

Article

# Distinguish between Stochastic and Chaotic Signals by a Local Structure-Based Entropy

Zelin Zhang <sup>1,2,†</sup> , Jun Wu <sup>1,2</sup>, Yufeng Chen <sup>3</sup>, Ji Wang <sup>4</sup> and Jinyu Xu <sup>3,\*,†</sup> 

<sup>1</sup> School of Mathematics, Physics and Optoelectronic Engineering, Hubei University of Automotive Technology, Shiyan 442002, China

<sup>2</sup> Hubei Key Laboratory of Applied Mathematics, Hubei University, Wuhan 430061, China

<sup>3</sup> School of Electrical and Information Engineering, Hubei University of Automotive Technology, Shiyan 442002, China

<sup>4</sup> School of Liberal Arts and Humanities, Sichuan Vocational College of Finance and Economics, Chengdu 610101, China

\* Correspondence: xujinyu\_2007@126.com

† These authors contributed equally to this work.

**Abstract:** As a measure of complexity, information entropy is frequently used to categorize time series, such as machinery failure diagnostics, biological signal identification, etc., and is thought of as a characteristic of dynamic systems. Many entropies, however, are ineffective for multivariate scenarios due to correlations. In this paper, we propose a local structure entropy (LSE) based on the idea of a recurrence network. Given certain tolerance and scales, LSE values can distinguish multivariate chaotic sequences between stochastic signals. Three financial market indices are used to evaluate the proposed LSE. The results show that the  $LSE^{FSTE100}$  and  $LSE^{S\&P500}$  are higher than  $LSE^{SZI}$ , which indicates that the European and American stock markets are more sophisticated than the Chinese stock market. Additionally, using decision trees as the classifiers, LSE is employed to detect bearing faults. LSE performs higher on recognition accuracy when compared to permutation entropy.

**Keywords:** multivariate time series; information entropy; chaotic sequence; random signal; machinery fault diagnose



**Citation:** Zhang, Z.; Wu, J.; Chen, Y.; Wang, J.; Xu, J. Distinguish between Stochastic and Chaotic Signals by a Local Structure-Based Entropy. *Entropy* **2022**, *24*, 1752. <https://doi.org/10.3390/e24121752>

Academic Editors: Ravi P. Agarwal and Maria Alessandra Ragusa

Received: 4 November 2022

Accepted: 27 November 2022

Published: 30 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The local structure and its integration are crucial components in defining the vast system as a whole. For instance, the personal abilities of football players on the field and their teamwork play key roles in a competition [1], or the seismic capacities of buildings rely on materials and frame structures [2]. However, people tend to use local structures to identify objects in machine learning work [3,4]. Numerous examples can be found in feature selection [5,6], pattern recognition [7–9], and so on. Is it possible that a local structure is enough to accomplish identification? If it is possible, it means that other methods, such as global–local structure-based learning [10], are not necessary. So, no one can deny that it depends on specific issues. People tend to agree that only when local structures contain enough information for judging can the local structure-based strategies take effect. Therefore, finding the amount of information contained in the local structures of samples is important for further exploration.

In the time series analysis field, entropies were used to reflect the amount of information contained in observed data. The first one can be traced back to Shannon entropy [11], which is defined as,

$$H(X) = - \sum_{x \in X} p(x) \ln(p(x)) \quad (1)$$

where  $X$  is a random variable and  $p(x)$  is its probability density. For a time series, Equation (1) is valid under the assumption of stationarity despite whether  $X$  is univariate

or multivariate. Interested readers can refer to other entropies based on Equation (1) for univariate cases, such as sample entropy [12,13], multiscale entropy [14,15], permutation entropy (PE) [16–18], etc. These methods have one thing in common—the observed time series is embedded into a phase space in order to reflect the autocorrelation. However, the above-mentioned entropies, except for PE, cannot be directly generalized to multivariate cases. With regard to PE, although it naturally provides a multivariate version [19–21], it does not reflect correlations between components. Although the Kullback–Leibler entropy [22] is a classical tool to assess the correlation between two components, it is asymmetrical and cannot be employed on a time series whose dimension is greater than 2. In general, correlations between different components together with autocorrelation in every component, are considered double-edged swords, which provide positive affection where the times series can be predicted, as well as a negative influence, where it is difficult to measure the complexity of a multi-dimensional time series. Researchers managed to introduce other entropies to characterize multivariate sequences, such as matrix-based entropy [23,24], multiscale entropy [25–27], and estimation methods [28]. For instance, Azami proposed a refined composite multivariate multiscale fuzzy entropy that demonstrates long-range, within, or cross-channel correlations [29]. Han proposed a multivariate multi-scale weighted permutation entropy to illustrate oil–water two-phase flow instability [30]. Mao constructed a complexity–entropy causality plane based on the normalized Shannon entropy and multivariate permutation entropy, thus characterizing various chaotic systems [31]. With the complex network method, Shang put forward another complexity–entropy causality plane via multiscale entropy and the degree distribution of the vector visual graph to measure the complexity of the stock index [32]. The above-mentioned methods have at least one of the following drawbacks: (1) the observed sample numbers need to be large, otherwise the results are not reliable; (2) the computing times are long, so they cannot satisfy the real-time requirements of the application; (3) focusing on autocorrelation but neglecting relationships between components. To illuminate those shortcomings, we plan to use a complex network-like approach to represent the local structure.

In the past decades, we have witnessed the development of complex network science [33–35], which is utilized to present complex systems such as climate dynamics [36] and human language [37]. Meanwhile, it is regarded as a newly advanced tool for time series analysis, such as for univariate cases [38–40] and multivariate cases [41]. Interested readers can find more information in the review articles [42,43]. These complex network-based methods can be divided into three classes: visibility graph (VG) [44,45], recurrence net (RN) [46,47], and transition net (TN) [48,49]. For VG, a vertex is obtained by the linear relationship contained in fragments. For TN, a vertex is usually generated by certain coarse-graining strategies. Links between pairs of adjacent vertexes in VG or TN are naturally produced according to the order. However, RN is quite different. Its edges are bridged by the similarities of observed values. The validities of these methods also indicate that local structures together with their combination forms could represent essential characteristics of a complex system. Inspired by the above means, we propose a local structure-based entropy (LSE) for analyzing a multivariate time series. This can be regarded as a multivariate version of the Shannon Entropy defined on a discrete distribution of the local structure; the recurrence of the observed vector is employed to represent local structure information. In fact, RN is constructed by the similarity between all observed data. If we just consider the reappearances of the initial state in the sliding window, then its distribution can be viewed as a manifestation of local structure.

The rest of this paper is organized as follows. Section 2 introduces the process of computing LSE; Section 3 tests LSE on the fractional-order chaotic time series; LSE is applied to the stock market index and bearing fault signals; finally, we draw our conclusions in Section 5.

## 2. Method

In [50], Donner introduced the construction of RN, which takes individual values as vertices and indicators of recurrences as edges. Here, a pair of states whose values are close enough can be regarded as recurrences. Thus, RN represents the reappearance of system states over a long period. It is natural for one to (hope to) use the recurrent numbers of the initial states in sliding windows of different lengths to characterize the complexities of the systems.

Let  $x_t = (x_{1t}, x_{2t}, \dots, x_{mt})_{t \leq N}^T$  be an  $m$ -dimensional time series.  $\tau$  and  $i$  are integers with  $i + \tau - 1 \leq N$ . Extract fragment  $Y_i$  from  $x_t$  by

$$Y_i = \begin{pmatrix} x_{1,i} & \cdots & x_{1,(i+\tau-1)} \\ \vdots & \vdots & \vdots \\ x_{m,i} & \cdots & x_{m,(i+\tau-1)} \end{pmatrix} \tag{2}$$

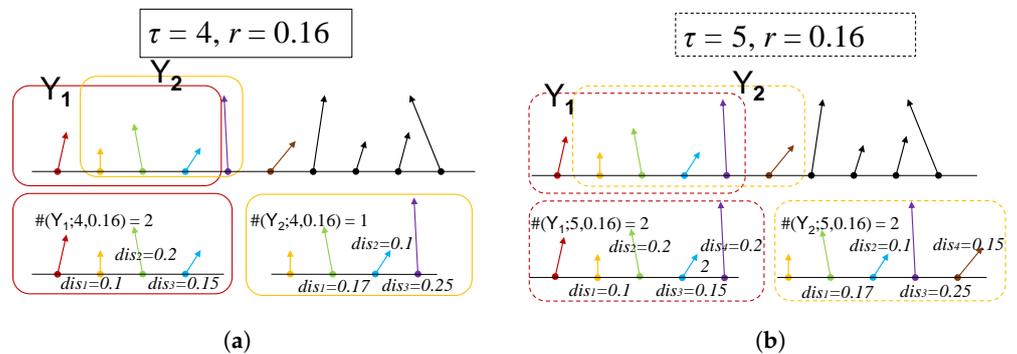
Let  $j \in \{0, 1, 2, \dots, \tau - 1\}$ ,  $y_j^i$  denotes the  $(j + 1)$ th column vector of  $Y_i$ . Define

$$dis_j = |y_0^i, y_{1+j}^i| \tag{3}$$

where  $|\cdot, \cdot|$  is a distance in  $\mathbb{R}^m$  (in this paper, we use the Euclidean distance). Given a tolerant threshold  $r$  and the number of  $j$ , such that  $dis_j \leq r$  is called the recurrent number of  $Y_i$  with the scale  $\tau$ , and the tolerance  $r$  is denoted by  $\#(Y_i; \tau, r)$ . That is,

$$\#(Y_i; \tau, r) = \sum_{j \in \{1, 2, \dots, \tau-1\}} \text{sign}(dis_j \leq r) \tag{4}$$

where  $\text{sign}(\cdot)$  is the indicator function. Two examples of the above procedure are demonstrated in Figure 1.



**Figure 1.** Schematic diagram of  $\#(Y_i; \tau, r)$  procedure. (a) tolerance  $r = 0.16$ , scale  $\tau = 4$ , (b)  $r = 0.16$ ,  $\tau = 5$ .

Now, the LSE of  $x_t$  under scale  $\tau$  and tolerance  $r$  is defined as

$$LSE(x_t; \tau, r) = \sum g(\#(Y_i; \tau, r)) \ln(g(\#(Y_i; \tau, r))) \tag{5}$$

where  $g(\cdot) = p(\cdot)/a_{\tau, r}$ ,  $p(\cdot)$  represents the density function, and  $a_{\tau, r}$  is the average of  $\#(Y_i; \tau, r)$ .

**NOTE 1** The value of LSE depends on the selection of  $\tau$  and  $r$ . When their ranges are given, they can be chosen by the maximal LSE. Namely,

$$(\tau, r) = \underset{\tau, r}{\operatorname{argmax}} LSE(x_t; \tau, r) \tag{6}$$

In general, the range of  $\tau$  can be decided by the length of the time series, such as 1/100 to 1/10 of the raw data. The range of  $r$  can be selected by referring to the standard deviation of the multivariate standard Gaussian series. See Appendix A.1.

**NOTE 2** Given a set of values for  $\tau$  and  $r$ , respectively, such as  $\tau \in \{\tau_1, \tau_2, \dots, \tau_q\}$  and  $r \in \{r_1, r_2, \dots, r_s\}$ , the values of  $LSE(x_t; \tau_i, r_j)$  form a matrix, which presents the variation of the complexity of  $x_t$  over different scales and tolerances. Therefore, it could be used to characterize  $x_t$ .

**NOTE 3** We adopt two strategies for further exploration. To designate the corresponding LSE values, we utilize the variables  $LSE_A$  and  $LSE_B$ . **METHOD A:** normalize each component of  $x_t$  and specify a tolerance range as described above, which leads to  $LSE_A$  being immune to linear transformations; **METHOD B:** do not normalize  $x_t$  but set tolerance range by the total of component variances, thus taking into account the extents of various channel volatilities. Broadly speaking,  $LSE_A$  and  $LSE_B$  are not equivalent, and they can be seen as two different ways of describing the initial sequence. See Appendix A.2 for more information.

### 3. Numerical Simulations

In this section, we test LSE on several multivariate deterministic and stochastic signals. At first, we analyze fractional-order chaotic systems and random series and find out ranges of  $\tau$  and  $r$  in Section 3.1, then we test LSE on integer-order, fractional-order chaotic systems and random sequences in Section 3.2.

#### 3.1. Fractional-Order Chaotic Systems and Random Series

Derivatives and integrals of fractional orders are employed to describe objects with memory properties, such as power law nonlocality or long-range dependence [51] and, thus, can model real-world systems more accurately than the classical integer calculus. Many fractional-order dynamical systems [52–54] with total order of less than three can exhibit chaos while continuous nonlinear systems with the total order of less than three cannot under the concept of the usual integer order. In this section, we simulate three multivariate fractional-order dynamical models and three random signals to test LSE.

For many classes of functions, the three most well-known fractional-order derivatives (the Grünwald–Letnikov (GL), Riemann–Liouville, and Caputo) are equivalent under some conditions [55]. Here, we use the GL definition, i.e.,

$$D_t^\alpha f(t) = \lim_{h \rightarrow 0} \frac{1}{h^\alpha} \sum_{j=0}^{\infty} (-1)^j \binom{\alpha}{j} f(t - jh) \tag{7}$$

Then, for the fractional-order differential equation

$${}_a D_t^\alpha y(t) = f(y(t), t) \tag{8}$$

a general numerical solution has the form of [56]

$$y(t_k) = f(y(t_k), t_k) h^\alpha - \sum_{j=v}^k c_j^{(\alpha)} y(t_{k-j}) \tag{9}$$

where

$$c_0^{(\alpha)} = 1, \quad c_j^{(\alpha)} = \left(1 - \frac{1 + \alpha}{j}\right) c_{j-1}^{(\alpha)} \tag{10}$$

Then, several different dynamic systems, including the fractional order chaotic system, multivariate vector autoregression moving-average process (VARMA), white Gaussian noise (WGN), and 1/f noise, are generated and then analyzed by LSE to showcase its effectiveness. The method in [56] is adopted to generate numerical solutions of Equations (11) to (13); they are chaotic time series, which are cross-validated with the largest Lyapunov exponent [54]. The simulation timespan is 0:0.005:30. (Start at 0, end at

30, the step is 0.005). To avoid the influences of the initial values, the first 10% of data are discarded and the sampled series (500 × 3, 500 simulated three-dimensional vectors) start from a random position. Figure 2 displays examples of these signals.

1. Fractional-order Chen system [52]

$$\begin{cases} D^{\alpha_1}x = a(y - x) \\ D^{\alpha_2}y = dx - xz + cy \\ D^{\alpha_3}z = xy - bz \end{cases} \quad (11)$$

$\alpha_1 = \alpha_2 = \alpha_3 = 0.9, a = 35, b = 3, c = 28, d = c - a = -7$ , initial values  $(x_0, y_0, z_0) = (-9, -5, 14)$ .

2. Fractional-order Rössler system [53]

$$\begin{cases} D^{\alpha_1}x = y - z \\ D^{\alpha_2}y = x + ay \\ D^{\alpha_3}z = b + z(x - c) \end{cases} \quad (12)$$

$\alpha_1 = 0.9, \alpha_2 = 0.85, \alpha_3 = 0.95, a = 0.5, b = 0.2, c = 10, (x_0, y_0, z_0) = (0.5, 1.5, 0.1)$ .

3. Fractional-order financial system [54]

$$\begin{cases} D^{\alpha_1}x = z + (y - a)x \\ D^{\alpha_2}y = 1 - by - x^2 \\ D^{\alpha_3}z = -x - cz \end{cases} \quad (13)$$

$\alpha_1 = 1, \alpha_2 = 0.95, \alpha_3 = 0.99, a = 1, b = 0.1, c = 1, (x_0, y_0, z_0) = (2, -1, 1)$ .

4. Multivariate vector autoregression moving-average process

$$y_t = c + \sum_{j=1}^p \Phi_j y_{t-j} + \sum_{k=1}^q \Theta_k \epsilon_{t-k} + \epsilon_t \quad (14)$$

$$\Phi_1 = \begin{pmatrix} 0.2 & -0.1 & 0 \\ 0.1 & 0.2 & 0.05 \\ 0 & 0.1 & 0.3 \end{pmatrix}, \Phi_2 = \begin{pmatrix} 0.3 & 0 & 0 \\ 0.1 & 0.4 & 0.1 \\ 0 & 0 & 0.2 \end{pmatrix}, \Phi_3 = \begin{pmatrix} 0.4 & 0.1 & -0.1 \\ 0.2 & -0.5 & 0 \\ 0.05 & 0.05 & 0.2 \end{pmatrix},$$

$$\Theta_1 = \begin{pmatrix} 0.2 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.2 \end{pmatrix}, \Theta_2 = \begin{pmatrix} 0.3 & 0.2 & 0.1 \\ 0.2 & 0.4 & 0 \\ 0.1 & 0 & 0.5 \end{pmatrix}, c = (0.05, 0, -0.05)^T, \text{ covariance}$$

matrix is the unit matrix,  $\epsilon$  is a standard Gaussian White noise, and the length equals to the above numerical solutions. This procedure is completed by the ARMA2AR and VARMA function of MATLAB2020b.

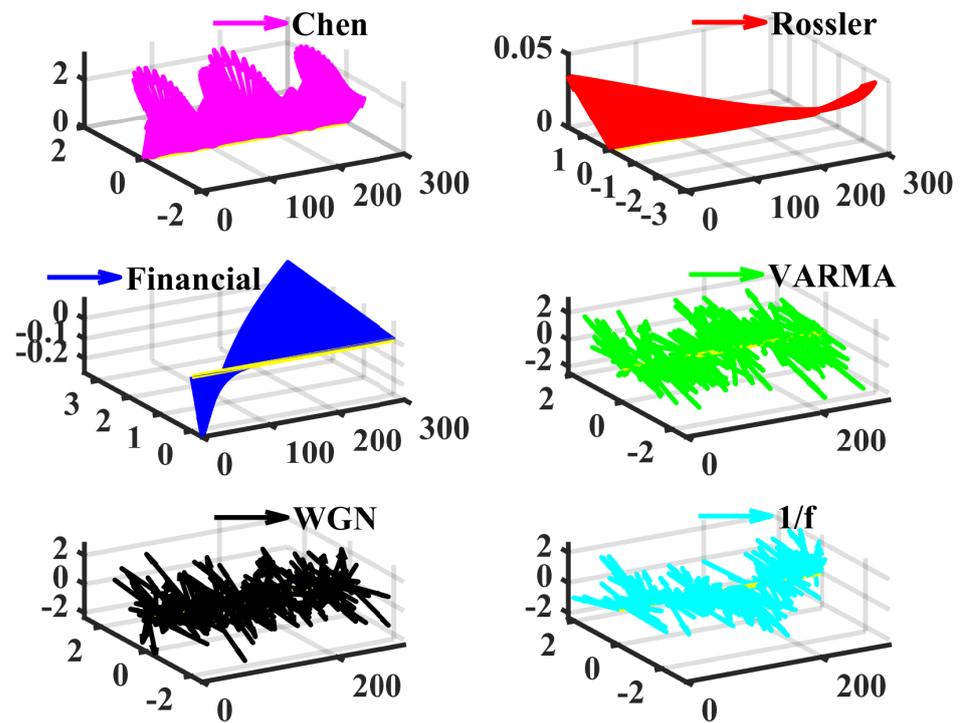
5. White Gaussian noise

We use the NORMRND function of MATLAB2020b to simulate WGN (500 × 3). Its components are independent with zero mean and unit standard deviation.

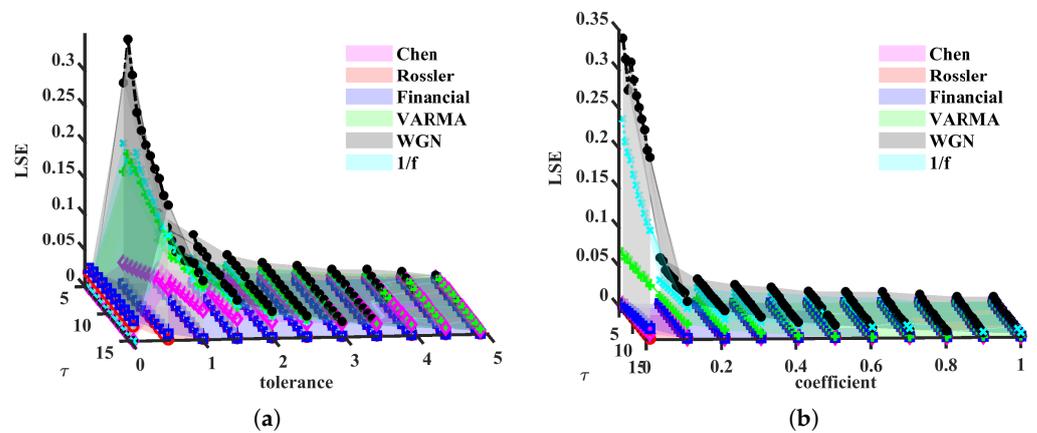
6. 1/f noise

On the basis of the algorithm in [57], the procedure of generating 1/f noise consists of three basic steps: (i) simulating white noise whose length is 1500, we obtain  $y_0$ ; (ii) DFT (discrete Fourier transformation) on  $y_0$ , multiplied by  $f^{-\frac{1}{2}}$  and symmetrized for a real function, then IDFT (inverse discrete Fourier transformation), adjust the mean and standard deviations, yielding  $y_1$ ; (iii) resize  $y_0$  and  $y_1$  to 500 × 3 matrix; finally, the 1/f noise series is composed as  $[y_0(:, 1), y_1(:, 2), \frac{1}{2}(y_0(:, 3) + y_1(:, 3))]$ .

Now, we set the simulation parameters as follows:  $\tau \in \{5, 6, \dots, 15\}, r \in \{0.05 : 0.041 : 0.5\} \times \sqrt{24}$  for METHOD A,  $r \in \{0.05 : 0.041 : 0.5\} \times \Sigma$  for METHOD B, where  $\Sigma = \sum_{i=1}^3 \sigma_i^4$  and  $\sigma_i$  is the standard deviation of  $i$ th component. More details can be found in Appendix A.2. LSE values are shown in Figure 3.



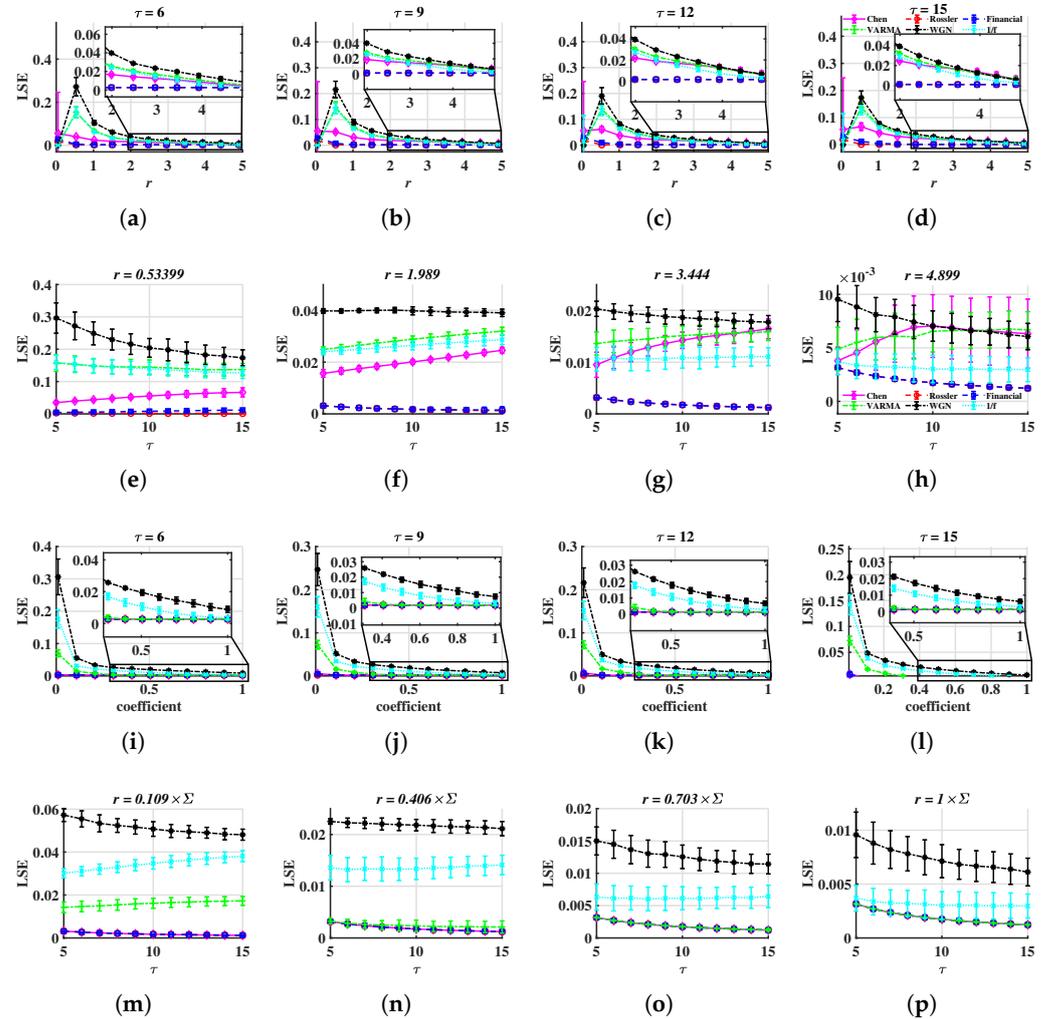
**Figure 2.** Examples of the chaotic and stochastic series. In fact, we view the multivariate time as a vector sequence. The first axis is the time,  $(t,0,0)$  is regarded as the original point for plotting the vectors  $(x(t),y(t),z(t))$ . The length is 300, and the time step is 0.005. Note that the series of Rössler and the financial systems appear to be very simple variations.



**Figure 3.** LSE(Local structure based entropy) values, as functions of  $\tau$  and  $r$ . (a) METHOD A, (b) METHOD B.

From Figure 3, we found both METHOD A and METHOD B can distinguish multivariate chaotic signals from stochastic ones. For a fixed  $(\tau, r)$ , the LSE of random time series is higher than that of a chaotic one. As shown in Figure 3a, WGN has the most complex signals, the  $1/f$  noise and VARMA series appear to have similar trends, and the other three chaotic series lie under 0.1. They keep the same orders for most  $(\tau, r)$ . In Figure 3b, LSE values are significantly different and hardly overlap, except for Chen and Rössler systems. The reason these two systems are indistinguishable can be because the range of coefficients of  $r$  is rather small. Next, we change the lower and upper bounds to 0.01 and 1. Moreover, the above simulation will be repeated 100 times (100 trajectories for each system) to check the robustness of LSE. In each simulation, we add small disturbances to initial values ( $Disturb_{initial} = 0.3 \times \epsilon$ ). Time series generated from Equations (11) to (13) are still chaos

with these disturbances; refer to [56]. Note that the intersections of LSE surfaces and  $\tau = \tau_i$  (or  $r = r_i$ ), other than surfaces themselves, are utilized to verify the validity of our methods. Subjecting to the space, only part of the results are demonstrated, see Figures 4 and 6.



**Figure 4.** LSE values, as functions of  $\tau$  or  $r$ . The time step is 0.005 for Equations (11) to (13). The marker presents the average and the error bar presents the standard deviation of LSE for fixed  $\tau$  and  $r$ . (a–h) METHOD A, (i–p) METHOD B.

In Figure 4, the LSE of random series is higher than that of chaotic ones, thus we believe LSE can be treated as an efficient tool to characterize multivariate time series. When  $\tau$  is fixed, LSE values of different dynamics show their own features. In Figure 4a–d, LSE values of chaotic systems show similar trends where  $LSE_A^{chaos}$  oscillates heavily when  $r = r_1 = 0.01 \times \sqrt{24}$ . Then,  $LSE_A^{chaos}$  decreases as tolerance increases. For  $LSE_A^{random}$ , they reach their highest values at  $r_2$  and then keep reducing as  $r$  increases. However, for METHOD B, all LSE values are decreasing, see Figure 4i–l. When  $r$  is fixed, except for the Rössler and Financial systems, other series can be distinguished by LSE values over different scales. For instance, in Figure 4e,  $LSE_A^{chen} < 0.08$  but  $LSE_A^{random} > 0.1$ , in Figure 4f,  $LSE_A^{WGN}$  lie at the highest level, and Figure 4g,h,m,n,  $LSE_A^{1/f}$  locate higher than  $LSE_A^{VARMMA}$ .

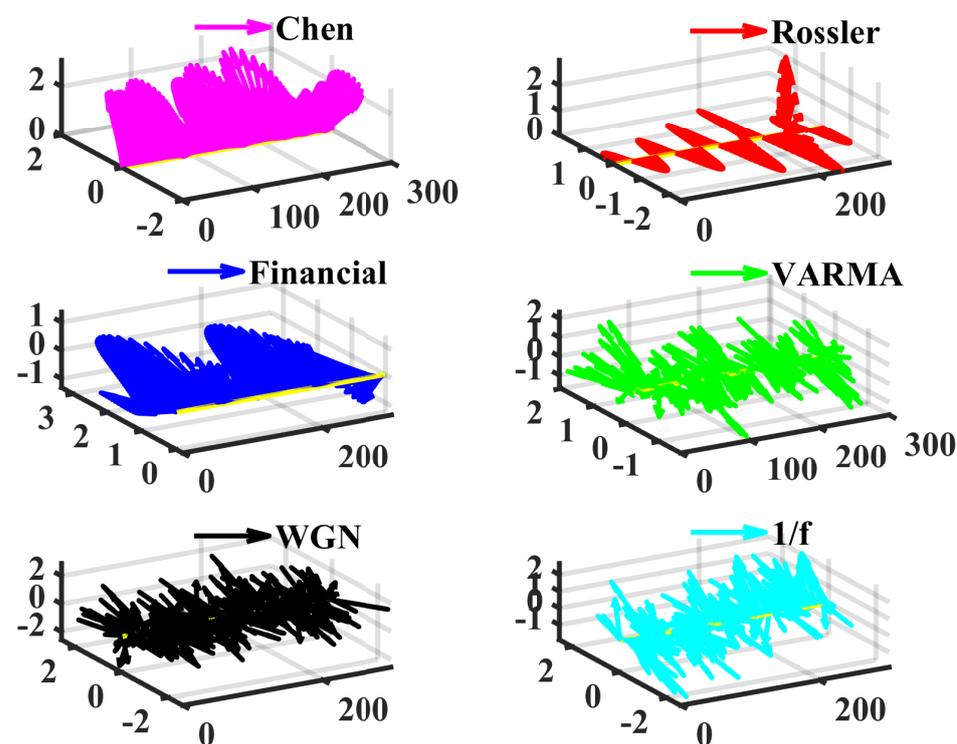
Compare METHOD A with METHOD B, when  $r$  is small, the later performs better at distinguishing multivariate stochastic series from chaotic ones, as well as more stable LSE values for chaos (basically the same). For instance, all  $LSE_A^{chaos}$  are overlapped and smaller than  $LSE_A^{random}$  when the coefficient is 0.109 in Figure 4m. However, in Figure 4e,f,  $LSE_A^{chen}$  is close to  $LSE_A^{random}$ . Moreover, the former shows better discrimination on chaotic

systems. This is because of the standardization process of METHOD A, which weights all components at an equal level. It may amplify the LSE if the system’s variation is caused by only one or a small group of components.

Based on the simulation findings mentioned above, the following reference range of parameters can be provided:

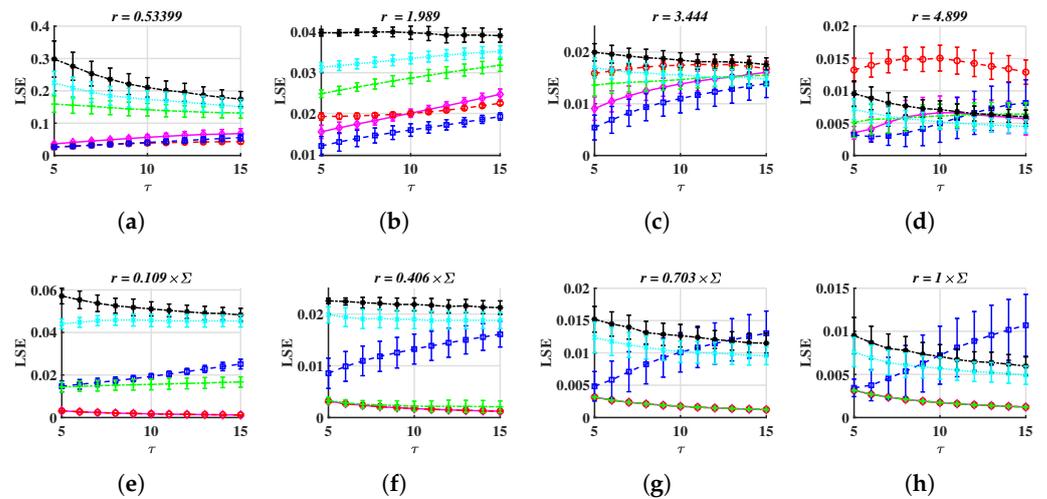
- $5 \leq \tau \leq 15$  is acceptable. In Figure 4a–d and in Figure 4i–l, when  $r$  is taken appropriately, these curves have essentially the same shape, making it possible to distinguish between different signals. This suggests that the distinction is not dependent on the value of  $\tau$ , and the corresponding time complexity can be taken into consideration when designing this parameter.
- For METHOD A,  $0.5 \leq r \leq 2$ . For METHOD B,  $0.1 \leq coefficient \leq 0.4$ . For example, when  $r = 0.53399$ ,  $LSE^{chaos}$  is less than 0.1 but  $LSE^{random}$  is higher than 0.1. Therefore, the stochastic sequence can be distinguished from chaotic ones. See Figure 4e. If  $r$  is too small,  $LSE$  has a large standard deviation, as is shown in Figure 4a,i. If  $r$  is too large,  $LSE^{chaos}$  and  $LSE^{random}$  are overlapped, thus, it is difficult to tell apart various signals. See Figure 4g,o.

However, one problem is that  $LSE^{rossler}$  and  $LSE^{financial}$  cannot be distinguished, as they are coincident in Figure 4. The reason for this is that the simulated series of Equations (12) and (13) is short and the time step is too small, thus they cannot reflect the feature of the whole system. So, we change the time step to 0.05 and the time span to [0:0.05:300] for the above two systems in order to verify whether LSE functions or not. Examples of the above six kinds of signals are shown in Figure 5.



**Figure 5.** Examples of chaotic and stochastic series. The time step is 0.005 for the Chen system, and 0.05 for the Rössler and financial systems. Compare to Figure 2, the Rössler and financial series show significant oscillations.

After repeating 100 times, similar to what we have done in previous tasks, we perceive that LSE works for many tolerance values and some results are drawn in Figure 6. Here, we just show several LSE- $\tau$  figures for fixed tolerances or coefficients, such as Figure 4e,n, because LSE- $r$  charts appear too similar to disparate different series.



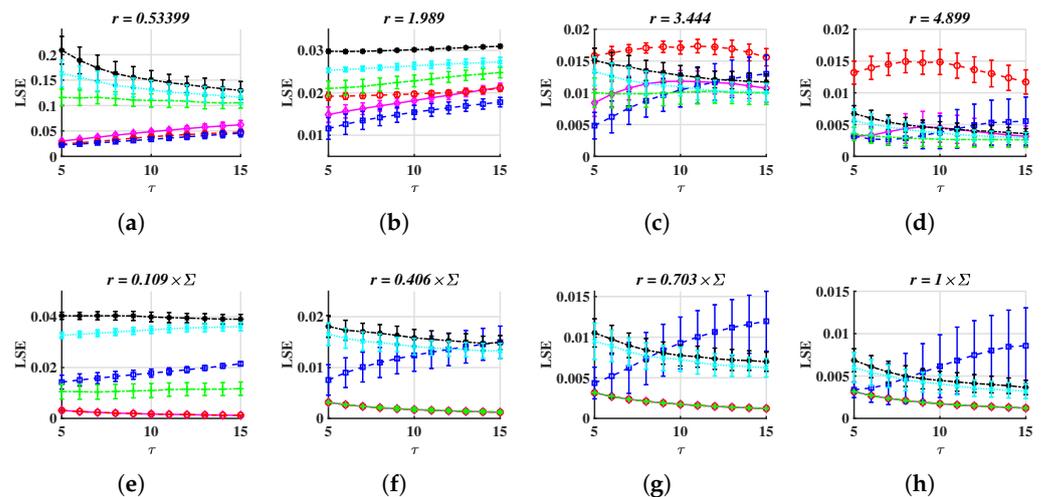
**Figure 6.** LSE values, as functions of  $\tau$  for several fixed tolerances or coefficients. The time step is 0.05 for the Rössler and Financial systems and 0.005 for the Chen system. (a–d) METHOD A, (e–h) METHOD B.

From Figure 6, one can easily distinguish six series. For example, in Figure 6d,  $LSE_A^{rossler}$  lies at the top, in Figure 6b, the LSE values of the other five series do not intersect, and the order is WGN, 1/f noise, VARMA, Chen, and financial, from the top to the bottom. However, the coefficient should be small, such as it is around 0.4 (Figure 6b,f). Otherwise, LSE curves are overlapped (Figure 6d,h).

Moreover, to test the dependence on the distance function, we replace Equation (3) with the distance derived from the  $l_\infty$  norm, i.e.,

$$dis_j = |y_0^i, y_{1+j}^i|_\infty = \max |y_0^i - y_{1+j}^i| \tag{15}$$

Repeat the above simulations, the corresponding results are drawn in Figure 7.



**Figure 7.** LSE values, where the distance function is derived from the infinite norm, Equation (15). (a–d) METHOD A, (e–h) METHOD B.

Compare Figure 7 to Figure 6, one can see that LSE with the same range of parameters can still effectively distinguish between different signals when using the  $l_\infty$  norm-derived distance, as Figure 7a,b,f show. Meanwhile, the range of  $r$  should be carefully chosen, or the LSE curve may be overlapped, such as Figure 7d,h.

To compare with the multivariate permutation entropy (PE) [17,58], Figure 8 shows PE curves for the above six series.

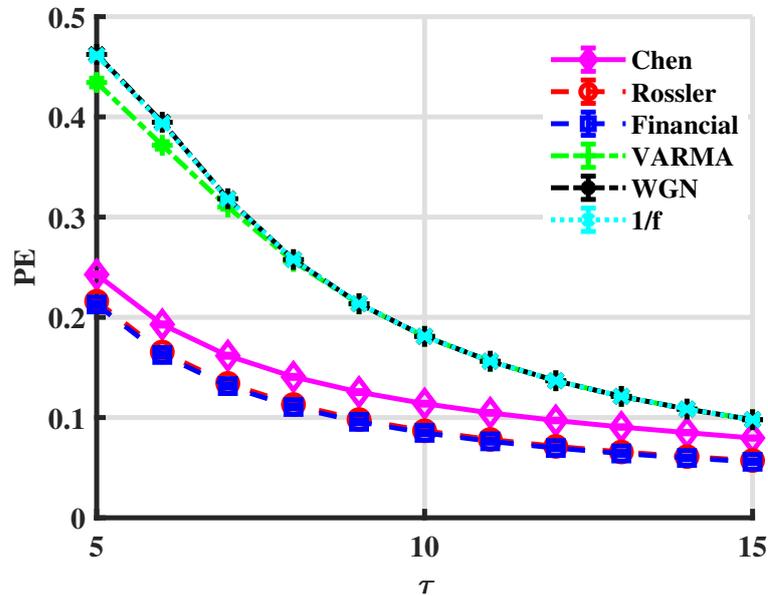


Figure 8. PE values. The embedding dimensions are {5, 6, . . . , 15}.

Chaotic signals can be distinguished from stochastic ones via PE since  $PE^{chaos} < PE^{random}$  for all  $\tau$  in Figure 8. However, PE cannot discriminate random series, which have correlations in different degrees. Moreover, it fails in the Rössler and financial cases. It means that PE is not sensitive to tiny differences in the complexity of systems. However, LSE can distinguish those systems, which indicates LSE is a more accurate measure of complexity.

3.2. Integer-Order Chaos, Fractional-Order Chaos, and Random Series

In this subsection, we test LSE on the integer-order chaotic systems, fractional-order chaotic systems, and stochastic sequences. For integer-order chaos, we use Chen, Rössler, and financial systems. The timespans are 0:0:005:30, 0:0.05:300, and 0:0.05:300, respectively. In each simulation, small disturbances are added to the initial values. See Equations (16) to (18).

1. Integer-order Chen system [52]

$$\begin{cases} \dot{x} = a(y - x) \\ \dot{y} = dx - xz + cy \\ \dot{z} = xy - bz \end{cases} \tag{16}$$

$a = 35, b = 3, c = 28, d = c - a = -7$ , initial values  $(x_0, y_0, z_0) = (-9, -5, 14) + \epsilon$ .

2. Integer-order Rössler system [53]

$$\begin{cases} \dot{x} = y - z \\ \dot{y} = x + ay \\ \dot{z} = b + z(x - c) \end{cases} \tag{17}$$

$a = 0.4, b = 2, c = 4, (x_0, y_0, z_0) = (0.5, 1.5, 0.1) + \epsilon$ .

3. Integer-order financial system [54]

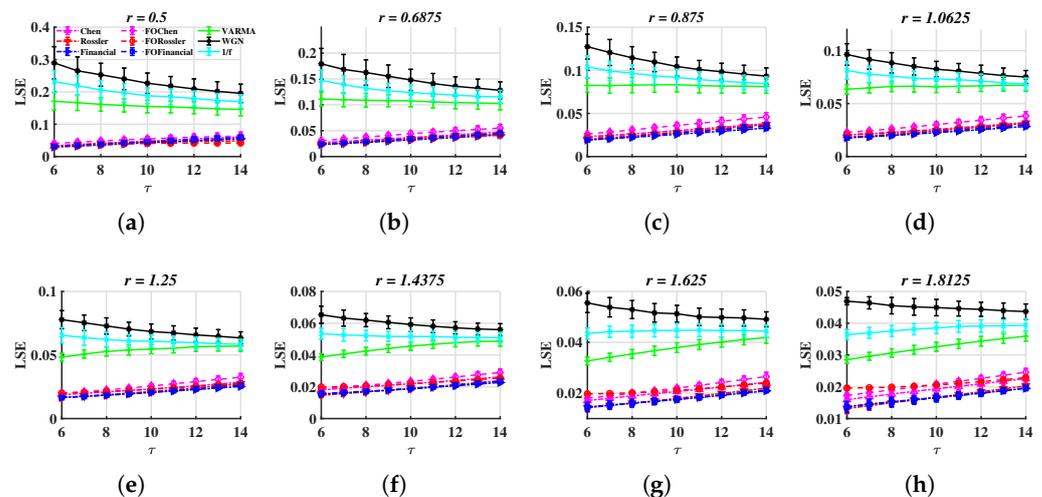
$$\begin{cases} \dot{x} = z + (y - a)x \\ \dot{y} = 1 - by - x^2 \\ \dot{z} = -x - cz \end{cases} \tag{18}$$

$$a = 1, b = 0.1, c = 1, (x_0, y_0, z_0) = (2, -1, 1) + \epsilon.$$

With the given parameters, the above three systems produce three-dimensional chaotic series.

For fractional-order chaos, Equations (11) to (13) are used to generate chaotic time series, such as in Section 3.1. Additionally, we add a small disturbance ( $Disturb_{\alpha} = \pm 0.02 \times \epsilon$ ) to parameters ( $\alpha_1, \alpha_2,$  and  $\alpha_3$ ), in each simulation. Moreover, we use the same methods as in Section 3.1 to generate random signals. All the simulated series have the same size ( $500 \times 3$ ).

Since we aim to illustrate the effectiveness of LSE, here, we only employ METHOD A to distinguish between different types of signals. Similar to Section 3.1, we also repeat the simulation procedure 100 times;  $\tau \in \{6, 7, \dots, 14\}$  and  $r \in [0.5, 1.81]$ . The results are drawn in Figure 9.

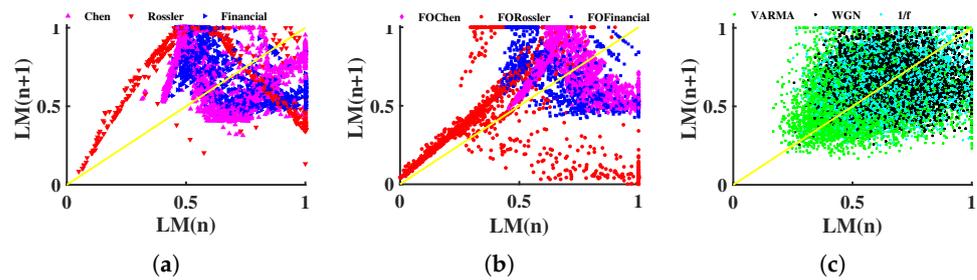


**Figure 9.** LSE values with  $\tau \in \{6, 7, \dots, 14\}$  and different tolerances. (a–d)  $r = 0.5, 0.6875, 0.8750, 1.0625$ , (e–h)  $r = 1.2500, 1.4375, 1.6250, 1.8125$ .

In Figure 9, chaotic signals can be distinguished from random ones. For instance,  $LSE^{random}$  is greater than 0.1 but  $LSE^{chaos}$  is less than 0.1 in Figure 9a. A similar situation can be seen in the images corresponding to other  $r$  values.

The return map (RM) can be used to characterize the nonlinear process by the transformation of the local maxima (or minima) of the signal [59–62]. We tested RM on the length sequences associated with the above nine multivariate time series and the results are drawn in Figure 10.

The difference between the Rössler system and the other signals is relatively obvious in both the integer-order and fractional-order cases, as shown in Figure 10a,b. However, it is difficult to distinguish between the Chen system and financial system. In addition, both the Chen and Financial systems are mostly located in  $[0.5, 1] \times [0.5, 1]$ , overlapping with the position where random signals are located, see Figure 10c. Thus, Figure 10 shows that several signals are muddled and difficult to differentiate apart. This failure could be related to the time series’ length, since RM commonly calls for a sizable amount of data points.

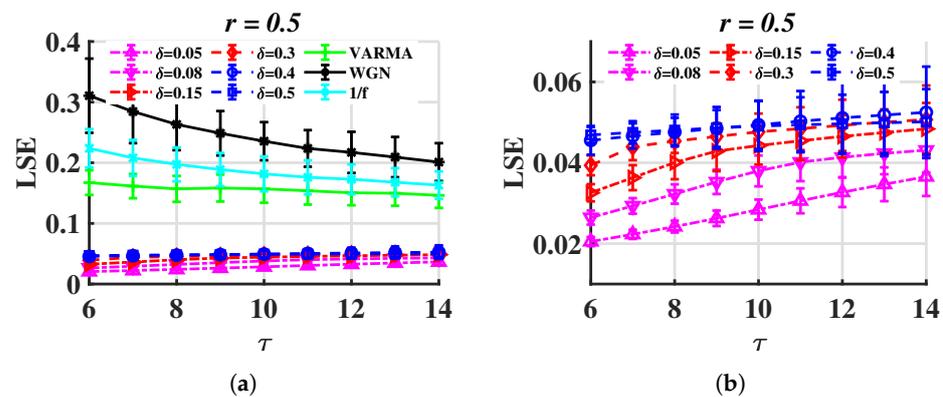


**Figure 10.** Return map. LM presents the local maxima of each length sequence, which is mapped into [0, 1] by the min–max method. (a) Integer-order chaotic systems, (b) fractional-order chaotic systems, (c) random signals.

Three shortcomings noted in Section 1 have been somewhat mitigated by the LSE method.

1. Multi-scale entropy and complex networks often necessitate a substantial number of data nodes to obtain useful findings, but LSE can process relatively short time series. The length of the test cases in this section is only 500.
2. According to the calculation process in Section 2, it is simple to know that the time complexity of the LSE method is linear ( $O(\tau n)$ ) and appropriate for handling real-time jobs.
3. LSE is more effective at depicting the short-term autocorrelation and the correlation between various components of multidimensional time series because it takes advantage of the similarity of vectors in sliding windows.

In Section 3.1, the time steps for Rössler and financial systems are 0.005 and 0.05, and the LSE values vary remarkably. Here, we test LSE on the Rössler system with different steps. The results are shown in Figure 11. It can be seen that the increase in the step length causes LSE to increase, but the gap between  $LSE^{rössler}$  and  $LSE^{random}$  is still clear.



**Figure 11.** LSE for trajectories of the Rössler system with different time step sizes ( $\delta$ ). (a)  $LSE^{rössler}$  for different time step sizes and  $LSE^{random}$ , (b)  $LSE^{rössler}$  only.

In order to assess the robustness of the proposed method, the original signal is supplemented with Gaussian white noise at different levels, and the LSE values are recorded accordingly. The original time series is produced by the chaotic Rössler system. The signal-to-noise ratio (SNR) is used to measure the level of background noise, which is defined by

$$SNR = 10 \times \log_{10}\left(\frac{P_S}{P_N}\right) \tag{19}$$

where  $P_S$  is the power of the signal and  $P_N$  is that of noise. See Figure 12.

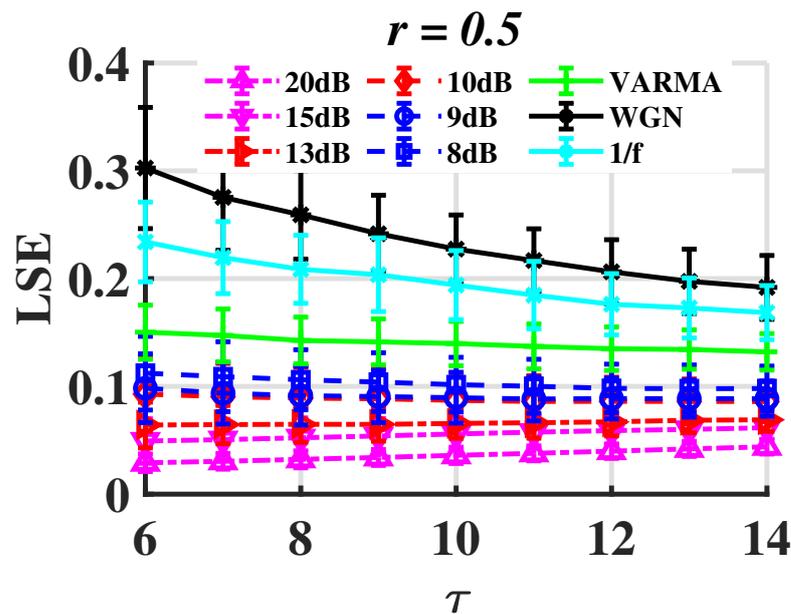


Figure 12. LSE of the Rössler system with WGN.

Noise tolerance is present in LSE to some extent. The  $LSE^{Rössler}$  is less than 0.1 when the SNR is higher than 10 dB. As a result, the random signals differ significantly. LSE, however, is unable to properly discriminate between chaotic and random signals as SNR continues to drop.

The above analysis indicates that the proposed LSE, based on its variations in different scales for certain tolerances (or coefficients), can be regarded as an efficient tool to identify multivariate time series. In the next section, we attempt to apply LSE to real-world data, such as the financial market index and machinery data.

#### 4. Application on Real-World Data

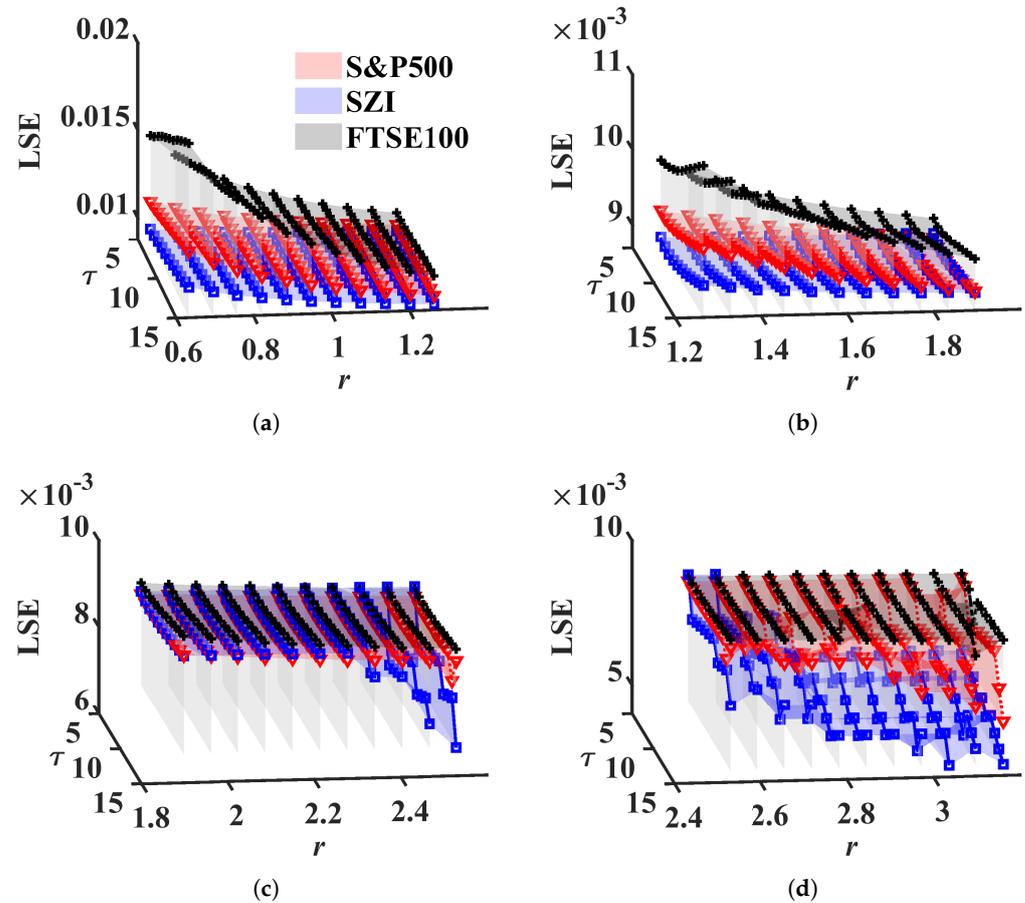
As we introduced in Section 1, multivariate time series conceal the characters in the autocorrelation of each component, as well as the cross-correlation between some channels, which makes it difficult to extract suitable features for further exploration. Two real-world applications, one for the financial market and another for fault diagnosing, will be discussed below.

##### 4.1. Financial Market Index

Financial time series are typical signals with high complexities. How do we illustrate the discrimination between financial markets in different regions? Here, we use LSE to quantify the complexities of three important indices, S&P500, FTSE100, and Shenzhen Securities Component Index (SZI). Their values can be attained from Yahoo.com [63]. The time period is between 25/06/2017 and 24/06/2022. In this experiment, daily OHLC (open, high, low, close) prices and volume are considered. In order to avoid the result being manipulated by the volume only, METHOD A is employed to explore the complexity of the three indices, because volume values have much higher standard deviations than OHLC prices. For more details, see Table 1. The scales are {5, 6, . . . , 15} and the tolerance is between 0.63 and 3.16. The results are drawn in Figure 13.

Table 1. Financial market index information.

Index	Length	Region
S&P500	1256	American
SZI	1220	China
FTSE 100	1263	Europe



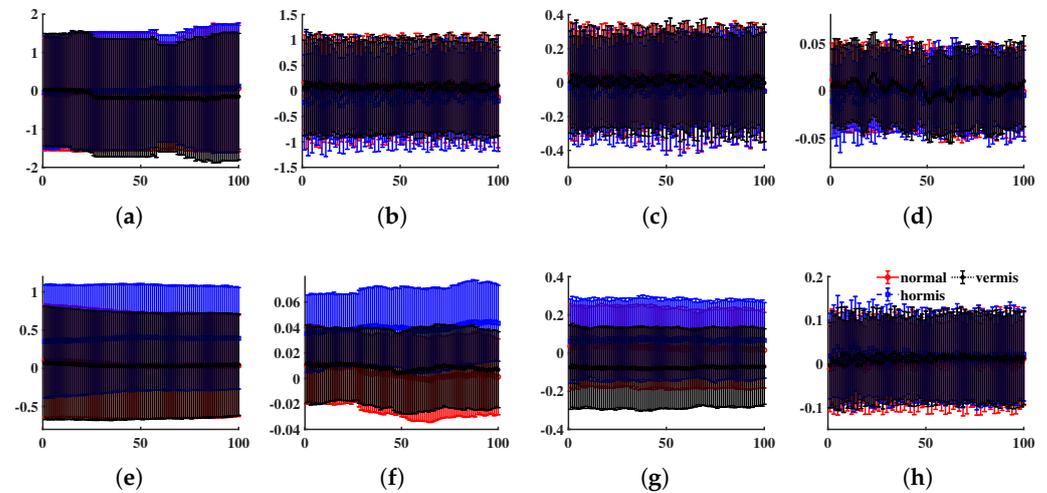
**Figure 13.** LSE values for stock market index. (a)  $r = [0.6325 : 0.6325 : 1.2649]$  (Start:time step:End), (b)  $r = [1.2649 : 0.6325 : 1.8974]$ , (c)  $r = [1.8974 : 0.6325 : 2.5298]$ , (d)  $r = [2.5298 : 0.6325 : 3.1623]$ .

From Figure 13, we can see that  $LSE^{FTSE100}$  lies at the top,  $LSE^{SZI}$  at the bottom, and  $LSE^{S\&P500}$  is between them when  $r < 1.8$ . That is, the European stock market shows higher complexity than the other two markets; mature financial markets, European, and American stock markets display more complexity and stability than the Chinese stock market, which is in accordance with Cao [64]. Note that  $LSE^{SZI}$  lies between  $LSE^{FTSE100}$  and  $LSE^{S\&P500}$  for  $1.8 < r < 2.4$ . It contradicts the above result. The reason for this can be attributed to the value of  $r$ . We are aware that the range  $r$  is determined by multi-dimensional WGN, but since the financial market index actually has a significant correlation, the tolerance should be lower than usual.

#### 4.2. Machinery Fault Recognition

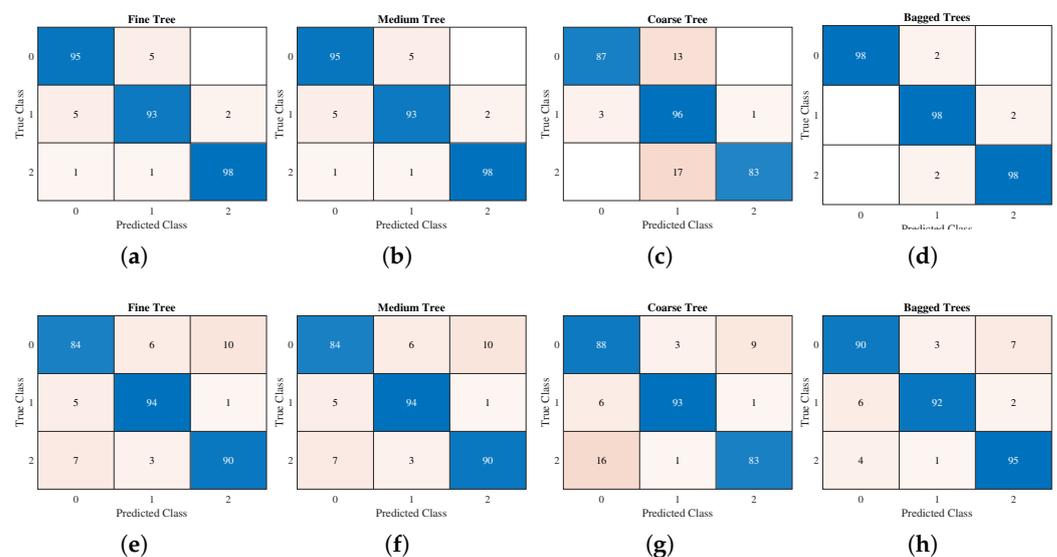
In this part, we examine the ability of the LSE to recognize vibration signals produced by normal or faulty mechanical systems. The bearing dataset is from the machinery fault database (MAFAULDA), which is kindly provided by the Signals Multimedia and Telecommunications Laboratory (SMT) of the Federal University of Rio de Janeiro (UFRJ) [65]. MAFAULDA collects multivariate time series recorded by sensors on a SpectraQuest’s machinery fault simulator (MFS) alignment–balance–vibration (ABVT) and comprises six different simulated states. We tested LSE for three states: normal function as well as horizontal and vertical misalignment faults. The data acquisition system is composed of several sensors: one Monarch Instrument MT-190 analog tachometer, three Industrial IMI Sensors accelerometers (Model 601A01), one IMI sensors triaxial accelerometer (model 604b31), and a Shure SM81 microphone. Each sequence has 8 columns sampled at 50 KHz for 5 s, namely a  $250,000 \times 8$  matrix. We randomly intercept 100 fragments (each has 3000 rows)

for every state from the database as the test dataset. Figure 14 shows the distribution of each component of the test data (first 100 rows).



**Figure 14.** The average and standard deviation of each component (first 100 observations), where the x-axis represents time and the y-axis the readings. (a) Tachometer signal, (b–d) underhang bearing accelerometer: axial, radial, and tangential directions, (e–g) overhang bearing accelerometer: axial, radial, and tangential directions, (h) microphone.

Then, we employed METHOD A to compute the LSE values. The scale set is  $\{5, 6, \dots, 15\}$  and the tolerance is 0.2457. Moreover, the decision tree method is applied to classify LSE values of different states. A decision support tool known as a decision tree employs a tree-like paradigm to represent options and their outcomes. It consists of nodes, branches, and leaves, each of which displays a property, a rule, or an outcome [66]. In this study, three alternative classification algorithms based on the decision tree theory, fine tree (about 100 leaves make fine distinctions between classes), medium tree (less than 20 leaves with medium flexibility), and coarse tree (less than 4 splits), were used to categorize the vibration signals [67]. In addition, an ensemble bagging tree classifier was added, which is processed by creating numerous decision trees during training and outputs the majority of these tree choices for classification tasks [68]. The ten-fold cross-validation was employed in this test and the confusion matrix was plotted in Figure 15a–d. Nevertheless, if we replace LSE by PE, the accuracy is significantly lower; see Figure 15e–h.



**Figure 15.** Confusion matrix. Class0–normal, class1–horizontal misalignment, class2–vertical misalignment. (a–d) LSE, (e–h) PE.

From Figure 15, bagged trees based on LSE had the highest accuracy at 98%, while that of PE was 92.3%. Moreover, the accuracies of the LSE-based fine tree, medium tree, and coarse tree (95.3%, 95.3%, and 88.7%) were higher than those of PE (89.3%, 89.3%, 88%). In short, as the LSE values (as features) contribute higher accuracies than PE, the proposed LSE can be an efficient tool for distinguishing multivariate real-world time series.

## 5. Conclusions

In this paper, we proposed a local structure-based entropy (LSE), which reflects recurrence conditions in certain scales. It can be regarded as an index of complexity for multivariate time series. Depending on whether or not the components are normalized, we suggest two strategies for using LSE: one shows greater discrimination while the other is more stable but easily ignores the effects of slightly varying components. When the tolerance is small, LSE values of fractional chaotic time series are significantly lower than those of stochastic ones. Moreover, the LSE method also has some resilience to noise. When the SNR is higher than 10 dB, accurate classification is obtained for the task of differentiating chaotic signals from random ones, but the accuracy declines when SNR drops. With suitable tolerance, LSE (in certain scales) can be considered a feature of a dynamical system. Regarding real-world data, it was applied to a financial market index, indicating that European and American financial markets are more complex and stable than Chinese markets. Furthermore, we tested LSE on MAFAULDA; it resulted in a higher accuracy than PE-based classification.

The results are sensitive to the parameters (especially for  $r$ ), but the optimal range is provided by the simulation other than the theoretical calculation. Therefore, a more comprehensive examination of the parameters and LSE values based on basic information, such as dimensionality, correlation, and time series length, is required. A distance function that can eliminate the impact of the correlation can also be utilized to increase the applicability of LSE.

**Author Contributions:** Conceptualization, methodology, software, writing—original draft preparation, Z.Z.; validation, visualization, Z.Z. and J.W. (Jun Wu); resources, J.X. and J.W. (Ji Wang); writing—review and editing, J.X.; supervision, J.X.; project administration, Y.C.; funding acquisition, Y.C. and J.W. (Jun Wu). All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the Doctoral Fund of Hubei University of Automotive Technology (grant no. BK201703) and the Hubei Key Laboratory of Applied Mathematics (grant no. HBAM202105).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** MAFAULDA, in Section 4.2, is available at [http://www02.smt.ufjf.br/~offshore/mfs/page\\_01.html](http://www02.smt.ufjf.br/~offshore/mfs/page_01.html), and we have accessed on 15 July 2022.

**Acknowledgments:** We thank Ivo Petras and Gaoxiang Ouyang for the coding fractional-order chaotic system and PE algorithm, as they kindly shared their works on the MATLAB Central File Exchange.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

### Appendix A.1

In NOTE 1, we suggest that the range of tolerance  $r$  can be chosen by referring to the standard normal distribution. In fact, we adopt this strategy in METHOD A. For two  $n$ -dimensional random variables whose components are  $N(0,1)$  iid, Proposition A1 shows the variance of the square of the Euclidean distance is  $8n$ .

**Proposition A1.** Assume  $x_i \sim N(0, 1)$  and  $y_i \sim N(0, 1)$  are iid, let  $SS = \sum_{i=1}^n (x_i - y_i)^2$ , then  $D(SS) = 8n$ .

**Proof.** Let  $z_i = \frac{x_i - y_i}{\sqrt{2}}$ , then  $z_i \sim N(0, 1)$ .

$S = \sum_{i=1}^n (\frac{x_i - y_i}{\sqrt{2}})^2 = \sum_{i=1}^n (z_i)^2 \sim \chi^2(n)$ , then  $D(S) = 2n$ .

Since  $S = \frac{1}{2}SS$ , we have  $D(SS) = D(2S) = 8n$ , thus  $\sigma(SS) = 2\sqrt{2n}$ .  $\square$

For **METHOD A** in Section 3, the scale is  $\tau \in [5 : 1 : 15]$ ; thus, the set of  $r$ , denoted by  $R$ , is defined as

$$R = [0.05 : \frac{0.5 - 0.05}{11} : 0.5] \times \sqrt{8 \times n} \quad (\text{A1})$$

$R$  can be regarded as a sequence of threshold values multiplied by a constant. Each element of this sequence ( $[\cdot : \cdot : \cdot]$  part) is called a coefficient.

The standardization is required for **METHOD A**, but it can shuffle the positions of the observed data. To overcome this drawback, **METHOD B** is introduced. Instead of standardization, different ranges of  $r$  are utilized to fit different time series. These ranges are selected according to the variance of the observed time series. In fact,  $R$  is defined as

$$R = [0.05 : \frac{0.5 - 0.05}{11} : 0.5] \times \Sigma \quad (\text{A2})$$

where  $\Sigma$  is introduced in Proposition A2.

#### Appendix A.2

**Proposition A2.** Let  $\mathbf{x} = (x_i)_{i \leq n}$ , where  $x_i \sim N(\mu_i, \sigma_i^2)$ , be a  $n$ -dimensional random variable and assume its components are independent;  $\mathbf{x}$  and  $\mathbf{y}$  are iid. Denote  $SS = \sum_{i=1}^n (x_i - y_i)^2$ , then  $\Sigma = D(SS) = 8 \sum_{i=1}^n \sigma_i^4$ .

**Proof.** Assume  $X \sim N(\mu, \sigma^2)$ . It is easy to know that  $E(X^2) = \sigma^2 + \mu^2$  and  $E(X^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$ .

Then,

$$\begin{aligned} D(X^2) &= E[(X^2)^2] - [E(X^2)]^2 \\ &= \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 - (\sigma^2 + \mu^2)^2 \\ &= 4\mu^2\sigma^2 + 2\sigma^4 \end{aligned} \quad (\text{A3})$$

Note that  $(x_i - y_i) \sim N(0, 2\sigma_i^2)$ , by Equation (A3) we have  $D[(x_i - y_i)^2] = 8\sigma_i^4$ . Then,

$$\begin{aligned} D(SS) &= D\left(\sum_{i=1}^n (x_i - y_i)^2\right) \\ &= \sum_{i=1}^n D[(x_i - y_i)^2] \\ &= 8 \sum_{i=1}^n \sigma_i^4 \end{aligned} \quad (\text{A4})$$

$\square$

## References

1. Sulistiyono, S.; Akhiruyanto, A.; Primasoni, N.; Arjuna, F.; Santoso, N.; Yudhistira, D. The effect of 10 weeks game experience learning (gel) based training on teamwork, respect attitude, skill and physical ability in young football players. *Teoriâ ta Metod. Fizičnogo Vihovannâ* **2021**, *21*, 173–179. [[CrossRef](#)]
2. Follesa, M.; Fragiaco, M.; Casagrande, D.; Tomasi, R.; Piazza, M.; Vassallo, D.; Canetti, D.; Rossi, S. The new provisions for the seismic design of timber buildings in Europe. *Eng. Struct.* **2018**, *168*, 736–747. [[CrossRef](#)]
3. Gao, T.; Fadnis, K.; Campbell, M. Local-to-global Bayesian network structure learning. In Proceedings of the International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; pp. 1193–1202.

4. Friedman, N.; Goldszmidt, M. Learning Bayesian networks with local structure. In *Learning in Graphical Models*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 421–459.
5. Lei, C.; Zhu, X. Unsupervised feature selection via local structure learning and sparse learning. *Multimed. Tools Appl.* **2018**, *77*, 29605–29622. [[CrossRef](#)]
6. Li, J.; Wen, G.; Gan, J.; Zhang, L.; Zhang, S. Sparse nonlinear feature selection algorithm via local structure learning. *Emerg. Sci. J.* **2019**, *3*, 115–129. [[CrossRef](#)]
7. Liao, S.; Yi, D.; Lei, Z.; Qin, R.; Li, S.Z. Heterogeneous face recognition from local structures of normalized appearance. In *Proceedings of the International Conference on Biometrics*, Alghero, Italy, 2–5 June 2009; Springer: Berlin/Heidelberg, Germany, 2009; pp. 209–218.
8. Qian, J.; Yang, J.; Xu, Y. Local structure-based image decomposition for feature extraction with applications to face recognition. *IEEE Trans. Image Process.* **2013**, *22*, 3591–3603. [[CrossRef](#)] [[PubMed](#)]
9. Heo, L.; Feig, M. High-accuracy protein structures by combining machine-learning with physics-based refinement. *Proteins Struct. Funct. Bioinform.* **2020**, *88*, 637–642. [[CrossRef](#)] [[PubMed](#)]
10. Zhang, L.; Du, G.; Liu, F.; Tu, H.; Shu, X. Global-local multiple granularity learning for cross-modality visible-infrared person reidentification. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**. [[CrossRef](#)]
11. Shannon, C.E. A mathematical theory of communication, 1948. *Bell Syst. Tech. J.* **1948**, *27*, 3–55. [[CrossRef](#)]
12. Richman, J.S.; Lake, D.E.; Moorman, J.R. Sample entropy. In *Methods in Enzymology*; Elsevier: Amsterdam, The Netherlands, 2004; Volume 384, pp. 172–184.
13. Yentes, J.M.; Hunt, N.; Schmid, K.K.; Kaipust, J.P.; McGrath, D.; Stergiou, N. The appropriate use of approximate entropy and sample entropy with short data sets. *Ann. Biomed. Eng.* **2013**, *41*, 349–365. [[CrossRef](#)]
14. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of biological signals. *Phys. Rev. E* **2005**, *71*, 021906. [[CrossRef](#)]
15. Busa, M.A.; van Emmerik, R.E. Multiscale entropy: A tool for understanding the complexity of postural control. *J. Sport Health Sci.* **2016**, *5*, 44–51. [[CrossRef](#)] [[PubMed](#)]
16. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [[CrossRef](#)] [[PubMed](#)]
17. Keller, K.; Mangold, T.; Stolz, I.; Werner, J. Permutation entropy: New ideas and challenges. *Entropy* **2017**, *19*, 134. [[CrossRef](#)]
18. Zhang, Z.; Xiang, Z.; Chen, Y.; Xu, J. Fuzzy permutation entropy derived from a novel distance between segments of time series. *AIMS Math.* **2020**, *5*, 6244–6260. [[CrossRef](#)]
19. Morabito, F.C.; Labate, D.; La Foresta, F.; Bramanti, A.; Morabito, G.; Palamara, I. Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer’s disease EEG. *Entropy* **2012**, *14*, 1186–1202. [[CrossRef](#)]
20. He, S.; Sun, K.; Wang, H. Multivariate permutation entropy and its application for complexity analysis of chaotic systems. *Phys. A Stat. Mech. Its Appl.* **2016**, *461*, 812–823. [[CrossRef](#)]
21. Ying, W.; Tong, J.; Dong, Z.; Pan, H.; Liu, Q.; Zheng, J. Composite multivariate multi-Scale permutation entropy and laplacian score based fault diagnosis of rolling bearing. *Entropy* **2022**, *24*, 160. [[CrossRef](#)]
22. Romera, E.; Nagy, Á. Density functional fidelity susceptibility and Kullback–Leibler entropy. *Phys. Lett. A* **2013**, *377*, 3098–3101. [[CrossRef](#)]
23. Wang, J.; Zheng, N.; Chen, B.; Chen, P.; Chen, S.; Liu, Z.; Wang, F.Y.; Xi, B. Multivariate Correlation Entropy and Law Discovery in Large Data Sets. *IEEE Intell. Syst.* **2018**, *33*, 47–54. [[CrossRef](#)]
24. Yu, S.; Giraldo, L.G.S.; Jenssen, R.; Principe, J.C. Multivariate Extension of Matrix-Based Rényi’s  $\alpha$ -Order Entropy Functional. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *42*, 2960–2966. [[CrossRef](#)]
25. Wang, Z.; Shang, P. Generalized entropy plane based on multiscale weighted multivariate dispersion entropy for financial time series. *Chaos Solitons Fractals* **2021**, *142*, 110473. [[CrossRef](#)]
26. Wang, X.; Si, S.; Li, Y. Variational Embedding Multiscale Diversity Entropy for Fault Diagnosis of Large-Scale Machinery. *IEEE Trans. Ind. Electron.* **2022**, *69*, 3109–3119. [[CrossRef](#)]
27. Yin, Y.; Wang, X.; Li, Q.; Shang, P. Generalized multivariate multiscale sample entropy for detecting the complexity in complex systems. *Phys. A Stat. Mech. Its Appl.* **2020**, *545*, 123814. [[CrossRef](#)]
28. Berrett, T.B.; Samworth, R.J.; Yuan, M. Efficient multivariate entropy estimation via k-nearest neighbour distances. *Ann. Stat.* **2019**, *47*, 288–318. [[CrossRef](#)]
29. Azami, H.; Escudero, J. Refined composite multivariate generalized multiscale fuzzy entropy: A tool for complexity analysis of multichannel signals. *Phys. A Stat. Mech. Its Appl.* **2017**, *465*, 261–276. [[CrossRef](#)]
30. Han, Y.F.; Jin, N.D.; Zhai, L.S.; Ren, Y.Y.; He, Y.S. An investigation of oil–water two-phase flow instability using multivariate multi-scale weighted permutation entropy. *Phys. A Stat. Mech. Its Appl.* **2019**, *518*, 131–144. [[CrossRef](#)]
31. Mao, X.; Shang, P.; Li, Q. Multivariate multiscale complexity-entropy causality plane analysis for complex time series. *Nonlinear Dyn.* **2019**, *96*, 2449–2462. [[CrossRef](#)]
32. Shang, B.; Shang, P. Complexity analysis of multiscale multivariate time series based on entropy plane via vector visibility graph. *Nonlinear Dyn.* **2020**, *102*, 1881–1895. [[CrossRef](#)]
33. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [[CrossRef](#)]
34. Cong, J.; Liu, H. Approaching human language with complex networks. *Phys. Life Rev.* **2014**, *11*, 598–618. [[CrossRef](#)]
35. Ruths, J.; Ruths, D. Control profiles of complex networks. *Science* **2014**, *343*, 1373–1376. [[CrossRef](#)] [[PubMed](#)]

36. Albert, R.; Jeong, H.; Barabási, A.L. Error and attack tolerance of complex networks. *Nature* **2000**, *406*, 378–382. [[CrossRef](#)] [[PubMed](#)]
37. Donges, J.F.; Zou, Y.; Marwan, N.; Kurths, J. Complex networks in climate dynamics. *Eur. Phys. J. Spec. Top.* **2009**, *174*, 157–179. [[CrossRef](#)]
38. Mutua, S.; Gu, C.; Yang, H. Visibility graphlet approach to chaotic time series. *Chaos* **2016**, *26*, 053107. [[CrossRef](#)]
39. Marwan, N.; Kurths, J. Complex network based techniques to identify extreme events and (sudden) transitions in spatio-temporal systems. *Chaos* **2015**, *25*, 097609. [[CrossRef](#)]
40. Zhang, Z.; Xu, J.; Zhou, X. Mapping time series into complex networks based on equal probability division. *AIP Adv.* **2019**, *9*, 015017. [[CrossRef](#)]
41. Zhao, Y.; Peng, X.; Small, M. Reciprocal characterization from multivariate time series to multilayer complex networks. *Chaos* **2020**, *30*, 013137. [[CrossRef](#)]
42. Small, M.; Zhang, J.; Xu, X. Transforming time series into complex networks. *Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng.* **2009**, *5 LNICST*, 2078–2089. [[CrossRef](#)]
43. Silva, V.F.; Silva, M.E.; Ribeiro, P.; Silva, F. Time series analysis via network science: Concepts and algorithms. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2021**, *11*, e1404. [[CrossRef](#)]
44. Lacasa, L.; Luque, B.; Ballesteros, F.; Luque, J.; Nuno, J.C. From time series to complex networks: The visibility graph. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4972–4975. [[CrossRef](#)]
45. Lacasa, L.; Just, W. Visibility graphs and symbolic dynamics. *Phys. D Nonlinear Phenom.* **2018**, *374–375*, 35–44. [[CrossRef](#)]
46. Donner, R.V.; Small, M.; Donges, J.F.; Marwan, N.; Zou, Y.; Xiang, R.; Kurths, J. Recurrence-based time series analysis by means of complex network methods. *Int. J. Bifurc. Chaos* **2011**, *21*, 1019–1046. [[CrossRef](#)]
47. Donges, J.F.; Heitzig, J.; Donner, R.V.; Kurths, J. Analytical framework for recurrence network analysis of time series. *Phys. Rev. E* **2012**, *85*, 046105. [[CrossRef](#)] [[PubMed](#)]
48. Ruan, Y.; Donner, R.V.; Guan, S.; Zou, Y. Ordinal partition transition network based complexity measures for inferring coupling direction and delay from time series. *Chaos Interdiscip. J. Nonlinear Sci.* **2019**, *29*, 043111. [[CrossRef](#)]
49. Guo, H.; Zhang, J.Y.; Zou, Y.; Guan, S.G. Cross and joint ordinal partition transition networks for multivariate time series analysis. *Front. Phys.* **2018**, *13*, 130508. [[CrossRef](#)]
50. Donner, R.V.; Zou, Y.; Donges, J.F.; Marwan, N.; Kurths, J. Recurrence networks—A novel paradigm for nonlinear time series analysis. *New J. Phys.* **2010**, *12*, 033025. [[CrossRef](#)]
51. Cattani, C.; Srivastava, H.M.; Yang, X.J. *Fractional Dynamics*; Walter de Gruyter GmbH & Co KG: Berlin, Germany, 2015; pp. 166–173.
52. Li, C.; Chen, G. Chaos in the fractional order Chen system and its control. *Chaos Solitons Fractals* **2004**, *22*, 549–554. [[CrossRef](#)]
53. Zhang, W.; Zhou, S.; Li, H.; Zhu, H. Chaos in a fractional-order Rössler system. *Chaos Solitons Fractals* **2009**, *42*, 1684–1691. [[CrossRef](#)]
54. Chen, W.C. Nonlinear dynamics and chaos in a fractional-order financial system. *Chaos Solitons Fractals* **2008**, *36*, 1305–1314. [[CrossRef](#)]
55. Podlubny, I. An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications. *Math. Sci. Eng.* **1999**, *198*, 340.
56. Petráš, I. *Fractional-Order Nonlinear Systems: Modeling, Analysis and Simulation*; Springer Science & Business Media: New York, NY, USA, 2011; pp. 285–290.
57. Zhivomirov, H. A method for colored noise generation. *Rom. J. Acoust. Vib.* **2018**, *15*, 14–19.
58. Ouyang, G.; Li, J.; Liu, X.; Li, X. Dynamic characteristics of absence EEG recordings with multiscale permutation entropy analysis. *Epilepsy Res.* **2013**, *104*, 246–252. [[CrossRef](#)] [[PubMed](#)]
59. Mukherjee, S.; Zawar-Reza, P.; Sturman, A.; Mittal, A.K. Characterizing atmospheric surface layer turbulence using chaotic return map analysis. *Meteorol. Atmos. Phys.* **2013**, *122*, 185–197. [[CrossRef](#)]
60. Chidori, K.; Yamamoto, Y. Effects of the lateral amplitude and regularity of upper body fluctuation on step time variability evaluated using return map analysis. *PLoS ONE* **2017**, *12*, e0180898. [[CrossRef](#)]
61. Rybin, V.; Butusov, D.; Rodionova, E.; Karimov, T.; Ostrovskii, V.; Tutueva, A. Discovering chaos-based communications by recurrence quantification and quantified return map analyses. *Int. J. Bifurc. Chaos* **2022**, *32*, 2250136. [[CrossRef](#)]
62. Voznesensky, A.; Butusov, D.; Rybin, V.; Kaplun, D.; Karimov, T.; Nepomuceno, E. Denoising Chaotic Signals using Ensemble Intrinsic Time-Scale Decomposition. *IEEE Access* **2022**. Available online: <https://ieeexplore.ieee.org/abstract/document/9932609> (accessed on 20 November 2022).
63. Yahoo. Yahoo Finance. Available online: <https://www.yahoo.com/author/yahoo-finance> (accessed on 6 June 2022).
64. Cao, H.; Li, Y. Unraveling chaotic attractors by complex networks and measurements of stock market complexity. *Chaos Interdiscip. J. Nonlinear Sci.* **2014**, *24*, 013134. [[CrossRef](#)]
65. Ribeiro, F.M.L. MAFAULDA—Machinery Fault Database [Online], 2021. Available online: [http://www02.smt.ufrj.br/~offshore/mfs/page\\_01.html](http://www02.smt.ufrj.br/~offshore/mfs/page_01.html) (accessed on 15 July 2022).
66. Song, Y.Y.; Ying, L. Decision tree methods: Applications for classification and prediction. *Shanghai Arch. Psychiatry* **2015**, *27*, 130.

- 
67. Ren, G.; Wang, Y.; Ning, J.; Zhang, Z. Using near-infrared hyperspectral imaging with multiple decision tree methods to delineate black tea quality. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2020**, *237*, 118407. [[CrossRef](#)]
  68. Himeur, Y.; Alsalemi, A.; Bensaali, F.; Amira, A. Robust event-based non-intrusive appliance recognition using multi-scale wavelet packet tree and ensemble bagging tree. *Appl. Energy* **2020**, *267*, 114877. [[CrossRef](#)]