# Refined Composite Multiscale Permutation Entropy to Overcome Multiscale Permutation Entropy Length Dependence

Anne Humeau-Heurtier, Chiu-Wen Wu, and Shuen-De Wu

***Abstract*—Multiscale permutation entropy (MPE) has recently been proposed to evaluate complexity of time series. MPE has numerous advantages over other multiscale complexity measures, such as its simplicity, robustness to noise and its low computational cost. However, MPE may loose statistical reliability as the scale factor increases, because the coarse-graining procedure used in the MPE algorithm reduces the length of the time series as the scale factor grows. To overcome this drawback, we introduce the refined composite MPE (RCMPE). Through applications on both synthetic and real data, we show that RCMPE is much less dependent on the signal length than MPE. In this sense, RCMPE is more reliable than MPE. RCMPE could therefore replace MPE for short times series or at large scale factors.**

***Index Terms*—Complexity, entropy, fractal, multiscale entropy, nonlinear dynamics, permutation entropy.**

## I. Introduction

SEVERAL entropic measures have been proposed to quantify the regularity of systems: approximate entropy [1], sample entropy [2], and permutation entropy (PE) [3], to cite some of the most well-known. Over approximate entropy, sample entropy has the advantage of being less dependent on the signal length. Moreover, over other complexity measures, PE has the advantage of being conceptually simple, robust to observational and dynamical noise, and computationally fast [3], [4]. Furthermore, it is invariant with respect to nonlinear monotonous transformations and can be applied to any type of time series (regular, chaotic, noisy, or reality based). We will therefore concentrate on PE in what follows. PE relies on the comparison of neighboring values to estimate the complexity of time series.

Time series from physiological systems or mechanical systems contain multiple temporal scale structures. Therefore, traditional algorithms–as PE–that consider only a single scale, have limited performances in assessing these systems complexity. This is why multiscale permutation entropy (MPE) has been introduced [5]. MPE calculate PE over multiple scales, through a coarse-graining procedure. For scale $\tau$, the coarse-graining procedure averages the samples of the time series inside consecutive but non overlapping windows of length $\tau$. For MPE, each coarse-grained time series is then processed with the PE algorithm [5]. The coarse-graining procedure has the advantage of leading to multiple scale time series from the original signal. However, it has the drawback of leading to time series with reduced length [6]: at a scale factor $\tau$, the coarse-grained time series has a length equal to the one of the original time series divided by $\tau$. Therefore, the larger the scale factor $\tau$, the shorter the coarse-grained time series. Unfortunately, PE requires a minimum number of samples to lead to measures with statistical reliability. Thus, to achieve reliable statistics and proper discrimination between stochastic and deterministic dynamics, PE requires that the length $N$ of the time series satisfies $N \gg m!$, where $m$ is the embedding dimension (see below) [7]. For practical purposes, Bandt and Pompe suggested to work with $3 \leq m \leq 7$ [3].

All this leads to drawbacks for MPE, at large scale factors (short coarse-grained time series). In order to overcome these drawbacks, we herein introduce two new algorithms: the composite MPE (CMPE) and the refined CMPE (RCMPE), as described thereafter.

In what follows, MPE, CMPE, and RCMPE are introduced. Each algorithm is then applied on synthetic and real data. The results are finally compared and discussed.

## II. Composite Multiscale Permutation Entropy and Refined Composite Multiscale Permutation Entropy

### A. Multiscale Permutation Entropy

The MPE algorithm relies on a two-step procedure: (i) a coarse-graining procedure that leads to multiple scale time series from the original signal; (ii) application of the PE algorithm on each of the coarse-grained time series [5].

For a signal of length $N\{x_1, \ldots, x_i, \ldots, x_N\}$, the coarse-grained time series $\mathbf{y}^{(\tau)}$ is computed as

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \qquad 1 \leq j \leq N/\tau, \tag{1}$$

where $\tau$ is the scale factor.

Manuscript received July 23, 2015; revised September 01, 2015; accepted September 23, 2015. Date of publication September 28, 2015; date of current version October 02, 2015. This work was supported in part by the Ministry of Science and Technology, Taiwan, under Grant NSC 102-2218-E-003 -001 -MY2. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Manuel Martinez-Ramon.

A. Humeau-Heurtier is with Université d'Angers, LARIS - Laboratoire Angevin de Recherche en Ingénierie des Systèmes, 49000 Angers, France (e-mail: anne.humeau@univ-angers.fr).

C. W. Wu and S. D. Wu are with the Department of Mechatronic Engineering, National Taiwan Normal University, Taipei 10610, Taiwan (e-mail: im113qr@gmail.com; sdwu@ntnu.edu.tw).

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.2015.2482603

In the general case of a time series $\{y_1, \ldots, y_i, \ldots, y_N\}$ the PE value is computed as follows: the $m$-dimensional delay embedding vector is computed as

$$\mathbf{y}_i^m = \left[y_i, y_{i+L}, \ldots, y_{i+(m-1)L}\right], \quad (2)$$

where $L$ is the delay time, and $i = 1, 2, \ldots, N-(m-1)L$ where $N$ is the length of the coarse-grained time series. Then, for each $i$, the $\mathbf{y}_i^m$ vector is arranged in an increasing order. There are $m!$ possible patterns $\pi$ (also called permutations) for an $m$-tuple vector. Let $f(\pi)$ denotes the frequency of the permutation $\pi$ in the time series, and $p(\pi)$ its relative frequency. In the general case of a signal $\mathbf{y}$, and for a dimension $m$ and a delay time $L$, the PE value of the time series $\mathbf{y}$ is defined as [3]

$$PE(\mathbf{y}, m, L) = -\sum_{i=1}^{m!} p(\pi)\ln(p(\pi)). \quad (3)$$

Then, the normalized permutation entropy $PE_{\mathrm{norm}}(\mathbf{y}, m, L)$ can be defined as

$$PE_{\mathrm{norm}}(\mathbf{y}, m, L) = \frac{PE(\mathbf{y}, m, L)}{\ln(m!)}. \quad (4)$$

Therefore, $0 \leq PE_{\mathrm{norm}}(\mathbf{y}, m, L) \leq 1$. When $PE_{\mathrm{norm}}$ is equal to 1, all permutations have equal probability. When $PE_{\mathrm{norm}}$ is zero, the time series is very regular. The smaller the value of $PE_{\mathrm{norm}}$, the more regular the time series.

### *B. Composite Multiscale Permutation Entropy and Refined Composite Multiscale Permutation Entropy*

As mentioned above, in order to obtain PE values with statistical reliability, the length $N$ of the time series should satisfy $N \gg m!$, where $m$ is the embedding dimension. In order to overcome this constraint, we first introduce the CMPE algorithm. For this purpose, and for a discrete time series $\mathbf{x} = \{x_1, \ldots, x_i, \ldots, x_N\}$, the $k$th coarse-grained time series for a scale factor $\tau$ is defined as $\mathbf{y}_k^{(\tau)} = \{y_{k,1}^{(\tau)}, y_{k,2}^{(\tau)} \ldots y_{k,p}^{(\tau)}\}$ where [8]

$$y_{k,j}^{(\tau)} = \frac{1}{\tau}\sum_{i=(j-1)\tau+k}^{j\tau+k-1} x_i, \qquad 1 \leq j \leq \frac{N}{\tau}, 1 \leq k \leq \tau. \quad (5)$$

Then, we define CMPE as

$$CMPE(\mathbf{x}, \tau, m, L) = \frac{1}{\tau}\sum_{k=1}^{\tau} PE(\mathbf{y}_k^{(\tau)}, m, L), \quad (6)$$

where $PE(\mathbf{y}_k^{(\tau)}, m, L)$ corresponds to the PE for the time series $\mathbf{y}_k^{(\tau)}$, calculated with the embedding dimension $m$ and delay time $L$. Therefore, at a scale factor $\tau$, the PE values of all coarse-grained time series are taken into account and the CMPE value is computed as the mean of the $\tau$ entropy values. By opposition, in the MPE algorithm, only the first coarse-grained time series is considered.

As will be shown thereafter, CMPE leads to more accurate values than MPE (lower variance of estimated entropy values) but does not resolve the decrease of PE as the length of the time series decreases (when this length does not satisfy $N \gg m!$). This is why we also introduce the RCMPE algorithm.

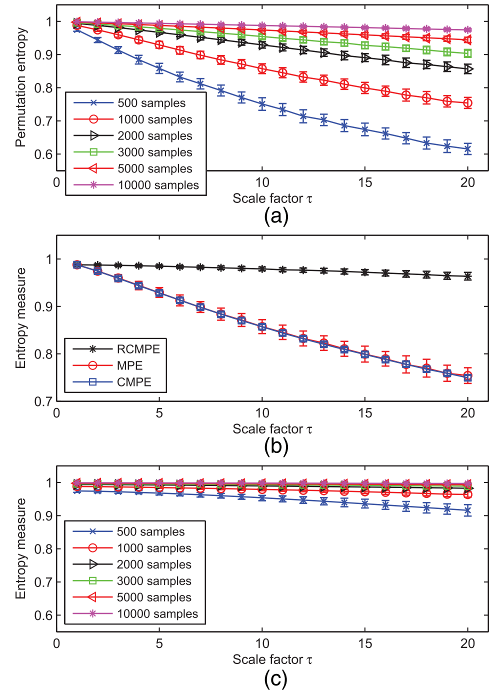


Fig. 1. (a) Multiscale permutation entropy (MPE) values for 200 realizations of Gaussian white noise; results for different lengths of data are shown; (b) MPE, composite MPE (CMPE), and refined CMPE (RCMPE) for 200 realizations of Gaussian white noise (data with 1000 samples); (c) RCMPE values for 200 realizations of Gaussian white noise; results for different lengths of data are shown. For (a), (b), and (c) the embedding dimension $m$ was set to 5 and the delay time $L$ was equal to 1.

We define RCMPE as

$$RCMPE(\mathbf{x}, \tau, m, L) = -\sum_{1}^{m!} \overline{p^{(\tau)}}(\pi)\ln(\overline{p^{(\tau)}}(\pi)), \quad (7)$$

where $\overline{p^{(\tau)}}(\pi) = \frac{1}{\tau}\sum_{k=1}^{\tau} p_k^{(\tau)}(\pi)$ with $p_k^{(\tau)}(\pi)$ the relative frequency of the permutation $\pi$ in the time series $\mathbf{y}_k^{(\tau)}$.

## III. Experimental Results

For the study of CMPE and RCMPE behaviors, both synthetic and real signals are herein processed: white noise, a signal from the well-known Mackey-Glass equation, a biomedical signal, and vibration data.

When processing white noise, the PE should be kept at unity, for all scales, because the coarse-grained time series of white noise is still white noise. However, it has been previously shown that, for white noise, PE values decrease with scales, when $N \gg m!$ is not satisfied (see, e.g., [5]). This is illustrated in Fig. 1(a) where we report the PE values computed from two hundred white noise realizations. Several lengths for the white noises are studied ($N$ = 500, 1000, 2000, 3000, 5000, and 10000 samples). Moreover, it has been reported that the standard deviation values of PE increase with the decrease of the length of the time series [5]. This is illustrated in Fig. 1(a) and Table I where the PE value of white noise data at a scale factor $\tau = 20$ is calculated using several data lengths.

Regarding CMPE on white noise, Fig. 1(b) shows that CMPE leads to the same mean values of entropy as MPE, for all the scales studied. However, CMPE has reduced standard deviation values compared to MPE. In this sense, CMPE leads to more accurate values than MPE. Nevertheless, as MPE, CMPE shows a

TABLE I
STANDARD DEVIATIONS OF MPE, CMPE, AND RCMPE FOR 200 REALIZATIONS OF WHITE NOISE AT SCALE FACTOR $\tau = 20$, FOR AN EMBEDDING DIMENSION $m = 5$ AND A DELAY TIME $L = 1$

| Algorithm | Data length (in samples) | | | | | |
|---|---|---|---|---|---|---|
| | 500 | 1000 | 2000 | 3000 | 5000 | 10000 |
| MPE | 0.0168 | 0.0165 | 0.0142 | 0.0119 | 0.0080 | 0.0040 |
| CMPE | 0.0052 | 0.0053 | 0.0043 | 0.0036 | 0.0023 | 0.0011 |
| RCMPE | 0.0172 | 0.0080 | 0.0037 | 0.0024 | 0.0013 | 0.0007 |

decreasing trend as the length of the time series decreases (when this length does not satisfy $N \gg m!$). In this work, we therefore also introduced RCMPE. Figs. 1(b) and (c) show that RCMPE leads to almost constant values for entropy of white noise, for all the scales studied, and for all the time series lengths studied. Thus, for a scale factor $\tau = 20$ and white noises with a length of 500 samples, the mean for the RCMPE value is 0.9159 (instead of a theoretical value of 1), whereas it is of 0.6156 for MPE. From Figs. 1(a) and (c), we also observe that for time series sufficiently long ($N \geq 10000$ samples), the mean value of the entropies given by MPE and RCMPE are nearly equal. However, there are differences for short time series. Therefore, RCMPE over-performs MPE and CMPE in terms of independence of data length. As a result, RCMPE is more reliable than MPE and CMPE.

We also processed a signal computed from the well-known Mackey-Glass equation [9]

$$\frac{dx}{dt} = -x + \frac{ax(t-\tau_s)}{1 + x^c(t-\tau_s)}, \tag{8}$$

with $t$ a dimensionless time, $a$ the feedback strength, $c$ the degree of nonlinearity, and $\tau_s$ the time-delay feedback. In our work, we choose the typical values $a = 2$, $c = 10$, and $\tau_s = 2$ for which the system is in a chaotic mode. From our results, we observe that the PE values obtained with the corresponding signal show an increasing trend with scale factors $\tau$, see Fig. 2(a). However, this trend is different when the condition $N \gg m!$ is not satisfied (see Fig. 2(a)): at the largest scale factors (shortest time series), the curves show a decreasing trend with scale factors $\tau$.

When processing the signal given by the Mackey-Glass equation with CMPE, we arrive at the same conclusion as for white noise realizations (see Fig. 2(b)): CMPE gives values of entropy close to the ones given by MPE, for all the scales studied. Furthermore, our work also reveals that RCMPE is much less length-dependent than MPE or CMPE, see Fig. 2(c). Thus, for a given length of a (short) time series, the values for the entropy measures given by RCMPE are closer to the values reached with longer time series than the ones given by MPE (or CMPE).

In addition to the processing of synthetic data, we also processed different kinds of *real data*: a biomedical signal and vibration data. For the biomedical field, we used a laser Doppler flowmetry (LDF) signal recorded on the forearm of a healthy subject, at rest. The sampling frequency was 18 Hz. The LDF technique reflects the microvascular perfusion and has the advantage of having good temporal resolution [10]. The computation of MPE, CMPE and RCMPE for the LDF signal leads to the same conclusion as for the synthetic data: PE values are length-dependent when $N \gg m!$ is not satisfied, see Fig. 3(a). Moreover, CMPE leads to the same mean values of entropy as MPE, for all the scales studied. Furthermore, as for MPE, CMPE is length-dependent when the condition $N \gg m!$ is not satisfied (see Fig. 3(b)). By opposition, RCMPE is much less length-dependent than MPE or CMPE, see Fig. 3(c). We also observe in

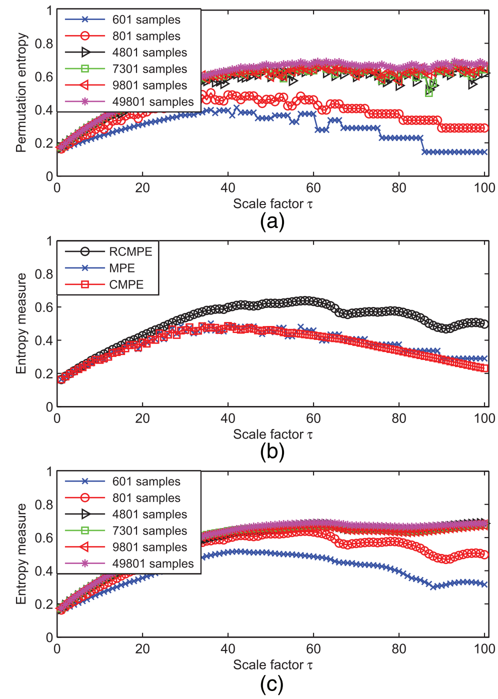


Fig. 2. (a) Multiscale permutation entropy (MPE) values for a signal computed from the Mackey-Glass equation; results for different lengths of data are shown; (b) MPE, composite MPE (CMPE), and refined CMPE (RCMPE) for a signal computed from the Mackey-Glass equation (data with 801 samples); (c) RCMPE values for a signal computed from the Mackey-Glass equation; results for different lengths of data are shown. For (a), (b), and (c) the embedding dimension $m$ was set to 5 and the delay time $L$ was equal to 1.

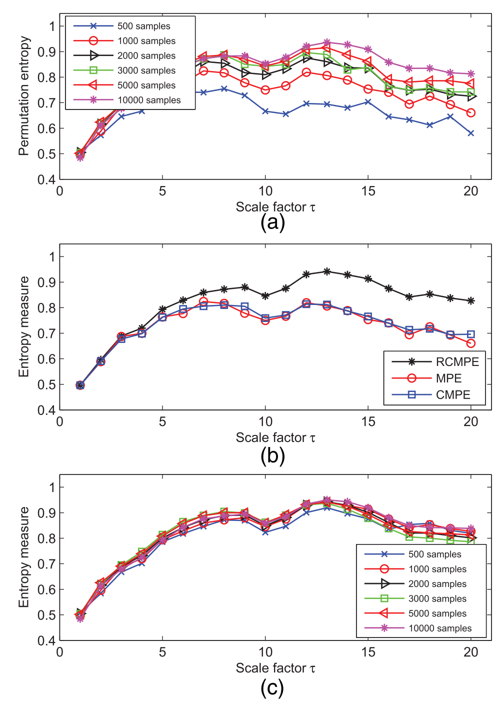


Fig. 3. (a) Multiscale permutation entropy (MPE) values for a LDF signal truncated at different lengths; (b) MPE, composite MPE (CMPE), and refined CMPE (RCMPE) for a LDF signal of 1000 samples; (c) RCMPE values for a LDF signal truncated at different lengths. For (a), (b), and (c) the embedding dimension $m$ was set to 5 and the delay time $L$ was equal to 1. The sampling frequency for the LDF signal was 18 Hz.

Fig. 3(c) that RCMPE shows a local minimum for scale factor $\tau$ around 17 (i.e. $T\tau = 1/18 \times 17$ s $= 0.94$ s). This could be due to the cardiac activity: the regularity of the heart beats could lead to lower values for entropy.

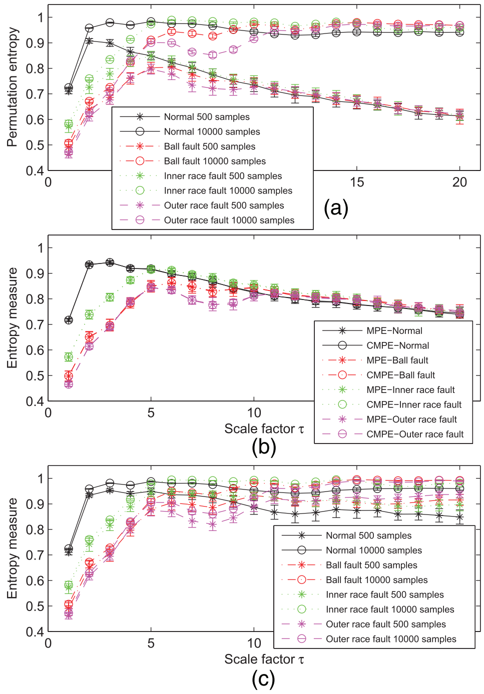


Fig. 4. (a) Multiscale permutation entropy (MPE) values for vibration data truncated at different lengths (see text for the details); (b) MPE and composite MPE (CMPE) for vibration data of 1000 samples; (c) RCMPE values for vibration data truncated at different lengths. For (a), (b), and (c) the embedding dimension $m$ was set to 5 and the delay time $L$ was equal to 1. The sampling frequency for the data was 48 kHz.

Experimental analyses on bearing faults have also been conducted in our work. The bearing fault data were obtained from the Case Western Reserve University (CWRU) Bearing Data Center [11]. The time-domain vibration signals of bearings were collected from the normal case, the ball fault case, the inner race fault case, and the outer race fault case located at the 3 o'clock position. The sampling frequency of the vibration signal was 48 kHz, and the rotation speed of the motor was 1730 rpm. The fault diameter used in the experiments was 7 mils. Fig. 4(a) shows that PE values are length-dependent, when $N \gg m!$ is not satisfied. Moreover, as before, CMPE leads to the same mean values of entropy as MPE, for all the scales studied, and therefore CMPE is also length-dependent when this length does not satisfy $N \gg m!$ (see Fig. 4(b)). However, CMPE has reduced standard deviation values compared to MPE, see Fig. 4(b). Furthermore, RCMPE is much less length-dependent than MPE or CMPE, see Fig. 4(c).

We also compared the computational cost of MPE, CMPE and RCMPE. We show that CMPE and RCMPE necessitate longer computational time than MPE (see Table II). Therefore, in order to balance the trade-off between accuracy and computational cost, an hybrid algorithm could be proposed: for coarse-grained time series satisfying $N \gg m!$, MPE could be used, whereas for shorter data length (large scales), RCMPE would be more appropriated [12].

From the best of our knowledge, only one paper proposes a modification of the original MPE algorithm [13]. The latter work introduces the weighted MPE (WMPE). WMPE is a modified version of MPE as WMPE incorporates amplitude information, which MPE does not do. WMPE and MPE have therefore different definition. By opposition, our work proposes a new algorithm (RCMPE) that improves MPE and that relies on the same background as MPE.

TABLE II
EXECUTION TIME (IN S) NEEDED TO ANALYZE A WHITE NOISE REALIZATION FROM SCALE FACTOR $\tau = 1$ TO 20, FOR THE MPE, CMPE, AND RCMPE ALGORITHMS. ALL THE SIMULATIONS HAVE BEEN RUN ON A PC WITH AN INTEL(R) CORE(TM) I7-4800MQ CPU AT 2.70 GHZ WITH THE MATLAB (R2012A) PLATFORM, WITH AN EMBEDDING DIMENSION $m = 5$ AND A DELAY TIME $L = 1$

| Algorithm | Data length (in samples) | | | | | |
|---|---|---|---|---|---|---|
| | 500 | 1000 | 2000 | 3000 | 5000 | 10000 |
| MPE | 0.037 | 0.065 | 0.120 | 0.174 | 0.284 | 0.463 |
| CMPE | 0.236 | 0.393 | 0.703 | 1.026 | 1.658 | 2.549 |
| RCMPE | 0.166 | 0.315 | 0.609 | 0.908 | 1.520 | 2.310 |

## IV. CONCLUSION

The PE algorithm has the advantage of being conceptually simple, computationally efficient and artifact resistant. However, to obtain statistical reliability, the time series length $N$ should satisfy the condition $N \gg m!$, where $m$ is the embedding dimension. For multiscale studies, we herein proposed the RCMPE algorithm to overcome this drawback. Based on a modified coarse-graining procedure, compared to MPE, and to a new PE computation, we show that RCMPE is less dependent on data length than MPE. In this sense, RCMPE over-performs MPE. Now, other works based on RCMPE could be conducted, such as generalization to multivariate data [14]. Moreover, it could also be interesting to analyze how PE behaves as a function of time delay [15].

## REFERENCES

[1] S. M. Pincus, "Approximate entropy as a measure of system complexity," in *Proc. Nat. Acad. Sci.*, USA, 1991, vol. 88, pp. 2297–2301.
[2] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 278, pp. H2039–H2049, 2000.
[3] C. Bandt and B. Pompe, "Permutation entropy: A natural complexity measure for time series," *Phys. Rev. Lett.*, vol. 88, 2002, article id. 174102.
[4] C. Bandt, G. Keller, and B. Pompe, "Entropy of interval maps via permutations," *Nonlinearity*, vol. 15, pp. 1595–1602, 2002.
[5] W. Aziz and M. Arif, "Multiscale permutation entropy of physiological time series," in *9th Int. Multitopic Conf., IEEE INMIC 2005*, 2005, pp. 1–6.
[6] A. Humeau-Heurtier, "The multiscale entropy algorithm and its variants: A Review," *Entropy*, vol. 17, pp. 3110–3123, 2015.
[7] O. A. Rosso, H. A. Larrondo, M. T. Martin, A. Plastino, and M. A. Fuentes, "Distinguishing noise from chaos," *Phys. Rev. Lett.*, vol. 99, 2007, Art. id. 154102.
[8] S. D. Wu, C. W. Wu, S. G. Lin, C. C. Wang, and K. Y. Lee, "Time series analysis using composite multiscale entropy," *Entropy*, vol. 15, pp. 1069–1084, 2013.
[9] M. C. Mackey and L. Glass, "Oscillation and chaos in physiological control systems," *Science*, vol. 197, pp. 287–289, 1977.
[10] A. Humeau-Heurtier, E. Guerreschi, P. Abraham, and G. Mahé, "Relevance of laser doppler and laser speckle techniques for assessing vascular function: State of the art and future trends," *IEEE Trans. Biomed. Eng.*, vol. 60, pp. 659–666, 2013.
[11] Case western reserve university bearing data center website [Online]. Available: http://csegroups.case.edu/bearingdatacenter/pages/download-data-file
[12] S. D. Wu, C. W. Wu, S. G. Lin, K. Y. Lee, and C. K. Peng, "Analysis of complex time series using refined composite multiscale entropy," *Phys. Lett. A*, vol. 378, pp. 1369–1374, 2014.
[13] Y. Yin and P. Shang, "Weighted multiscale permutation entropy of financial time series," *Nonlin. Dynam.*, vol. 78, pp. 2921–2939, 2014.
[14] M. U. Ahmed and D. P. Mandic, "Multivariate multiscale entropy analysis," *IEEE Signal Process. Lett.*, vol. 19, pp. 91–94, 2012.
[15] L. Zunino, M. C. Soriano, and O. A. Rosso, "Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach," *Phys. Rev. E*, vol. 86, 2012, article id. 046210.