

The ordinal Kolmogorov-Sinai entropy: A generalized approximation

J. S. Armand Eyebe Fouda^a, Wolfram Koepf^b, Sabir Jacquir^c

^a*Department of Physics, Faculty of Science, University of Yaoundé I, P.O. Box 812, Yaoundé, Cameroon*

^b*Institute of Mathematics, University of Kassel, Heinrich-Plett Str. 40, 34132 Kassel, Germany*

^c*LE2I UMR 6306, CNRS, Arts et Métiers, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France*

Abstract

We introduce the multi-dimensional ordinal arrays complexity as a generalized approximation of the ordinal Komogorov-Sinai entropy. The ordinal arrays entropy (OAE) is defined as the Shannon entropy of a series of m -ordinal patterns encoded symbols, while the ordinal arrays complexity (OAC) is defined as the differential of the OAE with respect to m . We theoretically establish that the OAC provides a better estimate of the complexity measure for short length time series. Simulations were carried out using discrete maps, and confirm the efficiency of the OAC as complexity measure from a small data set even in a noisy environment.

Keywords: complexity, entropy, ordinal pattern, ordinal array

1. Introduction

Measuring the complexity from series of observation is crucial for various fields of research. Existing theoretical complexity measures such as Lyapunov exponents, Kolmogorov-Sinai (KS) entropy, correlation dimension and many others are difficult to estimate from a finite data set [1–6], while empirical measures like the renormalized entropy or the approximate entropy often lack a theoretical foundation or are not well interpretable [7–9].

Considering the above limitations of the existing complexity measures, the permutation entropy (PE), an ordinal pattern analysis (OPA) based method, was introduced by Bandt and Pompe [10]. The PE is easily estimated from any type of data set and its theoretical foundation has been clearly established. Despite its well known performances, the PE does not well approximate the KS entropy for a finite data set. Therefore, the method cannot clearly state whether a dynamics is regular or not. Based on this observation, various modifications as well as other OPA based approaches have been developed [11–18]. Despite the improvements they brought, these methods have not yet clearly reached a zero

Email addresses: efoudajsa@yahoo.fr, Tel: +237670836099 (J. S. Armand Eyebe Fouda), koepf@mathematik.uni-kassel.de, Tel: +49 561 804 4207 (Wolfram Koepf), sjacquir@u-bourgogne.fr, Tel: +33380395886 (Sabir Jacquir)

URL: <http://www.mathematik.uni-kassel.de/~koepf/> (Wolfram Koepf)

complexity for regular dynamics as it should be. Indeed, although OPA based methods are very promising for applications of the dynamical system theory to real-world systems, they nevertheless present a fundamental limitation as their accuracy needs large embedding dimension and infinite time series length. In practice, such requirements are difficult to satisfy. Thus, it is judicious to try to improve the existing results for finite data sets. Recently the conditional entropy of ordinal patterns (CPE) was proposed where the average diversity of ordinal patterns succeeding a given one is characterized [19]. The CPE provides a much better practical estimation of the KS entropy than the permutation entropy, while having the same computational efficiency.

However, similarly as the PE, the CPE is estimating the complexity from the frequency occurrence of ordinal patterns. Indeed, the idea behind OPA based methods is to consider order relations between values of time series instead of the values themselves. The original time series is firstly converted into a sequence of ordinal patterns of an order n , each of them describing order relations between n successive points of the time series. The entropy is then estimated using the frequencies of the various ordinal patterns. Using the ordinal pattern transformation undoubtedly reduces the number of patterns of the same order while considering the values of the time series. In fact two different embedding vectors derived from the time series may lead to the same ordinal pattern, hence directly calculating the entropy from the ordinal patterns frequencies leads to some information loss.

In this paper, we suggest to apply the OPA to the values of the series of ordinal patterns derived from the time series. As the effective number of symbols can be evaluated for ordinal patterns of order n , the entropy of the new source of symbols can be well estimated. This approach allows to increase the number of patterns which has been reduced by the conversion into ordinal patterns, therefore to better approximate the entropy of the time series. For instance, for a given real-world time series with Λ distinct symbols or values, the number of distinct ordinal patterns of order n is $n!$, while the number of distinct embedding vectors of the same order derived from the time series values is Λ^n . Now considering the series of ordinal patterns of order n , the number of possible distinct embedding vectors of order n which can be derived from the time series values increases as $(n!)^n \gg n!$.

The paper presents first a brief recall of the permutation entropy in section 2. The ordinal arrays entropy (OAE) is introduced in section 3 followed by the definition of the ordinal arrays complexity (OAC) in section 4. In section 5 some simulation results are given, followed by concluding remarks in section 6.

2. Brief recall of the permutation entropy

Let $\{x_t\}_{t=0,1,\dots,T-1}$ be a time series of length T where t is the time index. The PE of order n is defined as a measure of the probabilities of permutations of order n [10]. Permutations of order n are obtained from the comparison of neighboring values (increasing order) in embedding vectors $\mathbf{x}_k = (x_{k\tau_0}, x_{k\tau_0+\tau}, \dots, x_{k\tau_0+l\tau}, \dots, x_{k\tau_0+(n-1)\tau})$, where $k \in \mathbb{N}$, n is the embedding dimension (number of values in $\mathbf{x}_k$), $\tau_0 \in \mathbb{N}_{\geq 1}$ is the delay time of the embedding vectors, $\tau \in \mathbb{N}_{\geq 1}$ is the delay time of samples and $l+1$ the index of $x_{k\tau_0+l\tau}$ in $\mathbf{x}_k$, $l \in \mathbb{N}$. Let P_k be the permutation derived from $\mathbf{x}_k$, with $\tau_0 = 1$.

$P_k = \left(\frac{1, 2, 3, \dots, n}{5, n, 1, \dots, 3} \right)$ for example is obtained by sorting the values of $\mathbf{x}_k$ in ascending order, with $x_{k+4\tau} < x_{k+(n-1)\tau} < x_k < \dots < x_{k+2\tau}$. Identical values are sorted by ascending order of their time index. The permutation entropy of order n is thus given by

$$H_p(n) = - \sum p(\theta) \cdot \ln(p(\theta)), \quad (1)$$

where

$$p(\theta) = \frac{\#\{k \mid k \leq T - n\tau, P_k = \theta\}}{T - n\tau + 1}, \quad (2)$$

is the probability of the permutation θ and $\#$ denotes the cardinality [10].

Definition 1 [20]:

A time series $\{x_t\}$ is called period- L cycle or simply L -periodic, if there exists a basic pattern of length q samples containing L distinct values ($L \leq q$) periodically repeated, independently of the time origin. q is known as the time space period and L as the phase space period.

Choosing $\tau > 1$ can lead to some misinterpretations in complexity values in the case of regular dynamics. Indeed, we showed in [20] that the number of distinct permutations in the case of L -periodic dynamics is fluctuating between $\delta = \frac{L}{\gcd(L, \tau)}$ and L , depending on the ordering of the time series. It therefore results that the PE is such that $\ln(\delta) \leq H_p(n) \leq \ln(L)$. Considering this dependence on the ordering of the time series, the bias between the PE and the largest Lyapunov exponent in the case of regular dynamics cannot be determined rigorously.

3. Ordinal arrays entropy

3.1. Designing of M -dimensional ordinal arrays

Measuring the Shannon entropy of an information source requires that the number of symbols is known and that the set of symbols (often called as the alphabet $\mathcal{A}$) is of finite size; otherwise, the Shannon entropy equals infinity. It is common to define the information entropy in bits. In the case of continuous random variables such as the speech signal for example, a uniform quantizer can be used to convert the corresponding continuous alphabet into a discrete one. Symbols of the discrete alphabet thus obtained are encoded with a finite number of bits. Similarly, the ordinal pattern analysis approach which considers the order relations between values of the time series instead of the values themselves can be seen as a quantizer. It then converts the continuous alphabet of the source into an ordinal pattern alphabet. Each ordinal pattern is encoded using positive integers similarly as bits are commonly used. The set of ordinal patterns can then be considered as a pseudo-source of information. If n is the length of ordinal patterns, then the cardinality of the ordinal pattern alphabet is equal to $n!$, while defining patterns of the same length using the values of the time series would have led to an Λ^n -length alphabet, Λ being the number of distinct values of the time series. It appears that considering the order relation between values instead of the values themselves reduces the number of accessible patterns. This reduction of the number of accessible patterns also leads to a reduction of the diversity of the original time series, hence the reduction of its complexity. This observation may justify why the permutation entropy of finite order n does not well approximate the KS entropy.

Although the PE converges to the KS entropy as n tends to infinity, the permutation entropy of finite order can be either much greater or much smaller than the KS entropy [19]. In order to reduce the difference between the PE and the KS entropy, the CPE has been defined as the average diversity of ordinal patterns succeeding a given one. The CPE characterizes the diversity of the successors of the given ordinal patterns, whereas the PE characterizes the diversity of the ordinal patterns themselves. In our approach, we propose to convert the pseudo-source of ordinal patterns into a set of M -dimensional arrays of ordinal patterns or simply M -dimensional ordinal arrays (OA) and to measure their diversity. Indeed, given the set $\{P_k\}$ of ordinal patterns, we define another set $\{\mathbf{S}_{t_1}^{(1)}\}$ of OA such that $\mathbf{S}_{t_1}^{(1)} = (P_{t_1}^T, P_{t_1+\tau}^T, \dots, P_{t_1+(d_1-1)\tau}^T)$, where $P_{t_1}^T$ is a column vector (the transpose of P_{t_1}). P_{t_1} is an n -length vector of positive integers in $\{1, 2, 3, \dots, n\}$, while $\mathbf{S}_{t_1}^{(1)}$ is an $n \times d_1$ matrix of the same integer set, and d_1 the number of permutations considered to form $\mathbf{S}_{t_1}^{(1)}$.

Let us clarify this presentation by considering for example a period-5 cycle orbit obtained by generating 5 distinct random numbers (0.8147, 0.9058, 0.1270, 0.9134, 0.6324) and repeating this basic sequence K -times ($K > 2$). The five distinct 5-order permutations obtained by sorting the values of vectors $\mathbf{x}_k$, $k=0$ to 4, are the following: $P_0 = \begin{pmatrix} 1,2,3,4,5 \\ 3,5,1,2,4 \end{pmatrix}$; $P_1 = \begin{pmatrix} 1,2,3,4,5 \\ 2,4,5,1,3 \end{pmatrix}$; $P_2 = \begin{pmatrix} 1,2,3,4,5 \\ 1,3,4,5,2 \end{pmatrix}$, $P_3 = \begin{pmatrix} 1,2,3,4,5 \\ 5,2,3,4,1 \end{pmatrix}$ and $P_4 = \begin{pmatrix} 1,2,3,4,5 \\ 4,1,2,3,5 \end{pmatrix}$. The corresponding set of ordinal patterns for the whole time series is $\{P_{t_1}\} = \{P_0, P_1, P_2, P_3, P_4, P_0, P_1, P_2, P_3, P_4, P_0, P_1, P_2, \dots\}$. From $\{P_{t_1}\}$, one can deduce for example the set of 5×3 OA as $\{\mathbf{S}_{t_1}^{(1)}\} = \{\mathbf{S}_0^{(1)}, \mathbf{S}_1^{(1)}, \mathbf{S}_2^{(1)}, \mathbf{S}_3^{(1)}, \mathbf{S}_4^{(1)}, \mathbf{S}_0^{(1)}, \dots\}$,

$$\text{where } \mathbf{S}_0^{(1)} = \begin{pmatrix} 3 & 2 & 1 \\ 5 & 4 & 3 \\ 1 & 5 & 4 \\ 2 & 1 & 5 \\ 4 & 3 & 2 \end{pmatrix}, \mathbf{S}_1^{(1)} = \begin{pmatrix} 2 & 1 & 5 \\ 4 & 3 & 2 \\ 5 & 4 & 3 \\ 1 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}, \mathbf{S}_2^{(1)} = \begin{pmatrix} 1 & 5 & 4 \\ 3 & 2 & 1 \\ 4 & 3 & 2 \\ 5 & 4 & 3 \\ 2 & 1 & 5 \end{pmatrix},$$

$$\mathbf{S}_3^{(1)} = \begin{pmatrix} 5 & 4 & 3 \\ 2 & 1 & 5 \\ 3 & 2 & 1 \\ 4 & 3 & 2 \\ 1 & 5 & 4 \end{pmatrix}, \text{ and } \mathbf{S}_4^{(1)} = \begin{pmatrix} 4 & 3 & 2 \\ 1 & 5 & 4 \\ 2 & 1 & 5 \\ 3 & 2 & 1 \\ 5 & 4 & 3 \end{pmatrix}. \text{ Therefore, } n = 5 \text{ and } d_1 = 3,$$

$\{\mathbf{S}_{t_1}^{(1)}\}$ is a set of 2-dimensional OA (matrices) and is also 5-periodic as the generating time series $\{x_t\}$. Choosing any other value of $d_1 \geq 1$ will output a period-5 series of OA.

Considering the previous series $\{\mathbf{S}_{t_1}^{(1)}\}$ of matrices, a series $\{\mathbf{S}_{t_2}^{(2)}\}$ of 3-dimensional OA, with $\mathbf{S}_{t_2}^{(2)} = (\mathbf{S}_{t_2}^{(1)}, \mathbf{S}_{t_2+\tau}^{(1)}, \dots, \mathbf{S}_{t_2+(d_2-1)\tau}^{(1)})$ can be constructed. From $\{\mathbf{S}_{t_2}^{(2)}\}$ can be constructed a series $\{\mathbf{S}_{t_3}^{(3)}\}$ of 4-dimensional OA with $\mathbf{S}_{t_3}^{(3)} = (\mathbf{S}_{t_3}^{(2)}, \mathbf{S}_{t_3+\tau}^{(2)}, \dots, \mathbf{S}_{t_3+(d_3-1)\tau}^{(2)})$. The process can be continued until the desired series $\{\mathbf{S}_{t_{(M-1)}}^{(M-1)}\}$ of M -dimensional OA is obtained. $\mathbf{S}_{t_{(M-1)}}^{(M-1)}$ is a $d_0 \times d_1 \times d_2 \times \dots \times d_j \times \dots \times d_{M-1}$ structure, with $0 \leq j \leq M-1$ and $d_0 = n$. We set $\tau = 1$ for the rest of the paper.

In the previous example, only 5 permutations are obtained, given the periodic nature of the time series. In the general case, the maximum number of permutations which can be derived from a given times series is $n!$. Considering

this possible maximum number of permutations, the possible (maximum) number of ordinal matrices that can be derived from the series of ordinal patterns is $(n!)^{d_1}$. In the case of 2-dimensional series, each ordinal matrix is encoded using d_1 permutations similarly as bits are commonly used in the information theory. In general, for a given time series of length T , the length of a resulting $A = n \times d_1 \times d_2 \times \cdots \times d_{M-1}$ series of OA, is equal to $N = T - (n + r - 1)$, where

$$r = \sum_{j=1}^{M-1} (d_j - 1), d_j > 0. \quad (3)$$

For a T -length 1-D OA series with embedding dimension n , the length of the resulting OA series is $N = T - n + 1$. Extending this calculation to multi-dimensional OA (M-D OA), the equivalent embedding dimension allowing to obtain a M-D OA series can be given by

$$\rho = n + r. \quad (4)$$

In the ordinal pattern analysis approach, symbols are permutations and are encoded using natural numbers while in the OA based analysis we are defining, symbols are arrays of permutations encoded using positive integers. So, the entropy by the proposed approach can be seen as the Shannon entropy of M-D OA derived from the permutation series. The method combines both the ordering relations between values and the values themselves. The idea is to consider the diversity of the OA as well as their ranking in the time series. The ordering relation is used for determining the permutation series while the values of OA themselves are used for determining the entropy of M-D OA.

3.2. Quantization of the ordinal arrays

The ordinal patterns (permutations) are defined as in the PE from a given time series $\{x_t\}$. In order to avoid misinterpretations, the delay times of sample and embedding vectors should be set as $\tau = 1$ and $\tau_0 = 1$, respectively [20]. Storing permutations in memory and computational time costly. To overcome such limitations, permutations are quantized and stored as real numbers. Indeed, each permutation P_k of order n is converted into a real number $z_k^{(0)}$ such that

$$z_k^{(0)} = \sum_{i=1}^n a^{-\frac{i}{i+1}} P_k(i), \quad (5)$$

where $1 < a \leq 2$ is chosen to avoid infinite values of $z_k^{(0)}$ for large n . If $\{P_k\}$ is an L -periodic time series, then $\{z_k^{(0)}\}$ is also a time series of the same period. This approach gives the possibility of considering large permutation orders. For a given value of a , the maximum value of n depends on the precision of the computer. In our algorithm, we fixed $a = 2$.

By replacing the permutations by real numbers, the ordinal matrix $\mathbf{S}_{t_1}^{(1)}$ becomes a line vector $\mathbf{z}_{t_1}^{(1)}$ of real numbers. Although the relations order in the time series $\{x_t\}$ may differ from that in $\{\mathbf{S}_{t_1}^{(1)}\}$, the diversity of $\{P_k\}$ is

preserved by the transform. Similarly as for the series of permutations, each ordinal matrix $\mathbf{S}_{t_1}^{(1)}$ is replaced by a real number $z_{t_1}^{(1)}$ such that

$$z_{t_1}^{(1)} = \frac{1}{\sum_{i=1}^{d_1} a^{-\frac{i}{i+1}}} \cdot \sum_{i=1}^{d_1} a^{-\frac{i}{i+1}} \mathbf{z}_{t_1}^{(1)}(i), \quad (6)$$

where $\mathbf{z}_{t_1}^{(1)} = (z_{t_1}^{(0)}, z_{t_1+1}^{(0)}, \dots, z_{t_1+(d_1-1)}^{(0)})$. Applying once more the same process to $\{z_{t_1}^{(1)}\}$, the third dimension is reduced and $\{z_{t_2}^{(2)}\}$ is obtained. The process is to be continued until $\{z_{t_{(M-1)}}^{(M-1)}\}$, so that the whole M-D OA is converted into a scalar. The general term can then be given as:

$$z_{t_j}^{(j)} = \frac{1}{\sum_{i=1}^{d_j} a^{-\frac{i}{i+1}}} \cdot \sum_{i=1}^{d_j} a^{-\frac{i}{i+1}} \mathbf{z}_{t_j}^{(j)}(i), \quad j \geq 1, \quad (7)$$

where $\mathbf{z}_{t_j}^{(j)} = (z_{t_j}^{(j-1)}, z_{t_j+1}^{(j-1)}, \dots, z_{t_j+(d_j-1)}^{(j-1)})$, t_j is a positive integer such that $0 \leq t_j \leq T - \rho$. Repeating the procedure for the whole series of M-D OA allows to obtain a 1-D time series.

3.3. Ordinal arrays entropy

Definition 2:

The ordinal arrays entropy (OAE) is defined as the Shannon entropy of an information source whose alphabet elements are permutation-encoded OA. Let us assume $m = d_{M-1}$, $z_t = z_{t_{(M-1)}}^{(M-1)}$ and the equivalent embedding dimension s of the basic (M-1)-dimensional OA such that

$$s = n + \sum_{j=1}^{M-2} (d_j - 1). \quad (8)$$

The OAE of M-D OA is defined as

$$H_m(s) = - \sum p(z) \cdot \ln(p(z)), \quad (9)$$

where

$$p(z) = \frac{\#\{t \mid t \leq T - s - m + 1, z_t = z\}}{T - s - m + 2}. \quad (10)$$

In the particular case of 1-D OA series, the 1-D OA can be considered as an $n \times m$ OA with $m = 1$. The OAE is thus reduced to the PE, according to the similarity between Eq. (2) and Eq. (10). The OA series is different from a 1-D series iff $n > 1$ and $m > 1$. In the case of a 2-D OA series, setting $z_t = z_{t_1}^{(1)}$, and $m = d_1$, the OAE is defined as:

$$H_m(n) = - \sum p(z) \cdot \ln(p(z)), \quad (11)$$

where

$$p(z) = \frac{\#\{t \mid t \leq T - n - m + 1, z_t = z\}}{T - n - m + 2}. \quad (12)$$

The z values are numeric conversions of permutation-encoded words, in a similar way as bit-encoded symbols are usually converted into decimal or real

number values. The size of the OA alphabet being $(n!)^m$, the maximum value of $H_m(n)$ is equal to $m \ln(n!)$. Computing the Shannon entropy of OA allows to better exploit the information provided by the ordinal patterns, hence to approximate better the complexity of the original time series as the number of OA increases as a power of m .

$H_m(s)$ is the OAE obtained by considering m -length vectors derived from an initial series of quantized $(M-1)$ -dimensional OA, whose equivalent embedding dimension is equal to s .

Theorem 1:

Let $\{x_t\}$ a real-valued period- L cycle time series of length T . If $L = q$, i.e no redundant value occurs in the basic period of $\{x_t\}$, and the delay time of the embedding vectors $\tau_0 = 1$, then the OAE does not depend on the embedding dimension m of the OA.

Proof:

In the case of 1-D OA, $\{x_t\}$ can be considered as a series of $n \times m$ OA with $m = 1$ ($s = n$). For $n = 1$, the patterns are reduced to the time series values. $\{x_t\}$ being L -periodic, there are L distinct patterns, also equal to the number of source symbols, each of them appearing with the same probability $p(\theta_i) = \frac{1}{L}$, $i = 1, 2, \dots, L$, if the number of patterns explored is a whole multiple of L . The OAE in that case is equal to $H_m(s) = \ln(L)$.

For $n > 1$, we also showed in [20] that the set $\{\mathbf{x}_k\}$ of embedding vectors is L -periodic if the delay time of the embedding vectors is $\tau_0 = 1$. Indeed, as all the values in the basic period are assumed to be distinct, each embedding vector (pattern) is repeated after L steps: $\mathbf{x}_k = \mathbf{x}_{k+L}$. As the new patterns are embedding vectors $\mathbf{x}_k$, it then results that the number of distinct patterns is once more equal to the number of source symbols and all of them are equiprobable, depending on the observation length, hence $H_m(s) = \ln(L)$.

In the case of M -dimensional arrays ($M \geq 2$), it can be also verified that the period of the original time series is preserved in the series of ordinal arrays. Indeed, in that case the time series is assumed to be a series of vectors or arrays $\{\mathbf{x}_k\}$. Let us assume that $\{\mathbf{x}_k\}$ is a L -periodic series of n -length vectors ($n > 1$). The first quantized ordinal arrays series $\{z_{t_1}^{(1)}\}$ is derived from $\{\mathbf{x}_k\}$ by considering d_1 -length ($d_1 > 1$) embedding vectors $\mathbf{z}_{t_1}^{(1)} = (\mathbf{x}_{t_1}, \mathbf{x}_{t_1+1}, \dots, \mathbf{x}_{t_1+(d_1-1)})$ of n -length vectors ($s = n$, $m = d_1$). We have already shown at the beginning of the proof that for any value of $n \in \mathbb{N}_{\geq 1}$, $\{\mathbf{x}_k\}$ and $\{x_t\}$ have the same period. By the same principle, $\{z_{t_1}^{(1)}\}$, hence $\{z_{t_1}^{(1)}\}$ and $\{\mathbf{x}_k\}$ also have the same period L . $z_{t_1}^{(1)}$ is the quantized value of the vector of scalars derived from $\{\mathbf{x}_k\}$ or simply a quantized ordinal matrix (2-D arrays). The number of distinct symbols in the basic period of $\{z_{t_1}^{(1)}\}$ being equal to L , its entropy is $H_m(s) = \ln(L)$. Now let us consider the ordinal arrays series $\{z_{t_2}^{(2)}\}$ derived from $\{z_{t_1}^{(1)}\}$ and the corresponding quantized series $\{z_{t_2}^{(2)}\}$. $z_{t_2}^{(2)}$ is the quantized value of a vector of ordinal matrices (3-D arrays). The series $\{z_{t_2}^{(2)}\}$ is obtained by quantizing vectors $\mathbf{z}_{t_2}^{(2)} = (z_{t_2}^{(1)}, z_{t_2+1}^{(1)}, \dots, z_{t_2+(d_2-1)}^{(1)})$ of quantized ordinal matrices. As the series of ordinal matrices is L -periodic without redundant values, there are only L distinct possible embedding vectors of ordinal matrices which repeats

each one after L steps if $\tau_0 = 1$, hence $\{z_{t_2}^{(2)}\}$ is L -periodic and its entropy is $H_m(s) = \ln(L)$, with $s = n + d_1 - 1$ and $m = d_2$. By continuing the process, the same result is obtained for $\{z_{t_3}^{(3)}\}$ derived from $\{z_{t_2}^{(2)}\}$ and from step to step, the period of the series $\{z_{t_{(M-1)}}^{(M-1)}\}$ derived from $\{z_{t_{(M-2)}}^{(M-2)}\}$ is also equal to L . The M-D array transform is achieved after M successive transforms with embedding dimensions d_j , where $0 \leq j \leq M - 1$ refers to the step index, with $d_0 = n$. As each transform step preserves the period of the corresponding input time series, from step to step the initial period is preserved in the whole M-D transform. Given that there is no redundant value in the basic period of the time series and the delay time of the embedding vectors is $\tau_0 = 1$, the entropy of the arrays series is also equal to $H_m(s) = \ln(L)$. It can then be concluded that the OAE of a L -periodic time series without redundant values in the basic period is equal to $H_m(s) = \ln(L)$, independently of m value, which finishes the proof.

Remark: According to the above proof, it appears that $H_n(1) = H_1(n)$. $H_n(1)$ is the OAE computed from the series of scalars by considering n -length embedding vectors, while $H_1(n)$ corresponds to the OAE computed from the series of n -length vectors by considering each individual basic element (each n -length vector) as embedding vector. $H_m(s)$ is the Shannon entropy obtained by m -length embedding vectors from the initial $(M - 1)$ -dimensional OA series whose equivalent embedding dimension is s . If $m = 1$, the $(M - 1)$ -dimensional basic series remains unchanged, its dimension too. As a consequence, each Q -dimensional OA A can be considered as M -dimensional ($M > Q$), such that $A = d_1 \times d_2 \times \dots \times d_Q \times d_{Q+1} \dots \times d_M$, where $d_j > 1, \forall 1 \leq j \leq Q$ and $d_j = 1, \forall j > Q$.

4. Ordinal arrays complexity

4.1. Usefulness of M -dimensional ordinal arrays

When redundant values occur in the basic period ($L < q$), the OAE depends on their number as well as their ordering. We showed in [20] that to avoid such dependence, embedding vectors should include all the redundant values and at least one non-redundant value. It then results in the case of 1-D time series that the embedding dimension $d_0 = n$ of the scalars should be chosen such that all the patterns are equiprobable. The equiprobability property can be observed for $n > R$ where R is the number of redundant values. In that case, the OAE is equal to $H_1(n) = \ln(q)$. Let us consider the following 3-periodic time series $\{x_t\} = \{a_0, a_1, a_0, a_1, a_2, a_0, a_1, a_0, a_1, a_2, a_0, a_1, a_0, a_1, a_2, \dots\}$ for example. $\{x_t\}$ presents a time period $q = 5$ and is obtained by periodically repeating its basic period $x^0 = \{a_0, a_1, a_0, a_1, a_2\}$. This basic period presents $R = 2$ redundant values (one redundant value a_0 and one redundant value a_1). If $n = 2$, the equiprobability cannot be obtained as the probability distribution is: $p_{(a_0, a_1)} = \frac{2}{5}$; $p_{(a_1, a_0)} = \frac{1}{5}$; $p_{(a_1, a_2)} = \frac{1}{5}$ and $p_{(a_2, a_0)} = \frac{1}{5}$. In that case, the OAE depends on m . For $n = 3$, the following equiprobability distribution is observed: $p_{(a_0, a_1, a_0)} = \frac{1}{5}$; $p_{(a_1, a_0, a_1)} = \frac{1}{5}$; $p_{(a_0, a_1, a_2)} = \frac{1}{5}$; $p_{(a_1, a_2, a_0)} = \frac{1}{5}$ and $p_{(a_2, a_0, a_1)} = \frac{1}{5}$. It can be easily verified that the same equiprobability property is obtained for all $n > 3$. This example thus confirms that choosing $n > R$ leads to an equiprobability distribution and that the OAE equals $H_1(n) = \ln(5)$. In

the case where redundancies are not reduced, the OAE may take different values, depending to the embedding dimension, although the time series is periodic.

Indeed, for some periodic time series, the number R of redundancies may be very large. Therefore, 2-D symbols can be considered for reducing their effect. Thus, there are two trivial solutions with two combined embedding dimensions n and $d_1 = m$ ($n > 1$ and $m > 1$): the first one is to consider $n > R$ and any small value of m ; and the second one is to consider large values of $m > R$ and a small value of n . Due to practical limitations, considering large values of embedding dimensions ($n = 20$ for example) is not interesting. It then results that even considering 2-D symbols is not enough as we may require large values for the first and second dimensions n and m . Another simple possibility is to increase the dimension of the ordinal pattern so that the redundancy is suppressed in the final series of ordinal patterns from which the entropy is evaluated. Let us for example consider an L -periodic time series with 16 consecutive identical values per period. The corresponding series of 9-order permutations ($n = 9$) is also L -periodic with 8 consecutive identity permutations per period. For such a redundancy to be canceled, we need to choose at least $n = 16$ or to convert the series of permutations into a 2-D time series with at least $d_1 = 8$. Let us consider $d_1 = 3$, then the resulting 2-D time series is periodic with 6 consecutive identical matrices per period. As the redundancy persists, a third dimension is required. Let us consider $d_2 = 4$ for the third dimension. Then the series of ordinal arrays obtained is periodic with 3 consecutive values per period. The process then needs to be continued until the redundancy is resorbed. A last dimension can then be considered with $d_3 = 3$. Finally, the series with no redundant value is a set of $S = n \times D$ arrays, with $n = d_0$ and $D = d_1 \times d_2 \times d_3$ array. The possible number of distinct OA is then equal to $(n!)^{d_1 \cdot d_2 \cdot d_3}$. Replacing each permutation by a real number reduces the dimension of the $n \times D$ OA by one. So, each M-D OA can be reduced into a one dimensional vector by the successive quantization as above described, see Eqs. (5) and (6).

Considering M-D OA is equivalent to decomposing a ρ -length embedding vector into an $n \times D \times m$ embedding array, where $\rho = s + m - 1$ according to Eqs. (3) and (8), and $D = d_1 \times d_2 \times \dots \times d_{M-2}$. Thus, this approach contributes to improve the statistical analysis of the time series by creating new criteria for discerning between patterns which initially look similar. As we showed previously in [19, 20], even when redundancies occur in the basic period of the time series ($L < q$), considering embedding dimensions $n > q$ is enough to get a q -periodic time series of the embedding vectors, hence an OAE which does not depend on m for regular dynamics.

Consequence:

According to theorem 1, the OAE of periodic dynamics does not depend on the embedding dimension, if the latter is chosen such that all the patterns are equiprobable. Thus, we suggest to use the difference between two OAE computed with two suitably chosen embedding dimensions as a complexity measure. By this approach, periodic dynamics are detected with a zero complexity, provided that equivalent embedding dimensions are chosen such to avoid any redundancy in the resulting series of OA.

4.2. Ordinal arrays complexity

Definition 3:

We propose to define the ordinal arrays complexity (OAC) as:

$$h_{m_1, m_2}(s) = \frac{H_{m_2}(s) - H_{m_1}(s)}{m_2 - m_1}, \quad (13)$$

where m_1 and m_2 are two embedding dimensions applied to the series of $S = n \times D$ OA, with $1 \leq m_1 < m_2$. The Shannon entropy $H_m(s)$ of the OA of size $S \times m$ is computed from the time series of size S OA. The corresponding embedding dimension for the whole process is $\rho = s + m - 1$. The OAC is simply the differential of the OAE and measures the diversity of OA as the embedding dimension m of $(M - 1)$ -dimensional OA increases.

The OAC is equal to zero for periodic dynamics if $s > q$. In the case of non-regular dynamics, considering two different embedding dimensions leads to two different distributions of probability, hence to $h_{m_1, m_2}(s) > 0$. The number of possible distinct $S = n \times d_1 \times \cdots \times d_{M-2}$ OA which can be derived from a basic time series containing no redundant value is equal to $(n!)^\Gamma$, where Γ is such that

$$\Gamma = \prod_{j=1}^{M-2} d_j. \quad (14)$$

$D = d_1 \times d_2 \times \cdots \times d_{M-2}$ is the array of permutations. For a given time series converted into a series of size S OA, the possible number of $S \times m$ OA is equal to $(n!)^{\Gamma \cdot m}$. It then results that $0 \leq H_m(s) \leq \Gamma \cdot m \ln(n!)$, which also implies that $H_{m_1}(s) < H_{m_2}(s)$ for $m_1 < m_2$, hence $0 \leq h_{m_1, m_2}(s) \leq \Gamma \cdot \ln(n!)$. This number expresses the large diversity of OA which can be derived from the time series, whereas only $n! \ll (n!)^{\Gamma \cdot m}$ ordinal patterns can be described by the PE.

Consequence:

Considering the equivalent embedding dimension $\rho = s + m - 1$, the time series length should be such that $T \gg \rho!$. Similarly, considering the number of possible OA suggests that $T \gg (n!)^{\Gamma \cdot m}$. For $\Gamma > 1$ and $m > 1$, it is obvious that $\rho! \ll (n!)^{\Gamma \cdot m}$. Considering the gap between the smallest value of the required time series length $(n!)^{\Gamma \cdot m}$ and the smallest time series length $\rho!$ due to the equivalent embedding dimension ρ , it comes that the OAC may give satisfactory results even for data lengths smaller than the required one. Indeed, although the number of possible patterns is increased as a power of the basic number of ordinal patterns, the corresponding embedding dimension just slightly increases. For a given time series of length T , the length of the resulting $S = n \times D$ ($D = d_1 \times d_2 \times \cdots \times d_{M-2}$) series of OA is equal to $N = T - (s - 1)$. Let us consider $T = 100$, $n = 4$, and $D = 3 \times 3 \times 5$, the equivalent embedding dimension is $s = 12$ while the length of the OA time series is $N = 89$. The possible number of distinct OA is $(24)^{45} = 1.2868 \cdot 10^{62}$, which is too large as compared to $s! = 479001600$. It results from this observation that by the OA approach, given a time series with fixed length, the number of patterns can be significantly increased, therefore improving the statistical analysis. Large embedding dimensions are necessary for detecting regular dynamics with large periods, but conversely require infinite data length for estimating the complexity of non-regular dynamics. So there is a compromise between the choice of the embedding dimension and the time series length, which appears to be well balanced by the OAC algorithm. The above OAC in **definition 3** thus appears as an improvement of the ordinal KS algorithm [19].

4.3. Relationship between the ordinal arrays complexity and the Kolmogorov Sinai entropy

Let us define the parametric function $f_m : [0, \Gamma \ln(n!)]$ by:

$$f_m(s) = h_{m,m+1}(s), m \in \mathbb{N}_{\geq 1}. \quad (15)$$

Considering the definition of the ordinal representation of the KS entropy given in [19], the relationship between the KS entropy and the OAC can be established as:

$$h_{KS} = \lim_{s \rightarrow \infty} \lim_{m \rightarrow \infty} f_m(s). \quad (16)$$

Eq. (16) can be seen as the generalized approximation of the ordinal representation of the KS entropy. For 2-D arrays ($r = 0$ and $s = n$), it is equivalent to:

$$h_{KS} = \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} f_m(n) \quad (17)$$

which is the relation defined in [19].

The CPE $h_c(n)$ of order n was defined as the first element of the sequence $f_m(n)$, hence $h_c(n) = f_1(n) = h_{1,2}(n)$. Similarly, $f_1(s)$ can be seen as the conditional entropy of OA of size $S = n \times D$. The CPE characterizes the diversity of successors of the given permutations ($\rho = n$) while the conditional entropy of OA characterizes the diversity of successors of given OA ($\rho > n$).

Let us define the function family g_ρ of OAC of order ρ such that

$$g_\rho = \{f_m(s) | \rho = m + s - 1, 2 \leq s \leq \rho\}. \quad (18)$$

g_ρ is the set of parametric functions having the same embedding dimension ρ and whose elements are $g_\rho(s)$. $g_\rho(\rho)$ is the last element of the family g_ρ and is equivalent to $f_1(\rho)$, thus corresponds to the CPE of order ρ . We agree to call g_ρ the OAC of order ρ and its number of elements depends on the dimension of the OA.

5. Results and discussion

The CPE has been presented in [19] as the best approximation of the KS entropy for periodic signals, as compared to other existing ordinal patterns based algorithms. In this section, we compare its performance with the OAC algorithm. Simulations are made only for 2-D ($n \times m$) OA for simplification purposes. We recall that the CPE of order n corresponds to $f_1(\rho)$ or $g_\rho(\rho)$.

5.1. Effectiveness of the OAC for chaos detection

The logistic map whose equation is $x_{t+1} = \mu x_t(1 - x_t)$, $t \in \mathbb{N}$, is commonly used for testing the efficiency of time series analysis algorithms due to its rich nature [21]. For almost all $\mu \in [0, 4]$ the KS entropy either coincides with the Lyapunov exponent if it is positive or is equal to zero otherwise [22, 23]. It is also shown that the Lyapunov exponent for the logistic map can be estimated rather accurately [24]. Applying the OAC algorithm to the logistic map gives satisfactory results, as shown in Fig. 1. The OAC is compared to the CPE and the Lyapunov exponent (λ_{Lyap}). We computed $f_7(9)$ for the OAC and $f_1(9)$ for the CPE, with $T = 4 \cdot 10^6$ as time series length. It is observed that for the

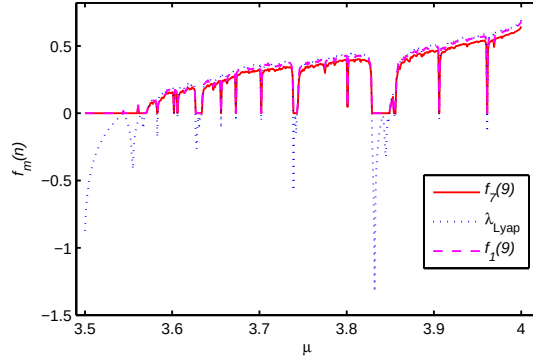


Figure 1: Comparison of the OAC with the CPE and the Lyapunov exponent, in the case of the logistic map.

whole spectrum, $f_7(9)$ values are close to the Lyapunov exponent, taking zero values for all the regular dynamics corresponding to negative values of λ_{Lyap} . The fact that the Lyapunov exponent remains slightly greater than $f_7(9)$ attests that the OAC uniformly tends to the KS limit as T tends to infinity. Conversely the CPE is greater than the Lyapunov exponent for some values of μ . It can also be observed that $f_1(9) = 0.063$ for $\mu = 3.602$, while it should be zero as the corresponding dynamics is regular, according to the Lyapunov exponent value $\lambda = -0.0085$.

5.2. Dependence of the CPE and the OAC on the data length T and the permutation order n

We theoretically established that the OAC algorithm depends less on the observation time than the CPE. In Fig. 2 the behaviors of the OAC and the CPE for various values of T are plotted. The CPE values are computed from $f_1(13)$, while those of the OAC are obtained from $f_9(5)$. The time series length is incremented from $T = 10000$ to $T = 500000$, by step size of $\Delta T = 10000$. As it can be observed from this figure, the result of the OAC does not depend

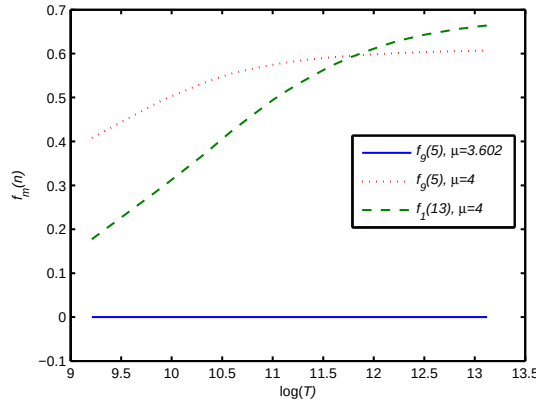


Figure 2: Dependence of the OAC and CPE results on the data length.

on the time series length in the case of regular dynamics. We verified the same

result for the CPE. However, for $\mu = 4$ corresponding to a chaotic motion, the OAC is uniformly increasing from 0.4080 up to 0.6068, corresponding to $\Delta f_9(5) = 0.1988$; while the CPE is uniformly increasing from 0.1772 up to 0.6642, corresponding to $\Delta f_1(13) = 0.4869$. Therefore, for the same embedding dimension $\rho = n + m - 1 = 13$, the OAC remains closer to its limit value, even for $T \ll (5!)^9$, than the CPE which strongly requires that the condition $T \gg 13!$ is satisfied. Using OA as patterns thus helps to increase the number of distinct futures to be analyzed, hence to improve the statistical analysis. This observation is important as it confirms that the OAC algorithm may be suitable for the analysis of short length time series. Conversely, the CPE converges faster to the KS entropy than the OAC. This result may be due to the fact that the CPE combines two OAE computed from two time series with different dimensions: $H_1(\rho)$ computed from a 1-D time series and $H_2(\rho)$ computed from a 2-D time series.

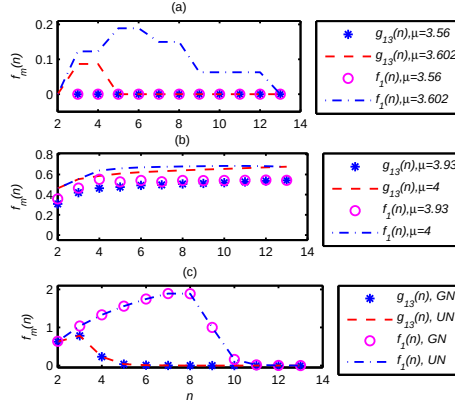


Figure 3: Relationship between the embedding dimension and the data length in the case of (a) regular dynamics, (b) non-regular dynamics and (c) noise.

We further investigated the impact of the embedding dimension on the complexity measure, but for $T = 10^6$. We have chosen three types of dynamics: regular dynamics ($\mu = 3.56$ and $\mu = 3.602$); chaotic dynamics ($\mu = 3.93$ and $\mu = 4$) and stochastic dynamics with uniform distribution (UN) obtained from the Matlab *rand* function, and Gaussian distribution (GN) obtained from the Matlab *randn* function. The CPE has been evaluated as $f_1(n)$ and the OAC as $f_m(n)$, where $m = 14 - n$ and $2 \leq n \leq 13$. The corresponding result is shown in Fig. 3. This figure shows that for regular dynamics, the CPE requires large embedding dimensions, independently of the data length; while the OAC rapidly converges to zero, thus allowing the detection of regular dynamics with better accuracy than the CPE.

In the case of chaotic dynamics, $f_1(n)$ for the CPE increases with the embedding dimension. Similarly, $g_{13}(n)$ for the OAC uniformly increases with the permutations order n . In that case, although the embedding dimension $\rho = 13$ remains constant, for $n > 2$ the required data length $(n!)^m$ becomes too large as compared to T . Observing an increment of the complexity for the OAC just confirms that the pattern distribution strongly depends on m so that $f_m(n)$ is

different from zero even when the time series length is too short as compared to the required data length.

In the case of noise, it is observed that $f_1(n)$ is the same for both the uniform noise (UN) and the Gaussian noise (GN). Similarly, $f_m(n)$ is the same for the two distributions. However, the CPE increases with n for T such that $T \gg n!$, and decreases until zero when this condition is no longer satisfied. The same observation is made for the OAC which increases as $T \gg (n!)^m$ and decreases to zero otherwise. These two results imply that the number of patterns rapidly increases in the case of noise and does not change anymore when the data length is smaller than the required value, thus leading to zero values for the complexity.

5.3. Impact of the equivalent embedding dimension ρ

For some values of the equivalent embedding dimension ρ , the detection result for regular dynamics may be wrong. An example is given for the logistic map with $r = 3.602$. For this dynamics, both the CPE and the OAC are giving a nonzero complexity for $\rho < 13$, independently of the time series length. Conversely, increasing the embedding dimension will impose to increase also the time series length, otherwise the complexity will reduce. In practice, the data length sometimes cannot be extended and an appropriate compromise between the data length and the embedding dimension is required to avoid any confusion between regular and non-regular dynamics.

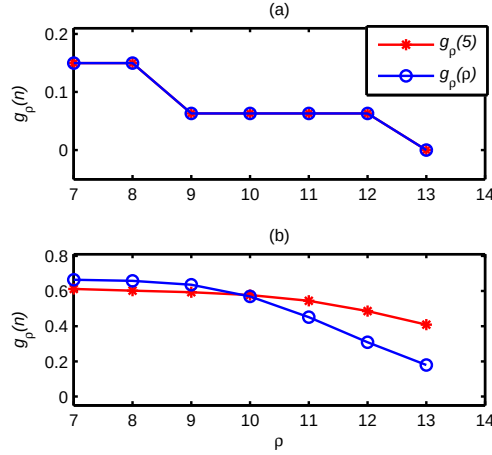


Figure 4: Impact of the embedding dimension on the OAC and the CPE algorithms: (a) regular dynamics and (b) non-regular dynamics.

In Fig. 4 the performances of the CPE and OAC algorithms are compared for a given data length ($T = 10000$ samples), with varying equivalent embedding dimension ($7 \leq \rho \leq 13$). We fixed the permutation order to $n = 5$ for the OAC algorithm. The dynamics to be analyzed is generated from the logistic map with $r = 3.602$ corresponding to a regular dynamics, and $r = 4$ corresponding to a chaotic dynamics. According to the results obtained, it appears that the OAC well balances the compromise between the embedding dimension and the data length. Indeed, the data length should be chosen such that $T \gg \rho!$. This requirement is not respected for both the OAC and the CPE.

However, one observes in the case of the chaotic dynamics that the value of the complexity by the CPE ($g_\rho(\rho)$) drastically reduces from $g_7(7) = 0.6643$ to $g_{13}(13) = 0.1788$ when $T < \rho!$; while the OAC result ($g_\rho(5)$) slightly decreases from $g_7(5) = 0.6114$ to $g_{13}(5) = 0.4091$. In the case of regular dynamics, the two algorithms behave similarly, although it should be pointed out that there is a smallest value of n ($n = 5$ for the dynamics corresponding to $r = 3.602$ as shown in Fig. 3(a)) under which the complexity of the regular dynamics is no longer zero, even for large values of ρ . It can then be concluded that the required data length for the OAC algorithm strongly depends on the permutation order n than the equivalent embedding dimension. Hence, choosing $T \gg n!$ is enough for the OAC algorithm, while the CPE algorithm requires $T \gg \rho!$. Moreover, the permutation order required for detecting regular dynamics with a zero complexity is significantly reduced in the OAC algorithm, as compared to the CPE algorithm. Although the required embedding dimension ρ is the same for both the CPE and the OAC, it appears that its decomposition by the OA approach contributes to reduce the permutation order required by the CPE for efficient complexity measure, hence to reduce the required data length. These observations thus attest that the OAC algorithm is more suitable for discerning between regular and non-regular dynamics from a small data set than the CPE algorithm.

5.4. Robustness against noise

According to the results in Fig. 3, the OAC is assumed to be more robust against noise than the CPE. Indeed, the complexity $g_\rho(n)$ of a random distribution uniformly tends to zero as soon as $(n!)^m \gg T$, where $m = \rho - n + 1$ and T is the data length. In Fig. 5 the spectrum of the logistic map contaminated with a Gaussian observational noise is shown. The noisy time series is obtained as $\phi_t = x_t + \sigma\eta_t$, where x_t is given by the logistic equation. Two noise levels $\sigma = 0.001$ and $\sigma = 0.01$ were considered. The Gaussian noise η_t with standard deviation $\sigma_0 = 1$ and mean value $\bar{\eta} = 0$ is obtained by using the Matlab *randn* function. As it can be seen, the OAC effectively performs well in presence of noise. Particularly, periodic windows are better detected than with the CPE. It is also noticed that the noise effect increases with the permutation order n .

5.5. Case of periodic dynamics with large period

We discussed the case of periodic dynamical systems with small period and we verified that there is a lower limit for the permutation order n under which the dynamics cannot be detected as regular. In practice, the period is not prior known and the embedding dimension ρ or the permutation order n in some cases may be badly chosen. Due to the limited data length, it is also common to choose small embedding dimensions. It is well known that the KS entropy of a periodic dynamical system is equal to zero. However, the CPE as well as the OAC in that case may output nonzero values. Such a result is illustrated in Fig. 6. We considered for this example the rotation map $u_\alpha(\omega) = (\omega + \alpha) \bmod 1$ on the interval $[0, 1)$. When α is a rational number, u_α provides a periodic behavior [19]. We used $T = 10000$ as data length and α varying from 0 to 0.5, with step size $\Delta\alpha = 0.001$. As shown on Fig. 6, all the three functions of the families g_8 and g_{13} behave similarly. We verified that g_8 and g_{13} go to 0 at rational numbers whose denominators are smaller than or equal to 8 and 13

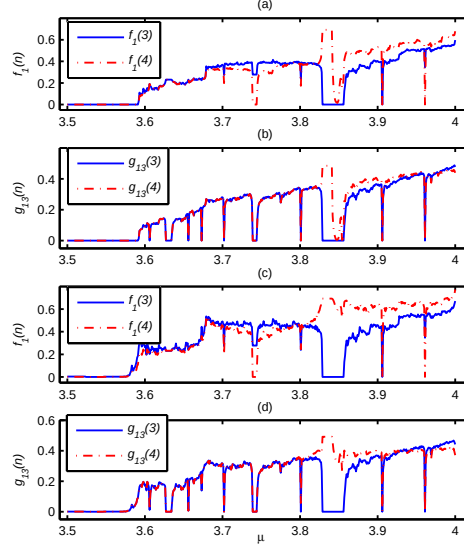


Figure 5: Spectra of the CPE and the OAC for the logistic map in presence of Gaussian observational noise, $T = 10000$. The noise level is (a)-(b) $\sigma = 0.001$ and (c)-(d) $\sigma = 0.01$.

respectively. For the set of α values considered in Fig. 6, g_8 goes to zero for $\alpha \in \{0, \frac{1}{8}, \frac{1}{5}, \frac{1}{4}, \frac{3}{8}, \frac{2}{5}, \frac{1}{2}\}$, while g_{13} goes to zero for $\alpha \in \{0, \frac{1}{10}, \frac{1}{8}, \frac{1}{5}, \frac{1}{4}, \frac{3}{10}, \frac{3}{8}, \frac{2}{5}, \frac{1}{2}\}$. It is also observed that $g_{13} < g_8$, which attests that the complexity converges to zero as the embedding dimension tends to infinity. The result might become zero by increasing the permutation order n and the embedding dimension ρ . Unfortunately, a significant increment of these two parameters is not possible due to the limited data length. Another important observation can be made: for periodic dynamical systems g_ρ depends only on the embedding dimension ρ for $n \ll L$, and therefore slowly converges to zero. We verified this observation for all the elements of the families g_8 and g_{13} . Taking into account this observation, it can be conjectured that the difference of two parametric functions $f_m(s)$ belonging to the same family g_ρ is equal to zero; which allows to easily detect regular dynamics with large periods.

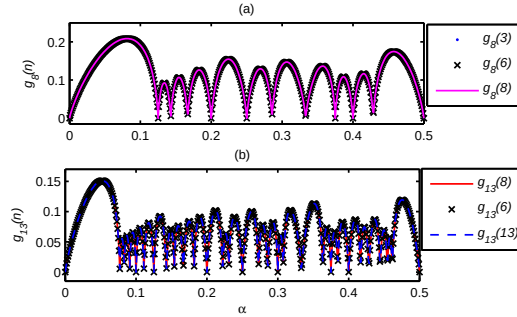


Figure 6: $g_\rho(n)$ -spectrum of the rotation map, $T = 10000$. The embedding dimensions are respectively (a) $\rho = 8$ and (b) $\rho = 13$.

5.6. Detection of the quasi-periodic route to chaos

The interest for this type of dynamical system is to confirm the non-dependence of the behavior of g_ρ on the permutation order n in the case of regular dynamics, as observed in the previous subsection. For this purpose, we have considered the sine circle map defined as $\theta_{t+1} = \left(\Omega + \theta_t + \frac{\gamma}{2\pi} \sin(2\pi\theta_t) \right) \bmod 1$, $t \in \mathbb{N}$. Ω is the frequency ratio parameter and γ the nonlinearity parameter. For $\Omega = \frac{-1+\sqrt{5}}{2}$, the system is known to exhibit quasi-periodic dynamics for $0 \leq \gamma \leq 1$ [25, 26]. As in the case of the rotation map, we considered sequences of length $T = 10000$ samples and the control parameter $0 \leq \gamma \leq 2.5$ varying by step size $\Delta\gamma = 0.01$. We also choose $\theta_0 = 0.5$ as initial condition. The corresponding result is shown in Fig. 7 from where it is confirmed that g_ρ does not depend on n in the case of regular (periodic and quasi-periodic) dynamics.

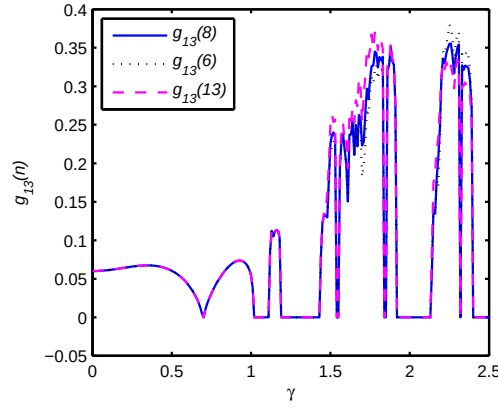


Figure 7: $g_\rho(n)$ -spectrum of the sine circle map, $T = 10000$. The embedding dimensions is $\rho = 13$.

In summary, it appears that given an equivalent embedding dimension ρ , the OAC $g_\rho(n)$ does not depend on n in the case of regular dynamics, for $n \ll L$. In the case of chaotic dynamics, $g_\rho(n)$ uniformly increases with n as long as $T \gg \rho!$. For stochastic dynamics, $g_\rho(n)$ depends on the time series length T : it uniformly increases as $T \gg (n!)^{(\rho-n+1)}$ and decreases to zero when this condition is no longer satisfied.

6. Conclusion

Large embedding dimensions are necessary for the detection of regular dynamics with zero complexity, but conversely require infinite data length for the estimation of non-regular dynamics complexity. This compromise between the embedding dimension and the data length is well balanced by the OAC algorithm, presented in this paper as a generalized approximation of the ordinal KS entropy. For fixed values of embedding dimensions ρ and finite data length T , we showed that $g_\rho(n)$ does not depend on n in the case of regular dynamics; it uniformly increases up to a maximum value $g_\rho(\rho)$ corresponding to the CPE for chaotic dynamics; while it increases fast up to a maximum value and rapidly decreases to zero for the noise. We also verified that the OAC algorithm is more

robust against noise than the CPE algorithm. Further investigations will be made for determining properties of the algorithm for more than 2-D OA and for applying it for distinguishing between deterministic and stochastic data.

Acknowledgements

This work is supported by the Alexander von Humboldt Foundation.

References

- [1] G. Choe, Computational Ergodic Theory, Springer, Berlin, 2005.
- [2] H. D. I. Abarbanel, Analysis of Observed Chaotic Data, Springer, New York, 1996.
- [3] J.-P. Eckmann, D. Ruelle, Ergodic theory of chaos, Rev. Mod. Phys. 57 (1985) 617.
- [4] P. Grassberger, I. Procaccia, Estimation of the kolmogorov entropy from a chaotic signal, Phys. Rev. A 28 (1983) 2591–2593.
- [5] M. Ding, C. Grebogi, E. Ott, T. Sauer, J. A. Yorke, Estimating correlation dimension from a chaotic time series: when does it occur?, Phys. Rev. Lett. 70 (1993) 3872.
- [6] J. B. Gao, J. Hu, W. W. Tung, Y. H. Cao, Distinguishing chaos from noise by scale-dependent Lyapunov exponent, Phys. Rev. E. 74 (2006) 066204.
- [7] S. M. Pincus, Approximate entropy as a measure of system complexity, Proc. Natl. Acad. Sci. 88 (1991) 2297–2301.
- [8] O. Afsar, G. B. Bagei, U. Tirnakli, Renormalized entropy for one dimensional discrete map: Periodic and quasi-periodic route to chaos and their robustness, Eur. Phys. J. B 86 (2013) 307–320.
- [9] J. Richman, J. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol-Heart C. 278 (2000) H2039–H2049.
- [10] C. Bandt, B. Pompe, Permutation entropy: A natural complexity measure for time series, Physical review letters 88 (2002) 174102.
- [11] J. M. Amigó, Permutation complexity in dynamical systems, Springer, 2010.
- [12] C. Bian, C. Qin, Q. D. Y. Ma, Q. Shen, Modified permutation-entropy analysis of heartbeat dynamics, Phys. Rev. E 85 (2012) 021906.
- [13] B. Fadlallah, J. Principe, B. Chen, A. Keil, Weighted-permutation entropy: An improved complexity measure for time series, Phys. Rev. E 87 (2013) 022911.
- [14] Z. Li, G. Ouyang, D. Li, X. Li, Characterization of the causality between spike trains with permutation conditional mutual information, Phys. Rev. E 84 (2011) 021929.

- [15] X. Li, S. Cui, L. Voss, Using permutation entropy to measure the electroencephalographic effects of sevoflurane, *Anesthesiology* 109 (2008) 448–456.
- [16] X. Li, G. Ouyang, D. Richards, Predictability analysis of absence seizures with permutation entropy, *Epilepsy Res.* 77 (2007) 70–74.
- [17] A. Bruzzo, B. Gesierich, M. Santi, C. Tassinari, N. Bir-baumer, G. Rubboli, Permutation entropy to detect vigilance changes and preictal states from scalp EEG in epileptic patients: a preliminary study, *Neurol. Sci.* 29 (2008) 3–9.
- [18] Y. Cao, W. Tung, J. Gao, V. Protopopescu, L. Hively, Detecting dynamical changes in time series using the permutation entropy, *Phys. Rev. E* 70 (2004) 046217.
- [19] V. A. Unakafov, K. Keller, Conditional entropy of ordinal patterns, *Physica D* 269 (2014) 94–102.
- [20] 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.
- [21] P. Collet, J.-P. Eckmann, *Iterated Maps on the Interval as Dynamical Systems*, Birkhäuser, Boston, 1980.
- [22] L.-S. Young, *Entropy in Dynamical Systems*, Princeton University Press, Princeton, 2003.
- [23] M. Martens, T. Nowicki, Invariant measures for lebesgue typical quadratic maps, *Astérisque* 261 (2000) 239–252.
- [24] J. Sprott, *Chaos and Time-Series Analysis*, Oxford University Press, Oxford, 2003.
- [25] J. S. A. E. Fouda, W. Koepf, Efficient detection of the quasi-periodic route to chaos by the three-state test, *Nonlinear Dyn.* 78 (2014) 14771487.
- [26] O. Afsar, G. Bagei, U. Tirnakli, Renormalized entropy for one dimensional discrete map: Periodic and quasiperiodic route to chaos and their robustness, *Eur. Phys. J.* 86 (2013) 307320.