Entropy of Difference: A New Tool for Measuring Complexity

Pasquale Nardone * and Giorgio Sonnino

Physics Department, Université Libre de Bruxelles, 50 av F. D. Roosevelt, 1050 Bruxelles, Belgium; giorgio.sonnino@ulb.be
* Correspondence: pasquale.nardone@ulb.be; Tel.: +32-2-6505515

Abstract: We propose a new tool for estimating the complexity of a time series: the entropy of difference (ED). The method is based solely on the sign of the difference between neighboring values in a time series. This makes it possible to describe the signal as efficiently as prior proposed parameters, such as permutation entropy (PE) or modified permutation entropy (mPE). Firstly, this method reduces the size of the sample that is necessary to estimate the parameter value, and secondly it enables the use of the Kullback–Leibler divergence to estimate the “distance” between the time series data and random signals.

Keywords: entropy; complexity measure; random signal

PACS: 05.45.-a; 05.45.Tp; 05.45.Pq; 89.75.-k; 87.85.Ng

MSC: 65P20; 65Z05; 91B70

1. Introduction

Permutation entropy (PE), introduced by Bandt and Pompe [1], as well as its modified version [2], are both efficient tools for measuring the complexity of chaotic time series. Both methods propose to analyze time series $X = (x_1, x_2, \cdots, x_N)$ by first choosing an embedding dimension $m$ to split the original data in a subset of $m$-tuples, $X_{m} = \{ x_1, x_2, \cdots, x_m \}$, and to then substitute the $m$-tuples values by the rank of the values, resulting in a new symbolic representation of the time series. For example, consider the time series $X = (0.2, 0.1, 0.6, 0.4, 0.1, 0.2, 0.4, 0.8, 0.5, 1.0, 0.3, 0.1, \cdots)$. Choosing, for example, an embedding dimension $m = 4$, will split the data in a set of 4-tuples: $X_4 = \{ (0.2, 0.1, 0.6, 0.4), (0.1, 0.6, 0.4, 0.1), (0.6, 0.4, 0.1, 0.2), \cdots \}$. The Bandt–Pompe method will associate the rank of the value with each 4-tuple. Thus, in $(0.2, 0.1, 0.6, 0.4)$, the lowest element 0.1 is in Position 2, the second element 0.6 is in Position 1, 0.4 is in Position 4, and finally 0.6 is in Position 3. Thus, the 4-tuple $(0.2, 0.1, 0.6, 0.4)$ is rewritten as $(2, 1, 4, 3)$. This procedure thus results in each $X_4$ being rewritten as a symbolic list: $(2, 1, 4, 3), (1, 4, 3, 2), (3, 4, 2, 1), \cdots$. Each element is then a permutation $\pi$ of the set $(1, 2, 3, 4)$. Next, the probability of each permutation $\pi$ in $X_{m}$ is then computed, $p_m(\pi)$, and finally, the PE for the embedding dimension $m$ is defined as $PE_m(X) = -\sum_{\pi} p_m(\pi) \log(p_m(\pi))$. The modified permutation entropy (mPE) just deals with those cases in which equal quantities may appear in the $m$-tuples. For example, for the $m$-tupple $(0.1, 0.6, 0.4, 0.1)$, computing PE will produce $(1, 4, 3, 2)$, while computing mPE will associate $(1, 1, 3, 2)$. Both methods are widely used due to their conceptual and computational simplicity [3–8]. For random signals, PE leads to a constant probability $q_m(\pi) = 1/m!$ (for white Gaussian noise), which does not make it possible to evaluate the “distance” between the probability found in the signal, $p_m(\pi)$, and the probability produced by a random signal, $q_m$, with the Kullback–Leibler (KL) divergence [9,10]: $KL_m[p(\pi) = \sum_{\pi} p_m(\pi) \log(p_m(\pi)/q_m(\pi))$. Furthermore, the number $K_m$ of $m$-tuples is $m!$ for PE and even greater for mPE [2], thus requiring then a large data sample to perform a significant statistical estimation of $p_m$. 

Citation: Nardone, P.; Sonnino, G. Entropy of Difference: A New Tool for Measuring Complexity. Axioms 2024, 13, 130. https://doi.org/10.3390/axioms13020130
2. The Entropy of Difference Method

The entropy of difference (ED) method proposes to substitute the \( m \)-tuples with strings \( s \) containing the sign (“+” or “−”), representing the difference between subsequent elements in the \( m \)-tuples. For the same \( X_t \), \((0.2, 0.1, 0.6, 0.4), (0.1, 0.6, 0.4, 0.1), (0.6, 0.4, 0.1, 0.2), \cdots \) this leads to the representation (“− + −”, “− − +”, “− − − +”, “− − − − +”, \cdots \). For an \( m \) value, we have \( 2^{m−1} \) strings \( s \) from “++ + · · · +” to “− − − · · · −”. Again, we compute, in the time series, the probability distribution \( q_m(s) \) of these strings \( s \) and define the entropy of difference of order \( m \) as \( ED_m = −\sum s q_m(s) \log_2 q_m(s) \). The number of elements, \( k_m \), to be treated, for an embedding \( m \), is smaller for ED compared with the number of permutations \( \pi \) in PE or to the elements in mPE (see Table 1).

Table 1. \( k \) values, for different \( m \)-embeddings.

| \( m \) | 3 | 4 | 5 | 6 | 7 |
|-------|---|---|---|---|---|
| \( k_{PE} \) | 6 | 24 | 120 | 720 | 5040 |
| \( k_{mPE} \) | 13 | 73 | 501 | 4051 | 37633 |
| \( k_{ED} \) | 4 | 8 | 16 | 32 | 64 |

Furthermore, the probability distribution for a string \( s \), in a random signal, \( q_m(s) \), is not constant and could be computed through the recursive equation. Indeed, let \( P(x \leq X_t \leq x + dx) = p(x)dx \) be the probability density for the signal variable \( X_t \) at time \( t \), and let \( F(x) \) be the corresponding cumulative distribution function \( (F(x) = \int_0^x p(x')dx) \). Consider the hypothesis that the signal is not correlated in time, which means that the join probability is only the product of the probability \( P(x_1 \leq X_{t_1} \leq x_1 + dx_1, x_2 \leq X_{t_2} \leq x_2 + dx_2) = P(x_1 \leq X_{t_1} \leq x_1 + dx_1)P(x_2 \leq X_{t_2} \leq x_2 + dx_2) \). Under these conditions, we can easily evaluate the \( q_m(s) \). For example, \( m = 3 \). We therefore have three data \( x_1, x_2, x_3 \), and we can have 4 possibilities. \( q_3(+) \) is the probability of having \( x_3 > x_2 > x_1 \), \( q_3(+) \) is that of having \( x_2 > x_1 \) and \( x_3 > x_2 \), \( q_3(−) \) is that of having \( x_1 > x_2 \) and \( x_3 > x_2 \), and finally \( q_3(−) \) is that of having \( x_1 > x_2 > x_3 \). Using the Heaviside step function \( \theta(x) \), \( \theta(x) = 1 \) if \( x \geq 0 \), and \( \theta(x) = 0 \) if \( x < 0 \), and noting the cumulative distribution function \( F(x) = P(X \leq x) \), we can evaluate \( q_3(+) \):

\[ q_3(+) = \int \int \int dx_1 dx_2 dx_3 p(x_1)p(x_2)p(x_3)\theta(x_3 - x_2)\theta(x_2 - x_1) \]
\[ = \int dx_3 p(x_3) \int_{x_3}^{x_2} dx_2 p(x_2) \int_{x_2}^{x_1} dx_1 p(x_1) \]
\[ = \int dx_3 p(x_3) \int_{x_3}^{x_2} dx_2 p(x_2)F(x_2) \]
\[ = \int dx_3 p(x_3) F(x_3) \frac{1}{2}F(x_3)^2 \]
\[ = \frac{1}{6} \] (1)

For \( q_3(−) \), we need to integrate \( \theta(x_1 - x_2)\theta(x_3 - x_2) \). Using the obvious \( \theta(x_1 - x_2) = 1 - \theta(x_2 - x_1) \), we have

\[ q_3(−) = \int \int \int dx_1 dx_2 dx_3 p(x_1)p(x_2)p(x_3)\theta(x_3 - x_2)\theta(x_1 - x_2) \]
\[ = \int \int dx_2 dx_3 p(x_2)p(x_3)\theta(x_3 - x_2) - q_3(+) \]
\[ = \int dx_3 p(x_3) F(x_3) - \frac{1}{6} = \int dx_3 F'(x_3) F(x_3) - \frac{1}{6} \]
\[ = \frac{1}{2} - \frac{1}{6} = \frac{2}{6} \] (2)
This result is totally independent of the probability density \( p(x) \) provided that the signal is not correlated in time. We can proceed in the same way for any \( q_m(s) \) and thus obtain a recurrence on \( q_m(s) \) (see Appendix A) (in the following equations, \( x \) and \( y \) are strings made of “+” and “−”):

\[
q_2(+) = q_2(−) = \frac{1}{2}
\]

\[
q_{m+1}(+,+,+,\cdots,+) = \frac{1}{(m+1)!}
\]

\[
q_{m+1}(−,x) = q_m(x) - q_{m+1}(+,x)
\]

\[
q_{m+1}(x,−) = q_m(x) - q_{m+1}(x,+)
\]

\[
q_{m+1}(x,−,y) = q_{a+1}(x)q_{b+1}(y) - q_{m+1}(x,+,y) \text{ with } a + b + 1 = m
\]

leading to a complex probability distribution for \( q_m(s) \). For example, for \( m = 9 \), we have \( 2^8 = 256 \) strings with the highest probability for the “−−−−−−−−−” string (and its symmetric “−−−−−−−−−”) : \( q_9(\text{max}) = \frac{62}{256} \approx 0.02187 \) (see Figure 1). These probabilities \( q_m(s) \) could then be used to determine the KL-divergence between the time series probability \( p_m(s) \) and the random uncorrelated signal.

To each string \( s \), we can associate an integer number, its binary representation, through the substitutions \( − \rightarrow 0 \) and \( + \rightarrow 1 \). Therefore, for \( m = 4 \), we have “−−−−” = 0, “−−+” = 1, “−+−” = 2, “−+++” = 3, and so on, up to “+++−” = 7 (see Tables 2 and 3).

![Figure 1. The 2^8 values for the probability of q_9(s), from s = −−−... ≡ 0 to s = +++... ≡ 255.](image_url)

| \( s \) | \( 0 \) | \( 1 \) | \( 2 \) | \( 3 \) | \( 4 \) | \( 5 \) | \( 6 \) | \( 7 \) | \( 8 \) | \( 9 \) | \( 10 \) | \( 11 \) | \( 12 \) | \( 13 \) | \( 14 \) | \( 15 \) |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| \( 6 \) q_3 | 1   | 2   | 2   | 1   | 1   | 2   | 2   | 1   | 1   | 2   | 2   | 1   | 1   | 2   | 2   | 1   |
| \( 24 \) q_4 | 1   | 3   | 5   | 3   | 5   | 3   | 5   | 3   | 1   | 1   | 2   | 2   | 1   | 1   | 2   | 2   | 1   |
| \( 120 \) q_5 | 1   | 4   | 9   | 6   | 9   | 16  | 11  | 4   | 4   | 11  | 16  | 9   | 6   | 9   | 4   | 1   |
Axioms 2024, 13, 130

Table 3. ED\(_m\) values, for different m-embeddings.

| m | ED\(_2\) | ED\(_3\) | ED\(_4\) | ED\(_5\) |
|---|---|---|---|---|
| 1 | 1 | \(1 + \log_2(3) = 1.9183\) | \(3 + \frac{1}{2} \log_2(3) - \frac{3}{2} \log_2(5) = 2.82501\) | \(\frac{47}{30} + \frac{3}{10} \log_2(3) + \log_2(5) - \frac{1}{11} \log_2(11) = 3.72985\) |

The recurrence gives some specific \(q_m\). To simplify the notations, we can write \(a_+\), a set of \(a\) successive “+”. For example, the second and third rules gives

\[
q_{m+1}(a_+, -) = q_m(a_+) - q_{m+1}(a_+, +) = \frac{1}{m!} - \frac{1}{(m+1)!}
\]

\[
q_{m+1}(a_+, +) = q_m(a_+) = \frac{m}{(m+1)!}
\]

(4)

then

\[
q_{m+1}(a_+, -b_+) = q_{a+1}(a_+)q_{b+1}(b_+) - q_{m+1}(a_+, +b_+) =
\]

\[
q_{m+1}(a_+, -b_+) = \frac{1}{(m+1)!} \binom{m+1}{a_+1} \text{ with } b + a + 1 = m
\]

(5)

We can also write

\[
q_{m+1}(a_+, -b_+, -, c_+) = q_{a+1}(a_+)q_{b+c+2}(b_+, -c_+) - q_{m+1}(a_+, +b_+, -, c_+) =
\]

\[
q_{m+1}(a_+, -b_+, -, c_+) = \frac{1}{(a+1)!} \frac{1}{(m-a)!} \binom{m+1}{b+c+1} - \frac{1}{(m+1)!} \binom{m+1}{a+c+1} =
\]

\[
= \frac{1}{(a+1)!(b+1)!(c+1)!} - \frac{1}{(a+1)!(b+c+2)!} - \frac{1}{(a+1)!(b+c+3)!}
\]

(6)

This equation is also valid when \(b = 0\) and thus for \(q_{m+1}(a_+, -, -, c_+)\) (with \(m = a + c + 2\) or for \(c = 0\). We can continue in this way and determine the general values of \(q_{m+1}(a_+, -, b_+, -, c_+, -d_+)\) and so on.

In the case, where the data are integers, we can avoid the situation where two successive data are equal (\(x_i = x_{i+1}\)) by adding a small amount of random noise. For example, we take the first \(10^8\) decimal of \(\pi\) (and we add a small amount of noise \(\epsilon \in [-0.01, 0.01]\)) and we have the following (See Tables 4 and 5):

Table 4. \(q_m\) values for \(\pi\), for different m-embeddings.

| m | \(q_3\) | \(q_4\) | \(q_5\) | \(q_6\) |
|---|---|---|---|---|
| 6 | 0.982 | 0.201 | 0.201 | 0.991 |
| 24 | 0.924 | 0.300 | 0.505 | 3.000 | 3.000 | 3.000 | 0.960 |
| 120 | 0.756 | 3.86 | 9.10 | 5.92 | 9.23 | 16.0 | 11.0 | 4.03 | 3.86 | 11.1 | 16.2 | 9.10 |

Table 5. ED\(_m\) values for \(\pi\), for different m-embeddings.

| m | ED\(_2\) | ED\(_3\) | ED\(_4\) | ED\(_5\) |
|---|---|---|---|---|
| 1 | 0.999998 |
| 2 | 1.91361 |
| 3 | 2.81364 |
| 5 | 3.71059 |

Despite the complexity of \(q_m(s)\), the Shannon entropy for a random signal, \(ED_m = -\sum q_m(s) \log_2 q_m(s)\), increases linearly with \(m\) (see Figure 2): \(ED_m = -0.799574 + 0.905206 m\). If the \(m\)-tuples are equiprobable, it will lead to \(- \log_2(2) + m \log_2(2) = m - 1\).
3. Periodic Signal

We will now see what happens with a period 3 data \( X = (x_1, x_2, x_3, x_1, x_2, x_3, \cdots ) \). To evaluate \( q_m \), we only have 3 types of 2-tuples. For example, for \( q_2 \), we have \( ((x_1, x_2), (x_2, x_3), \text{ and } (x_3, x_1)) \). We have only two possible strings, “+” or “−”, so the probabilities must be \( q_2(+) = 2/3, q_2(−) = 1/3 \) or \( q_2(+) = 1/3, q_2(−) = 2/3 \). For \( q_3 \), again we have only 3 types of 3-tuples: \( ((x_1, x_2, x_3), (x_2, x_3, x_1), \text{ and } (x_3, x_1, x_2)) \). We have \( 2^2 \) possible strings \( (+, +), (+, −), (−, +), \) and \( (−, −) \). The consistency of the inequalities between \( x_1, x_2, \) and \( x_3 \) reduces the number of possible strings to 3. For example, if \( (x_1, x_2, x_3) \) gives \( (+, +) \), then \( (x_2, x_3, x_1) \) must be \( (+, −) \), and \( (x_3, x_1, x_2) \) must be \( (−, +) \). Due to Period 3, these \( (x, y) \) values will appear \( 1/3 \) times. To evaluate \( q_4 \), we have again only 3 types of 4-tuples: \( ((x_1, x_2, x_3, x_1), (x_2, x_3, x_1, x_2), \text{ and } (x_3, x_1, x_2, x_3)) \), and again these will appear \( 1/3 \) times in the data. This reasoning can be generalized to a signal of period \( p, q_p = 1/p \); consequently, \( ED_p = \log_2(p) \), and this remains constant for \( m \geq 2 \). Obviously, since we are only using the differences between the \( x_i \)’s, the periodicity in terms of signs \( x_{i+1} − x_i \) may be smaller than the periodicity \( p \) of the data, so \( ED_p \leq \log_2(p) \).

4. Chaotic Logistic Map Example

Let us illustrate the use of ED on the well known logistic map [11] \( Lo(x, λ) \) driven by the parameter \( λ \).

\[
x_{n+1} = Lo(x_n, λ) = λx_n(1 − x_n)
\] (7)

It is obvious that, for a range of values of \( λ \), where the time series reaches a periodic behavior (any cyclic oscillation between \( n \) different values), the ED will remain constant. The evaluation of the ED could thus be used as a new complexity parameter to determine the behavior of the time series (see Figure 3).

Figure 3. The Shannon entropy of \( q_m(s) \): ED increases linearly with \( m \), and the fit \(-0.799574 + 0.905206 m \) gives a sum of squared residuals of \( 1.7 \times 10^{-3} \) and a \( p \)-value = \( 1.57 \times 10^{-12} \) and \( 1.62 \times 10^{-30} \) on the fit parameter respectively.
For $\lambda = 4$, we know that the data are randomly distributed with a probability density given by \[ p_{Lo}(x) = \frac{1}{\pi \sqrt{(1-x)x}} \] (8)

However, the logistic map produces correlations in the data, so we expect a deviation from the uncorrelated random $q_{m}$.

We can then compute exactly the ED for an $m$-embedding as well as the KL-divergence from a random signal. For example, for $m = 2$, we can determine $q_{2}^{Lo}(+)$ and $q_{2}^{Lo}(-)$ by solving the inequality $x < Lo(x)$ and $x > Lo(x)$, respectively, which implies that $0 < x < 3/4$ and $3/4 < x < 1$. Then,

$$
q_{2}^{Lo}(+) = \int_{0}^{3/4} dx \ p_{Lo}(x) = \frac{2}{3} \\
q_{2}^{Lo}(-) = \int_{3/4}^{1} dx \ p_{Lo}(x) = \frac{1}{3}
$$

(9)

In this case, the logistic map produces a signal that contains twice as many increasing pairs “+” than decreasing pairs “−”. Thus,

$$
ED_{2} = - \left( \frac{2}{3} \log_{2} \frac{2}{3} + \frac{1}{3} \log_{2} \frac{1}{3} \right) = \log_{2} \frac{3}{2^{7/3}} \approx 0.918 \\
KL_{2} = \frac{1}{3} \log_{2} \frac{32}{27} \approx 0.082
$$

(10)

For $m = 3$, we can perform the same calculation. We have, respectively,

$$
x_{1} < x_{2} < x_{3} \quad \rightarrow \quad (+, +) : 0 < x < \frac{1}{4} \\
x_{1} < x_{3} < x_{2} \quad \rightarrow \quad (+, -) : \frac{1}{4} < x < \frac{1}{8}(5 - \sqrt{5}) \\
x_{3} < x_{1} < x_{2} \quad \rightarrow \quad (+, -) : \frac{1}{8}(5 - \sqrt{5}) < x < \frac{3}{4} \\
x_{2} < x_{1} < x_{3} \quad \rightarrow \quad (-, +) : \frac{3}{4} < x < \frac{1}{8}(5 + \sqrt{5}) \\
x_{2} < x_{3} < x_{1} \quad \rightarrow \quad (-, +) : \frac{1}{8}(5 + \sqrt{5}) < x < 1
$$

Graphically we have:

$$
q_{3}^{Lo}(++) = \frac{1}{3} \\
q_{3}^{Lo}(+-) = \frac{1}{3} \\
q_{3}^{Lo}(-+) = \frac{1}{3} \\
q_{3}^{Lo}(--) = 0
$$

(11)

$$
ED_{3} = \log_{2} 3 \approx 1.58 \\
KL_{3} = \frac{1}{3} \approx 0.33
$$

Effectively, the logistic map with $\lambda = 4$ forbids the string “−−” where $x_{1} > x_{2} > x_{3}$. For strings of length 3, we have

$$
q_{4}^{Lo}(+++) = q_{4}^{Lo}(+ + -) = q_{4}^{Lo}(− + +) = q_{4}^{Lo}(− + −) = \frac{1}{6} \\
q_{4}^{Lo}(− + +) = \frac{2}{6}
$$

$$
ED_{4} = \log_{2} 108 \approx 2.25 \\
KL_{4} = \log_{2} \left( \frac{16,384}{1125} \right)^{1/6} \approx 0.64
$$

(12)

The probability of difference $q_{m}(s)$ for some string length $m$ versus $s$, the string binary value, where “+”→ 1 and “−”→ 0, gives us the “spectrum of difference” for the distribution $q$ (see Figures 4 and 5).
Figure 4. The $\text{ED}_{13}$ (strings of length 12) is plotted versus $\lambda$, with the bifurcation diagram, and the value of the Lyapunov exponent, respectively [13]. The constant value appears when the logistic map enters into a periodic regime.

Figure 5. From $x_1$ (blue), the first iteration of the logistic map (gray) gives $x_2$, and the second iteration (black) gives $x_2$. The respective positions of $x_1, x_2, x_3$ allow us to determine $q_3$.

5. $\text{KL}_m(p|q)$ Divergences Versus $m$ on Real Data and on Maps

The manner in which the $\text{KL}_m(p|q)$ evolves with $m$ is another parameter of the complexity measure. $\text{KL}_m(p|q)$ measures the loss of information when the random distribution $q_m$ is
used to predict the distribution $p_m$. Increasing $m$ introduces more bits of information in the signal, and the behavior versus $m$ shows how the data diverge from a random distribution.

The graphics (see Figure 6) show the behavior of $KL_m$ versus $m$ for two different chaotic maps and for real financial data [14]: the opening value of the nasdaq100, be120 every day from 2000 to 2013. For maps, the logarithmic map $x_{n+1} = \ln(|a|x_n|)$ and the logistic map are shown (see Figure 6 for the logarithmic map (Figure 7)).

Figure 6. The spectrum of $q_{13}$ (black) versus the string binary value (from 0 to $2^{12} - 1$) for the logistic map at $\lambda = 4$ and the one from a random distribution $q_{13}$ (red).

Figure 7. The ED$_{13}$ versus $a$ for the logarithm map $x_{n+1} = \ln(|a|x_n|)$.
For maps, the simulation starts with a random number between 0 and 1 and first iterates 500 times to avoid transients. Starting with these seeds, 720 iterates were kept, and $KL_m$ was computed. It can be seen that the Kullback–Leibler divergence from the logistic map at $\lambda = 4$ to the random signal is fitted by a quadratic function of $m$: $KL_m = -0.4260 + 0.2326 m + 0.0095 m^2$ (p-value $\approx 2 \times 10^{-7}$ for all parameters), while the logarithmic map behavior is linear in the range $a \in [0.4, 2.2]$. Financial data are also quadratic $KL_m(\text{nasdaq}) = 0.1824 - 0.0973 m + 0.0178 m^2$, $KL_m(\text{bel20}) = 0.1587 - 0.0886 m + 0.0182 m^2$ with a higher curvature than the logistic map due to the fact that the spectrum of the probability $p_m$ is compatible with a constant distribution (see Figure 6), rendering the prediction of an increase or decrease signal completely random, which is not the case in any true random signal (See Figures 8 and 9).

Figure 8. The KL-divergence for the data.

Figure 9. The spectrum of $q_8$ versus the string binary value (from 0 to $2^7 - 1$) for the bel20 financial data.
6. Conclusions

The simple property of increases or decreases in a signal makes it possible to introduce the entropy of difference $ED_m$ as a new efficient complexity measure for chaotic time series. This new technique is numerically fast and easy to implement. It does not require complex signal processing and could replace the evaluation of the Lyapunov exponent (which is far more time-consuming). For a random signal, we have determined the value of $ED_m$, which is independent of the probability of distribution of this signal. This makes it possible to calculate the “distance” between the analyzed signal and a random signal (independent of its distribution probability). As “distance”, we evaluate the Kullback–Leibler divergence versus the number of data $m$ used to build the difference string. This $KL_m$ shows different behavior for different types of signal and can also be used also to characterize the complexity of a time series. Since the only assumption for a random signal is that it is uncorrelated, this method makes it possible to determine the correlated nature of signals, even in chaotic regimes.

Author Contributions: Conceptualization and methodology, formal analysis, P.N.; writing, review and editing, P.N. and G.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

The Mathematica program for the probability $q(s)$:

\[
\begin{align*}
\text{P["+"\]} &= \text{P["-"\]} = 1/2; \\
\text{P["-", x___\]} &= \text{P[x\]} - \text{P["+", x\]}; \\
\text{P[x___, "-"\]} &= \text{P[x\]} - \text{P[x, "+"]}; \\
\text{P[x___, ",", y___\]} &= \text{P[x\]} \text{P[y\]} - \text{P[x, ",", y\]}; \\
\text{P[x___\]} &= 1/(StringLength[StringJoin[x]] + 1);
\end{align*}
\]

References

1. Bandt, C.; Pompe, B. Permutation Entropy: A Natural Complexity Measure for Time Series. Phys. Rev. Lett. 2002, 88, 174102. [CrossRef]
2. Bian, C.; Qin, C.; Ma, Q.D.Y.; Shen, Q. Modified permutation-entropy analysis of heartbeat dynamics. Phys. Rev. E 2012, 85, 021906. [CrossRef]
3. Zunino, L.; Pérez, D.G.; Martín, M.T.; Garavaglia, M.; Plastino, A.; Rosso, O.A. Permutation entropy of fractional Brownian motion and fractional Gaussian noise. Phys. Lett. A 2008, 372, 4768. [CrossRef]
4. Li, X.; Ouyang, G.; Richard, D.A. Predictability analysis of absence seizures with permutation entropy. Epilepsy Res. 2007, 77, 70. [CrossRef]
5. Li, X.; Cui, S.; Voss, L.J. Using permutation entropy to measure the electroencephalographic effects of sevoflurane. Anesthesiology 2008, 109, 448. [CrossRef] [PubMed]
6. Frank, B.; Pompe, B.; Schneider, U.; Hoyer, D. Permutation entropy improves fetal behavioural state classification based on heart rate analysis from biomagnetic recordings in near term fetuses. Med. Biol. Eng. Comput. 2006, 44, 179. [CrossRef] [PubMed]
7. Olofson, E.; Sleigh, J.W.; Dahan, A. Permutation entropy of the electroencephalogram: A measure of anaesthetic drug effect. Br. J. Anaesth. 2008, 101, 810. [CrossRef] [PubMed]
8. Rosso, O.A.; Zunino, L.; Perez, D.G.; Figliola, A.; Larrondo, H.A.; Garavaglia, M.; Martin, M.T.; Plastino, A. Extracting features of Gaussian self-similar stochastic processes via the Bandt–Pompe approach. Phys. Rev. E 2007, 76, 061114. [CrossRef] [PubMed]
9. Kullback, S.; Leibler, R.A. On Information and Sufficiency. Ann. Math. Statist. 1951, 22, 79. [CrossRef]
10. Roldán, E.; Parrondo, J.M.R. Entropy production and Kullback–Leibler divergence between stationary trajectories of discrete systems. Phys. Rev. E 2012, 85, 031129. [CrossRef] [PubMed]
11. May, R.M. Simple mathematical models with very complicated dynamics. Nature 1976, 261, 459. [CrossRef] [PubMed]
12. Jakobson, M. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. Commun. Math. Phys. 1981, 81, 39–88. [CrossRef]
13. Ginelli, F.; Poggi, P.; Turchi, A.; Chate, H.; Livi, R.; Politi, A. Characterizing Dynamics with Covariant Lyapunov Vectors. *Phys. Rev. Lett.* **2007**, *99*, 130601. [CrossRef] [PubMed]

14. Available online: [http://www.wessa.net/](http://www.wessa.net/) (accessed on 1 February 2024).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.