

Multivariate Multiscale Entropy Analysis

Mosabber Uddin Ahmed and Danilo P. Mandic

Abstract—Multivariate physical and biological recordings are common and their simultaneous analysis is a prerequisite for the understanding of the complexity of underlying signal generating mechanisms. Traditional entropy measures are maximized for random processes and fail to quantify inherent long-range dependencies in real world data, a key feature of complex systems. The recently introduced multiscale entropy (MSE) is a univariate method capable of detecting intrinsic correlations and has been used to measure complexity of single channel physiological signals. To generalize this method for multichannel data, we first introduce multivariate sample entropy (MSampEn) and evaluate it over multiple time scales to perform the multivariate multiscale entropy (MMSE) analysis. This makes it possible to assess structural complexity of multivariate physical or physiological systems, together with more degrees of freedom and enhanced rigor in the analysis. Simulations on both multivariate synthetic data and real world postural sway analysis support the approach.

Index Terms—Multivariate embedding, long term correlation, multivariate multiscale entropy (MMSE), multivariate sample entropy (MSampEn), multivariate system complexity.

I. INTRODUCTION

REAL world phenomena are characterized by long term dependencies in their dynamical behaviors, and a number of criteria have been developed to assess structural complexity of the underlying signal generating mechanisms. Such criteria are typically based on local predictability, irregularity, self-similarity, and synchrony [1]. Time delay embedded reconstruction is particularly popular, as it allows for the estimation of invariant quantities (in terms of smooth transformations in state space) of the original system, such as attractor dimensions, Lyapunov exponents and various entropy measures [1].

Traditional entropy measures, such as Shannon entropy, Kolmogorov–Sinai (KS) entropy, approximate entropy (ApEn) [2], and sample entropy (SampEn) [3], are maximized for completely random processes, and are used to quantify the regularity of univariate time series on a single scale, by e.g., evaluating repetitive patterns [4]. As a result, using entropy to measure complexity of physiological data, which exhibit high degree of structural richness, yields lower entropy than for their randomized surrogates, formed by shuffling the original data samples and thus destroying any structure present. This is

counterintuitive for a measure of complexity, and the greater entropy of the uncorrelated surrogate series also highlights a lack of a straightforward correspondence between regularity and complexity. Neither completely predictable (e.g., periodic) nor completely unpredictable (e.g., uncorrelated random) signals are truly complex, since at a global level none is structurally rich. Instead, time series observed from dynamical physical and physiological systems generally exhibit long-range correlations at multiple spatial and temporal scales.

The multiscale entropy (MSE) method proposed by Costa *et al.* [4] explicitly quantifies the amount of structure (correlation) in real world time series, that is, the underlying system complexity. The method evaluates sample entropy of coarse grained (averaged over increasing segment lengths) univariate time series; the underlying idea is that coarse graining defines temporal scales, hence a system without structure would exhibit a rapid decrease in entropy with an increase in time scale. The existing MSE algorithm has been proven to be able to distinguish between physiological time series with different degrees of complexity and its extensions have included more rigorous definitions of time scales [5]. Applications include the analysis of pathologic heartbeat conditions like erratic cardiac arrhythmia and congestive heart failure [4], electroencephalogram (EEG) changes in patients with Alzheimer’s disease [6], postural sway dynamics analysis [7], and complex dynamics of human red blood cells [8]. All the above results strongly support the general “loss of complexity” behavior when a living system undergoes change from its normal “unconstrained” state (healthy) to that under stress, e.g., due to ageing and disease [9].

Recent developments in sensor technology have enabled routine recording of multivariate time series from both physical and biological systems, yet when assessing their complexity the standard MSE analysis can only consider every data channel separately. This is only appropriate if the components of a multivariate signal are statistically independent or at the very least uncorrelated (which is usually not the case). For instance, the embedding theorem [10] establishes that the dynamics of a multivariate system can be reconstructed from sufficiently long time delay embedded vectors of a single time series (seen as a one-dimensional projection of the multivariate system trajectory), however, in practice this is not reliable for systems exhibiting more than two degrees of freedom. Indeed, based on measurements of the z -coordinate of the Lorenz system we cannot reconstruct its dynamics, as embedding based on the z -coordinate does not resolve the x - y symmetry [11].

In this work, we illustrate substantial advantages in simultaneously analyzing dynamical complexity of several variables observed from the same system, especially if there is a large degree of uncertainty or coupling underlying the measured system dynamics. To this end, we first introduce multivariate sample

Manuscript received November 03, 2011; revised December 10, 2011; accepted December 13, 2011. Date of publication December 20, 2011; date of current version January 12, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Z. Jane Wang.

The authors are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: mosabber.ahmed@imperial.ac.uk; d.mandic@imperial.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2011.2180713

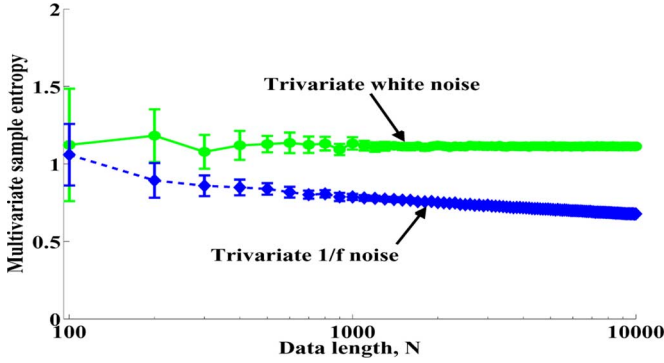


Fig. 1. Multivariate sample entropy as a function of data length N , for $r = 0.20$ and $m_k = 2$ in each data channel. Shown are the mean values for 30 simulated trivariate time series containing white and $1/f$ noise.

entropy, and evaluate its evolution at different temporal scales. The so introduced multivariate multiscale entropy (MMSE) is shown to operate on any number of data channels, and to provide a robust relative complexity measure for multivariate data. The approach is supported by simulations on both synthetic and real world multivariate processes.

II. MULTIVARIATE MULTISCALE ENTROPY

To introduce the multivariate multiscale entropy (MMSE) analysis, we propose the following steps.

- 1) Define temporal scales by averaging a p -variate time series $\{x_{k,i}\}_{i=1}^N$, $k = 1, 2, \dots, p$ over non-overlapping time segments of increasing length (coarse graining), where N is the number of samples in every channel. This way, for scale ϵ , a coarse grained multivariate time series is obtained as $y_{k,j}^\epsilon = \frac{1}{\epsilon} \sum_{i=(j-1)\epsilon+1}^{j\epsilon} x_{k,i}$, where $1 \leq j \leq \frac{N}{\epsilon}$ and the channel index $k = 1, \dots, p$.
- 2) Evaluate multivariate sample entropy, MSampEn, for each coarse-grained multivariate $y_{k,j}^\epsilon$, and plot MSampEn as a function of the scale factor ϵ .

However, methods to calculate multivariate entropy are still lacking. To this end, we next propose a multivariate sample entropy measure, and use it to introduce the MMSE method.

A. The MSampEn Method

To calculate MSampEn, recall from multivariate embedding theory [11], that for a p -variate time series $\{x_{k,i}\}_{i=1}^N$, $k = 1, 2, \dots, p$, observed through p measurement functions $h_k(y_i)$, the multivariate embedded reconstruction is based on a composite delay vector

$$X_m(i) = [x_{1,i}, x_{1,i+\tau_1}, \dots, x_{1,i+(m_1-1)\tau_1}, x_{2,i}, x_{2,i+\tau_2}, \dots, x_{2,i+(m_2-1)\tau_2}, \dots, x_{p,i}, x_{p,i+\tau_p}, \dots, x_{p,i+(m_p-1)\tau_p}] \quad (1)$$

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

For a p -variate time series $\{x_{k,i}\}_{i=1}^N$, $k = 1, 2, \dots, p$, the MSampEn method is introduced in Algorithm 1, and represents a natural extension of standard univariate sample entropy [3].

Algorithm 1 The Multivariate Sample Entropy (MSampEn)

- 1: Form $(N - n)$ composite delay vectors $X_m(i) \in \mathbb{R}^m$, where $i = 1, 2, \dots, N - n$ and $n = \max\{\mathbf{M}\} \times \max\{\boldsymbol{\tau}\}$.
- 2: Define the distance between any two vectors $X_m(i)$ and $X_m(j)$ as the maximum norm, that is, $d[X_m(i), X_m(j)] = \max_{l=1, \dots, m} \{|x(i+l-1) - x(j+l-1)|\}$.
- 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, $B_i^m(r) = \frac{1}{N-n-1} P_i$, and define $B^m(r) = \frac{1}{N-n} \sum_{i=1}^{N-n} B_i^m(r)$.
- 4: Extend the dimensionality of the multivariate delay vector in (1) from m to $(m + 1)$. This can be performed in p different ways, as from a space with the embedding vector $\mathbf{M} = [m_1, m_2, \dots, m_k, \dots, m_p]$ the system can evolve to any space for which the embedding vector is $[m_1, m_2, \dots, m_k + 1, \dots, m_p]$ ($k = 1, 2, \dots, p$). Thus, a total of $p \times (N - n)$ 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 .
- 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, $B_i^{m+1}(r) = \frac{1}{p(N-n)-1} Q_i$, and define $B^{m+1}(r) = \frac{1}{p(N-n)} \sum_{i=1}^{p(N-n)} B_i^{m+1}(r)$.
- 6: Finally, for a tolerance level r , estimate MSampEn as

$$\text{MSampEn}(\mathbf{M}, \boldsymbol{\tau}, r, N) = -\ln \left[\frac{B^{m+1}(r)}{B^m(r)} \right]. \quad (2)$$

B. Effect of Data Length on MSampEn

It has been suggested in [2] 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 trivariate white as well as $1/f$ noise as a function of sample size N , where for each channel the embedding dimension was $m_k = 2$ and the threshold $r = 0.2$. Fig. 1 shows that for both the white and $1/f$ trivariate noise, MSampEn estimates were consistent for data length $N \geq 300$, illustrating suitability of MSampEn for the analysis of real world data. This way, the standard deviation of multivariate sample entropy estimates (error bars) related to the length of the simulated series is likely to be smaller than the deviations related to experimental errors as well as inter- and intra-subject variability for most practical applications.

Physically, for the standard univariate sample entropy, the increase in sample entropy values with an increase in the 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. For instance, for the case in Fig. 1 and $m = 6$, we are not taking all the six points from the same univariate series, instead we are taking two points (as $m_k = 2$) from each of the three channels of the trivariate series in hand.

Besides, in MSampEn calculation we do not give preference to any particular m_k over m_j when increasing the embedding dimension from m_k to $m + 1$. Instead, we create embedded vectors in all the p subspaces and compare them within and across these subspaces to find the r -neighbors. This also effectively increases the total number of available delay vectors p times, and makes the proposed MSampEn robust to the variation in both the parameter m and data length N .

C. Selection of Parameter Values

There are several parameters that need to be chosen in MSampEn calculation, each introducing their own constraints. For instance, for multimodal data coming from heterogeneous data channels, the individual channels are likely to exhibit different embedding parameters m_k and τ_k . There are several methods [12] for determining the optimal embedding parameters m_k and τ_k simultaneously, for a single channel. Future developments should test for all the m_k and τ_k parameters from the different data channels simultaneously, but these are not yet available. Also, standard sample entropy introduces a practical constraint that $m_k < 5$, which is inherited in MSampEn. This did not influence the results in this work and is sufficient for the majority of real world applications. In MSampEn 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 15%) of $tr(S)$, which is similar to taking $r = 0.15$ in each channel.

D. Multivariate Complexity Analysis

The multivariate MSE (MMSE) method assesses relative complexity of normalized multichannel temporal data by plotting multivariate sample entropy as a function of the scale factor. 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 structurally complex. For instance, the standard univariate MSE analysis in [4] showed that for random white noise (uncorrelated) the sample entropy values monotonically decrease with scale, whereas for $1/f$ noise (long-range correlated) the sample entropy remains constant over multiple scales. This has a physical justification, as by design $1/f$ noise is structurally more complex than uncorrelated random noise.

To illustrate the corresponding behavior for the proposed multivariate approach, we considered a 4-variate time series, where originally all the data channels were fed with mutually independent white noise. The number of noise channels was then gradually decreased (from 4 to 0) with a simultaneous increase in the number of data channels that represent independent $1/f$ noise (from 0 to 4). Fig. 2 shows the MMSE curves for all the possible cases: notice that, as desired, when the number

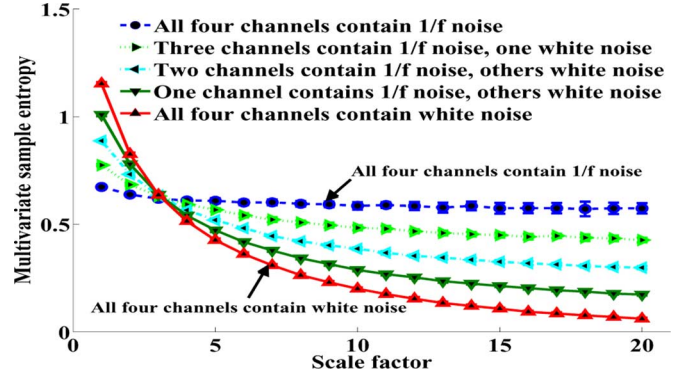


Fig. 2. MMSE analysis for 4-channel data containing independent white and $1/f$ noise. Each data channel had 10,000 data points, and the plots represent an average of 20 independent realizations and error bars the standard deviation (SD).

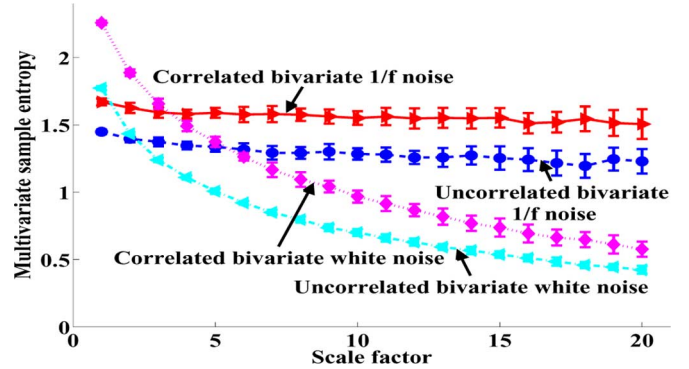


Fig. 3. Multivariate multiscale entropy (MMSE) analysis for bivariate white and $1/f$ noise. Each data channel had 10,000 data points, and the plots represent an average of 20 independent realizations and error bars the standard deviation (SD).

of channels representing $1/f$ noise increased, the MSampEn at higher scales also increased, and when all the four data channels contained $1/f$ noise, the complexity at larger scales was the highest (cf. smallest at scale 1).

Cross-channel correlations are a key aspect in multivariate analysis and in the next set of simulations, we analyzed the ability of MMSE to capture cross-channel properties. Fig. 3 shows that, as desired, the proposed multivariate extension of MSE caters for both within- and cross-channel correlations: the complexity of the correlated bivariate $1/f$ noise at large scales was the highest, followed by the uncorrelated $1/f$ noise, and correlated and uncorrelated white noise.

III. EXPERIMENTAL RESULTS

Multivariate complexity analysis of real world postural sway dynamics (time series of the center of pressure (COP) displacement) is next performed for young and elderly subjects during quiet standing, and has been compared with the existing univariate complexity analysis [7]. COP displacement was recorded simultaneously for the mediolateral (side-to-side) and anteroposterior (front-to-back) direction [13] for 15 healthy young and 12 healthy elderly volunteers at 60 Hz.

Since the postural sway time series exhibits high frequency fluctuations superimposed on low frequency trends, the data was first detrended using the multivariate empirical mode decomposition technique [14]. The values of the parameters used to calculate MSampEn were $m_k = 2$, $\tau_k = 1$, and

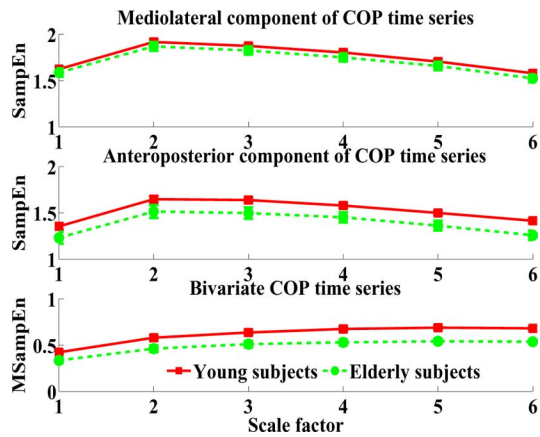


Fig. 4. Multiscale entropy analysis of COP time series for young (red—solid line) and elderly (green—dotted line) subjects. Top: Univariate MSE for the mediolateral component. Middle: Univariate MSE for the anteroposterior component. Bottom: Bivariate MMSE analysis. The plots represent mean values of SampEn/MSampEn for all subjects across all the trials and error bars represent standard error.

$r = 0.15 \times (\text{standard deviation of the normalized time series})$ for each data channel; these parameters were chosen on the basis of previous studies indicating good statistical reproducibility for the univariate SampEn [3], [7]. Since the original time series had 1.8×10^3 data points, the highest achievable scale factor ($\epsilon = 6$) had 300 data points; this was sufficient for accurate analysis, as shown in Fig. 1.

The top and middle panel in Fig. 4 show the results obtained by the univariate MSE performed for the mediolateral (ML) and anteroposterior (AP) component of the COP time series, respectively. The bottom panel shows that when both the ML and AP component were considered within the multivariate approach, the proposed MMSE was able to discriminate between young and elderly subjects more efficiently, as indicated by the better separation of the MMSE curves and the complexity index (area under the MSE curve) being higher for healthy young subjects than for their elderly counterparts¹. Table I shows the statistical significance of the results obtained by the one-tailed t-test with unequal variances. It is evident that the difference in complexity between the young and elderly subjects was statistically significant in the AP direction for a 5% significance level and not significant in the ML direction. If, for rigor, we take a 1% significance level, then none of the univariate analyses gave statistically significant results. On the contrary, in the bivariate case we observed significant differences, even for the significance level of $p = 6.65 \times 10^{-4}$. This illustrates significant advantages of using 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 [9]. The Matlab code for the MMSE analysis can be downloaded from [15].

IV. CONCLUSION

This work has generalized the recently introduced univariate multiscale entropy (MSE) method to the multivariate case, to

¹Due to ageing and the associated constraints, the complexity of postural sway for the elderly is lower than for the young.

TABLE I
COMPLEXITY INDICES AND P VALUES OBTAINED USING ONE-TAILED STUDENT'S T-TEST WITH UNEQUAL VARIANCE

	AP(mean \pm SD) univariate	ML(mean \pm SD) univariate	AP & ML bivariate
Young subjects	7.75 \pm 1.67	8.90 \pm 1.52	3.13 \pm 1.53
Elderly subjects	7.07 \pm 2.23	8.66 \pm 1.62	2.48 \pm 1.20
p value	0.0181	0.1598	6.65×10^{-4}

provide complexity analysis of real world biological and physical systems which are typically of multivariate, correlated and noisy natures. The proposed multivariate multiscale entropy (MMSE) method has been shown to be naturally suited to reveal long range within- and cross-channel correlations present in multichannel data. The MMSE method has been validated on both illustrative benchmark signals and on a simultaneous analysis of ML and AP component of postural sway dynamics from the young and elderly subjects.

REFERENCES

- [1] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
- [2] S. M. Pincus and A. L. Goldberger, "Physiological time-series analysis: What does regularity quantify?," *AJP—Heart Circ. Physiol.*, vol. 266, no. 4, pp. H1643–1656, Apr. 1994.
- [3] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *AJP—Heart Circ. Physiol.*, vol. 278, no. 6, pp. H2039–2049, 2000.
- [4] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Phys. Rev. Lett.*, vol. 89, no. 6, pp. 068102–068102, 2002.
- [5] J. Valencia, A. Porta, M. Vallverdú, F. Claria, R. Baranowski, E. Orłowska-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. Biomed. Eng.*, vol. 56, no. 9, pp. 2202–2213, 2009.
- [6] R. Hornero, D. Absolo, J. Escudero, and C. Gomez, "Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease," *Philos. Trans. R. Soc. A*, vol. 367, no. 1887, pp. 317–336, 2009.
- [7] 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," *Europhys. Lett.*, vol. 77, no. 6, pp. 68008–68008, 2007.
- [8] 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*, vol. 78, no. 2, pp. 020901–020901, 2008.
- [9] 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. Nat. Acad. Sci. USA*, vol. 99, no. Suppl 1, pp. 2466–2472, 2002.
- [10] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick 1980*, ser. Lecture Notes in Mathematics, D. Rand and L.-S. Young, Eds. : Springer-Verlag, vol. 898, pp. 366–381.
- [11] L. Cao, A. Mees, and K. Judd, "Dynamics from multivariate time series," *Phys. D: Nonlinear Phenomena*, vol. 121, no. 1–2, pp. 75–88, 1998.
- [12] T. Gautama, D. P. Mandic, and M. M. Van Hulle, "A differential entropy based method for determining the optimal embedding parameters of a signal," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP '03)*, 2003, vol. 6, pp. 29–32.
- [13] A. A. Priplata, J. B. Niemi, J. D. Harry, L. A. Lipsitz, and J. J. Collins, "Vibrating insoles and balance control in elderly people," *Lancet*, vol. 362, no. 9390, pp. 1123–1124, 2009.
- [14] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proc. R. Soc. A: Math. Phys. Eng. Sci.*, vol. 466, no. 2117, pp. 1291–1302, 2010.
- [15] [Online]. Available: http://www.commsp.ee.ic.ac.uk/~mandic/research/Complexity_Stuff.htm