Article

Multivariate Multiscale Cosine Similarity Entropy and Its Application to Examine Circularity Properties in Division Algebras †

Hongjian Xiao 1,*, Theerasak Chanwimalueang 2 and Danilo P. Mandic 1,*

1 Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK
2 Faculty of Engineering, Srinakharinwirot University, Nakhon Nayok 26120, Thailand
* Correspondence: d.mandic@imperial.ac.uk; Tel.: +44(0)-207-594-6271
† This paper is an extended version of our paper published in https://doi.org/10.1109/ICASSP43922.2022.9747282.

Abstract: The extension of sample entropy methodologies to multivariate signals has received considerable attention, with traditional univariate entropy methods, such as sample entropy (SampEn) and fuzzy entropy (FuzzyEn), introduced to measure the complexity of chaotic systems in terms of irregularity and randomness. The corresponding multivariate methods, multivariate multiscale sample entropy (MMSE) and multivariate multiscale fuzzy entropy (MMFE), were developed to explore the structural richness within signals at high scales. However, the requirement of high scale limits the selection of embedding dimension and thus, the performance is unavoidably restricted by the trade-off between the data size and the required high scale. More importantly, the scale of interest in different situations is varying, yet little is known about the optimal setting of the scale range in MMSE and MMFE. To this end, we extend the univariate cosine similarity entropy (CSE) method to the multivariate case, and show that the resulting multivariate multiscale cosine similarity entropy (MMCSE) is capable of quantifying structural complexity through the degree of self-correlation within signals. The proposed approach relaxes the prohibitive constraints between the embedding dimension and data length, and aims to quantify the structural complexity based on the degree of self-correlation at low scales. The proposed MMCSE is applied to the examination of the complex and quaternion circularity properties of signals with varying correlation behaviors, and simulations show the MMCSE outperforming the standard methods, MMSE and MMFE.

Keywords: cosine similarity entropy; angular distance; multi-channel system; detection of circularity; complex circularity; quaternion circularity; multivariate entropy

1. Introduction

The investigation of entropy has been a considerably important topic in nonlinear analysis, especially in relation to the quantification of the degree of irregularity and multiple correlations within time series. Entropy-based complexity studies have been implemented in many physical and physiological studies, mostly based on the widely accepted complexity loss theory (CLT) [1]. The CLT states that the physiological responses of the human body in a pathological state (resulting from, for example, stress, illness, or aging) exhibit a loss in structural complexity, while the highest degree of irregularity is associated with healthy and young bio-systems [1,2].

The concept of entropy derives from thermodynamics to measure the probability of micro-states in physics [3]. When it comes to Shannon’s information theory, entropy is employed to describe the amount of information and the generation rate of new information [4,5]. Among numerous methods to characterize signals in terms of entropy, two commonly used techniques are approximate entropy [6] and sample entropy [7]. Both employ the Chebyshev distance and utilize estimates of the conditional probability of
similar patterns within the time series. Sample entropy is usually considered an enhancement of approximate entropy, as it gives less biased estimation and higher robustness [7]. Subsequently, sample entropy (SampEn) has found applications in both physical and physiological systems [8–13]. For example, research based on SampEn of the electrocardiogram (ECG) has been utilized to estimate different degrees of atrial fibrillation [14], or the analysis of stress in public speaking [15]. A further improvement of SampEn, called fuzzy entropy (FuzzyEn), employs fuzzy functions as a substitute for the hard thresholding within SampEn in the definition of similar patterns [16]. Given a continuous relation between the similarity and the obtained distance, FuzzyEn is capable of giving a more reliable and well-defined estimation, while requiring shorter data length compared to SampEn [16]. Moreover, FuzzyEn has been applied in research on the dynamics of gaze complexity during virtual sailing [17].

Despite the broad use of SampEn and FuzzyEn, their shortcomings still remain when it comes to real-world data. The problems include the following:

- The distance measure utilized in SampEn and FuzzyEn is based on the amplitude of the time series, where the unlimited range would lead to undefined values.
- Due to the signal related parameter setting, when the tolerance is set according to the value of standard deviation, the data normalization is required in advance, which would re-scale the original signal and cause information loss.
- Considering the requirement of re-scaling, this affects the noise robustness of SampEn and FuzzyEn, in particular, when processing short time series.
- Long data length is required for traditional SampEn and FuzzyEn analyses when exploring complexity with a high embedding dimension.

Recently, a new theory associated with self-correlation states that pathology is simultaneously followed with an increase in structural complexity (e.g., correlation within time series), which can be considered a complementary theory to the traditional complexity loss theory (CLT) [18]. This promises to help resolve the problems with SampEn- and FuzzyEn-based methodologies, which fail to give a holistic quantification of the self-correlation of the target signals. Unlike other existing entropy methods, cosine similarity entropy (CSE) was introduced to quantify the structural complexity of time series in terms of self-correlations based on angular distance [18].

The CSE provides a robust and more meaningful estimation of the structural complexity of chaotic systems, in contrast to SampEn and FuzzyEn [18]. The extension of CSE to multivariate case is introduced in this paper, whereby the proposed multivariate multiscale cosine similarity entropy (MMCSE) is based on the idea of the composite delay vector introduced in multivariate multiscale sample entropy (MMSE) [19] and the rule of angular distance applied in CSE [18]. The performance of MMCSE is first evaluated over all parameter settings, and then applied to the detection of complex and quaternion non-circularity in multi-channel systems, whereby different correlation behaviors were examined and discussed by multivariate entropy methods. The virtues of MMCSE are further illustrated by correctly quantifying the meaningful structural complexity for varying conditions of correlation degrees and power in data channels.

The reminder of the paper is organized as follows. Related work and development of SampEn and FuzzyEn-based entropy are reviewed in Section 2. To contrast the structure of the cosine similarity entropy method, traditional sample entropy and fuzzy entropy along with their multivariate extended forms are presented and reviewed in Section 3. The details of the extension of cosine similarity entropy to the multivariate case are given in Section 4. Section 5 discusses the selection of parameters and provides the suggested values for the manually set parameters when applying multivariate multiscale cosine similarity entropy. Section 6 illustrates the application of MMCSE in the detection of circularity in terms of varying correlation behaviors. Finally, the conclusions are given in the last section.
2. Related Work and Development of Existing Sample Entropy and Fuzzy Entropy

Although FuzzyEn exhibits a higher consistency at the expense of computational load, single-scale processing has shown limitations in real-world entropy analysis. Hence, the ‘coarse-graining process’ (CGP) was subsequently introduced in traditional entropy methodologies to yield multiscale entropy [20], which allows for the examination of information at high scales. This is achieved by taking the average over neighboring data points. However, it was realized that CGP is not the optimal method to obtain the scaled signal because the traditional CGP in the frequency domain operates as a low-pass filter with large stop band ripples, resulting in biased scaled signals containing artifacts [21]. In addition, the averaging procedure of CGP requires a large data set and comes with a loss of the high-frequency component in data [22]. To overcome the first drawback of CGP, refined multiscale entropy was introduced by employing a low-pass Butterworth filter to generate unbiased scaled signals [21]. As a refinement of CGP in terms of the second flaw, several improved methods have been proposed. For instance, composite multiscale sample entropy operates by implementing CGP at various starting points and taking the average over the resulting entropy values [23], while generalized multiscale entropy employs higher moments (e.g., variance) to give the dynamics over multiple scales in place of the mean [24]. Another multiscale strategy that has been applied in entropy analysis is named intrinsic mode entropy (IMEn), and employs intrinsic mode functions (IMFs) obtained by empirical mode decomposition (EMD), whereby the scaled signals in IMEn are obtained by cumulative sums of IMFs to introduce robustness toward low-frequency changes [25]. However, the additional processes used to improve CGP inevitably increase the computational time. Despite these drawbacks of CGP, it is widely used in real-world applications since it balances computation efficiency with reasonable entropy estimates.

Multivariate multiscale sample entropy (MMSE) was proposed as an extension to SampEn in order to simultaneously process information from multiple channels and relax the requirement of long data length for uni-variate entropy [19]. For multi-channel data, MMSE provides a more powerful and accurate estimation, even for limited data length [19]. Multivariate multiscale fuzzy entropy (MMFE) was subsequently introduced, which maintains the merits of FuzzyEn in contrast to SampEn [26], in the multivariate setting. Both FuzzyEn- and SampEn-based methods are built based on the concept of amplitude distance, whereby outliers manifest themselves as spurious peaks and have a severe impact on the accuracy of estimation of structural complexity [18]. Further, the complexity quantified by SampEn or FuzzyEn is based on randomness and irregularity, which is one aspect of complex systems, while other properties, such as long-range correlations, fail to be demonstrated by single-scale SampEn and FuzzyEn [27]. Note that the scale of interest in multiscale entropy, based on amplitude distance, has no unified measure indicators as criteria when the scale is guaranteed to be set as a certain value to exhibit the long-range correlation of the signal.

3. Multivariate Multiscale Sample Entropy and Multivariate Multiscale Fuzzy Entropy

In this section, two traditional entropies, sample entropy (SampEn) and fuzzy entropy (FuzzyEn), are introduced and compared. Starting from a uni-variate single-scale process to the multivariate multi-scale approach, both SampEn and FuzzyEn are developed based on the amplitude distance and the conditional probability of similar patterns.

3.1. Sample Entropy and Fuzzy Entropy

The details of sample entropy are given in Algorithm 1. In standard SampEn, the original signal is first reconstructed in the phase space following Takens’ embedding theory [28]. Each embedding vector is jointly described by the embedding dimension, \( m \), and time delay, \( l \). Then, the Chebyshev distance is applied to pairs of the embedding vectors to give the degree of similarity, where a similar pair is defined by the Heaviside function with the tolerance coefficient, \( r \). In the final stage, the probability of similar patterns is produced, and the estimated result is generated by the ratio of the probabilities given in increased phase spaces.
Similarly to sample entropy, fuzzy entropy employs the Chebyshev distance to give the similarity between paired embedding vectors. Here, the embedding matrix needs to be centered prior to the reconstruction. Next, the embedding matrix, \( X_m(i) \), is formed and the distance matrix \( D_m(i, j) \) is computed. The details of FuzzyEn are presented in Algorithm 2. The main modification of FuzzyEn from SampEn is the involvement of a continuous transformation from the distance, \( d_m(i, j) \), to the similarity, \( s_m(i, j) \), through the fuzzy membership functions as given in the Step 3 in Algorithm 2. Benefiting from the employment of the fuzzy membership functions, undefined estimation due to the absence of matching pairs is largely avoided. However, a higher computational load is required compared to SampEn, and the performance of FuzzyEn generally shows improved but also similar trends as with SampEn.

Algorithm 1. Sample Entropy

Given a univariate data set \( \{ x(i) \}_{i=1}^{N} \) of length \( N \), the parameters involved are the embedding dimension, \( m \), tolerance, \( r_s \), and time delay, \( l \).

1. Construct the embedding matrix, \( X_m \), derived from the original signal, \( \{ x(i) \}_{i=1}^{N} \), in the form \( x_m(i) = (x(i), x(i+1), \ldots, x(i+(m-1)l)) \).
2. Compute the distance between all the pairwise embedding vectors, \( x_m(i) \& x_m(j) \), based on the Chebyshev distance, \( d_m(i, j) = \max \{ x_m(i+k) - x_m(j+k) \} | 0 \leq k \leq m-1, i \neq j \} \).
3. Compute the number of matching patterns, \( B_{r_s}^m(i) \), defined as the similar pairs that satisfy the criterion \( d_m(i, j) \leq r_s \).
4. Compute an estimate of the local probability of \( B_{r_s}^m(i) \) as \( C_m^r(i) = \frac{B_{r_s}^m(i)}{N-n-1} \), \( n = (m-1)^*l \).
5. Compute an estimate of the global probability of \( B_{r_s}^m(i) \) as \( \Phi_m^r = \frac{\sum_{i=1}^{N-n-1} C_m^r(i)}{N-n} \).
6. Repeat step 1–5 with an increased embedding dimension, \( m+1 \), and obtain the updated global probability, denoted as \( \Phi_{m+1}^r = \frac{\sum_{i=1}^{N-n-1} C_{m+1}^r(i)}{N-n} \), \( n = m^*l \).
7. Sample entropy is defined as \( SE(m, l, r_s, N) = -\ln \left[ \frac{\Phi_{m+1}^r}{\Phi_m^r} \right] \).

Algorithm 2. Fuzzy entropy.

Given a univariate data set \( \{ x(i) \}_{i=1}^{N} \) of length \( N \), the parameters involved are the embedding dimension, \( m \), tolerance, \( r_f \), and time delay, \( l \). The fuzzy function applied here is the Gaussian function with a chosen order, \( \eta \).

1. Form the centered embedding matrix, \( U_m \), derived from the original signal, \( \{ x(i) \}_{i=1}^{N} \), by \( u_m(i) = x_m(i) - \mu_m(i) \), where \( \mu_m(i) = \frac{1}{m} \sum_{k=1}^{m} x_m(i+k) \).
2. Compute the distance between all the pairwise embedding vectors, \( u_m(i) \& u_m(j) \), based on the Chebyshev distance, \( d_m(i, j) = \max \{ u_m(i+k) - u_m(j+k) \} | 0 \leq k \leq m-1, i \neq j \} \).
3. Convert the distance matrix, \( D_m \), to the similarity matrix, \( S_m \), calculated through the Gaussian function as \( s_m(i, j) = e^{-\frac{d_m(i, j)^\eta}{r_f^\eta}} \).
4. Compute the estimated local probability of \( S_m(i) \) as \( C_m^f(i) = \frac{\sum_{i=1}^{N-n-1} s_m^f(i)}{N-n-1} \), \( n = (m-1)^*l \).
5. Compute the estimated global probability of \( S_m(i) \), by \( \Phi_m^f = \frac{\sum_{i=1}^{N-n-1} C_m^f(i)}{N-n} \).
6. Repeat step 1–5 with an increased embedding dimension, \( m+1 \), and obtain the updated global probability as \( \Phi_{m+1}^f = \frac{\sum_{i=1}^{N-n-1} C_{m+1}^f(i)}{N-n} \), with \( n = m^*l \).
7. Fuzzy entropy is defined as \( FE(m, l, r_f, N) = -\ln \left[ \frac{\Phi_{m+1}^f}{\Phi_m^f} \right] \).
3.2. Multiscale Entropy

The information provided by single-scale entropy is limited, as it does not account for long-range correlations and temporal fluctuations over multiple scales [20]. To this end, by employing the consecutive coarse-graining process (CGP) with a new parameter denoted as the scale factor, $\tau$, Costa et al. proposed multiscale entropy [20], given in Algorithm 3.

Algorithm 3. Multiscale sample entropy and multiscale fuzzy entropy.

Assume that a univariate data set, $\{x(i)\}_{i=1}^{N}$, is of length, $N$, and the coarse graining scale factor is donated as $\tau$.

1. Obtain the scaled time series, $\{y(j)\}_{j=1}^{N/\tau}$, by the coarse graining process as

$$y(j) = \frac{1}{\tau} \sum_{i=j-\tau/2}^{j+\tau/2-1} x(i), \quad 1 \leq j \leq \frac{N}{\tau}.$$

2. Apply the scaled data set, $y(j)$, as the input into sample entropy given in Algorithm 1 or fuzzy entropy in Algorithm 2 to obtain the complexity estimation based on multiscale sample entropy (MSE) and multiscale fuzzy entropy (MFE), respectively.

3.3. Multivariate Multiscale Entropy

The implementation of CGP reveals the hidden information buried in multiple scales. However, the averaging process requires multiple times larger data size. To this end, multivariate multiscale entropy based on SampEn (MMSE) was proposed [19], followed by multivariate multiscale fuzzy entropy [26]. There are two main modifications in multivariate entropy algorithms based on amplitude distance. One is that the embedding vector in the univariate SampEn is replaced by the composite delay vector (CDV) composed by the embedding vectors from each channel in multivariate entropy. Figure 1 shows an example of composite delay vector construction for $m = 3$ and $l = 1$. In addition, due to the different range of amplitudes across channels, the obtained distance between pairwise composite delay vectors can be biased. Therefore, the second modification of multivariate entropy is the requirement of normalization of input data sets. As stated in [26], the performance of MMFE is proved to be more consistent than MMSE, particularly in high-dimensional phase spaces. The processes of MMSE and MMFE are summarized in Algorithm 4.

$m = 3, \quad l = 1$

![Figure 1. Example of composite delay vectors construction in multiscale entropy](image-url)
Algorithm 4. Multivariate multiscale sample entropy and multivariate multiscale fuzzy entropy

Given a multivariate data set with \( P \) channels \( \{x_{k,i}\}_{i=1}^{N} \), \( 1 \leq k \leq p \), of length \( N \). Parameters involved are the embedding dimension, \( M = [m_1, m_2, \ldots, m_p] \), tolerance, \( r \), time delay, \( L = [l_1, l_2, \ldots, l_p] \), and the scale factor, \( \tau \).

1. Standardize the original multivariate data sets by subtracting the mean and dividing by standard deviation for each channel.
2. Coarse graining is next applied to the normalized datasets, \( \{y^{(\tau)}_{k,i}\}_{i=1}^{N/\tau} \), for each channel following Algorithm 3.
3. Form the composite delay matrix, \( Y_M(i) \), according to the embedding dimension, \( M \), and the time delay, \( L \), in the form

\[
Y_M(i) = \begin{bmatrix}
y_{1,i}, y_{1,i+l_1}, \ldots, y_{1,i+(m_1-1)l_1} r \\
y_{2,i}, y_{2,i+l_2}, \ldots, y_{2,i+(m_2-1)l_2} r \\
\vdots \\
y_{p,i}, y_{p,i+l_p}, \ldots, y_{p,i+(m_p-1)l_p} r 
\end{bmatrix}
\]

4. Apply the scaled composite delay matrix, \( Y_M(i) \), as the input into the sample entropy presented in Algorithm 1 or fuzzy entropy in Algorithm 2 to obtain the measure of multivariate multiscale sample entropy (MMSE) and multivariate multiscale fuzzy entropy (MMFE), respectively.

4. Cosine Similarity Entropy and Multi-Variate Approach

Multiscale cosine similarity entropy (MCSE) was proposed [18], whereby instead of amplitude-based distance, CSE employs the angular distance in phase space to define the difference among embedding vectors. The angular distance offers advantages, especially regarding the sensitivity to outliers or sharp changes in time series that amplitude-distance-based calculation suffers from. Additionally, the angular distance, restricted by the maximal length of \( 2\pi \), is more robust, and is less prone to generating undefined estimates in the presence of noise. In amplitude-based entropy calculation, because of the large distance that is obtained by amplitude, the tolerance, \( r \), is set as a ratio associated with the standard deviation of the input data set, where parameters driven by data can be unstable for signals with high variance. Given that cosine similarity entropy is restricted to values within \( [0, 1] \), the tolerance for CSE is selected, irrelevant to the variance of original signal. Hence, the process of CSE exhibits enhanced stability, especially when dealing with highly dynamical signals. The virtue of CSE is therefore that entropies can be assessed more stably in multivariate analysis when the data sets across channels have different variance.

The extension of MCSE to multivariate scenarios employs the construction of the composite delay vector in MMSE [19] and follows the main procedure of CSE [18]. Note that multivariate multiscale cosine similarity entropy (MMCSE) is sensitive to the DC offset, as it requires the origin coordination to project the angular distance within the tolerance angle. Therefore, the DC offset or a long-term trend needs to be removed in MMCSE. Here, we apply the zero-median approach to remove the global trend because of its robustness against outliers. It is worth highlighting that the definition of CSE is developed based on Shannon entropy, where the calculation only relies on the probability of phase space, \( m \), without the involvement of phase space \( m+1 \) [18], which is less computationally expensive. The proposed MMCSE is presented in Algorithm 5.
Algorithm 5. Multivariate multiscale cosine similarity entropy.

Given a multivariate data set with $P$ channels $\{x_{i,k}\}_{i=1}^{N}$, $1 \leq k \leq p$, of length $N$. Parameters involved are designated as the embedding dimension, $M = [m_1, m_2, \ldots, m_P]$, tolerance, $r_c$, time delay, $L = [l_1, l_2, \ldots, l_P]$, and the scale factor, $\tau$.

1. Remove the DC offset of the original data sets $\{x_{i,k}\}_{i=1}^{N}$ by subtracting the median for each channel.
2. Obtain the scaled time series $\{y_{i,k}\}_{i=1}^{N}$ following the instruction of Algorithm 3.
3. Reconstruct composite delay vectors $Y_M(i)$ as demonstrated in Algorithm 4.
4. Calculate the angular distance based on cosine similarity for all pairwise composite delay vectors, $Y_M(i) \& Y_M(j)$, as
   $$d_M(i, j) = \frac{1}{\tau} \cos^{-1} \left( \frac{Y_M(i) \cdot Y_M(j)}{|Y_M(i)||Y_M(j)|} \right), \quad i \neq j$$
5. Compute the number of similar patterns defined as similar pair, $B^r_M(i)$, that satisfy the criterion $d_M(i, j) \leq r_c$.
6. Compute the estimated local probability of $B^r_M(i)$ by $C^c_M(i) = \frac{B^c_M(i)}{N-\tau}$, where
   $$n = \max(M) \ast \max(L)$$
7. Compute the estimated global probability of $B^r_M(i)$ as $\Phi^r_M = \frac{\sum_{i=1}^{n} C^c_M(i)}{N-n}$
8. Cosine similarity entropy is defined as
   $$MMCSE(M, L, r_c, N) = -(|\Phi^c_M| \log_2 \Phi^c_M + (1 - \Phi^c_M) \log_2 (1 - \Phi^c_M)).$$

5. Value of Parameters

As pointed out in the MMCSE algorithm, several parameters need to be manually selected. This section provides a comprehensive discussion regarding the choice of the parameters, including the tolerance, $r$, embedding dimension, $m$, and data length, $N$. The corresponding performances of tri-variate MMSE and MMFE are examined against MMCSE. Throughout the analysis, the time delay was set to $L = 1$ for all simulations to avoid unknown influence and also make the temporal span fully controlled by the modification of the embedding dimension. Unless otherwise stated, the default values of parameters for MMSE and MMFE are set to $r = 0.15 \ast tri(S)$, $M = 2$ and $N = 10,000$, where $tri(S)$ refers to the total variance of the covariance matrix, $S$. The mean complexity curves in each subsection are plotted with error bars representing the standard deviation. Both the mean and standard deviation were calculated over outcomes of 10 independent realizations for each system.

5.1. Tolerance, $r$

Unlike MMSE and MMFE, the selection of tolerance, $r$, in MMCSE is independent of the variance of the input data sets. However, due to multiple channels, the influence of the number of variate needs to be considered when selecting the value of tolerance in MMCSE. To explore the relationship of the number of channels, $p$, and the tolerance, $r$, multi-channel white Gaussian noise (WGN) was utilized as the reference signal due to its complete randomness and simplest structure. The performances of multivariate entropy were first evaluated by varying tolerance, and then the mathematical relation was given by the best fit of the curve between $r$ and $p$.

The left panel of Figure 2a shows the outcomes of multivariate single-scale CSE based on WGN as a function of tolerance for a varying number of input channels, from a univariate to 16-variate case. The tolerance was initially set to vary from 0.01 to 1 at intervals of 0.05 and was then linearly interpolated using values ranging from 0 to 1, at 0.01 intervals. Observe, in all curves, a rise in entropy with tolerance from 0 to 0.5 and a decrease in entropy with tolerance from 0.5 to 1. Due to this symmetry, only the range of 0 to 0.5 was considered to give the common relationship that the larger the tolerance, the more similar patterns are
found shown as higher entropy. As discussed in [18], the tolerance for univariate multiscale cosine similarity entropy was empirically set to \( r = 0.07 \), whereby the performance of MCSE applied to WGN resulted in the entropy value of 0.365. Therefore, the relationship between the number of variates and tolerance, \( r \), was estimated by setting the entropy as 0.365 as a standard (the black dashed line shown in the left panel in Figure 2a), as revealed in the right panel of Figure 2b using piece-wise cubic interpolation fitting curve. It is found that the best fit was for

\[
 r_{cse} = -0.4(p^{-0.71}) + 0.47, \tag{1}
\]

where \( p \) donates the number of variates.

![Figure 2](image-url)  
**Figure 2.** Choice of parameters in MCSE. (a) Behaviour of multivariate single-scale CSE on the estimation of white Gaussian noise (WGN), as a function of the tolerance, \( r \). The default parameters were set as \( N = 10,000 \), \( M = 2 \), and \( L = 1 \) for all channels. The error bars designate the standard deviation over 10 realizations. (b) The optimal tolerance, \( r \), as a function of the number of variates.

Therefore, the tolerance in MMCSE is decided as a function of the number of variates. This relationship is given in Equation (1) to maintain the performance of multivariate CSE for a varying channel number. For the commonly applied situation when the channel number is 2 (bivariate data), the typical setting value of tolerance is \( r = 0.225 \) and for the tri-variate case, it is \( r = 0.287 \), according to the equation.

### 5.2. Embedding Dimension, \( m \), and Data Length, \( N \)

The embedding dimension, \( m \), in amplitude-distance-based methods, is usually restricted by the data length, \( N \). As a rule of thumb, \( N \) should scale as \( 10^m \) [29,30]. For fuzzy entropy, the limitation is relaxed such that \( N \) should exceed 50 samples to give a defined estimate. That is due to the continuous boundary of the transfer function between similarity and tolerance at the expense of computational load [16]. As a result, limited by the sample size and computational time, the structure of complexity in high embedding dimensions is usually poorly estimated by current entropy methods. The trade-off between the embedding dimension and sample length means that the entropy methods are difficult to implement in real time. The influences of the embedding dimension, \( m \), and the data length, \( N \), of these multivariate entropy methods are discussed here.

The performance of the three multivariate entropy algorithms, MMSE, MMFE, and MMCSE, was tested on white Gaussian noise (WGN) and autoregressive (AR) models AR(1), AR(2), and AR(3) in three-channel multivariate systems. A total of 10 independent realizations were generated for each model with a constant sample size of \( N = 2000 \) points for all operations, and the results are shown in Figure 3. The mean entropies are shown, and error bars designated the standard deviation. Coefficients of the considered AR models are given in Table 1.

The mean values of entropy are shown as a function of the embedding dimension, \( m \), in Figure 3. The given embedding dimension ranges from 1 to 10 with increments of 1 for MMSE and MMFE, and the range of used embedding dimension for MMCSE was set to 2 to 10, due to the minimum requirement of 2 when calculating angular distance. Observe that in the left panel of Figure 3, the selection of \( m \) for MMSE is
limited to lower than 7 to give a defined estimation for the most complex signal AR(3). On the other hand, for uncorrelated WGN, the maximum embedding dimension with a defined positive MMSE value is \( m = 2 \). In the second panel, MMFE produces a more stable estimation at higher scales, in contrast to MMSE. Yet, the differences between the four synthetic signals decrease with increasing the embedding dimension, and when \( m \) comes to 9, the complexity estimates overlap. In addition, resulting from the single-scale analysis, both MMSE and MMFE fail to order the complexities of the four simulated signals. The highest complexity among the given models ought to be assigned to AR(3). In comparison, the MMCSE in the last panel gives a positive, defined estimate for all the scenarios with the correct rank based on the structural complexity of the signals, where complexity(AR(3)) > complexity(AR(2)) > complexity(AR(1)) > complexity(WGN). Moreover, MMCSE illustrates a consistent estimation with a good separation among the four models. Therefore, MMCSE exhibits a more stable and consistent performance when applying complexity evaluation under high embedding dimensions in these simulations.

Table 1. Coefficients of the AR models used.

| Coefficients | \( a_1 \) | \( a_2 \) | \( a_3 \) |
|--------------|----------|----------|----------|
| AR (1)       | 0.9      |          |          |
| AR (2)       | 0.7      | 0.25     |          |
| AR (3)       | 0.6      | 0.25     | 0.125    |

Figure 3. Behavior of multivariate single-scale SE, FE and CSE as a function of the embedding dimension, \( m \). The default parameters were set as \( N = 2000, L = [1,1,1], r = 0.45 \) for multivariate SE/FE, and \( r = 0.287 \) for multivariate CSE.

The second significant parameter that controls the performance of the complexity estimation based on entropy is the measured data length, \( N \), which is difficult to control for many real-world signals. Figure 4 exhibits the performance of the three single-scale multivariate entropy methods. The default parameters were set to \( M = [2,2,2], L = [1,1,1], \) and the tolerance values were selected following the Equation (1). In the left panel in Figure 4, multivariate sample entropy performs poorly when \( N < 60 \) on the most complex signal here, AR(3). Generally it gives wide error bars with little separation between the four models. Multivariate fuzzy entropy and multivariate cosine similarity entropy exhibit a more stable estimate and a less strict requirement for the minimal data length compared to that given by multivariate sample entropy. However, in the middle panel and left panel, the most structurally complex model, AR(3), is wrongly assigned the lowest degree of complexity. The improvement of fuzzy entropy from sample entropy comes at the expense of calculation efficiency with more time consumption required. The cosine similarity entropy-based algorithm manages to reduce the minimal data length required without increasing the computational load.

5.3. Complexity Profile of MMCSE, MMSE and MMFE

The complexity profiles of multivariate multiscale entropy algorithms were first calculated for simulated linear and nonlinear signals. The same auto-regressive (AR) models were applied together with a commonly encountered nonlinear signal, the 1/f noise. With the default parameters set to \( M = [2,2,2] \) and \( N = 10,000 \), the mean complexity of
10 independent realizations with three channels are illustrated in Figure 5. Error bars represent the standard deviation over the 10 realizations.

**Figure 4.** Behavior of multivariate single-scale SE, FE and CSE as a function of the data length, \(N\). The default parameters were set as \(M = [2,2,2], L = [1,1,1], r = 0.45\) for multivariate SE/multivariate FE, and \(r = 0.287\) for multivariate CSE.

From the left to the right panel in Figure 5, the complexity profiles of the multivariate multiscale sample entropy (MMSE), multivariate multiscale fuzzy entropy (MMFE) and multivariate multiscale cosine similarity entropy (MMCSE) are presented. With the same setting, MMSE is observed to be capable of distinguishing signals with self-correlations, including AR models and \(1/f\) noise, from the completely uncorrelated WGN (in blue line). However, the random WGN is wrongly assigned to exhibit a structural complexity at low scales, and the separation among correlated signal is also difficult to be achieved as overlapped error bars shown across all the scales by MMSE. In contrast with MMSE, MMFE correctly demonstrates the complexity difference of the five models (AR models from WGN and \(1/f\) noise from WGN respectively) with a narrower variance.

Observe from the middle panel in Figure 5, that the \(1/f\) noise is expected to be a truly complex signal across all scales, which is correctly given by MMFE. However, the problem of the wrong estimation of the complexity of uncorrelated WGN still remains by applying MMFE at small scales. When estimated by MMCSE in the right panel in Figure 5, the complexity measures based on WGN remain the lowest across all scales, reflecting the uncorrelated structure of an ideal random signal. In regard to the AR processes with different orders, a system with the higher order means more coefficients involved when generating signals associated with a higher degree of freedom. Hence, the correct relation of structural complexity of simulated AR processes is supposed to be \(AR(3) > AR(2) > AR(1)\), which can be only observed in the graph given by MMCSE. Meanwhile, the estimation of the complexity given by MMCSE is consistent in the comparison between AR processes with various orders, and between \(1/f\) noise and WGN based on self-correlation. To this end, the proposed multivariate multiscale cosine similarity entropy yields stable estimates at a high coarse-graining scale, thus making it possible to examine structural complexity of real world processes with long-range correlations.

**6. Detection of Circularity**

We next examine the application of multivariate entropy in the detection of complex and quaternion non-circularity, that is, the rotation dependence of probability density function (pdf). The non-circularity manifests itself through a degree of correlation and
power imbalance of the system channels, which can be reflected in a scatter plot as the inclined angle and dispersed degree of the distribution. The long-range correlation of a time series measured by multiscale entropy reflects the influence of past states on the generation of future signals [31], while the correlation of a multi-channel system is a measure of association among channels, a summary of dependence strength for multivariate systems, and a feature that could reflect the structural richness of the system [32,33]. A strongly correlated system is expected to show a high structural complexity, while a less correlated system would approach the performance of WGN to give a low structural complexity. As for the other critical property, the power of the signal has been widely used as an indicator of dynamics in complex systems, such as the analysis based on heart rate variability in physical systems [34], where the ratio of the low-band frequency power and high-band frequency power was used to reveal the sympatho-vagal balance of the autonomous nervous system [35].

The ability of standard MMSE, MMFE and the proposed MMCSE in the detection of circularity of multivariate signals is next simulated and discussed. We consider the case of a tri-variate (pure quaternion) signal, while the analysis based on bi-variate (complex valued) system can be found in Appendix A. Systems constructed by tri-variate correlated WGN, $Y$, are discussed here and generated by the correlation matrix, $C$, shown in Equations (2) and (3). The correlation behavior is jointly controlled by the coefficients, $p$ and $q$, while the correlated power is determined by the factor, $q$, that is

$$ Y = X \cdot \text{chol}(C) $$

(2)

$$ C = \begin{bmatrix}
1 & pq & pq \\
pq & q^2 & pq^2 \\
pq & pq^2 & q^2
\end{bmatrix}, \quad \text{chol}(C) = \begin{bmatrix}
1 & pq & pq \\
0 & q\sqrt{1-p^2} & q\sqrt{1-p^2} \\
0 & 0 & q\sqrt{1-2p^2+1+p^2}
\end{bmatrix} $$

(3)

Here, $X$ denotes a tri-variate uncorrelated WGN system, $X = [\epsilon_1, \epsilon_2, \epsilon_3], \epsilon(i) \sim \mathcal{N}(0,1)$, and Chol($C$) refers to the Cholesky decomposition based on the correlation matrix, $C$.

The tri-variate systems with varying correlation degree, $p$, and power, $q$, were analyzed by MMSE, MMFE, and the proposed MMCSE. The default parameters were set to $M = [2,2,2], L = [1,1,1], N = 1000$, and the maximum scale factor was chosen as $\tau = 10$. The tolerance, $r$, for MMSE and MMFE was set as $0.15 \ast \text{tr}(S)$, where tr$(S)$ refers to the total variance of the input signals. The choice of tolerance for MMCSE was calculated according to Equation (1). The error bars designate the average and standard deviation over 30 realizations for each model.

6.1. Correlated WGN with Equal Power

Tri-variate input signals associated with five different correlation degrees and equal power were first generated as scatter plots shown in Figure 6. The first model in the left graph with $p = 0$, $q = 1$, gave the uncorrelated WGN system. The four considered cases were set as the fixed power, $q = 1$, and the correlated relations among input channels, $Y$, were hence determined by one varying coefficient, $p \in \{0.4, 0.6, 0.75, 0.99\}$. Observe the varying degree of non-circularity, from the rotation-invariant (circular) case on the left panel through to a high degree of non-circularity (narrow scatter plot) on the right-most panel.

Figure 7 illustrates the performance of the multivariate multiscale analysis based on the sample entropy, fuzzy entropy and cosine similarity entropy on the task of the detection of the degree of non-circularity in tri-variate (quaternion) data. Observe that MMSE and MMCSE, in the left and right panels respectively, were capable of distinguishing between the varying settings of coefficients, $p$, shown as different trends of lines, while MMFE in the middle panel was only prone to separating the case with maximum correlation (in amber) from the rest of the scenarios. As $p$ increases, the structural richness of the system (degree of non-circularity) rises, as shown in Figure 6. Hence,
the structural complexity for signals with large $p$ is supposed to be greater than that of signals generated from small $p$, which is correctly evaluated by both MMSE and MMCSE. However, as scale, $\tau$, increases, MMSE exhibits a descending trend and the gaps among curves narrow down accordingly, whereby the ability of MMSE to classify varying correlations is negatively influenced by the increasing scale. On the other hand, for MMCSE in the right panel, the increasing scale has no impact on the separation of different models as shown by the flat lines.

Figure 6. Scatter diagrams of tri-variate (quaternion) input correlated WGN (pure quaternion) with equal powers in the three data channels, but different degrees of correlation among them. Observe an increase in the degree of non-circularity from left to right.

Figure 7. Behavior of standard MMSE, MMFE, and the proposed MMCSE applied to correlated tri-variate (quaternion) WGN with equal power.

6.2. Uncorrelated WGN with Unequal Power

Next, five models of tri-variate (pure quaternion) uncorrelated WGN with unequal power were simulated and analyzed by multivariate entropy. Figure 8 demonstrates the scatter diagrams of five uncorrelated input systems, where the dynamics in one channel is independent of the change of others. The coefficient, $p$, was fixed as zero (to make the $\text{chol}(C)$ a diagonal matrix) and $q$ was varying in this case among $\{1, 0.6, 0.45, 0.35, 0.1\}$ to control the power of the system. As shown, the above two channels exhibit no correlation, with flat scatter diagrams and the power imbalance increase from the left to the right when the power reaches its maximum and is fully balanced, which is the case of white noise as given in the first graph. Figure 9 shows the performance of the three multivariate entropy methods.

The left and middle panels of Figure 9 give the estimation by MMSE and MMFE, which fail to reflect the effects of different powers of the signals due to data normalization, as given in the first step of amplitude distance-based algorithms shown in Algorithm 4. In other words, the information about power imbalance is lost when normalizing the data for SampEn and FuzzyEn-based analyses, thus resulting in a decrease in entropy values approaching the behaviors of random white noise. However, data normalization is one of the key pre-processing steps for amplitude-based entropy analyses, which cannot be avoided. In contrast, the MMCSE in the right panel is observed to detect the discrepancy of powers among the data channels of the uncorrelated systems, resulting from non-unified variance towards the input multivariate data sets based on the angular distance. In addition, the non-descending trends in the right panel remain constant, irrespective of the varying power, where the expected performance of structural complexity estimation can be observed at single scale by multivariate CSE.
6.3. Correlated WGN with unequal power

Next, we jointly considered the two above-discussed scenarios for systems of tri-variate (pure quaternion) correlated WGN associated with unequal powers in data channels. In terms of the correlated input, the scatter diagrams are illustrated in Figure 10, where \( q \) is varying within \( \{1, 0.5, 0.35, 0.1\} \), given \( p \) a non-zero value as \( p = 0.6 \). For comparison, the uncorrelated WGN, \( p = 0 \) \& \( q = 1 \), was also produced as the model with least structural complexity (circular) as shown in the first diagram in Figure 10. Figure 11 demonstrates the complexity estimation based on the models generated.

With a fixed, non-zero coefficient \( p \), the increasing power imbalance (non-circularity) yields a higher structural complexity with stronger correlation as shown in the scatter plots from the left to the right in Figure 10. Regarding the complexity measures, observe that MMSE in the left panel of Figure 11 could only separate the completely uncorrelated signals (in blue) from the correlated signals, while MMFE fails to detect the significant discrepancy of structural complexity within all the simulated models. In terms of the proposed MMCSE, signals with various powers and correlation structures (different degrees of non-circularity) are separated in a consistent way in the right panel in Figure 11. To further explore the ability of complexity estimation for signals with jointly varying coefficients \( p \) and \( q \), the heatmaps of entropy estimated via multivariate single-scale sample entropy, multivariate single-scale fuzzy entropy and multivariate single-scale cosine similarity entropy are displayed in Figure 12, where \( p \) in the \( y \)-axis was set to range from 0.05 to 0.99 at 0.05 intervals, and power, \( q \), in the \( x \)-axis ranged from 0.05 to 1 at 0.05 intervals. Observe that SampEn- and FuzzyEn-based multivariate analyses are sensitive to the correlation coefficient \( p \), but are insensitive to signal power, \( q \). Comparing the two amplitude entropy methods, multivariate SE exhibits higher sensitivity to changes of \( p \) than multivariate FE. In contrast, the proposed
CSE-based multivariate method manages to detect changes in both $p$ and $q$ as shown in the right heatmap in Figure 12.

![Heatmap of complexity estimation based on standard multivariate SE, multivariate FE, and the proposed multivariate CSE in a function of varying coefficients $p$ and $q$.](image.png)

Figure 11. Behavior of standard MMSE, MMFE, and the proposed MMCSE applied to correlated tri-variate (quaternion) WGN with unequal power.

Figure 12. Heatmap of complexity estimation based on standard multivariate SE, multivariate FE, and the proposed multivariate CSE in a function of varying coefficients $p$ and $q$.

7. Conclusions

This work extended the univariate cosine similarity entropy to the multivariate case. It was shown that the proposed multivariate multiscale cosine similarity entropy (MMCSE) method is capable of estimating structural complexity based on self-correlation with higher stability at large scales and lower requirement on data length than the existing methods. In contrast with standard MMSE and MMFE, the MMCSE has also exhibited higher consistency in multiscale analysis, where the long-range correlation of signals can be correctly measured at low scales. The performance of MMCSE was first examined on five benchmark signals to reveal the improved estimation against MMSE and MMFE, and then tested on multivariate correlated systems associated with unequal power among data channels (complex and quaternion non-circularity), to detect the properties of circularity including the degree of correlation among channels and the strength of power in the system. As desired, due to the angle-based distance and the relaxation of the requirement of data normalization, the proposed MMCSE exhibited noticeable improvement in detecting the correlation degree of the multivariate system and also showed the unique feature to detect the change of power in the system, unlike MMSE and MMFE. The results indicated the wide range of potential applications of MMCSE in complexity science for practical scenarios. However, regarding the limitation of the coarse-graining average, the multiscale procedure can be further improved by employing the enhanced scaling process. It should be noted that this study is restricted to the analysis on synthetic signals. Therefore, future work will focus on the analysis of real-world data sets and involve statistical methodologies.

Author Contributions: Conceptualization: H.X., T.C. and D.P.M.; methodology, software, formal analysis: H.X. and T.C.; data curation, validation, investigation, resources, visualization, writing—original draft preparation: H.X.; project administration, supervision, writing—review and editing: T.C. and D.P.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.
Data Availability Statement: The data that support the findings of this study are openly available in figshare at https://doi.org/10.6084/m9.figshare.21084355.v1 (accessed on 11 August 2022).

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

Appendix A. Examination of Non-Circularity of Bi-Variate Systems

To examine the ability of MMCSE, similarly to Section 5, simulations on bivariate (complex valued) data were classified into three groups: (a) correlated WGN with equal power, (b) uncorrelated WGN with unequal power and (c) correlated WGN with unequal power. The dual-channel inputs, \( Y \), were generated according to Equation (A1), where \( X = [x_1, x_2] = [\epsilon_1, \epsilon_2] \), \( \epsilon(i) \sim \mathcal{N}(0,1) \), and \( \text{Chol}(C) \) refers to the Cholesky decomposition based on the correlation matrix, \( C \). Thus, the dual channel of the generated system, \( Y = [y_1, y_2] \), becomes
\[
y_1 = x_1 \quad \text{and} \quad y_2 = q(p x_1 + (\sqrt{1-p^2})x_2).
\]
We modified the value of coefficients, \( p \) and \( q \), to control the behaviors of the system correlation during the examination, that is
\[
Y = X \cdot \text{chol}(C), \quad C = \begin{bmatrix} 1 & pq \\ pq & q^2 \end{bmatrix}
\]

For each setting, 30 random realizations were generated with sample size \( N = 1000 \), to give the performance of estimation shown as error bar plots based on the standard deviation. The default parameters of multivariate multiscale entropy were set to \( M = [2,2] \), \( L = [1,1] \), \( r_{se} = r_{fe} = 0.15 * \text{tri}(S) \), and \( r_{cse} = 0.225 \) according to Equation (1). The synthetic data generations and simulations were implemented in MATLAB, where the graphs based on circularity statistics were plotted by using online MATLAB toolbox referred in [36].

Appendix A.1. Correlated WGN with Equal Power

We considered five systems with equal power, \( q = 1 \), but different correlation levels, where \( p = [0, 0.6, 0.8, 0.9, 0.99] \) to give different degrees of complex non-circularity. To illustrate the behaviors of the synthetic systems, angular densities plotted on the unit circle are shown in the first row of Figure A1 as well as angular histograms are in the second row. The number in blue in each angular histogram designates the frequency density of the largest bin in the graph. The system in the first column represents the situation of a random (circular) system, where two channels are independent of each other as the points on the unit circle is loosely distributed with flat histogram in Figure A1. With fixed power, the correlation increases from the left to the right system as the density gradually concentrates on specific non-zero angles, \( \pi/4 \) and \( -3\pi/4 \). Therefore, the structural complexity is expected to increase as \( p \) becomes larger, and the bivariate distribution more non-circular.

Scatter plots are plotted in the first row of Figure A2 to give the space distribution (non-circularity of pdf) of each bi-variate system. The second row of Figure A2 exhibits the complexity estimation given by MMSE, MMFE and MMCSE accordingly. Complexity profiles are shown as entropy in a function of scale factor. Observe that MMSE and MMCSE (in blue and black) present ascending levels with increasing \( p \), while a limited difference can be seen based on MMFE with varying correlations (in red). Although both MMSE and MMCSE are capable of demonstrating the different structural richness, MMSE drops down with increasing scale, while MMCSE exhibits a consistent estimation across all scales.

Appendix A.2. Uncorrelated WGN with Unequal Power

We considered five uncorrelated dual-channel (complex valued) WGN systems, \( p = 0 \), with different power with \( q = [1, 0.4, 0.3, 0.2, 0.1] \). The density and histogram are shown in Figure A3, while the corresponding scatter plots are given in Figure A4. The data points are gathered around horizontal line and the angular density approaches zero as power decreases. Therefore, according to the space distribution and angular density plots, the structure richness of the system (degree of non-circularity) is ascending as \( q \) decreases.
Figure A1. Angular density shown as points on the unit circle and angular histogram of bi-variate (complex valued) input correlated WGN with equal powers in data channels. Observe the increase in the degree of complex non-circularity, from left to right.

Figure A2. Scatter diagrams and entropy-based complexity (non-circularity) estimations of bi-variate input (complex valued) correlated WGN with equal powers in data channels.

Figure A3. Angular density shown as points on the unit circle and angular histogram of bi-variate (complex valued) input uncorrelated WGN with unequal powers in data channels.

The estimates of multivariate multiscale entropy are accordingly shown in Figure A4 below the scatter plot of each system. Both MMSE and MMFE (in blue and red) failed to reflect the change of power with identical complexity and tendency. In contrast, MMCSE (in black) exhibited an increasing structural richness from the left to the right, thus correctly reflecting the change in the degree of non-circularity.

Appendix A.3. Correlated WGN with Unequal Power

Next, we applied the entropy-based estimation on four correlated WGN systems, \( p = 0.6 \), with unequal power, \( q = [1, 0.5, 0.3, 0.1] \) and uncorrelated random system was included as reference, \( p = 0 \) and \( q = 1 \). Figure A5 shows that, despite the fixed correlation coefficient, the decreasing power pulls the angular density back to the horizontal level. Therefore, the structural richness enhances either with an increasing correlation, \( p \), or decreasing power, \( q \), which is also reflected in the scatter plots in Figure A6.
Figure A4. Scatter diagrams and entropy-based complexity estimations of bi-variate (complex valued) input uncorrelated WGN with unequal powers in data channels.

Figure A5. Angular density shown as points on the unit circle and angular histogram of bi-variate (complex valued) input correlated WGN with unequal powers in data channels.

Figure A6. Scatter diagrams and entropy-based complexity estimations of bi-variate (complex valued) input correlated WGN with unequal powers in data channels.

Figure A6 plots the results of entropy in the second row. The same conclusion can be drawn as for tri-variate systems in Section 5 that MMSE and MMFE were able to separate completely random system in the left most panel from the rest correlated systems. However, due to the need of normalization, the power difference could not be detected by the complexity estimation given by either MMSE or MMFE. In contrast, the proposed MMCSE managed to evaluate the structural change of the system either caused by the modification of power or the behavior of correlation.

References

1. Li, Y.; Wang, X.; Liu, Z.; Liang, X.; Si, S. The entropy algorithm and its variants in the fault diagnosis of rotating machinery: A review. *IEEE Access* 2018, 6, 66723–66741. [CrossRef]

2. 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]

3. Marinescu, D.C. *Classical and Quantum Information*; Academic Press: Cambridge, MA, USA, 2011.

4. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* 1948, 27, 379–423. [CrossRef]
5. Blahut, R.E. *Principles and Practice of Information Theory*; Addison-Wesley Longman Publishing Co., Inc.: Boston, MA, USA, 1987.

6. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [CrossRef]

7. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.-Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [CrossRef]

8. Humeau-Heurtier, A.; Wu, C.W.; Wu, S.D. Refined composite multiscale permutation entropy to overcome multiscale permutation entropy length dependence. *IEEE Signal Process. Lett.* **2015**, *22*, 2364–2367. [CrossRef]

9. Ni, L.; Cao, J.; Wang, R. Analyzing EEG of quasi-brain-death based on dynamic sample entropy measures. *Comput. Math. Methods Med.* **2013**, *2013*, 618743. [CrossRef]

10. Govindan, R.; Wilson, J.; Eswaran, H.; Lowery, C.; Preißl, H. Revisiting sample entropy analysis. *Phys. A Stat. Mech. Its Appl.* **2007**, *376*, 158–164. [CrossRef]

11. Looney, D.; Adjei, T.; Mandic, D.P. A novel multivariate sample entropy algorithm for modeling time series synchronization. *Entropy* **2018**, *20*, 82. [CrossRef]

12. Xiao, H.; Mandic, D.P. Variational embedding multiscale sample entropy: A tool for complexity analysis of multichannel systems. *Entropy* **2021**, *24*, 26. [CrossRef]

13. Hemakom, A.; Goverdovsky, V.; Looney, D.; Mandic, D.P. Adaptive-projection intrinsically transformed multivariate empirical mode decomposition in cooperative brain–computer interface applications. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2016**, *374*, 20150199. [CrossRef] [PubMed]

14. Alcaraz, R.; Abásolo, D.; Hornero, R.; Rietas, J. Study of Sample Entropy ideal computational parameters in the estimation of atrial fibrillation organization from the ECG. In Proceedings of the IEEE Conference on Computing in Cardiology, Belfast, UK, 26–29 September 2010; pp. 1027–1030.

15. Chanwimalueang, T.; Aufegger, L.; von Rosenberg, W.; Mandic, D.P. Modelling stress in public speaking: Evolution of stress levels during conference presentations. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 814–818.

16. Chen, W.; Wang, Z.; Xie, H.; Yu, W. Characterization of surface EMG signal based on fuzzy entropy. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2007**, *15*, 266–272. [CrossRef] [PubMed]

17. Bradley, M.M.; Miccoli, L.; Esgrig, M.A.; Lang, P.J. The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* **2008**, *45*, 602–607. [CrossRef]

18. Chanwimalueang, T.; Mandic, D.P. Cosine Similarity Entropy: Self-Correlation-Based Complexity Analysis of Dynamical Systems. *Entropy* **2017**, *19*, 652. [CrossRef]

19. Ahmed, M.U.; Mandic, D.P. Multivariate multiscale entropy: A tool for complexity analysis of multichannel data. *Phys. Rev. E* **2011**, *84*, 061918. [CrossRef] [PubMed]

20. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* **2002**, *89*, 068102. [CrossRef] [PubMed]

21. Valencia, J.F.; Porta, A.; Vallverdu, M.; Claris, F.; Baranowski, R.; Orłowska-Baranowska, E.; Caminal, P. Refined multiscale entropy: Application to 24-h holter recordings of heart period variability in healthy and aortic stenosis subjects. *IEEE Trans. Biomed. Eng.* **2009**, *56*, 2202–2213. [CrossRef]

22. Ahmed, M.; Rehman, N.; Looney, D.; Rutkowski, T.; Mandic, D. Dynamical complexity of human responses: A multivariate data-adaptive framework. *Bull. Pol. Acad. Sci. Tech. Sci.* **2012**, *60*, 433–445. [CrossRef]

23. Wu, S.D.; Wu, C.W.; Lin, S.G.; Wang, C.C.; Lee, K.Y. Time series analysis using composite multiscale entropy. *Entropy* **2013**, *15*, 1069–1084. [CrossRef]

24. Costa, M.D.; Goldberger, A.L. Generalized multivariate entropy analysis: Application to quantifying the complex volatility of human heartbeat time series. *Entropy* **2015**, *17*, 1197–1208. [CrossRef]

25. Amoud, H.; Snoussi, H.; Hewson, D.; Doussot, M.; Duchene, J. Intrinsische mode decomposition in cooperative brain–computer interface applications with application to uterine EMG complexity analysis. *Entropy* **2017**, *19*, 2. [CrossRef]

26. Costa, M.D.; Peng, C.K.; Goldberger, A.L. Multiscale analysis of heartbeat time series: Entropy and time irreversibility measures. *Cardiowasc. Eng.* **2008**, *8*, 88–93. [CrossRef] [PubMed]

27. Taken, F. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980*; Springer: Berlin/Heidelberg, Germany, 1981; pp. 366–381.

28. Hausdorff, J.M. Gait dynamics in Parkinson’s disease: Common and distinct behavior among stride length, gait variability, and fractal-like scaling. *Chaos Interdiscip. J. Nonlinear Sci.* **2009**, *19*, 026113. [CrossRef] [PubMed]

29. Gao, T.; Mandic, D.P.; Van Hulle, M.M. A differential entropy based method for determining the optimal embedding parameters of a signal. In Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP’03), Hong Kong, China, 6–10 April 2003; Volume 6, p. VI-29.

30. Menayo, R.; Manzanares, A.; Segado, F. Complexity, Regularity and Non-Linear Behavior in Human Eye Movements: Analyzing the Dynamics of Gaze in Virtual Sailing Programs. *Symmetry* **2018**, *10*, 528. [CrossRef]

31. Pincus, S. Approximate entropy as an irregularity measure for financial data. *Econom. Rev.* **2008**, *27*, 329–362. [CrossRef]
33. Paluš, M. Testing for nonlinearity using redundancies: Quantitative and qualitative aspects. *Phys. D Nonlinear Phenom.* 1995, 80, 186–205. [CrossRef]
34. Malik, M.; Bigger, J.T.; Camm, A.J.; Kleiger, R.E.; Malliani, A.; Moss, A.J.; Schwartz, P.J. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Eur. Heart J.* 1996, 17, 354–381. [CrossRef]
35. von Rosenberg, W.; Chanwimalueang, T.; Adjei, T.; Jaffer, U.; Goverdovsky, V.; Mandic, D.P. Resolving ambiguities in the LF/HF ratio: LF-HF scatter plots for the categorization of mental and physical stress from HRV. *Front. Physiol.* 2017, 8, 360. [CrossRef]
36. Berens, P. CircStat: A MATLAB toolbox for circular statistics. *J. Stat. Softw.* 2009, 31, 1–21. [CrossRef]