BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 60, No. 3, 2012
DOI: 10.2478/v10175-012-0055-0

Dynamical complexity of human responses: a multivariate data-adaptive framework

M.U. AHMED[1*], N. REHMAN[1,2], D. LOONEY[1], T.M. RUTKOWSKI[3], and D.P. MANDIC[1]

[1] Department of Electrical and Electronic Engineering, Imperial College London, UK
[2] COMSATS Institute of Information Technology, Park Road, Chak Shahzad, Islamabad, Pakistan
[3] University of Tsukuba & RIKEN Brain Science Institute, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan

Abstract. Established complexity measures typically operate at a single scale and thus fail to quantify inherent long-range correlations in real-world data, a key feature of complex systems. The recently introduced multiscale entropy (MSE) method has the ability to detect fractal correlations and has been used successfully to assess the complexity of univariate data. However, multivariate observations are common in many real-world scenarios and a simultaneous analysis of their structural complexity is a prerequisite for the understanding of the underlying signal-generating mechanism. For this purpose, based on the notion of multivariate sample entropy, the standard MSE method is extended to the multivariate case, whereby for rigor, the intrinsic multivariate scales of the input data are generated adaptively via the multivariate empirical mode decomposition (MEMD) algorithm. This allows us to gain better understanding of the complexity of the underlying multivariate real-world process, together with more degrees of freedom and physical interpretation in the analysis. Simulations on both synthetic and real-world biological multivariate data sets support the analysis.

Key words: multivariate sample entropy, multivariate empirical mode decomposition (MEMD), multivariate multiscale entropy, complexity analysis, multivariate complexity, postural sway analysis, stride interval analysis, brain consciousness analysis, alpha-attenuated EEG data.

1. Introduction

Real-world phenomena typically exhibit complex dynamical behavior, and a number of criteria have been proposed to characterize the underlying signal-generating mechanisms from the observed univariate or multivariate time series. These diverse descriptors include complexity, local predictability, irregularity, self-similarity, and synchrony [1], and operate on a single time scale – that defined through the data acquisition. However, the very essence of complex dynamical behavior are intrinsic correlations (of different natures), not only within single data channels, but even more importantly across data channels and over a range of temporal scales defining system dynamics. Entropy is particularly useful as a measure of the structure of time series, as it reflects the degree of regularity/irregularity of a sequence of data. Despite a number of established entropy measures and the fact that they are well understood, the problem in using entropies in their original form is that they achieve their maximum for signals with no structure (random) and are defined only for a single scale. To assess the structural dynamics of a system across the different time-scales, Costa *et al.* introduced the multiscale entropy (MSE) measure which performs multiple coarse-graining operations on the data (thus defining temporal scales) and calculates sample entropy for each so-defined scale [2, 3]. This way, the MSE method quantifies signal complexity, which remains hidden when using standard methods where temporal scales of a signal are not processed separately.

While the MSE measure has been successfully applied to distinguish between different real-world physiological time series based on their dynamical complexity [4–7], it also has some limitations stemming from the deterministic way of generating multiple scales of input data. The method uses the so-called coarse graining process which, owing to its low-pass filtering characteristics, is unsuitable for the extraction of high frequency components and also results in aliasing (see also Fig. 1), causing potential systemic artifacts which inhibit the analysis. More critically, the coarse graining process reduces the input data length to half its original size for each successive data scale and the 'deterministic' temporal scales do not necessarily match the intrinsic dynamical scales defined by the signal-generating system. As a result, only input data of 'sufficient' length and 'regular' scales can be reliably processed by the MSE method; in turn, for real-world data we can evaluate the MSE only over a limited range of temporal scales. To alleviate these problems, a class of Butterworth filters were used to circumvent the aliasing observed in the original coarse graining process of MSE [8]. However, this does not circumvent the need for data driven scales, over which to perform the analysis.

It was recently proposed to employ a data-driven method, the empirical mode decomposition (EMD) [9], to generate intrinsic multiple data scales from input data, to be used for the subsequent MSE analysis [10, 11]. The resulting EMD-based MSE method produced improved results owing to the fully data-driven nature of EMD and also due to the fact that it operates locally based on the extrema of the (univariate) in-

*e-mail: mosabber.ahmed@imperial.ac.uk

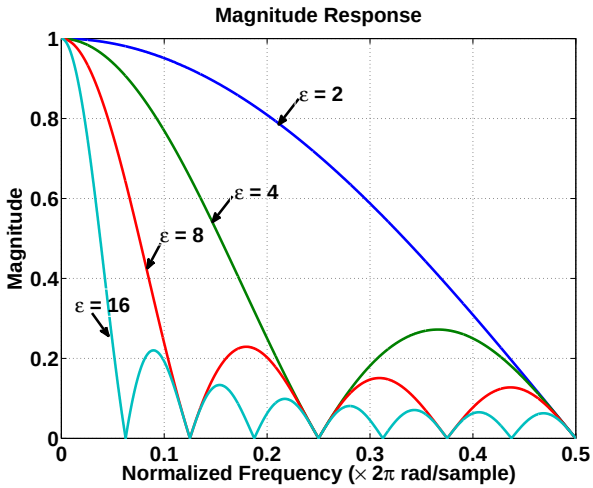


Fig. 1. Magnitude response of the equivalent FIR filter in the coarse graining process for different scale factor ϵ

put signal, yielding well defined narrowband scales intrinsic to the input data. Another benefit of using EMD in conjunction with the MSE method is that the standard MSE fails to cater for nonstationary signals, that is, if a signal contains one or more pronounced trends, little can be inferred from sample entropy – trends tend to dominate other interesting features. From the statistical perspective, it is therefore imperative to remove any trend before meaningful interpretation can be made from MSE analysis. Since EMD decomposes data into narrow-band *quasi-stationary* signals [9], subsequent MSE analysis on each EMD output component promises to facilitate MSE based complexity analysis. For instance, EMD naturally captures a trend in the input data in its residue (last extracted component) which can be removed prior to the MSE analysis.

Recently, advances in sensor and data acquisition technologies have made it possible to record in a coherent way real-world signals containing multiple data channels, with possibly large differences in the dynamics across the channels. For such signals, assessment of cross-statistical properties between multiple input channels is vital for a complete understanding of the underlying signal-generating system. This calls both for the development of multivariate extensions of existing signal processing algorithms in order to directly process multiple channels of input data and, in the process, employ both within- and cross-channel information, and to design data-adaptive algorithms which define multiple intrinsic time scales in multivariate data, both subject of this work.

Recent multichannel extensions of both EMD [12] and MSE algorithm [13–15] have been shown to outperform their standard, univariate, counterparts in the analysis of real-world multivariate signals. The availability of these extensions provides an opportunity to develop a robust framework for the complexity analysis of multivariate data. More specifically, we propose to employ our recently developed multivariate EMD (MEMD) to generate intrinsic data scales for the subsequent multivariate MSE (MMSE) analysis of input multichannel data, and to benefit from the mode alignment property of MEMD, yielding the following advantages:

- This ensures that the scales generated for each data channel are same in number and belong to the same frequency band (monocomponent), which makes their comparison meaningful;
- This way, the limitation of sufficient input data length due to coarse graining process is alleviated since EMD/MEMD generates temporal data scales of same length as the length of the input signal;
- The so-generated time scales are data adaptive and fully suited to the dynamics of the signal in hand, unlike the currently used coarse graining techniques;
- The proposed multivariate MSE complexity assessment method operates on nonstationary data thus bypassing the main limitations of current methods - requirement of stationary data sources.

The presented MEMD-enhanced MMSE method is fully multivariate, unlike e.g. the method given in [11] where MEMD was used to generate multiple scales from the input data and univariate sample entropy was subsequently applied for complexity analysis. The implication of employing univariate sample entropy was that MEMD had to be operated across multiple trials rather than multiple channels, thus, not making use of the full potential of MEMD. Owing to the multivariate nature of the introduced multivariate multiscale entropy measure in conjunction with MEMD, our proposed approach is fully multivariate – it analyzes both the within- and cross-channel information simultaneously and uniquely, it operates on nonstationary multivariate data with large discrepancies in channel dynamics, as demonstrated on the complexity analysis of real-world multivariate biological data.

2. Multivariate empirical mode decomposition

In its original formulation, the empirical mode decomposition (EMD) algorithm decomposes an arbitrary nonlinear and nonstationary signal into a set of simpler monocomponent (narrow-band) functions, and is univariate, that is, it operates only on single channel data. To obtain monocomponent bases common for all the data channels within multivariate signals in a data-adaptive manner, a generalized multivariate extension of EMD (MEMD) has been developed recently [16], which operate directly on multivariate signals, containing any number of data channels.

The operation of the MEMD algorithm rests on the estimation of the local mean of multivariate signals, a key step in EMD-based algorithms which is achieved in single-channel EMD by taking the average of the upper and lower envelopes of the signal in hand, obtained by interpolating the local maxima (upper envelope) and the local minima (lower envelope). For multivariate signals, however, the notion and locations of local maxima and minima cannot be defined directly (especially for complex- and quaternion-valued data). To this end, MEMD operates by taking multiple univariate signal projections, along different directions in p-dimensional spaces,

which are then averaged to obtain the local mean. For this purpose, projection direction vectors are most conveniently generated by uniformly sampling a $p-$sphere[1], for instance, by quasi-Monte Carlo sequences. A suitable set of direction vectors on a $(p-1)$-sphere is typically generated using the low-discrepancy Hammersley sequence, and univariate projections of the multivariate input signal are calculated along this set. The extrema of such projected signals are then interpolated component-wise to yield multidimensional envelopes of a multivariate signal. The multiple envelope curves, each corresponding to a particular direction vector, are then averaged to obtain the multivariate signal mean.

More specifically, consider a sequence of p-dimensional vectors $\mathbf{s}(t) = \{s_1(t), s_2(t), \ldots, s_p(t)\}$, representing a multivariate signal with p components (data channels), and the symbol $\mathbf{x}_{\theta_v} = \{x_1^v, x_2^v, \ldots, x_p^v\}$ denoting a set of $v = 1, 2, \ldots, V$ direction vectors along the directions given by angles $\theta_{\mathbf{v}} = \{\theta_{v_1}, \theta_{v_2}, \ldots \theta_{v_{p-1}}\}$ in $\mathbb{R}^p$. Then, the multivariate extension of EMD (MEMD), suitable for operating on general nonlinear and nonstationary multivariate time series, is summarized in Algorithm 1.

Algorithm 1 . **Multivariate EMD**

1: *Sampling criterion:* Choose a suitable point set for sampling a $(p-1)$-sphere (uniform, equiangular, etc.);
2: *Univariate spatial projections:* Calculate a multidimensional projection, denoted by $q_{\theta_v}(t)$, of the p-variate input signal $\mathbf{s}(t)$ along the direction vector $\mathbf{x}_{\theta_v}$ for all $v \in \theta_v$ (the whole set of direction vectors), giving the set of projections $q_{\theta_v}(t)\}_{v=1}^{V}$;
3: *Extrema finding:* Find the time instants $\{t_{\theta_v}^i\}_{v=1}^{V}$ corresponding to the maxima for every member of the set of projected signals $q_{\theta_v}(t)\}_{v=1}^{V}$;
4: *Envelope detection:* Interpolate $[t_{\theta_v}^i, \mathbf{s}(t_{\theta_v}^i)]$ to obtain multivariate envelope curves $\mathbf{e}_{\theta_v}(t)\}_{v=1}^{V}$;
5: *Local mean calculation:* For a set of V direction vectors, the mean $\mathbf{m}(t)$ of the envelope curves is calculated as:

$$\mathbf{m}(t) = \frac{1}{V}\sum_{v=1}^{V} \mathbf{e}_{\theta_v}(t) \tag{1}$$

6: *Sifting process:* Extract 'detail' $\mathbf{d}(t)$ using $\mathbf{d}(t) = \mathbf{s}(t) - \mathbf{m}(t)$. If $d(t)$ fulfills the stoppage criterion for a multivariate IMF, apply the above procedure to $\mathbf{s}(t) - \mathbf{d}(t)$, otherwise apply it to $\mathbf{d}(t)$.

Once the first IMF is extracted, it is subtracted from the input signal and the same process (steps 1–5 in Algorithm 1) is applied to the resulting signal yielding the second IMF and so on; the process is repeated until all the IMFs are extracted and only the residual is left; in the multivariate case, the residual corresponds to a signal whose projections do not contain enough extrema to form a meaningful multivariate envelope. The sifting process for a multivariate IMF can be stopped when all the projected signals fulfill any of the stoppage criteria adopted in standard EMD. One popular stopping criterion used in EMD is to stop the sifting process when the number of extrema and the zero crossings differ at most by one for S consecutive iterations of the sifting algorithm [17].

Since the multichannel algorithm, MEMD, operates directly on multivariate signals with scales defined by the data, it provides intrinsic information regarding interaction between multiple data channels. This offers us physical insight into the structure of the real-world data and a convenient interpretation:

- The first extracted IMF is the highest frequency component in a signal, containing plenty of detail;
- The subsequent IMFs are ideally narrowband and monocomponent, whereby the characteristic frequency decreases with the IMF number;
- The last IMF – the trend – can often contain the signal power and little signal detail, so that it is typically omitted from analysis.
- Even if the original signal is non-stationary, the IMFs are much better conditioned and are typically quasi-stationary;
- The IMFs are locally orthogonal, providing a parsimonious representation with minimum artifacts in decomposition and reconstruction; the local orthogonality facilitates the 'processing' of non-stationary signals.

Advantages offered by MEMD over the univariate (single-channel) EMD include:

1. Direct processing of multichannel data gives the same number of IMFs for all data channels facilitating the analysis of their properties at each (orthogonal) scale, independently;
2. MEMD automatically *aligns* common scales, present across multiple channels, within its multivariate IMFs; a desirable property that is hard to achieve by applying univariate EMD channel-wise on multivariate data, as shown later in Fig. 4.

Figure 2 demonstrates the advantages of MEMD, on the bivariate decomposition of a complex real-world wind signal with the north-south velocity as its real component and the east-west velocity as its imaginary component, sampled at 50 Hz. Observe that the multivariate (bivariate) extension of EMD generated an *equal number of IMFs* for the real and imaginary parts of the data and therefore admits associating a physical meaning to the components. Moreover, the residual signal r clearly captures the overall 'trend' of the real and imaginary components, the single most misleading feature in traditional dynamical complexity analysis.

Similarly, the property of *alignment of common scales* in multivariate data via MEMD can be illustrated by its quasi-dyadic filter bank structure for multivariate white Gaussian noise (WGN) inputs. For that purpose, simulations were carried out on multiple independent realizations of an 8-channel mutually independent WGN process. The average spectra of the IMFs obtained from $D = 500$ realizations of 8-channel WGN are plotted in Fig. 3, both for standard EMD (bottom

[1] A p-sphere, or equivalently a hypersphere, can be considered as an extension of the ordinary sphere to an arbitrary dimension.

plot) and multivariate EMD (top plot). Observe that statistically, for a given number of noise realizations D, standard EMD failed to accurately align the band-pass filters associated with the sequence of multivariate IMFs shared across the eight noise channels. Although for univariate EMD this alignment is expected to become better with an increase in the number of noise realizations, MEMD-based spectra achieved much better results with the same number of ensembles[2].

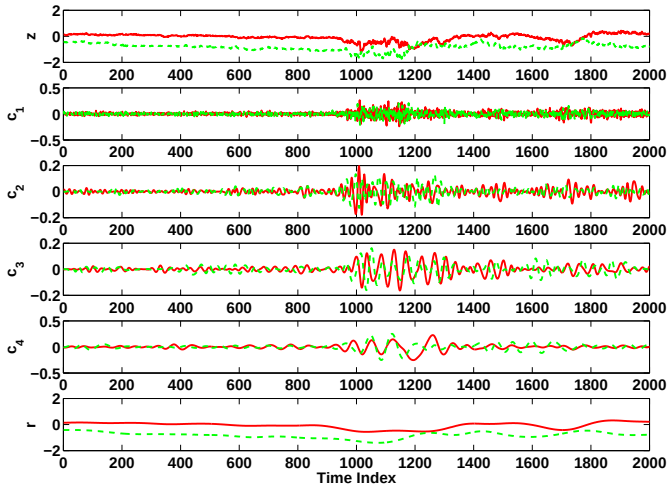


Fig. 2. Complex wind data (denoted by z in the first row) decomposed by bivariate EMD (IMFs denoted by $c_1 - c_4$ in the lower subplots and 'trend' denoted by r in the bottom subplot). Real and imaginary components are depicted by solid and dashed lines, respectively

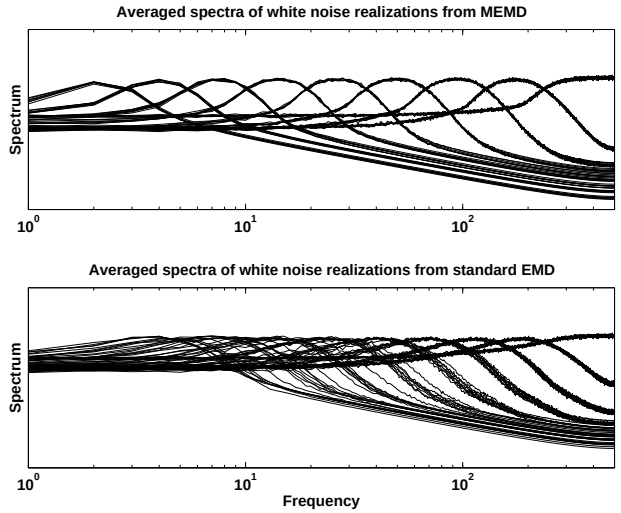
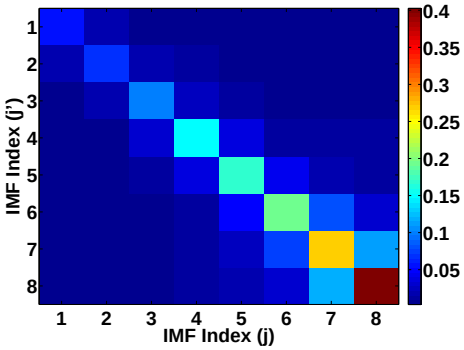


Fig. 3. Averaged spectra of IMFs obtained from $D = 500$ realizations of 8-channel independent white Gaussian noise via MEMD (top) and the standard EMD (bottom). Overlapping of the frequency bands corresponding to the same-index IMFs is improved in both cases after averaging, with MEMD bands showing much better alignment

a) Correlation of IMFs obtained by MEMD



b) Correlation of IMFs obtained by channel-wise EMD

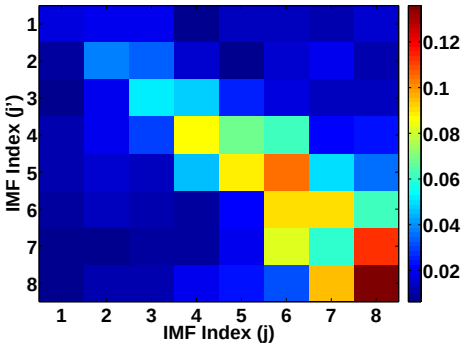


Fig. 4. Normalized IMF cross-correlation from multiple channels of WGN obtained via (a) MEMD and (b) EMD channel-wise. The distribution of higher values of the cross-correlation measure along the diagonal line in the case of MEMD (panel (a)) indicates the excellent mode-alignment between IMFs from multiple channels. The IMF indices grow from left to right and from top to the bottom

A quantitative evaluation of the mode-alignment observed for MEMD-based filter banks is shown in Fig. 4, which shows the normalized intrinsic cross-correlation between IMFs obtained from MEMD and standard EMD, applied on bivariate WGN. In the simulation, $D = 500$ bivariate WGN realizations were used, each of length $N = 1000$. As expected, due to an excellent alignment of the spectra of the corresponding (same-indexed) IMFs from multiple data channels and near-perfect local orthogonality of consecutive IMFs, obtained using MEMD, the cross-correlation matrix across IMFs had a pronounced diagonal ($j = j'$), in Fig. 4(a). In the EMD-based decomposition, however, due to fluctuations in the alignment and local orthogonality of IMFs, non-negligible cross-correlation was observed off-diagonal for ($j \neq j'$), as shown in Fig. 4(b)[3]

3. Multivariate sample entropy

Richman & Moorman [18] introduced sample entropy (SampEn) as a conditional probability that two sequences of *m*break consecutive data points, which are similar to within a tolerance level *r*, will remain similar when the next data point is included, provided that self-matches are not considered in

[2]Data alignment according to their temporal scales across multichannel data is a single most critical feature in multisensor data fusion.

[3]More detail and MATLAB code for MEMD can be found at http://www.commsp.ee.ic.ac.uk/∼mandic/research/emd.htm

calculating the probability. This measure is very stable for data lengths of over several hundred samples [15] and has become a work horse in practical entropy estimation from real-world data. Multivariate sample entropy enables entropy calculation for multichannel data, by taking into account both within- and cross-channel dependencies, and was introduced only recently in [14]. To calculate multivariate sample entropy (MSampEn), recall from multivariate embedding theory [19], that for a *p*-variate time series $\{x_{k,i}\}_{i=1}^{N}, k = 1, 2, \ldots, p$, observed through p measurement functions $h_k(y_i)$, the multivariate embedded reconstruction is based on the composite delay vector

$$\begin{aligned} X_m(i) = [&x_{1,i}, x_{1,i+\tau_1}, \ldots, x_{1,i+(m_1-1)\tau_1}, x_{2,i}, \\ &x_{2,i+\tau_2}, \ldots, x_{2,i+(m_2-1)\tau_2}, \ldots, x_{p,i}, \\ &x_{p,i+\tau_p}, \ldots, x_{p,i+(m_p-1)\tau_p}], \end{aligned} \tag{2}$$

where $\mathbf{M} = [m_1, m_2, \ldots, m_p] \in \mathbb{R}^p$ is the embedding vector, $\boldsymbol{\tau} = [\tau_1, \tau_2, \ldots, \tau_p]$ the time lag vector, the composite delay vector $X_m(i) \in \mathbb{R}^m$ (where $m = \sum_{k=1}^{p} m_k$).

It is important to note that multivariate data do not necessarily have the same range among the data channels, so that the distances calculated on such embedded vectors are typically biased towards the variates with largest ranges. In our proposed formulation of multivariate sample entropy (MSampEn), we scale the data to the range $[0, 1]$. For a *p*-variate time series $\{x_{k,i}\}_{i=1}^{N}, k = 1, 2, \ldots, p$, the multivariate sample entropy (MSampEn) is calculated in Algorithm 2.

3.1. Geometric interpretation of MSampEn. Figure 5 illustrates the principle behind multivariate sample entropy calculation. Consider a real-world bivariate center of pressure (COP) displacement time series in Fig. 5a, where the mediolateral (side-to-side) component is denoted by $x(n)$ (dotted red line) and the anteroposterior (front-to-back) component by $y(n)$ (solid blue line). For illustration, assume the time lag vector $\boldsymbol{\tau} = [1, 1]$ and the embedding vector $\mathbf{M} = [1, 1]$; then the composite bivariate delay vectors are $[x(n), y(n)]$ as shown in Fig. 5b, where n denotes the sample index. In the process of MSampEn calculation, for any such vector (e.g. $[x(64), y(64)]$), we need to count the number of neighbors which are within a distance r (tolerance level), illustrated by a circle centered at $[x(64), y(64)]$ with radius r in Fig. 5b. For an *m*-dimensional space, the set of neighboring vectors would be enclosed by an *m*-sphere if the distance is calculated using the Euclidean norm and by an *m*-cube if we use a maximum distance norm. The average number of composite delay vectors that are within a fixed threshold r (so called r-neighbors) in this two-dimensional space is next calculated, which serves as an estimate of the local probability density, and is also a measure of their joint probability, as all the m-components of the neighboring vector (as maximum distance norm is used) have to be simultaneously similar to those of the vector in hand. By increasing the embedding dimension m, we therefore inherently involve joint probabilities covering larger time spans.

Algorithm 2 . The Multivariate sample Entropy (MSampEn)

1: Form $(N - \delta)$ composite delay vectors $X_m(i) \in \mathbb{R}^m$, where $i = 1, 2, \ldots, N - \delta$ and $\delta = \max\{M\} \times \max\{\boldsymbol{\tau}\}$.

2: Define the distance between any two composite delay vectors $X_m(i)$ and $X_m(j)$ as the maximum norm, that is, $d[X_m(i), X_m(j)] = \max_{l=1,\ldots,m}\{|x(i + l - 1) - x(j + l - 1)|\}$.

3: For a given composite delay vector $X_m(i)$ and a threshold r, count the number of instances P_i for which $d[X_m(i), X_m(j)] \leq r$, $j \neq i$, then calculate the frequency of occurrence, $B_i^m(r) = \frac{1}{N-\delta-1} P_i$, and define

$$B^m(r) = \frac{1}{N - \delta} \sum_{i=1}^{N-\delta} B_i^m(r). \tag{3}$$

4: Extend the dimensionality of the multivariate delay vector in (3) from m to $(m+1)$. This can be performed in *p* different ways, as from a space with the embedding vector $\mathbf{M} = [m_1, m_2, \ldots, m_k, \ldots, m_p]$ the system can evolve to any space for which the embedding vector is $[m_1, m_2, \ldots, m_k + 1, \ldots, m_p]$ (k=1, 2, ..., p). Thus, a total of $p \times (N - \delta)$ vectors $X_{m+1}(i)$ in $\mathbb{R}^{m+1}$ are obtained, where $X_{m+1}(i)$ denotes any embedded vector upon increasing the embedding dimension from m_k to $(m_k + 1)$ for a specific variable *k*. In the process, the embedding dimension of the other data channels ($\neq k$) is kept unchanged, so that the overall embedding dimension of the system undergoes the change from m to $(m + 1)$.

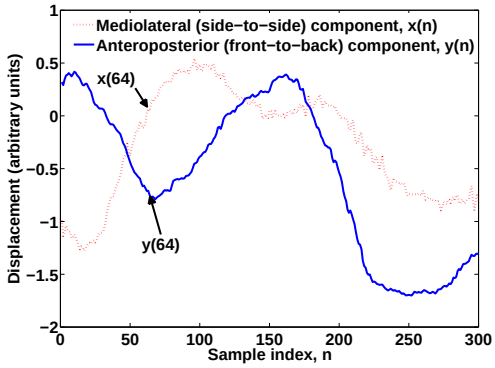
5: For a given $X_{m+1}(i)$, calculate the number of vectors Q_i, such that $d[X_{m+1}(i), X_{m+1}(j)] \leq r$, where $j \neq i$, then calculate the frequency of occurrence, $B_i^{m+1}(r) = \frac{1}{p(N-\delta)-1} Q_i$, and define

$$B^{m+1}(r) = \frac{1}{p(N - \delta)} \sum_{i=1}^{p(N-\delta)} B_i^{m+1}(r). \tag{4}$$

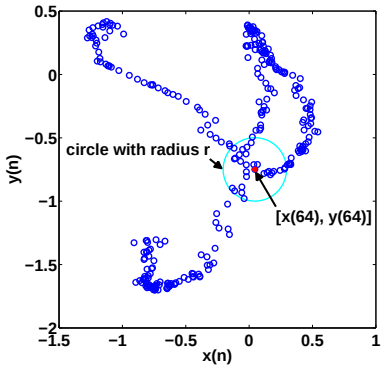
6: Finally, for a tolerance level *r*, estimate $MSampEn$ as

$$MSampEn(\mathbf{M}, \boldsymbol{\tau}, r, N) = -ln \left[\frac{B^{m+1}(r)}{B^m(r)} \right]. \tag{5}$$

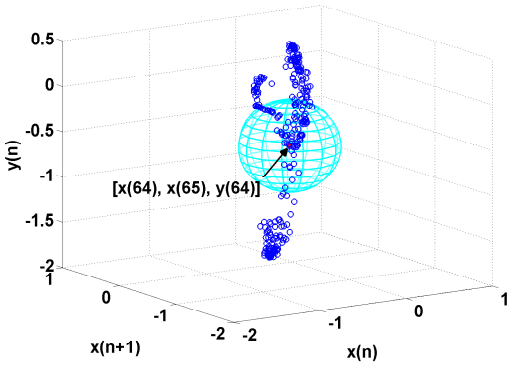
a) COP time series



b) A 2D scatter plot of vectors $[x(n), y(n)]$



c) A 3D-plot of vectors $[x(n), x(n+1), y(n)]$



d) A 3D-plot of vectors $[x(n), y(n), y(n+1)]$

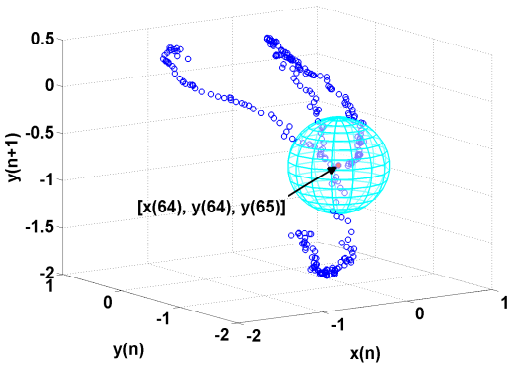


Fig. 5. Geometry behind the calculation of MSampEn

For the example in Fig. 5, upon increasing the embedding dimension from two to three, we have two possible subspaces of dimension three: (i) the subspace of all the vectors $[x(n), x(n+1), y(n)]$ shown in Fig. 5c, and (ii) the subspace of all the vectors $[x(n), y(n), y(n+1)]$ shown in Fig. 5d. A naive approach would be to calculate the number of vectors that are within a fixed threshold r in each three-dimensional subspace and then average over both subspaces [13]. Instead, we employ a rigorous approach and compare composite delay vectors (to find the neighbors) not only within each subspace but also across all the subspaces, thus fully catering for both within- and cross-channel correlations. This allows us to calculate the conditional probability that two sequences of *m* data points (or two composite delay vectors in *m*-dimensional space), which are similar to within a tolerance level *r*, will remain similar in the same sense, when the next data point is included (or equivalently the dimension of the composite delay vector is increased by one), provided that self-matches are not considered. A negative logarithm of this conditional probability defines the multivariate sample entropy (see Eq. (5) in Algorithm 2)[4].

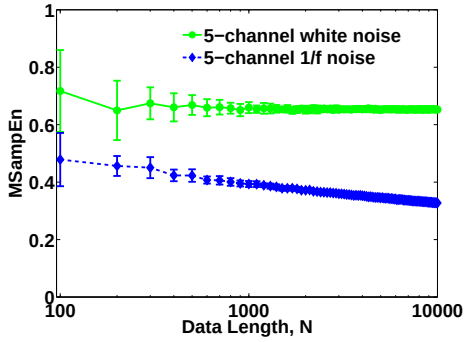
3.2. Effect of data length and parameters on MSampEn. It has been suggested in [20] that $10^m - 20^m$ data samples are sufficient to robustly estimate univariate approximate entropy or sample entropy. To assess the sensitivity of the proposed multivariate sample entropy to the data length parameter, we evaluated multivariate sample entropy of a 5-channel white as well as $1/f$ noise as a function of sample size N, where for each channel the embedding dimension $m_k = 2$ and the threshold $r = 0.20$ was taken. Figure 6a shows that for both the white and $1/f$ 5-channel noise, MSampEn estimates were consistent for data length $N \geq 300$. Note that, the greater the value of N, the more robust the MSampEn estimates as seen from the errorbars in Fig. 6a. MMSE (MSE) calculates MSampEn (SampEn) for different scales generated by the coarse-graining process and the length of each coarse-grained time series is equal to the length of the original time series divided by the scale factor, ϵ. The coarse graining procedure of the standard MMSE approach thus imposes the constraint that the highest scale should have enough data points (at least 300 points) to be able to calculate a valid entropy estimate. This somewhat limits the applicability of coarse graining based MMSE for very short real-world data. Our proposed MEMD-enhanced MMSE overcomes this limitation as the decomposed IMFs have the same lengths as that of the original signal.

Physically, for the standard univariate sample entropy, the increase in sample entropy values with an increase in embedding dimension m is due to progressively fewer delay vectors to compare as m increases. On the contrary, for MSampEn the increase in m does not reduce the number of the available delay vectors, as the composite multivariate embedded vectors are constructed in parallel. Figure 6b presents an MSampEn vs m_k plot for a 5-channel white and 1/f noise with 40,000

[4]The MATLAB code for MMSE can be found at http://www.commsp.ee.ic.ac.uk/∼mandic/research/Complexity_Stuff.htm

points in each channel, and illustrates that the result does not depend on m_k (where the number of channels, p, is 5 and m_k is taken up to 4, resulting in a maximum total embedding dimension of 20). For large m_k, the length of the time series required for a valid entropy estimate would be prohibitively large (for $m_k = 5$, we need at least 10^5 data points in each channel), and therefore for practical purposes, $m_k = 1$ or 2 is usually used in the literature. In SampEn calculation, the threshold parameter r is set to some percentage of the standard deviation; for MSampEn, we used its multivariate generalization - the total variation, $tr(S)$, where S is the covariance matrix. To maintain the same total variation for all the multivariate series, the individual data channels were normalized to unit variance so that the total variation equals the number of channels/variables. This way, we take r as a percentage (say 20%) of $tr(S)$, which is similar to taking $r = 0.20$ in each channel.

a) Sensitivity to data length, N



b) Sensitivity to embedding parameter, m_k

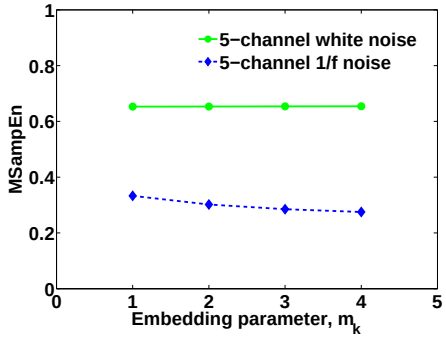


Fig. 6. Multivariate sample entropy as a function of: (a) data length N where $r = 0.20$ and $m_k = 2$ in each data channel and (b) embedding parameter m_k, where each channel has 40,000 samples and $r = 0.20$ in each data channel. Shown are the mean values for 30 simulated 5-channel time series containing white and $1/f$ noise

4. Multivariate complexity analysis

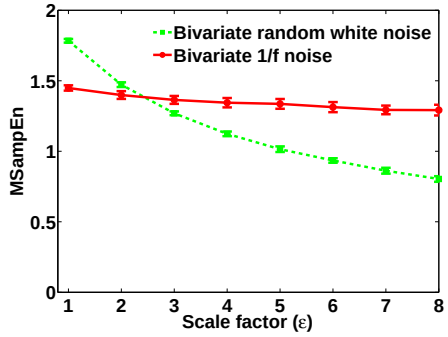
In this section, the standard coarse graining based multivariate multiscale entropy is described first, followed by the proposed MEMD-enhanced multivariate multiscale entropy. The MMSE curves generated by both these methods for multichannel white Gaussian noise as well as $1/f$ noise are compared and their interpretation is illuminated.

4.1. Standard multivariate multiscale entropy. The coarse graining based multivariate MSE (MMSE) method assesses relative complexity of normalized multi-channel temporal data by plotting multivariate sample entropy as a function of the scale whereby:

- A multivariate time series is considered more structurally complex than another if, for the majority of time scales, its multivariate entropy values are higher than those of the other time series;
- A monotonic decrease in multivariate entropy values with the scale factor reveals that the signal in hand only contains information at the smallest scale, and is thus not dynamically complex;
- A constant MSE curve over all the scales indicates long term correlations in the data, a signature of truly complex systems.

For instance, Fig. 7a shows the standard multivariate MSE analysis [14, 15] for bivariate random white noise (uncorrelated), conforming with the interpretation that the MSampEn values monotonically decrease with scale, whereas for $1/f$ noise (long-range correlated) the MSampEn remains constant over multiple scales. This has a physical justification, as by design $1/f$ noise is structurally more complex than uncorrelated white noise.

a) Multivariate MSE



b) MEMD-enhanced Multivariate MSE

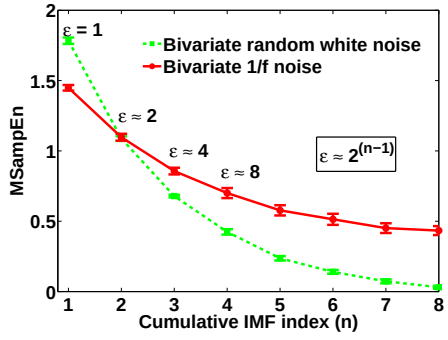


Fig. 7. Multivariate multiscale entropy (MMSE) analysis for bivariate white and 1/f noise, each with 5,000 data points using: a) coarse graining based standard multivariate MSE, and b) MEMD-enhanced multivariate MSE. The curves represent an average of 20 independent realizations and error bars the standard deviation (SD)

4.2. A note on complexity of biological data. Traditional entropy measures, such as Shannon entropy [21], Kolmogorov-Sinai (KS) entropy [22], approximate entropy (ApEn) [23] and sample entropy (SampEn) [18], are maximized for completely random processes, and are frequently used to quantify the regularity (predictability) of time series on a single scale, by e.g. evaluating repetitive patterns [2]. This, in turn, makes it difficult to use standard entropies as complexity measures, since:

- Physiological data, which exhibit a high degree of structural richness, have lower entropy than their randomized surrogates, formed by shuffling the original data;
- The greater entropy of the so-generated surrogate series (which are less complex) also signifies a lack of simple correspondence between regularity and complexity;
- Neither completely predictable (e.g. periodic) signals nor completely unpredictable (e.g. uncorrelated random) signals are truly complex, since at a global level they do admit a simple description.

The behavior of truly complex systems is far from perfect regularity or complete randomness, and instead time series coming from dynamical physical and physiological systems generally exhibit long-range correlations on multiple spatial and temporal scales. To that end, multiscale entropy analysis aims at quantifying the interdependence between entropy and scale, achieved by evaluating sample entropy of univariate time series coarse grained at multiple temporal scales. This facilitates the assessment of the dynamical complexity of a system; in biology this is associated with the ability of living systems to adjust to a changing environment. The underlying integrative multiscale functionality is interpreted by non-diminishing entropy values across increasing time scales.

4.3. MEMD-enhanced multivariate multiscale entropy. To alleviate the problems of trend and shortening of available data with the temporal scale, we propose to use MEMD to generate multiple scales of a given multivariate data, adaptively, and subsequently perform multivariate entropy analysis on separate or cumulative[5] IMFs (scales). For this cause, fully aligned scales from input multivariate data are first obtained by applying MEMD both *across multiple channels and multiple conditions of the input data*. Next, multivariate sample entropy estimates are calculated for the so-defined 'scales' of the multivariate input data to reveal its long-range correlation structure; the algorithm is listed in Algorithm 3.

To illustrate the performance of the proposed method, it was applied to a synthetically generated bivariate white noise and bivariate $1/f$ noise. The $1/f$ noise possesses long-range correlations and its standard entropy (at scale 1) is lower than that of white noise, however, the $1/f$ noise is structurally complex whereas the bivariate white noise is not, and any complexity measure should be higher for $1/f$ noise at increasing scales. Observe from Fig. 7b that though bivariate white noise has higher complexity than $1/f$ noise for the first scale, the complexity becomes lower than $1/f$ noise for higher scales. This example on synthetic data illustrates, that by design, $1/f$ noise is structurally more complex than uncorrelated random noise, a result consistent with standard MSE/MMSE [2, 14, 15] as shown in Fig. 7a.

Algorithm 3. MEMD-enhanced multivariate multiscale entropy

1: Generate multiple scales from J IMFs obtained by applying MEMD to a given multivariate time series $\{x_{k,i}\}_{i=1}^{N}$ for $k = 1, 2, \ldots, p$, where p denotes the total number of variates (channels) and N represents the total number of samples in each variate which does not change across MEMD-based scales.
2: Define data-driven 'scales' of x as the cumulative sum of IMFs either by $\mathbf{c}_n = \sum_{j=n}^{J} \mathbf{c}_j$ (Approach 1) or by $\mathbf{c}_n = \sum_{j=1}^{J-n+1} \mathbf{c}_j$ (Approach 2), where $n \in [1, J]$ denotes the cumulative IMF index, and $\mathbf{c}_j$ denotes the j-th IMF. *Only Approach 1 is used in the sequel.*
3: Calculate and plot multivariate sample entropy measure, given in (5), for each scale n.

Remark 1. A direct comparison is often not possible between the scales of MEMD-enhanced MSE and those of standard MSE as, by design, the frequency ranges of the cumulative IMFs adapts to the data. In the case of white noise, however, the dyadic filter bank property of MEMD is well known [12]. Disregarding elements of coarse graining[6], the averaging operation at scale ϵ is equivalent to low pass filtering with a cutoff frequency (normalized) of $f_c = 0.5/\epsilon$. Thus for the nth cumulative IMF index (Approach 1) of white noise, the equivalent coarse grained scale factor is given by $\epsilon \approx 2^{n-1}$. For insight, the equivalent scale factors for white noise are shown for cumulative IMF indexes in Fig. 7b.

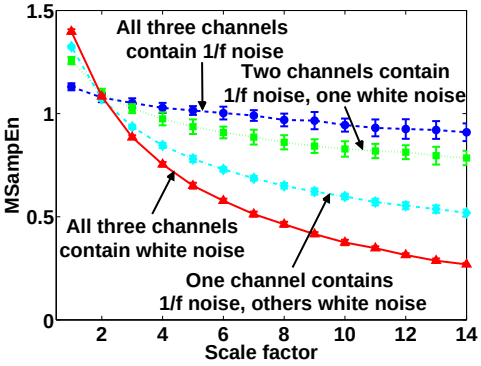
To further illustrate this behavior for the multicahannel scenario where different channels contain different noise realizations, we next generated a trivariate time series, where originally all the data channels were realizations of mutually independent white noise. We then gradually decreased the number of variates that represent white noise (from 3 to 0) and simultaneously increased the number of data channels that represent independent $1/f$ noise (from 0 to 3), so that the total number of variates was always three. Figure 8a shows the standard coarse graining based MMSE curves and Fig. 8b shows MEMD-enhanced MMSE curves for the cases considered; notice that as the number of variates representing $1/f$ noises increased, MSampEn at higher scales also increased, and when all the three data channels contained $1/f$ noise, the complexity at larger scales was the highest. The analysis

[5] Due to their narrowband nature, an alternative option is to additionally apply coarse graining to the IMF-scales themselves with minimal risk of aliasing.

[6] The filtering operation equivalent to coarse graining is characterized by a very slow roll-off as well as large sidelobes which introduce aliasing artifacts [8]. The equivalent relationship between scale factor and cumulative IMF index given in the paper assumes a considerably faster roll-off as well as the absence sidelobes.

in Fig. 8 therefore confirms that, as desired, the more variables/channels within a multivariate time series exhibit long range correlations, the higher the overall complexity of the underlying multivariate system.

a) Multivariate MSE



b) MEMD-enhanced Multivariate MSE

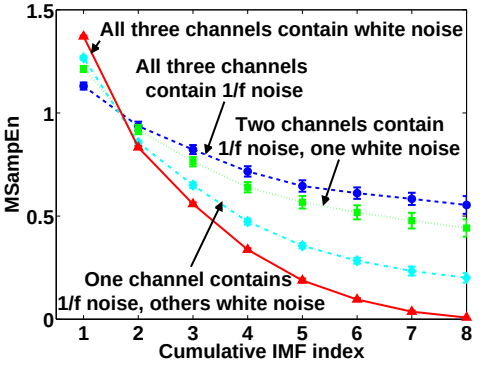


Fig. 8. Multivariate multiscale entropy (MMSE) analysis for 3-channel data containing white and 1/f noise, each with 10,000 data points using: a) coarse graining based standard multivariate MSE, and b) MEMD-enhanced multivariate MSE. The curves represent an average of 20 independent realizations and error bars the standard deviation (SD)

Remark 2. Unlike EMD/MEMD based sample entropy methods given in [10] and [11] which employ univariate sample entropy, the proposed method is fully multivariate as it calculates directly multivariate sample entropy estimates, thereby catering for linear/nonlinear correlations both within and between the data channels.

Remark 3. Standard MMSE calculates MSampEn for different scales generated by the coarse-graining process and the length of each coarse-grained time series is equal to the length of the original time series divided by the scale factor, ϵ. Thus, one needs to ensure that the highest scale has enough data points to be able to calculate a valid entropy estimate. In practical cases, if the data length is very short, applying coarse graining-based MMSE is not feasible. On the contrary, MEMD-enhanced MMSE can be applied for input signals having very short data lengths.

5. Experimental results

The multivariate multiscale entropy analysis is next evaluated for several multivariate real-world recordings: human postural sway analysis, stride interval time series analysis, alpha-attenuated EEG data analysis and brain consciousness analysis.

5.1. Postural sway analysis. Multivariate complexity analysis of real-world postural sway dynamics time series, measuring center of pressure (COP) displacement, was performed for young and elderly subjects during quiet standing. It is expected that young healthy subjects exhibit unconstrained movement with high complexity, whereas the postural sway of the elderly is expected to be constrained, thus having lower complexity, as compared to the young subjects. This conforms with the analysis in Subsec. 4.2.

The data used in the simulations was the measured COP displacement from 12 young and 12 elderly, all healthy, volunteers at 60 Hz; it was recorded simultaneously for the mediolateral (side-to-side) and anteroposterior (front-to-back) direction [24] to form a bivariate time series. Since the postural sway time series exhibits high frequency fluctuations superimposed on low frequency trends, the data need first to be detrended using methods such as wavelet transform and/or empirical mode decomposition [9]. It was shown in the literature [15, 25] that only after detrending (removing last few IMFs), a valid complexity analysis could be performed using multivariate multiscale entropy method.

To illustrate this point, we first present the multivariate complexity analysis of the COP data by applying coarse graining-based MMSE without any detrending, as shown in Fig. 9a. Unless otherwise stated, the values of the parameters used to calculate MSampEn were $m_k = 2$, $\tau_k = 1$, and $r = 0.15\times$(standard deviation of the normalized time series) for each data channel for all the experimental results. The parameters were chosen on the basis of previous studies indicating good statistical reproducibility for the univariate SampEn [18, 25]. Since the original time series had 1.8×10^3 data points, the highest achievable scale factor ($\epsilon = 6$) in standard coarse graining-based MMSE contained 300 data points; this was sufficient for accurate analysis, as shown in [15].

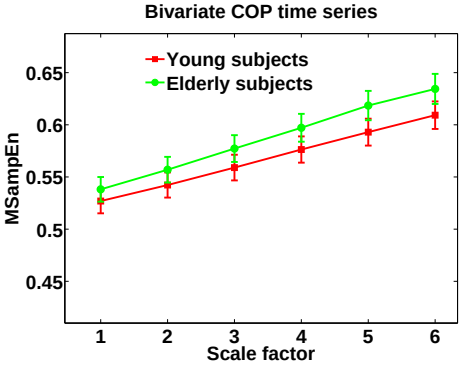
It is evident from Fig. 9a that the standard coarse graining-based MMSE, without detrending, was not able to discriminate between young and elderly subjects as the error bars overlapped. Moreover, the mean MMSE curve was lower for healthy young subjects than for their elderly counterparts, which is contrary to the intuition and the underlying physics[7], and is attributed to the shortcomings of the scale generation method (coarse graining) of standard MMSE.

Next, Fig. 9b shows the complexity analysis results obtained by applying MEMD-enhanced MMSE method taking only the first five IMFs. From the figure, it is evident that MEMD-enhanced MMSE was able to discriminate between young and elderly subjects more effectively, as indicated by a better separation of their MMSE curves. Moreover, the com-

[7]Due to ageing and the associated constraints, the complexity of postural sway for the elderly should be lower than for the young.

plexity was higher for healthy young subjects than for their elderly counterparts which is physically intuitive, as for young subjects there exist correlation between mediolateral and anteroposterior components of COP time series which degrades in elderly subjects. Moreover, the difference in complexity between the young and elderly subjects was found to be statistically significant ($p < 0.01$) over scales 1–4 using one-tailed t-test with unequal variances. This illustrates significant advantages of using MEMD-enhanced MMSE when assessing relative complexity of real-world multivariate data, and supports the more general concept of multiscale complexity loss with ageing and disease or when a system is under constraints, as those factors reduce the adaptive capacity of biological organization at all levels [26].

a) Multivariate MSE



b) MEMD-enhanced Multivariate MSE

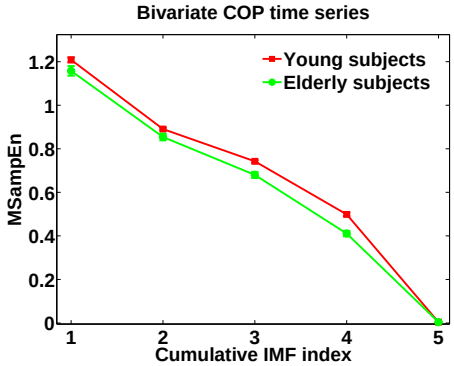
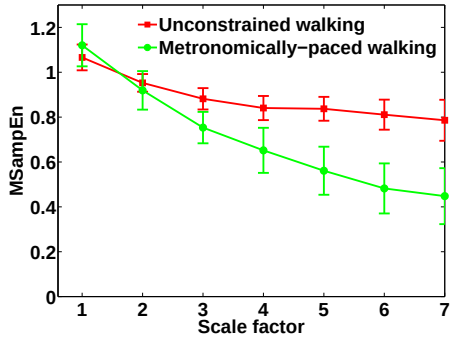


Fig. 9. Bivariate multiscale entropy analysis of COP time series for young (red square) and elderly (green circle) subjects using: a) coarse graining based standard multivariate MSE, and b) MEMD-enhanced multivariate MSE. The plots represent mean values of MSampEn for all subjects across all the trials and error bars represent standard error

5.2. Stride interval analysis. In order to reveal long-range correlations in stride interval dynamics, a signature that suggests cooperation within the different bodily subsystems at different time scales, stride intervals from human gait [27] data were analyzed next. Stride interval fluctuations were recorded from ten healthy subjects who walked for 1 hour at normal, slow, and fast paces and also walked following a metronome set to each participant's mean stride interval [27]. Three walking conditions (from the data available from [27]) were considered as different variables from the same system, and MSampEn values were calculated for different scales (cumulative IMFs) generated using MEMD and in this way we were able to discriminate between the 'self-paced' and 'metronomically-paced' walk. Figure 10a shows the results obtained by the standard coarse-graining based MMSE method and Fig. 10b for the proposed MEMD-enhanced method. Both methods found that self-paced 'unconstrained' walk has higher complexity, and thereby exhibits greater long-range correlations compared to constrained 'metronomically-paced' walk.

a) Multivariate MSE



b) MEMD-enhanced Multivariate MSE

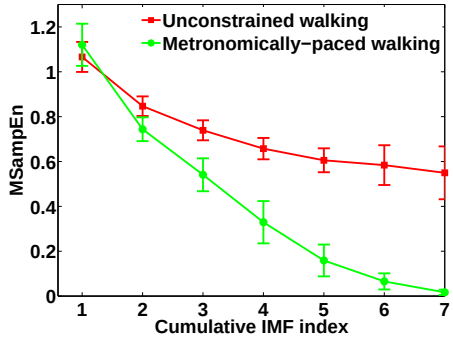


Fig. 10. MMSE analysis for self-paced (red square) vs metronomically-paced (green circle) stride interval time series: using: a) coarse graining based standard multivariate MSE, and b) MEMD-enhanced multivariate MSE. The curves represent an average over 10 subjects, and the error bars the SD

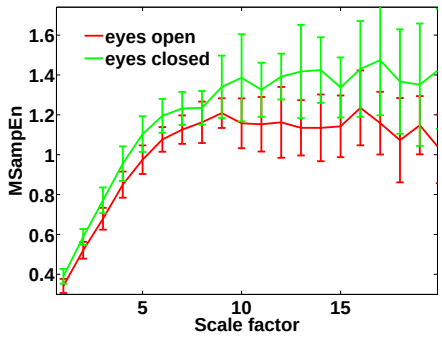
The statistical difference of the entropy statistics of self-paced and metronomically-paced sets were evaluated using the Student's t-test and the Mann-Whitney U test. Both these tests revealed significant differences ($p < 0.01$) at all IMF-defined scales except the first two for standard coarse graining as well as MEMD-enhanced MMSE method. Observe that the first scale corresponds to the raw signal and MSampEn measures cannot discriminate between self-paced and metronomically-paced walk in either method. Moreover, as desired the separation between the MMSE curves of unconstrained and metronomically-paced walk was higher for the MEMD-enhanced method (Fig. 10b), as indicated by much smaller error bars. Thus, using cumulative IMFs as data-adaptive scales offers a significant improvement over the coarse-graining based MMSE. These results also agree with

the more general concept of multiscale complexity loss with ageing and disease or when a system is under constraints (metronomically-paced walk), which all reduce the adaptive capacity of biological organization at all levels [26].

5.3. Structural complexity of different brain states. The proposed algorithm was applied to multivariate electroencephalogram (EEG) signals to detect changes in brain states. The alpha component in EEG (8–12 Hz) increases with the closing of the eyes and is closely linked to the state of alertness. Correlations are introduced by the alpha response, therefore destroying the $1/f$ nature of the standard EEG response and causing an increase in complexity at the corresponding scales. Several EEG recordings were made on the same subject for the states of 'eyes open' and 'eyes closed' at a sampling frequency of 512 Hz from electrodes Cz and POz based on the 10–20 system[8]. Recording trials for both states can be represented as a composite $2 \times N_e \times N_{tr}$-variate vector where $N_e = 2$ denotes the number of electrode channels, $N_{tr} = 7$ denotes the number of trials, and the length of the vector equals the sample length of each recording. A single MEMD operation was performed on the composite delay vector, ensuring aligned scales across: trials; electrode channels; and different brain states.

The average complexity analysis over all the trials is shown in Fig. 11a using standard coarse graining-based MMSE and in Fig. 11b using MEMD-enhanced MMSE. Note that while separation between the 'eyes closed' and 'eyes open' states of alertness was not possible using standard MMSE (the error bars overlap for every scale in Fig. 11a), the adaptive nature of MEMD-enhanced MMSE enabled a clear separation between the states. Conforming with theory, in Fig. 11b the alpha response caused a significant increase in complexity at scales corresponding to the alpha frequency range (2 and 3), such that the error bars do not overlap.

a) Multivariate MSE



b) MEMD-enhanced Multivariate MSE

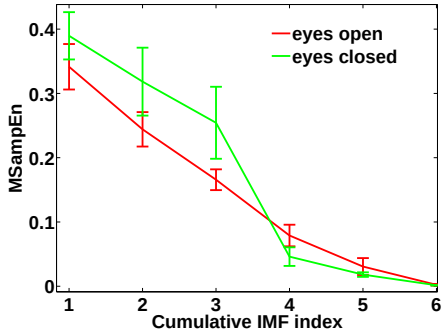


Fig. 11. MMSE analysis for 'eyes open' and 'eyes closed' EEG using: a) coarse graining based standard multivariate MSE, and b) MEMD-enhanced multivariate MSE. The curves represent an average of 7 trials and the error bars the SD

5.4. Brain consciousness analysis. Next, we evaluated standard as well as MEMD-enhanced multivariate multiscale entropy (MMSE) for the characterization of brain consciousness, particularly, the coma and quasi-brain-death state. The legal definition of brain death is an irreversible loss of forebrain and brainstem functions [28], however, brain death diagnosis procedures are complicated, and some tests require temporary disconnection from medical support. An initial prognosis of quasi-brain-death (QBD) is given based on various methods used for studying brain states using electroencephalogram (EEG) [29]. Studies have shown that large activity in the alpha band reflects the alertness of a patient [30], however, standard spectral analyzes are unable to yield information of the brain's inherent nonlinear complex dynamics [31], an important feature for brain states diagnosis. It is natural to assume that a brain in the states of coma and quasi-brain-death would have different degrees of complexity, and that the more stressed the system (QBD) the lower the complexity. As a result, methods from nonlinear dynamics theory such as MMSE are a natural choice in this context [32].

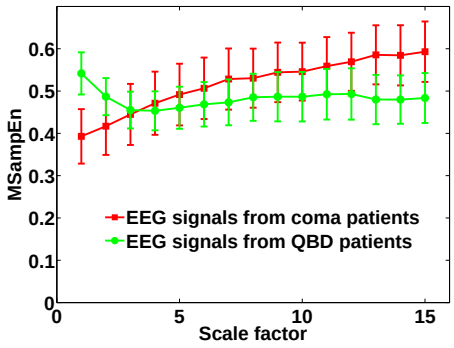
The EEG data were recorded in the intensive care unit in Hua Shan Hospital, Shanghai, China using a standardized 10–20 system. The measured voltage signal was digitized via a portable EEG recording instrument with a sampling frequency of 1000 Hz. The data was then bandpass filtered (FIR filter) to retain frequencies within the range 1–40 Hz and then downsampled by a factor of 10. Experimental data were obtained from 10 patients in coma, and 10 in the quasi-brain-death (QBD) state. For each patients, 50s segment are taken from the EEG signal.

The coarse graining based multivariate multiscale entropy (MMSE) method was first applied over all the six electrodes (FP1, FP2, F3, F4, F7, F8) from both coma and quasi-brain-death patients. Figure 12a shows the coarse graining based standard MMSE results for EEG signals from both coma and quasi-brain-death patients. Observe that, although the coma patients had higher complexity in EEG than the quasi-brain-death patients, the error bars overlapped. Figure 12b shows the result for the MEMD-enhanced MMSE method considering only first seven IMFs. In this case, separation in error bars are observed in cumulative IMF index 4, 5 and 6. In both cases, this indicates a reduction in the intra-cortical information flow and lower neuronal process in the brain for the QBD patients. This also shows inhibition of previously active networks or a loss of dynamical brain responsiveness to the environmental

[8]The data was recorded using a g.tec g.USBamp biosignal amplifier and then bandpass filtered to retain frequencies within the range 1–45 Hz.

conditions and supports the more general concept of multiscale complexity loss with aging and disease [26].

a) Multivariate MSE



b) MEMD-enhanced Multivariate MSE

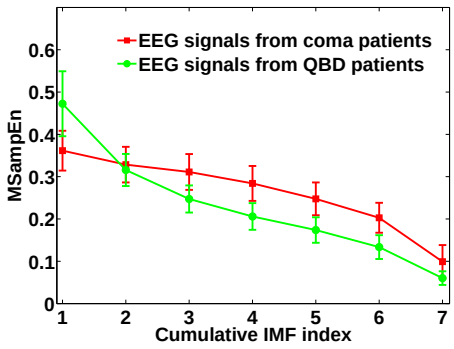


Fig. 12. MMSE analysis for EEG signals from coma (red square) vs QBD (green circle) patients: using: a) coarse graining based standard multivariate MSE, and b) MEMD-enhanced multivariate MSE. The curves represent an average over 10 subjects, and the error bars represent standard error

6. Conclusions

By introducing two critical algorithm enhancements – multivariate sample entropy and rigorous account of data adaptive scales, a robust framework for complexity analysis of multivariate time series has been developed. The proposed algorithm has been shown to be both suitable to alleviate problems arising from data nonstationarity and to detect and quantify complex dynamics within the multiple data channels, prerequisites for complexity analysis of real-world systems which are typically of multivariate, coupled and noisy natures. Simulations on different multivariate physiological signals have shown the effectiveness of the proposed approach in revealing long-range spatio-temporal correlations.

Acknowledgements. We wish to thank Prof Jianting Cao from Saitama Institute of Technology, Japan for providing us with the EEG datasets from coma and quasi-brain-death (QBD) patients used in the analysis.

REFERENCES

[1] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, 2000.

[2] M. Costa, A.L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series", *Phys. Rev. Lett.* 89 (6), 068102 (2002).

[3] M. Costa, A.L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of biological signals", *Phys. Rev.* E 71 (2), 021906 (2005).

[4] R. Hornero, D. Abásolo, J. Escudero, and C. Gomez, "Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease", *Phil. Trans. R. Soc.* A 367 (1887), 317–336 (2009).

[5] M. Costa, C.-K. Peng, A.L. Goldberger, and J.M. Hausdorff, "Multiscale entropy analysis of human gait dynamics", *Physica A: Statistical Mechanics and Its Applications* 330 (1–2), 53–60 (2003).

[6] T. Takahashi, R.Y. Cho, T. Murata, T. Mizuno, M. Kikuchi, K. Mizukami, H. Kosaka, K. Takahashi, and Y. Wada, "Age-related variation in EEG complexity to photic stimulation: A multiscale entropy analysis", *Clinical Neurophysiology* 120 (3), 476–483 (2009).

[7] M. Costa, I. Ghiran, C.-K. Peng, A. Nicholson-Weller, and A.L. Goldberger, "Complex dynamics of human red blood cell flickering: Alterations with in vivo aging", *Phys. Rev.* E 78 (2), 020901 (2008).

[8] J.F. Valencia, A. Porta, M. Vallverdu, F. Claria, R. Baranowski, E. Orlowska-Baranowska, and P. Caminal, "Refined multiscale entropy: Application to 24-h holter recordings of heart period variability in healthy and aortic stenosis subjects", *IEEE Trans. on Biomedical Engineering* 56 (9), 2202–2213 (2009).

[9] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, Nai-Chyuan Yen, C.C. Tung, and H.H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proc. Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 454 (1971), 903–995 (1998).

[10] H. Amoud, H. Snoussi, D. Hewson, M. Doussot, and J. Duchene, "Intrinsic mode entropy for nonlinear discriminant analysis", *IEEE Signal Processing Letters* 14 (5), 297–300 (2007).

[11] M. Hu and H. Liang, "Adaptive multiscale entropy analysis of multivariate neural data", *IEEE Trans. on Biomedical Engineering* 59 (1), 12–15 (2012).

[12] N. Rehman and D.P. Mandic, "Filter bank property of multivariate empirical mode decomposition", *IEEE Trans. on Signal Processing* 59 (5), 2421–2426 (2011).

[13] M.U. Ahmed, L. Li, J. Cao, and D.P. Mandic, "Multivariate multiscale entropy for brain consciousness analysis", *Proc. IEEE Int. Conf. Engineering in Medicine and Biology Society (EMBC'11)* 4, 810–813 (2011).

[14] M.U. Ahmed and D.P. Mandic, "Multivariate multiscale entropy: A tool for complexity analysis of multichannel data", *Phys. Rev.* E 84, 061918 (2011).

[15] M.U. Ahmed and D.P. Mandic, "Multivariate multiscale entropy analysis", *IEEE Signal Processing Letters* 19 (2), 91–94 (2012).

[16] N. Rehman and D.P. Mandic, "Multivariate empirical mode decomposition", *Proc. Royal Society A: Mathematical, Physical and Engineering Science* 1, CD-ROM (2009).

[17] N.E. Huang, M.C. Wu, S.R. Long, S.S.P. Shen, W. Qu, P. Gloersen, and K.L. Fan, "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis", *Proc. Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 459 (2037), 2317–2345 (2003).

[18] J.S. Richman and J.R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy", *AJP – Heart and Circulatory Physiology* 278 (6), H2039–2049 (2000).

[19] L. Cao, A. Mees, and K. Judd, "Dynamics from multivariate time series", *Physica D: Nonlinear Phenomena* 121 (1–2), 75–88 (1998).

[20] S.M. Pincus and A.L. Goldberger, "Physiological time-series analysis: what does regularity quantify?", *AJP – Heart and Circulatory Physiology* 266 (4), H1643–1656 (1994).

[21] C.E. Shannon, "A mathematical theory of communication", *Bell Syst. Tech. J.* 27 (2093), 379–423, 623–656 (1948).

[22] P. Grassberger and I. Procaccia, "Estimation of the Kolmogorov entropy from a chaotic signal", *Phys. Rev.* A 28 (4), 2591–2593 (1983).

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

[24] A.A. Priplata, J.B. Niemi, J.D. Harry, L.A. Lipsitz, and J.J. Collins, "Vibrating insoles and balance control in elderly people", *The Lancet* 362 (9390), 1123–1124 (2003).

[25] M. Costa, A.A. Priplata, L.A. Lipsitz, Z. Wu, N.E. Huang, A.L. Goldberger, and C.-K. Peng, "Noise and poise: enhancement of postural complexity in the elderly with a stochastic-resonance–based therapy", *Europhysics Letters* 77 (6), 68008 (2007).

[26] A.L. Goldberger, L.A.N. Amaral, J.M. Hausdorff, P.C. Ivanov, C.-K. Peng, and H.E. Stanley, "Fractal dynamics in physiology: Alterations with disease and aging", *Proc. Natl. Acad. Sci. USA* 99 (Suppl 1), 2466–2472 (2002).

[27] J.M. Hausdorff, P.L. Purdon, C.-K. Peng, Z. Ladin, J. Y. Wei, and A. L. Goldberger, "Fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuations", *J. Applied Physiology* 80 (5), 1448–1457 (1996).

[28] H.K. Beecher, "A definition of irreversible coma: Report of the ad hoc committee of the Harvard medical school to examine the definition of brain death", *J. American Medical Association* 205 (6), 337–340 (1968).

[29] L. Li, Y. Xia, B. Jelfs, J. Cao, and D. P. Mandic, "Modelling of brain consciousness based on collaborative adaptive filters", *Int. Sym. on Neural Networks* 1, CD-ROM (2010).

[30] D.P. Mandic, M. Golz, A. Kuh, D. Obradovic, and T. Tanaka, *Signal Processing Techniques for Knowledge Extraction and Information Fusion*, Springer, Berlin, 2008.

[31] T. Gautama, D.P. Mandic, and M.M. Van Hulle, "A novel method for determining the nature of time series", *IEEE Trans. on Biomedical Engineering* 51 (5), 728–736 (2004).

[32] L. Li, Y. Xia, B. Jelfs, J. Cao, and D. P. Mandic, "Modelling of brain consciousness based on collaborative adaptive filters", *Neurocomputing* 76 (1), 36–43 (2012).