Eur. Phys. J. Special Topics **222**, 249–262 (2013)

DOI: 10.1140/epjst/e2013-01862-7

THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS

Review

Practical considerations of permutation entropy

A tutorial review

M. Riedl[a], A. Müller[b], and N. Wessel

Department of Physics, Humboldt-Universität zu Berlin, Berlin, Germany

Received 27 March 2013 / Received in final form 25 April 2013
Published online 25 June 2013

Abstract. More than ten years ago Bandt and Pompe introduced a new measure to quantify complexity in measured time series. During these ten years, this measure has been modified and extended. In this review we will give a brief introduction to permutation entropy, explore the different fields of utilization where permutation entropy has been applied and provide a guide on how to choose appropriate parameters for different applications of permutation entropy.

1 Introduction

Quantifying the complexity of a given time series is an important challenge in data analysis giving us a deeper insight into the mechanisms governing the regarded processes. Usually this is done employing some measure of entropy, e.g. Shannon entropy [1], Kolmogorov-Sinai entropy [2] or Approximation entropy [3]. Most of the classical methods ignore the temporal order of the values in a time series, e.g. the classical Shannon entropy [1], and require very long time series while being computationally expensive, e.g. Kolmogorov-Sinai entropy [2]. In 2002 Bandt and Pompe [4] combined the concepts of entropy and symbolic dynamics to create a new measure of complexity, termed permutation entropy (PE), which was able to overcome the problems mentioned above while still being simple and robust. Since then, more than ten years have passed and the PE has been established in several new fields of application, extensions have been developed and new concepts based on PE have emerged. There are already excellent reviews about the application of PE in biomedical and econophysic questions [5] and about symbolic dynamics in general [6]. A detailed theoretical background can be found in Amigó [7] for example. This review follows a more heuristic approach and aims at scientists who might not be familiar with the concept of PE but want to be able to put PE to a fast yet sophisticated use. In principle, the application of PE in data analysis can be reduced to a comparison of values for different time series to compare the complexity of the underlying processes. Unfortunately, the values of PE depend on the chosen parameters, thus

[a] e-mail: maik.riedl@physik.hu-berlin.de
[b] e-mail: andreas.mueller@physik.hu-berlin.de

rapidly leading to a multiple testing problem when regarding the multiple parameter pairs. Therefore, the two most important questions for beginners when applying PE are:

1. What are the optimal parameters for calculating PE?
2. How can the parameters be interpreted and what is their influence on the values of PE?

To avoid a complicated theoretical consideration of the first question, we performed an exhaustive literature research summarizing and classifying the experiences of experts when applying PE. We provide an overview of applications of PE in theoretical as well as in practical tasks and a summarization about parameter selection and interpretation in relation to the problem. Based on that, we formulate a recommendation about the use of PE for beginners. This work also shows that the answer to the second question is strongly related to the first one.

2 Definition of permutation entropy

When quantifying complexity for a given time series $X = \{x_i\}_{i=1...N}$, entropy measures (cf. e.g. [1]) are often the first choice, where classical methods like Fourier transformation fail. But, common approaches neglect any effects stemming from the temporal order of the values in the successive x_i. To take into account this causal information, the time series can be encoded into sequences of symbols, based on the theory of symbolic dynamics (see e.g. [8–14]). Combining these two approaches Bandt and Pompe [4] proposed a natural encoding which reflects the rank order of successive x_i in sequences of length n and thus defining permutation entropy by

$$H_n = -\sum_{j=1}^{n!} p'_j log_2(p'_j). \tag{1}$$

The p'_j represent the relative frequencies of the possible patterns of symbol sequences, termed permutations (see Fig. 1). The permutation entropy per symbol is given by

$$h_n = -1/(n-1) \sum_{j=1}^{n!} p'_j log_2(p'_j), \tag{2}$$

in order to able to compare permutation entropies with different n. Beside the normalization using $n-1$ there are also other approaches like e.g. normalization with $log_2(n!)$ [15] to get $0 \leq h_n \leq 1$.

3 Calculation of permutation entropy

PE (see Eqs. (1) and (2)) essentially measures information based on the occurrence or absence of certain permutation patterns of the ranks of values in a time series. To compute the PE for a given time series $\{x_i\}$ of length N, there are seven steps to be followed [4]:

1. Define the order of permutation n. That leads to the possible permutation patterns π_j $(j = 1, .., n!)$ which are build from the numbers $1, ..., n$. Their dynamical representation is shown in Fig. 1 for $n = 3$.

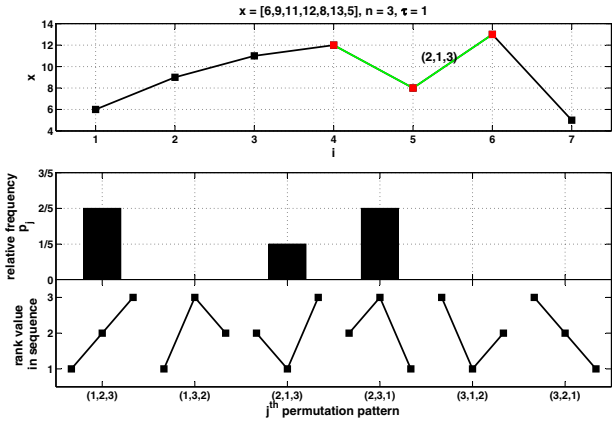

Fig. 1. Example of the encoding for the calculation of the permutation entropy. (top) Time series with exemplarily colour coded sequence of three successive values which is represented by the permutation pattern (2,1,3). (middle) The relative frequencies of all possible permutation patterns for $n = 3$ (bottom) for this particular time series.

2. Initialize $i = 1$ as the index of the considered time series $\{x_i\}_{i=1,...,N}$ and the counter $z_j = 0$ for each π_j.
3. Calculate the ranks of the values in the sequence $x_i, ..., x_{i+n-1}$ which leads to the rank sequence $r_i, ..., r_{i+n-1}$. The ranks are the indices of the values in ascending sorted order.
4. Compare the rank sequence of step 3 with all permutations pattern and increase the counter of the equal pattern $\pi_k = r_i, ..., r_{i+n-1}$ by one ($z_k = z_k + 1$).
5. If $i \leq N - n$ then increase i by one ($i = i + 1$) and start from step 3 again. If $i > N - n$ go to the next step.
6. Calculate the relative frequency of all permutations π_j by means of $p'_j = z_j / \sum z_k$ as an estimation of their probability p_j.
7. Select all values of p'_j greater than 0 (since empty symbol classes yield $0log0 = 0$) and calculate the permutation entropy by Eqs. (1) and (2).

In the following example the algorithm is illustrated. Consider the time series $x = 6, 9, 11, 12, 8, 13, 5$ with $N = 7$.

Step 1: The permutation order is set to $n = 3$. Therefore $n! = 6$ permutation are possible with $\pi_1 = 1, 2, 3$, $\pi_2 = 1, 3, 2$, $\pi_3 = 2, 1, 3$, $\pi_4 = 2, 3, 1$, $\pi_5 = 3, 1, 2$, and $\pi_6 = 3, 2, 1$.
Step 2: Initialize $i = 1$ and $z_{j=1,...,6} = 0$.
Step 3: The rank sequence of the selected values $6, 9, 11$ is $1, 2, 3$.
Step 4: It is equal to π_1, therefore z_1 is increased to 1.
Step 5 and *Step 3:* Because $i = 1 < 7 - 3$, the next values $9, 11, 12$ are selected which have the rank sequence $1, 2, 3$.
Step 4: It is again equal to π_1, therefore z_1 is increased to 2. The loop between *Step 3* and *Step 5* is then passed through three more times, which leads in the end to $z_1 = 2$, $z_2 = 0$, $z_3 = 1$, $z_4 = 2$, $z_5 = 0$, and $z_6 = 0$.
Step 6: The values of the counters are divided by $sum = 5$ which leads to $p'_1 = 2/5$, $p'_2 = 0$, $p'_3 = 1/5$, $p'_4 = 2/5$, $p'_5 = 0$, and $p'_6 = 0$.
Step 7: On the basis of the non-zero p'_j, the permutation entropy of order 3 is $H_3 = -(2/5log_2(2/5) + 1/5log_2(1/5) + 2/5log_2(2/5)) \approx 1.5$ (see Eq. (1)) and the permutation entropy per symbol is $h_3 = H_3/2 \approx 0.76$ (see Eq. (2)).

To handle equal values in a time series, there are three strategies:

1. The ranks of these values are determined in accordance to their order in the sequence.
2. The identity is eliminated by adding white noise with the strength of the stochastic term being smaller than the smallest distance between values.
3. The values get the same rank number within the regarded sequence, for example $3, 6, 2, 3, 8$ leads to the rank sequence $2, 4, 1, 2, 5$ [16].

The last modification should be preferred if there is a high frequency of equal values which might e.g. be caused by low digitalization. Unfortunately, the number of possible "permutations" is then greater than $n!$ [16]. Another reason for equal values is oversampling where states are measured in multiple successive repetitions [6]. The common approach to overcome this redundancy is to lengthen the time interval between the values of the considered sequence $x_i, ..., x_{i+n-1}$ (*Step 3*) and the step size $i = i + 1$ (*Step 5*) by means of $x_i, x_{i+\tau}, ..., x_{i+\tau(n-1)}$ and $i = i + \tau$, respectively. The term $x_i, x_{i+\tau}, ..., x_{i+\tau(n-1)}$ is known from the phase space reconstruction by means of time delay embedding where τ is the delay time and n is the embedding dimension. In that approach the time step remains $i = i + 1$ [17]. This extension of the permutation entropy is another interpretation of the ordered patterns which describe partitions of the reconstructed phase space, that means classes of states [18].

An implementation for MATLAB (The Math Works, Inc., Natick, ME, USA) of the PE algorithm, including the different mentioned strategies to overcome the problem of equal values, can be found in the TOCSY-toolbox (`http://tocsy.pik-potsdam.de/`).

4 Choice of algorithmic parameters: Experience of experts

The first practical question of a scientist applying permutation entropy is the determination of the parameters n, determining the length of the sequences and thus the number of possible permutation patterns, and τ, describing the time delay between successive points in the symbol sequences. In some approaches the tuning of one of these parameters is a constructive element of the application of PE. Variants of this tuning are for example [5]: changing the length N of the regarded time series while keeping n and τ fixed to distinguish noise from chaos [19] and changing τ with N and n fixed to identify different time scales [20] and main period lengths [21]. In the majority of applications, PE is used only to quantify the complexity of a time series. Due to the strong dependence of PE on the parameters n and τ a consideration of several combinations of parameter values, rapidly leads to the problem of multiple testing. So, the choice of the optimal parameters depending on the field of application is crucial for the use of PE. For a first differentiation of the diverse fields of application let us follow Daw and Finney [6] who distinguish:

i) measurements of dynamical processes being observed at a fixed rate per unit time, which are best modeled by differential equations,
ii) measurements of processes that possess an inherent cycle represented by triggering events, which are more naturally described by discrete models.

The latter case comprises not only iterative maps, e.g. the logistic map, but also discrete stochastic processes like autoregressive models, where the random innovation in each step displays a constructive element of the observed dynamics, in contrast to neglectable additive disturbances in case i). Taking this classification into account we can arrange the reviewed papers utilizing PE as shown in Tables 1, 2, 3 and 4 for various fields of research, in order to simplify finding the entries related to the user's

Table 1. Theoretic systems: The table shows the application of PE theoretical systems as classified in Sect. 4 (AR – autoregressive process; MA – moving average process).

Ref.	Topic	Order	Time delay	Amount of data
	case i)			
[17]	Lorenz system	$5-7$	10	1024
[22]	bidirectionally coupled Rössler systems	$6-7$	150	2^{19}
[35]	Lang-Kobayashi equation	$4-8$	1000	$2 \cdot 10^6$
[15]	coupled Lorenz and Rössler system	$3-10$	3	$100-12000$
[36]	model of feed back coupled laser	6	≈ 150	5000
[20]	Mackey-Glass oscillator	$4-8$	$1-700$	10^6
	case ii)			
[26]	2D Henon map with additive noise	$4-7$	1	$100-8000$
[17,21,37–40]	logistic map	$2-16$	$1-2$	10^3-10^7
[29]	noise generators	5	1	$>10^6$
[41]	delayed logistic map with linear/nonlinear feedback	$3,6$	1	10^6
[42]	logistic map	$3-5$	$1-5$	$60,120,500,1000$
[38,39]	three-way Bernoulli map	6	1	$>10^6$
[31]	logistic map with coloured noise	6	1	10^3-10^5
[43]	chaotic signal (Lang-Kobayashi model) with a hidden small amplitude digital message	6	1	10000
[21]	MA(30)	4	$1-20$	1000
[28,44–46]	fractional Brownian and Gaussian noise	$3-6$	1	$100-2^{15}$
[42]	linear and nonlinear AR and MA processes of order $1-3$	$3-5$	$1-5$	$60,120,500,1000$
[47]	stochastic resonance of a Brownian particle in a sinusodially modulated double-well potential and in the Fitz-Hugh-Nagumo model	$4-6$	1	60000
[25]	geometric Brownian motions	6	$1-500$	$4673,5000,50000,500000$

Table 2. Medicine: The table shows the application of PE to normal EEG (top), EEG during anaesthesia (second), EEG during seizures (middle), local field potentials (fourth) and heart rate time series (bottom).

Ref.	Topic	Order	Time delay	Amount of data
	case ii)			
[48,49]	EEG	3, 7	1	512, 768
[50]	EEG	3	$1-2$	3000, 3840
[51–53]	EEG anaesthesia	$2-7$	$1-3$	$1000-2560$
[15,17]	EEG seizures	5	3	1024
[54]	EEG seizures	20	1	2000
[30,37,55–58]	EEG epileptic seizure	$3-7$	1	$100-64512$
[59]	local field potentials	3	1	2000
[16]	heart rate	$3-7$	$1-4$	$500-8000$
[60]	fetal heart rate	3	4	$\sim$500
[61]	heart beat-to-beat intervals of heart patients	$2-18$	$1-20$	2000, 100000

Table 3. Physical systems: The table shows the application of PE to physical systems as classified in Sect. 4.

Ref.	Topic	Order	Time delay	Amount of data
	case i)			
[23]	delayed opto-electronic oscillator (Ikeda scenario)	6	$1-200$	–
[35]	single-mode laser with moderate delayed feedback	8	1540	
[24]	coupled lasers	6	$1-1200$	20000
	case ii)			
[62,63]	zero-th mode contribution of a strong external field to the production of charged meson pairs–classical quantum transition	5	1	$>= 5000$
[64]	consecutive interdropout intervals of a laser diode with optical feedback from an external cavity	$3-5$	1	> 10000
[32]	delayed coupled semiconductor lasers encoded time interval between each pair of synchronized dropouts	8	1	~ 3000

own scientific problems. Within the tables the applications are divided according to the cases of i) and ii) except in Table 2, where only ii) is present and the items are grouped according to the field of research. Beside the references, the used parameters n and τ are displayed. Comparing Tables 1, 2, 3 and 4, the more seldom case i) is found in [15,17,20–25]. Here, one can observe a large variance in the values of the time delay employed. This results from the dependence of the delay τ on sampling rate and/or integration step size. A more detailed comparison of these references showed, that PE is used to answer three different questions. Regarding the first, to use PE for determining the complexity of the main oscillation the optimal parameters seem to be a τ-value equal to the period length and a dimension of $n > 3$. For these values the PE exhibits a local maximum. The second question concerns using PE to determine the optimal embedding parameters for a given time series. However, in this approach

Table 4. Economics and Environmental studies: The table shows the application of PE to systems as classified in Sect. 4.

Ref.	Topic	Order	Time delay	Amount of data
	case i)			
[21]	Southern oscillation index	4	$1-20$	453
[21]	Fish recruitment	4	$1-20$	453
[21]	Canadian lynx	4	$1-20$	114
[25]	20 Dow Jones UBS subindexes	$4-6$	$1-500$	4673
	case ii)			
[65]	Sedimentary data	5	1	1000
[45]	Annual minimal water levels	2	1	~ 659
[34]	World stock indices	$4-6$	1	~ 3138
[66]	Stock market indices of 32 countries	$4-5$	1	~ 3000
[67]	Daily values of bond indices	$4-6$	1	2047

the optimal n and τ strongly depend on the system. A more intuitive formulation of the optimal parameter selection will be given by means of simulations in the next section. The third question deals with the detection of a delay in a given system or a coupling between two systems, characterized by the minimum of PE. Comparing the results in the tables, higher values of n seem to allow for a better detection of the right τ. This last application works for cases i) and ii). Otherwise, when analyzing time series of the ii)-type (see the remaining references in Tables 1, 2, 3 and 4), a lag of $\tau = 1$ is commonly used. Regarding the necessary lengths of the time series for both cases, it is typically recommended to satisfy $N > 5n!$ (see e.g. [26]), which essentially leads to a maximum of $n = 7$ in most applications.

5 Interpretation of the parameters and their influence on permutation entropy

In the following, we will apply the aforementioned ideas to three well-known model systems: the Lorenz system, the van-der-Pol-oscillator and the logistic map. The Lorenz system is given by

$$\dot{x} = \sigma(y - x) \qquad \dot{y} = x(\rho - z) - y \qquad \dot{z} = xy - \beta z \tag{3}$$

and the parameters used are $\sigma = 10$, $\rho = 28$ and $\beta = 8/3$ which is well into the chaotic regime of the system. The van-der-Pol-oscillator can be described by

$$\dot{x} = \mu(x - \frac{1}{3}x^3 - y) \qquad \dot{y} = \frac{1}{\mu}x \tag{4}$$

and for $\mu = 10$ the system exhibits periodic behaviour. The logistic map is given by

$$x_{n+1} = r x_n (1 - x_n), \tag{5}$$

where r controls the dynamic behaviour, for example: periodic behaviour for $r = 3.55$, intermittency near period 3 for $r = 3.8282$ and deterministic chaos for $r = 4$. As an example of a random process, we chose an autoregressive model

$$x_i = 0.3 x_{i-9} + \epsilon_i, \tag{6}$$

where ϵ_i is a series of independently normally distributed ($N(0, 1)$) random numbers.

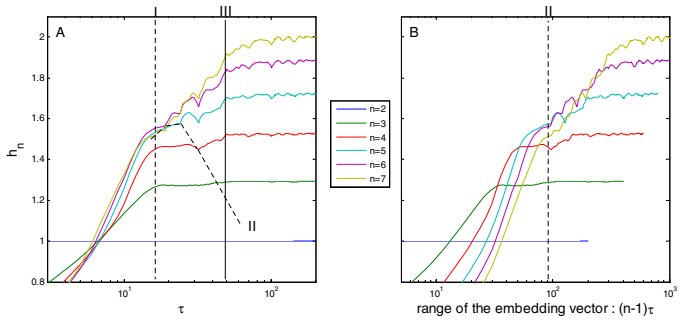

Fig. 2. (A) Normalized permutation entropy h_n of the first coordinate of the Lorenz system (integration step: 0.0170; $N = 59393$ to be well above the required data length for the maximum n) for different values of τ and n. The dotted lines *(I,II,III)* mark changes of the relation between h_n and τ and separate the range of τ into four sections. (B) The same values of h_n in relation to the range of the time delay embedding vector (embedding dimension n and delay τ). In contrast to panel (A), the marker *(II)* is a vertical line in this representation.

Considering the case i) from Sect. 4, the parameters for the calculation of PE can be interpreted either in the meaning of time delay embedding or the display of dominant periods. The crucial point in this basic technique of nonlinear dynamics is the choice of the embedding parameters allowing for the best reconstruction of the phase space of the dynamical system. In Fig. 2 we consider the normalized permutation entropy resulting from the time series of the first coordinate of the Lorenz system, an exemplary workhorse of nonlinear dynamics. Obviously, h_n depends on the selected delay τ and the order n in a non-uniform way. There are four phases:

1. Before *(I)* (see Fig. 2A), there is a linear rise of the value for increasing τ. It reflects the unfolding of the trajectory and shows the loss of correlation between the values in the embedding vector (i.e. ranked sequence). DeMicco et al. [27] called the influence of this correlation "redundancy effect", where a high correlation leads to reconstructed trajectories which only visit single parts of the phase space. Decreasing values of correlation represent an unfolding of the trajectories, thus maximizing the number of participating parts of the phase space.
2. After *(I)* (see Fig. 2A) there is a small epoch of saturation which ends at *(II)* (see Fig. 2B). The limit *(II)* does not depend on τ but on the range of the embedding vector $(n-1)\tau$ and indicates the begin where the last value of the vector becomes independent from the first one.
3. From *(II)* on, this independence involves more and more values of the vector until *(III)*, where the first and the second entry of the embedding vector are no longer correlated and the maximum entropy is reached.
4. The remaining part after *(III)* is called "irrelevance effect" by De Micco et al. [27].

Fig. 2 also shows an underlying oscillation in the region after *(I)*, which reflects the influence of typical oscillation lengths and allows a fine tuning of embedding (see Fig. 2A). This is more impressive for the van-der-Pol-oscillator as an example of regular oscillation where the phases after *(II)* do not exist (see Fig. 3). In the case of $n = 7$ a maximum value of h_n can be found for the values of τ equal to the main period length, which was also reported by Bandt et al. [21] for other regular oscillations. This maximum reflects the complexity of the variation of the main oscillation, resulting from numerical errors, as it is not displayed for $n = 3$ due to the smaller number of included cycles in the rank sequence. This finding is in accordance with the applications shown in Table 4, where the optimum values of τ are equal to the annual cycles and the dimension n is not smaller than 4. In the meaning of time delay

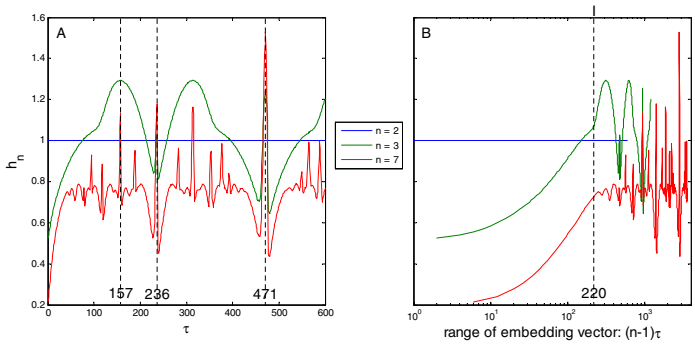

Fig. 3. (A) Normalized permutation entropy h_n of the first coordinate of the van-der-Pol-oscillator (integration step: 0.0405; $N = 49329$; period: 471 data points) for different values of τ and n. (B) The same values of h_n in relation to the range of the time delay embedding vector (embedding dimension n and delay τ). The dotted lines mark prominent points of the function h_n: the first local maximum for $n = 3$ ($\tau = 157$), the first local minimum for $n = 3$ ($\tau = 236$) and the location of marker *(I)* at $(n-1)\tau = 220$ (cf. Fig. 2).

embedding, the "redundancy effect" is overcome for a range of the embedding vector of $(n-1)\tau = 220$ which is slightly smaller than half of the main period length (see Fig. 3B). At this value of τ the PE is highest for $n = 3$, which indicates the sufficient embedding dimension of the system. For this dimension, the first local maximum is located at τ equal to about the third of the main period length. This supports the finding of Groth et al. [18] that the trajectories are most unfolded for the n-th part of the main period length (Fig. 3A). In contrast to the chaotic dynamics of the Lorenz system, the first threshold *(I)* for the van-der-Pol-oscillator depends on the range of the embedding vector.

Coming back to the regions after marker *(II)*, h_n increases linearly with τ until the maximum. We assumed the decreasing correlation between the components of the embedding vector being the cause of that. Following an idea of De Micco et al. [27], we consider a resampled version of the time series (see Fig. 4A). The new sampling step is set to τ. If we vary the starting point of this resampling from the index 1 to τ of the original signal, we get an ensemble of time series for each τ, for which we can estimate the distribution of h_n (delay 1 and dimension n). This distribution can be displayed by e.g. boxplots (see Fig. 4A). The comparison of the original function of h_n and the median of the ensemble values of the h_n-distribution shows that the location of *(I)* remains the same. In contrast, the increase between *(II)* and *(III)* disappears. The reason can be found in the decreasing number of sequences (reconstructed vectors) caused by the resampling (see e.g. Fig. 4). This decrease leads to the phenomenon of forbidden words also in random processes, where it is used to distinguish chaos from noise [23,26,28–34], and thus proves the assumption that the loss of correlation leads to rising values of h_n. The initial length of the original time series used here with $N = 6025$ is of course higher than necesseary for $n = 5$. But in order for the resampling to provide reasonable results for small values of τ a longer time series was chosen. The effects of too few data points become visible for higher τ.

As a last point, additive noise of moderate level does neither change the location of *(I)* nor the functional structure after it (see Fig. 4B). Summarizing the last results, the normalized permutation entropy is suitable to characterize dynamical processes being observed at a fixed rate per unit time. As the best choice of n, we recommend the value which leads to the highest h_n at the point *(I)* (saturation after "redundancy effect"). It is $n = 6$ and $n = 3$ in the cases of the Lorenz system (see Fig. 2A) and the

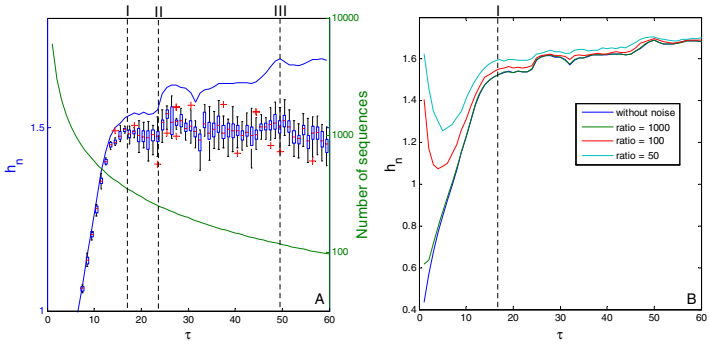

Fig. 4. (A) The blue line shows the normalized permutation entropy h_n of the first coordinate of the Lorenz system (integration step: 0.0170; $N = 6025$) for different values of τ and n. Dotted lines display markers from Fig. 2. Additionally, there is a boxplot for each τ which results from a resampling of the time series using a sampling step τ. The variation of the starting point of this resampling from the index 1 to τ of the original time series, leads to τ realizations of h_n. The normalized PE h_n is calculated using a delay of $\tau = 1$ and a dimension of $n = 5$ for each member of this ensemble. The red line in the boxes marks the median of h_n in ensemble. The upper and lower boundaries of the box are the upper and the lower quartiles, respectively. Black whiskers indicate the range of values and the red crosses mark assumed outliers. The green line shows the decreasing number of sequences (reconstructed vectors) caused by the resampling process. (B) The PE h_n (delay τ and $n = 5$) for different signal to noise ratios.

van-der-Pol-oscillator (see Fig. 2B), respectively. The optimal value of τ is either the location of *(I)*, respectively $\tau = 16$ and $\tau = 157$ in the Lorenz system and the van-der-Pol-oscillator, or the location of the first prominent local maximum, i.e. $\tau = 22$ and $\tau = 220$. The estimated embedding dimension for the van-der-Pol-oscillator is conform with the condition of Taken's delay embedding theorem [2] where a delay map size of two times the box counting dimension plus one is sufficient. In the case of the Lorenz attractor the box counting dimension is 2.06 which leads to the estimated value of an embedding dimension of 6.

Let us now come back to the second class ii) of the analysed processes, which is based on the observation of events as well as the time between their occurrence. In Fig. 4B the effect on the normalized permutation entropy is e.g. displayed by means of additive white noise. You see that there is an overlapping of influences of both the Lorenz system and the stochastic process where the stochastic process leads to a local maximum at $\tau = 1$. In contrast to the highly correlated successive sample points of the Lorenz system ("redundancy effect") each step in the noise process reflects an innovation less predictable than in the Lorenz system alone. This event based class not only includes stochastic processes but also nonlinear iterative maps, because of an innovation in each step. Therefore, $\tau = 1$ is mostly used in the case of these event-based processes, as can be seen in the tables 1, 2, 3, and 4. In contrast to case (i), a "redundancy effect" is not possible. When looking at the logistic map in different regimes, for example (Fig. 5), the phase before mark *(*I) (Figs. 2A, 3B, and 4) is skipped as expected. Beginning from $\tau = 1$, there is either an oscillatory modulation on a constant level (blue line in Fig. 5) or an increase of h_n until a maximum value (green and red line in Fig. 5) like in the case (i) after mark *(II)*. The constant level corresponds to the plateau between *(I)* and *(II)* (see Fig. 2A, 3, and 4A). The minimum values are located at integer multiples of half the mean period (blue line Fig. 4). In the case of intermittence we not only find the increase of h_n but also local minima at integer multiples of the period 3 (green line). The chaotic regime is characterized by an abrupt

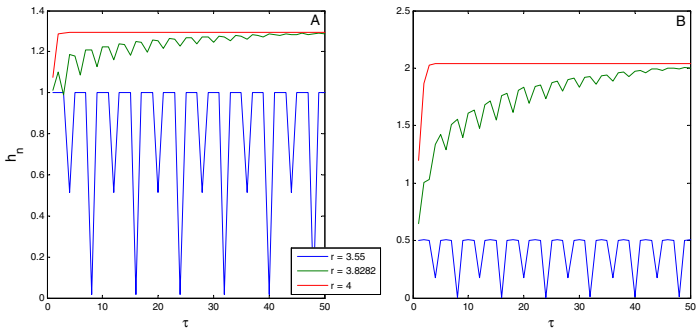

Fig. 5. Normalized permutation entropy of the logistic map for different values of τ and for order (A) $n = 3$ and (B) $n = 7$. The colours represent the different regimes: period 8 (blue), intermittence near period 3 (green) and deterministic chaos (red). In the case of $\tau = 1$ a higher dimension leads to clearer distinction between the different regimes.

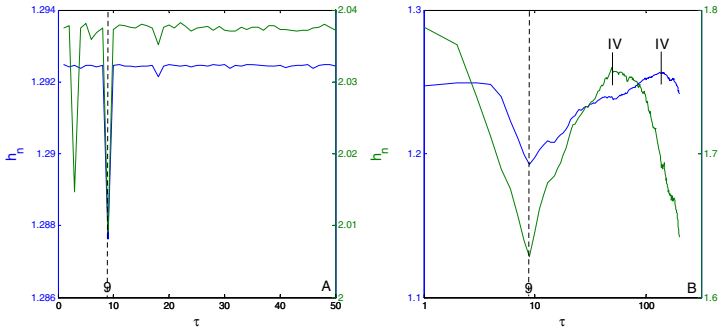

Fig. 6. Normalized permutation entropy for different values of τ and orders (blue: $n = 3$, green: $n = 7$) of a realization of (A) the autoregressive process $x_i = 0.3x_{i-9} + \epsilon_i$ (ϵ_i - white noise) and (B) the integration (cumulative sum) of this realization as an example of a random walk, where the stepwise increments are correlated. The long dotted lines mark the location of the defined lags. The lines marked *(IV)* are at the same location for $n = 3$ and $n = 7$, if h_n is plotted over the range of the sequences $(n-1)\tau$.

change after $\tau = 1$ to the value of h_n for white noise. A higher order does not essentially change this behaviour. Only the differentiation between the regimes improves for $\tau = 1$.

Finally, we consider an autoregressive process (see Fig. 6) and its integrated version which represents a random walk with correlated steps, as examples for random processes. The first one is regarded as a representative of a model commonly used in medical applications whereas the second one stands for economic models. Calculating h_n for a realization of the autoregressive process, the defined delay is characterized by a prominent local minimum for both $n = 3$ and $n = 7$. For most of the other values of τ, h_n is nearly constant and similar to the maximal value of h_n. Additionally, there is second local minimum for $n = 7$ which is an artifact of the enlargement of the sequences (see Fig. 6A). The defined lag 9 is also found for the resulting realization of the random walk (see Fig. 6B), again represented by a characteristic local minimum. But in contrast to the realization of the autoregressive process, there is a decrease of h_n after the line marked *(IV)* (see Fig. 6B). This mark scales with the range of the sequence $(n-1)\tau$. Therefore, the decreasing value of h_n reflects the increasing effect of the observed trend in this realization.

6 Summary

Combining the findings of the exhaustive literature research and the shown examples, we propose the following parameter selection procedure:

Decide which kind of system is considered:

i) Measurements observed at a fixed rate per unit time:
   a) Complexity of the main oscillation:
   Choose τ as the period length of the main oscillation and a dimension preferably with $n > 3$, which leads to a maximum of PE; the higher the dimension the better resolution of differences in the complexity, (see Table 4 and Fig. 3). This case is similar to ii)a) where the onset of each cycle represents a new event.
   b) Determining embedding dimension:
   Choose the n where the normalized PE (Eq. (2)) has its highest value at *(I)* (saturation after "redundancy effect") and τ either as the location of *(I)* or the location of the first prominent local maximum in the range between *(I)* and *(II)* (see Table 1 and Figs. 2 and 3).
   c) Determining delay/coupling:
   Choose τ as the location of the minimum of PE for a maximum of n (see Tables 1 and 3).

ii) Measurements of processes that possess an inherent cycle represented by triggering events:
   a) Complexity of the time series:
   Choose $\tau = 1$ (see Tables 1, 2, 3 and 4 and Figs. 5 and 6) and the maximal dimension; the higher the dimension the better resolution of differences in the complexity (see Fig. 5).
   b) Determining delay/coupling:
   Choose τ as the location of the minimum of PE for a maximum of n; (see Figs. 5 and 6).

The most important artefact in the calculation of permutation entropy is connected to a too small sample size in relation to the dimension/order of permutation n. Therefore, we advise to follow the recommendation in [26] to choose the maximum n according to $N > 5n!$. This decision tree is only valid for the classical PE as introduced by Bandt and Pompe [4]. Since the introduction many extensions for PE and other applications of symbolic dynamics, cf. e.g. [9–14], have been proposed. But these extensions require additional consideration, which is out of scope of this review. Nonetheless, the classical PE provides a simple first fast step for analyzing complexity in time series, as shown in the reviewed literature and our simulations. But, as with all time series analysis methods, meaningful interpretation requires inclusion of existing *a priori* knowledge about the system. We also suggest relating own findings to results gained via established methods.

References

1. C.E. Shannon, Bell Syst. Techn. J. **27**, 379 (1948)
2. H. Kantz, T. Schreiber, *Nonlinear time series analysis* (Cambridge University Press, 2003)
3. S.M. Pincus Proc. National Acad. Sci. USA **88**, 2297 (1991)
4. C. Bandt, B. Pompe Phys. Rev. Lett. **88**, 1 (2002)
5. M. Zanin, L. Zunino, O.A. Rosso, D. Papo Entropy **14**, 1553 (2012)
6. C.S. Daw, C.E.A. Finney, Rev. Scientific Instru. **74**, 915 (2003)

7. J. Amigó *Permutation Complexity in Dynamical Systems* (Springer Series in Synergetics, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010)
8. J. Kurths, A. Voss, P. Saparin, A. Witt, H.J. Kleiner, N. Wessel, Chaos (Woodbury, N.Y.) **5**, 88 (1995)
9. N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Physics Reports **438**, 237 (2007)
10. S. Schinkel, N. Marwan, J. Kurths, Cognitive Neurodyn. **1**, 317 (2007)
11. M. Staniek, K. Lehnertz, Phys. Rev. Lett. **100**, 158101 (2008)
12. X. Sun, Y. Zou, V. Nikiforova, J. Kurths, D. Walther, BMC Bioinform. **11**, 607 (2010)
13. L. Zunino, D.G. Pérez, A.M. Kowalski, M.T. Martín, M. Garavaglia, A. Plastino, O.A. Rosso, Physica A: Stat. Mech. Appl. **387**, 6057 (2008)
14. N. Wessel, A. Suhrbier, M. Riedl, N. Marwan, H. Malberg, G. Bretthauer, T. Penzel, J. Kurths, Europhys. Lett. **87**, 10004 (2009)
15. M. Staniek, K. Lehnertz, Int. J. Bif. Chaos **17**, 3729 (2007)
16. C. Bian, C. Qin, Q.D.Y. Ma, Q. Shen, Phys. Rev. E **85**, 021906 (2012)
17. Y. Cao, W.-w. Tung, J. Gao, V. Protopopescu, L. Hively, Phys. Rev. E **70**, 1 (2004)
18. A. Groth, Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Related Interdisc. Topics **72**, 1 (2005)
19. J.M. Amigó, L. Kocarev, J. Szczepanski, Phys. Lett. A **355**, 27 (2006)
20. L. Zunino, M. Soriano, I. Fischer, O.A. Rosso, C.R. Mirasso, Phys. Rev. E **82**, 1 (2010)
21. C. Bandt, Ecol. Modell. **182**, 229 (2005)
22. R. Monetti, W. Bunk, T. Aschenbrenner, F. Jamitzky, Phys. Rev. E **79**, 046207 (2009)
23. M.C. Soriano, L. Zunino, L. Larger, I. Fischer, C.R. Mirasso, Opt. lett. **36**, 2212 (2011)
24. J.-G. Wu, Z.-M. Wu, G.-Q. Xia, G.-Y. Feng, Opt. Exp. **20**, 1741 (2012)
25. L. Zunino, B.M. Tabak, F. Serinaldi, M. Zanin, D.G. Pérez, O.A. Rosso, Physica A: Stat. Mech. Appl. **390**, 876 (2011)
26. J.M. Amigó, S. Zambrano, M.A.F. Sanjuán, EPL (Europhys. Lett.) **83**, 60005 (2008)
27. L. De Micco, J.G. Fernández, H. Larrondo, A. Plastino, O. Rosso, Physica A: Stat. Mech. Appl. **391**, 2564 (2012)
28. L.C. Carpi, P.M. Saco, O.A. Rosso, Physica A: Stat. Mech. Appl. **389**, 2020 (2010)
29. H.A. Larrondo, M.T. Martín, C.M. González, A. Plastino, O.A. Rosso, Phys. Lett. A **352**, 421 (2006)
30. G. Ouyang, X. Li, C. Dang, D.A. Richards, Phys. Rev. E **79**, 041146–1 (2009)
31. O.A. Rosso, L.C. Carpi, P.M. Saco, M. Gómez Ravetti, A. Plastino, H.A. Larrondo, Physica A: Stat. Mech. Appl. **391**, 42 (2012)
32. J. Tiana-Alsina, J.M. Buldú, M.C. Torrent, J. García-Ojalvo, Phil. Trans. Ser. A, Math., Phys. Eng. Sci. **368**, 367 (2010)
33. M. Zanin, Chaos (Woodbury, N.Y.) **18**, 013119 (2008)
34. L. Zunino, M. Zanin, B.M. Tabak, D.G. Pérez, O.A. Rosso, Physica A: Stat. Mech. Appl. **388**, 2854 (2009)
35. M.C. Soriano, L. Zunino, O.A. Rosso, I. Fischer, C.R. Mirasso, IEEE J. Quant. Elect. **47**, 252 (2011)
36. S. Xiang, W. Pan, B. Luo, L. Yan, X. Zou, N. Jiang, L. Yang, H. Zhu, Opt. Commun. **284**, 5758 (2011)
37. K. Keller, M. Sinn, Physica A: Stat. Mech. Appl. **356**, 114 (2005)
38. L. De Micco, C. González, H.A. Larrondo, M.T. Martín, A. Plastino, O.A. Rosso, Physica A: Stat. Mech. Appl. **387**, 3373 (2008)
39. L. De Micco, H.A. Larrondo, A. Plastino, O.A. Rosso, Phil. Trans. Ser. A, Math., Phys. Eng. Sci. **367**, 3281 (2009)
40. O.A. Rosso, L. De Micco, H.a. Larrondo, M.T. Martín, A. Plastino, Int. J. Bifurcation Chaos **20**, 775 (2010)
41. C. Masoller, O.A. Rosso, Phil. Trans. Ser. A, Math., Phys. Eng. Sci. **369**, 425 (2011)
42. M. Matilla-García, M. Ruiz Marín, Econ. Lett. **105**, 1 (2009)
43. O.A. Rosso, R. Vicente, C.R. Mirasso, Phys. Lett. A **372**, 1018 (2008)
44. O.A. Rosso, L. Zunino, D.G. Pérez, A. Figliola, H.A. Larrondo, M. Garavaglia, M.T. Martín, A. Plastino, Phys. Rev. E **76**, 061114 (2007)

45. M. Sinn, K. Keller, Comput. Stat. Data Anal. **55**, 1781 (2011)
46. L. Zunino, D.G. Pérez, M.T. Martín, M. Garavaglia, A. Plastino, O.A. Rosso, Phys. Lett. A **372**, 4768 (2008)
47. O.A. Rosso, C. Masoller, Phys. Rev. E **79**, 040106–1 (2009)
48. K. Keller, M. Sinn, J. Emonds, Stoch. Dyn. **07**, 247 (2007)
49. K. Keller, H. Lauffer, M. Sinn, Chaos Compl. Lett. **2**, 247 (2007)
50. E. Olofsen, J.W. Sleigh, A. Dahan, British J. Anaesthesia **101**, 810 (2008)
51. D. Jordan, G. Stockmanns, E.F. Kochs, S. Pilge, G. Schneider, Anesthesiology **109**, 1014 (2008)
52. X. Li, S. Cui, L.J. Voss, Anesthesiology **109**, 448 (2008)
53. A. Silva, S. Campos, J. Monteiro, C. Venâncio, B. Costa, P. Guedes de Pinho, L. Antunes, Anesthesiology **115**, 303 (2011)
54. S. Faul, G. Boylan, S. Connolly, W. Marnane, G. Lightbody, Entropy **2**, 381 (2005)
55. A.A. Bruzzo, B. Gesierich, M. Santi, C.A. Tassinari, N. Birbaumer, G. Rubboli, Neurological Sci.: Official J. Italian Neurol. Soc. Italian Soc. Clinical Neurophys. **29**, 3 (2008)
56. K. Keller, K. Wittfeld, Int. J. Bifurcation Chaos **14**, 693 (2004)
57. X. Li, G. Ouyang, D.A. Richards, Epilepsy Res. **77**, 70 (2007)
58. N. Nicolaou, J. Georgiou, Expert Syst. Appl. **39**, 202 (2012)
59. A. Silva, H. Cardoso-Cruz, F. Silva, V. Galhardo, L. Antunes, Anesthesiology **112**, 355 (2010)
60. B. Frank, B. Pompe, U. Schneider, D. Hoyer, Med. Biol. Eng. Comput. **44**, 179 (2006)
61. U. Parlitz, S. Berg, S. Luther, A. Schirdewan, J. Kurths, N. Wessel, Comput. Biol. Med. **42**, 319 (2012)
62. A.M. Kowalski, M.T. Martín, A. Plastino, O.A. Rosso, Physica D: Nonlinear Phenomena **233**, 21 (2007)
63. A.M. Kowalski, M.T. Martín, A. Plastino, O.A. Rosso, M. Casas, Entropy **13**, 1055 (2011)
64. J. Tiana-Alsina, M.C. Torrent, O.A. Rosso, C. Masoller, J. Garcia-Ojalvo, J. Phys. Rev. A **82**, 1 (2010)
65. P.M. Saco, L.C. Carpi, A. Figliola, E. Serrano, O.A. Rosso, Physica A: Stat. Mech. Appl. **389**, 5022 (2010)
66. L. Zunino, M. Zanin, B.M. Tabak, D.G. Pérez, O.A. Rosso, Physica A: Stat. Mech. Appl. **389**, 1891 (2010)
67. L. Zunino, A. Fernández Bariviera, M.B. Guercio, L.B. Martinez, O.A. Rosso, Physica A: Stat. Mech. Appl. **391**, 4342 (2012)