Permutation Entropies and Chaos

Russell Hawkins
Department of Physics
UC Davis
rrhawkins@ucdavis.edu

June 11, 2015

Abstract

Permutation entropy is a metric used to quantify the regularity of a time series. Here I report on its application to the problem of characterizing the behavior of a few nonlinear dynamical systems, specifically the Duffing oscillator and the Tent and Logistic maps. Numerical results indicate that the permutation entropy can in principle be used to detemine the entropy rate of a system, and thus whether it is chaotic or not, but that in practice the convergence is slow. Additionally, comparisons are made between the excess entropy of the tent and logistic maps and a recently defined analagous quantity for permutations, the permutation excess entropy.

1 Introduction

Automatically classifying the behavior of a nonlinear dynamical system as periodic or chaotic from time series data alone is challenging. One approach involves a quantity called the Permutation Entropy [1] [2] [3]. The per symbol permutation entropy converges to the entropy rate of a process for a broad class of systems. Nonzero entropy rate is indicative of stochasticity in discrete time discrete state processes, and chaos in continuous deterministic dynamical systems, when applied to their symbolic dynamics. This suggests that permutation entropy can in principle be used to distinguish regular and chaotic behavior in dynamical systems.

Permutation entropy has a number of advantages over other approaches to this problem. It is simple to implement, relatively fast, and robust to observational noise. The convergence described above, however, is not guarenteed to be within practical limits. The focus of this report is to determine how practical the permutation entropy is when employed in this way.

Results of measurements of permutation entropy are shown for three systems: the Duffing oscillator, a canonical nonlinear oscillator, the Tent map, and the Logistic map. For each of these systems, the permutation entropy rate is compared to the entropy rate over a range of parameters. It is apparent that the permutation entropy rate captures qualitatively the behavior of the entropy rate for all these examples, however the quantitative convergence is inconveniently slow.

Additionally, the Permutation Excess Entropy is examined, for the Tent and Logistic maps. The permutation excess entropy is defined analogously to the standard excess entropy, and converges to it for the same class of systems for which the permutation entropy rate converges to the entropy rate [2]. Again, qualitative behavior is captured, however quantitative convergence is complicated by the partition dependence of the standard excess entropy, making the interpretation of the comparison less clear.

2 Background

Permutation entropy is best defined through a simple example. Let our time series data be the following:

$$\mathbf{x} = (5, 9, 2, 4, 6, 8, 9).$$

Now we examine the time series through a length 3 sliding window. Our first window is $(5, 9, 2)$. We map this window to the permutation 132 because $x_1 < x_3 < x_2$. The next window is $(9, 2, 4)$, which maps to 231. The next window, $(2, 4, 6)$, maps to 123, as do the final two windows, $(4, 6, 8)$ and $(6, 8, 9)$. The permutation entropy of order 3 is then

$$H_3^* = 2 \cdot -1/5 \cdot \log_2 1/5 - 3/5 \cdot \log_2 3/5$$

$$H_3^* \approx 1.37 \text{ bits}$$

More generally, to compute the permutation entropy of order n we consider all $n!$ permutations of the orderings of elements in windows of length n. We define the frequency of a permutation as

$$p(\pi) = \frac{\#\text{ of windows of permutation } \pi}{T - n + 1}$$

where T is the length of the time series, that is, the total number of elements. The normalization $T - n + 1$ is simply the number of length n windows in a time series of length T. This frequency converges to the true probability in the $T \to \infty$ limit. With these probabilities in hand, the permutation entropy of order n is then

$$H_n^* = -\sum_{\pi} p(\pi) \log_2 p(\pi).$$

The per symbol permutation entropy is defined as

$$h_n^* = \frac{H_n^*}{n-1}.$$

There is an $n-1$ in the denominator as opposed to an n because H_1^* is not defined. This in turn leads us to define the permutation entropy rate of a process as

$$h^* = \lim_{n\to\infty} h_n^*.$$

It is this quantity that can be shown to converge to the Shannon entropy rate of a process for a broad class of systems. The results of this paper will compare how the lenght n approximate to h^*, h_n^*, converges to h_μ, the Shannon entropy rate, which will be computed by other means.

Another quantity of interest is the Permutation Excess Entropy. This is defined in terms of H_n^* and h^* in the same way that the conventional excess entropy is defined in terms of the length l Shannon entropy H_l and the Shannon entropy rate h_μ. Explicitly, the excess entropy E is defined as

$$E = \lim_{l\to\infty} H_l - l \cdot h_\mu.$$

There are other equivalent expressions for the E, which are helpful for understanding and interpretation, but here we will require only this form. The permutation excess entropy is defined as

$$E^* = \lim_{n\to\infty} H_n^* - n \cdot h^*.$$

For the systems analyzed, h_μ can be substituted for h^*. The convergence of the length n approximation to the permutation excess entropy,

$$E_n^* = H_n^* - n \cdot h_\mu,$$

is compared with the E computed from a binary partition of data from the tent and logistic maps.

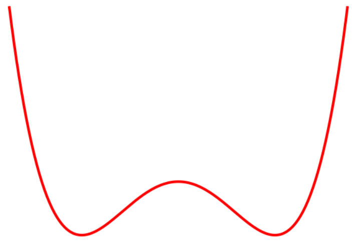

Figure 1: Duffing Potential

3 The System(s)

The systems reported on here are canonical examples of chaotic continuous and discrete time dynamical systems.

3.1 The Duffing Oscillator

The Duffing oscillator is a nonlinear, continuous time dynamical system whose equations of motion is

$$\ddot{x} + \gamma\dot{x} - \alpha x + \beta x^3 = G\cos(\omega t),$$

in the damped, driven case. Physically speaking, this is the equation of motion for a cosinusoidally driven particle in a quadratic potential well subject to viscous damping. Here we will consider the double well case, whose potential is illustrated in figure 1.

When considering the chaotic dynamics of the damped driven Duffing equation, it is important to emphasize that the equation above can be recast as 3 first order equations:

$$\begin{aligned}\dot{x} &= y\\ \dot{y} &= -\gamma y + \alpha x - \beta x^3 + G\cos\theta\\ \dot{\theta} &= \omega.\end{aligned}$$

From this it can be seen that the true phase space of the system is 3 dimensional, which in turn means the system can exhibit chaotic dynamics. This form is also essential for the calculation of the Lyapunov spectrum of the system, as will be described in the Methods section. In my analysis, I will vary G and ω, the driving frequency and amplitude, while holding the damping constant.

3.2 Tent and Logistic Maps

The tent and logistic maps are nonlinear, discrete time maps of the unit interval onto itself. The tent map is defined as follows:

$$x_{n+1} = \begin{cases} ax_n, & 0 \leq x_n \leq \frac{1}{2} \\ a(1-x_n), & \frac{1}{2} < x \leq 1 \end{cases}$$

The logistic map is defined as:

$$x_{n+1} = rx_n(1 - x_n)$$

Both of these maps exhibit periodic and chaotic behavior as their control parameters are varied. 1 dimensional maps are computationally very easy and efficient to implement, which greatly facilitates the analysis of the convergence of h_n^*.

4 Methods

4.1 Duffing Oscillator Entropy Rate

The entropy rate computed for the Duffing oscillator is the Kolmogorov-Sinai entropy rate, which is the analog of the Shannon entropy rate for continuous time systems. A result known as Pesin's theorem states that the KS entropy rate is given by the sum of the positive Lyapunov exponents of the system. Due to dimensional constraints, for a 3 dimensional dynamical system, there can only be one positive Lyapunov exponent, therefore for the Duffing oscillator the KS entropy rate is equal to the largest Lyapunov exponent.

Time series data for the Duffing oscillator was generated using built in ODE solvers in the Python packages SciPy and NumPy. To compute the Lyapunov exponents of the system, the system must be integrated in its full three dimensional, autonomous form. Additionally, time evolution via the Jacobian of the system is required, giving a total of 12 ODEs that need to be evolved in time simultaneously.

The essence of the algorithm for computing the Lyapunov spectrum goes as follows. Begin with a set of orthonormal vectors, with the number of vectors equal to the dimension of the dynamical system. Evolve this set of vectors in time using the Jacobian of the dynamical system. The Jacobian of the system determines how small displacements from a trajectory evolve in time, so the evolution of these vectors reflect how neighboring trajectories diverge from one another. After evolving for a period of time, the vectors have been stretched and rotated. The extent to which the vectors have stretched reflects the Lyapunov spectrum of the system.

To get a quantitative measurement of the Lyapunov spectrum, the orthonormal set of vectors is evolved forward by the Jacobian and periodically reorthonormalized via application of the Graham-Schmidt procedure. At each reorthonormalization, the logarithm of the change in length of each vector is recorded. Asymptotically, the time average of the logartithm of the changes in length of all of the vectors gives the complete Lyapunov spectrum.

4.2 Duffing Oscillator Permutation Entropy Rate

When calculating permutation entropy of a dynamical system it is only necessary to focus on a single degree of freedom of the system [3]. For the Duffing oscillator, I chose to focus on the position variable.

The position time series must be discretize by sampling at some frequency. This frequency was chosen to be twice the Nyquist frequency, with the justification being that given that in principle the complete continous signal can be reconstructed from such a discretization, this discretization carries the important information of the continuous signal.

The Nyquist frequency is the frequency at which the Fourier transform of a signal goes to zero. For the Duffing oscillator, in the chaotic regime at least, the Fourier transform never decays completely to zero, so a cutoff must be decided upon. From inspection of Fourier transforms of Duffing signals from multiple parameters, it seemed that 3ω, three times the driving frequency, is a good cutoff, and we used that value for all the Duffing oscillator permutation entropy calculations.

4.3 1-D Map Entropy Rate

Computation of the entropy rate for the 1-D maps is much more straightforward than for the Duffing system. In this case, Pesin's theorem guarentees that the Kolmogorov-Sinai entropy rate is equal to the Lyapunov exponent of the system. For 1-D maps, the Lyapunov exponent is given by

$$\lambda = \lim_{N\to\infty} \sum_{n=0}^{N-1} \log_2 |f'(x_n)|.$$

For the tent map, this can be evaluated in closed form, since everywhere on the interval $|f'(x)| = a$. The Lyapunov exponent, and the Kolmogorov-Sinai entropy rate, is simply

$$\lambda = \log_2 a.$$

For the logistic map, the expression for λ must be approximated from the x_n values of the trajectory.

4.4 1-D Map Excess Entropy

The analysis of the permutation excess entropy can only be carried out on the 1-D maps, because there is no known way of computing the standard excess entropy for continuous time systems. Excess entropy can be computed for 1-D maps from their symbolic dynamics. I chose to compare the permutation excess entropy with the excess entropy derived from the binary partition of the tent and logistic maps, the binary partition being the simplest generating partition for symmetric unimodal maps. From the binary time series, the length l Shannon entropy H_l is computed, and then the length l approximate to the excess entropy is

$$E_l = H_l - h_\mu \cdot l.$$

A large l value is taken as the excess entropy of the map, and various length approximations of the permutation excess entropy are compared to it.

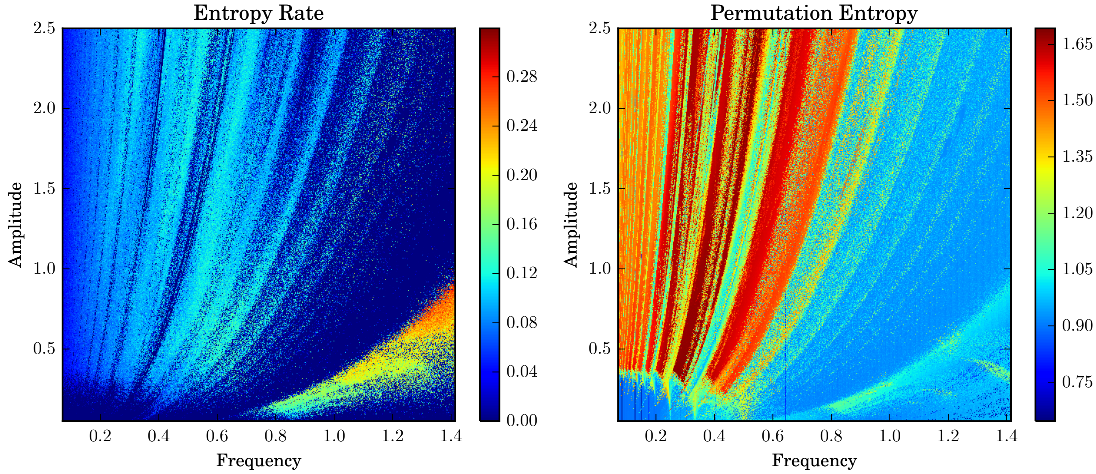

Figure 2: Duffing oscillator permutation entropy rate comparison

5 Results

5.1 Duffing Oscillator Permutation Entropy Rate

Figure 2 shows the Kolmogorov-Sinai entropy rate and the permutation entropy rate of the Duffing oscillator for a range of dimensionless driving amplitude and frequency values. It is clear that the permutation entropy rate shares the same qualitative features as the entropy rate, but quantitatively it does not match. The extent of the qualitative match is impressive when you consider the fact that the permutation entropy was computed with no reference to the equations of motion whatsoever. Also impressive is how little refinement was needed in the permutation entropy calculation. The Nyquist frequency was taken to be 3ω for each simulation, which is a rather crude estimate, and the permutation entropy rate was only taken to order $n = 5$. Yet despite that, the qualitative agreement is undeniable.

It was the desire to understand the quantitative convergence of h_n^* to h_μ that lead to the analysis of 1 dimensional maps, for which generating time series, calculating Lyapunov exponents, and calculating permutation entropies is far more computationally efficient and straightforward.

5.2 1-D Maps Permutation Entropy Rate

Before making a comparsion between the permutation entropy rate and the entropy rate for the 1-D maps, the data requirements for permutation entropies of different orders were assessed. Figures 3 and 4 plot the order $n = 2$ through 15 permutation entropy rate estimates for different data lengths for the tent and logistic maps, respectively.

When, for a given order n, points lie on top of each other, it means that going to longer data lengths does not improve the estimate. Therefore, it is safe to use the shortest

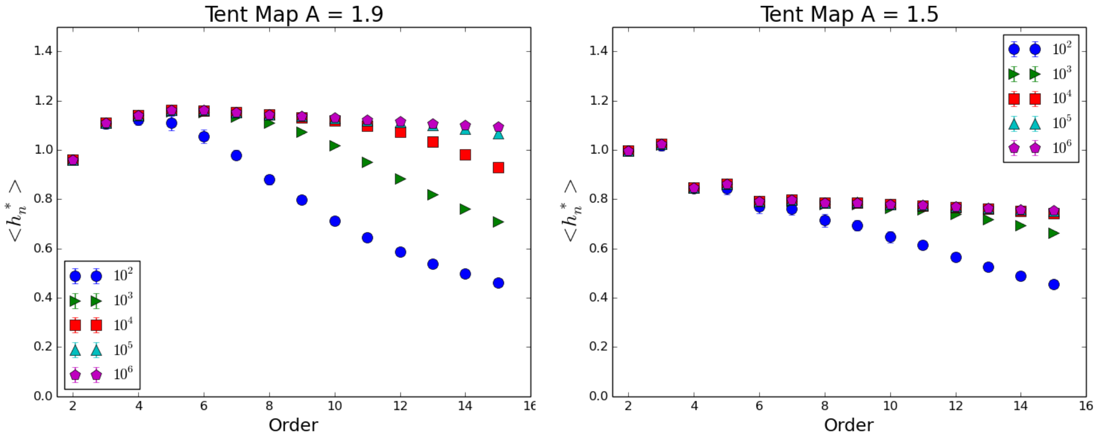

Figure 3: Tent map permutation entropy rate convergence

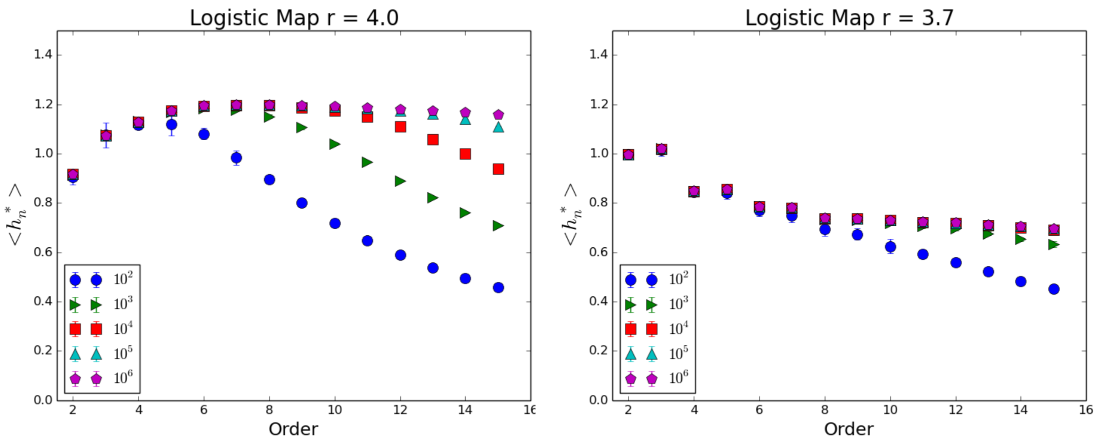

Figure 4: Logistic map permutation entropy rate convergence

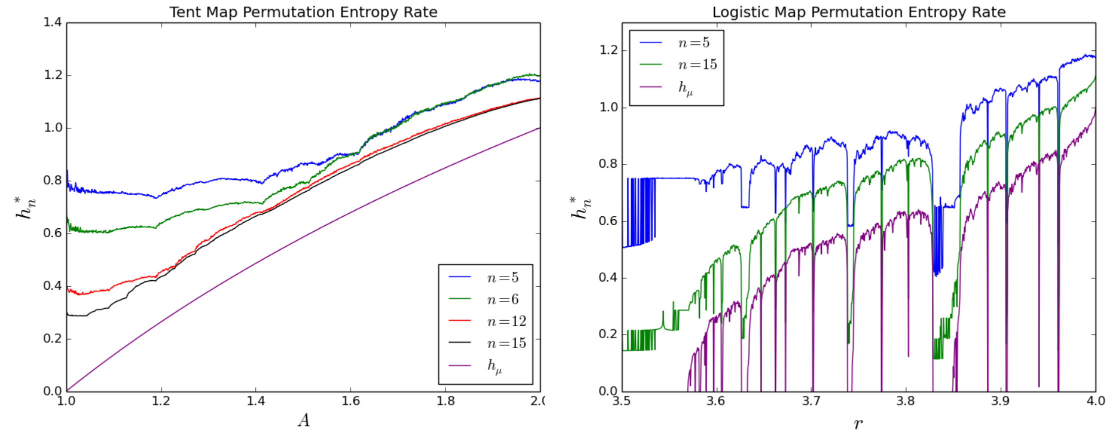

Figure 5: Tent and Logistic map permutation excess entropy comparison

data length that is among those points that overlap. From the results in figures 3 and 4 I concluded that data lengths of 10^4 are sufficient for $n \leq 12$. For $12 < n \leq 15$, length 10^5 was used. Note how much structure must be being picked up by the permutations given that, despite there being $15! \approx 1.3$ trillion permutations of 15 elements, data lengths of 10^5 and 10^6 are consistent with one another. This suggests that only a very small fraction of possible permutations are represented in the data.

Figure 5 shows h_n^* to h_μ over a range of map parameter values for increasing orders n. In both the tent and logistic maps, we can see a monotonically decreasing trend toward h_μ as n is increased. At large n values, we again see that the permutation entropy rate captures qualitatively the behavior of h_μ quite well, but does not match quantitatively.

Note that in the periodic windows of the logistic map the permutation entropy rate exhibits a kind of bistability. This is a well understood effect that is guaranteed to vanish as $n \to \infty$.

5.3 1-D Map Permutation Excess Entropy

Figures 6 and 7 compare the binary partition derived excess entropy and the permutation excess entropy for the tent and logistic maps. Again, there seems to be an interesting degree of qualitative agreement between the binary excess entropy and the various E_n^*, but there is no longer any sense of monotonic convergence. This is actually understandable, given the recently discovered fact that the excess entropy measured for a 1-D map is partition dependent. That is, different generating partitions, partitions that are equally "good" in the sense of having unique symbol sequences for each initial condition and have Shannon entropy rate equal to the Kolmogorov-Sinai entropy rate, give different values for the excess entropy.

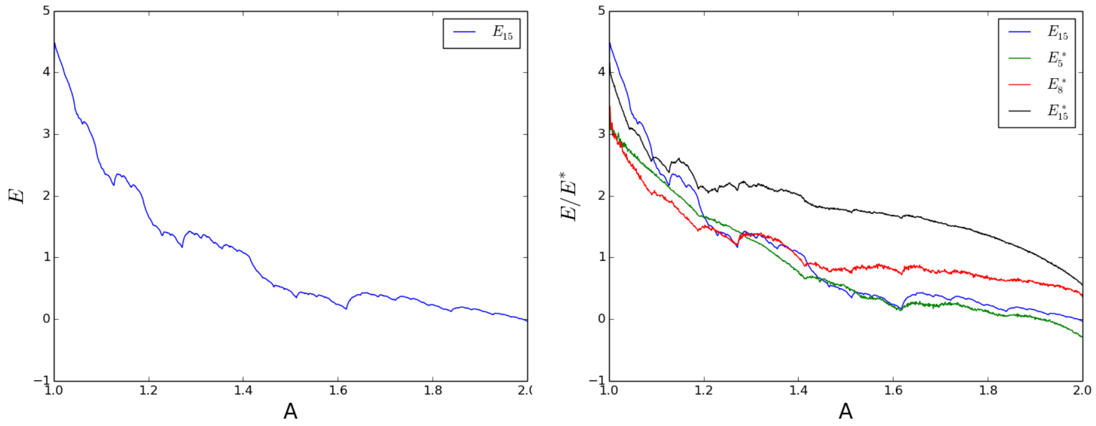

Figure 6: Tent map permutation excess entropy comparison

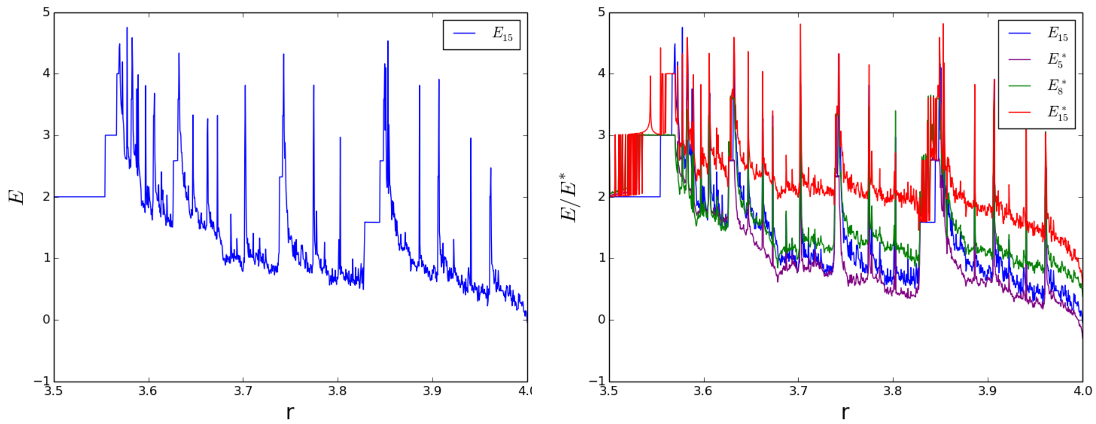

Figure 7: Logistic map permutation excess entropy comparison

With this insight in mind, it makes sense that the excess entropy derived from permutations, which are a kind of partition that becomes generating in the $n \to \infty$ limit, would give a different result than that derived from the binary partition. The permutation excess entropy should not be used as a method to compute some "standard" excess entropy, but should be seen as an informative quantity in its own right. Viewed this way, the comparisons in figures 6 and 7 show that while the values derived from the different partitions are different, they have similar qualitative features.

6 Conclusions

The results shown indicate that the permutation entropy does give some indication as to whether the behavior of a system is periodic or chaotic. It does not, however, give a clear binary answer to this question, as does the entropy rate, which is much more difficult to calculate. Moreover, as the parameters of the dynamical system are varied, the permutation entropy rate has very similar behavior as the entropy rate, but is not quantitatively the same. Quantitative agreement is possible in principle, by going to very high order, but does not appear to be practical. The results show the permutation entropy rate approaching the quantitative value of the entropy rate, but at a slow and diminishing rate.

The permutation excess entropy result is a similar but more complicated story. The results show a qualitative similarity between the binary partition excess entropy and the permutation excess entropy, but due to the partition dependence of the excess entropy, we shouldn't expect them to converge. The result then indicates that the excess entropy from the partition induced by permutation is structurally similar to that induced by the binary partition. This is good news, since there are many systems for which simple generating partitions do not exist, and this suggests that in these cases permutation excess entropy could be a useful proxy.

References

[1] C. Bandt, B. Pompe, Permutation entropy: a natural complexity measure for time series. Physical Review Letters 88, 174102, 2002

[2] T. Haruna, K. Nakajima, Permutation excess entropy and mutual information between past and future. Int. J. Comput. Ant. Sys. , 2012

[3] J. M. Amigó, Permutation Complexity in Dynamical Systems. Springer-Verlag Berlin Heidelberg, 2010.