtain the following results: overall accuracy  $A_{cc} = 98.08\%$ , sensitivity  $S_e = 99.07\%$ , specificity  $S_p = 70.71\%$ , positive predictive value  $P^+ = 73.42\%$ , and negative predictive value  $P^- = 98.94\%$ . Our method outperforms a recently proposed machine learning approach [36], and our results are close to those obtained in [37] where PE was combined with morphological quantifiers. By combining the COPPS measure with other features and methods as in [38, 39, 40], these results could be further advanced, and a multi-class classification envisaged. We also observed that abnormal beats are preceded in general by normal beats with low complexity as compared to the mean value, which could help to predict their occurrences.

#### 4. Conclusion

We present in this paper the new COPPS measure. The proposed approach is directly applicable to any type of time series, and shows a highly robustness against noise. The  $\lambda_s(n)$  measure allows to distinguish between periodic, quasi-periodic, chaotic and stochastic dynamics. Chaos-chaos transitions are detected as the complexity level  $s$  is low. Quasi-periodic dynamics and periodic dynamics with large periods are detected with  $\lambda_s = 0$  while considering the OPS encoding, but the complexity of nonregular dynamics is well evaluated with the OPPS encoding. The application of the COPPS measure to ECG data confirms its efficiency for analyzing short length data; and reveals that arrhythmia beats are characterized by large variations in the complexity, and preceded by normal beats with low complexity. Thus, the method could help to early predict extreme events from observations, hence to predict catastrophes.

#### References

- [1] C. Masoller, Y. Hong, S. Ayad, F. Gustave, S. Barland, A. Pons, S. Gómez, A. Arenas, Quantifying sudden changes in dynamical systems using symbolic networks, New J. Phys. 17 (2015) 023068. doi:10.1088/1367-2630/17/2/023068.

- [2] R. Bartsch, J. W. Kantelhardt, T. Penzel, S. Havlin, Experimental evidence for phase synchronization transitions in the human cardiorespiratory system, *Phys. Rev. Lett.* 98 (2007) 054102. doi:10.1103/PhysRevLett.98.054102.
- [3] C. Bandt, Ordinal time series analysis, *Ecol. Modell.* 182 (2005) 229 – 238. doi:10.1016/j.ecolmodel.2004.04.003.
- [4] C. Bandt, B. Pompe, Permutation entropy: A natural complexity measure for time series, *Phys. Rev. Lett.* 88 (2002) 174102. doi:10.1103/PhysRevLett.88.174102.
- [5] J. S. A. E. Fouda, The matching energy: a novel approach for measuring complexity in time series, *Nonlinear Dynamics* 86 (2016) 2049–2060. doi:10.1007/s11071-016-3014-8.
- [6] S. M. Pincus, Approximate entropy (ApEn) as a complexity measure, *Chaos* 5 (1995) 110–117. doi:10.1063/1.166092.
- [7] O. Afsar, G. B. Bagei, U. Tirnakli, Renormalized entropy for one dimensional discrete map: periodic and quasiperiodic route to chaos and their robustness, *Eur. Phys. J. B* 86 (2013) 307–320. doi:10.1140/epjb/e2013-40184-2.
- [8] J. S. Richman, J. R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, *Circulatory Physiology* 278 (2000) H2039–H2049. doi:10.1152/ajpheart.2000.278.6.H2039.
- [9] N. Marwan, J. F. Donges, Y. Zou, R. V. Donner, J. Kurths, Complex network approach for recurrence analysis of time series, *Physics Letters A* 373 (2009) 4246 – 4254. doi:10.1016/j.physleta.2009.09.042.
- [10] R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths, H. Scheingraber, A comparative classification of complexity measures, *Chaos, Soliton & Fractals* 4 (1994) 133–173. doi:10.1016/0960-0779(94)90023-X.
- [11] P. E. Rapp, C. J. Cellucci, K. E. Korlund, T. A. Watanabe, M. A. Jimenez-Montano, Effective normalization of complexity measurements for epoch length and sampling frequency, *Phys. Rev. E* 64 (2001) 016209. doi:10.1103/PhysRevE.64.016209.
- [12] N. Marwan, M. C. Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems, *Physics Reports* 438 (2007) 237–329. doi:10.1016/j.physrep.2006.11.001.
- [13] N. Marwan, How to avoid potential pitfalls in recurrence plot based data analysis, *International Journal of Bifurcation and Chaos* 21 (2011) 1003–1017. doi:10.1142/S0218127411029008.

- [14] M. Kennel, M. Buhl, Estimating good discrete partitions from observed data: Symbolic false nearest neighbors, *Phys. Rev. Lett.* 91 (2003) 084102. doi:10.1103/PhysRevLett.91.084102.
- [15] Y. Hirata, K. Judd, D. Kilminster, Estimating a generating partition from observed time series: Symbolic shadowing, *Phys. Rev. E* 70 (2004) 016215. doi:10.1103/PhysRevE.70.016215.
- [16] M. Buhl, M. Kennel, Statistically relaxing to generating partitions for observed time-series data, *Phys. Rev. E* 71 (2005) 046213. doi:10.1103/PhysRevE.71.046213.
- [17] Y. Zou, R. V. Donner, N. Marwan, J. F. Donges, J. Kurths, Complex network approaches to nonlinear time series analysis, *Physics Reports* 787 (2019) 1–97. doi:10.1016/j.physrep.2018.10.005.
- [18] S. Schinkel, N. Marwan, J. Kurths, Order patterns recurrence plots in the analysis of erp data, *Cognitive Neurodynamics* 1 (2007) 317–325. doi:10.1007/s11571-007-9023-z.
- [19] J. S. A. E. Fouda, W. Koepf, Detecting regular dynamics from time series using permutations slopes, *Commun. Nonlinear Sci. Numer. Simulat.* 27 (2015) 216 – 227. doi:10.1016/j.cnsns.2015.03.008.
- [20] J. S. A. E. Fouda, Applicability of the permutation largest slope entropy to strange nonchaotic attractors, *Nonlinear Dyn.* 87 (2017) 1859 – 1871. doi:10.1007/s11071-016-3158-6.
- [21] J. S. A. E. Fouda, B. Bodo, G. M. D. Djeufa, S. L. Sabat, Experimental chaos detection in the Duffing oscillator, *Commun. Nonlinear Sci. Numer. Simulat.* 33 (2016) 259 – 269. doi:10.1016/j.cnsns.2015.09.011.
- [22] J. S. A. E. Fouda, J. Y. Effa, M. Kom, M. Ali, The three-state test for chaos detection in discrete maps, *Applied Soft Computing* 13 (2013) 4731 – 4737. doi:10.1016/j.asoc.2013.07.020.
- [23] D. Cuesta-Frau, A. Molina-Picó, B. Vargas, P. González, Permutation entropy: Enhancing discriminating power by using relative frequencies vector of ordinal patterns instead of their shannon entropy, *Entropy* 21 (2019) 1013. doi:10.3390/e21101013.
- [24] D. Cuesta-Frau, Using the information provided by forbidden ordinal patterns in permutation entropy to reinforce time series discrimination capabilities, *Entropy* 22 (2020) 1013. doi:10.3390/e22050494.
- [25] O. A. Rosso, L. C. Carpi, P. M. Saco, M. G. Ravetti, H. A. Larrondo, A. Plastino, The amigó paradigm of forbidden/missing patterns: a detailed analysis, *The European Physical Journal B* 85 (2012). doi:10.1140/epjb/e2012-30307-8.

- [26] J. M. Amigó, L. Kocarev, J. Szczepanski, Order patterns and chaos, *Phys. Lett. A* 355 (2006) 27–31. doi:10.3390/e22050494.
- [27] J. M. Amigó, S. Zambrano, M. A. F. Sanjuán, True and false forbidden patterns in deterministic and random dynamics, *Europhys. Lett.* 79 (2007) 50001. doi:10.3390/e22050494.
- [28] L. C. Carpi, P. M. Saco, O. Rosso, Missing ordinal patterns in correlated noises, *Physica A: Statistical Mechanics and its Applications* 389 (2010) 2020–2029. doi:<https://doi.org/10.1016/j.physa.2010.01.030>.
- [29] J. S. A. E. Fouda, W. Koepf, Complexity measure by ordinal matrix growth modeling, *Nonlinear Dynamics* 89 (2017) 1385–1395. doi:10.1007/s11071-017-3523-0.
- [30] J. S. A. E. Fouda, W. Koepf, S. Jacquir, The ordinal Kolmogorov-Sinai entropy: A generalized approximation, *Commun Nonlinear Sci Numer Simulat* 46 (2017) 103–115. doi:10.1016/j.cnsns.2016.11.001.
- [31] A. M. Unakafov, K. Keller, Conditional entropy of ordinal patterns, *Physica D: Nonlinear Phenomena* 269 (2014) 94–102. doi:10.1016/j.physd.2013.11.015.
- [32] M. Carrara, L. Carozzi, T. J. Moss, M. de Pasquale, S. Cerutti, M. Ferrario, D. E. Lake, J. R. Moorman, Heart rate dynamics distinguish among atrial fibrillation, normal sinus rhythm and sinus rhythm with frequent ectopy, *Physiol. Meas.* 36 (2015) 1873–1888. doi:10.1088/0967-3334/36/9/1873.
- [33] Y. Huang, H. Li, X. Yu, A novel time representation input based on deep learning for ecg classification, *Biomedical Signal Processing and Control* 83 (2023) 104628. doi:10.1088/0967-3334/36/9/1873.
- [34] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, H. E. Stanley, Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals, *Circulation* 101 (2000) e215–e220. doi:10.1161/01.cir.101.23.e215.
- [35] G. Moody, R. Mark, The impact of the mit-bih arrhythmia database, *IEEE Engineering in Medicine and Biology Magazine* 20 (2001) 45–50. doi:10.1109/51.932724.
- [36] S. Singh, S. K. Pandey, U. Pawar, R. R. Janghel, Classification of ECG arrhythmia using recurrent neural networks, *Procedia Computer Science* 132 (2018) 1290–1297. doi:10.1016/j.procs.2018.05.045.

- [37] J. B. B. à Mougoufan, J. S. A. E. Fouda, M. Tchuenta, W. Koepf, Adaptive ECG beat classification by ordinal pattern based entropies, *Commun. Nonlinear Sci. Numer. Simulat.* 84 (2020) 105–156. doi:10.1016/j.cnsns.2019.105156.
- [38] S. K. pandey, A. Shukla, S. Bhatia, T. R. Gadekallu, A. Kumar, A. Mashat, M. A. Shah, R. R. Janghel, Detection of arrhythmia heartbeats from ecg signal using wavelet transform-based cnn model, *Int J Comput Intell Syst* 16 (2023) –. doi:10.1007/s44196-023-00256-z.
- [39] M. Carrara, L. Carozzi, T. J. Moss, M. de Pasquale, S. Cerutti, M. Ferrario, D. E. Lake, J. R. Moorman, Heart rate dynamics distinguish among atrial fibrillation, normal sinus rhythm and sinus rhythm with frequent ectopy, *Physiol. Meas.* 36 (2015) 1873–1888. doi:10.1088/0967-3334/36/9/1873.
- [40] B. Chen, D. M. Maslove, J. D. Curran, A. H. P. R. Laird, P. Mousavi, S. Sibley, A deep learning model for the classification of atrial fibrillation in critically ill patients, *Intensive Care Medicine Experimental* 11 (2023) –. doi:10.1186/s40635-022-00490-3.