

Complexity

Measuring Complexity of Biomedical Signals

Lead Guest Editor: Gastón Schlotthauer

Guest Editors: Anne Humeau-Heurtier, Javier Escudero, and Hugo L. Rufiner





Measuring Complexity of Biomedical Signals

Complexity

Measuring Complexity of Biomedical Signals

Lead Guest Editor: Gastón Schlotthauer

Guest Editors: Anne Humeau-Heurtier, Javier Escudero,
and Hugo L. Rufiner



Copyright © 2018 Hindawi. All rights reserved.

This is a special issue published in “Complexity.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

- José A. Acosta, Spain
Carlos F. Aguilar-Ibáñez, Mexico
Tarek Ahmed-Ali, France
Basil M. Al-Hadithi, Spain
Juan A. Almendral, Spain
Diego R. Amancio, Brazil
David Arroyo, Spain
Mohamed Boutayeb, France
Arturo Buscarino, Italy
Guido Caldarelli, Italy
Eric Campos-Canton, Mexico
Mohammed Chadli, France
Diyi Chen, China
Giulio Cimini, Italy
Danilo Comminiello, Italy
Sara Dadras, USA
Manlio De Domenico, Italy
Pietro De Lellis, Italy
Albert Diaz-Guilera, Spain
Thach Ngoc Dinh, France
Jordi Duch, Spain
Marcio Eisencraft, Brazil
Joshua Epstein, USA
Mondher Farza, France
Thierry Floquet, France
Mattia Frasca, Italy
Lucia Valentina Gambuzza, Italy
Bernhard C. Geiger, Austria
- Carlos Gershenson, Mexico
Peter Giesl, UK
Sergio Gómez, Spain
Lingzhong Guo, UK
Xianggui Guo, China
Sigurdur F. Hafstein, Iceland
Chittaranjan Hens, Israel
Giacomo Innocenti, Italy
Sarangapani Jagannathan, USA
Mahdi Jalili, Australia
Jeffrey H. Johnson, UK
M. Hassan Khooban, Denmark
Vincent Labatut, France
Lucas Lacasa, UK
Qingdu Li, Germany
Chongyang Liu, China
Xiaoping Liu, Canada
R. M. Lopez Gutierrez, Mexico
Vittorio Loreto, Italy
Didier Maquin, France
Eulalia Martínez, Spain
Marcelo Messias, Brazil
Ana Meštrović, Croatia
Ludovico Minati, Japan
Ch. P. Monterola, Philippines
Marcin Mrugalski, Poland
Roberto Natella, Italy
Nam-Phong Nguyen, USA
- Beatrice M. Ombuki-Berman, Canada
Irene Otero-Muras, Spain
Yongping Pan, Singapore
Daniela Paolotti, Italy
Cornelio Posadas-Castillo, Mexico
Mahardhika Pratama, Singapore
Luis M. Rocha, USA
Miguel Romance, Spain
Avimanyu Sahoo, USA
Matilde Santos, Spain
Hiroki Sayama, USA
Michele Scarpiniti, Italy
Enzo Pasquale Scilingo, Italy
Dan Selișteanu, Romania
Dimitrios Stamovlasis, Greece
Samuel Stanton, USA
Roberto Tonelli, Italy
Shahadat Uddin, Australia
Gaetano Valenza, Italy
Dimitri Volchenkov, USA
Christos Volos, Greece
Zidong Wang, UK
Yan-Ling Wei, Singapore
Honglei Xu, Australia
Xinggang Yan, UK
Massimiliano Zanin, Spain
Hassan Zargazadeh, USA
Rongqing Zhang, USA

Contents

Measuring Complexity of Biomedical Signals

Gastón Schlotthauer , Anne Humeau-Heurtier , Javier Escudero , and Hugo L. Rufiner 
Editorial (3 pages), Article ID 5408254, Volume 2018 (2018)

Complexity-Based Discrepancy Measures Applied to Detection of Apnea-Hypopnea Events

R. E. Rolón , I. E. Gareis , L. E. Di Persia, R. D. Spies, and H. L. Rufiner
Research Article (18 pages), Article ID 1435203, Volume 2018 (2018)

Linear and Complex Measures of Heart Rate Variability during Exposure to Traffic Noise in Healthy Women

Myrela Alves, David M. Garner , Anne M. G. G. Fontes, Luiz Vinicius de Alcantara Sousa ,
and Vitor E. Valenti 
Research Article (14 pages), Article ID 2158391, Volume 2018 (2018)

Invariant Measures Based on the U-Correlation Integral: An Application to the Study of Human Voice

Juan F. Restrepo  and Gastón Schlotthauer
Research Article (9 pages), Article ID 2173640, Volume 2018 (2018)

A Comparison between Theoretical and Experimental Measures of Consciousness as Integrated Information in an Anatomically Based Network of Coupled Oscillators

Antonio J. Ibáñez-Molina  and Sergio Iglesias-Parro 
Research Article (8 pages), Article ID 6101586, Volume 2018 (2018)

Complexity Measures for Quantifying Changes in Electroencephalogram in Alzheimer's Disease

Ali H. Hussein Al-Nuaimi , Emmanuel Jammeh , Lingfen Sun , and Emmanuel Ifeakor 
Research Article (12 pages), Article ID 8915079, Volume 2018 (2018)

Multifractal-Multiscale Analysis of Cardiovascular Signals: A DFA-Based Characterization of Blood Pressure and Heart-Rate Complexity by Gender

Paolo Castiglioni , Davide Lazzeroni, Paolo Coruzzi, and Andrea Faini
Research Article (14 pages), Article ID 4801924, Volume 2018 (2018)

Identification of Alcoholism Based on Wavelet Renyi Entropy and Three-Segment Encoded Jaya Algorithm

Shui-Hua Wang, Khan Muhammad , Yiding Lv, Yuxiu Sui , Liangxiu Han , and Yu-Dong Zhang 
Research Article (13 pages), Article ID 3198184, Volume 2018 (2018)

Efficient Computation of Multiscale Entropy over Short Biomedical Time Series Based on Linear State-Space Models

Luca Faes, Alberto Porta, Michal Javorka, and Giandomenico Nollo
Research Article (13 pages), Article ID 1768264, Volume 2017 (2018)

Editorial

Measuring Complexity of Biomedical Signals

Gastón Schlotthauer ^{1,2}, **Anne Humeau-Heurtier** ³, **Javier Escudero** ⁴,
and Hugo L. Rufiner ^{1,5}

¹*Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Oro Verde, Argentina*

²*Instituto de Investigación y Desarrollo en Bioingeniería y Bioinformática (IBB), UNER-CONICET, Argentina*

³*Univ Angers, LARIS - Laboratoire Angevin de Recherche en Ingénierie des Systèmes, Angers, France*

⁴*School of Engineering, Institute for Digital Communications, University of Edinburgh, Edinburgh, UK*

⁵*Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i),*

Facultad de Ingeniería y Ciencias Hídricas-Universidad Nacional del Litoral/CONICET, Argentina

Correspondence should be addressed to Gastón Schlotthauer; gschlotthauer@conicet.gov.ar

Received 12 July 2018; Accepted 13 July 2018; Published 20 September 2018

Copyright © 2018 Gastón Schlotthauer et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is well known that biomedical signals, such as heart rate variability (HRV), electrocardiogram (ECG), electroencephalogram (EEG), and voice, arise from complex nonlinear dynamical systems, as the cardiovascular, nervous, or phonatory ones. Information extracted from these signals provides insights regarding the status of the underlying physiology. Complexity measures are helpful to quantitatively describe nonlinear biomedical systems and to detect changes in their dynamics that can be associated with physiological or pathological events [1–5]. These measures on biomedical signals and images can be used in a wide field of applications as pathology detection, decision support systems, treatment monitoring, and temporal segmentation. They can also be used to characterize biomedical systems that gave rise to those images and time series. However, in practice, many challenges emerge when these complexity measures are applied, such as the influence of the noise, the quantization effects, the lengths of the available data, or the parameter tuning. Many of these issues are still unsolved [6–8]. How to cope with these difficulties and how to obtain tools that can be employed in clinical practice are the subjects of this special issue. It is focused not only on the application of existing complexity measures on biomedical signals and images but also on the development of new complexity measure algorithms. Some interesting complexity-based works are also associated with machine learning-based strategies, automatization in parameter setting, and applications in

pattern recognition problems, as well as developments and applications of novel complexity estimators for multivariate, multiscale, or multimodal data [9, 10].

In this context, different proposals that explore theory and applications of complexity-based measures related to biomedical signal problems were selected. After a rigorous review process, 8 papers have been accepted for this special issue.

M. Alves et al. contribute the paper entitled “Linear and Complex Measures of Heart Rate Variability during Exposure to Traffic Noise in Healthy Women” on the evaluation of the complexity of HRV during traffic noise exposure. They analyze healthy female students exposed to traffic noise through an earphone for a period of 20 minutes. The traffic noise was recorded from a very busy city street (71–104 dB). They observe no significant changes in the linear analysis approach. Chaotic global analysis and Shannon, Renyi, and Tsallis entropies are also considered. They conclude that traffic noise under laboratory conditions increases the complexity of HRV through chaotic global analysis and some measures of entropy in healthy females.

In their article “Invariant Measures Based on the U-Correlation Integral: An Application to the Study of Human Voice”, J. F. Restrepo and G. Schlotthauer use nonlinear measures such as the correlation dimension, the correlation entropy, and the noise level to characterize normal and pathological voices. Following this line, the reader can identify two original contributions. The first one is an

automated algorithm, based on the recently proposed U-correlation integral, to calculate the aforementioned invariants. The second one is the idea of using the noise level to quantitatively differentiate between voice types. These two proposals are tested using the “Massachusetts Eye & Ear Infirmary” Voice Disorders Database, distributed by Kay Elemetrics. The results show that the voice dynamics has a low dimension which increases during pathological states. This suggests the presence of the inherent stochastic perturbation that strengthens along with the voice type. Finally, the authors conclude that the voice production dynamical system increases in complexity in the presence of a pathology and with the type of voice.

A. J. Ibáñez-Molina and S. Iglesias-Parro explore the interesting problem of consciousness measures in their work “A Comparison between Theoretical and Experimental Measures of Consciousness as Integrated Information in an Anatomically Based Network of Coupled Oscillators.” Theories of consciousness have led to new measures to detect consciousness in a system. The integrated information theory (IIT) is possibly one of the best mathematical rooted attempts to quantify the level of consciousness in a system. From the experimental point of view, perturbational complexity index (PCI) was introduced trying to detect conscious states having almost perfect classification accuracy. In this study, the authors explore the statistical correspondence between theoretical and experimental measures through a well-known neurocomputational model. This model consists of coupled oscillators that can be artificially perturbed. Their results show that both measures are statistically related, but, in principle, this relationship is far to be perfect. These results are discussed in the context of the model of coupled oscillators employed, which mainly focuses on the dynamics of collective synchronization between subsets of brain areas.

The paper entitled “Complexity Measures for Quantifying Changes in Electroencephalogram in Alzheimer’s Disease” by A. H. Hussein Al-Nuaimi et al. explores the application of complexity-based measures as possible biomarkers for Alzheimer’s Disease (AD). This is a progressive disorder that starts many years before its clinical manifestations. For this reason, it is important to find accurate, low-cost, and easy-to-use biomarkers that provide quantitative measures of changes in the brain in the early stages of AD. This is a cross-sectional study that aims to demonstrate the usefulness of EEG complexity measures in this context. Tsallis entropy (TsEn), Higuchi fractal dimension (HFD), and Lempel-Ziv complexity (LZC) methods were analyzed for this purpose. In this study, the complexity measures are derived from EEG frequency bands because of their relationship with brain states. The results show that, for specific EEG frequency bands/channels, the complexity-based measures here evaluated and provided useful information to detect AD.

P. Castiglioni et al., in their paper “Multifractal-Multiscale Analysis of Cardiovascular Signals: A DFA-Based Characterization of Blood Pressure and Heart-Rate Complexity by Gender,” present a new method based on detrended fluctuation analysis for joint multifractal/multiscale analysis which is designed for overcoming the limitations of the current methods for evaluating the multiscale

or multifractal features of cardiovascular signals. In this work, the authors study the multifractal/multiscale nature of the interbeat intervals, systolic blood pressure, and diastolic blood pressure, in 42 female and 42 male healthy volunteers. Their results shown that both scale coefficients and degree of multifractality depend on the type of cardiovascular signal and on the temporal scale, with significant gender differences.

The paper titled “Identification of Alcoholism Based on Wavelet Renyi Entropy and Three-Segment Encoded Jaya Algorithm” by S.-H. Wang et al. describes a processing pipeline of structural magnetic resonance imaging (MRI) data where features are extracted via wavelet Renyi entropy and a classifier is used to distinguish between people with alcoholism and healthy controls. In addition to describing the implementation of the classification approach in Java, this manuscript illustrates the potential of nonlinear metrics to be applied to clinical and biomedical data other than time series. By considering two-dimensional objects such as images, the authors show how entropy-based techniques (Renyi entropy in this case) have potential to be applied in more settings than the traditional analysis of time series.

In their paper entitled “Efficient Computation of Multiscale Entropy over Short Biomedical Time Series Based on Linear State-Space Models,” L. Faes et al. propose a new method to assess complexity, the so-called linear multiscale entropy (MSE). Compared to other MSE-based methods, such as the refined MSE (RMSE), linear MSE is based on theoretical rather than empirical grounds. Therefore, it can be analytically computed from the parametric representation of an observed stochastic process. For this purpose, linear MSE uses linear state-space (SS) models. Simulations and the analysis of short biomedical time series show that, for linear stochastic processes, linear MSE outperforms RMSE.

The paper entitled “Complexity-Based Discrepancy Measures Applied to Detection of Apnea-Hypopnea Events” by R. E. Rolón et al. explores the use of complexity-based discrepancy measures in the context of biomedical signal classification. The authors proposed a method to construct discriminative subdictionaries for sparse representations in the particular case of pulse oximetry signals applied to apnea-hypopnea event detection. Besides traditional discrepancy measures, they study a simple and recently introduced measure called difference of conditional activation frequency (DCAF). They also explore the effect of overcompleteness and redundancy of the dictionary as well as the sparsity level of the representation. Their results show that complexity-based measures are capable of adequately pointing out discriminative atoms of the dictionary. Moreover, DCAF yields competitive averaged detection accuracy rates at low computational cost. This shows that discriminative subdictionary construction methods for sparse representations of pulse oximetry signals constitute a valuable tool for apnea-hypopnea screening.

Conflicts of Interest

The editors declare that they have no conflicts of interest regarding the publication of this special issue.

Acknowledgments

The guest editorial team would like to thank all the authors for their interest in selecting this special issue and for their valuable contributions. The editors would also like to thank all the anonymous reviewers who have contributed to improve the papers.

Gastón Schlotthauer
Anne Humeau-Heurtier
Javier Escudero
Hugo L. Rufiner

References

- [1] M. Costa, A. L. Goldberger, and C. K. Peng, “Multiscale entropy analysis of complex physiologic time series,” *Physical Review Letters*, vol. 89, no. 6, 2002.
- [2] M. Costa, A. L. Goldberger, and C. K. Peng, “Multiscale entropy analysis of biological signals,” *Physical Review E*, vol. 71, no. 2, 2005.
- [3] J. Courtiol, D. Perdikis, S. Petkoski et al., “The multiscale entropy: guidelines for use and interpretation in brain signal analysis,” *Journal of Neuroscience Methods*, vol. 273, pp. 175–190, 2016.
- [4] T. Takahashi, “Complexity of spontaneous brain activity in mental disorders,” *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 45, pp. 258–266, 2013.
- [5] A. C. Yang and S.-J. Tsai, “Is mental illness complex? From behavior to brain,” *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 45, pp. 253–257, 2013.
- [6] K. Kuntzelman, L. Jack Rhodes, L. N. Harrington, and V. Miskovic, “A practical comparison of algorithms for the measurement of multiscale entropy in neural time series data,” *Brain and Cognition*, vol. 123, pp. 126–135, 2018.
- [7] A. V. Lubetzky, D. Harel, and E. Lubetzky, “On the effects of signal processing on sample entropy for postural control,” *PLoS One*, vol. 13, no. 3, article e0193460, 2018.
- [8] W. Xiong, L. Faes, and P. C. Ivanov, “Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: effects of artifacts, nonstationarity, and long-range correlations,” *Physical Review E*, vol. 95, no. 6, 2017.
- [9] X. Li, Z. Zhu, W. Zhao et al., “Decreased resting-state brain signal complexity in patients with mild cognitive impairment and Alzheimer’s disease: a multi-scale entropy analysis,” *Biomedical Optics Express*, vol. 9, no. 4, pp. 1916–1929, 2018.
- [10] B. Marchand and N. Saito, “Earth mover’s distance-based local discriminant basis,” in *Multiscale Signal Analysis and Modeling*, X. Shen and A. I. Zayed, Eds., pp. 275–294, Springer, New York, NY, USA, 2013.

Research Article

Complexity-Based Discrepancy Measures Applied to Detection of Apnea-Hypopnea Events

R. E. Rolón ^{1,2} I. E. Gareis ^{1,3} L. E. Di Persia,¹ R. D. Spies,⁴ and H. L. Rufiner^{1,3}

¹Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i), UNL, CONICET, FICH, Santa Fe, Argentina

²Facultad Regional Paraná, Universidad Tecnológica Nacional, Paraná, Entre Ríos, Argentina

³Laboratorio de Cibernética, Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Oro Verde, Argentina

⁴Instituto de Matemática Aplicada del Litoral, IMAL, UNL, CONICET, FIQ, Santa Fe, Argentina

Correspondence should be addressed to R. E. Rolón; rrolon@sinc.unl.edu.ar

Received 2 February 2018; Accepted 30 May 2018; Published 26 August 2018

Academic Editor: Dimitri Volchenkov

Copyright © 2018 R. E. Rolón et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, an increasing interest in the development of discriminative methods based on sparse representations with discrete dictionaries for signal classification has been observed. It is still unclear, however, what is the most appropriate way for introducing discriminative information into the sparse representation problem. It is also unknown which is the best discrepancy measure for classification purposes. In the context of feature selection problems, several complexity-based measures have been proposed. The main objective of this work is to explore a method that uses such measures for constructing discriminative subdictionaries for detecting apnea-hypopnea events using pulse oximetry signals. Besides traditional discrepancy measures, we study a simple one called Difference of Conditional Activation Frequency (DCAF). We additionally explore the combined effect of overcompleteness and redundancy of the dictionary as well as the sparsity level of the representation. Results show that complexity-based measures are capable of adequately pointing out discriminative atoms. Particularly, DCAF yields competitive averaged detection accuracy rates of 72.57% at low computational cost. Additionally, ROC curve analyses show averaged diagnostic sensitivity and specificity of 81.88% and 87.32%, respectively. This shows that discriminative subdictionary construction methods for sparse representations of pulse oximetry signals constitute a valuable tool for apnea-hypopnea screening.

1. Introduction

Although it is widely used and accepted, the notion of complexity has very often avoided a rigorous formalization. It is therefore not surprising that no universally accepted measure exists yet for quantifying such a concept. In particular, within information theory, the complexity of any element of a code, or of any feature of a signal representation in the context of signal processing, is known to be strongly related to the information it carries or, more precisely, to the value of its entropy. It is important to point out however that, in the context of signal classification, the more informative features (in terms of classification) are not necessarily the ones with larger entropy. Hence, more “ad hoc” measures are needed. In fact, any appropriate complexity measure corresponding to a given feature should be instead, strongly related to the amount of information about class membership provided

by such a feature. One could then think of using as measure of complexity the conditional entropy of the class given the feature. However, features providing the most discriminative information regarding a class are almost always those with lower conditional entropy values, and hence, the best features for classification purposes will be the least complex ones.

Information theory was originally based on the engineering of noisy communication channels, and it is closely associated to a large number of disciplines such as signal processing, artificial intelligence, complex systems, and pattern recognition, to name only a few. We are particularly interested in the latter. Pattern recognition is a discipline which is mainly oriented to the generation of algorithms or methods that can decide an action based upon certain recognized similarities (patterns) in the input data. Within signal classification, which is perhaps one of the most important subfields of pattern recognition, several discrepancy

measures have been used in problems coming from a wide variety of areas such as machine learning [1], image and speech processing [2], neural networks [3], and biomedical signal processing [4, 5]. Among them, the most commonly used is probably the Kullback-Leibler (KL) divergence [6, 7]. This divergence, also known as relative entropy, was used as a discriminative measure for selecting, from a large collection of orthonormal bases, the one attaining maximum information [1]. A more recent approach was introduced by Gupta et al. [8] who used this divergence as a discrepancy measure in the traditional k-nearest neighbor (k-NN) algorithm, yielding competitive classification performances in the context of raw electroencephalographic signal classification. Although it provides certain computational and theoretical advantages, the lack of symmetry of the KL divergence has motivated the development of several symmetric versions such as the so-called J divergence [9] and the well-known and widely used Jensen-Shannon divergence [10].

Sparse representation of signals constitutes a useful technique which has drawn wide interest in recent years due to its success in many applications such as signal and image processing [11]. This technique allows the analysis of the signals by means of only a few well-defined basic waveforms. Due to its advantages, such as robustness to noise and dimension reduction, sparse representation has acquired a large popularity in the area of biomedical signal processing. For example, this technique has been successfully applied to several problems including the estimation of the human respiratory rate [12] and electrocardiographic signal processing, both for signal enhancement and QRS complex detection, for improving heart disease analysis and diagnosis [13]. It is timely to point out however that, up to our knowledge, no applications of discrepancy measures to sparse representation for signal classification are known yet.

All reconstructive methods, such as principal component analysis (PCA), independent component analysis (ICA), and the previously mentioned sparse representations [14], produce particular types of signal representations minimizing a given cost functional which usually involves both fidelity and regularization terms. These methods have been successfully applied in a wide variety of problems such as signal denoising, missing data, and outliers. On the other hand, discriminative methods such as linear discriminant analysis (LDA) are oriented to find optimal decision boundaries to be used for classification tasks. It is well known that for signal classification, which is our main interest in this work, discriminative methods generally outperform reconstructive methods. It is mainly for this reason that several authors have recently developed supervised approaches based on sparse representation which are simultaneously reconstructive and discriminative [15, 16].

The obstructive sleep apnea-hypopnea (OSAH) syndrome [17] is one of the most common sleep disorders and more often than not it remains undiagnosed and therefore not treated. This syndrome is caused by repeated events of partial or total blockage of the upper airway during sleeping, which correspond to events of hypopnea and apnea, respectively. To evaluate the severity degree of the OSAH syndrome, medical physicians have created the so-called

apnea-hypopnea index (AHI), which is defined as the average number of apnea-hypopnea events per hour of sleep. In terms of this index, OSAH is classified as normal, mild, moderate, or severe depending on whether such an index falls in the interval $[0, 5)$, $[5, 15)$, $[15, 30)$ or $[30, \infty)$, respectively. The gold standard test for OSAH diagnosis is a study called polysomnography (PSG). However, PSG is both costly and lengthy and the accessibility to this type of study is limited. Additionally, PSG studies require information coming from a variety of physiological signals such as electroencephalography (EEG), airflow and pulse oximetry (SaO_2). It is known however that cessations of breathing associated with apnea-hypopnea events are always accompanied by a drop in the oxygen saturation level in the SaO_2 signal record, although quite often such a drop is very small and almost impossible to detect by a human observer.

The main objective of this work is precisely to develop a technique based on sparse representations and the use of appropriate discriminative information that be able to accurately and efficiently detect apnea-hypopnea events by using only the SaO_2 signal. Several ways exist for combining discriminative information and sparse representations within the context of signal classification. We shall follow one consisting of using the discriminative information for detecting those atoms having the most frequent activations in order to provide them as input for a classifier. This approach was initially introduced in [4] where two methods using the absolute value of the activation differences of the atoms as a measure of the discriminative information for the detection of OSAH were presented. In this work, a rigorous formalization of such a measure is introduced and compared with several other discrepancy measures for classifying apnea-hypopnea events. Also, the combined effect of using different sizes of nonredundant dictionaries and different sparsity degrees is explored in detail. Results show clearly that the proposed measure is capable of adequately pointing out discriminative atoms in a full dictionary, yielding competitive accuracy rates in the detection of individual apnea-hypopnea events. Additionally, this new approach is computationally very cheap. In fact, it has proved to be at least twice faster than those associated to all other discrepancy measures.

The rest of this article is organized as follows: in Section 2, the obstructive sleep apnea-hypopnea syndrome is explained. Sparse representation of signals is introduced in Section 3. The problem of finding discriminative sub-dictionaries is described in Section 4 while several discriminative information measures are presented in Section 5. Section 6 contains a detailed description about the performed experiments. Results and discussions are introduced in Section 7 while conclusions are presented in Section 8.

2. Sleep Apnea-Hypopnea

Apnea-hypopnea events occur as a consequence of a functional-anatomic disturbance of the upper airway producing its partial or total blockage. At the end of an apnea-hypopnea event, a pronounced desaturation of the blood hemoglobin commonly occurs. These desaturations generate

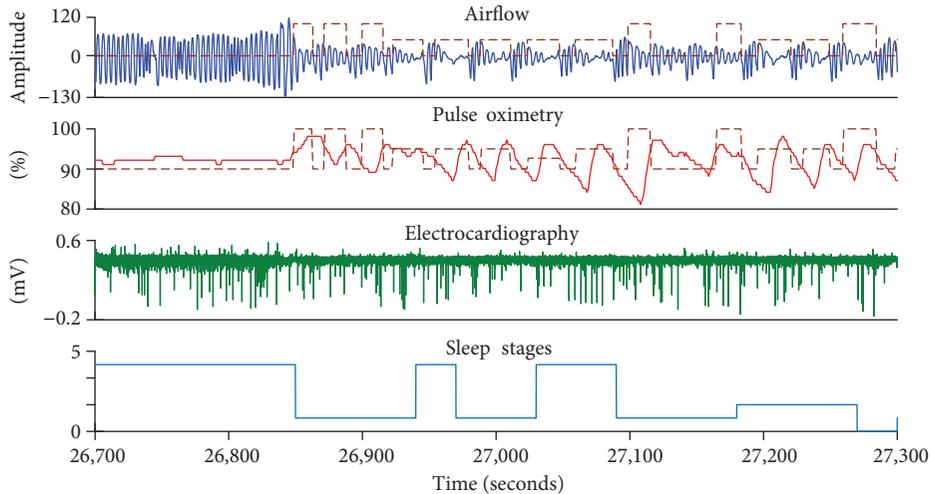


FIGURE 1: A portion of a few number of physiological signals coming from a full PSG. Dashed lines (brown) are apnea-hypopnea labels introduced by the medical expert.

characteristic patterns in the pulse oximetry record known as intermittent hypoxemias. The hypoxemia-reoxygenation cycles promote oxidative stress, angiogenesis, and tumor growth and favor the sympathetic activation with increment of blood pressure and systemic and vascular inflammation with endothelial dysfunction which contributes to multi-organic chronic morbidity, metabolic abnormalities, and cognitive impairment [18]. Additionally, strong correlations between neoplastic diseases and the OSAH syndrome have been described in [19]. Also, a recent study among male mice suggests that OSAH's intermittent hypoxia can be associated to fertility reduction [20]. Currently, this pathology affects more than 4% of the human population around the world [21]. Additionally, it was found that aging, male gender, snoring, and obesity are all risk factors for OSAH syndrome [22].

Although very limited in many countries, overnight polysomnography (PSG) is currently the gold standard tool for diagnosing OSAH syndrome. As previously mentioned, a full PSG consists of the simultaneous measurement of several physiological signals such as EEG, electrocardiography (ECG), respiratory effort, airflow, SaO_2 , and electrical activity produced by skeletal muscles (EMG). Mainly due to its ease of acquisition, we are particularly interested in the SaO_2 signal. Figure 1 shows a typical temporal plot of just a few physiological signals coming from a full PSG. This figure also depicts a portion of an original raw airflow signal as well as the corresponding portion of the SaO_2 signal. The corresponding labels of apnea-hypopnea events (dashed lines) are also shown. Finally, at the bottom of this figure, the electrical activity of the heart as well as the sleep stages are shown. In a typical PSG study, after a normal period of sleep, the recorded signals are provided to medical experts who analyze the whole record and mark the apnea-hypopnea events and sleep stages, needed for the posterior evaluation the AHI index. Due to its complexity and cost, a few alternatives to PSG have been adopted. One of the most popular ones is the so-called home respiratory polygraphy (HRP) [23] which requires no neurophysiological signals. Although

studies have shown that there exists a high correlation between AHI values generated by HRP and PSG studies [24], HRP still needs of several physiological signals, whose acquisition strongly affects the normal sleeping of the person. It is therefore highly desirable to develop a reliable OSAH screening system which makes use of as few as possible physiological signals. In this regard, pulse oximetry, being a cheap and noninvasive technique, has become a suitable alternative for screening purposes [25].

In this work, we shall develop a method for the detection of apnea-hypopnea events that uses only the SaO_2 signals. Our approach leads to a binary classification problem whose main purpose is the detection of the presence (or not) of events of apnea and hypopnea. It is timely to point out that although our method does take into consideration an appropriate fidelity term, we are by no means interested in achieving accurate signal representation.

3. Sparse Representations

As previously mentioned, one of the most popular reconstructive methods is based on sparse representations of the signals involved. Sparsity can be enforced by including upper bounds for the number of nonzero coefficients in the representation of the given signals in terms of atoms in a dictionary.

Formally, the problem of sparse representations of signals can be separated into two subproblems, the so-called sparse coding problem and the dictionary learning problem. We shall now proceed to describe in detail each one of these subproblems. To be more precise, let $\mathbf{x} \in \mathbb{R}^N$ be a discrete signal and let $\Phi \in \mathbb{R}^{N \times M}$ (generally with $M \geq N$) be a dictionary whose columns $\phi_j \in \mathbb{R}^N$ are atoms that we want to use for obtaining a representation of \mathbf{x} of the form $\mathbf{x} = \Phi \mathbf{a}$. Here, and in the sequel, we shall refer to the vector $\mathbf{a} = [a_1 \ a_2 \ \dots \ a_M]^T \in \mathbb{R}^M$ as a “representation” of \mathbf{x} . Sparsity consists essentially of obtaining a representation with as

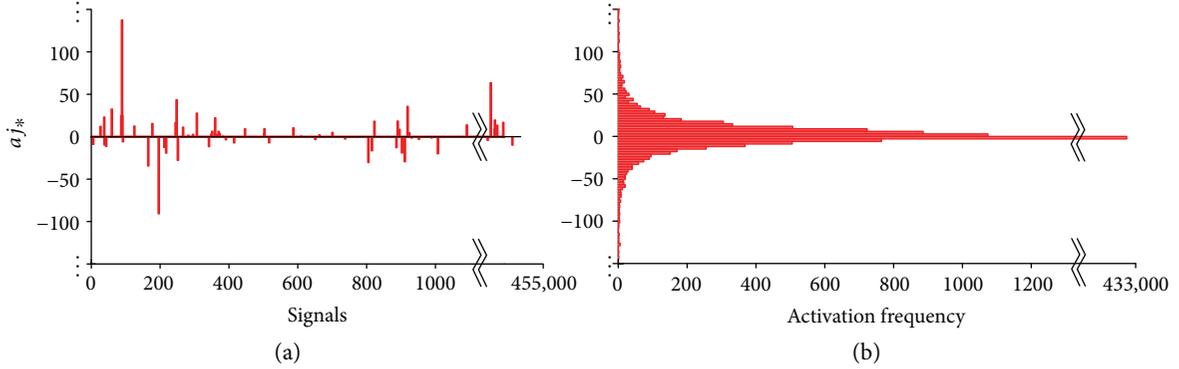


FIGURE 2: The values of the activations of a particular atom for each signal (a) and the corresponding histogram of activations (b).

few nonzero elements as possible. A way of obtaining such a representation consists of solving the following problem:

$$(P_0): \min_{\mathbf{a}} \|\mathbf{a}\|_0 \quad (1)$$

subject to $\mathbf{x} = \Phi\mathbf{a}$,

where $\|\mathbf{a}\|_0$ denotes the l_0 pseudonorm, defined as the number of nonzero elements of \mathbf{a} .

Several questions regarding problem (P_0) immediately arise. Among them are the following: (i) does there exist an exact representation $\mathbf{x} = \Phi\mathbf{a}$?, (ii) if an exact representation exists, is it unique?, (iii) in the case of nonuniqueness, how do we find the “sparsest” representation? and (iv) how difficult is it, from the computational point of view, to solve problem (P_0) ?. Although it is not an objective of this article to get into details about the answers to these questions, it turns out that imposing exact representation is most often a too restrictive and therefore inappropriate constrain and, on the other hand, solving (P_0) is generally an NP-hard problem yielding this approach highly unsuitable for most applications. For more details, we refer the reader to ([26], § 1.8).

In order to overcome some of the difficulties which entail solving problem (P_0) , several relaxed versions of it have been considered. One of them consists of allowing a small representation error while imposing an upper bound on the l_0 pseudonorm of the representation:

$$(P_0^q): \min_{\mathbf{a}} \|\mathbf{x} - \Phi\mathbf{a}\|_2 \quad (2)$$

subject to $\|\mathbf{a}\|_0 \leq q$,

where q is a prescribed integer parameter. This formulation takes into account the existence of possible additive noise terms; in other words, it assumes that $\mathbf{x} = \Phi\mathbf{a} + \mathbf{e}$, where $\mathbf{e} \in \mathbb{R}^N$ is a small energy noise term. Thus, this approach is particularly suitable in most real applications (such as biomedical signal processing) where measured signals are always contaminated by noise. Several greedy strategies have been proposed for solving problem (P_0^q) [27, 28]. Among them, orthogonal matching pursuit (OMP) [28] is perhaps the most commonly used strategy. This greedy algorithm guarantees convergence to the projection of \mathbf{x} into the span of the dictionary atoms, in no more than q iterations. Figure 2 shows an example of the values of a particular coefficient a_{j^*} associated

to the atom ϕ_{j^*} obtained by applying the OMP algorithm for a large number (almost half a million) of segments of SaO_2 signals and its corresponding activation histogram.

Although preconstructed dictionaries, such as the well-known wavelet packets [29], typically lead to fast sparse coding, they are almost always restricted to certain classes of signals. It is mainly for this reason that new approaches introducing data-driven dictionary learning techniques emerged. A Dictionary Learning (DL) problem consists of simultaneously finding a dictionary Φ and representations of n signals \mathbf{x}_i , $1 \leq i \leq n$, (in terms of atoms of such a dictionary) complying with a sparsity constraint for each one of the n signals, while minimizing the total representation error. The (DL) problem associated to the data: $q, M, N \in \mathbb{N}$, $M \geq N$, and n signals in \mathbb{R}^N , $\mathbf{x}_1, \dots, \mathbf{x}_n$, can be formally written as

$$(DL): \min_{\substack{\Phi \in \mathbb{R}^{N \times M} \\ \mathbf{a}_i \in \mathbb{R}^M, \|\mathbf{a}_i\|_0 \leq q, 1 \leq i \leq n}} \sum_{i=1}^n \|\mathbf{x}_i - \Phi\mathbf{a}_i\|_2 \quad (3)$$

The first data-based dictionary learning algorithms were originally developed almost three decades ago [30–32]. Some of them have their roots in probabilistic frameworks by considering the observed data as realizations of certain random variables [30, 31]. In [31] for example, the authors developed an algorithm for finding a redundant dictionary that maximizes the likelihood function of the probability distribution of the data. In that work, an analytic expression for the likelihood function was derived by approximating the posterior distribution by Gaussian functions. An iterative approach for dictionary learning, known as the “method for optimal directions” (MOD), was presented in [32]. The sparse coding stage of this method makes use of the OMP algorithm followed by a simple dictionary updating rule. A new iterative algorithm was recently proposed by Aharon et al. in [14]. This new approach, called “K singular value decompositions” (K-SVD), consists mainly of two stages: a sparse coding stage and a dictionary learning stage. The OMP algorithm is used for the sparse coding stage, which is followed by a dictionary updating step where the atoms are updated one at a time and the representation coefficients are allowed to change in order to minimize the total representation error.

4. Discriminative Subdictionary Construction

Although data-driven dictionary learning algorithms produce sparse representations of signals which are robust against noise and missing data, such representations turn out to be unsuitable if the final objective is signal classification. This is mainly so because those algorithms do not take into account any a priori or available information concerning class membership. In order to overcome this difficulty, some strategies which incorporate appropriate class information have been proposed [4, 16, 33]. In [33], for instance, the authors developed a discriminative dictionary learning method by efficiently integrating a single predictive linear classifier into the cost function of the K-SVD algorithm. A method incorporating a discriminative term into the cost function of the standard K-SVD algorithm is presented in [16]. This method finds an optimal dictionary which is simultaneously representative and discriminative for face recognition tasks. In this work, we make use of a simple approach for detecting discriminative atoms from a previously learned dictionary and using them to build a new subdictionary. This approach, which is originally presented in [4], consists of solving two problems, namely, (i) the above mentioned full (*DL*) problem and (ii) a discriminative subdictionary (*DSD*) construction problem. We shall now proceed to describe problem (iii). One way to obtain discriminative subdictionaries consists of maximizing an appropriate discriminative value functional $G(\cdot)$. Given a data matrix $\mathbf{X} \in \mathbb{R}^{N \times n}$, a class label vector $\mathbf{c} \in \mathcal{C}^n$ (where \mathcal{C} is the set of all classes; in the binary case $\mathcal{C} = \{c_1, c_2\}$), a dictionary $\Phi \in \mathbb{R}^{N \times M}$ and $p \in \mathbb{N}$ (with $p < M$), the most discriminative subdictionary $\hat{\Phi}^{\mathbf{d}} \in \mathbb{R}^{N \times p}$, according to an appropriate prescribed discriminative value functional $G_{\mathbf{X}, \mathbf{c}, \Phi} : \mathbb{R}^{N \times p} \rightarrow \mathbb{R}_0^+$, is defined as

$$(DSD): \hat{\Phi}^{\mathbf{d}} = \arg \max_{\substack{\mathbf{d} \doteq [i_1 i_2 \dots i_p] \\ i_j \in \{1, 2, \dots, M\} \\ i_j \neq i_k \forall j \neq k}} G_{\mathbf{X}, \mathbf{c}, \Phi}(\Phi^{\mathbf{d}}) \quad (4)$$

where for $\mathbf{d} \doteq [i_1 i_2 \dots i_p]$, $\Phi^{\mathbf{d}}$ denotes the $N \times p$ matrix whose j th column is the i_j th column of Φ . The function G , which must be provided, quantifies the discriminative power of each subdictionary $\Phi^{\mathbf{d}}$. Thus, large values of G correspond to highly discriminative subdictionaries while small values of G are associated to subdictionaries with low discriminability.

Several questions concerning problem (*DSD*) clearly emerge. Among them are the following: (i) how do we find an appropriate discriminative value function G ?, (ii) given the functional G , does problem (*DSD*) have a solution?, (iii) if it does, is it unique?, (iv) in the case of nonuniqueness, how do we decide which subdictionary, among the optimizers, is the best for our classification purposes? and (v) how difficult is it, in terms of computational cost, to solve problem (*DSD*)?. Although this problem has not been extensively studied, it is known that solving (*DSD*) is

computationally very challenging for $p > 1$, mainly due to the combinatorial explosion problem. A way to overcome the computational complexities entailed by problem (*DSD*) consists of defining an appropriate discriminative value functional G for $p = 1$. In that way G is independently evaluated at each one of the atoms (columns) of Φ and the discriminative subdictionary $\Phi^{\mathbf{d}} \in \mathbb{R}^{N \times p^*}$ is constructed by stacking side-by-side the first p^* ranked columns of Φ with largest G values. This simplification is based on the assumption that each atom in the dictionary is used to model specific characteristics that are not completely modeled by the other atoms. Thus, the discriminative information provided by a particular atom will be different from the information contributed by other atoms.

5. Discriminative Value Functions for Atom Selection

Several ways for appropriately constructing discriminative value functions G exists. In this section, we present two different approaches to define such a function, namely, (i) using traditional discrepancy measures and (ii) using a new discriminative measure to which we shall refer as the ‘‘Difference of Conditional Activation Frequency’’ (DCAF). We shall previously need to introduce an appropriate setting and terminology regarding probability density functions (PDFs) in the context of sparse representations for signal classification.

Here, and in the sequel, we shall consider the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ as realizations of a particular random vector \mathbf{X} . Any sparse representation of those vectors will result in the PDFs of each coefficient a_j (associated to the atom ϕ_j) showing a very concentrated peak at zero with heavy tails (as depicted in Figure 2). In the context of binary signal classification, it is reasonable to think that if a given atom ϕ_j is highly discriminative, then the conditional PDFs $\pi(a_j | c_1)$ and $\pi(a_j | c_2)$ will be significantly different. Thus, if a dictionary Φ is poorly discriminative, then one should expect $\pi(a_j | c_1) \approx \pi(a_j | c_2)$ for all j .

Although the elements a_j of the representation vector \mathbf{a} are in general real numbers, for practical reasons, it is appropriate to discretize them. That can be done in the usual way by partitioning the real line \mathbb{R} into intervals $I_k \doteq ((k - 1/2)\Delta, (k + 1/2)\Delta]$, $k \in \mathbb{Z}$, of length Δ and the associated discretized random variable $\mathcal{X}_j \doteq \sum_{k \in \mathbb{Z}} k \chi_{I_k}(a_j)$. The corresponding probability mass function (PMF) is $p_{\mathcal{X}_j}(k) = P(a_j \in I_k) = \int_{I_k} \pi(a_j) da_j$, $k \in \mathbb{Z}$. Figure 3 shows the estimated PMF and the corresponding conditional PMFs (given each one of the two classes), both for a nondiscriminative and a discriminative atom using SaO₂ signals.

We shall now proceed to define how we compute the discriminative value function G . Given the data matrix $\mathbf{X} \in \mathbb{R}^{N \times n}$, the corresponding class label vector $\mathbf{c} \in \mathcal{C}^n$ and a full dictionary $\Phi \in \mathbb{R}^{N \times M}$, the first step consists of obtaining the sparse matrix $\mathbf{A} \doteq [\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_n] \in \mathbb{R}^{M \times n}$ by applying the OMP algorithm. The j th row of this sparse matrix is then used for estimating the conditional PMFs $p_{\mathcal{X}_j}(\cdot | c_1)$ and $p_{\mathcal{X}_j}(\cdot | c_2)$. Finally, the value of G at the atom ϕ_j is computed as the

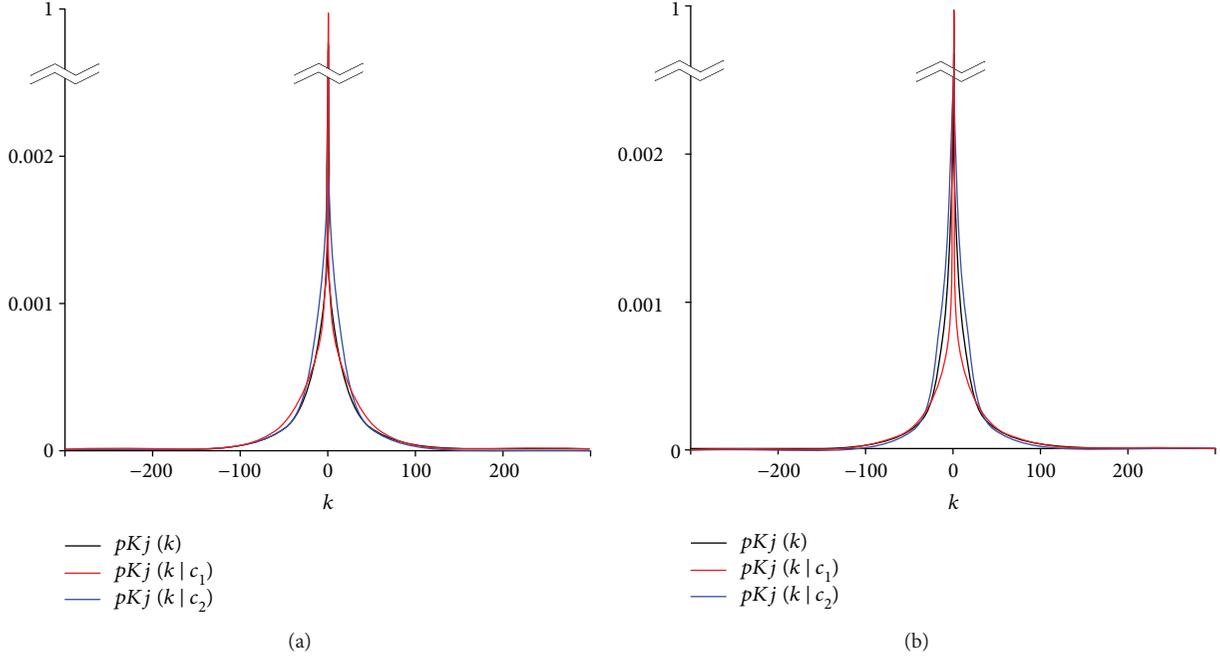


FIGURE 3: Estimated probability mass functions for a nondiscriminative atom ϕ_j (a) and a discriminative one (b).

discrepancy (as quantified by an appropriate discrepancy measure) between these two PMFs. In what follows, we introduce the discrepancy measures that we shall use in this work.

5.1. Traditional Discrepancy Measures. A great diversity of measures whose purpose is performing comparisons between probability distributions exists [34]. In this work, the best known and more commonly used ones are compared in terms of their performance for selecting the most discriminative atoms in a dictionary. The KL, J, and JS divergence measures were utilized, along with the Fisher score (F).

The KL divergence [7] is probably the most widely used information “distance” measure from a theoretical framework, and it was successfully applied in numerous problems for signal classification [1, 35, 36]. To compare the two conditional PMFs associated with the activation of the j th atom, the KL distance was used as follows:

$$\text{KL}(p_{\mathcal{X}_j}(\cdot | c_1), p_{\mathcal{X}_j}(\cdot | c_2)) \doteq \sum_{k \in \mathbb{Z}} p_{\mathcal{X}_j}(k | c_1) \log \left(\frac{p_{\mathcal{X}_j}(k | c_1)}{p_{\mathcal{X}_j}(k | c_2)} \right), \quad (5)$$

assuming that $0 \log(0) \doteq 0$.

Despite the computational and theoretical properties provided by KL distance, what usually becomes a trouble in many problems of signal classification is its lack of symmetry. It can be easily seen that altering the order of the arguments in (5) can change the output value. To solve this issue, a symmetric version of the KL distance can be used such as the J divergence [9], which, even though was not initially created as a symmetric version of the KL distance, is the sum of the

two possible KL distances between probability distributions. In this article, the J divergence is defined as follows:

$$\text{J}(p_{\mathcal{X}_j}(\cdot | c_1), p_{\mathcal{X}_j}(\cdot | c_2)) \doteq \text{KL}(p_{\mathcal{X}_j}(\cdot | c_1), p_{\mathcal{X}_j}(\cdot | c_2)) + \text{KL}(p_{\mathcal{X}_j}(\cdot | c_2), p_{\mathcal{X}_j}(\cdot | c_1)). \quad (6)$$

Another symmetric smoothed version of the KL distance is the JS divergence [10]. For the problem of comparing the two conditional probabilities associated to each class it is defined as

$$\text{JS}(p_{\mathcal{X}_j}(\cdot | c_1), p_{\mathcal{X}_j}(\cdot | c_2)) \doteq w_1 \text{KL}(p_{\mathcal{X}_j}(\cdot | c_1), q_{\mathcal{X}_j}(\cdot)) + w_2 \text{KL}(p_{\mathcal{X}_j}(\cdot | c_2), q_{\mathcal{X}_j}(\cdot)), \quad (7)$$

where $q_{\mathcal{X}_j}(\cdot) = w_1 p_{\mathcal{X}_j}(\cdot | c_1) + w_2 p_{\mathcal{X}_j}(\cdot | c_2)$ and w_1 and w_2 are the weights associated to each of the conditional PMFs, with $w_1, w_2 \geq 0$ and $w_1 + w_2 = 1$. An interesting feature of the JS distance is the fact that different values of weights (w_1 and w_2) can be assigned to the probability distributions according to their importance. In this work, $w_1 = P(c_1)$ and $w_2 = P(c_2)$, that is, the weights are associated with the a priori probabilities of the classes. Note that computing the JS distance as defined here is the same as computing the mutual information between the class and the activations, that is, $\text{JS}(p_{\mathcal{X}_j}(\cdot | c_1), p_{\mathcal{X}_j}(\cdot | c_2)) = \text{MI}(\mathcal{X}_j, \mathcal{C})$.

Within signal classification problems, F is a measure which has been extensively used. Unlike the other measures presented here, that require estimations of the conditional PMFs, F uses just two parameters of the distributions (the means and standard deviations). This makes this measure

much less expensive computationally speaking, but implicitly assumes certain characteristics of the distribution under study (i.e., second-order characteristics). In the case of univariate binary problem at hand, the F can be defined as

$$F\left(p_{\mathcal{X}_j}(\cdot|c_1), p_{\mathcal{X}_j}(\cdot|c_2)\right) \doteq \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}, \quad (8)$$

where μ_ℓ and σ_ℓ^2 are the mean and standard deviation of $p_{\mathcal{X}_j}(\cdot|c_\ell)$ [37].

Although the abovementioned discrepancy measures provide, in a certain sense, “measures” of distance between two probability distribution functions, most of them (such as the KL divergence and those symmetric variants) are not strictly a metric. For instance, the KL divergence is a nonsymmetric discrepancy measure where the triangular inequality is not satisfied. Nevertheless, $\text{KL}(p_{\mathcal{X}_j}(\cdot|c_1), p_{\mathcal{X}_j}(\cdot|c_2))$ is a nonnegative measure, that is, $\text{KL}(p_{\mathcal{X}_j}(\cdot|c_1), p_{\mathcal{X}_j}(\cdot|c_2)) \geq 0$ and $\text{KL}(p_{\mathcal{X}_j}(\cdot|c_1), p_{\mathcal{X}_j}(\cdot|c_2)) = 0$ if and only if $p_{\mathcal{X}_j}(\cdot|c_1) = p_{\mathcal{X}_j}(\cdot|c_2)$.

5.2. Difference of Conditional Activation Frequency. In a previous work, a method called Most Discriminative Column Selection (MDCS) for the construction of a discriminative subdictionary was originally presented [4]. The sparse representations of the signals in terms of subdictionaries constructed using MDCS provided good performance in the detection of apnea-hypopnea events. In the mentioned work, the most discriminative atoms were identified by comparing the difference of conditional activation frequency.

The candidates to be considered as “most discriminative” according to [4] are those atoms with higher absolute difference between conditional activation probabilities given the class. That is, an atom is considered as highly discriminative if it is active, in proportion, more times for one of the classes. The use of this approach as a measure of discriminative power follows from the idea that one of the most expressive parameters regarding the importance of a given atom is its activation probability. Moreover, if certain atoms are active mostly for a given class, then it is assumed they represent features of importance in the description of that particular class.

Following this idea, DCAF is defined as

$$\text{DCAF}\left(\eta_1^j, \eta_2^j\right) \doteq \left| \eta_1^j - \eta_2^j \right|, \quad (9)$$

where

$$\eta_\ell^j \doteq \frac{\text{number of activations of the } j\text{th atom for } c_\ell}{\text{number of } c_\ell \text{ samples}}. \quad (10)$$

The measure defined in (9) is symmetric; its value is always ≥ 0 and is inexpensive in terms of computing (if the classes are balanced, the DCAF can be replaced just by simply counting, without the necessity of dividing with the number of samples).

It can easily be seen that the definition of η_ℓ^j in (10) is equal to the maximum likelihood estimation of the conditional probability of activation, that is,

$$p_{\mathcal{X}_j}(k \neq 0 | c_\ell) \approx \eta_\ell^j. \quad (11)$$

Replacing this expression in (9), we can write

$$\begin{aligned} \text{DCAF}\left(\eta_1^j, \eta_2^j\right) &\approx \left| p_{\mathcal{X}_j}(k \neq 0 | c_1) - p_{\mathcal{X}_j}(k \neq 0 | c_2) \right| \\ &\approx \left| \left(1 - p_{\mathcal{X}_j}(k = 0 | c_1)\right) - \left(1 - p_{\mathcal{X}_j}(k = 0 | c_2)\right) \right| \\ &\approx \left| p_{\mathcal{X}_j}(k = 0 | c_2) - p_{\mathcal{X}_j}(k = 0 | c_1) \right|, \end{aligned} \quad (12)$$

finally expressing the DCAF in terms of the complementary conditional probabilities that the atoms will not be activated. With the exception of the F, all the measures presented in Section 5.1 can be expressed as summations, where only one of the terms is computed using the probabilities that $k = 0$. However, due to the high sparsity of the representations the terms associated with $k = 0$ are particularly important. This fact allows us to expect some correlation between the results obtained with the different discrepancy measures and the DCAF.

Figure 4 shows a representation of the conditional PMFs associated to the activations of two different atoms (left side) as well as an illustration of such functions where the peaks centered at zero ($k = 0$) were discarded (middle). It is important to note that, when excluding the zero-centered peak from the graphic, a significant reduction in the magnitude of the y -axis scale is produced which highlights the importance of the activation probability of sparse representations. However, the discrepancy between the distributions is not only due to the atoms activation probability, since slight differences between the probability values for all $k \neq 0$ exist (zoom-in region). Additionally, the absolute values of these differences are represented by the gray regions. It is also important to point out that these area values shown in gray ($\sum_{k \neq 0} |p_{\mathcal{X}_\ell}(k | c_1) - p_{\mathcal{X}_\ell}(k | c_2)|$) are not necessarily equal to those corresponding to the DCAF values. Nevertheless, for symmetric PMFs with high kurtosis and heavy tails (such is the case of the PMFs used in this work), the conditional and a priori distributions are similar and therefore both area values are close to each other.

6. Experimental Setup

This section presents the proposed system and its configuration settings, aimed at detecting patients suspected of suffering from moderate to severe OSAH syndrome. It also describes the database used for training and testing the method along with the measures selected for assessing its performance.

The main objective of our research is to explore the effect of using discrepancy measures to rank the atoms according to their discriminative power. Also, the experiments are designed to determine the effect of using dictionaries with different degrees of overcompleteness (redundant dictionaries) for the detection of apnea-hypopnea events. Additionally, the performance of the system for different sizes of subdictionaries and sparsity degrees is analyzed.

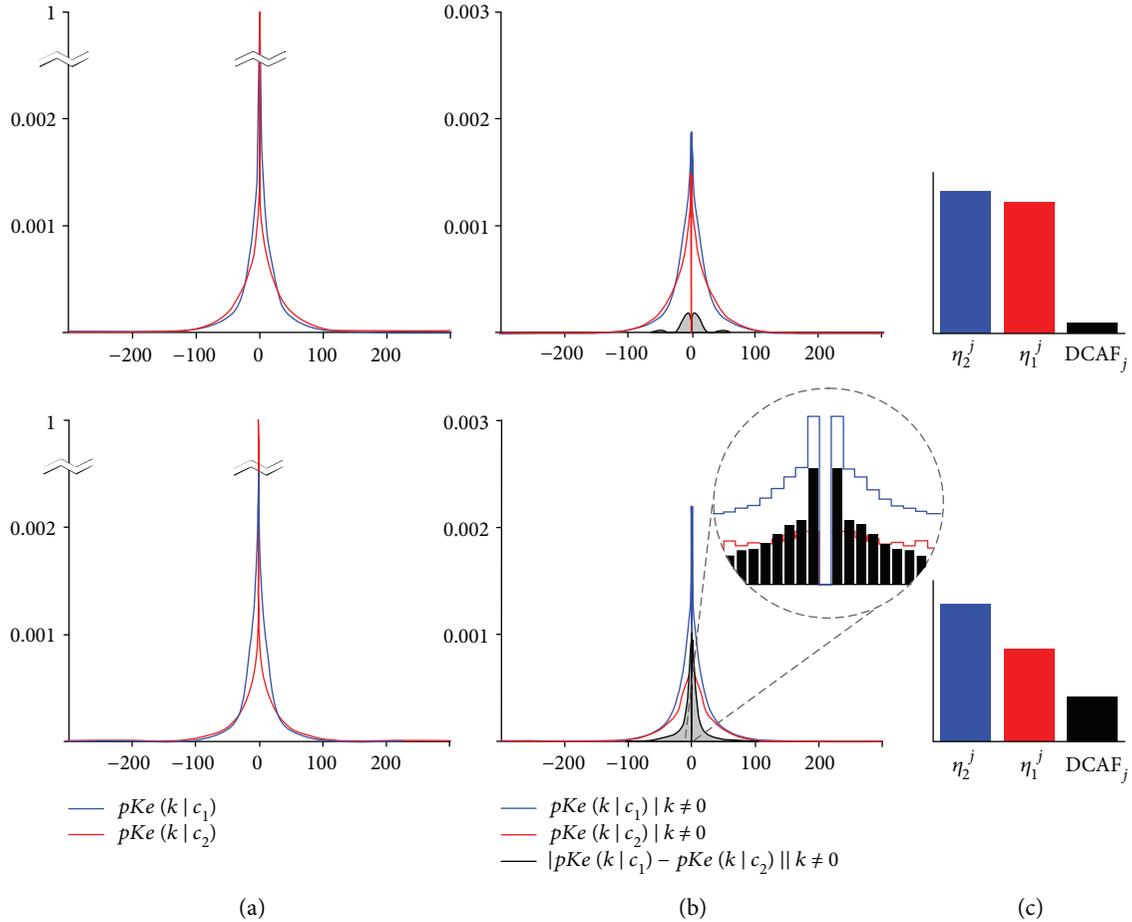


FIGURE 4: A representation of the conditional PMFs corresponding to the activations of two different atoms (a), the same functions excluding the peaks centered at zero ($k = 0$) and the absolute value of their differences (b), and a graphical interpretation of the DCAF (c). The top row corresponds to a nondiscriminative atom (ϕ_j) while the bottom row corresponds to a discriminative one (ϕ_j).

Figure 5 shows a simplified block diagram of the presented system. It can be observed that our system comprises a training phase (above) and a testing phase (below). To clarify the system's description, we divided it into three different stages, namely, stage I, stage II, and stage III. It can be seen that stages I and II are included into training and testing phases while stage III is only used during testing. Stage I is composed by a preprocessing block whose inputs are the raw SaO_2 signals, and its outputs are filtered segments of such signals, as described in Section 6.1. At the training phase, stage II receives segmented signals and finds an optimal discriminative subdictionary. During the testing phase, stage II obtains a sparse matrix in terms of the previously found subdictionary. These processes are thoroughly described in Section 6.2. Finally, the obtained sparse codes are used as input of stage III. This stage detects apnea-hypopnea events and estimates the AHI value, as described in Section 6.3.

6.1. Database and Signal's Preprocessing. The Sleep Heart Health Study (SHHS) dataset [38, 39] was originally designed to study correlations between sleep-disordered breathing and cardiovascular diseases. This dataset includes a large number

of PSG studies, each of them containing several physiological signals such as EEG, ECG, nasal airflow and SaO_2 . Medical expert annotations of sleep stages, arousals, and apnea-hypopnea events are also provided. In this work, only the SaO_2 signal (sampled at 1 Hz) and its corresponding apnea-hypopnea labels are considered for performing the experiments. In this article, the first online version of such a database (SHHS-2) is used. This version of the database contains a total of 995 freely available PSG studies (<https://physionet.org/physiobank/>).

The SaO_2 signals are mainly degraded by patient movements, baseline wander, disconnections, and the limited resolution of pulse oximeters, among other factors. When a disconnection occurs, the recording during the time interval where the sensor signal is blocked is lost. In order to overcome this inconvenience, the values of blood oxygen saturation during such an interval are linearly interpolated. To denoise the signals, a wavelet processing technique [40] is used. The denoising process is performed by zeroing the approximation coefficients at level 8, as well as the coefficients of the first three detail levels of the discrete dyadic wavelet transform with mother wavelet Daubechies 2. The signals are then synthesized using the modified wavelet

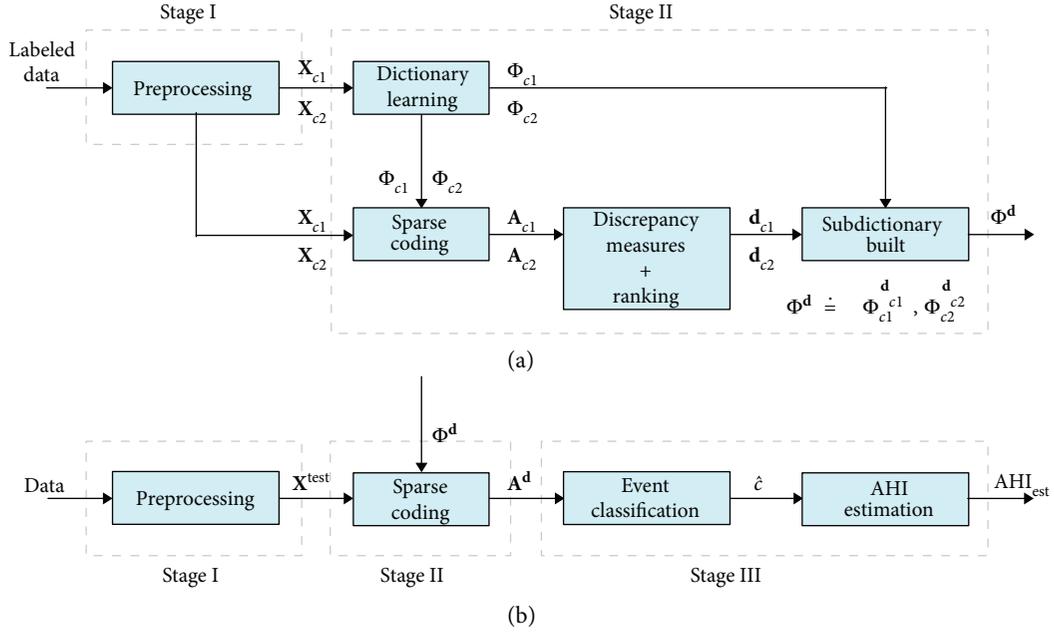


FIGURE 5: Block diagram of the proposed system during training (a) and testing (b).

coefficients by inverse discrete dyadic wavelet transform. The application of this wavelet decomposition technique has the effect of a band-pass filter where the baseline wander and both the low-frequency noise and the high-frequency noise, as well as the quantization noise are eliminated. Figure 6 shows a small fragment of the original raw SaO_2 signal (top) and its wavelet-filtered version (bottom). Labels of apnea-hypopnea events (dashed lines) introduced by the medical experts are also added. These labels were generated by medical experts using the airflow information and thus are not aligned to the desaturations, that is, there is a variable delay between the start time of an event and the corresponding desaturation.

The application of the sparse representation technique requires an appropriate segmentation of the signals. Segments of length $N = 128$ (corresponding to 128 seconds of the signal recording) with a 75% overlapping between two consecutive segments are taken. It is appropriate to point out that although several overlapping percentages were tested, the best system performances were yielded by a 75% overlapping. This redundancy prevents apnea-hypopnea events from being undetected. In this segmentation process, the time intervals where a disconnection occurs are discarded. The segments of pulse oximetry signals are then simultaneously arranged as column vectors $\mathbf{x}_i \in \mathbb{R}^N$ and labeled with ones (c_1) and minus ones (c_2), where a one corresponds to apnea-hypopnea events, and a minus one to the lack of it. Finally, a signal matrix \mathbf{X} is built by stacking side-by-side the column vectors \mathbf{x}_i , that is, the signal matrix is defined as $\mathbf{X} \doteq [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_n]$.

As mentioned above, the entire dataset used in this work contains 995 complete studies, 41 of which were not taken into account for performing the experiments since the size of the signal vectors differs from the corresponding vector

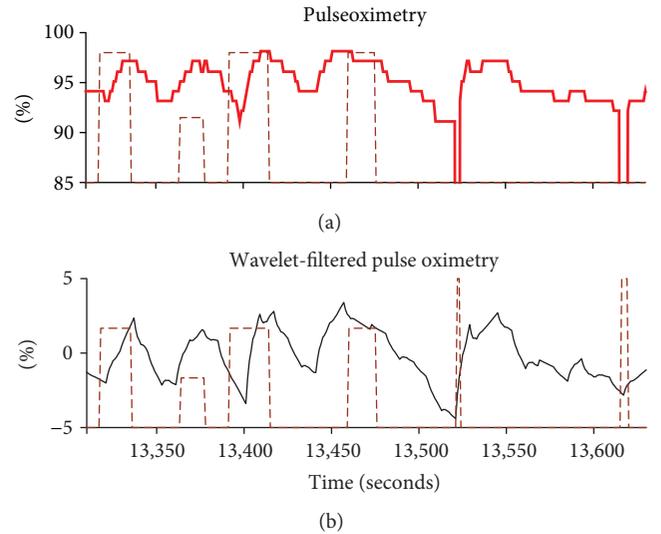


FIGURE 6: A small fragment of a pulse oximetry signal (a) and its wavelet-filtered version (b). Dashed lines represent labels of apnea-hypopnea events established by the medical expert.

of class labels. Among the remaining 954 studies, a subset of 667 (70%) studies were randomly selected and fixed for learning the dictionary and training the classifier. The remaining 287 (30%) studies were left out for the final test. The SaO_2 signals were filtered using wavelet filters and segmented as explained previously into column vectors of size 128. After performing the filtering and segmentation process, a signal matrix $\mathbf{X}^{\text{train}}$ of size 128×455515 is assembled by joining two previously constructed signal matrices, one for

each class, $\mathbf{X}^{\text{train}} \doteq [\mathbf{X}_{c_1}^{\text{train}} \ \mathbf{X}_{c_2}^{\text{train}}]$, which contain 183,163 and 272,352 segments, respectively. On the other hand, for each study included into the testing dataset, a testing matrix \mathbf{X}^{test} is built.

6.2. Sparse Coding and Subdictionary Construction. In our experiments, the learning of the dictionaries is performed by using the traditional K-SVD method [14]. Optimized MATLAB codes for dictionary learning using K-SVD as well as for sparse coding using the OMP algorithm are freely available for academic and personal use at the Ron Rubinstein's personal web page (<http://www.cs.technion.ac.il/~ronrubin/software.html>). At the beginning, the atoms assigned to conform the initial dictionary are randomly selected from the input signal matrix for training without taking into account any information about the classes. If the signal's space dimension is fixed, which should be the effect of constructing dictionaries with different overcompleteness degree?. To answer this question, three types of dictionaries denoted by $\Phi 1$ of size 128×128 , $\Phi 2$ of size 128×256 , and $\Phi 4$ of size 128×512 , corresponding to redundancy factors of 1, 2, and 4, respectively, were built. First, the dictionary $\Phi 1$ was constructed by joining two subcomplete dictionaries of sizes 128×64 denoted by $\Phi 1_{c_1}$ and $\Phi 1_{c_2}$ learned using a large number of training segments (a total of 100,000 segments for each of the classes) belonging to the classes c_1 and c_2 , respectively. Following the same idea, redundant dictionaries denoted by $\Phi 2$ (256 atoms) and $\Phi 4$ (512 atoms) were appropriately built. At the dictionary learning stage, the number of nonzero elements was selected and fixed as a percentage value of 12.5 of the atoms conforming the dictionary. Also, a total of 30 iterations of the K-SVD algorithm were performed.

Once the dictionary has already been trained, the sparse representation vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ corresponding to the input signals $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ are obtained by applying the OMP algorithm. In such a procedure, the nearest integer number to a percentage value of 12.5 of M is selected and fixed. The reason for having chosen this percentage value is because it presented the best trade-off between representativity and discriminability of the segments. Thus, sparsity values of $q = 16$, $q = 32$ and $q = 64$ are selected to represent the input signals for training in terms of the full dictionaries $\Phi 1$, $\Phi 2$ and $\Phi 4$, respectively.

Histograms are typically used to approximate data distributions. In this work, we make use of histograms of the atom's activations to approximate the PDFs. The discretization process was performed by using a Δ value of 0.5. The detection of the most discriminative atoms is obtained by maximizing the discrepancy between the conditional PMFs of the atom's activations given the classes. This objective is achieved using the proposed DCAF measure as well as those denoted by KL, J, JS, and F. The application of different discrepancy measures to the sparse vectors allows for the selection of different "discriminative atoms," which implies the construction of discriminative subdictionaries which are essentially different. The construction of subdictionaries, here denoted by $\Phi 1^d$, $\Phi 2^d$ and $\Phi 4^d$, is performed by selecting atoms from $\Phi 1$, $\Phi 2$, and $\Phi 4$, respectively. Once the most discriminative atoms are detected, the subdictionary is built and

consequently the feature vectors are obtained by applying the OMP algorithm. Finally, each feature vector is assigned to be the input of the ELM classifier.

6.3. Event Detection and AHI Estimation. Multilayer perceptron (MLP) neural networks trained for signal classification have proved to be a tool which provides quite good performances for OSAH syndrome detection [4]; however, the process of training this class of neural network becomes very costly mainly in terms of time. For this reason, in this work, we propose the use of extreme learning machine (ELM) [41] which is a type of single-hidden layer feedforward neural networks (SLFNs), instead of using MLP neural networks. Theoretically, this algorithm (ELM) results in providing a good generalization performance at extremely fast learning speed. The experimental results based on a few artificial and real benchmark function approximation and classification problems including large complex applications show that ELM can produce good generalization performance in most cases and can learn thousands times faster than conventional popular learning algorithms for feedforward neural networks [42].

Basic ELM classifier's MATLAB codes are available for download on the Guang-Bin Huang's web page (http://www.ntu.edu.sg/home/egbhuang/elm_codes.html). To train such a classifier, the main parameters to be fixed are the number of neurons in the hidden layer as well as the activation function of the neurons. In our experiments, the number of neurons in the hidden layer of the ELM corresponds to four times the feature vector dimension. Also, the well-known sigmoid activation function, which is the most common activation function in the nodes of the hidden and/or output layer, is chosen.

In order to evaluate the performance of the proposed classifier in the detection of individual apnea-hypopnea events (a local approach), or more specifically, in the identification of persons suspected of suffering from moderate to severe OSAH syndrome (a global approach), three performance measures are used. For the identification of single segments containing apnea-hypopnea events, the sensitivity (SE_{AH}) represents the total number of correctly classified segments of signals for which any apnea-hypopnea event occurred. Following the same idea, for the detection of individual segments of signals "not containing" any apnea-hypopnea event, the specificity (SP_{AH}) is defined as the total number of correctly classified segments for which any apnea-hypopnea is not present. The accuracy (AC_{AH}) is finally defined as follows:

$$AC_{AH} \doteq \frac{1}{n} \sum_{i=1}^n \delta(c_i, \hat{c}_i), \quad (13)$$

where n represents the total number of segments, c_i and \hat{c}_i denote the corresponding class label of the i th segment and the corresponding prediction of the classifier, respectively, and $\delta(x, y)$ represents the delta function whose output is true (one) if the condition $x = y$ is satisfied and false (zero) otherwise.

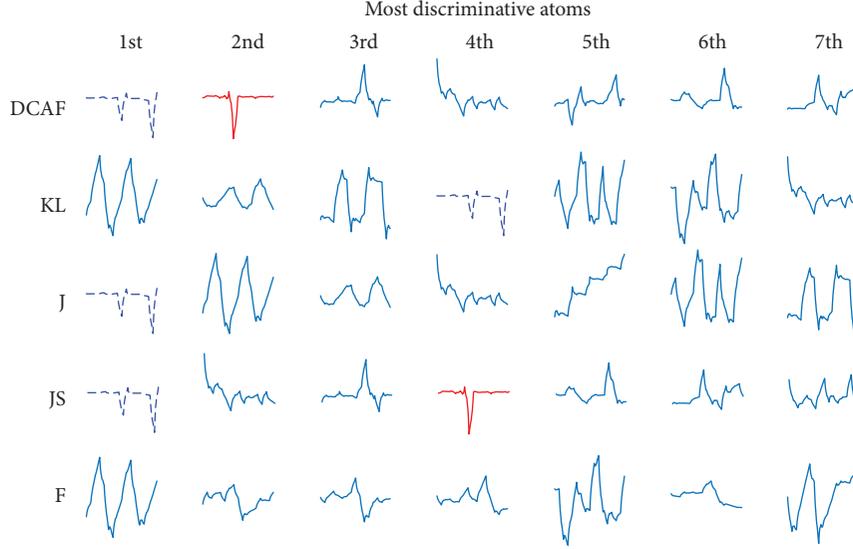


FIGURE 7: Waveforms corresponding to the first seven ranked atoms according to each one of the evaluated measures.

The differences in performance obtained for the event detection between each discrepancy measure were evaluated in order to test whether or not they are statistically significant. The test was performed assuming statistical independence of the classification errors for the different studies and approximating the error's binomial distribution by means of a normal distribution. These assumptions are reasonable due to the large number of SaO₂ signal segments available for each study (about 1100 segments per study, totaling 301,306 segments).

The estimated AHI index (AHI_{est}) is defined as the average number of predicted events per hour of study. This new index is used for OSAH syndrome detection. In this case, the sensitivity (SE_{OSAH}) is defined as the ratio of persons with OSAH syndrome for whom the final test is positive, and the specificity (SP_{OSAH}) is defined as the ratio of health patients for whom the final test is negative. Also, the area under the ROC curve (AUC) derived from a receiver operating characteristic (ROC) analysis [43] is used. A ROC analysis consists of computing the values of the sensitivity and specificity across all the possible detection threshold (DT) values. Then, the ROC curve is built by performing a plot of $1 - \text{specificity}$ versus sensitivity values. This curve has been widely used by medical physicians for evaluating diagnostic tests [44]. A comparison between two different methods can be effectively done by finding the “optimal” (in certain sense) cut-off point of the curve and evaluating their corresponding performances. Finally, the accuracy AC_{OSAH} is defined as follows:

$$AC_{OSAH} \doteq \frac{1}{m} \sum_{i=1}^m \delta \left(AHI_{est}^{(i)} > DT, AHI^{(i)} > 15 \right), \quad (14)$$

where m corresponds to the total number of studies coming from the testing dataset and “DT” is the detection threshold value which adjusts overestimation of the events produced in the segmentation process. The value of DT results in the best cut-off point of the ROC curve. This point, which

maximizes simultaneously sensitivity and specificity, corresponds to the minimum Euclidean distance (d_{min}) to the point (0,1) of the ROC curve.

7. Results and Discussion

In this section, results of the performed experiments are presented and discussed. This section is mainly separated into two subsections, namely, (i) the performance tuning section and (ii) the optimal system performance section.

7.1. Performance Tuning. This section presents results of the exploratory experiments performed to find optimal configurations of the proposed system. As explained in Section 6.2, three different full dictionaries called Φ_1 , Φ_2 , and Φ_4 were learned by applying the standard K-SVD algorithm. In this process, it is expected that most dictionary atoms would capture high-frequency oscillations and normal respiration cycles in SaO₂ signals. It is important to point out however that typical desaturations in signals associated to apnea-hypopnea events should be encoded by some atoms. Secondly, the sparse matrices A_1 , A_2 , and A_4 were obtained by applying the OMP algorithm. As described in Section 6.2, several measures were used to quantify the discriminative degree of individual atoms of each one of the studied dictionaries. Finally, the dictionary atoms were ranked in decreasing order of magnitude according to their discriminative power. Figure 7 shows the waveforms of the first seven ranked atoms of the dictionary Φ_1 according to our measure (first row) as well as the first seven ranked atoms of such a dictionary according to all other discrepancy measures (rows from two to five). It can be seen that the most discriminative atom selected by DCAF (dashed waveform) provides information about two well-defined desaturations in the signal. It is also important to point out that this atom corresponds to the most discriminative one when using J divergence or eventually when using the JS divergence. Moreover, one can

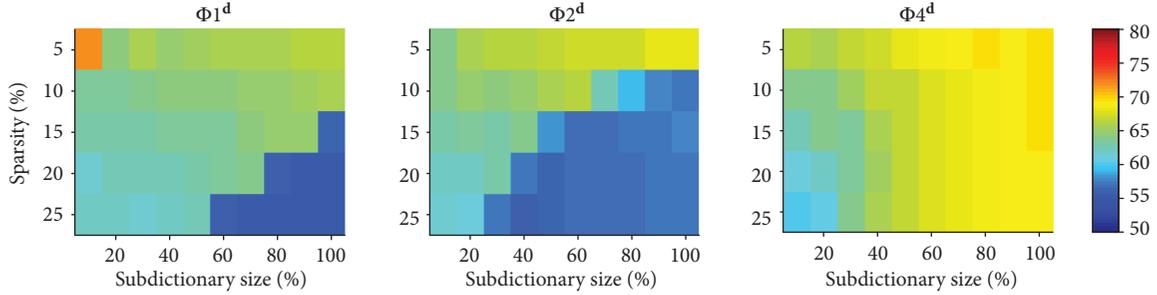


FIGURE 8: Averaged accuracy rates obtained by varying the percentages of the subdictionary size and the sparsity level according to a random ranking of the atoms.

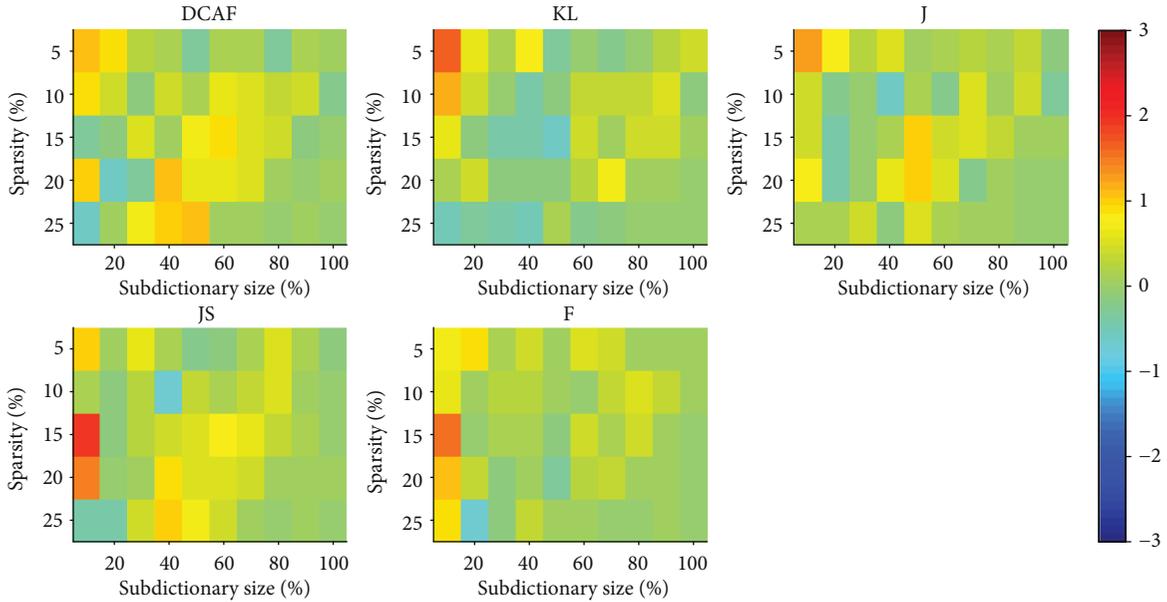


FIGURE 9: Five images representing differences between accuracy rates yielded by DCAF and all other discrepancy measures and random selection for Φ_1 .

clearly note that no highly discriminative atoms were taken when using Fisher score.

Discriminative subdictionaries called Φ_1^d , Φ_2^d , and Φ_4^d were built by stacking side-by-side the first p ranked atoms from Φ_1 , Φ_2 , and Φ_4 , respectively, according to their discriminative degree. It is appropriate to mention that the evaluation of several discrepancy measures leads to the construction of different discriminative subdictionaries. However, optimal values of p (subdictionary size) and q (sparsity level) are parameters that need to be tuned. In order to find optimal values of such hyperparameters, a grid search was performed.

The performance of our system was first tested by performing a Random Selection (RS) of the dictionary atoms. The involved results were fixed and appropriately used as reference. The random selection of the atoms was performed ten times. Additionally, for each one of the atoms' random selection, 60 iterations of the grid search were performed. Thus, the accuracy rate's variations introduced by the classifier were minimized. Figure 8 shows three images corresponding to averaged accuracy rates for each one of the

evaluated dictionaries. Averaged accuracy rates (reference values) obtained by using the dictionary Φ_1 for the detection of apnea-hypopnea events are shown on the left of this figure. It can be seen that sparse representations in terms of Φ_1 , using the smallest subdictionary size and the highest sparsity degree, result in better performance than the ones obtained by using all other configurations of Φ_1 and the overcomplete dictionaries Φ_2 and Φ_4 . In this way, two regions can be distinguished corresponding to a high-performance region and a low-performance one. The first one, which is of our interest, is yielded by simultaneously employing a small subdictionary size (10%) and a high sparsity degree (5%).

Next, DCAF and four other discrepancy measures were used for appropriately constructing discriminative subdictionaries. Then, a grid search of hyperparameters was performed by analyzing the performance that yields our system when using each one of the subdictionaries. Figure 9 shows five images corresponding to DCAF (upper left) and the other four discrepancy measures. These images represent the differences between accuracy rates obtained by using discriminative measures and the reference one (random

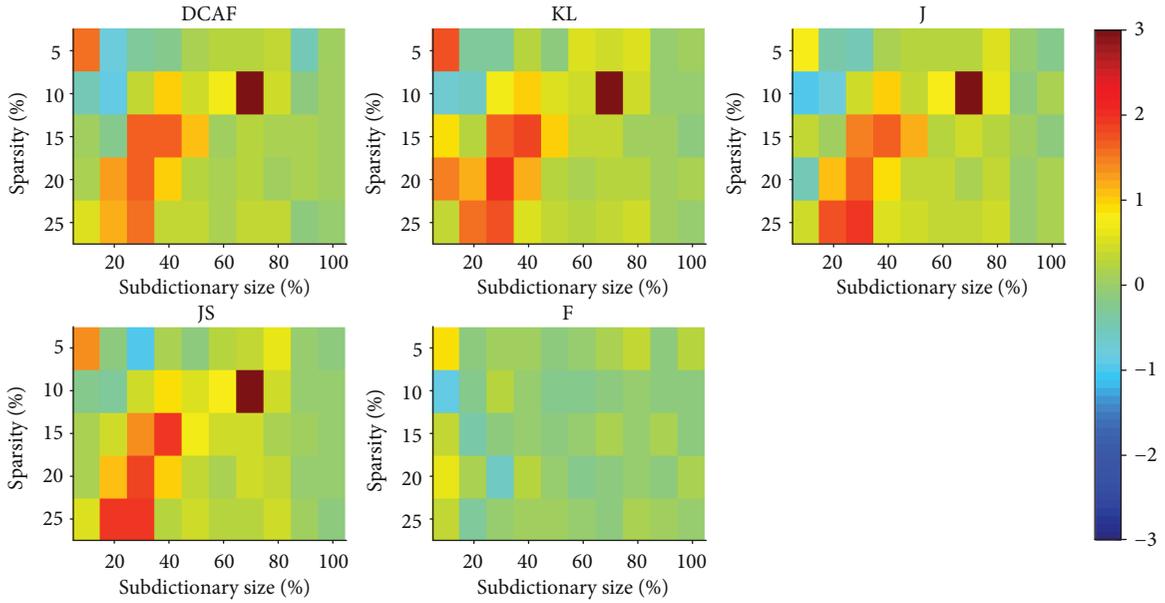


FIGURE 10: Five images representing differences between accuracy rates yielded by DCAF and all other discrepancy measures and random selection for Φ_2 .

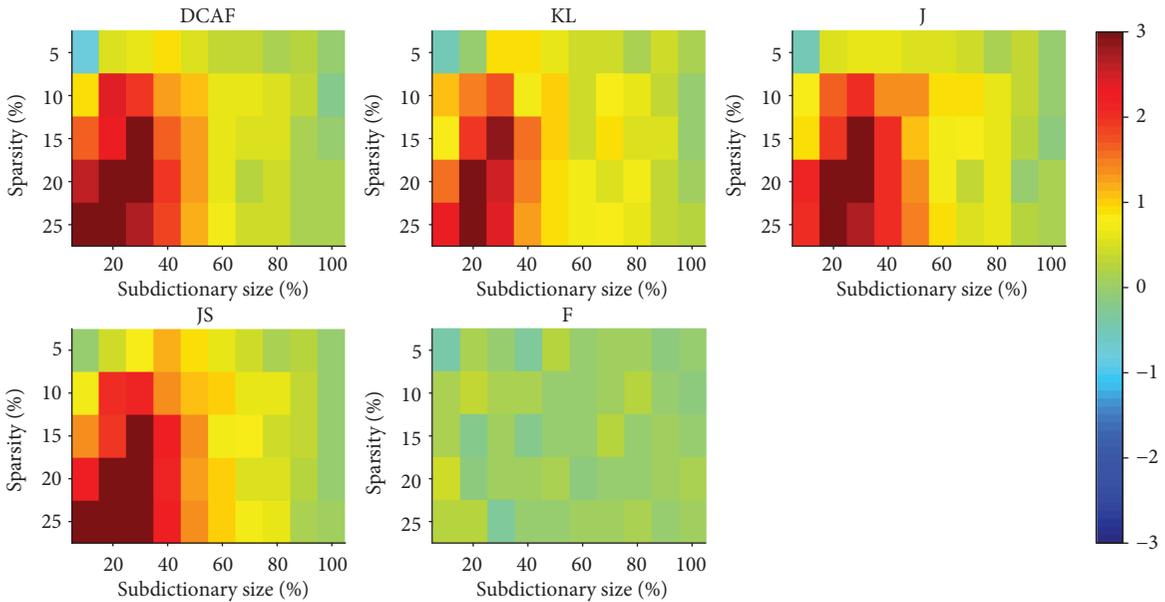


FIGURE 11: Five images representing differences between accuracy rates yielded by DCAF and all other discrepancy measures and random selection for Φ_4 .

selection) for Φ_1 . Also, each pixel of these images corresponds to particular percentages of subdictionary size and sparsity level. It can be observed that, independently of the discriminative measure, small percentages of subdictionary size yield good performances. It is appropriate to point out however that the effect of the dimension (subdictionary size) in the performance of the system is more important than the one induced by using discriminative measures.

Analogously, Figures 10 and 11 show five images which correspond to DCAF (upper left) and all other discrepancy measures. The images depicted in Figures 10 and 11

represent the differences between accuracy rates obtained by using these measures and the reference one for dictionaries Φ_2 and Φ_4 , respectively.

If we compare the results shown in Figures 9–11, then it can be concluded that the proposed system presents the best performance, in terms of accuracy rate in the detection of apnea-hypopnea events, when using the full dictionary Φ_1 . Although similar results were obtained applying the proposed DCAF measure and those traditional ones (see Figure 9), it is important to point out that the use of discrepancy measures resulted in a significantly high

TABLE 1: Averaged accuracy rates for subdictionary sizes of 10% regarding to each one of the evaluated full dictionaries.

Measure	$\Phi 1^d(128 \times 12)$		$\Phi 2^d(128 \times 24)$		$\Phi 4^d(128 \times 50)$	
	Max	Avg.	Max	Avg.	Max	Avg.
DCAF	72.62	64.68	65.20	63.15	65.19	64.21
KL	73.20	64.91	65.44	63.53	65.42	63.66
J	72.82	64.88	64.50	62.82	65.39	63.68
JS	72.55	64.10	65.02	63.18	65.87	64.01
F	72.23	65.21	64.57	63.04	65.64	62.71
Full dictionary	66.39	59.77	68.13	59.57	69.28	69.21

improvement with respect to a “random” selection of the atoms. As discussed above, the dimension reduction in the subdictionary size as well as high sparse levels yielded high accuracy rates. This is the reason for which a small subdictionary size (10%) and high sparse level (5%) were chosen to perform the final test.

System performance changes were analyzed by performing a comparison between averaged accuracy rates obtained by using discriminative subdictionaries and the ones obtained by using full dictionaries. Table 1 shows averaged accuracy percentages obtained by taken into account fixed discriminative subdictionary sizes (10%) while allowing the sparsity level to change (rows from 3 to 7). The last row of this table presents averaged accuracy percentages yielded by using full dictionaries for different sparsity levels. It can be observed that, in all of cases, discriminative subdictionaries outperform full dictionaries in the detection of apnea-hypopnea events.

The impact of sparsity degree in the performance of our system is illustrated in Table 2. These results were yielded by averaging accuracy rates obtained for a sparsity level of 5% and considering all possible subdictionary sizes (from 10% to 90%). For example, the second row shows averaged accuracy rates obtained by means of discriminative subdictionaries whose atoms were taken from $\Phi 1$, $\Phi 2$, and $\Phi 4$ by using DCAF measure.

7.2. Optimal System Performance. Optimal system configurations were selected and fixed to perform the final test. In the previous section, it was found that discriminative subdictionaries constructed by taken atoms from the dictionary $\Phi 1$ yield better performances than the ones constructed by selecting atoms from the dictionaries $\Phi 2$ and $\Phi 4$. Additionally, it was found that a discriminative subdictionary composed by only 12 atoms (10%) and a sparsity level of one (5%) yield in the best accuracy rate of our system.

In order to overcome the variance introduced by ELM predictors, 60 repetitions of the testing process were performed. Table 3 shows percentage values of minimum (Min), maximum (Max), average (μ), and standard deviation (σ) corresponding to obtained accuracy rates in the detection of apnea-hypopnea events. Although, DCAF performs similarly to the four other discrepancy measures, its performance is achieved with a relatively low computational cost. Additionally, results show that performances obtained by using discriminative measures for constructing subdictionaries always outperform the ones yielded by making use of randomly constructed subdictionaries.

TABLE 2: Averaged accuracy rates by considering a sparsity level of 5% regarding to all possible subdictionary sizes.

Measure	$\Phi 1$	$\Phi 2$	$\Phi 4$
DCAF	66.41	66.51	67.95
KL	66.49	66.72	67.98
J	66.60	66.56	67.98
JS	66.41	66.57	68.15
F	66.53	66.54	67.58

TABLE 3: Averaged accuracy rates for a subdictionary percentage of 10 for the detection of apnea-hypopnea events.

Measure	Min	Max	μ	σ
DCAF	71.72	73.14	72.57	0.345
Kullback-Leibler	72.06	73.78	73.26	0.390
Jeffrey	71.77	73.31	72.66	0.319
Jensen-Shannon	71.79	73.11	72.55	0.295
Fisher	71.01	72.77	72.18	0.325
Random Selection	70.01	71.51	70.91	0.372

TABLE 4: A summary of the performed statistical significance tests.

	RS	DCAF	KL	J	JS	F
RS	—	✓	✓	✓	✓	✗
DCAF	—	—	✗	✗	✗	✗
KL	—	—	—	✗	✗	✗
J	—	—	—	—	✗	✗
JS	—	—	—	—	—	✗
F	—	—	—	—	—	—

TABLE 5: Maximum cut-off points for testing accuracy for a subdictionary percentage of 10 for the detection of apnea-hypopnea events.

Measure	d_{\min}	SE	SP	ACC	AUC
DCAF	0.2211	81.88	87.32	84.60	0.9250
Kullback-Leibler	0.2242	81.46	87.39	84.43	0.9271
Jeffrey	0.2311	80.86	87.04	83.95	0.9283
Jensen-Shannon	0.2267	80.75	88.03	84.39	0.9244
Fisher	0.2280	80.66	87.91	84.29	0.9252

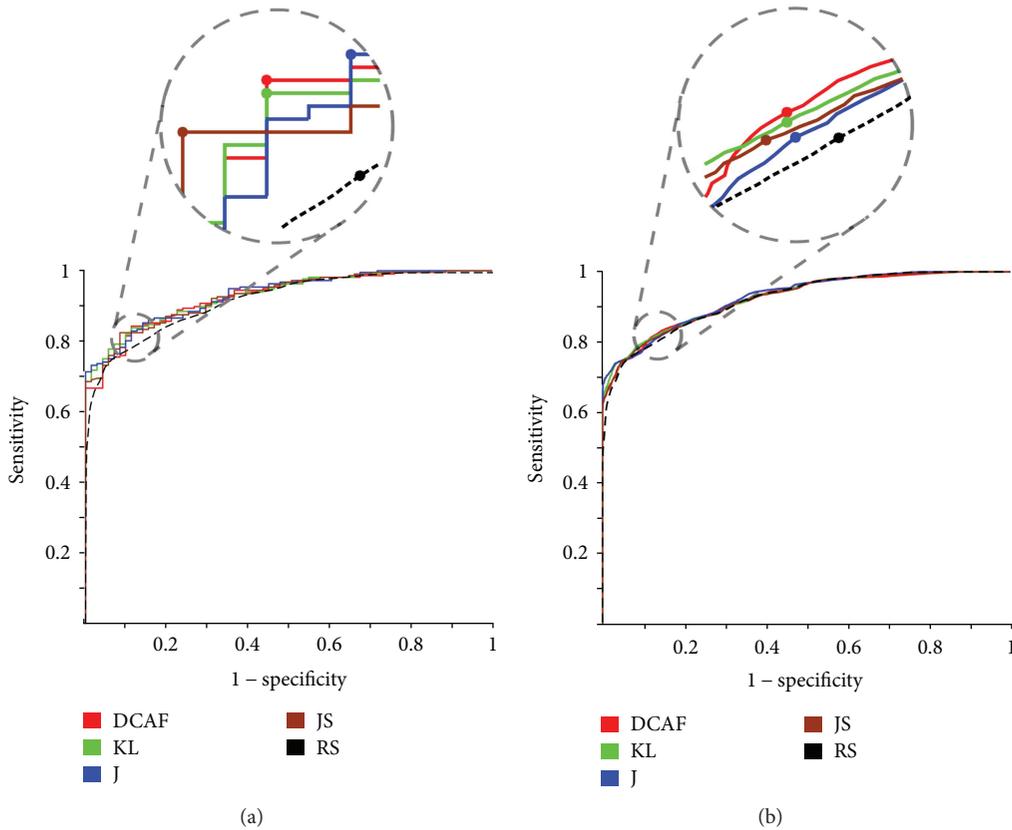


FIGURE 12: ROC curves corresponding to the performance measures described in Tables 5 and 6.

We have also evaluated the statistical significance of the results presented in Table 3 by computing the probability that using each one of the evaluated measures, including RS, yields in better classification performances than the others. In order to perform this test, we assumed the statistical independence of the classification errors for each study. Also, it was possible to approximate the error's binomial probability distribution by a normal distribution due to a wide availability of signals (301,306). Table 4 summarizes the results of the performed statistical significance tests by considering a p value of 0.01. It can be seen that DCAF and three other discrepancy measures (KL, J, and JS divergences) are significantly different with respect to random selection. Also, no significant difference was found between F score and random selection. Additionally, it was found that DCAF does not perform significantly better than that of the KL, J, and JS divergences.

To determine the severity degree of OSAH syndrome, a ROC curve analysis was successfully performed by considering a detection AHI of 15. This index was selected in order to identify patients suspected of suffering from moderate to severe OSAH syndrome. Table 5 shows the minimum operating (cut-off) point of the ROC curves and maximum percentages of sensitivity, specificity, and accuracy as well as maximum values of area under the ROC curve for AHI diagnostic threshold values of 15 (Figure 12(a)). It can be seen that DCAF resulted in a maximum area under the ROC curve of 0.9250 and sensitivity and

TABLE 6: Averaged cut-off points for testing accuracy for a subdictionary percentage of 10 for the detection of apnea-hypopnea events.

Measure	d_{\min}	SE	SP	ACC	AUC
DCAF	0.2211	81.88	87.32	84.60	0.9250
Kullback-Leibler	0.2242	81.46	87.39	84.43	0.9271
Jeffrey	0.2311	80.86	87.04	83.95	0.9283
Jensen-Shannon	0.2267	80.75	88.03	84.39	0.9244
Fisher	0.2280	80.66	87.91	84.29	0.9252
Random Selection	0.2396	80.85	85.60	83.23	0.9222

specificity percentages of 81.88 and 87.32, respectively. These are the maximum performance measures at which the minimum cut-off point of the ROC curve is attained. If we compare the performances attained between all of the evaluated measures, then the maximum SE and AUC value is yielded by J divergence. Also, JS divergence outperformed all the others in terms of ACC and DCAF resulted in the minimum cut-off point of the ROC curve.

We additionally performed a ROC curve analysis of the averaged performances of DCAF and all the other discrepancy measures, including (RS) (Figure 12(b)). Additionally, Table 6 shows the minimum operating (cut-off) point of the averaged ROC curves as well as the maximum percentages of sensitivity, specificity, and accuracy, including the

maximum values of AUC for the same OSAH syndrome diagnostic threshold. The results show that DCAF outperforms all the other discrepancy measures in terms of minimum optimal operating cut-off point of the ROC curve as well as in terms of sensitivity and accuracy rate. Also, KL divergence resulted in the best averaged area under the curve ROC and the maximum averaged specificity was yielded by JS divergence. A significant performance improvement was observed when using DCAF or any of the other discrepancy measures compared to random selection.

Several applications exist where it is desirable to maximize the sensitivity. For instance, if the primary purpose of the test is “screening,” that is, detection of early disease in a large numbers of apparently healthy persons, then a high sensitivity is generally desired. With this in mind, if a sensitivity of 98% is chosen in the ROC curves in Figure 12, for all used measures, the method achieves a specificity close to 45%. This fact shows that the analysis of pulse oximetry signals by means of the proposed method could be potentially applied as an efficient diagnostic screening tool in clinical practice.

In a previous work [4], it was shown that the MDCS method using DCAF to select discriminative atoms in a given dictionary provides good accuracy rates in the detection of apnea-hypopnea events. In that work, a comparative analysis of the performances yielded by MDCS and other methods [45–47] has shown that MDCS outperforms all the others. It was also observed that the computational cost of MDCS is slightly higher than those required by the other three methods. On the other hand, in this work, we show that MDCS using DCAF for selecting discriminative atoms performs similarly than MDCS using several other traditional discrepancy measures. It is important to highlight that DCAF is very easy to compute and yields competitive performance rates in the detection of apnea-hypopnea events at a low computational cost.

8. Conclusions

Sparse representations of signals constitute a powerful technique which yields high accuracy rates in the detection of apnea-hypopnea events. In this work, the difference of conditional activation frequency (DCAF) measure was successfully used for accurately pointing out discriminative atoms in a full dictionary. Additionally, we compared the performance of the DCAF with four widely used discrepancy measures. It was found that the DCAF and three other discrepancy measures (KL, J, and JS divergences) outperform the random selection of atoms, unlike F score. Additionally, DCAF is cheaper to compute. Discriminative subdictionaries were successfully constructed by taking the best ranked atoms of full dictionaries according to their discriminative power. Results show that sparse representations of signals in terms of discriminative subdictionaries result in better performances than the ones obtained in terms of full dictionaries in the detection of apnea-hypopnea events by using only pulse oximetry signals. In this context, it was found that more sparse solutions almost always yielded in better performances. Additionally, it was observed that larger dictionary overcompleteness worsens the performance of the

system. Future research lines include more analysis of the DCAF measure, the study of its properties, and an extension of such a measure to multiclass problems.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET, through PIP 2014–2016 no. 11220130100216-CO and PIP 2012–2014 no. 114 20110100284-KA4, by the Air Force Office of Scientific Research, AFOSR/SOARD, through Grant no. FA9550-14-1-0130, by the Universidad Nacional del Litoral through projects CAI+D PIC no. 504 201501 00098 LI (2016) and PIC no. 504 201501 00036 LI (2016), and by the Asociación Gremial de Docentes of the Universidad Tecnológica Nacional (FAGDUT), Paraná section. The authors would like to thank Dr. Luis D. Larrateguy, who is a specialist in sleep-related disorders, for his valuable comments and suggestions.

References

- [1] N. Saito and R. R. Coifman, “Local discriminant bases and their applications,” *Journal of Mathematical Imaging and Vision*, vol. 5, no. 4, pp. 337–358, 1995.
- [2] S. Tabibian, A. Akbari, and B. NaserSharif, “Speech enhancement using a wavelet thresholding method based on symmetric Kullback–Leibler divergence,” *Signal Processing*, vol. 106, pp. 184–197, 2015.
- [3] M. Sánchez-Gutiérrez, E. M. Albornoz, H. L. Rufiner, and J. G. Close, “Post-training discriminative pruning for RBMs,” *Soft Computing*, pp. 1–15, 2017.
- [4] R. E. Rolón, L. D. Larrateguy, L. E. Di Persia, R. D. Spies, and H. L. Rufiner, “Discriminative methods based on sparse representations of pulse oximetry signals for sleep apnea–hypopnea detection,” *Biomedical Signal Processing and Control*, vol. 33, pp. 358–367, 2017.
- [5] V. Peterson, H. L. Rufiner, and R. D. Spies, “Generalized sparse discriminant analysis for event-related potential classification,” *Biomedical Signal Processing and Control*, vol. 35, pp. 70–78, 2017.
- [6] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [7] S. Kullback and R. A. Leibler, “On information and sufficiency,” *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.
- [8] A. Gupta, S. Parameswaran, and C.-H. Lee, “Classification of electroencephalography (EEG) signals for different mental activities using Kullback Leibler (KL) divergence,” in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1697–1700, Taipei, Taiwan, April 2009.

- [9] H. Jeffreys, "An invariant form for the prior probability in estimation problems," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 186, no. 1007, pp. 453–461, 1946.
- [10] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 145–151, 2006.
- [11] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.
- [12] X. Zhang and Q. Ding, "Respiratory rate estimation from the photoplethysmogram via joint sparse signal reconstruction and spectra fusion," *Biomedical Signal Processing and Control*, vol. 35, pp. 1–7, 2017.
- [13] Y. Zhou, X. Hu, Z. Tang, and A. C. Ahn, "Sparse representation-based ECG signal enhancement and QRS detection," *Physiological Measurement*, vol. 37, no. 12, pp. 2093–2110, 2016.
- [14] M. Aharon, M. Elad, and A. Bruckstein, "rmK-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [15] D. S. Pham and S. Venkatesh, "Joint learning and dictionary construction for pattern recognition," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, Anchorage, AK, USA, June 2008.
- [16] Q. Zhang and B. Li, "Discriminative K-SVD for dictionary learning in face recognition," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern*, pp. 2691–2698, San Francisco, CA, USA, June 2010.
- [17] M. J. Sateia, "International classification of sleep disorders-third edition: highlights and modifications," *Chest*, vol. 146, no. 5, pp. 1387–1394, 2014.
- [18] N. A. Dewan, F. J. Nieto, and V. K. Somers, "Intermittent hypoxemia and OSA: implications for comorbidities," *Chest*, vol. 147, no. 1, pp. 266–274, 2015.
- [19] W. Kukwa, E. Migacz, K. Druc, E. Grzesiuk, and A. M. Czarnecka, "Obstructive sleep apnea and cancer: effects of intermittent hypoxia?," *Future Oncology*, vol. 11, no. 24, pp. 3285–3298, 2015.
- [20] M. Torres, R. Laguna-Barraza, M. Dalmases et al., "Male fertility is reduced by chronic intermittent hypoxia mimicking sleep apnea in mice," *Sleep*, vol. 37, no. 11, pp. 1757–1765, 2014.
- [21] T. Young, L. Evans, L. Finn, and M. Palta, "Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women," *Sleep*, vol. 20, no. 9, pp. 705–706, 1997.
- [22] J. Durán, S. Esnaola, R. Rubio, and A. Izutueta, "Obstructive sleep apnea-hypopnea and related clinical features in a population-based sample of subjects aged 30 to 70 yr," *American Journal of Respiratory and Critical Care Medicine*, vol. 163, no. 3, pp. 685–689, 2001.
- [23] R. Thurnheer, K. E. Bloch, I. Laube, M. Gugger, M. Heitz, and Swiss respiratory Polygraphy registry, "Respiratory polygraphy in sleep apnoea diagnosis. Report of the Swiss respiratory polygraphy registry and systematic review of the literature," *Swiss Medical Weekly*, vol. 137, no. 5-6, pp. 97–102, 2007.
- [24] E. García-Díaz, E. Quintana-Gallego, A. Ruiz et al., "Respiratory polygraphy with actigraphy in the diagnosis of sleep apnea-hypopnea syndrome," *Chest*, vol. 131, no. 3, pp. 725–732, 2007.
- [25] A. Yadollahi, E. Giannouli, and Z. Moussavi, "Sleep apnea monitoring and diagnosis based on pulse oximetry and tracheal sound signals," *Medical & Biological Engineering & Computing*, vol. 48, no. 11, pp. 1087–1097, 2010.
- [26] M. Elad, *Sparse and Redundant Representations*, Springer, New York, NY, USA, 2010.
- [27] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [28] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [29] R. R. Coifman, Y. Meyer, S. Quake, and M. V. Wickerhauser, "Signal processing and compression with wavelet packets," in *Wavelets and Their Applications. NATO ASI Series (Series C: Mathematical and Physical Sciences)*, J. S. Byrnes, J. L. Byrnes, K. A. Hargreaves, and K. Berry, Eds., vol. 442, pp. 363–379, Springer, Dordrecht, 1994.
- [30] M. S. Lewicki and B. A. Olshausen, "Probabilistic framework for the adaptation and comparison of image codes," *Journal of the Optical Society of America A*, vol. 16, no. 7, p. 1587, 1999.
- [31] M. S. Lewicki and T. J. Sejnowski, "Learning overcomplete representations," *Neural Computation*, vol. 12, no. 2, pp. 337–365, 2000.
- [32] K. Engan, S. O. Aase, and J. H. Husoy, "Method of optimal directions for frame design," in *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258)*, vol. 5, pp. 2443–2446, Phoenix, AZ, USA, March 1999.
- [33] Z. Jiang, Z. Lin, and L. Davis, "Label consistent K-SVD: learning a discriminative dictionary for recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2651–2664, 2013.
- [34] M. Basseville, "Distance measures for signal processing and pattern recognition," *Signal Processing*, vol. 18, no. 4, pp. 349–369, 1989.
- [35] W. Gersch, F. Martinelli, J. Yonemoto, M. D. Low, and J. A. McEwan, "Automatic classification of electroencephalograms: Kullback-Leibler nearest neighbor rules," *Science*, vol. 205, no. 4402, pp. 193–195, 1979.
- [36] P. J. Moreno, P. P. Ho, and N. Vasconcelos, "A Kullback-Leibler divergence based kernel for SVM classification in multimedia applications," in *Advances in Neural Information Processing Systems*, S. Thrun, L. K. Saul, and P. B. Schölkopf, Eds., vol. 16, pp. 1385–1392, MIT Press, 2004.
- [37] C. C. Aggarwal, *Data Classification: Algorithms and Applications*, CRC Press, 2014.
- [38] S. F. Quan, B. V. Howard, C. Iber et al., "The Sleep Heart Health Study: design, rationale and methods," *Sleep*, vol. 20, no. 12, pp. 1077–1085, 1997.
- [39] B. K. Lind, J. L. Goodwin, J. G. Hill, T. Ali, S. Redline, and S. F. Quan, "Recruitment of healthy adults into a study of overnight sleep monitoring in the home: experience of the Sleep Heart Health Study," *Sleep and Breathing*, vol. 7, no. 1, pp. 13–24, 2003.
- [40] F. Lestussi, L. Di Persia, and D. Milone, "Comparison of on-line wavelet analysis and reconstruction: with application to ECG," in *2011 5th International Conference on Bioinformatics*

and Biomedical Engineering, pp. 1–4, Wuhan, China, May 2011.

- [41] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: theory and applications,” *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.
- [42] J. Tang, C. Deng, and G. B. Huang, “Extreme learning machine for multilayer perceptron,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 4, pp. 809–821, 2016.
- [43] J. A. Swets, “ROC analysis applied to the evaluation of medical imaging techniques,” *Investigative Radiology*, vol. 14, no. 2, pp. 109–121, 1979.
- [44] K. Hajian-Tilaki, “Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation,” *Caspian Journal of Internal Medicine*, vol. 4, no. 2, pp. 627–635, 2013.
- [45] E. Chiner, J. Signes-Costa, J. M. Arriero, J. Marco, I. Fuentes, and A. Sergado, “Nocturnal oximetry for the diagnosis of the sleep apnoea hypopnoea syndrome: a method to reduce the number of polysomnographies?,” *Thorax*, vol. 54, no. 11, pp. 968–971, 1999.
- [46] J. C. Vázquez, W. H. Tsai, W. W. Flemons et al., “Automated analysis of digital oximetry in the diagnosis of obstructive sleep apnoea,” *Thorax*, vol. 55, no. 4, pp. 302–307, 2000.
- [47] G. Schlotthauer, L. E. Di Persia, L. D. Larrateguy, and D. H. Milone, “Screening of obstructive sleep apnea with empirical mode decomposition of pulse oximetry,” *Medical Engineering and Physics*, vol. 36, no. 8, pp. 1074–1080, 2014.

Research Article

Linear and Complex Measures of Heart Rate Variability during Exposure to Traffic Noise in Healthy Women

Myrela Alves,¹ David M. Garner ,² Anne M. G. G. Fontes,¹
Luiz Vinicius de Alcantara Sousa ,³ and Vitor E. Valenti ¹

¹Centro de Estudos do Sistema Nervoso Autônomo, Departamento de Fonoaudiologia, Faculdade de Filosofia e Ciências, UNESP, Marília, SP, Brazil

²Cardiorespiratory Research Group, Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brookes University, Headington, Oxford OX3 0BP, UK

³Laboratório de Epidemiologia e Análises de Dados, Faculdade de Medicina do ABC (FMABC), Santo Andre, SP, Brazil

Correspondence should be addressed to Vitor E. Valenti; vitor.valenti@marilia.unesp.br

Received 1 October 2017; Revised 7 February 2018; Accepted 20 February 2018; Published 26 June 2018

Academic Editor: Gastón Schlotthauer

Copyright © 2018 Myrela Alves et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Previous studies have described significant impact of different types of noise on the linear behavior of heart rate variability (HRV). However, there are few studies regarding the complexity of HRV during exposure to traffic noise. In this study, we evaluated the complexity of HRV during traffic noise exposure. We analyzed 31 healthy female students aged between 18 and 30 years. Volunteers remained at rest seated under spontaneous breathing during 10 minutes with an earphone turned off, and then they were exposed to traffic noise through an earphone for a period of 10 minutes. The traffic noise was recorded from a very busy city street and the sound was comprised of car, bus, and trucks engines and horn (71–104 dB). We observed no significant changes in the linear analysis of HRV. CFP3 (Cohen's $d = 1.28$, large effect size) and CFP6 (Cohen's $d = 1.11$, large effect size) parameters of chaotic global analysis and Shannon (Cohen's $d = 1.13$, large effect size), Renyi (Cohen's $d = 1.06$, large effect size), and Tsallis (Cohen's $d = 1.14$, large effect size) entropies significantly increased ($p < 0.005$) during traffic noise exposure. In conclusion, traffic noise under laboratory conditions increased the complexity of HRV through chaotic global analysis and some measures of entropy in healthy females.

1. Introduction

Noise may be considered an unpleasant sound, which may have effects on physiological variables. It is often found in hazardous situations due to industrialization and urbanization [1]. In this way, the research literature has previously investigated the effects of different types of noise on autonomic nervous system by analyzing heart rate variability (HRV) [2]. Lee et al. [3] noted that white noise above 50 dB influences spectral analysis of HRV, indicating significant correlation between frequency domain analysis and sound pressure level. Umemura and Honda [4] restated that this type of noise also encourages deviations in HRV. Yet, until now the research literature has only focused on traditional linear indices of HRV analysis [2, 4, 5].

The linear analysis of HRV in the time and frequency domains is not entirely suitable to provide information about the complex dynamics of heartbeat origination. This is because the mechanisms involved in cardiovascular physiology interact with each other in a nonlinear way [6]. Furthermore, methods related to nonlinear behavior of HRV were reported to present clinical relevance and to offer improved interpretation about these pathological mechanisms [7, 8].

Most recently, the European Society of Cardiology together with the European Heart Rhythm Association and coendorsed by the Asia Pacific Heart Rhythm Society drew attention to nonlinear methods for assessing HRV [9]. In this review, the authors address entropy and regularity, long-range correlation and fractal analysis, short-term complexity, nonlinear dynamical systems, and chaotic behavior generally.

Nevertheless, there is little in the research literature comparing HRV analysis with chaotic global analysis and Shannon, Renyi, and Tsallis entropies (see later section on nonlinear analysis).

This information related to chaos theory, fractal mathematics, and the dynamic complexity of HRV has not yet been fully applied in medical practice clinically. Yet, it is a productive area for research and development of knowledge in both health and disease [10]. Besides, the complex measurement of the intervals between consecutive heart beats (RR intervals) analysis during exposure to traffic noise has not been studied. Studies analyzing HRV and traffic noise exposed subjects to real traffic, which exposed subjects to multiple stimuli (visual, conversation, temperature, and humidity) that have a significant impact on the autonomic nervous system. Sensitive techniques to identify autonomic changes are necessary to avert possible physiological injury in the organism. Consequently, we aimed to evaluate the acute effects of traffic noise on the complexity of HRV under laboratory conditions alone.

2. Method

2.1. Study Population. We examined 31 apparently healthy female students aged between 18 and 30 years. All volunteers were informed about the procedures and objectives of the study and, after agreeing, signed a confidential consent form. All study procedures were approved by the Research Ethics Committee (REC) of the institution (case number 2011/382) and followed the Resolution 196/96 of the National Health Council. We excluded women under the following conditions: body mass index (BMI) > 30 kg/m², systolic blood pressure (SBP) > 140 mmHg or diastolic blood pressure (DBP) > 90 mmHg (at rest), and endocrine, cardiovascular, respiratory, and neurological related disorders or any condition that prevented the subject from performing the study. In order to avoid effects related to sexual hormones, we did not include women on the 11th to 15th and 21st to 25th days after the first day of the menstrual cycle [11].

2.2. Initial Assessment. The subjects were identified by collecting the following information: age, mass, height, and body mass index (BMI). Mass was measured using a digital scale (W200/5, Welmy, Brazil) with a precision of 0.1 kg. Height was determined using a stadiometer (ES2020, Sanny, Brazil) with a precision of 0.1 cm and being 220 cm long. The body mass index (BMI) was calculated by the subsequent formula: mass (kg)/height (m²). We measured heart rate and blood pressure. Heart rate was measured with the Polar RS800CX heart rate monitor (Polar Electro, Finland). Blood pressure was indirectly measured by auscultation through calibrated aneroid sphygmomanometer (Welch Allyn, New York, USA) and stethoscope (Littmann, St. Paul, USA) with all subjects seated.

2.3. Measurement of Auditory Stimulation. The measurements of equivalent sound levels were performed in a sound-proofed room, using an audio dosimeter SV 102 (Svantek,

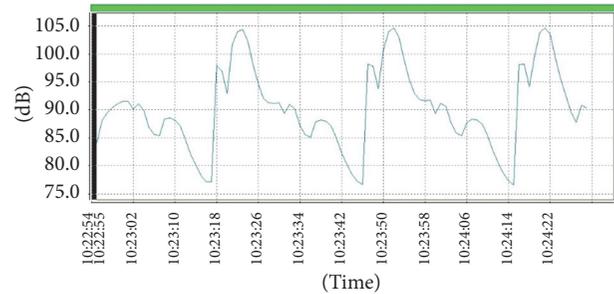


FIGURE 1: Equivalent sound level.

Finland). It was programmed in measuring circuit 7 in “A” weighting, slow response [12].

We used the MIRE earphone, which was placed inside the auditory canal of the subject and linked to a personal stereo. Prior to each measurement, the earphones were calibrated with the acoustic calibrator CR: Model 514 (Cirrus Research plc).

This tool was used to analyze the Leq (A), which is defined as the equivalent sound pressure level, and the sound level corresponds to the same constant time interval. It contained the same total sound energy, which also analyzed the spectrum of sound stimulation (eighth track) frequency [13] of traffic noise (71–104 dB) (Figure 1).

2.4. Experimental Protocol. Data collection was commenced at room temperature between 21°C and 25°C and with humidity between 50% and 60%. The subjects were instructed not to ingest alcohol or caffeine for 24 hours prior to evaluation. The data collection was achieved individually between 18:00 and 21:00 to avoid circadian influences. The volunteers were instructed to remain at rest and avoid conversation during the experiment.

After the initial evaluation, the heart monitor belt was placed over the thorax, aligned with the distal third of the sternum and the Polar RS800CX heart rate receiver (Polar Electro, Finland) was placed on the wrist. Subsequently, the volunteers remained at rest seated for 10 minutes with the headset off.

Next, the volunteers were exposed to traffic noise through an earphone for a period of 10 minutes. The traffic noise was recorded from a very busy street in Marília city, SP, Brazil. The sounds were produced by cars, buses, trucks engines, and horns.

2.5. Analysis of HRV. The RR intervals were recorded by the Polar RS800CX heart rate monitor with a sampling rate of 1000 Hz. They were then transferred to the Polar Precision Performance software (v. 3.0, Polar Electro, Finland). This software allowed the visualization of the HR and the extraction of a file relating to a cardiac period (RR-interval) in a “txt” file. After digital filtering supplemented with manual filtering to eliminate artefacts and premature ectopic beats, 500 RR intervals were applied for data analysis. Only series with more than 95% of sinus beats were included in the study. HRV was analyzed before and during traffic noise.

2.6. Linear Analysis of HRV. The time domain analysis was accomplished in terms of SDNN (standard deviation of normal-to-normal RR intervals), pNN50 (percentage of adjacent RR intervals with a difference of duration greater than 50 milliseconds), and RMSSD (root-mean square of differences between adjacent normal RR intervals in a time interval) [14].

To obtain the spectral indexes for HRV analysis in the frequency domain, the frequency recordings underwent mathematical processing, thus generating a tachogram that expressed the variation of RR intervals as a function of time. The tachogram contained a signal that varied with time and was processed by the mathematical Fast Fourier Transform (FFT) algorithm. Welch's periodogram method based on FFT using a window width of 256 seconds and an overlap of 50% was applied.

Low frequency (LF, ranging between 0.04 and 0.15 Hz) and high frequency (HF, ranging from 0.15 to 0.4 Hz) spectral components were selected in normalized units (nu). The ratio between these components in absolute values (LF/HF) represents the relative value of each spectral component in relation to the total potential minus the very low frequency (VLF) components. It is important to mention that the LF/HF index may provide significant information on autonomic regulation of sinus node under controlled conditions and short-term recordings [14].

For computation of the linear indices, we applied the HRV analysis software (Kubios HRV v.1.1 for Windows, Biomedical Signal Analysis Group, Department of Applied Physics, University of Kuopio, Finland).

2.7. Statistical Analysis of Linear Indices. Statistical methods of the linear indices were approved for the computation of means and standard deviations. Normal Gaussian distribution of the data was verified by the Shapiro-Wilk goodness-of-fit test (z value > 1.0).

To enable a comparison of the variables between control and traffic noise exposure, we applied the unpaired Student t -test for parametric distribution and the Mann-Whitney test for nonparametric distributions. Level of significance was set at $p < 0.005, 0.5\%$.

2.8. Nonlinear Analysis

2.8.1. Detrended Fluctuation Analysis (DFA). Detrended fluctuation analysis (DFA) [15] may be applied to datasets where parameters such as mean, variance, and autocorrelation vary with time. DFA computes the correlation *within* the signal. It quantifies how the fluctuations of a signal scale with the number of samples of that signal. According to Donaldson et al. [16], the time series of length k was manipulated as shown:

$$y(k) = \sum_{i=1}^k (\text{RR}(i) - \text{mean}(\text{RR})). \quad (1)$$

The integrated time series was then divided into equally sized and nonoverlapping windows of length w . A linear regression line was fitted through the data in each window and the time series manipulated by subtracting the regression line from the data.

The root-mean square fluctuation $F(w)$ of the integrated and detrended time series was calculated for different values of w , as follows:

$$F(w) = \left[\frac{1}{N} \sum_{k=1}^N [y(k) - y_w(k)]^2 \right]^{1/2}. \quad (2)$$

The scaling exponent (α) was obtained as the slope of a straight line fit to $F(w)$ against w on a log-log plot:

$$F(w) \propto w^\alpha. \quad (3)$$

DFA is a technique extensively imposed in variability analysis. It has been applied to the evaluation of posture [17], exercise [18] and sleep stage classification [19], and classification of asthma [20] and COPD [16, 21, 22].

2.8.2. Chaotic Global Analysis. Multitaper Method (MTM) [23] is useful for spectral estimation and signal reconstruction, of a time series of a spectrum that may contain broadband and line components. MTM lessens the variances of spectral estimates by using a small set of tapers (windows). Data is premultiplied by orthogonal tapers created to minimize the spectral leakage owing to the finite length of the time series. A set of approximations of the power spectrum are calculated. These functions identified as Discrete Prolate Spheroidal Sequences (DPSS) sometimes called Slepian Sequences [24] are a set of functions which optimize these tapers. They are defined as eigenvectors of a Rayleigh-Ritz minimization problem [25].

2.8.3. High Spectral Entropy. High spectral entropy (*hsEntropy*) [26] is a function of the irregularity of amplitude and frequency of the power spectra peaks. It is derived by applying Shannon entropy to the MTM power spectrum (see Figure 2). Then, we calculate an intermediate parameter which is the median Shannon entropy of the value obtained from three different power spectra using the MTM power spectra under three test conditions: (a) a perfect sine wave, (b) uniformly distributed random variables, and finally (c) the experimental oscillating signal. These values are normalized mathematically so that the sine wave gives a value of zero, uniformly random variables give unity, and the experimental signal gives values between zero and unity. It is the final value that corresponds to *hsEntropy*.

2.8.4. High Spectral DFA. As stated before, the DFA [26] algorithm can be applied to datasets where statistics such as mean, variance, and autocorrelation vary with time. The *high spectral* detrended fluctuation analysis (*hsDFA*) algorithm is where the DFA is applied to the frequency rather than time on the horizontal axis (Figure 2). So, the x -axis is frequency and the y -axis is amplitude. To obtain *hsDFA*, we calculate the spectral adaptation in exactly the same manner as for *hsEntropy* applying a MTM power spectrum with the same settings, but DFA rather than Shannon entropy is the algorithm enforced.

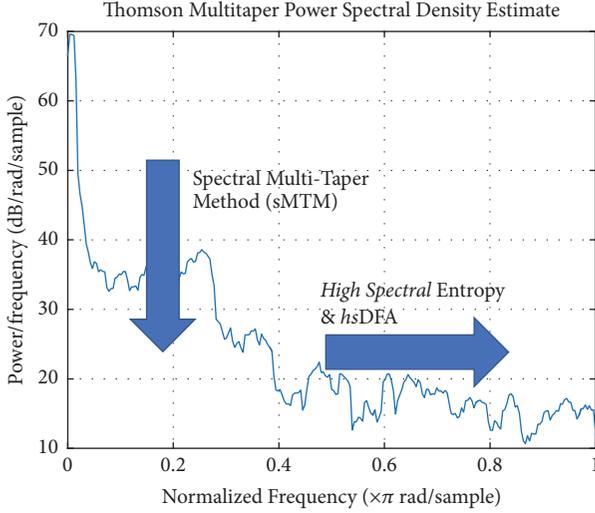


FIGURE 2: A MTM power spectrum of a time series of 500 RR intervals in a traffic noise exposure subject. sMTM is the area beneath the spectrum, yet above the baseline created by broadband noise as the signal becomes chaotic. *High spectral entropy* and *high spectral DFA* are derived by applying the Shannon entropy and detrended fluctuation analysis (DFA) functions to the MTM power spectrum. Parameters for the MTM power spectra as set at (i) sampling frequency of 1 Hz, (ii) time bandwidth for the DPSS at 3, (iii) FFT length of 256, and (iv) Thomson's "adaptive" nonlinear combination method to combine individual spectral estimates.

2.8.5. Spectral Multitaper Method. Spectral Multitaper Method (sMTM) [27] is founded on the increased intensity of broadband noise in power spectra generated by irregular and chaotic signals. sMTM is the area between the MTM power spectrum and the baseline (see Figure 2).

2.8.6. Chaotic Forward Parameters (CFP 1 to CFP7). The parameters (CFP 1–7) are referred to as chaotic forward parameters (CFP) for the functions CFP1 to CFP7 below where they are applied to normal and traffic noise exposure subjects' RR-interval time series. Since *hsDFA* responds to chaos inversely to the others, we subtract its value from unity. In this study, all three chaotic global values have weightings of unity.

$$\begin{aligned} \text{CFP1} &= \left[\left(\left[\frac{hsEntropy}{\max(hsEntropy)} \right] \right)^2 \right. \\ &\quad + \left(\left[\frac{sMTM}{\max(sMTM)} \right] \right)^2 \\ &\quad \left. + \left(1 - \left[\frac{hsDFA}{\max(hsDFA)} \right] \right)^2 \right]^{1/2} \\ \text{CFP2} &= \left[\left(\left[\frac{hsEntropy}{\max(hsEntropy)} \right] \right)^2 \right. \\ &\quad \left. + \left(1 - \left[\frac{hsDFA}{\max(hsDFA)} \right] \right)^2 \right]^{1/2} \end{aligned}$$

$$\begin{aligned} \text{CFP3} &= \left[\left(\left[\frac{hsEntropy}{\max(hsEntropy)} \right] \right)^2 \right. \\ &\quad \left. + \left(\left[\frac{sMTM}{\max(sMTM)} \right] \right)^2 \right]^{1/2} \\ \text{CFP4} &= \left[\left(\left[\frac{sMTM}{\max(sMTM)} \right] \right)^2 \right. \\ &\quad \left. + \left(1 - \left[\frac{hsDFA}{\max(hsDFA)} \right] \right)^2 \right]^{1/2} \\ \text{CFP5} &= \left| \left(1 - \left[\frac{hsDFA}{\max(hsDFA)} \right] \right) \right| \\ \text{CFP6} &= \left| \left(\left[\frac{sMTM}{\max(sMTM)} \right] \right) \right| \\ \text{CFP7} &= \left| \left(\left[\frac{hsEntropy}{\max(hsEntropy)} \right] \right) \right|. \end{aligned} \tag{4}$$

2.8.7. Shannon Entropy. Shannon entropy [28] is represented by the degree of ambiguity associated with the occurrence of the result. A higher value of entropy gives a more uncertain outcome and is more difficult to predict.

Shannon entropy may be used globally, applying to the time series wholly or nearby around specific points. This measure can provide extra evidence about specific events such as outliers or intermittent events. In contrast to Tsallis [29] and Renyi [30] entropies, Shannon entropy is additive. Hence, if the probabilities can be factorised into independent factors, the entropy of the joint process is the sum of the entropies of the distinct processes.

2.8.8. Renyi Entropy. Renyi entropy is a general statement of Shannon entropy that is dependent on a specified parameter. Renyi entropy depends on the entropic order α (which we set to 0.25). Renyi entropy approaches Shannon entropy as $\alpha \rightarrow 1$ which can be derived by l'Hôpital's rule [31, 32]. As entropic order increases, the procedures become more sensitive to the values occurring at higher probabilities and less sensitive to those of lower probabilities. Renyi entropy is described fully in studies by Zyczkowski [33] and Lenzi et al. [30].

2.8.9. Tsallis Entropy. Tsallis entropy is a general statement of the standard Shannon-Boltzmann-Gibbs entropy. It was introduced in the application of statistical mechanics and is used in computer sciences for pattern recognition. Tsallis entropy is dependent on the specified parameter termed entropic index q (which we set to 0.25); Tsallis entropy becomes the Shannon-Boltzmann-Gibbs entropy, as the entropic index $q \rightarrow 1$. Tsallis entropy is discussed further in the publications by dos Santos [29], A. R. Plastino and A. Plastino [34], and Mariz [35].

2.8.10. Approximate Entropy. Approximate Entropy (ApEn) was discussed by Pincus [36]. It is a procedure required

to evaluate the level of uniformity and the unpredictability of changes over time series. ApEn is the logarithmic ratio of component-wise matching sequences from the signal length, N . Other parameters include r , tolerance, and m , the embedding dimension. Here we set the parameters of m to 2 and r to 20% of the standard deviation of the data. The disadvantages of ApEn are that it is very dependent on the length of the time series and is often lower than expected on shorter time series. Finally, it is disadvantageous because it lacks “relative consistency” [37].

A minimum value of zero for ApEn would indicate a totally predictable time series, while a maximum value of one would specify an entirely unpredictable time series. Most of the time, the values are between these two values.

ApEn is mathematically described as in the Kubios HRV Analysis Manual [38].

First a set of length m vectors u_j is formed; note the embedding dimension, m , and N , the number of RR intervals.

$$u_j = (RR_j, RR_{j+1}, \dots, RR_{j+m-1}), \quad (5)$$

$$j = 1, 2, \dots, N - m + 1.$$

The distance between these vectors is the maximum absolute difference between the corresponding elements; hence,

$$d(u_j, u_k) = \max \{ |RR_{j+n} - RR_{k+n}| \mid n = 0, \dots, m-1 \}. \quad (6)$$

Next for each u_j the relative number of vectors u_k for which $d(u_j, u_k) \leq r$ is calculated. This index is denoted with $C_j^m(r)$ and can be written in the form

$$C_j^m(r) = \frac{\text{number of } \{u_k \mid d(u_j, u_k) \leq r\}}{N - m + 1} \quad \forall k. \quad (7)$$

Due to the normalization, the value of $C_j^m(r)$ is always smaller than or equal to 1. Note that the value is, however, at least $1/(N - m + 1)$ since u_j is also included in the count. Then, take the natural logarithm of each $C_j^m(r)$ and average over j to yield

$$\Phi^m(r) = \frac{1}{N - m + 1} \sum_{j=1}^{N-m+1} \ln C_j^m(r). \quad (8)$$

Finally, the ApEn is obtained as $\text{ApEn}(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r)$.

2.8.11. Sample Entropy. Sample entropy (SampEn) [37–39] is analogous to ApEn but there are two significant modifications in its computation. For ApEn, in the computation of the number of vectors u_k for which $d(u_j, u_k) \leq r$, also the vector u_j itself is contained within. This ensures that $C_j^m(r)$ is always greater than zero and the logarithm can be calculated. Regrettably, it makes ApEn biased. SampEn was formulated to lessen this bias. Yet again, the embedding dimension is m

and the tolerance parameter r . We set m to 2 and r to 20% of the standard deviation of the time series. Equally, ApEn and SampEn are estimations for the negative natural logarithm of the conditional probability that data of length N , having repeated itself within a tolerance r for m points, will also repeat itself for $m + 1$ points.

SampEn is also described as in the Kubios HRV Analysis Manual [38].

In SampEn, the self-comparison of u_j is eliminated by calculating $C_j^m(r)$ as

$$C_j^m(r) = \frac{\text{number of } \{u_k \mid d(u_j, u_k) \leq r\}}{N - m} \quad \forall k \neq j. \quad (9)$$

Now the value of $C_j^m(r)$ will be between 0 and 1. Then, the values of $C_j^m(r)$ are averaged to yield

$$C^m(r) = \frac{1}{N - m + 1} \sum_{j=1}^{N-m+1} C_j^m(r). \quad (10)$$

SampEn is described mathematically as $\text{SampEn}(m, r, N) = \ln(C^m(r)/C^{m+1}(r))$.

2.8.12. Higuchi Fractal Dimension (HFD). Fractal systems exhibit a characteristic termed self-similarity. A self-similar object upon close examination is comprised of smaller versions of itself. There are several algorithms which can be applied to measure fractal dimension. There are those by Higuchi [40], Katz [41], and Castiglioni [42]. Here, we apply the technique formulated by Higuchi viewed frequently as the most robust technique.

Higuchi derived this new algorithm to measure the fractal dimension of discrete time sequences. It is a technique that is enforced directly to the RR intervals. There is no power spectrum step involved. As the reconstruction of the attractor phase space is unnecessary, the algorithm is simpler and faster than the Correlation Dimension [43, 44]. Khoa et al. [45] describe the algorithm mathematically, adapted below.

It is based on a measure of length, $L(k)$, of the curve that represents the considered time series while using a segment of k samples as a unit, if $L(k)$ scales like

$$L(k) \sim k^{-D_f}. \quad (11)$$

The curve is said to show fractal dimension D_f because a simple curve has dimension equal to 1 and a plane has dimension equal to 2; value of D_f is always between 1 (simple curve) and 2 (curve which almost fills out the whole plane). D_f measures complexity of the curve and so of the time series this curve represents on a graph.

From a given time series, $RR(1), RR(2), \dots, RR(N)$, the algorithm constructs k new time series:

$$RR_{km} = \left\{ RR(m), RR(m+k), RR(m+2k), \dots, \right. \\ \left. RR\left(m + \text{int}\left(\frac{(N-m)}{k}\right) \cdot k\right) \right\} \quad \text{for } m = 1, 2, \dots, k, \quad (12)$$

TABLE 1: Body mass index (BMI), age, height, and mass of the volunteers. m: meters; kg: kilograms; bpm: beats per minute; ms: milliseconds; mmHg: millimeters of mercury.

Variable	Value
Age (years)	20.5 ± 1.4
Height (m)	1.62 ± 0.5
Mass (kg)	59.9 ± 12.1
BMI (kg/m ²)	22.7 ± 4.5

where m is initial time value, k indicates the discrete time interval between points (hence the delay, k_{\max} , is the maximum interval time), and $\text{int}(a)$ is integer part of a real number a .

For each of the time series RR_{km} constructed, the average length $L_m(k)$ is then computed as

$$L_m(k) = \frac{1}{k} \left[\left(\sum_{i=1}^{\text{int}((N-m)/k)} |\text{RR}(m+i \cdot k) - \text{RR}(m+(i-1) \cdot k)| \right) \right] \quad (13)$$

$$\times \frac{N-1}{\text{int}((N-m)/k) \cdot k},$$

where N is total number of RR intervals. Afterwards, the length of the curve for time interval k is expressed as the sum value over k sets of $L_m(k)$ as illustrated by the following equation:

$$L(k) = \frac{1}{k} \sum_{m=1}^k L_m(k). \quad (14)$$

Finally, the slope of the curve $\ln(L(k))/\ln(1/k)$ is estimated using least squares linear best fit and the resulting slope is the HFD. To select a suitable value for k_{\max} , HFD values are plotted against a range of k_{\max} . The point at which the fractal dimension plateaus is considered a saturation point. That k_{\max} value should be selected. No saturation point is achieved with the data we measured here.

2.8.13. Effect Size. To quantify the magnitude of difference between protocols for significant differences, the effect size was calculated using Cohen's d for significant differences ($p < 0.005$). Effect size was considered large for values ≥ 0.9 , medium for values between 0.9 and 0.5, and small for values between 0.5 and 0.25 [46].

3. Results

Table 1 illustrates the values for mass, height, and BMI of the volunteers; all values were within normal physiological standards.

According to Figures 3 and 4, we illustrate that traffic noise did not induce significant changes in linear indices of HRV analysis. There was no significant change in the time (heart rate, SDNN, Mean RR, pNN50, and RMSSD) and frequency domain (LF and HF in absolute and normalized units and LF/HF ratio) indices of HRV.

3.1. Chaotic Global Analysis. In Table 2 and Figure 5, we display mean values and standard deviation for the chaotic forward parameters (CFP1 to CFP7) for the normal and traffic noise exposure subjects. There are 500 RR intervals throughout and both the parametric one-way analysis of variance (ANOVA1) and the nonparametric Kruskal-Wallis tests of significance are applied. The following are the inconclusive tests of normality (see below).

There are seven permutations of the three chaotic global parameters. All chaotic global values have equal weighting. The chaotic forward parameter (CFP) enables different combinations of chaotic globals to be applied to ensure that we have the best combination to be verified later by a multivariate analysis. It is anticipated that the CFP which applies all three should be the most robust. This is because it takes the information and processes it in three different ways. The summation of the three would be expected to deviate greater than single or double permutations. The potential analytical hazard here is that since we are only calculating spectral components, the phase information is lost.

When implementing parametric statistics, normal distribution of data is assumed. To test this assumption, we apply the Anderson-Darling and Lilliefors tests. In the case of the Anderson-Darling test, an empirical cumulative distribution function is applied, while the Lilliefors test is beneficial when the number of subjects is low. The results from both tests reveal similar numbers of nonnormal and normal distributions, so we apply both the Kruskal-Wallis and ANOVA1 tests of significance.

3.2. Principal Component Analysis. Principal Component Analysis (PCA) is a multivariate technique for analyzing the complexity of high-dimensional datasets. PCA is useful when (1) sources of variability in the data need to be explained and (2) reducing the complexity of the data and through this assessing the data with less dimensions. The primary goal of PCA is to rationalize the sources of variability in the data and to represent the data with fewer variables while sustaining the majority of the total variance (Figure 6).

CFP1t has the First Principal Component (PC1) of 0.358 and the Second Principal Component (PC2) of -0.406 . However, CFP3t has PC1 of 0.191 and PC2 of -0.540 . Only the first two components need be considered due to the steep scree plot. The cumulative influence as a percentage is 58.1 percent for the PC1 and 99.5 percent for the cumulative total of the PC1 and PC2. PC2 has an influence of 41.3 percent. So, CFP1 which applies all three chaotic global techniques is the optimal and most robust overall combination regarding influencing the correct outcome (Figure 6).

Table 3 illustrates the relevant Principal Component Analysis for CFPt for 7 groups of 31 traffic noise exposure subjects. The CFP values are deduced from RR-interval time series and with the chaotic global algorithms enforced.

3.3. Higuchi Fractal Dimension (HFD). The descriptive statistics of the Higuchi fractal dimension from the control subjects ($N = 31$) for 500 RR intervals are presented in Table 4. The parameter was calculated repeatedly for values of K_{\max} between 10 and 150 at intervals of 10.

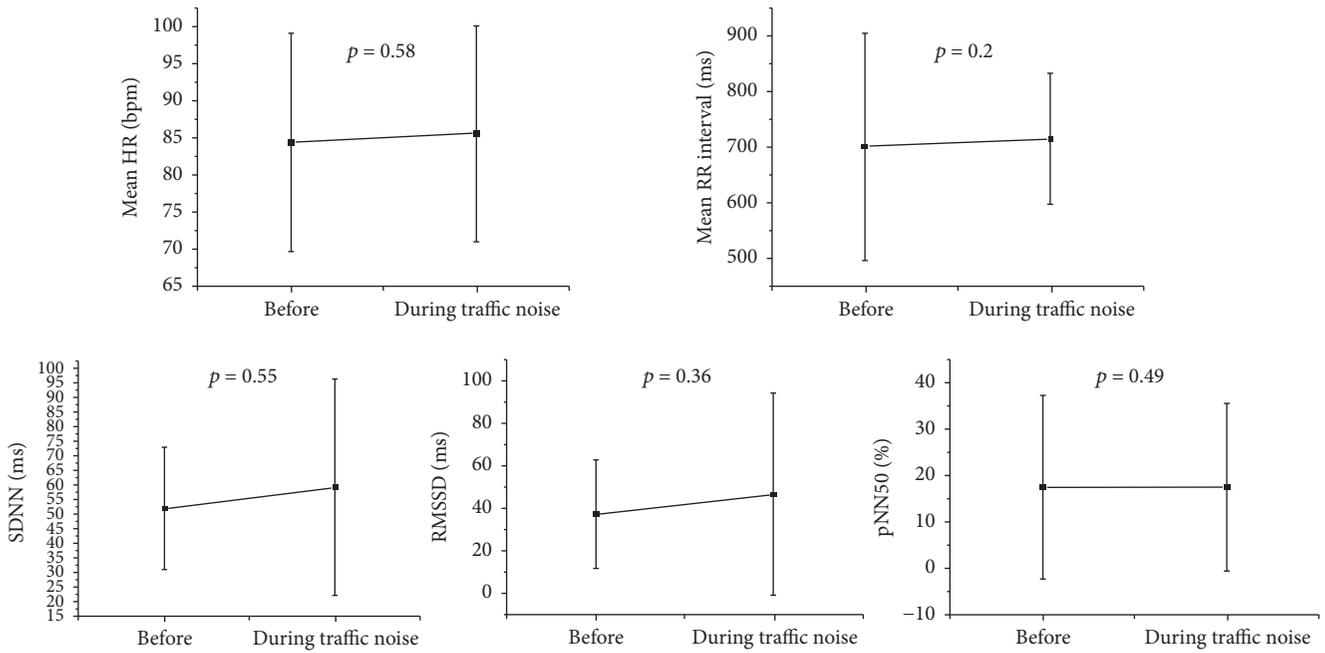


FIGURE 3: Mean heart rate and HRV analysis before (control) and during traffic noise exposure. pNN50: the percentage of adjacent RR intervals with a difference of duration greater than 50 ms; RMSSD: root-mean square of differences between adjacent normal RR intervals in a time interval; SDNN: standard deviation of normal-to-normal RR intervals; HR: heart rate; p : level of significance.

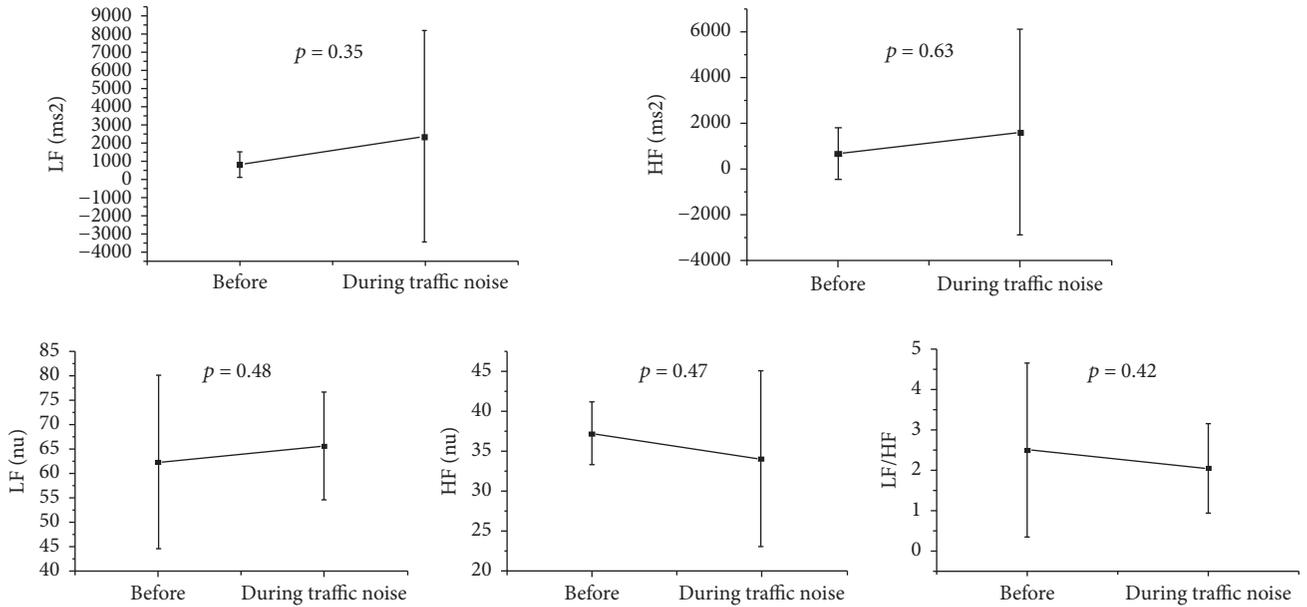


FIGURE 4: HRV analysis before (control) and during traffic noise exposure. LF: low frequency; HF: high frequency; LF/HF: low frequency/high frequency ratio; n.u.: normalized units; ms: milliseconds; HR: heart rate; p : level of significance.

The descriptive statistics of the Higuchi fractal dimension from the traffic noise exposure subjects ($N = 31$) for 500 RR intervals are presented in Table 5. The parameter was calculated repeatedly for values of K_{max} between 10 and 150 at intervals of 10.

Figure 7 illustrates the box-and-whiskers plot for Higuchi fractal dimension of RR intervals of the control subjects (a)

and the traffic noise exposure subjects (b), calculated multiple times from 10 to 150 in equidistant units for different levels of K_{max} . The point closest to the zero is the minimum and the point farthest away is the maximum. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median (not the mean), and the boundary of the box farthest from zero indicates the 75th percentile.

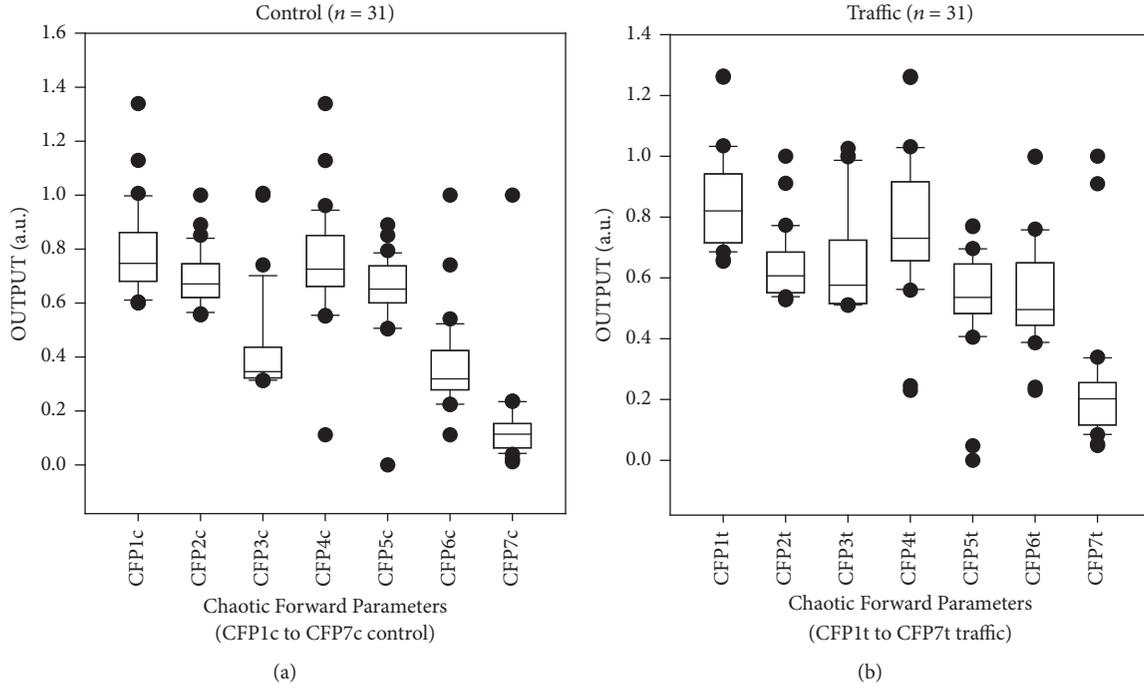


FIGURE 5: The boxplots illustrate the values of chaotic forward parameters one to seven (CFP1 to CFP7) for control (a) and traffic noise exposure (b) subjects with 500 RR intervals throughout. The point closest to the zero is the minimum and the point farthest away is the maximum. The boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median (not the mean), and the boundary of the box farthest from zero indicates the 75th percentile. The difference between these points is the interquartile range (IQR). Whiskers (or error bars) above and below the box indicate the 90th and 10th percentiles, respectively.

TABLE 2: Mean values and standard deviation for the chaotic forward parameters (CFP) for the normal and traffic noise exposure subjects.

Chaotic global	Mean \pm SD Normal ($n = 31$)	Mean \pm SD Traffic ($n = 31$)	ANOVA1 (p value)	Kruskal-Wallis (p value)	Effect size
CFP1	0.7853 ± 0.1602	0.8491 ± 0.1620	0.1243	0.0704	-
CFP2	0.6889 ± 0.1003	0.6325 ± 0.1105	0.0395	0.0043	0.53 (medium)
CFP3	0.4182 ± 0.1790	0.6368 ± 0.1598	<0.0001	<0.0001	1.28 (large)
CFP4	0.7439 ± 0.2022	0.7749 ± 0.2281	0.5739	0.4182	-
CFP5	0.6428 ± 0.1524	0.5376 ± 0.1676	0.0121	0.0011	0.65 (medium)
CFP6	0.3620 ± 0.1646	0.5517 ± 0.1764	<0.0001	<0.0001	1.11 (large)
CFP7	0.1440 ± 0.1698	0.2316 ± 0.2084	0.0748	0.0016	0.46 (small)

TABLE 3: Principal Component Analysis for CFPt for 7 groups of 31 traffic noise exposure subjects.

Chaotic global	PC1	PC2
CFP1t	0.358	-0.406
CFP2t	0.066	-0.577
CFP3t	0.191	-0.540
CFP4t	0.490	0.086
CFP5t	0.446	0.253
CFP6t	0.494	-0.023
CFP7t	-0.384	-0.371

TABLE 4: Higuchi fractal dimension statistics through K_{\max} between 10 and 150 at intervals of 10 in the control protocol.

Property		Higuchi fractal dimension Statistics (control)						
K_{\max}	Mean	SE mean	StDev	Minimum	Q1	Median	Q3	Max
10	1.6768	0.0309	0.1722	1.1992	1.5644	1.6971	1.7999	1.9369
20	1.7446	0.0274	0.1526	1.2605	1.7057	1.7664	1.8479	1.9496
30	1.7783	0.0259	0.1441	1.3436	1.7341	1.8055	1.8879	1.9604
40	1.8017	0.0253	0.1408	1.3984	1.7661	1.8355	1.9092	1.9723
50	1.8194	0.0242	0.1350	1.4385	1.7921	1.8589	1.9115	1.9709
60	1.8333	0.0233	0.1295	1.4514	1.8088	1.8785	1.9219	1.9711
70	1.8436	0.0224	0.1246	1.4686	1.8302	1.8824	1.9283	1.9688
80	1.8509	0.0216	0.1205	1.4868	1.8348	1.8883	1.9298	1.9664
90	1.8573	0.0211	0.1173	1.5063	1.8488	1.8995	1.9315	1.9675
100	1.8618	0.0207	0.1151	1.5267	1.8565	1.9007	1.9334	1.9702
110	1.8659	0.0203	0.1133	1.5473	1.8655	1.9080	1.9344	1.9697
120	1.8709	0.0203	0.1128	1.5659	1.8799	1.9131	1.9358	1.9705
130	1.8760	0.0202	0.1126	1.5682	1.8928	1.9216	1.9388	1.9695
140	1.8808	0.0201	0.1119	1.5743	1.8944	1.9298	1.9437	1.9715
150	1.8852	0.0201	0.1117	1.5769	1.8915	1.9311	1.9484	1.9755

TABLE 5: Higuchi fractal dimension statistics through K_{\max} between 10 and 150 at intervals of 10 in the traffic noise protocol.

Property		Higuchi fractal dimension statistics (traffic noise exposure)						
K_{\max}	Mean	SE mean	StDev	Minimum	Q1	Median	Q3	Max
10	1.6971	0.0279	0.1555	1.2606	1.6284	1.7171	1.8040	1.9496
20	1.7644	0.0265	0.1477	1.2952	1.7010	1.8077	1.8462	1.9383
30	1.7898	0.0256	0.1428	1.3544	1.7507	1.8447	1.8762	1.9450
40	1.8082	0.0242	0.1347	1.3912	1.7775	1.8533	1.8811	1.9579
50	1.8240	0.0227	0.1264	1.4132	1.8103	1.8590	1.9000	1.9703
60	1.8358	0.0217	0.1208	1.4290	1.8232	1.8741	1.9082	1.9691
70	1.8446	0.0214	0.1193	1.4368	1.8208	1.8879	1.9161	1.9722
80	1.8507	0.0212	0.1178	1.4372	1.8269	1.8934	1.9140	1.9762
90	1.8568	0.0210	0.1171	1.4337	1.8414	1.8922	1.9194	1.9804
100	1.8613	0.0209	0.1163	1.4326	1.8497	1.8990	1.9243	1.9837
110	1.8660	0.0207	0.1151	1.4383	1.8569	1.9076	1.9267	1.9824
120	1.8694	0.0203	0.1130	1.4474	1.8613	1.9121	1.9286	1.9811
130	1.8727	0.0200	0.1115	1.4597	1.8701	1.9103	1.9338	1.9831
140	1.8769	0.0197	0.1099	1.4732	1.8809	1.9070	1.9397	1.9810
150	1.8806	0.0193	0.1075	1.4888	1.8854	1.9066	1.9380	1.9795

TABLE 6: Higuchi fractal dimension at varying levels of K_{\max} between 10 and 150 at equidistant intervals of 10.

Property	Higuchi fractal dimension statistics (control versus traffic)	
	ANOVA1 (p value)	Kruskal-Wallis (p value)
K_{\max}		
10	0.6269	0.6322
20	0.6059	0.5543
30	0.7534	0.6523
40	0.8539	0.9215
50	0.8904	0.8880
60	0.9367	1.0000
70	0.9742	0.9215
80	0.9940	0.8108
90	0.9855	0.7675
100	0.9860	0.7568
110	0.9981	0.7354
120	0.9585	0.5082
130	0.9082	0.3041
140	0.8911	0.2910
150	0.8683	0.3175

The difference between these points is the interquartile range (IQR). Whiskers (or error bars) above and below the box indicate the 90th and 10th percentiles, respectively.

The levels of significance for parametric ANOVA1 and nonparametric Kruskal-Wallis test of significance for values of the Higuchi fractal dimension at varying levels of K_{\max} between 10 and 150 at equidistant intervals of 10 are displayed in Table 6.

3.4. Five Entropies and DFA

3.4.1. ANOVA1 and Kruskal-Wallis Tests. Once more, we apply the Anderson-Darling and Lilliefors tests to the data to assess the normality. The results from both tests reveal similar numbers of nonnormal and normal distributions. So again we apply the Kruskal-Wallis and ANOVA1 tests of significance.

Table 7 reveals the mean values and standard deviation for the five entropic measures and DFA for the control and traffic noise exposure subjects RR intervals. The number of RR intervals is 500. ANOVA1 and Kruskal-Wallis test of significance were applied to results.

3.4.2. Principal Components Analysis. Here again we must complete a multivariate analysis. Shannon entropy has the First Principal Component (PC1) of 0.470, the Second Principal Component (PC2) of 0.258, and the Third Principal Component (PC3) of -0.245 . But, Renyi entropy has the PC1 of 0.485, PC2 of 0.187, and PC3 of -0.200 . However, Tsallis entropy has the PC1 of 0.472, PC2 of 0.249, and PC3 of -0.242 .

Only the first three components need be considered due to the relatively steep scree plot. The cumulative influence as a percentage is 65.4 percent for the PC1 and 95.4 percent for the cumulative total of the PC1 and PC2. Finally, it is 99.3 percent for the cumulative total of the PC1, PC2, and PC3.

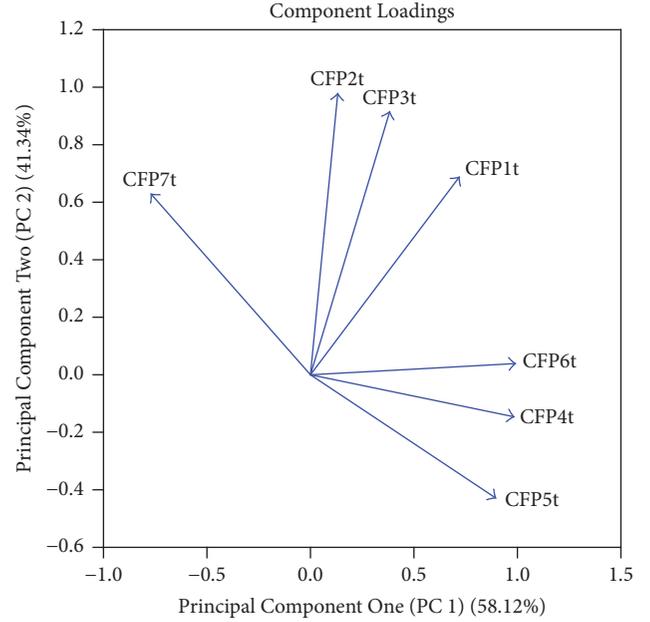


FIGURE 6: The plot illustrates the component loadings CFP1 to CFP7 for the 500 RR intervals of 31 traffic noise exposure subjects. The CFP values are deduced by using the MTM spectra throughout. The properties of the MTM spectra are as follows: sampling frequency 1Hz, DPSS of 3, FFT length of 256, and Thomson's nonlinear combination at "adaptive." CFP1 and CFP3 are the most influential components when assessed by PCA.

PC2 has an influence of 30.0 percent. PC3 has an influence of 3.9 percent. So, Shannon, Renyi, and Tsallis are the optimal and most robust statistically overall combination regarding influencing the correct outcome. This is the case by means of the ANOVA1, Kruskal-Wallis, and the multivariate technique, hence PCA.

Table 8 illustrates the relevant Principal Component Analysis for five entropies and DFA of 31 traffic noise exposure subjects. The five entropy values and DFA are again deduced from 500 RR-interval time series.

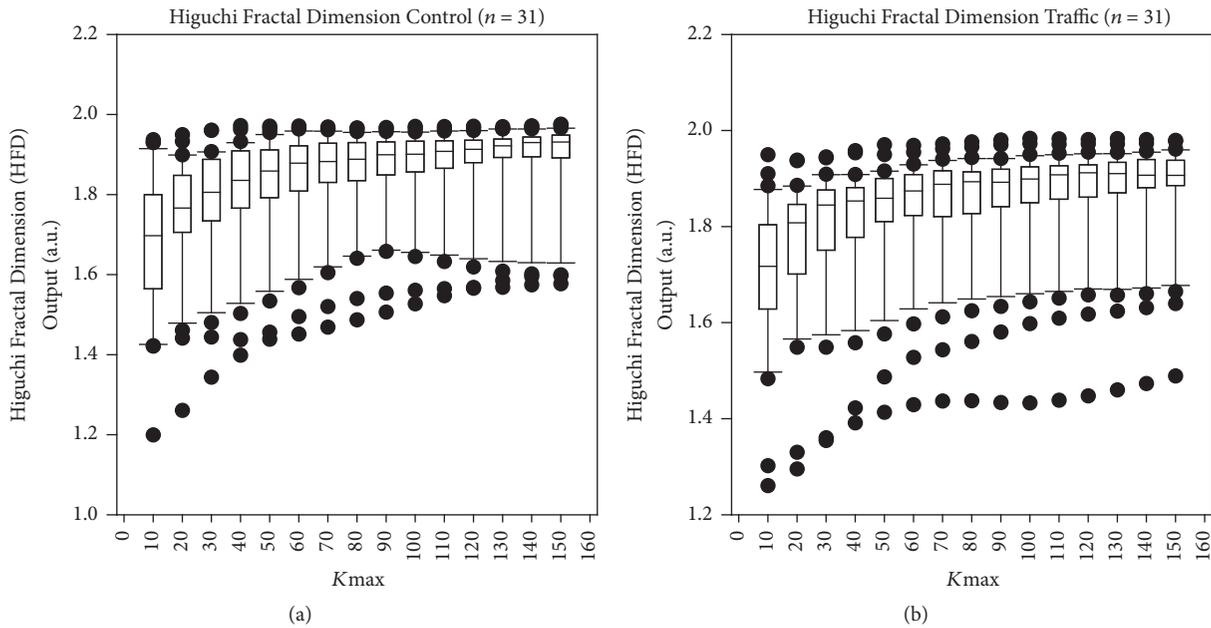
4. Discussion

To provide further evidence regarding the interaction between auditory processing and the autonomic nervous system, we attempted to investigate whether acute exposure to traffic noise influenced the complexity of HRV. As a main outcome, we noticed that the traditional linear indices of HRV were unchanged during traffic noise exposure while some nonlinear approaches evidenced that the complexity of heart rate autonomic control increased during exposure to traffic noise.

In this context, previous studies suggest that noise exposure affects the sympathetic component of heart rate autonomic control [47, 48]. Tzaneva et al. [47] exposed subjects to 135 min of noise with Leq 95 dB (A) sound pressure and analyzed HRV before, during, and after noise exposure. They revealed an increase in the sympathetic regulation of heart rate under noise exposure. Björ et al. [48] investigated

TABLE 7: Entropic measures for the control and traffic noise exposure subjects RR intervals.

Entropy or DFA	Mean \pm SD Control ($n = 31$)	Mean \pm SD Traffic noise ($n = 31$)	ANOVA1 (p value)	Kruskal-Wallis (p value)	Effect size
Approximate	0.7443 \pm 0.2354	0.7890 \pm 0.2235	0.4465	0.3107	-
Sample	0.6923 \pm 0.2300	0.7047 \pm 0.2165	0.8278	0.8658	-
DFA	0.1722 \pm 0.1943	0.1483 \pm 0.1871	0.6222	0.6831	-
Shannon	0.5564 \pm 0.1289	0.7017 \pm 0.1272	<0.0001	<0.0001	1.13 (large)
Renyi	0.9840 \pm 0.0058	0.9898 \pm 0.0051	<0.0001	<0.0001	1.06 (large)
Tsallis	0.5981 \pm 0.1193	0.7322 \pm 0.1159	<0.0001	<0.0001	1.14 (large)

FIGURE 7: Box-and-whiskers plot for Higuchi fractal dimension of RR intervals of the control subjects (a) and the traffic noise exposure subjects (b), calculated multiple times from 10 to 150 in equidistant units for different levels of K_{\max} .

healthy men and women and also noted increased values of the LF/HF ratio during noise exposure, indicating increased sympathetic control of heart rate.

Yet, an important point to be highlighted in their studies is the limitation of the LF/HF ratio to provide information regarding the sympathetic modulation of heart rate. The sympathovagal balance index that was added to their investigation, calculated by the LF/HF ratio, has been demonstrated to be theoretically flawed and empirically unsupported. Though many criticisms of this measure abound, the most serious concern is that LF index does not represent the sympathetic component. Thus, there is a lack of rationale and/or compelling evidence that its strength in relation to the HF index component would indicate relative strength of vagal and sympathetic signaling. Furthermore, the physiological significance of LF/HF ratio is erroneous and represents a superficial understanding of autonomic regulatory mechanisms [49–51]. We therefore emphasize that spectral analyses of HRV under controlled situations are the most effective

markers of heart rate autonomic modulation. Yet, they do not accurately measure neural traffic or autonomic activity (i.e., pupil dilation, salivation, facial vasodilation, etc.).

Equally, Sim et al. [2] evaluated the effects of different noises on linear HRV. The authors enrolled 40 healthy men (23.9 ± 1.8 years old, and average BMI being 23.7 ± 2.1 kg/m²) and submitted them to self-made traffic noise composed by aircraft and road traffic noise. The authors observed that traffic noise exposure increased SDNN and HF band in absolute units, indicating that traffic noise acutely increased HRV.

Although we did not observe any significant effects of traffic noise on time and frequency domain indices of HRV, we reported significant changes in the nonlinear parameters of HRV during traffic noise exposure. Entropic and chaotic global analysis of HRV revealed that the complexity of heart rate autonomic control increased during traffic noise exposure, suggesting increasing randomness in the system.

According to our findings, Shannon entropy values increased (large effect size) during traffic noise exposure.

TABLE 8: The relevant Principal Component Analysis for five entropies and DFA of 31 traffic noise exposure subjects.

Entropy (or DFA)	PC1	PC2	PC3
Approximate	0.397	-0.415	0.497
Sample	0.405	-0.419	0.307
DFA	0.007	0.700	0.708
Shannon	0.470	0.258	-0.245
Renyi	0.485	0.187	-0.200
Tsallis	0.472	0.249	-0.242

Entropy is theoretically related to the amount of disorder of particles in a system; if the entropy decreases, the predictability of the process increases and the system becomes less complex [52]. The Shannon entropy quantifies the complexity of a system by means of an average information content [52]. In a recent study, heart beat time series were quantified by Shannon entropy and decreased values were associated with increased severity of pathological condition [53]. Also, decreased Shannon entropy values were found in leprosy victims when HRV was investigated [54].

We also revealed that Renyi entropy values were higher during exposure to traffic noise (large effect size). The Renyi entropy generalizes the Shannon entropy and considers the Shannon entropy as a singular case [55]. The Renyi entropy was previously reported to identify cardiac autonomic neuropathy [56]. It was recently shown as an effective method in real-time monitoring of atrial fibrillation patients and for prediction and diagnosis of paroxysmal atrial fibrillation [57].

Based on our data, Tsallis entropy analysis confirmed that the complexity of HRV increased during traffic noise exposure and Cohen's d calculation exhibited large effect size. This nonlinear approach is not chiefly used in HRV analysis; Eduardo Virgilio Silva and Otavio Murta [58] applied Tsallis entropy in time series and suggested it as a potential method for complexity system analysis, thus supporting our conclusions.

Our results demonstrated through chaotic global analysis of HRV that CFP3 and CFP6 significantly increased (large effect size) during traffic noise exposure, indicating higher complexity of RR-intervals oscillations during auditory stimulation. A previous study reported that chaotic global analysis was unable to identify HRV changes during mental task [59]. Another research study investigated chaotic global analysis in RR intervals during exposure to heavy metal music [60]. The authors failed to reveal influences of this music style on the complexity of HRV.

Nonlinear analysis of HRV is a complex issue owing to its physiological interpretation. Conversely, the literature shows that decreased complexity of HRV represents a physiological impairment. Accordingly, our data points to an interesting interpretation that acute traffic noise exposure in a laboratory situation does not cause stressful autonomic responses. An elegant systematic review reported that the majority of studies performed at the roadside evidenced stressful effects of traffic noise on cardiovascular, respiratory, and metabolic health [61]. However, in view of our results, we deduced that

the stress induced by exposure to road traffic noise is not only due to the auditory stimulus but due to the roadside environmental situation.

The interaction between auditory processing and heart rate autonomic control has been reviewed before [62]. Nakamura et al. [63] reported that auditory stimulation influenced renal sympathetic nerve activity and blood pressure in anesthetized rats. The same researchers observed that vagal gastric nerve activity was similarly influenced by music [64]. The authors indicated that the suprachiasmatic nucleus of the hypothalamus is involved in this process [63].

Amongst the important points to be addressed in our study, we allow for the laboratory conditions the volunteers were exposed to. This is because we intended to discard the influence of the traffic environmental impact on HRV, that is, pollution, visual stimulation, and conversation. We investigated only women in order to avoid influence of sexual hormones. We believe that a combination of different factors during traffic noise stimulus would induce tougher effects on HRV, since the ANS is sensitive to innumerable exogenous elements [14].

The luteal and follicular phase of the menstrual cycle were also controlled, since there is previous evidence of its influence on nonlinear HRV [11].

Another fact worth highlighting is that, in our study, nonlinear methods of HRV were more sensitive at detecting changes in the RR-interval fluctuations. This is possibly because some information may be erroneous if only linear analysis is undertaken. Nonlinear analysis was revealed to be a more powerful approach to identify complex systems [9].

5. Conclusion

Traffic noise exposure did not significantly alter linear indices of HRV. Higuchi fractal dimension, DFA, and Approximate and Sample entropies were similarly significantly unaffected. Yet, it significantly changed chaotic global analysis (combinations CFP3 and CFP6) and Shannon, Renyi, and Tsallis entropies. Our results indicate that traffic noise acutely enhances the complexity of heart rate autonomic control in healthy women.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study received financial support from FAPESP (Process no. 2012/01366-6 and 2018/02664-7).

References

- [1] T. B. Lonsdorf and C. J. Merz, "More than just noise: Inter-individual differences in fear acquisition, extinction and return of fear in humans-biological, experiential, temperamental factors, and methodological pitfalls," *Neuroscience & Biobehavioral Reviews*, vol. 80, pp. 703–728, 2017.
- [2] C. S. Sim, J. H. Sung, S. H. Cheon, J. M. Lee, J. W. Lee, and J. Lee, "The effects of different noise types on heart rate variability in men," *Yonsei Medical Journal*, vol. 56, no. 1, pp. 235–243, 2015.
- [3] G.-S. Lee, M.-L. Chen, and G.-Y. Wang, "Evoked response of heart rate variability using short-duration white noise," *Autonomic Neuroscience: Basic and Clinical*, vol. 155, no. 1-2, pp. 94–97, 2010.
- [4] M. Umemura and K. Honda, "Influence of music on heart rate variability and comfort," *Journal of Human Ergology*, vol. 27, pp. 30–38, 1998.
- [5] O. K. A. Lee, Y. F. L. Chung, M. F. Chan, and W. M. Chan, "Music and its effect on the physiological responses and anxiety levels of patients receiving mechanical ventilation: A pilot study," *Journal of Clinical Nursing*, vol. 14, no. 5, pp. 609–620, 2005.
- [6] M. Fernandes de Godoy, "Nonlinear Analysis of Heart Rate Variability: A Comprehensive Review," *Journal of Cardiology and Therapy*, vol. 3, no. 3, pp. 528–533, 2016.
- [7] A. Voss, S. Schulz, R. Schroeder, M. Baumert, and P. Caminal, "Methods derived from nonlinear dynamics for analysing heart rate variability," *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 367, no. 1887, pp. 277–296, 2009.
- [8] B. Francesco, B. Maria Grazia, and G. Emanuele, "Linear and nonlinear heart rate variability indexes in clinical practice," *Computational and Mathematical Methods in Medicine*, vol. 2012, Article ID 219080, 5 pages, 2012.
- [9] R. Sassi, S. Cerutti, F. Lombardi et al., "Advances in heart rate variability signal analysis: joint position statement by the e-Cardiology ESC Working Group and the European Heart Rhythm Association co-endorsed by the Asia Pacific Heart Rhythm Society," *EP Europace*, vol. 17, no. 9, pp. 1341–1353, 2015.
- [10] H. V. Huikuri, T. H. Mäkikallio, and J. Perkiömäki, "Measurement of heart rate variability by methods based on nonlinear dynamics," *Journal of Electrocardiology*, vol. 36, pp. 95–99, 2003.
- [11] X. Bai, J. Li, L. Zhou, and X. Li, "Influence of the menstrual cycle on nonlinear properties of heart rate variability in young women," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 297, no. 2, pp. H765–H774, 2009.
- [12] A. L. Roque, V. E. Valenti, H. L. Guida et al., "The effects of different styles of musical auditory stimulation on cardiac autonomic regulation in healthy women," *Noise & Health*, vol. 15, no. 65, pp. 281–287, 2013.
- [13] A. L. Roque, V. E. Valenti, H. L. Guida et al., "The effects of auditory stimulation with music on heart rate variability in healthy women," *Clinics*, vol. 68, no. 7, pp. 960–967, 2013.
- [14] A. J. Camm, M. Malik, J. T. Bigger et al., "Heart rate variability: standards of measurement, physiological interpretation and clinical use," *Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation*, vol. 93, pp. 1043–1065, 1996.
- [15] C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 5, no. 1, pp. 82–87, 1995.
- [16] G. C. Donaldson, T. A. R. Seemungal, J. R. Hurst, and J. A. Wedzicha, "Detrended fluctuation analysis of peak expiratory flow and exacerbation frequency in COPD," *European Respiratory Journal*, vol. 40, no. 5, pp. 1123–1129, 2012.
- [17] P. Castiglioni, L. Quintin, A. Civijian, G. Parati, and M. Di Rienzo, "Local-scale analysis of cardiovascular signals by detrended fluctuations analysis: Effects of posture and exercise," in *Proceedings of the Engineering in Medicine and Biology Society, EMBS 29th Annual International Conference of the IEEE*, pp. 5035–5038, 2007.
- [18] R. Karasik, N. Sapir, Y. Ashkenazy et al., "Correlation differences in heartbeat fluctuations during rest and exercise," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 66, no. 6, Article ID 062902, p. 062902/4, 2002.
- [19] R. De Leon-Lomeli, J. S. Murguia, I. Chouvarda et al., "Relation between heart beat fluctuations and cyclic alternating pattern during sleep in insomnia patients," in *Proceedings of the Engineering in Medicine and Biology Society (EMBC), 36th Annual International Conference of the IEEE*, pp. 2249–2252, 2014.
- [20] C.-M. Liao, N.-H. Hsieh, and C.-P. Chio, "Fluctuation analysis-based risk assessment for respiratory virus activity and air pollution associated asthma incidence," *Science of the Total Environment*, vol. 409, no. 18, pp. 3325–3333, 2011.
- [21] R. C. Rossi, F. M. Vanderlei, A. F. Bernardo et al., "Effect of pursed-lip breathing in patients with COPD: Linear and non-linear analysis of cardiac autonomic modulation," *Copd-Journal of Chronic Obstructive Pulmonary Disease*, vol. 11, no. 1, pp. 39–45, 2014.
- [22] J. Annegarn, M. Spruit, H. Savelberg et al., "Stride time fluctuations during the six minute walk test in COPD patients," in *Proceedings of the Rehabilitation: Mobility, Exercise, and Sports: 4th International State-of-the-Art Congress*, p. 149, IOS Press, 2010.
- [23] M. Ghil, "The SSA-MTM toolkit: Applications to analysis and prediction of time series," *Applications of Soft Computing*, pp. 216–230, 1997.
- [24] D. Slepian, "Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty—V: The Discrete Case," *Bell System Technical Journal*, vol. 57, no. 5, pp. 1371–1430, 1978.
- [25] S. H. Gould, *Variational methods for eigenvalue problems: an introduction to the methods of Rayleigh, Ritz, Weinstein, and Aronszajn*, Courier Dover Publications, 1995.
- [26] R. Wajnsztein, T. D. De Carvalho, D. M. Garner et al., "Heart rate variability analysis by chaotic global techniques in children with attention deficit hyperactivity disorder," *Complexity*, vol. 21, no. 6, pp. 412–419, 2016.
- [27] D. M. Garner and B. W.-K. Ling, "Measuring and locating zones of chaos and irregularity," *Journal of Systems Science and Complexity*, vol. 27, no. 3, pp. 494–506, 2014.
- [28] C. E. Shannon, "A mathematical theory of communication," *Bell Labs Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.
- [29] R. J. dos Santos, "Generalization of Shannon's theorem for Tsallis entropy," *Journal of Mathematical Physics*, vol. 38, no. 8, pp. 4104–4107, 1997.
- [30] E. K. Lenzi, R. S. Mendes, and L. R. Da Silva, "Statistical mechanics based on Renyi entropy," *Physica A: Statistical Mechanics and its Applications*, vol. 280, no. 3, pp. 337–345, 2000.

- [31] A. E. Taylor, "L'Hospital's rule," *The American Mathematical Monthly*, vol. 59, pp. 20–24, 1952.
- [32] I. Pinelis, "L'Hospita Type rules for monotonicity, with applications," *Journal of Inequalities in Pure and Applied Mathematics*, no. 3, 2002.
- [33] K. Zyczkowski, "Rényi extrapolation of Shannon entropy," *Open Systems and Information Dynamics*, vol. 10, no. 3, pp. 297–310, 2003.
- [34] A. R. Plastino and A. Plastino, "Stellar polytropes and Tsallis' entropy," *Physics Letters A*, vol. 174, no. 5–6, pp. 384–386, 1993.
- [35] A. M. Mariz, "On the irreversible nature of the Tsallis and Rényi entropies," *Physics Letters A*, vol. 165, no. 5–6, pp. 409–411, 1992.
- [36] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.
- [37] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.
- [38] M. P. Tarvainen, J.-P. Niskanen, J. A. Lipponen, P. O. Ranta-aho, and P. A. Karjalainen, "Kubios HRV—heart rate variability analysis software," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 210–220, 2014.
- [39] J. S. Richman, D. E. Lake, and J. R. Moorman, "Sample Entropy," *Methods in Enzymology*, vol. 384, pp. 172–184, 2004.
- [40] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Physica D: Nonlinear Phenomena*, vol. 31, no. 2, pp. 277–283, 1988.
- [41] M. J. Katz, "Fractals and the analysis of waveforms," *Computers in Biology and Medicine*, vol. 18, no. 3, pp. 145–156, 1988.
- [42] P. Castiglioni, "What is wrong in Katz's method? Comments on: 'A note on fractal dimensions of biomedical waveforms'," *Computers in Biology and Medicine*, vol. 40, no. 11–12, pp. 950–952, 2010.
- [43] J. E. Skinner, C. M. Pratt, and T. Vybiral, "A reduction in the correlation dimension of heartbeat intervals precedes imminent ventricular fibrillation in human subjects," *American Heart Journal*, vol. 125, no. 3, pp. 731–743, 1993.
- [44] P. Van Leeuwen, H. Bettermann, U. An der Heiden, and H. C. Kummell, "Circadian aspects of apparent correlation dimension in human heart rate dynamics," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 269, no. 1, pp. H130–H134, 1995.
- [45] T. Q. D. Khoa, V. Q. Ha, and V. V. Toi, "Higuchi fractal properties of onset epilepsy electroencephalogram," *Computational and Mathematical Methods in Medicine*, vol. 2012, 2012.
- [46] D. S. Quintana, "Statistical considerations for reporting and planning heart rate variability case-control studies," *Psychophysiology*, vol. 54, no. 3, pp. 344–349, 2017.
- [47] L. Tzaneva, S. Danev, and R. Nikolova, "Investigation of noise exposure effect on heart rate variability parameters," *Central European Journal of Public Health*, vol. 9, no. 3, pp. 130–132, 2001.
- [48] B. Björ, L. Burström, M. Karlsson, T. Nilsson, U. Näslund, and U. Wiklund, "Acute effects on heart rate variability when exposed to hand transmitted vibration and noise," *International Archives of Occupational and Environmental Health*, vol. 81, no. 2, pp. 193–199, 2007.
- [49] G. E. Billman, "The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance," *Frontiers in Physiology*, vol. 4, article 26, 2013.
- [50] J. A. J. Heathers, "Sympathovagal balance from heart rate variability: An obituary," *Experimental Physiology*, vol. 97, no. 4, pp. 556–556, 2012.
- [51] D. S. Goldstein, O. Benthoo, M.-Y. Park, and Y. Sharabi, "Low-frequency power of heart rate variability is not a measure of cardiac sympathetic tone but may be a measure of modulation of cardiac autonomic outflows by baroreflexes," *Experimental Physiology*, vol. 96, no. 12, pp. 1255–1261, 2011.
- [52] T. M. Cover and J. A. Thomas, *Elements of information theory*, 2nd edition, 2006.
- [53] M. Vallverdu, F. Claria, U. Melia, A. B. De Luna, and P. Caminal, "Measuring heart rate variability by means of information entropies based on Choi-Williams distribution," in *Proceedings of the Engineering in Medicine and Biology Society (EMBC), 37th Annual International Conference of the IEEE*, pp. 1797–1800, 2015.
- [54] M. C. D. S. Santos, L. C. D. L. Silveira, S. C. G. Moura-Tonello, A. Porta, A. M. Catai, and G. D. S. Souza, "Heart rate variability in multibacillar leprosy: Linear and nonlinear analysis," *PLoS ONE*, vol. 12, no. 7, Article ID e0180677, 2017.
- [55] R. Alfréd, "On measures of information and entropy," in *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, pp. 547–561, 1960.
- [56] D. J. Cornforth, M. P. Tarvainen, and H. F. Jelinek, "How to Calculate Rényi Entropy from Heart Rate Variability, and Why it Matters for Detecting Cardiac Autonomic Neuropathy," *Frontiers in Bioengineering and Biotechnology*, vol. 2, 2014.
- [57] Y. Xin, Y. Zhao, Y. Mu, Q. Li, and C. Shi, "Paroxysmal atrial fibrillation recognition based on multi-scale Rényi entropy of ECG," *Technology and Health Care*, vol. 25, no. 1, pp. S189–S196, 2017.
- [58] L. Eduardo Virgilio Silva and L. Otavio Murta, "Evaluation of physiologic complexity in time series using generalized sample entropy and surrogate data analysis," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 22, no. 4, Article ID 043105, 2012.
- [59] A. M. G. Fontes, D. M. Garner, L. C. de Abreu et al., "Global chaotic parameters of heart rate variability during mental task," *Complexity*, vol. 21, no. 5, pp. 300–307, 2016.
- [60] M. L. Nogueira, D. M. Garner, E. Osório, L. C. Abreu, and V. E. Valenti, "Globally chaotic analysis of Heart Rate Variability during acute auditory stimulus by heavy metal music," *Medical Express*, vol. 2, no. 5, 2015.
- [61] A. Recio, C. Linares, J. R. Banegas, and J. Díaz, "Road traffic noise effects on cardiovascular, respiratory, and metabolic health: An integrative model of biological mechanisms," *Environmental Research*, vol. 146, pp. 359–370, 2016.
- [62] V. E. Valenti, H. L. Guida, A. C. F. Frizzo, A. C. V. Cardoso, L. C. M. Vanderlei, and L. C. de Abreu, "Auditory stimulation and cardiac autonomic regulation," *Clinics*, vol. 67, no. 8, pp. 955–958, 2012.
- [63] T. Nakamura, M. Tanida, A. Nijima, H. Hibino, J. Shen, and K. Nagai, "Auditory stimulation affects renal sympathetic nerve activity and blood pressure in rats," *Neuroscience Letters*, vol. 416, no. 2, pp. 107–112, 2007.
- [64] T. Nakamura, M. Tanida, A. Nijima, and K. Nagai, "Effect of auditory stimulation on parasympathetic nerve activity in urethane-anesthetized rats," *In Vivo*, vol. 23, no. 3, pp. 415–420, 2009.

Research Article

Invariant Measures Based on the U-Correlation Integral: An Application to the Study of Human Voice

Juan F. Restrepo ^{1,2} and Gastón Schlotthauer^{1,2}

¹Laboratorio de Señales y Dinámicas no Lineales, Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Ruta Prov. 11, Km 10, Oro Verde, Entre Ríos, Argentina

²Instituto de Investigación y Desarrollo en Bioingeniería y Bioinformática, Oro Verde, Entre Ríos, Argentina

Correspondence should be addressed to Juan F. Restrepo; jrestrepo@bioingenieria.edu.ar

Received 27 September 2017; Revised 20 February 2018; Accepted 12 March 2018; Published 17 April 2018

Academic Editor: Roberto Tonelli

Copyright © 2018 Juan F. Restrepo and Gastón Schlotthauer. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nonlinear measures such as the correlation dimension, the correlation entropy, and the noise level were used in this article to characterize normal and pathological voices. These invariants were estimated through an automated algorithm based on the recently proposed U-correlation integral. Our results show that the voice dynamics have a low dimension. The value of correlation dimension is greater for pathological voices than for normal ones. Furthermore, its value also increases along with the type of the voice. The low correlation entropy values obtained for normal and pathological type 1 and type 2 voices suggest that their dynamics are nearly periodic. Regarding the noise level, in the context of voice signals, it can be interpreted as the power of an additive stochastic perturbation intrinsic to the voice production system. Our estimations suggest that the noise level is greater for pathological voices than for normal ones. Moreover, it increases along with the type of voice, being the highest for type 4 voices. From these results, we can conclude that the voice production dynamical system is more complex in the presence of a pathology. In addition, the presence of the inherent stochastic perturbation strengthens along with the voice type. Finally, based on our results, we propose that the noise level can be used to quantitatively differentiate between type 3 and type 4 voices.

1. Introduction

The human voice is the most important means of communication among individuals. Thanks to vocal communication, activities like asking for help are apparently trivial in our daily routine. Thus, a voice disorder can limit our ability to cover our most basic needs, producing a negative impact on our quality of life. For this reason, it is very important not only to increase our knowledge about the mechanism of voice production but also to characterize its dynamics in normal and pathological conditions.

In the literature, several methodologies to assess human voices can be found. However, their reliability depends on the nature of the studied voice. Consequently, Titze proposed a qualitative classification for voice signals [1]. The scheme proposed by Titze divides the signals into three types: type 1 signals are nearly periodic voice signals, type 2 signals have strong modulation or subharmonics, and type 3 signals are

characterized by a very irregular or even chaotic behavior. Sprecher et al. proposed a modification to this scheme. They redefined type 3 voices as being deterministic chaotic signals, adding a fourth type that is characterized by a dominant random-like behavior [2].

Figure 1 shows the time series, the state space reconstruction, and the spectrogram of each type of voice. A normal voice (first column) is characterized for a quasiperiodic time representation and a smooth attractor in the reconstructed state space. Moreover, from its spectrogram, one can easily observe the fundamental frequency and its harmonics. A pathological type 1 voice (second column) displays a more irregular time series than the normal voice. Notice that although their attractors are similar, pathological type 1 attractor is not so smooth. Furthermore, both spectrograms are also similar in the sense that one can still distinguish a fundamental frequency and its harmonics. Nevertheless, the noise content blurs the harmonics at high frequencies.

In the case of a pathological type 2 voice (third column), it is possible to observe a less regular time series, compared to the pathological type 1 voice. The volume occupied by its attractor has been reduced and it is more difficult to distinguish a smooth shape. In its spectrogram, one can find subharmonic frequencies. The pathological type 3 voice (fourth column) is characterized by an irregular time series. The state space reconstruction is similar to the one of white Gaussian noise. Its spectrogram shows the fundamental frequency but its harmonics are rapidly blurred. Finally, a pathological type 4 voice (fifth column) has an irregular time series representation and a regular state space reconstruction, like the type 3 voice. In its spectrogram, one can find a fundamental frequency but not the harmonics.

These representations are useful to differentiate between some kinds of voices, for example, between normal and pathological type 3 voices. However, it is very difficult to differentiate between a normal and a pathological type 1 voice or between a type 3 and a type 4 voice. The importance of this classification is based on the fact that traditional perturbation measures like jitter and shimmer give reliable results if they are applied to type 1 and type 2 signals. In contrast, nonlinear dynamics concepts should be used to characterize type 3 voices, whereas they are unreliable when applied to type 4 signals. Moreover, until now, the classification of type 4 voices has been done subjectively by visual inspection [2].

Over the last three decades, strong evidence about the nonlinear behavior of the voice production mechanism has been collected [3–6], leading the scientific community to develop concepts and methods based on nonlinear dynamics and chaos theories. There exist an extensive set of publications in which those concepts have been used to characterize healthy and pathological voices. The correlation dimension (D) and the correlation entropy (K_2) are two quantities that are used to characterize the complexity of a dynamical system. The former can be thought of as an estimation of the number of variables involved in the dynamics (degrees of freedom) [7]. The latter measures the rate at which information about the dynamics is lost over time. More complex systems are commonly characterized by having higher dimensions and entropy values [8].

The correlation integral is the quantity used to estimate D and K_2 [9]. It has been widely used in the biomedical field since it was proposed by Grassberger and Procaccia in the early 1980s [9].

According to what we know so far, this correlation integral has been used in all studies that have estimated D and K_2 from voice signals. However, it is well known that the Grassberger and Procaccia approach is not robust in the presence of noise, and it needs special conditions to converge [10]. More robust variants of the correlation integrals have been proposed, like the Gaussian kernel correlation integral (GCI) [11] and the U-correlation integral [12].

The objective of this article is to characterize the dynamics of normal and pathological voices through the correlation dimension, the correlation entropy, and the noise level (σ). Furthermore, we seek to statistically analyze these invariants for the four types of voices. As a novelty, we obtain these invariants using U-correlation integral through an automated

algorithm [12] which allowed us to avoid subjective judgments. Finally, we propose a new quantitative methodology to differentiate between type 3 and type 4 voices.

2. Materials and Methods

The correlation dimension and the correlation entropy are invariants that characterize the natural measure of a dynamical system [13]. These invariants can be easily computed from the correlation integral which is obtained from indirect temporal measures of one of the variables of the system [9]. For an observed stationary time series of length N , $\{x_n\}_{n=1}^N$, the reconstructed m -dimensional delay vectors $\mathbf{x}_i = (x_i, x_{i+\tau}, \dots, x_{i-(m-1)\tau})$, $i = 1, 2, \dots, L = N - (m-1)\tau$, must be formed. The correlation integral $G_m(h)$, in its general form, is defined as the probability that the distance $\bar{z} = \|\mathbf{x}_i - \mathbf{x}_j\|$ between two randomly selected m -dimensional delay vectors \mathbf{x}_i and \mathbf{x}_j is smaller than a value h [13]:

$$G_m(h) = \int g(h, \bar{z}) f_{\bar{z}}^m(\bar{z}) d\bar{z}, \quad (1)$$

where $g(h, \bar{z})$ is a kernel function and $f_{\bar{z}}^m(\bar{z})$ is the probability density function (pdf) of the distance between pairs of delay vectors. In this article, we used the Euclidean distance.

The Grassberger and Procaccia correlation integral $C_m(h)$ uses as a kernel function $g(h, \bar{z}) = H(1 - \bar{z}/h)$, where $H(\cdot)$ is the Heaviside step function [9]. On the other hand, the GCI proposed by Diks et al., $T_m(h)$, adopts $g(h, \bar{z}) = \exp(-\bar{z}^2/4h^2)$ [11]. The main advantage of the GCI over the Grassberger and Procaccia correlation integral is that the former allows us to model the influence of additive noise on the scaling law.

For a zero-mean time series of variance σ_c^2 with additive white Gaussian noise of variance σ_n^2 , the scaling rule is [11, 14, 15]

$$T_m(h) \sim h^m (h^2 + \sigma^2)^{(D-m)/2} e^{-m\pi K_2} \quad (2)$$

for $m \rightarrow \infty$, $\sqrt{h^2 + \sigma^2} \rightarrow 0$,

where

$$\sigma = \frac{\sigma_n}{\sqrt{\sigma_c^2 + \sigma_n^2}}. \quad (3)$$

It is important to say that all time series used in this work were rescaled to have unitary standard deviation. In this sense, σ is the noise level after rescaling; that is, $\sigma \rightarrow 0$ corresponds to clean time series and $\sigma \rightarrow 1$ implies a pure stochastic process. The signal-to-noise ratio (SNR) can be calculated as

$$\text{SNR} = 10 \log_{10} \left(\frac{1 - \sigma^2}{\sigma^2} \right) \text{dB}. \quad (4)$$

In practical applications, it is very important to be able to quantify σ , the noise level in the time series. This is because it allows us to correct the estimations of D and K_2 . Moreover, the conclusions driven by the interpretation of the invariants

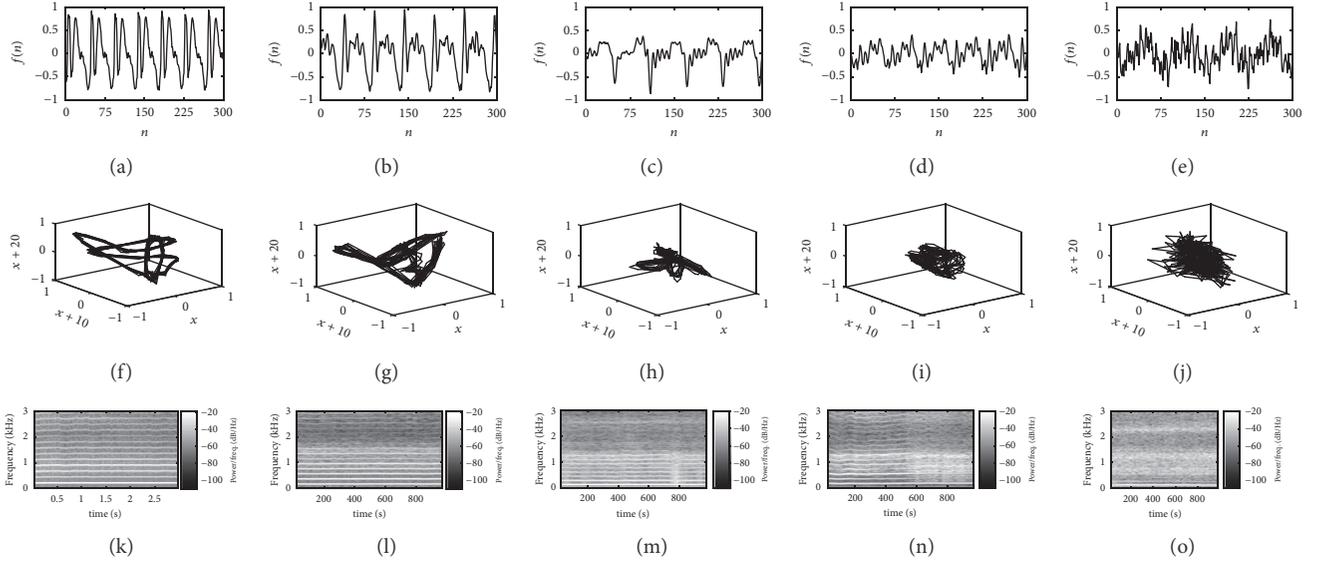


FIGURE 1: Time series, state space reconstruction, and spectrogram of normal and pathological voices. First column: normal voice; second column: pathological type 1 voice; third column: pathological type 2 voice; fourth column: pathological type 3 voice; fifth column: pathological type 4 voice.

calculated in the presence of high levels of noise should be taken carefully. For this reason, an estimate of the noise level must be also reported, allowing the readers to be aware of the limitations of the estimates. On the other hand, when the data is taken under controlled conditions, like the voices in the database analyzed herein, a high noise level can be seen as an indicator of a dominant underlying additive stochastic process.

The main disadvantage of the GCI is that it requires high values of m to obtain a convergent estimate of K_2 [10, 15]. This is because the pdf of the interpoint distance changes with the embedding dimension, whereas the GCI's kernel does not (see (1)). As a result, the convergence of the GCI to K_2 is slowed down [12]. The last statement also applies to the Grassberger and Procaccia correlation integral. As a solution, we have recently proposed the U-correlation integral (UCI) [12]:

$$U_m^\beta(h) = \int \frac{\Gamma(\beta/2, z/h^2)}{\Gamma(\beta/2)} \hat{f}_m(\sigma; z) dz, \quad (5)$$

where $\Gamma(a, t)$ is the upper incomplete Gamma function, $\Gamma(a)$ is the Gamma function, and $\hat{f}_m(\sigma; z)$ is the pdf of the squared interpoint distance. There are two important aspects about this correlation integral that deserve to be mentioned. First, note that the UCI's kernel function, given by

$$g(h, \beta, z) = \frac{\Gamma(\beta/2, z/h^2)}{\Gamma(\beta/2)}, \quad (6)$$

has a parameter β that is used to incorporate information about the embedding dimension. In other words, this kernel function is able to change according to m . Second, we are now working with the square of the interpoint distance, that is, $z = \bar{z}^2$, to reduce the computational cost.

The scaling law for the UCI is [12]

$$U_m^\beta(h) = \frac{\hat{\phi}}{2} (2\sigma)^D \frac{\Gamma(D/2) \Gamma((\beta+m)/2)}{\Gamma(\beta/2) \Gamma(m/2+1)} \cdot e^{-m\tau K_2} \left(\frac{h^2}{4\sigma^2} \right)^{m/2} \cdot F\left(\frac{\beta+m}{2}, \frac{m-D}{2}; \frac{m+2}{2}; -\frac{h^2}{4\sigma^2} \right), \quad (7)$$

where $\hat{\phi}$ is a normalization constant and $F(a, b; c; t)$ is the Gauss hypergeometric function. The UCI is approximated with the U-correlation sum $\widehat{U}_m^\beta(h)$ which, in this article, is calculated as follows: first, we obtain the squared distances between each pair of m -dimensional delay vectors $z_\omega = \|\mathbf{x}_i - \mathbf{x}_j\|^2$, where $\omega = \{(i, j) / i = 1, 2, \dots, L, j = 1, 2, \dots, L, i \neq j\}$, and then

$$\widehat{U}_m^\beta(h) = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} g(h, \beta, z_\omega) = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} \frac{\Gamma(\beta/2, z_\omega/h^2)}{\Gamma(\beta/2)}, \quad (8)$$

where $\Omega = L(L-1)$. The evaluation of the function $g(h, \beta, z_\omega)$ requires numerical integration which is computationally expensive. Instead, we performed the evaluation through spline interpolation of the function $\hat{g}(t) = \Gamma(\beta/2, t)/\Gamma(\beta/2)$, where $t \in [10^{-6}, 10^2]$ for each value of β .

Once the correlation sum is calculated, we estimated D , K_2 , and σ using coarse-grained estimators. These are explicit expressions for D , K_2 , and σ as functions of m and h [13]. Such functions are useful because they allow us to estimate the invariants and to visually confirm a scaling regime. In [16], we have presented three coarse-grained functions: $D_m^U(h)$, $K_m^U(h)$, and $\sigma_m^U(h)$ which are based on the UCI.

TABLE 1: Analyzed subjects: number and age.

Condition	Number		Age (years) mean \pm SD	
	Male	Female	Male	Female
Healthy	21	32	38.8 \pm 8.5	34.1 \pm 7.8
Pathologic	67	96	41.2 \pm 9.2	37.6 \pm 8.3

As an advantage, they can be calculated in an automated manner from two U-correlation integrals. Moreover, we have proposed an automated algorithm to estimate D , K_2 , and σ from these coarse-grained estimators [16].

2.1. Coarse-Grained Estimators Based on UCI. Here, we present the coarse-grained estimators used in this study. A more detailed description can be found in [12, 16].

We define the noise level functional as [12]

$$\Delta_m^U(h) = \frac{1}{2} [\dot{D}_{m+2}^{\beta=m}(h) - \dot{D}_m^{\beta=m}(h)] \approx \frac{4\sigma^2}{h^2 + 4\sigma^2} \quad (9)$$

for $m \gg D$,

where

$$\dot{D}_m^\beta(h) = \frac{d \ln U_m^\beta(h)}{d \ln h}. \quad (10)$$

This functional is the difference between the logarithmic derivatives of two U-correlation integrals: $U_m^{\beta=m}(h)$, which is the UCI calculated with squared distances z_ω coming from m -dimensional delay vectors and a kernel with the parameter $\beta = m$, and $U_{m+2}^{\beta=m}(h)$. The last one is the correlation integral obtained from distances between pairs of $(m+2)$ -dimensional delay vectors and a kernel with the parameter $\beta = m$.

We have shown in [12] that $\Delta_m^U(h)$ is a function that decreases monotonically from 1 to 0 on a scale proportional to the value of σ . From this noise functional, we can define the noise level coarse-grained estimator as [12]

$$\sigma_m^U(h) = \frac{h}{2} \sqrt{\frac{\Delta_m^U(h)}{1 - \Delta_m^U(h)}}. \quad (11)$$

The coarse-grained estimators for D and K_2 can be read as [16]

$$D_m^U(h) = \dot{D}_m^{\beta=m}(h) + \frac{\Delta_m^U(h)}{1 - \Delta_m^U(h)} \left[\dot{D}_m^{\beta=m}(h) + 2(m - 1) \left(\frac{U_m^{\beta=m-2}(h)}{U_m^{\beta=m}(h)} - 1 \right) \right], \quad (12)$$

$$K_m^U(h) = -\frac{1}{2\tau} \left\{ \ln \left(\frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} \right) + \ln \left(\frac{D_m^U(h)}{m} + 1 \right) \right\} + \frac{1}{2\tau} \cdot \ln \left[\Delta_m^U(h) \left(\frac{m - \dot{D}_m^{\beta=m}(h)}{m - D_m^U(h)} \right) + (1 - \Delta_m^U(h)) \left(\frac{\dot{D}_m^{\beta=m}(h)}{m} + 1 \right) \right], \quad (13)$$

respectively.

The calculation of $D_m^U(h)$ requires the estimation of $\Delta_m^U(h)$ and two correlation integrals: $U_m^{\beta=m}(h)$ and $U_m^{\beta=m-2}(h)$. On the other hand, $K_m^U(h)$ depends on the correlation integrals $U_{m+2}^{\beta=m+2}(h)$ and $U_m^{\beta=m}(h)$, $\Delta_m^U(h)$, the logarithmic derivative $\dot{D}_m^{\beta=m}(h)$, and the coarse-grained estimator $D_m^U(h)$.

2.2. Database. This study was conducted using the *Massachusetts Eye & Ear Infirmary* (MEEI) Voice Disorders Database, distributed by *Kay Elemetrics* [17]. We have chosen a subset of 219 records (53 normal and 166 pathological) of sustained phonation of vowel /a/ [18]. Each record was down-sampled to 10 kHz and the subjects' descriptive statistics are presented in Table 1. From this set of voices, we excluded 3 records because they contained less than 8000 sample points, ending with a total of 216 records to be analyzed.

For each record, a centered window of $N = 8000$ data points was selected and normalized to have zero mean and unitary standard deviation. Then, the U-correlation sums were computed for $m = \{4, 6, \dots, 20\}$, $\tau = 10$, and $h \in [e^{-5}, e^2]$. Moreover, the nearest 10 temporal neighbors of each delay vector were discarded [7]. Finally, Algorithm 1 was applied. We must clarify that the range of values for the embedding dimension was selected taking into account the studies conducted in [6, 19–21], where a low-dimensional vocal system is suggested ($D < 5$ for pathological voices). On the other hand, the embedding lag was selected as the average lag (over all voices) where the first minimum of the mutual information function occurs [7, 10].

In a preliminary study, the pathological records were classified into Titze's scheme based on the visual observation of the time series and its spectrograms. The classification resulted in 74 type 1 records and 81 type 2 records. The remainder 11 voices were taken apart since these techniques cannot differentiate between type 3 and type 4 voices.

- (1) Calculate $U_m^{\beta=m}(h)$ and $U_m^{\beta=m-2}(h)$ using Eq. (8) for $m > 2$.
- (2) Calculate $\Delta_m^U(h)$ using Eq. (9) and the UCIs obtained in step (1).
- (3) Compute $\sigma_m^U(h)$ with Eq. (11) and obtain the function $F_\sigma(h)$. Estimate σ within a range of h centered at the value h where $F_\sigma(h)$ is minimum.
- (4) Use $U_m^{\beta=m}(h)$, $U_m^{\beta=m-2}(h)$ and $\Delta_m^U(h)$ to calculate $D_m^U(h)$ (Eq. (12)) and obtain the function $F_D(h)$. Estimate D within a range of h centered at the value h where $F_D(h)$ is minimum.
- (5) Obtain $K_m^U(h)$ (Eq. (13)) using $U_m^{\beta=m}(h)$, $\Delta_m^U(h)$ and $D_m^U(h)$. Calculate the function $F_{K_2}(h)$. Estimate K_2 within a range of h centered at the value h where $F_{K_2}(h)$ is minimum.

ALGORITHM 1: Automated estimation of attractors' invariants.

Following the definition of type 4 voices given by Sprecher et al. in [2], we calculated the coarse-grained estimators $D_m^U(h)$ and $K_m^U(h)$ for these 11 voices. Then, we grouped together the voices that did show a scaling regime (finite correlation dimension and entropy) as type 3 and those that did not as type 4.

3. Results

The first result of the simulations can be observed in Figure 2. It shows the coarse-grained estimators $\sigma_m^U(h)$, $D_m^U(h)$, and $K_m^U(h)$ for a normal voice (Figures 2(a), 2(f), and 2(k)), a pathological type 1 voice (Figures 2(b), 2(g), and 2(l)), a pathological type 2 voice (Figures 2(c), 2(h), and 2(m)), a pathological type 3 voice (Figures 2(d), 2(i), and 2(n)), and a pathological type 4 voice (Figures 2(e), 2(j), and 2(o)).

It is important to observe that all coarse-grained estimators present a scaling region, except the ones for correlation dimension and K_2 entropy of type 4 voices. This highlights the suitability of these estimators to analyze normal and pathological (type 1, type 2, and type 3) voices.

As it can be observed in the first row of Figure 2, the noise level is greater for the pathological voices than for the normal voice and, in the pathological case, it increases along with the type. This means that, for both normal and pathological voices, there is an underlying stochastic component and its level increases in the presence of pathology.

The behavior of the estimator $D_m^U(h)$ is presented in the second row of Figure 2. As it can be seen, this estimator suggests a value of $D \approx 1.25$ for normal voices and slightly greater values for pathological type 1 and type 2 voices. However, it is difficult to say whether there is any difference between the analyzed normal and pathological voice. Nevertheless, these results suggest that the voice production system has a relatively low dimension.

Regarding the estimator $K_m^U(h)$, it can be observed from the third row of Figure 2 that, for the voices analyzed herein, the estimator converges to values close to zero. This suggests the presence of a strong harmonic component. Furthermore, there exists a small increase of K_2 entropy from normal voice to pathological type 1 voice and from the latter to pathological type 2 voice. This reflects an increasing degree of irregularity which is often associated with an increase in complexity.

For the type 3 voice, it can be observed in the fourth column of Figure 2 that the noise level coarse-grained estimator converges to a higher value than the one corresponding to the type 2 voice. However, it is still possible to observe a scaling range in both $D_m^U(h)$ and $K_m^U(h)$. Moreover, these estimators converge to higher values than the ones for type 2 voice. On the other hand, for the type 4 voice (fifth column of Figure 2), it is not possible to find a scaling range either in dimension or in the entropy estimator. This behavior is expected since, by definition, these types of voices have an infinite value of dimension and entropy.

As we mentioned before, this kind of nonlinear techniques should not be used with type 4 voices. Note that, from the curves of $D_m^U(h)$ and $K_m^U(h)$ (Figures 2(j) and 2(o), resp.), it is not easy to decide whether there is a scaling range. In this sense, an untrained person could erroneously determine its existence, resulting in misleading estimations.

On the other hand, a scaling range can be found for $\sigma_m^U(h)$ (Figure 2(e)), and it converges to a very high noise level value $\sigma \approx 0.5$ (SNR ≈ 4.77 dB). This suggests that the dynamics are mostly ruled by the underlying stochastic component. It is important to mention that this behavior is consistent in all voices that did not have a scaling region on $D_m^U(h)$ and $K_m^U(h)$ (type 4 voices). Based on this result, we suggest that a high value of σ could be a quantitative indicator of type 4 voices.

In order to find a noise level value that allows us to differentiate between type 3 and type 4 voices, we select by visual inspection all voices that did not show a scaling region on the estimators $D_m^U(h)$ and $K_m^U(h)$. Then, for each signal, we estimate the noise level value from $\sigma_m^U(h)$ and select the minimum of these values ($\sigma_{\text{thr}} = 0.4$) as a threshold.

3.1. Automated Estimation of Invariants. In order to obtain reliable estimations of D , K_2 , and σ , it is essential to verify the existence of a scaling behavior over the coarse-grained estimators. Once it is found, one has to choose a range of h values from which to estimate the invariants. This choice is critical since, in practical application, the coarse-grained estimators strongly vary as a function of the scale. To avoid subjective judgements, we have proposed an algorithm (see Algorithm 1) for the automated estimation of D , K_2 , and σ based on the coarse-grained estimators $D_m^U(h)$, $K_m^U(h)$, and $\sigma_m^U(h)$, respectively [16].

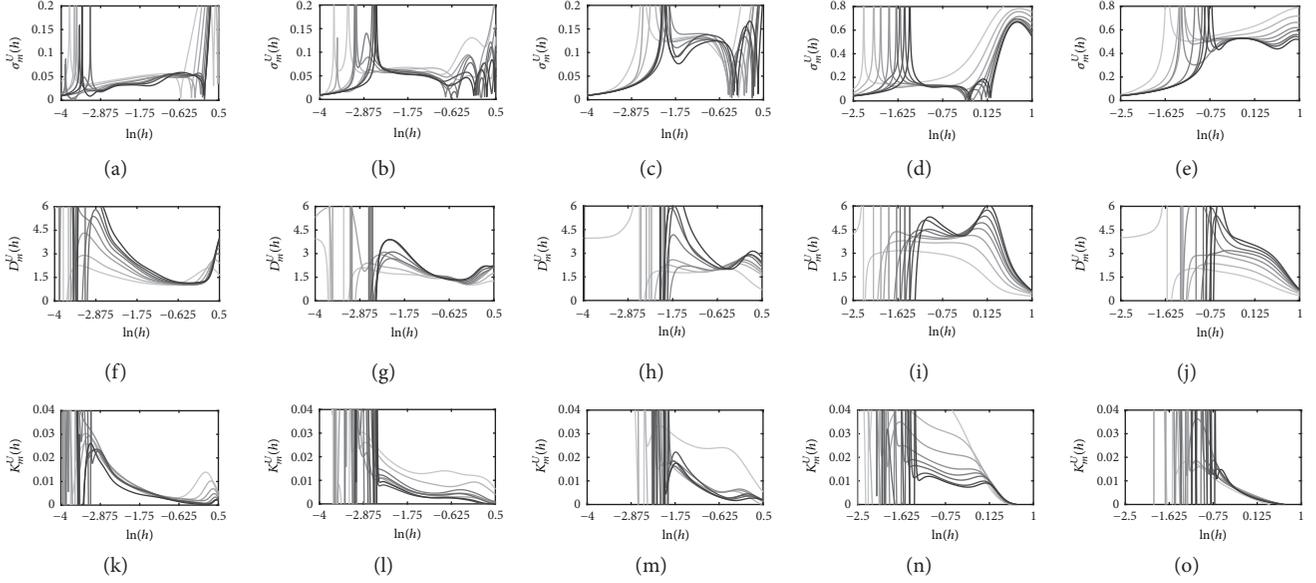


FIGURE 2: Coarse-grained estimators $\sigma_m^U(h)$, $D_m^U(h)$, and $K_m^U(h)$ for normal and pathological voices. First column: normal voice; second column: pathological type 1 voice; third column: pathological type 2 voice; fourth column: pathological type 3 voice; fifth column: pathological type 4 voice. The curves for $m = \{4, 6, \dots, 20\}$ are color-coded in grayscale where the lightest gray corresponds to $m = 4$.

This algorithm selects the scaling range where those invariants should be estimated based on the next criteria: (i) the coarse-grained estimator must be constant for a range of scale values and (ii) the value of the invariant should converge as the embedding dimension increases.

The algorithm begins by approximating the U-correlation integrals $U_m^{\beta=m}(h)$ and $U_m^{\beta=m-2}(h)$ for different m values ($m > 2$) using (8). Note that $U_m^{\beta=m-2}(h)$ must be calculated for $m > 2$ since the shape parameter of the incomplete Gamma and Gamma functions must be greater than zero. Next, the logarithmic derivatives $\dot{D}_m^{\beta=m}(h)$ and $\dot{D}_{m+2}^{\beta=m}(h)$ must be computed in order to obtain $\Delta_m^U(h)$ (see (9)). Observe that the correlation integral $U_{m+2}^{\beta=m}(h)$ is equal to the correlation integral $U_{\widehat{m}}^{\beta=\widehat{m}-2}(h)$ evaluated at $\widehat{m} = m + 2$. In this article, the logarithmic derivatives were obtained using a wavelet transform approach [22]. This produces a smooth version of the derivatives allowing us to better estimate the invariants.

The noise level must be calculated from a range of h where the coarse-grained estimator $\sigma_m^U(h)$ is nearly constant and its variation across the m values is the smallest. In this sense, for the noise level, we define the functions [16]

$$\begin{aligned}
 A_\sigma(h) &= \frac{1}{M} \sum_{i=1}^M \frac{d\sigma_{m_i}^U(h)}{d \ln h}, \\
 V_\sigma(h) &= \frac{1}{M-1} \sum_{i=1}^M (\sigma_{m_i}^U(h) - \bar{\sigma}^U(h))^2, \\
 F_\sigma(h) &= A_\sigma(h) V_\sigma(h),
 \end{aligned} \tag{14}$$

where $m \in \{m_1, m_2, \dots, m_i, \dots, m_M\}$ and $\bar{\sigma}^U(h)$ is the average of $\sigma_m^U(h)$ across m . $A_\sigma(h)$ is the average over m of the derivative of $\sigma_m^U(h)$ with respect to $\ln h$, $V_\sigma(h)$ gives the variation of $\sigma_m^U(h)$ across m , and $F_\sigma(h)$ is the product of the two aforementioned functions. We propose to estimate σ within a range of h centered at the h value at which $F_\sigma(h)$ is minimum. This way, σ is estimated in a range of h centered in a plateau region (scaling region) of $\sigma_m^U(h)$, and its value is consistent through the parameter m .

The correlation dimension and the correlation entropy can be determined using the coarse-grained estimators $D_m^U(h)$ (see (12)) and $K_m^U(h)$ (see (13)), respectively. To find a range of h values to estimate D and K_2 , we use a similar approach, but from functions $F_D(h)$ and $F_{K_2}(h)$, respectively. The functions $F_D(h)$ and $F_{K_2}(h)$ can be computed similarly to $F_\sigma(h)$ but using the coarse-grained estimators $D_m^U(h)$ and $K_m^U(h)$, respectively.

The estimation of each invariant for the whole group of signals is presented in Figure 3. Figure 3(a) shows a box plot of the noise level estimation for normal (N), pathological type 1 (PT1), pathological type 2 (PT2), pathological type 3 (PT3), and pathological type 4 (PT4) voices. As it can be observed, σ is greater for pathological voices than for normal ones. However, it is not possible to establish a statistical difference between normal and pathological type 1 and type 2 voices. On the other hand, we can differentiate normal from pathological type 3 and type 4 voices. It is very important to say that the group labeled as TP4 was selected because they have a noise level value $\sigma > 0.4$. Note that, using this threshold, we can separate type 3 from type 4 voices.

The box plot of the correlation dimension is presented in Figure 3(b). There is an increase of D from normal to

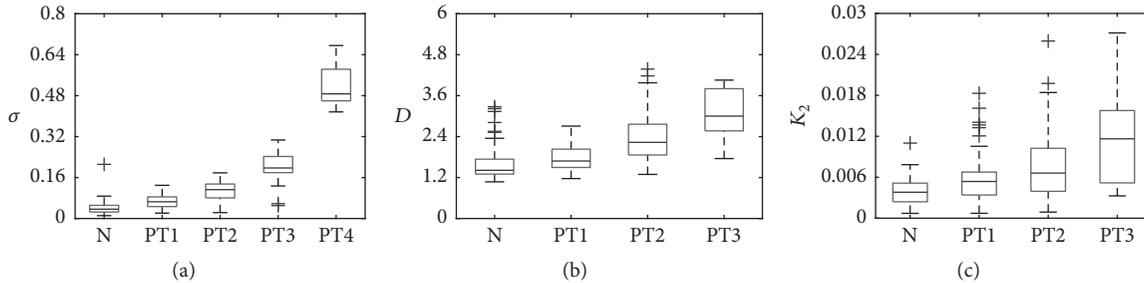


FIGURE 3: Invariant estimation box plot: (a) noise level, (b) correlation dimension, and (c) correlation entropy. Normal (N), pathological type 1 (PT1), pathological type 2 (PT2), pathological type 3 (PT3), and pathological type 4 (PT4) voices.

pathological voices. For normal voices, the median is $D = 1.4$; for pathological type 1 voices, $D = 1.68$; for type 2 voices, $D = 2.23$; and for type 3 voices, $D = 2.99$. From these results, we can establish that the vocal system has a low dimension, even in the presence of a pathology. The group TP4 is not shown in this plot since, by definition, its dimension is infinite. Regarding the correlation entropy, it is shown in Figure 3(c) that there is a small increase of K_2 from normal to pathological voices. However, it is not possible to find a statistical difference between normal, pathological type 1, and pathological type 2 voices. The median correlation entropy estimated for normal voices is $K_2 = 0.0038$, for pathological type 1 voices is $K_2 = 0.0053$, for pathological type 2 voices is $K_2 = 0.0068$, and for pathological type 3 voices is $K_2 = 0.012$.

4. Discussion

Bearing in mind that the used records were taken in a controlled environment, we could associate the noise level with the strength of an additive stochastic component that coexists with the dynamic producing the voice. There is a tendency of σ to increase its value from normal voices to pathological ones. This can be seen as an increase of the power of the stochastic component caused by the presence of a pathology. Moreover, in pathological voices, σ increases its value along with the voice type.

The definition of type 4 voices given by Sprecher et al. [2] states that these signals are characterized by pure stochastic oscillations; therefore, their dimension is infinite [10]. In practice, it is not possible to measure an infinite correlation dimension value since it is bounded by the embedding dimension used to calculate the correlation integral. Instead, an infinite correlation dimension is inferred if there is not a scaling regime in the coarse-grained estimator of D . In previous studies, this was done through visual inspection, which is always a subjective judgement [2, 19].

One interesting aspect of the coarse-grained estimator $\sigma_m^U(h)$ is that, in the type 4 signals analyzed here, it always presents a scaling range from which to estimate σ , although $D_m^U(h)$ and $K_m^U(h)$ have no scaling regions. These results led us to think that the noise level can be used as an objective measure to discriminate between type 3 and type 4 voices.

The threshold proposed here ($\sigma_{\text{thr}} = 0.4$) was set using the voice signals that did not present a scaling regime in $D_m^U(h)$ and $K_m^U(h)$. With this threshold, we were able to well separate type 3 and type 4 voices. Nevertheless, we are aware that this threshold was selected based on observations from this database. This finding must be validated with a more extensive study involving a larger number of records and voice care professionals.

Our estimations of correlation dimension are in concordance with other studies [6, 19–21]. In [23], Choi et al. conducted a very similar simulation over the Kay Elemetrics database. They reported a mean value of $D = 1.57$ for normal voices and an increased dimension value for pathological ones. Moreover, they obtained a decreasing SNR (it was calculated according to [24]) with the type of the voice, being the lowest for type 3 voices. However, they did not analyze type 4 voices. Another study by Zhang and Jiang conducted over the Kay Elemetrics database reported a mean correlation dimension of $D = 1.51$ for normal voices and $D = 3.17$ for a group with vocal tremor [25]. As far as we know, all researches conducted on normal and pathological voices have used the Grassberger and Procaccia correlation integral to estimate D and K_2 . This methodology has the disadvantage that its estimations are sensitive to noise presence, requires large m values to converge, and does not give an estimation of the noise level [10, 13]. However, other variants like the Gaussian correlation integral have not been used.

Regarding the correlation entropy, our results suggest that pathological type 1 and type 2 voices have a slightly greater value than normal ones. Furthermore, these three types of voices have a K_2 value close to zero, suggesting nearly periodic dynamics. For type 3 voices, the values of K_2 are the greatest, meaning more irregular and unpredictable dynamics. These values are comparable with those presented by Yan et al. [26]. They reported an estimation of K_2 for normal (mean value $K_2 = 0.014$) and esophageal phonation (mean value $K_2 = 0.023$) subjects. In [27], Calawerts et al. calculated the largest Lyapunov exponent for type 1, type 2, and type 3 voices using the Kay Elemetrics database. Their results suggest an increase of the value of the largest Lyapunov exponent along with the type of the voice. Our results are in concordance with this study since we obtain an estimation of correlation entropy that not only is greater for pathological

voices than for normal ones but also increases with the type of the voice.

5. Conclusions

In this article, we have studied normal and pathological voices through the correlation dimension, the correlation entropy, and the noise level. These invariants were estimated using an automated algorithm based on coarse-grained estimators derived from the U-correlation integral. The results suggest that the voice production dynamical system has a low dimension. The value of D is greater for pathological voices than for normal ones. Moreover, its value also increases along with the type of the voice. Regarding the correlation entropy, its value is very low for normal and for type 1 and type 2 pathological voices. Although a more extensive study is still needed, this finding suggests that the system dynamics have a harmonic oscillatory behavior. On the other hand, pathological type 3 voices present higher values of K_2 , implying a more complex behavior which is reflected in a more irregular dynamic. The noise level can be interpreted as the power of a stochastic perturbation intrinsic to the voice production system. Our results show that σ is greater for pathological voices than for normal ones. Furthermore, it increases along with the type of voice, being the highest for type 4 voices. This means that the presence of the stochastic component is stronger in pathological voices. Based on these results, in this work, we have proposed a quantitative criterion that can be used to differentiate between type 3 and type 4 voices. We are aware of the limitation of the reduced samples of type 3 and type 4 voices. In this sense, these preliminary results will be validated with a more extensive study involving a larger number of records and voice care professionals.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Scientific and Technical Research Council (CONICET) of Argentina, the National University of Entre Ríos (UNER), IBB, and Grants PID-6136 (UNER) and PIO-14620140100014CO (CITER-CONICET). The authors would like to thank Professor Cristina Jackson-Menaldi, Ph.D., and Juliana Codino, M.S., CCC-SLP, for the fruitful discussions.

References

- [1] I. R. Titze, *Workshop on Acoustic Voice Analysis: Summary Statement*, vol. 24, National Center for Voice and Speech, 1995.
- [2] A. Sprecher, A. Olszewski, J. J. Jiang, and Y. Zhang, "Updating signal typing in voice: Addition of type 4 signals," *The Journal of the Acoustical Society of America*, vol. 127, no. 6, pp. 3710–3716, 2010.
- [3] I. R. Titze, R. J. Baken, H. Herzel, and P. Arts, "Evidence of chaos in vocal fold vibration," in *Vocal fold physiology: new frontiers in basic science*, vol. 1, pp. 143–188, 1993.
- [4] H. Herzel, D. Berry, I. R. Titze, and M. Saleh, "Analysis of vocal disorders with methods from nonlinear dynamics," *Journal of Speech, Language, and Hearing Research*, vol. 37, no. 5, pp. 1008–1019, 1994.
- [5] A. Giovanni, M. Ouaknine, B. Guelfucci, P. Yu, M. Zanaret, and J.-M. Triglia, "Nonlinear behavior of vocal fold vibration: The role of coupling between the vocal folds," *Journal of Voice*, vol. 13, no. 4, pp. 465–476, 1999.
- [6] J. J. Jiang, Y. Zhang, and C. McGilligan, "Chaos in voice, from modeling to measurement," *Journal of Voice*, vol. 20, no. 1, pp. 2–17, 2006.
- [7] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, 2nd edition, 2004.
- [8] J.-P. Eckmann and D. Ruelle, "Ergodic theory of chaos and strange attractors," *Reviews of Modern Physics*, vol. 57, no. 3, part 1, pp. 617–656, 1985.
- [9] P. Grassberger and I. Procaccia, "Characterization of strange attractors," *Physical Review Letters*, vol. 50, no. 5, pp. 346–349, 1983.
- [10] M. Small, *Applied nonlinear time series analysis: applications in physics, physiology and finance*, World Scientific, vol. 52, 2005.
- [11] C. Diks, "Estimating invariants of noisy attractors," *Physical Review E: Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 53, no. 5, pp. R4263–R4266, 1996.
- [12] J. F. Restrepo and G. Schlotthauer, "Noise-assisted estimation of attractor invariants," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 94, no. 1, Article ID 012212, 2016.
- [13] C. Diks, *Nonlinear time series analysis*, vol. 4 of *Nonlinear Time Series and Chaos*, World Scientific Publishing Co., Inc., River Edge, NJ, 1999.
- [14] D. Yu, M. Small, R. G. Harrison, and C. Diks, "Efficient implementation of the Gaussian kernel algorithm in estimating invariants and noise level from noisy time series data," *Physical Review E: Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 61, no. 4, pp. 3750–3756, 2000.
- [15] G. Nolte, A. Ziehe, and K.-R. Müller, "Noise robust estimates of correlation dimension and k_2 entropy," *Physical Review E: Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 64, no. 1, p. 10, 2001.
- [16] J. F. Restrepo and G. Schlotthauer, "Automatic estimation of attractor invariants," *Nonlinear Dynamics*, vol. 91, no. 3, pp. 1681–1696, 2018.
- [17] M. Eye and E. Infirmary, *Voice disorders database, version. 1.03 (cd-rom)*, Kay Elemetrics Corp, Lincoln Park, NJ.
- [18] V. Parsa and D. G. Jamieson, "Identification of pathological voices using glottal noise measures," *Journal of Speech, Language, and Hearing Research*, vol. 43, no. 2, pp. 469–485, 2000.
- [19] A. Behrman and R. J. Baken, "Correlation dimension of electroglottographic data from healthy and pathologic subjects," *The Journal of the Acoustical Society of America*, vol. 102, no. 4, pp. 2371–2379, 1997.
- [20] G. Vaziri, F. Almasganj, and R. Behroozmand, "Pathological assessment of patients' speech signals using nonlinear dynamical analysis," *Computers in Biology and Medicine*, vol. 40, no. 1, pp. 54–63, 2010.
- [21] N. Yan, M. L. Ng, D. Wang, L. Zhang, V. Chan, and R. S. Ho, "Nonlinear dynamical analysis of laryngeal, esophageal, and tracheoesophageal speech of cantonese," *Journal of Voice*, vol. 27, no. 1, pp. 101–110, 2013.
- [22] J. Luo, J. Bai, and J. Shao, "Application of the wavelet transforms on axial strain calculation in ultrasound elastography," *Progress*

- in Natural Science: Materials International*, vol. 16, no. 9, pp. 942–947, 2006.
- [23] S. H. Choi, Y. Zhang, J. J. Jiang, D. M. Bless, and N. V. Welham, “Nonlinear dynamic-based analysis of severe dysphonia in patients with vocal fold scar and sulcus vocalis,” *Journal of Voice*, vol. 26, no. 5, pp. 566–576, 2012.
- [24] P. Milenkovic, “Least mean square measures of voice perturbation,” *Journal of Speech, Language, and Hearing Research*, vol. 30, no. 4, pp. 529–538, 1987.
- [25] Y. Zhang and J. J. Jiang, “Nonlinear dynamic mechanism of vocal tremor from voice analysis and model simulations,” *Journal of Sound and Vibration*, vol. 316, no. 1-5, pp. 248–262, 2008.
- [26] N. Yan, M. L. Ng, D. Wang, V. Chan, and L. Zhang, “Nonlinear dynamics of voices in esophageal phonation,” in *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2011*, pp. 2732–2735, USA, September 2011.
- [27] W. M. Calawerts, L. Lin, J. C. Sprott, and J. J. Jiang, “Using Rate of Divergence as an Objective Measure to Differentiate between Voice Signal Types Based on the Amount of Disorder in the Signal,” *Journal of Voice*, vol. 31, no. 1, pp. 16–23, 2017.

Research Article

A Comparison between Theoretical and Experimental Measures of Consciousness as Integrated Information in an Anatomically Based Network of Coupled Oscillators

Antonio J. Ibáñez-Molina  and Sergio Iglesias-Parro 

Department of Psychology, University of Jaén, Jaén, Spain

Correspondence should be addressed to Antonio J. Ibáñez-Molina; anjoibanez@gmail.com

Received 14 September 2017; Revised 22 January 2018; Accepted 25 February 2018; Published 8 April 2018

Academic Editor: Gastón Schlotthauer

Copyright © 2018 Antonio J. Ibáñez-Molina and Sergio Iglesias-Parro. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The rise of mathematical developments in the theories of consciousness has led to new measures to detect consciousness in a system. The Integrated Information Theory (IIT) is one of the best mathematical rooted attempts to quantify the level of consciousness in a system with Φ as the effective information generated in a system above its parts. Recently, the IIT has inspired the Perturbational Complexity Index (PCI) to detect conscious states in patients with disorders of consciousness, and it has shown to have almost perfect classification accuracy. In this study, we explore the statistical correspondence between the theoretical Φ and the experimental PCI through a neurocomputational model of coupled oscillators that can be artificially perturbed, which mainly focuses on the dynamics of collective synchronization between subsets of brain areas. Our results reveal that both measures are statistically related but, in principle, this relationship is far to be perfect. These results are discussed in the context of the model of coupled oscillators, which mainly focuses on the dynamics of collective synchronization between subsets of brain areas.

1. Introduction

One of the most challenging and still in progress tasks in science is to objectively quantify to what extent a person is conscious. One possible reason is that consciousness itself cannot be operatively defined in an easy way since it is a subjective phenomenon, and hence, it cannot be directly observed [1]. There are different theoretical approaches to consciousness [2]. For example, from a philosophical perspective, it has been proposed that there is a high order thought associated with consciousness. A conscious thought would be composed of simple percepts associated with other thoughts that provide further semantic value. Within neuroscience fields, theories can be divided into biological and functionalistic theories. Biological theories state that consciousness is a biological state of the brain; consciousness is studied through the association of different cognitive states with particular brain regions. On the other hand, from a functionalistic perspective, the existence of consciousness only requires an abstract structure to exist. For example, in

silico structures could support conscious experiences as long as they obey the necessary conditions provided by the theory. A good example of functionalism is the Global Workspace account, which considers that conscious experience is the result of competition between functional networks in the brain [3] being conscious the winning network. However, within functionalistic accounts, the Integrated Information Theory (IIT) is probably the most solid and mathematically rooted attempt to define what consciousness might be and how it can be quantified [4–6]. Specifically, Tononi, in an early version of his theory [4], proposed that consciousness arises as integrated information in a system and provided Φ as a computable measure of how conscious any physical system might be. The main concept behind the calculation of Φ is to measure to what extent a system as a whole cannot be explained as a sum of its parts. As we will explain in the next section, in order to compute Φ , it is necessary to compare the information generated by the entire system with the information of the system considered as two subsystems (bipartition).

The main limitation in Φ computation is the necessary condition to find and calculate the information contained in all possible bipartitions of the system. This is a problem of complexity class NP, which cannot be exhaustively computed for systems with a large number of elements. Given that in the human brain the number of neurons is in the order of 10^{11} [7], even if we have the entire information from every single neuron, it would not be possible to obtain a value for Φ . The number of possible bipartitions for a brain would be in the order of a Stirling number of second kind that can be computed with $2^{n-1} - 1$ and would give a number of combinations in the order of $2^{10^{11}}$, which interestingly is much higher than the classic Eddington's number that estimates the amount of protons in the known universe ($\approx 10^{80}$) [8]. To solve this and other practical problems, a number of versions and estimators have been developed in different fields of research. One remarkable example was provided by Barrett and Seth [9] where they proposed a version of Φ for time series data. This version noted here as Φ_{TS} (TS stands for time series) can be applied to time series from a generative dynamical model, and it perfectly agrees with the IIT in its early version (and in the main concepts of the further versions).

In the work we present here, we will focus on this measure because it is theoretically well funded and easy to apply to time series data. However, Φ_{TS} shows the same limitation as the original Φ , which, as stated before, includes the computation of information from all possible bipartitions of the system, and hence, it is only possible to be applied to time series from systems with limited number of elements.

In a different line of research, one of the most important empirical estimators of Φ has been developed in the field of clinical assessment of unresponsive patients with disorder of consciousness. Casali et al. [10] presented a Perturbational Complexity Index (PCI) of integrated information to decide if a given patient is conscious or not. The PCI was designed to capture information and integration in the system. It quantifies the richness of information in the process of propagation of activity across the EEG channels right after discrete transcranial magnetic stimulation (TMS). The theoretical rationale of this measure is that a system with high Φ , when stimulated with TMS pulses, needs to show cortical propagation (reflecting integration of its elements) as well as diverse functional reactions at different areas of the cortex (differentiation between the elements \approx information). Given that Φ is related to information and integration in a system, PCI is proposed as an estimator of Φ for data collected from real patients.

Although PCI has been demonstrated to classify patients into conscious or unconscious with an almost perfect accuracy rate [10], it is an indirect measure of Φ and it is not clear whether it reflects the theoretical Φ calculated from the brain as a physical system. In this study, we explore the relationship between the experimental PCI and the theoretical Φ . To investigate this question, it would be necessary to obtain both measures from the same system, and this is not easy since the latter (Φ_{TS}) is designed for simple dynamical systems and the former (PCI) for real brains. The approach we take here is

to obtain Φ_{TS} from an accepted neurocomputational model of whole brain resting-state activity, that is, a variant of the Kuramoto of coupled oscillators [10–13], which can be artificially perturbed to simulate TMS pulses in order to obtain a PCI for the same model.

Henceforth, in this study, we will focus on two versions of Φ : (1) Φ_{TS} for simple dynamical systems and time series and (2) PCI estimator developed to measure the level of consciousness in patients with different disorders of consciousness. With this work we want to explore the possible relationship between these two measures, Φ_{TS} , theoretically well founded, and the PCI, with indisputable clinical results. To the best of our knowledge, this is the first work that directly addresses this problem.

2. Methods

To investigate the potential relationship between the experimental estimator of consciousness PCI and the more theoretical index Φ_{TS} calculated over the same system of Kuramoto oscillators, we followed the next steps.

We first designed a Kuramoto model to simulate resting dynamics of the cortex. Second, we calculated several realizations of the model and obtained Φ_{TS} from the model. Third, we perturbed the system to simulate TMS pulses and compute the PCI of the model. In the next three sections, we explain these steps in detail.

2.1. The Kuramoto Model to Simulate Resting-State Dynamics. A Kuramoto model can be defined as a set of coupled oscillators modelled as the evolution of its phases according to the following set of coupled delay differential equations:

$$\frac{d\theta_i}{dt} = \omega + k \sum_{j=1}^N a_{ij} c_{ij} \sin(\theta_j(t - \tau_{ij}) - \theta_i(t)), \quad (1)$$

$$i = 1, \dots, N,$$

where θ_i is the phase of the i th oscillator on its limit cycle and ω is its natural frequency in radians. The control parameter k is the global excitatory coupling strength, a parameter that scales all coupling strengths. N is the total number of oscillators. Importantly, c_{ij} is the connectivity strength between each pair of oscillators and τ_{ij} represents the time delays between these oscillators. Both adjacency matrices (connectivity strength and time delays) were obtained by Hagmann et al. [11] that defined the structure of a network coupled together according to human white matter tractography (in the work of Hagmann et al. following diffusion spectrum and MRI acquisitions, the segmented grey matter was partitioned into 66 anatomical regions according to anatomical landmarks. White matter tractography was used to determine which regions pairs were connected by putative white matter fiber tracts and to estimate their density and corresponding length, from which the structural connectivity and delays were obtained). $N = 66$ in the network of oscillators we use here (see Figure 1). In short, each oscillator represents a cortical region of the brain located in a three-dimensional space with different connections to all other



FIGURE 1: Spatial representation of the nodes that have been used in the present model. In red, the nodes whose connectivity was temporarily increased are indicated. PARC (paracentral lobe), PCUN (precuneus), and SP (superior parietal cortex). L: left and R: right.

oscillators. The term a_{ij} allows us to dynamically modify the structural connectivity. It is an important term in this study because it can generate perturbations in the model that aimed to simulate TMS. For a Kuramoto model, the degree of synchrony between oscillators is conveniently measured by an order parameter, $r(t)$, that satisfies

$$r(t) e^{i\psi(t)} = \frac{1}{N} \sum_{j=1}^N e^{i\theta_j(t)}, \quad (2)$$

where $0 \leq r(t) \leq 1$ measures the phase coherence or synchrony of the N oscillators population; i is the symbol for the imaginary operator; and ψ is the average phase of the collective [12, 13]. $r(t)$ indicates how coherent oscillators are in a given time and it qualifies if the phases of the collective are tightly clustered or widely distributed. We have chosen r instead of other possible order parameters, as the average phase ψ , because it does not describe with accuracy the oscillators' collective behavior (for a given ψ , for example, there are many possible phase distributions of oscillators).

According to several studies, the Kuramoto model shares dynamical similarities with resting-state brain functioning when it shows high metastability [14, 15]. This concept refers to high variance of $r(t)$ which in other words can be defined as the tendency of a system of oscillators to continuously migrate between a variety of transient synchronous states, allowing a dynamical organization between the elements of the network. The system continuously goes from ordered to disordered states [16]. Then, the values of the parameters in the model were selected so that the global dynamics showed high metastability. For $\omega = 60$ Hz (gamma rhythm), metastability was evident with $\hat{\tau} = 3$ ms, and $0.5 < k < 6.5$, with $\hat{\tau}$ as the mean value of τ_{ij} . Note that any change in $\hat{\tau}$ can be considered a change in the mean velocity of the conduction delays between oscillators, and other values of $\hat{\tau}$ with a different range of k produced similar behaviors. For example, we found approximately the same effects for $2.5 < \hat{\tau} < 5$.

The Kuramoto model was simulated for a wide range of k values. As indicated before, k represents the strength in the global connectivity of the model, and from a biological

point of view it could be seen as a parameter to characterize integration between oscillators. Each simulation consisted of a baseline of 65×10^3 ms. As in Cabral et al. [17], we used an Euler scheme in which the time step of numerical integrations was set to .1 ms.

It would be important to state that since no experimental data are provided here, our results are obtained for parameters of the model (k , r , and metastability) that have been shown, in previous works [18–20], to parallel key characteristics of brain functioning. Hence, excluding the parameter k , we did not manipulate the parameters in the coupled equations of the Kuramoto system.

2.2. Integrated Information in the Kuramoto Model. Integrated information was measured with the version Φ_{TS} and it is mainly based on the concept of effective information (φ) [9].

Let $X = [X^1, \dots, X^n]$ be a multivariate random variable that takes values in the space Ω_X . It is evident that the dimensions of Ω_X are the number of elements in the system that generates X^n . The effective information generated by a system in its current state X_t about the state $X_{t-\tau}$ with respect to a bipartition of it $B = \{M^1, M^2\}$ is defined by the mutual information generated by the entire system minus the sum of the mutual information of its parts in the bipartition:

$$\varphi[X, \tau, B] = I(X_{t-\tau}; X_t) - \sum_{k=1}^2 I(M_{t-\tau}^k, M_t^k). \quad (3)$$

Mutual information in bits can be calculated with the expression

$$I(X_{t-\tau}, X_t) = \sum_{x \in \Omega_X} P(X_{t-\tau}, X_t) \log_2 \frac{P(X_{t-\tau}, X_t)}{P(X_{t-\tau})P(X_t)}, \quad (4)$$

a measure that gives the average bits that can be predicted in X_t given the state $X_{t-\tau}$ [21]. The calculation of mutual information includes the calculation of probabilities and joint probabilities of any estate X_t and $X_{t-\tau}$. Integrated information Φ_{TS} is the effective information with respect to the minimum information bipartition (MIB):

$$\Phi_{TS}[X; \tau] = \varphi[X, \tau, B^{\text{MIB}}(X, \tau)], \quad (5)$$

where

$$B^{\text{MIB}} = \arg \min_B \left\{ \frac{\varphi[X, \tau, B]}{K(B)} \right\}, \quad (6)$$

with $\arg_B \min$ stating for “the minimum number in the set” and

$$K(B) = \min \{H(M^1), H(M^2)\} \quad (7)$$

is a normalization factor to correct excessive unbalanced bipartitions. H here stands for Shannon entropy.

To apply Φ_{TS} to the Kuramoto model described in the previous section, we divided the 66 regions of the original network into 6 clusters proposed by Hagmann et al. in the

original work [11]. This simplification allowed exploring all possible bipartitions in the system (the Stirling number for this case gives $2^{6-1} - 1 = 31$ possible bipartitions). In addition, and following [22], time series for each cluster were calculated as the synchrony between the oscillators belonging to that cluster ($r_c(t)$). Then, we characterized these series as synchronized or not synchronized by constructing new binary time series from each $r_c(t)$. In our study, we considered a synchronization threshold of $\gamma = 0.8$ and 1 was assigned for all $r_c(t) > \gamma$. We selected this value because it was the median of $r(t)$, and using the median for thresholding eliminates the possible influence of extreme values, due to its robust properties. In addition, there was a theoretical reason for this election. The value of $\gamma = 0.8$; it was the one used in [22], which would allow us to compare the results we obtained.

Finally, Φ_{TS} was calculated to these binary series and the result was taken as the complexity value of the Kuramoto model. We set $\tau = 150$ ms. Values of $\tau < 150$ ms gave negative estimations of Φ and for $\tau > 150$ ms the pattern of results for different k values did not change.

2.3. Estimation of PCI in the Kuramoto Model. In order to calculate an estimation of the PCI in the Kuramoto model, it was necessary to solve two problems in the simulation process. The first one was to perturb or stimulate the system from an external source (to emulate the effects of TMS), and the second difficulty we found was to calculate reliable ERPs for the final PCI calculation. The problem of the stimulation was easily solved since it has been done in other studies. For example, Hellyer et al. [23] simulated external stimulation to a similar Kuramoto model we present here by increasing the connectivity between some nodes of the network. Similarly, Ibáñez-Molina and Iglesias-Parro [24] stimulated another Kuramoto version by transiently increasing the connectivity of key oscillators in the network. Hence, in our study we followed these studies and perturbed the system by transient increases in the connectivity between six oscillators located in the parietal cortex (see Figure 1). We arranged this procedure by introducing $a_{ij} = 10$ for these oscillators during short periods of time of 5 ms. We randomly repeated this stimulation 15 times for each numerical integration of the model.

The next step was to build reliable ERPs with the resulting phases from the oscillators. To achieve this goal we simulated EEG series, and then, ERPs were calculated for each model by averaging the segments associated with each period of stimulation. We explain this procedure in detail in the next two sections.

2.3.1. EEG and ERPs Simulation. The EEG activity from 32 sensors was simulated for each condition in agreement with the following weighted sum of the activity in each oscillator:

$$x_k(t) = \sum_{j=1}^N w_{kj} \sin(\theta_j(t)), \quad k = 1, \dots, P, \quad (8)$$

where $x_k(t)$ is the time series from sensor k th and w_{kj} is the weighted contribution of source j th in sensor k th. Each w_{kj} was calculated using a standard forward model algorithm [25] applied according to the Talairach coordinates of the

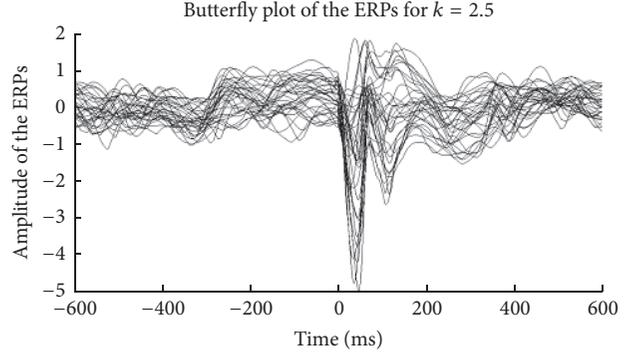


FIGURE 2: Butterfly plot from simulated ERPs with the model. In this case, the ERPs were produced for $k = 2.5$.

oscillators. After that, each oscillator was considered a cortical source. Second, the weights of these sources were normalized to a maximum value of 1. EEG signals obtained with this procedure gave a set of signals that changed in amplitude and frequency variations around 60 Hz (natural frequency of the oscillators). Because in the ERPs the interesting information is in the amplitude, we calculated the envelope of $x_k(t)$ with the Hilbert transform. Envelopes of the signal were then used to construct the ERPs with the average of all segments in each realization of the model. Formally, ERPs were built with the analytical signal (in the complex plane) of $x_k(t)$ which is

$$x_{ka}(t) = x_k(t) + j\widehat{x}_k(t) = a_k(t) e^{i\varphi_k(t)}, \quad (9)$$

where $x_k(t)$ is the original signal and represents the real part of the new complex series and $\widehat{x}_k(t)$ is the imaginary part from the Hilbert transform with i as the imaginary operator. In the right side of the expression, $\varphi_k(t)$ and $a_k(t)$ stand for the angle and modulus representing the complex values in Euler's notation. The modulus is the amplitude or the analytic power of the signal and can be easily calculated with

$$a_k(t) = [x_k(t)^2 + \widehat{x}_k(t)^2]^{1/2}. \quad (10)$$

These new series $a_k(t)$ were considered the activity from each sensor and the ERPs were built with segments extracted from them. For each sensor k ,

$$\text{ERP}_k = \frac{1}{N_s} \sum_{N_s} \mathbf{a}_k, \quad (11)$$

where N_s is the number of segments in a single realization of the model and \mathbf{a}_k is a vector with $a_k(t)$ values from -600 to 600 ms after stimulation. $N_s = 15$ for all numerical integrations (see Figure 2 for a visual inspection of ERPs).

2.3.2. PCI Calculation. ERPs were the input signal for PCI calculation. PCI was obtained following the original algorithm in Casali et al. [10]. For each condition of the Kuramoto model, first, we built a binary source matrix with dimensions corresponding to sensors (k) and time steps (t) from 0 to 1200 ms after stimulus presentation ($SS(k, t)$). The signals

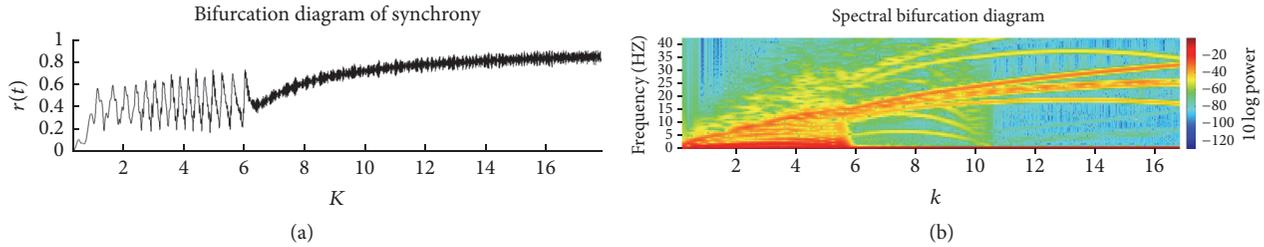


FIGURE 3: Bifurcation diagrams constructed to explore the evolution of synchrony in time, while k is increased ($k_t = k_{t+1} + 5 \times 10^{-4}$ for t in ms). As can be seen in the plot (a), metastability is abruptly reduced for $k \approx 6.5$. In the plot (b), we show the power spectra of synchrony. Slow components are reduced after the bifurcation in $k \approx 6.5$. An additional reduction in the components was found about $k \approx 11$. Note that although the frequency of oscillators is 60 Hz, the behavior of $r(t)$ exhibits multiple components.

were downsampled ten times to obtain a sampling rate similar to real data. $SS(k, t) = 1$ if the absolute value of the poststimulus simulated signal was higher than the absolute value of the maximum entry encountered in any sensor and any time step from the prestimulus baseline, and $SS(k, t) = 0$ otherwise. $SS(k, t)$ was used as input for the Lempel-Ziv measure [26] to estimate the algorithmic complexity (C_L). C_L gives the number of chains with nonredundant information contained in $SS(k, t)$. The algorithm seeks for the minimal number of patterns necessary to describe the sequence. For random sequences, the asymptotic behavior of the measure is $LH_B(L)\log_2 L$, where $H_B(L)$ is the binary entropy for length L

$$H_B(L) = -p_1 \log_2(p_1) - (1 - p_1) \log_2(1 - p_1), \quad (12)$$

where p_1 is the probability to find a “1” in the binary sequence of length L . PCI is defined as the normalized value of C_L :

$$PCI = C_L \frac{\log_2(L)}{LH_B(L)}. \quad (13)$$

3. Results and Discussion

The results for the Kuramoto simulations are characterized in the first place to understand the basic dynamics of the model. In addition, we include graphical descriptions of the ERPs to visualize the structure of the averaged waves at each sensor from simulated perturbations. Finally, we describe and compare Φ and PCI taking into account the values of k with high and low metastability, which as mentioned above is a necessary condition for brain dynamics at resting state.

3.1. Kuramoto Simulations in the Baseline Condition. In Figure 3(a) we show a diagram with the $r(t)$ behavior in the baseline condition at several values of its coupling parameter k . The most important property in the evolution of $r(t)$ is the metastability that can be estimated by the variability of $r(t)$. As can be seen, there is a bifurcation for $k \approx 6.5$ that indicates the end of high metastability and hence the dynamics of the model for $k > 6.5$ should be taken with caution since in principle, there is no functional correspondence with real cortical dynamics.

In addition, it is important to note that the frequency structure in $r(t)$ is not fixed for all k s. One can observe in

Figure 3(a) that the frequency of $r(t)$ seems to increase with the increase of k . To better understand this phenomenon, we include a spectral decomposition of the evolution of $r(t)$ in Figure 3(b). Surprisingly, we found a complex landscape in the oscillatory structure of $r(t)$. The general structure of the spectral diagram showed a resemblance with the bifurcation diagram of the classical logistic map. This similarity appeared because it showed a bifurcation-like proliferation of frequency components as the parameter k increased. The nature of this spectral structure, however, was not explored and goes beyond the goals of this study. If we inspect Figure 3(b), it can be stated that the end of slow oscillatory properties of high metastability was evident at $k \approx 6.5$. Above this value, the critical coupling strength of certain clusters is achieved, so their synchronization becomes stable, while the order parameter of the whole system remains < 1 . As k increases further, larger and larger synchronized clusters are formed, resulting in a reduced number of components, until $r(t)$ approaches 1, with ultimately only a single component as k tends to infinity [27]. Accordingly, in Figure 3(b), the main frequency of the signal slowly increased with k , and the components of $r(t)$ seemed to increase with k as well following a complicated pattern. It is also noteworthy that in Figure 3(b) another bifurcation-like region for $k \approx 11$ can be perceived that consists in a reduction in the number of components. Hence, by the end of the k landscape, $r(t)$ seems to be more simpler with less oscillatory properties and probably this could lead to low values of Φ_{TS} and PCI.

The shape of this diagrams led us to consider that Φ_{TS} and PCI could be sensitive to the bifurcation region in $r(t)$. If metastability is a necessary condition for brain functioning [16] it would be reasonable to think that Φ_{TS} should diminish in the low metastability region. The same should be true for PCI if this measure is closely related to Φ_{TS} . As we will show in the next sections, the reduction after metastability was only found for Φ_{TS} .

3.2. Comparisons between Φ_{TS} and PCI. The relation between Φ_{TS} and PCI was assessed using Pearson product-moment coefficient between the values of Φ_{TS} obtained for each of the k -levels (from .5 to 15) and the corresponding values of the PCI obtained for those same levels of k . The results showed a nonsignificant negative correlation ($r = -.21$, $p = .13$). Thus, apparently, Φ_{TS} and PCI are linearly independent. However,

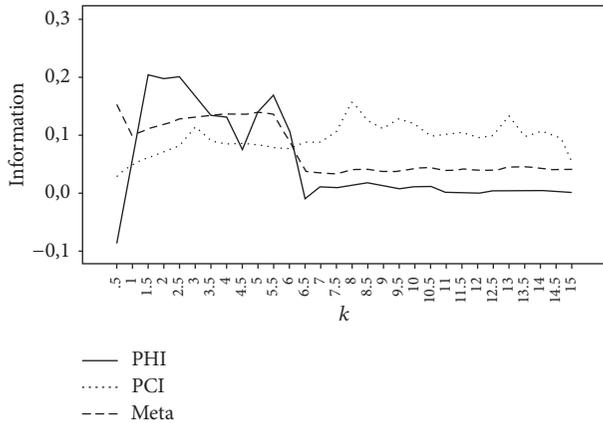


FIGURE 4: Evolution of Φ_{TS} and PCI and the metastability of the model for different values of global coupling strength (k) assessed.

taking into account metastability values, a bifurcation for $k = 6.5$ is apparent. We therefore divided the series into two parts according to the bifurcation, before (from $k = .5$ to 6) and after (from $k = 6.5$ to 15) the bifurcation, and recalculated the correlation between Φ_{TS} and PCI in each of those two parts. In this case, results showed a significant positive relation between Φ_{TS} and PCI before ($r = .64$, $p = .01$) and after ($r = .51$, $p = .01$) the bifurcation.

A graphical description of the evolution of Φ_{TS} and PCI and the metastability of the model can be observed in Figure 4. Values were calculated with $.5$ k -steps. A visual inspection of Figure 4 shows that both Φ_{TS} and metastability exhibited a big decrease around $k \approx 6.5$ indicating the dependence between Φ_{TS} and metastability. Actually, the correlation between metastability and Φ_{TS} over the whole range of k was significant ($r = .68$; $p < .01$). However, the PCI did not show a significant decrease in this region of k . In fact, the PCI evolution seems to progressively increase with k . The fact that there is no decrease in PCI does not mean that this measure is not related to Φ_{TS} ; as seen before, a closer exploration of both measures indicated a positive relation between them. Moreover, one can observe in Figure 4 that the PCI trend seems to stop after $k \approx 11$ which is in agreement with the bifurcation shown in Figure 3(b) showing that the synchrony dynamics are simpler for this region.

4. Conclusions

Under the assumption that conscious states come from integrated information in a system, various metrics have been proposed to try to quantify consciousness. In the present work, we tested two of them using a neurocomputational model. On the one hand, the theoretically well founded Φ has been proposed as a way to quantify the total amount of information that a conscious system can integrate [28]. On the other hand, the PCI distinguishes conscious versus unconscious states at a single patient precision [10]. Under the assumption that conscious states correspond to a distributed but nonuniform spatiotemporal pattern of current sources, Casali et al. applied a standard data compression scheme (the Lempel-Ziv algorithm) to distinguish between conscious and

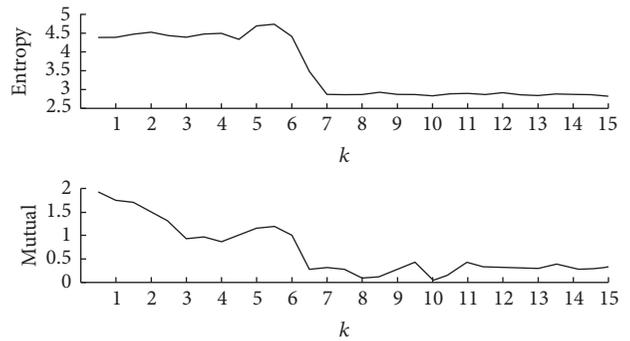


FIGURE 5: Evolution of entropy and mutual information (mutual) of the model as a function of the global coupling strength (k).

unconscious states. Despite the excellent results at applied level, the claim that the measure is theoretically grounded in a conceptual understanding of consciousness deserves a closer look. In the present work, we have tackled the possible relationship between these two measures of the degree of consciousness in a system.

As stated previously, according to the IIT, wakeful consciousness requires the ability to integrate information across multiple brain regions with a high degree of differentiated activity. Thus, loss of consciousness may result as a consequence of a loss of integration as well as a loss of differentiation (or both). Due to our manipulations in the present paper (i.e., increase of k -values), the reduction in consciousness indicated by Φ_{TS} values would reflect such stereotypical behavior across different oscillators. Increasing the mean field connectivity of the model results in an increase of synchrony, as revealed by the Kuramoto order parameter (r) and in a reduction of its variability (metastability). From this point of view, results obtained from metastability and Φ_{TS} converge with the theoretical predictions that suggest that metastability is a necessary condition for healthy brain functioning and consciousness [16].

In general, as k increases, the system as a whole is more coherent and hence, it is more integrated. When the system is above the bifurcation point ($k > 6.5$), synchrony is very high and the dynamics of the binary time series from the clusters tend to be 1 all the time ($r(t) > 0.8$). Hence, information will tend to be low. So what happens is that information is much lower as the system crosses the bifurcation point. One can objectively see this by observing the entropy and mutual information (in bits) of the system (see Figure 5). This descent in the entropy after the bifurcation point could be responsible for the apparent inability of the PCI to capture the dynamic of the system after the bifurcation. In this respect, [29] have suggested that the inability of Lempel-Ziv to compress efficiently low entropy sequences is due to the inability to cope with long runs of identical symbols. In this respect, the Perturbative Integration Latency Index that characterizes the latency of extinction of a massive stimulation perturbing a basal state without drawing upon Lempel-Ziv algorithm, recently proposed by [30], is a promising option that future works could explore.

An important finding in this study is the significant positive correlations between Φ_{TS} and PCI before and after

the bifurcation point. These correlations might indicate that there is a modulation in the PCI when Φ_{TS} changes due to connectivity manipulations. Hence, from the exploration we carried out in this study we can claim that the PCI is sensitive to Φ_{TS} modulations, and that this is true when the system is considered in a coherent region of metastability (high versus low).

One limitation in our study is that we have not found a critical value of k at which both measures reached a maximum. An outstanding correspondence between Φ_{TS} and PCI would have led to an optimum k parameter that characterizes the system in terms of integrated information for both theoretical and empirical estimators. It is evident that the reasons why we did not find this perfect convergence could be that our model is an oversimplification of brain functioning, or the procedure to calculate Φ_{TS} and PCI relies on many simplifications for the characterization of the system. Due to this oversimplification of the brain dynamics, the fact that the PCI after the bifurcation point tend to increase could be due to intrinsic characteristics of the model. However, to the best of our knowledge, Casali et al. did not test PCI when patients exhibit loss of consciousness due to global synchronization in the cortex (epileptic states, for example). Future works could explore the possibility that the loss of consciousness due to global synchronization cannot be fully captured by PCI. However, it is noteworthy that when we consider the metastable region of the model ($k < 6.5$) the maximum values for Φ_{TS} and PCI are found in the short range $2 < k < 3$, and the minimum values are found for $k = .5$ when the connectivity of the system is relatively low. These two findings might indicate that Φ_{TS} and PCI could have a better agreement for high levels of metastability, and if this is true, it is not surprising that PCI is a good clinical indicator of conscious states.

Another potential limitation in our study could be the algorithm used to estimate Φ_{TS} originally proposed by Barrett and Seth [9]. This algorithm produces negative values that could hinder the interpretation of the obtained results. These limitations have given rise to new versions of Φ in which the disadvantage of negative values is solved [31, 32]. Although both measures are highly correlated, future studies could include both estimators.

Disclosure

This study was partially reported as oral presentation in the local workshop “*Modelos Globales de Dinámica Cerebral*,” held in Seville on December 12, 2017.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work has been supported by the Spanish Government and the European Union (via ERDF funds) through the Research Project MTM2014-61312-EXP and by Junta de

Andalucía (Biomedical and Health Science Research Project PI-0386-2016).

References

- [1] D. J. Chalmers, “Facing up to the problem of consciousness,” *Journal of Consciousness Studies*, vol. 2, no. 3, pp. 200–219, 1995.
- [2] N. Block, “Comparing the major theories of consciousness,” in *The Cognitive Neurosciences IV*, M. Gazzaniga, Ed., pp. 1111–1123, 2009.
- [3] S. Dehaene and J.-P. Changeux, “Experimental and theoretical approaches to conscious processing,” *Neuron*, vol. 70, no. 2, pp. 200–227, 2011.
- [4] D. Balduzzi and G. Tononi, “Integrated information in discrete dynamical systems: motivation and theoretical framework,” *PLoS Computational Biology*, vol. 4, no. 6, Article ID e1000091, 2008.
- [5] M. Oizumi, L. Albantakis, and G. Tononi, “From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0,” *PLoS Computational Biology*, vol. 10, no. 5, Article ID e1003588, 2014.
- [6] G. Tononi, “Consciousness as integrated information: A provisional manifesto,” *The Biological Bulletin*, vol. 215, no. 3, pp. 216–242, 2008.
- [7] G. Buzsáki, *Rhythms of the Brain*, Oxford University Press, Oxford, NY, USA, 2011.
- [8] J. D. Barrow and F. J. Tipler, *The Anthropic Cosmological Principle*, Oxford University Press, 1986.
- [9] A. B. Barrett and A. K. Seth, “Practical measures of integrated information for time-series data,” *PLoS Computational Biology*, vol. 7, no. 1, Article ID e1001052, 2011.
- [10] A. G. Casali, O. Gosseries, M. Rosanova et al., “A Theoretically Based Index of Consciousness Independent of Sensory Processing and Behavior,” *Science Translational Medicine*, vol. 5, no. 198, pp. 198ra105–198ra105, 2013.
- [11] P. Hagmann, L. Cammoun, X. Gigandet et al., “Mapping the structural core of human cerebral cortex,” *PLoS Biology*, vol. 6, no. 7, article e159, 2008.
- [12] J. A. Acebrón, L. L. Bonilla, C. J. Perez-Vicente, F. Ritort, and R. Spigler, “The Kuramoto model: a simple paradigm for synchronization phenomena,” *Reviews of Modern Physics*, vol. 77, no. 1, pp. 137–185, 2005.
- [13] M. Breakspear, S. Heitmann, and A. Daffertshofer, “Generative models of cortical oscillations: neurobiological implications of the Kuramoto model,” *Frontiers in Human Neuroscience*, vol. 4, article 190, 2010.
- [14] W. H. Lee and S. Frangou, “Linking functional connectivity and dynamic properties of resting-state networks,” *Scientific Reports*, vol. 7, no. 1, 2017.
- [15] J. Cabral, E. Hugues, O. Sporns, and G. Deco, “Role of local network oscillations in resting-state functional connectivity,” *NeuroImage*, vol. 57, no. 1, pp. 130–139, 2011.
- [16] E. Tognoli and J. A. S. Kelso, “The Metastable Brain,” *Neuron*, vol. 81, no. 1, pp. 35–48, 2014.
- [17] J. Cabral, H. Luckhoo, M. Woolrich et al., “Exploring mechanisms of spontaneous functional connectivity in MEG: how delayed network interactions lead to structured amplitude envelopes of band-pass filtered oscillations,” *NeuroImage*, vol. 90, pp. 423–435, 2014.

- [18] J. Cabral, M. L. Kringelbach, and G. Deco, "Exploring the network dynamics underlying brain activity during rest," *Progress in Neurobiology*, vol. 114, pp. 102–131, 2014.
- [19] I. Chouvarda, E. Michail, A. Kokonozi, L. Staner, N. Domis, and N. Maglaveras, "Investigation of sleepiness induced by insomnia medication treatment and sleep deprivation," in *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, D. D. Schmorow, I. V. Estabrooke, and M. Grootjen, Eds., pp. 120–127, Springer, Berlin, Germany, 2009.
- [20] J. Escudero, A. Ibanez-Molina, and S. Iglesias-Parro, "Effect of the average delay and mean connectivity of the Kuramoto model on the complexity of the output electroencephalograms," in *Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2015*, pp. 7873–7876, August 2015.
- [21] J. Jeong, J. C. Gore, and B. S. Peterson, "Mutual information analysis of the EEG in patients with Alzheimer's disease," *Clinical Neurophysiology*, vol. 112, no. 5, pp. 827–835, 2001.
- [22] P. A. M. Mediano, J. C. Farah, and M. Shanahan, "Integrated information and metastability in systems of coupled oscillators," <https://arxiv.org/abs/1606.08313>.
- [23] P. J. Hellyer, M. Shanahan, G. Scott, R. J. S. Wise, D. J. Sharp, and R. Leech, "The control of global brain dynamics: Opposing actions of frontoparietal control and default mode networks on attention," *The Journal of Neuroscience*, vol. 34, no. 2, pp. 451–461, 2014.
- [24] A. J. Ibáñez-Molina and S. Iglesias-Parro, "Neurocomputational model of EEG complexity during mind wandering," *Frontiers in Computational Neuroscience*, vol. 10, article no. 20, 2016.
- [25] K. Hoehstetter, H. Bornfleth, D. Weckesser, N. Ille, P. Berg, and M. Scherg, "BESA source coherence: A new method to study cortical oscillatory coupling," *Brain Topography*, vol. 16, no. 4, pp. 233–238, 2004.
- [26] A. Lempel and J. Ziv, "On the Complexity of Finite Sequences," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 75–81, 1976.
- [27] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, "Synchronization reveals topological scales in complex networks," *Physical Review Letters*, vol. 96, no. 11, Article ID 114102, 2006.
- [28] G. Tononi, "An information integration theory of consciousness," *BMC Neuroscience*, vol. 5, no. 1, article 42, 2004.
- [29] S. R. Kosaraju and G. Manzini, "Compression of low entropy strings with Lempel-Ziv algorithms," in *Proceedings of the Compression and Complexity of SEQUENCES 1997 (Cat. No. 97TB100171)*, pp. 107–121, 1997.
- [30] G. Deco, J. Cabral, V. M. Saenger et al., "Perturbation of whole-brain dynamics in silico reveals mechanistic differences between brain states," *NeuroImage*, vol. 169, pp. 46–56, 2018.
- [31] M. Oizumi, S.-I. Amari, T. Yanagawa, N. Fujii, and N. Tsuchiya, "Measuring Integrated Information from the Decoding Perspective," *PLoS Computational Biology*, vol. 12, no. 1, Article ID e1004654, 2016.
- [32] M. Tegmark, "Improved Measures of Integrated Information," *PLoS Computational Biology*, vol. 12, no. 11, Article ID e1005123, 2016.

Research Article

Complexity Measures for Quantifying Changes in Electroencephalogram in Alzheimer's Disease

Ali H. Hussein Al-Nuaimi , Emmanuel Jammeh ,
Lingfen Sun , and Emmanuel Ifeakor 

Signal Processing and Multimedia Communication (SPMC) Research Group, Faculty of Science and Engineering, School of Computing, Electronics, and Mathematics, University of Plymouth, Plymouth, UK

Correspondence should be addressed to Ali H. Hussein Al-Nuaimi; ali.al-nuaimi@plymouth.ac.uk

Received 15 September 2017; Revised 26 December 2017; Accepted 4 February 2018; Published 13 March 2018

Academic Editor: Hugo Leonardo Rufiner

Copyright © 2018 Ali H. Hussein Al-Nuaimi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Alzheimer's disease (AD) is a progressive disorder that affects cognitive brain functions and starts many years before its clinical manifestations. A biomarker that provides a quantitative measure of changes in the brain due to AD in the early stages would be useful for early diagnosis of AD, but this would involve dealing with large numbers of people because up to 50% of dementia sufferers do not receive formal diagnosis. Thus, there is a need for accurate, low-cost, and easy to use biomarkers that could be used to detect AD in its early stages. Potentially, electroencephalogram (EEG) based biomarkers can play a vital role in early diagnosis of AD as they can fulfill these needs. This is a cross-sectional study that aims to demonstrate the usefulness of EEG complexity measures in early AD diagnosis. We have focused on the three complexity methods which have shown the greatest promise in the detection of AD, Tsallis entropy (TsEn), Higuchi Fractal Dimension (HFD), and Lempel-Ziv complexity (LZC) methods. Unlike previous approaches, in this study, the complexity measures are derived from EEG frequency bands (instead of the entire EEG) as EEG activities have significant association with AD and this has led to enhanced performance. The results show that AD patients have significantly lower TsEn, HFD, and LZC values for specific EEG frequency bands and for specific EEG channels and that this information can be used to detect AD with a sensitivity and specificity of more than 90%.

1. Introduction

Alzheimer's disease (AD) is an age-related progressive, neurodegenerative disorder that is characterized by loss of memory and cognitive decline [1, 2] and it is the main cause of disability among older people [3]. AD is ranked as the sixth leading cause of death in US [4]. The rapid increase in the number of people living with AD and other forms of dementia due to the ageing population represents a major challenge to health and social care systems worldwide [5]. Currently, there are over 46.8 million individuals with dementia worldwide with an annual cost of care estimated at US\$818 billion and is projected to reach 74.7 million by 2030 with an annual cost of US\$2 trillion [6]. The number of individuals with dementia worldwide is expected to exceed 131 million by 2050 which will have a huge economic impact [7]. However, many dementia sufferers do not receive early

diagnosis [7, 8]. It is estimated that up to 50% of people living with dementia may not have received formal diagnosis [8, 9]. In 2011, 28 million people of 36 million dementia sufferers did not receive a diagnosis worldwide [10].

Degeneration of brain cells due to AD starts many years before the clinical manifestations become clear [5, 11–15]. An early diagnosis of AD will contribute to the development of effective treatments that could slow, stop, or prevent significant cognitive decline [16–18]. An early diagnosis of AD could also be useful for identifying dementia sufferers who have not received a formal early diagnosis and this may provide an opportunity for them to access appropriate health care services [19–21].

A biomarker that can measure degeneration of brain cells due to AD at an early stage would be useful for its early diagnosis [2, 22–24]. But this may require dealing with large numbers of people as up to 50% of people living with

dementia may not have received a formal diagnosis. Therefore, there is a need for simple, noninvasive, low-cost, and reliable biomarkers for early diagnosis which can be accessed in clinical practice [5, 25, 26]. Recent guidelines promote the use of biochemical and neuroimaging biomarkers to improve the diagnosis of AD. Cerebral spinal fluid (CSF) testing for AD is not widely used in clinical practice because it requires lumbar puncture which is an invasive procedure [2, 27, 28]. Neuroimaging is expensive, available only in specialist centres [29], and may not be suitable for patients that have pacemakers or certain implants [30]. Blood-based biomarkers have shown promising results in AD diagnosis but they are not yet fully developed and low-cost biosensors to detect AD biomarkers in blood do not exist at present [2, 25, 31].

Potentially, the electroencephalogram (EEG) can play a valuable role in the early diagnosis of AD [11, 20, 21, 24, 32–34]. EEG is noninvasive, low-cost, has a high temporal resolution, and provides valuable information about brain dynamics in AD [20, 21, 33, 35, 36]. The fundamental utility of EEG to detect brain signal changes even in the preclinical stage of the disease has been demonstrated [33, 37, 38]. Thus, EEG biomarkers may be used as a first-line decision-support tool in AD diagnosis [11, 35] and could complement other AD biomarkers [26].

AD is characterized by loss of memory and cognitive decline resulting from damage to brain cells which influence brain activity [38]. AD causes changes in the features of the EEG [35, 38, 39] and EEG analysis may provide valuable information about brain dynamics due to AD [20, 21, 33, 35]. The most characteristic features in EEG caused by AD are slowing of EEG, a decrease in EEG coherence, and reduction in EEG complexity [33–35, 37, 38, 40, 41]. These changes in the EEG can be quantified as a biomarker of AD. A variety of linear and nonlinear methods are being developed to quantify changes in EEG as AD biomarkers [42, 43]. AD biomarkers based on the slowing in EEG and a decrease in EEG coherence are often derived using linear analysis methods (i.e., spectral analysis of the EEG signal) [37, 44, 45], while biomarkers extracted by analysing the complexity of the EEG are based on nonlinear approaches (e.g., entropy methods, fractal dimension, and Lempel-Ziv complexity). The EEG complexity approaches have shown promising results in AD diagnosis [11, 35, 46] and appear to be appropriate for AD diagnosis [38, 47, 48]. Complexity is a measure of the extent to which the dynamic behavior of a given sequence resembles a random one [49]. The cortical areas of the brain fire spontaneously and this dynamic behavior of the brain is complex [50, 51]. AD causes a reduction in neuronal activity of the brain [52] resulting in decreased capability of the brain to process information [53–55] and this may be reflected in the EEG signals [52]. EEG complexity can potentially be a good biomarker for AD diagnosis [38] as AD patients have a significant reduction in EEG complexity [38, 40, 41, 52, 56, 57]. Several studies have investigated EEG complexity as a potential AD biomarker using whole EEG record with the objective of achieving a high performance. Given the association of EEG activities (e.g., alpha, delta activities) with AD, we hypothesized that the derivation of EEG complexity based on EEG activities should lead to enhanced performance.

This is a cross-sectional study aimed at demonstrating the usefulness of EEG based complexity measures to detect AD. In this study, we investigated an important class of complexity measures, information theoretic methods, which offers a potentially powerful approach for quantifying changes in the EEG due to AD [58]. Information theoretic methods (i.e., TsEn and LZC) have emerged as a potentially useful complexity-based approach to derive robust EEG biomarkers of AD [47, 58–62]. They are attractive because of the potential natural link between information theory-based biomarkers and changes in the brain caused by AD [58]. Conceptually, information processing activities in the brain are thought to be reflected in the information content of the EEG.

In particular, TsEn approach has been shown to be one of the most promising information theoretic methods for quantifying changes in the EEG [62, 63]. It has also been shown to be a reliable analysis tool to use with working memory tasks. As its computation is fast, it can serve as a basis for a real-time decision-support tool for dementia diagnosis by both specialists and nonspecialists [64]. Sneddon et al. [65] investigated TsEn of the EEG and was able to detect mild dementia due to AD with a sensitivity of 88% and specificity of 94%. De Bock et al. [62] found TsEn of the EEG to be a highly promising potential diagnostic tool for mild cognitive impairment (MCI) and early dementia with a sensitivity and specificity of 82% and 73%, respectively. Using TsEn approach, Al-Nuaimi et al. [35] detected AD from normal subjects with a sensitivity and specificity of 85.8% and 70.9%, respectively. Garn et al. [66] investigated the use of TsEn to diagnose AD based on EEG analysis and achieved a p value < 0.0036 for channels T7 and T8 in discriminating between AD patients and normal subjects.

LZC is a nonparametric, nonlinear measure of complexity for finite length sequences [67]. It is a simple and powerful method which has been used in several biomedical applications [68]. LZC depends on a coarse-grain processing of the measurements [69] and can be applied directly on physiologic signal without preprocessing [70]. LZC has been applied extensively in analysing biomedical signals (e.g., EEG) to measure the complexity of discrete-time physiologic signals [67]. Furthermore, it is used to analyse brain function, brain information transmission, and EEG complexity in patients with AD [43]. The LZC approach produces a good biomarker for AD detection [70, 71]. Hornero et al. [72] used LZC to analyse EEG and magnetoencephalogram (MEG) in AD patients. They found that LZC provides a good insight into the EEG background activity characteristics and the changes associated with AD. Hornero et al. [73] found that LZC values were lower in AD patients and suggested that the most relevant differences are in the posterior region. In addition, they suggested that the MEG activity from AD patients is characterized by a lower degree of irregularity and complexity and that the LZC measures can be used to detect AD with a sensitivity and specificity values of 65% and 76.2%, respectively. McBride et al. [56] analysed EEG complexity based on the LZC method to discriminate between patients with early MCI, AD patients, and normal subjects. They found that EEG complexity features for specific EEG frequency bands with regional electrical activity provide promising results

in discriminating between MCI, AD, and normal subjects. Fernandez et al. [74] analysed MEG complexity for MCI patients, AD patients, and normal subjects based on LZC method for discriminating between the three groups. They found that a combination of age and posterior LZC scores allowed them to distinguish between AD patients and MCI patients with 94.4% sensitivity and specificity.

HFD is a fast computational method for obtaining the fractal dimension of time series signals [75–77] even when very few data points are available [75]. It can track changes in a biosignal from a measure of its complexity [75, 76] and it is suited to capturing region-specific neural changes due to AD [45, 77]. In addition, HFD provides a more accurate measure of the complexity of signals compared to other methods [75, 78, 79] and it has been shown to be an efficient method for discriminating between AD patients and normal subjects [31, 80]. HFD of the EEG is potentially a good biomarker of AD diagnosis as it is significantly lower in AD patients than in normal subjects [46, 80, 81]. Smits et al. found that HFD is sensitive to neural changes selectively related to AD patients and normal subjects. Al-Nuaimi et al. [46] investigated HFD of EEG for AD diagnosis and they found that HFD is a promising EEG biomarker that captures changes in the regions of the brain thought to be affected first by AD and it could be used to detect AD with sensitivity and specificity values of 100% and 80%, respectively.

It is widely accepted that AD causes a decrease in the power of high frequencies (alpha, beta, and gamma) and an increase in the power of low frequencies (delta and theta) [11, 33, 34, 38, 41]. We hypothesized that complexity measures based on the EEG frequency bands would provide better results than those derived directly from the whole EEG record. The aim was to enhance the performance of the complexity measures and to demonstrate their usefulness in quantifying changes in EEG due to AD.

Digital filters were used to extract the five EEG frequency bands (i.e., delta, theta, alpha, beta, and gamma). Complexity measures were then obtained for each of the five EEG frequency bands and for each channel using each of the three methods of computing complexity measures (TsEn, HFD, and LZC).

For each method, we computed a panel of 114 biomarkers (i.e., 19 biomarkers for the whole EEG record and 19 biomarkers for each of the five EEG frequency bands). The performance measures for each biomarker were computed (including the sensitivity and specificity).

The paper is arranged as follows. In Section 2, the materials and methods used in the study are described. In Section 3, the results and discussions are presented and the conclusions are presented in Section 4.

2. Materials and Methods

2.1. Materials. This study was based on EEG dataset that was recorded from 52 volunteers. All the volunteers underwent a strict protocol based on normal hospital practices at Derriford Hospital, Plymouth, UK [11]. The EEG recordings include several states such as hyperventilation, awake, drowsy, and alert, with periods of eyes closed and open. For

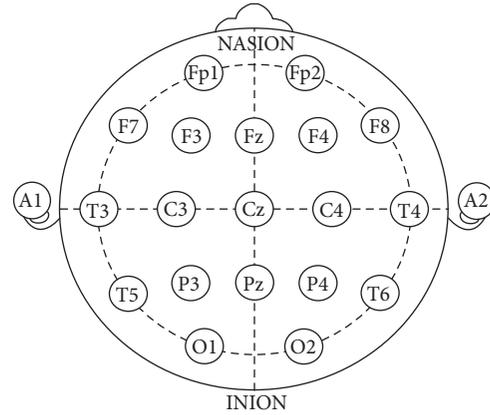


FIGURE 1: International 10–20 system.

storage reasons, the sampling rate was reduced from 256 Hz to 128 Hz by averaging two consecutive samples. The duration of each EEG signal is 4 minutes. Figure 1 shows the electrode locations using a 10–20 system. The letters F, C, P, O, and T refer to cerebral cortex lobes (F: frontal, C: central, P: parietal, O: occipital, and T: temporal) [82].

The EEG dataset consists of two subdatasets (A and B). Subdataset A includes 11 age matched subjects over 65 years old (3 AD patients and 8 normal subjects). Subdataset A was recorded using the traditional 10–20 system in a Common Reference Montage by using the average of all channels as reference and the EEG signals were converted to Common Average and Bipolar Montages using software. Subdataset B includes 41 subjects that were not perfectly age matched (24 normal subjects, 10 males and 14 females, have mean age 69.4 ± 11.5 years (from 40 to 84 years) and 17 were probable AD patients, 9 male and 8 female). The normal subjects have a mean age of 69.4 ± 11.5 years (40 to 84 years) and the probable AD subjects have mean age of 77.6 ± 10.0 years (from 50 to 93 years). Subdataset B was recorded using the modified Maudsley system. The conventional 10–20 system has a similar setting with the Maudsley electrode positioning system [83].

All patients were referred to the EEG department at Derriford Hospital from a specialist memory clinic. A battery of psychometric tests (including the MMSE [84], Rey Auditory Verbal Learning Test [85], Benton Visual Retention Test [86], and memory recall tests [87]) were performed on all patients at the memory clinic. The classification of subjects with dementia was based on the working diagnosis provided by the specialist memory clinic. All healthy volunteers and AD patients had their EEG confirmed by a consultant clinical neurophysiologist at the hospital as normal and probable mild AD, respectively [11].

2.2. Methods. In our approach, the complete recordings of the EEG including artefacts were used without a priori selection of elements for analyses. This enabled us to have an idea about the robustness and usefulness of the method in practice. Data from a fixed interval (61 s to 240 s) was used to avoid electrical artefacts, which regularly occur at the

beginning of a record, leaving a standard three-minute data to analyse.

The following steps outline the procedure that was used to derive the biomarkers for the three complexity methods (i.e., TsEn, HFD, and LZC)

- (1) The EEG signal was filtered using infinite impulse response (IIR) Chebyshev-II bandpass filter into five frequency bands (i.e., delta 0–4 Hz, theta 4–8 Hz, alpha 8–12 Hz, beta 12–30 Hz, and gamma 30–45 Hz). A low computational IIR filter was used to retain the computational efficiency of the derived complexity-based biomarkers [88].
- (2) The biomarkers were then derived first from the whole EEG record and then for each of EEG frequency bands for each of the three EEG complexity methods.
- (3) For each biomarker of the EEG complexity methods (i.e., TsEn, HFD, and LZC), p values were computed between AD patients and normal subjects using Student's t -test.
- (4) The performance of each complexity measure to detect AD is then assessed. For each complexity measure, a classification model, based on the support vector machine (SVM), was used to detect AD.

Tsallis Entropy (TsEn). TsEn [89] biomarker computation of an N -samples EEG data sequence $x(1), x(2), \dots, x(N)$ is based on the generalised measure of entropy, due to Tsallis:

$$\text{TsEn}_q = \frac{\left(\sum_{i=1}^k P_i - P_i^q\right)}{(q-1)}, \quad (1)$$

where TsEn_q is the Tsallis entropy value, k is the number of states that the amplitudes of the EEG are quantized into, P_i is a probability associated with the i th state, and q is Tsallis parameter ($k = 2200$ and $q = 0.5$).

Higuchi Fractal Dimension (HFD). To compute HFD biomarker [75, 77, 90] of an N -sample EEG data sequence $x(1), x(2), \dots, x(N)$, the data is first divided into a k -length subdata set as

$$x_k^m: x(m), x(m+k), x(m+2k), \dots, x\left(m + \left[\frac{N-m}{k}\right] \cdot k\right), \quad (2)$$

where $[\]$ is Gauss' notation, k is constant, and $m = 1, 2, \dots, k$. The length $L_m(k)$ for each subdata set is then computed as

$$L_m(k) = \left\{ \left[\sum_{i=1}^{\lfloor (N-m)/k \rfloor} |x(m+ik) - x(m+(i-1) \cdot k)| \right] \cdot ((N-1) / (\lfloor (N-m)/k \rfloor \cdot k)) \right\} \cdot (k)^{-1}. \quad (3)$$

The mean of $L_m(k)$ is then computed to find the HFD for the data as

$$\text{HFD} = \frac{1}{K} \sum_{M=1}^K L_m(k). \quad (4)$$

Lempel-Ziv Complexity (LZC). To compute the LZC [43, 49, 67, 68, 70] biomarker of an N -sample EEG data sequence $x(1), x(2), \dots, x(N)$, the EEG signal is first converted into a binary string as

$$x(i) = \begin{cases} 0 & \text{if EEG}(i) < M \\ 1 & \text{if EEG}(i) \geq M, \end{cases} \quad (5)$$

where $x(i)$ is the equivalent binary value of $\text{EEG}(i)$, i is the index of all values in the EEG signal, and M is the median value of each EEG channel. The median value is used to manage the outliers.

The binary string is then scanned from left to right until the end to produce new substrings. A complexity counter $c(N)$ is the number of new substrings. The upper bound of $c(N)$ is used to normalise $c(N)$ to get an independent value from the sequence of length N . The upper bound of $c(N)$ is $N/\log_2(N)$. $c(N)$ is then normalised by $b(N)$ as

$$C(N) = \frac{c(N)}{b(N)}, \quad (6)$$

where $C(N)$ is the normalised value of the LZC and $b(N)$ is the upper bound of the $c(N)$.

A panel of 114 biomarkers was computed (19 biomarkers for the whole EEG record and 19 biomarkers for each of EEG frequency band (i.e., delta, theta, alpha, beta, and gamma). To determine which features have a significant statistical association with AD, we computed p values between AD patients and normal subjects using Student's t -test. This allowed us to identify significant features that may be useful to discriminate between AD patients and normal subjects. The dataset was split into training and testing data (60% for training and 40% for testing) with subjects selected at random. We selected 32 subjects for training and 20 subjects for testing at random from the datasets, a ratio of 60 : 40. The training data includes 12 AD (two from dataset A and 10 from dataset B) and 20 normal subjects (six from dataset A and 14 from dataset B). The testing data includes 8 AD (one from dataset A and seven from dataset B) and 12 normal subjects (two from dataset A and 10 from dataset B). p values were computed using the training EEG dataset. Machine learning techniques were used to develop models based on the biomarkers. As a classifier, we used support vector machine (SVM) to model biomarkers extracted using TsEn, HFD, and LZC methods. SVM classifier was used because it is widely used in machine learning and has found application in dementia diagnosis. It has shown better performance in biomedical data analysis and in automatic AD diagnosis compared to other conventional classifiers (e.g., Euclidean distance classifier) and good capability to learn from experimental data [91, 92], and it has a stable classification performance [93].

It has also been shown to outperform other machine learning techniques (e.g., Naive Bayes, Multilayer Perceptron, Bayes Network, egging, Logistic Regression, and Random Forest,) in diagnosis of MCI and dementia [94]. We used the testing EEG dataset to test the performance of the models. For each complexity method, six performance tables were created (whole EEG record, and table for each EEG frequency band).

The performance of the TsEn, HFD, and LZC biomarkers for AD diagnosis was assessed in terms of sensitivity (Sen), specificity (Spec), accuracy (ACC), F -measure, error rate, true positive rate (TPR), false positive rate (FPR), positive predictive value (PPV), and negative predictive value (NPV). Matthew's correlation coefficient (MCC) was computed to measure the quality of the binary classification (AD and normal) between the actual and predicted results [95, 96].

3. Results and Discussions

3.1. Result. We analysed the performance of the three different complexity measures in quantifying changes in EEG due to AD. For this purpose, we examined the differences between the values of the complexity measures derived from EEG signals of AD subjects and those of normal subjects. Biomarkers that do not show significant differences between AD patients and normal subjects may not be suitable for quantifying changes in EEG due to AD as they may not be capable of being used to discriminate between AD patients and normal subjects.

We found that complexity measures derived from the EEG frequency bands for AD patients were significantly different to those of normal subjects compared to complexity measures derived from the whole EEG record. This suggests that they may be better suited to quantify changes in the EEG due to AD and potentially may provide better results in AD diagnosis.

Figure 2 shows the EEG biomarkers derived from whole EEG record (i.e., unfiltered) and those derived from the five EEG bands (delta, theta, alpha, beta, and gamma bands) using the TsEn method. The results show that TsEn values for AD patients are lower than those for normal subjects for whole EEG record. This is consistent with the findings in other studies [35, 58, 62, 64]. Figure 2 also shows that the differences between the TsEn values for AD patients and for normal subjects for the EEG bands (delta and theta bands in particular) are larger than those for whole EEG record. This is a desirable feature in a biomarker as it suggests that TsEn biomarkers derived from the EEG bands may provide better performance in detecting AD than those extracted from whole EEG record.

Figure 3 shows the EEG biomarkers derived from whole EEG record and those derived from the EEG bands (delta, theta, alpha, beta, and gamma bands in particular) using the HFD method. In this case, the results show that HFD values for AD patients are lower than those for normal subjects. This result is consistent with the finding in other studies [46, 80]. As with the TsEn, the differences between HFD biomarkers for AD patients and normal subjects for the EEG frequency bands (i.e., delta, theta, and alpha bands) were

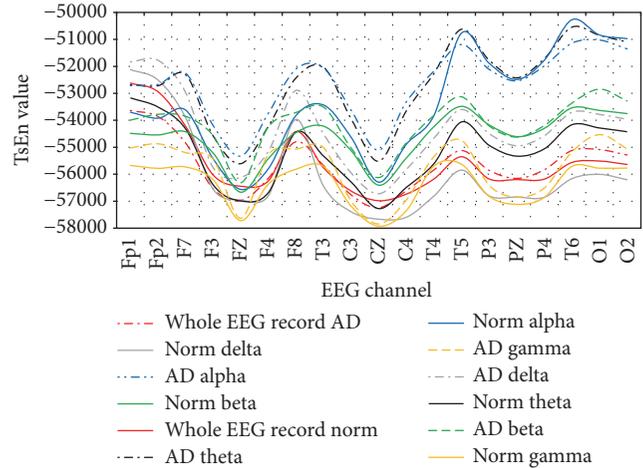


FIGURE 2: EEG biomarkers for TsEn.

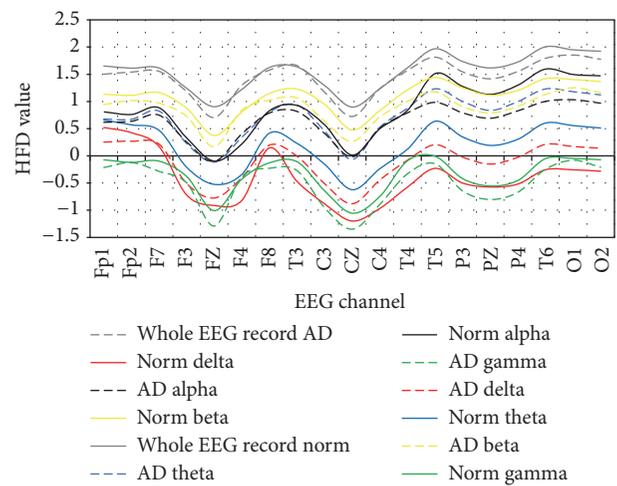


FIGURE 3: EEG biomarkers for HFD.

larger than those for the whole EEG record suggesting that the use of biomarkers derived from the frequency bands would be better at detecting AD than the use of whole EEG record.

Figure 4 shows similar results for the LZC method. In this case, the results show that LZC values for AD patients were lower than those for normal subjects and these are consistent with the finding in other studies [43, 97]. Again, the differences between the LZC biomarkers for AD patients and normal subjects for the five EEG frequency bands (the theta, beta, and gamma bands, in particular) were larger than those for the whole EEG record, suggesting that the use of biomarkers derived from the frequency bands would be better at detecting AD than the use of whole EEG record.

We analysed the complexity measures using p values to determine the statistical significance in detecting AD

Figure 5 shows p values of the differences in TsEn measures between AD patients and normal subjects for whole EEG record and those from the EEG frequency bands. The results show that TsEn biomarkers that were extracted from theta bands have the smallest p values, while the TsEn

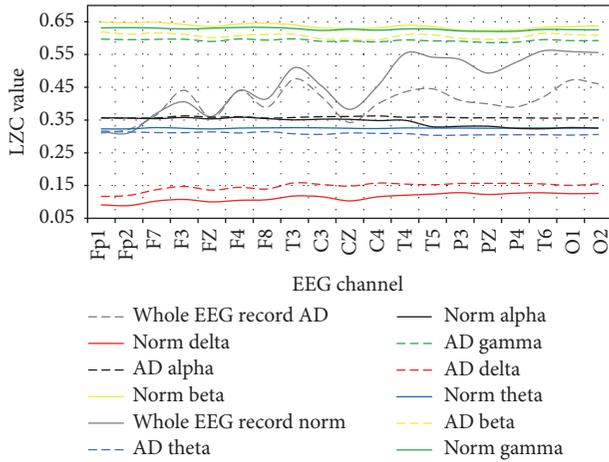
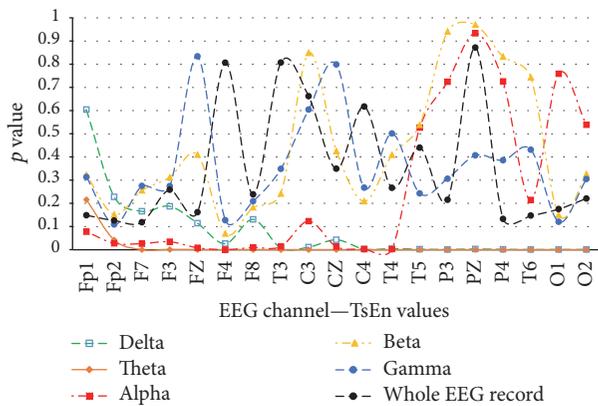
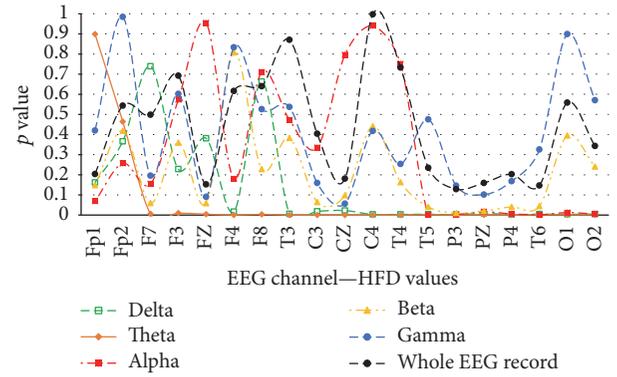
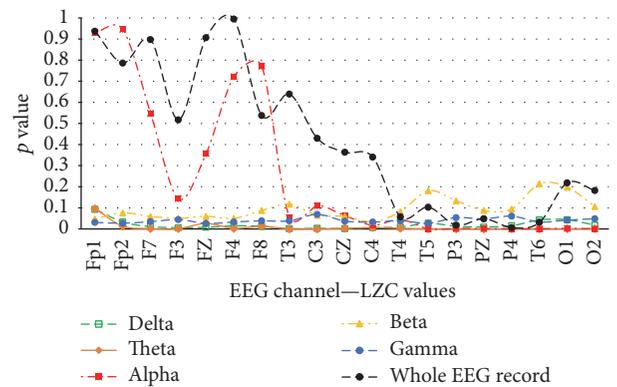


FIGURE 4: EEG biomarkers for LZC.

FIGURE 5: p values for TsEn between AD patients and normal subjects of the training EEG dataset.

biomarkers derived from gamma band have the maximum p value between AD patients and normal subjects. This suggests that biomarkers that are extracted from theta band may provide the best performance in AD diagnosis. Figure 5 also shows that biomarkers that were extracted from EEG frequency bands may have a more significant association with AD than the EEG biomarkers that are derived from whole EEG record based on p value analysis. Therefore, the complexity measures derived from the EEG frequency band may provide better results in the classification between AD patients and normal subjects.

Figures 6 and 7 depict the results of similar p value analysis for HFD and LZC measures, respectively. The results show that, in both HFD and LZC methods, the complexity measures derived from the EEG frequency bands, theta band have significantly smaller p values compared to those of measures derived from the whole EEG record. In both methods, complexity measures derived from the theta band gave the smallest p value. This implies that biomarkers derived from the frequency bands, the theta band in particular, may provide the best possible performance in AD diagnosis using the HFD and LZC methods.

FIGURE 6: p values for HFD between AD patients and normal subjects of the training EEG dataset.FIGURE 7: p values for LZC between AD patients and normal subjects of the training EEG dataset.

Looking across all the results (Figures 5, 6, and 7), the theta band has a minimum p value between AD patients and normal subjects for all three complexity methods (i.e., TsEn, HFD, and LZC). Thus, EEG biomarkers derived from EEG frequency bands are better than the biomarkers that were extracted from whole EEG record. The biomarkers derived from theta band may provide the best performance in AD diagnosis across all three methods.

3.2. The Performance of the EEG Complexity-Based Measures. Table 1 shows the performance of the SVM-based classification model using TsEn biomarkers for whole EEG record for the 19 EEG channels. In this case, the best sensitivity and specificity were 46.67% and 80%, respectively, for Fp2 and F7 EEG channels.

Similar performance indices were computed for each of the five EEG bands using the TsEn. As an example, Table 2 shows the performance indices for TsEn biomarkers for the delta band for the 19 EEG channels. The best sensitivity and specificity were 85.71% and 84.62%, respectively, for T4, O1, and O2 EEG channels.

Similar performance indices were computed for each of the five EEG bands using HFD and LZC methods. Table 3

TABLE 1: TsEn performance for whole EEG record.

EEG channel	Sen.%	Spec.%	Acc.%	<i>F</i> -measure%	Error rate	MCC	FPR%	FNR%	PPV%	NPV%
Fp1	43.75	75.00	50.00	58.33	0.50	0.153	25.00	56.25	87.50	25.00
Fp2	46.67	80.00	55.00	60.87	0.45	0.236	20.00	53.33	87.50	33.33
F7	46.67	80.00	55.00	60.87	0.45	0.236	20.00	53.33	87.50	33.33
F3	43.75	75.00	50.00	58.33	0.50	0.153	25.00	56.25	87.50	25.00
FZ	44.44	100.00	50.00	61.54	0.50	0.272	0.00	55.56	100.00	16.67
F4	44.44	100.00	50.00	61.54	0.50	0.272	0.00	55.56	100.00	16.67
F8	44.44	100.00	50.00	61.54	0.50	0.272	0.00	55.56	100.00	16.67
T3	37.50	50.00	40.00	50.00	0.60	-0.102	50.00	62.50	75.00	16.67
C3	35.71	50.00	40.00	45.45	0.60	-0.134	50.00	64.29	62.50	25.00
CZ	42.11	100.00	45.00	59.26	0.55	0.187	0.00	57.89	100.00	8.33
C4	44.44	100.00	50.00	61.54	0.50	0.272	0.00	55.56	100.00	16.67
T4	35.29	33.33	35.00	48.00	0.65	-0.229	66.67	64.71	75.00	8.33
T5	33.33	50.00	40.00	40.00	0.60	-0.167	50.00	66.67	50.00	33.33
P3	28.57	33.33	30.00	36.36	0.70	-0.356	66.67	71.43	50.00	16.67
PZ	37.50	50.00	40.00	50.00	0.60	-0.102	50.00	62.50	75.00	16.67
P4	35.71	50.00	40.00	45.45	0.60	-0.134	50.00	64.29	62.50	25.00
T6	26.67	20.00	25.00	34.78	0.75	-0.471	80.00	73.33	50.00	8.33
O1	27.27	44.44	35.00	31.58	0.65	-0.287	55.56	72.73	37.50	33.33
O2	30.00	50.00	40.00	33.33	0.60	-0.204	50.00	70.00	37.50	41.67

TABLE 2: TsEn performance for delta band of the EEG signal.

EEG channel	Sen.%	Spec.%	Acc.%	<i>F</i> -measure%	Error rate	MCC	FPR%	FNR%	PPV%	NPV%
Fp1	50.00	66.67	60.00	50.00	0.40	0.167	33.33	50.00	50.00	66.67
Fp2	50.00	62.50	60.00	33.33	0.40	0.102	37.50	50.00	25.00	83.33
F7	55.56	72.73	65.00	58.82	0.35	0.287	27.27	44.44	62.50	66.67
F3	80.00	73.33	75.00	61.54	0.25	0.471	26.67	20.00	50.00	91.67
FZ	50.00	62.50	60.00	33.33	0.40	0.102	37.50	50.00	25.00	83.33
F4	50.00	61.11	60.00	20.00	0.40	0.068	38.89	50.00	12.50	91.67
F8	57.14	69.23	65.00	53.33	0.35	0.257	30.77	42.86	50.00	75.00
T3	71.43	76.92	75.00	66.67	0.25	0.471	23.08	28.57	62.50	83.33
C3	60.00	66.67	65.00	46.15	0.35	0.236	33.33	40.00	37.50	83.33
CZ	100.00	63.16	65.00	22.22	0.35	0.281	36.84	0.00	12.50	100.00
C4	71.43	76.92	75.00	66.67	0.25	0.471	23.08	28.57	62.50	83.33
T4	85.71	84.62	85.00	80.00	0.15	0.685	15.38	14.29	75.00	91.67
T5	80.00	73.33	75.00	61.54	0.25	0.471	26.67	20.00	50.00	91.67
P3	75.00	83.33	80.00	75.00	0.20	0.583	16.67	25.00	75.00	83.33
PZ	100.00	75.00	80.00	66.67	0.20	0.612	25.00	0.00	50.00	100.00
P4	83.33	78.57	80.00	71.43	0.20	0.579	21.43	16.67	62.50	91.67
T6	83.33	78.57	80.00	71.43	0.20	0.579	21.43	16.67	62.50	91.67
O1	85.71	84.62	85.00	80.00	0.15	0.685	15.38	14.29	75.00	91.67
O2	85.71	84.62	85.00	80.00	0.15	0.685	15.38	14.29	75.00	91.67

summarises the best performance indices for the three complexity measures.

Figures 8, 9, and 10 summarise the performance indices of the TsEn, HFD, and ZLC methods.

The results show that TsEn, HFD, and ZLC EEG biomarkers derived from the EEG frequency bands provide better performance than EEG biomarkers derived from the whole EEG record.

3.3. Discussions. The results of this study show that EEG complexity-based measures provide a potentially useful way to detect AD. The most characteristic feature caused by AD is the reduction in EEG complexity [33–35, 37, 38, 40, 41] compared to normal subjects. This is consistent with other studies [35, 38, 43, 46, 56, 58, 62, 64, 80, 97, 98] and shows that EEG complexity measures are potentially a good biomarker for detecting AD.

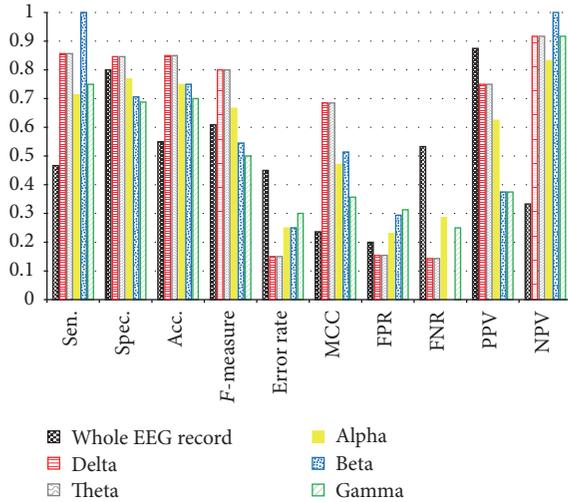


FIGURE 8: TsEn performance.

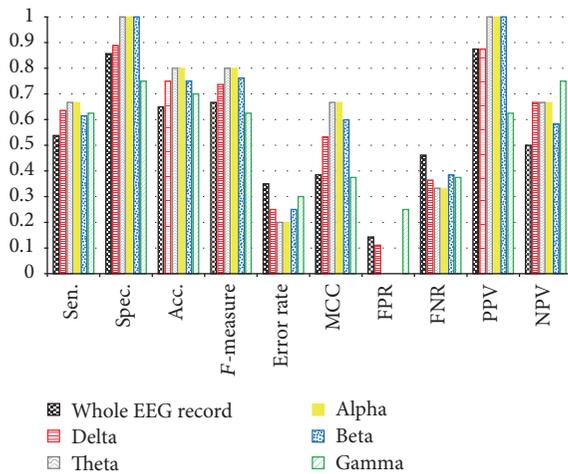


FIGURE 9: HFD performance.

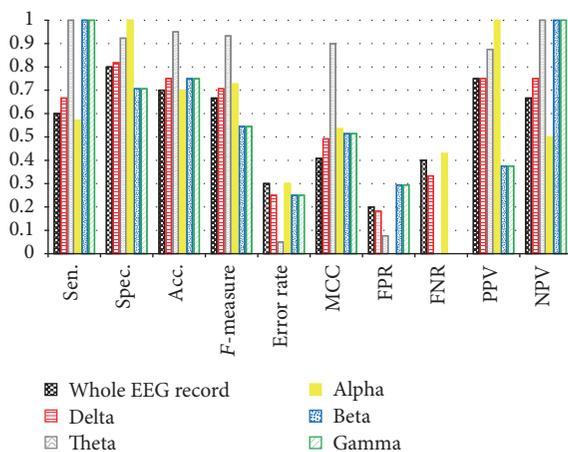


FIGURE 10: LZC performance.

TABLE 3: Summary of the best performance indices for the three complexity measures.

Method	TsEn		HFD		LZC
	Delta	Theta	Theta	Alpha	Theta
Feature	Delta	Theta	Theta	Alpha	Theta
EEG channel	T4, O1, O2	F4	C4	T5, P3	C3
Sen.%	85.71	85.71	66.67	66.67	100
Spec.%	84.62	84.62	100	100	92.31
Acc.%	85	85	80	80	95
F-measure%	80	80	80	80	93.33
Error rate	0.15	0.15	0.2	0.2	0.05
MCC	0.685	0.685	0.667	0.667	0.9
FPR%	15.38	15.38	0	0	7.69
FNR%	14.29	14.29	33.33	33.33	0
PPV%	75	75	100	100	87.5
NPV%	91.67	91.67	66.67	66.67	100

Unlike previous studies, we found that the complexity measures derived from the EEG frequency bands (i.e., delta, theta, alpha, beta, and gamma) provide significantly better performance in detecting AD than the complexity measures derived from whole EEG records. This comes from the greater differences between the complexity measures for AD patients and normal subjects when they are derived from the frequency bands compared to when they are derived from whole record which is a desirable property of a good biomarker.

In particular, we found that for the TsEn and HFD complexity measures derived from the delta and theta bands gave the best performance. For the delta band, three EEG channels (T4, O1, and O2) gave the best performance. For the theta band, F4 gave the best performance.

Similar results were obtained for the LZC complexity measures, except that the best EEG channel was C3 for the theta band. This is consistent with the findings of other studies which suggested that AD starts from the back of the brain and then spreads gradually to other parts of the brain [5, 46, 99–101]. This implies that it may be possible to use only a small number of EEG channels to detect AD.

The findings of this study have a number of implications for research to develop new and robust techniques for the analysis of EEG to increase the contributions EEG makes to the diagnosis of AD.

The results suggest that the three EEG complexity measures, derived from the EEG frequency bands, can detect AD reliably (with sensitivity and specificity of >90%). Thus, EEG complexity measures could provide a basis for developing an accurate, low-cost, and easy to use tool to detect AD. Although the results of the studies are consistent with previous studies, unlike previous studies, in this study, the complexity measures are derived from EEG frequency bands (i.e., delta, theta, alpha, beta, and gamma). The results suggest that deriving the complexity measures from the EEG frequency bands is an important step for achieving robust biomarkers.

We found that AD patients have significantly lower complexity measures for specific EEG frequency bands and for specific EEG channels than normal subjects. This is

consistent with findings in previous studies [33–35, 37, 38, 40, 41]. Thus, it may be possible to identify specific EEG channels and specific frequency bands that may provide the best biomarkers to detect AD. In situations where the number of available channels is limited (e.g., when portable EEG systems are used outside specialist centres), this may be exploited to achieve a good performance.

It may be possible to enhance the performance of the complexity-based approach further, by combining the three complexity measures into a composite model. Given that the three complexity measures are analysing different aspects of the signal (e.g., entropy and fractal measures), integrating them may lead to improved performance.

Our study has a number of limitations. At present, our methods have been applied only to the detection of AD, the most common form of dementia. A more detailed study is necessary to evaluate the methods using a much larger and diverse EEG datasets. This includes using the methods to differentiate between normal, MCI, and AD subjects [57, 63, 73].

This study shows that the abnormalities caused by AD can be detected by the complexity measures. However, similar changes may be caused by other neurodegenerative diseases, such as other types of dementia. To enhance the diagnostic usefulness of the methods, it may be necessary to develop them further to differentiate between dementias.

4. Conclusions

AD causes changes in the EEG due to loss of memory and cognitive decline and these changes are thought to be associated with functional disconnections among cortical areas resulting from the death of brain cells. Therefore, EEG analysis may provide valuable information about brain dynamics in AD. AD causes a reduction in neuronal activity of the brain and this may be reflected in EEG signals. Nonlinear methods based on EEG complexity approaches have shown promising results in detected changes in the EEG thought to be due to AD. Therefore, EEG complexity can potentially be a good biomarker for AD diagnosis. We investigated three complexity measures, TsEn, HFD, and LZC methods, derived from EEG frequency bands. We found that AD patients have significantly lower TsEn, HFD, and LZC values in specific EEG frequency bands and specific EEG channels compared to normal subjects. This may provide an effective way to discriminate between AD patients and normal subjects. Future work will evaluate the methods using larger and more diverse EEG datasets, including different types of dementia.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article

Acknowledgments

The first author would like to thank the Ministry of Higher Education and Scientific Research (MoHESR), Iraq, for their

financial support. Financial support by the EPSRC is also gratefully acknowledged.

References

- [1] T. Harmony, T. Fernández, J. Gersenowies et al., “Specific EEG frequencies signal general common cognitive processes as well as specific task processes in man,” *International Journal of Psychophysiology*, vol. 53, no. 3, pp. 207–216, 2004.
- [2] C. Humpel, “Identifying and validating biomarkers for Alzheimer’s disease,” *Trends in Biotechnology*, vol. 29, no. 1, pp. 26–32, 2011.
- [3] W. H. Organization, *WHO Mental Health Gap Action Programme (mhGAP) 2008*, 2014.
- [4] ADNI, <http://www.adni-info.org/>.
- [5] A. H. Al-Nuaimi, E. Jammeh, L. Sun, and E. Ifeachor, “Changes in the EEG amplitude as a biomarker for early detection of Alzheimer’s disease,” in *Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2016*, pp. 993–996, USA, August 2016.
- [6] M. Prince, A. Wimo, M. Guerchet, G.-C. Ali, Y. Wu, and M. Prina, *World Alzheimer Report*, 2015.
- [7] M. Prince, A. Comas-Herrera, M. Knapp, M. Guerchet, and M. Karagiannidou, “World Alzheimer Report 2016 Improving healthcare for people living with dementia,” *Coverage, Quality and costs now and in the future*, pp. 1–140, 2016.
- [8] E. Jammeh, C. Carroll, S. Pearson et al., “Using NHS primary care data to identify undiagnosed dementia,” *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 86, no. 11, pp. e4.134–e4, 2015.
- [9] B. Michalowsky, T. Eichler, J. R. Thyrian et al., “Healthcare resource utilization and cost in dementia: Are there differences between patients screened positive for dementia with and those without a formal diagnosis of dementia in primary care in Germany?” *International Psychogeriatrics*, vol. 28, no. 3, pp. 359–369, 2016.
- [10] M. Prince, R. Bryce, and C. Ferri, *The benefits of early diagnosis and intervention World Alzheimer Report 2011*, Elsevier, 2011.
- [11] G. Henderson, E. Ifeachor, N. Hudson et al., “Development and assessment of methods for detecting dementia using the human electroencephalogram,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 8, pp. 1557–1568, 2006.
- [12] A. L. Sutton, Ed., *Alzheimer Disease Sourcebook*, Omnigraphics: Peter E. Ruffner, Detroit, USA, 5th edition, 2011.
- [13] D. Galimberti and E. Scarpani, *Biomarkers for Early Diagnosis of Alzheimer’s Disease*, Nova Biomedical Books, 2008.
- [14] A. Association, *2017 Alzheimer’s disease facts and figures*, Alzheimer’s Dement, 2017.
- [15] K. A. Jellinger, B. Janetzky, J. Attems, and E. Kienzl, “Biomarkers for early diagnosis of Alzheimer disease: ‘ALzheimer ASsociated gene’ - A new blood biomarker?” *Journal of Cellular and Molecular Medicine*, vol. 12, no. 4, pp. 1094–1117, 2008.
- [16] K. Ritchie and S. Lovestone, “The dementias,” *The Lancet*, vol. 360, no. 9347, pp. 1759–1766, 2002.
- [17] R. A. Sperling, P. S. Aisen, L. A. Beckett et al., ““Toward defining the preclinical stages of Alzheimer’s disease: Recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease,” *Alzheimer’s Dement*, vol. 7, no. 3, pp. 333–334, 2011.

- [18] A. Association, "2016 Alzheimer's disease facts and figures," *Alzheimer's Dement*, vol. 12, no. 4, pp. 459–509, 2016.
- [19] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, "Functional connectivity in the resting brain: a network analysis of the default mode hypothesis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 1, pp. 253–258, 2003.
- [20] A. Gevins, M. E. Smith, L. McEvoy, and D. Yu, "High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice," *Cerebral Cortex*, vol. 7, no. 4, pp. 374–385, 1997.
- [21] K. Sasaki, A. Nambu, T. Tsujimoto, R. Matsuzaki, S. Kyuhou, and H. Gemba, "Studies on integrative functions of the human frontal association cortex with MEG," *Cognitive Brain Research*, vol. 5, no. 1-2, pp. 165–174, 1996.
- [22] T. Abeel, T. Helleputte, Y. Van de Peer, P. Dupont, and Y. Saeys, "Robust biomarker identification for cancer diagnosis with ensemble feature selection methods," *Bioinformatics*, vol. 26, no. 3, pp. 392–398, 2009.
- [23] A. Boucheham and M. Batouche, "Robust biomarker discovery for cancer diagnosis based on meta-ensemble feature selection," in *Proceedings of the 2014 Science and Information Conference, SAI 2014*, pp. 452–460, UK, August 2014.
- [24] A. H. Al-nuaimi, E. Jammeh, L. Sun, and E. Ifeachor, "Changes in the Electroencephalogram as a Biomarker of Alzheimer's Disease," in *Proceedings of the Biosense Dementia 2017 - international workshop on biosensors for dementia from 13 - 14 June 2017*, Plymouth University, Plymouth, UK, 2017.
- [25] A. Cedazo-Minguez and B. Winblad, "Biomarkers for Alzheimer's disease and other forms of dementia: clinical needs, limitations and future aspects," *Experimental Gerontology*, vol. 45, no. 1, pp. 5–14, 2010.
- [26] D. Ferreira, V. Jelic, L. Cavallin et al., "Electroencephalography Is a Good Complement to Currently Established Dementia Biomarkers," *Dementia and Geriatric Cognitive Disorders*, vol. 42, no. 1-2, pp. 80–92, 2016.
- [27] K. Blennow and H. Hampel, "CSF markers for incipient Alzheimer's disease," *The Lancet Neurology*, vol. 2, no. 10, pp. 605–613, 2003.
- [28] S. E. O'Bryant, M. M. Mielke, R. A. Rissman et al., "Blood-based biomarkers in Alzheimer disease: Current state of the science and a novel collaborative paradigm for advancing from discovery to clinic," *Alzheimer's & Dementia*, vol. 13, no. 1, pp. 45–58, 2017.
- [29] D. Ferreira, L. Perestelo-Pérez, E. Westman, L.-O. Wahlund, A. Sarrisa, and P. Serrano-Aguilar, "Meta-review of CSF core biomarkers in Alzheimer's disease: The state-of-the-art after the new revised diagnostic criteria," *Frontiers in Aging Neuroscience*, vol. 6, Article ID Article 47, 2014.
- [30] M. Weiner and Z. Khachaturian, *The use of MRI and PET for clinical diagnosis of dementia and investigation of cognitive impairment: a consensus report*, Alzheimer's Assoc., Chicago, USA, 2005.
- [31] B. Dubois, H. Hampel, H. H. Feldman et al., "Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria," *Alzheimer's & Dementia*, vol. 12, no. 3, pp. 292–323, 2016.
- [32] H. Jiang, M. P. White, M. D. Greicius, L. C. Waelde, and D. Spiegel, "Brain activity and functional connectivity associated with hypnosis," *Cerebral Cortex*, vol. 27, no. 8, pp. 4083–4093, 2017.
- [33] M. Signorino, E. Pucci, N. Belardinelli, G. Nolfé, and F. Angeleri, "EEG spectral analysis in vascular and Alzheimer dementia," *Electroencephalography and Clinical Neurophysiology*, vol. 94, no. 5, pp. 313–325, 1995.
- [34] C. Besthorn, R. Zerfass, C. Geiger-Kabisch et al., "Discrimination of Alzheimers disease and normal aging by EEG data," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 2, pp. 241–248, 1997.
- [35] A. H. Al-Nuaimi, E. Jammeh, L. Sun, and E. Ifeachor, "Tsallis entropy as a biomarker for detection of Alzheimer's disease," in *Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2015*, pp. 4166–4169, Italy, August 2015.
- [36] A. Horvath, A. Szucs, G. Csukly, A. Sakovics, G. Stefanics, and A. Kamondi, *EEG and ERP biomarkers of Alzheimer's disease: a critical review*, EEG and ERP biomarkers of Alzheimer's disease, a critical review.
- [37] D. Moretti, C. Fracassi, M. Pievani et al., "Increase of theta/gamma ratio is associated with memory impairment," *Clinical Neurophysiology*, vol. 120, no. 2, pp. 295–303, 2009.
- [38] J. Jeong, "EEG dynamics in patients with Alzheimer's disease," *Clinical Neurophysiology*, vol. 115, no. 7, pp. 1490–1505, 2004.
- [39] Y. A. L. Pijnenburg, Y. Vd Made, A. M. van Cappellen van Walsum, D. L. Knol, P. Scheltens, and C. J. Stam, "EEG synchronization likelihood in mild cognitive impairment and Alzheimer's disease during a working memory task," *Clinical Neurophysiology*, vol. 115, no. 6, pp. 1332–1339, 2004.
- [40] M. Ishikawa, K. Doya, H. Miyamoto, and T. Yamakawa, "Dynamical nonstationarity analysis of resting EEGs in Alzheimer's disease," in *Proceedings of the International Conference on Neural Information Processing*, pp. 921–929, 2007.
- [41] J. Dauwels, F.-B. Vialatte, and A. Cichocki, "On the early diagnosis of Alzheimer's disease from EEG signals: a mini-review," in *Advances in Cognitive Neurodynamics (II)*, pp. 709–716, Springer, 2011.
- [42] C. Babiloni, G. Binetti, E. Cassetta et al., "Sources of cortical rhythms change as a function of cognitive impairment in pathological aging: a multicenter study," *Clinical Neurophysiology*, vol. 117, no. 2, pp. 252–268, 2006.
- [43] D. Abásolo, R. Hornero, C. Gómez, M. García, and M. López, "Analysis of EEG background activity in Alzheimer's disease patients with Lempel-Ziv complexity and central tendency measure," *Medical Engineering & Physics*, vol. 28, no. 4, pp. 315–322, 2006.
- [44] C. Babiloni, C. Del Percio, F. Vecchio et al., "Alpha, beta and gamma electrocorticographic rhythms in somatosensory, motor, premotor and prefrontal cortical areas differ in movement execution and observation in humans," *Clinical Neurophysiology*, vol. 127, no. 1, pp. 641–654, 2016.
- [45] G. G. Yener, D. D. Emek-Savaş, R. Lizio et al., "Frontal delta event-related oscillations relate to frontal volume in mild cognitive impairment and healthy controls," *International Journal of Psychophysiology*, vol. 103, pp. 110–117, 2016.
- [46] A. H. Al-Nuaimi, E. Jammeh, L. Sun, and E. Ifeachor, "Higuchi fractal dimension of the electroencephalogram as a biomarker for early detection of Alzheimer's disease," in *Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2320–2324, Jeju Island, South Korea, July 2017.
- [47] D. Abásolo, R. Hornero, P. Espino, J. Poza, C. I. Sánchez, and R. De La Rosa, "Analysis of regularity in the EEG background

- activity of Alzheimer's disease patients with Approximate Entropy," *Clinical Neurophysiology*, vol. 116, no. 8, pp. 1826–1834, 2005.
- [48] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, 2nd edition, 2004.
- [49] A. Lempel and J. Ziv, "On the Complexity of Finite Sequences," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 75–81, 1976.
- [50] G. Tononi, G. M. Edelman, and O. Sporns, "Complexity and coherency: integrating information in the brain," *Trends in Cognitive Sciences*, vol. 2, no. 12, pp. 474–484, 1998.
- [51] R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths, and H. Scheingraber, "A comparative classification of complexity measures," *Chaos, Solitons & Fractals*, vol. 4, no. 1, pp. 133–173, 1994.
- [52] C. Besthorn, H. Sattel, C. Geiger-Kabisch, R. Zerfass, and H. Förstl, "Parameters of EEG dimensional complexity in Alzheimer's disease," *Electroencephalography and Clinical Neurophysiology*, vol. 95, no. 2, pp. 84–89, 1995.
- [53] A. P. Burgess, J. Rehman, and J. D. Williams, "Changes in neural complexity during the perception of 3D images using random dot stereograms," *International Journal of Psychophysiology*, vol. 48, no. 1, pp. 35–42, 2003.
- [54] F. Gu, X. Meng, E. Shen, and Z. Cai, "Can we measure consciousness with EEG complexities?" *International Journal of Bifurcation and Chaos*, vol. 13, no. 3, pp. 733–742, 2003.
- [55] H. Adeli and S. Ghosh-Dastidar, *Automated EEG-based diagnosis of neurological disorders: Inventing the future of neurology*, CRC Press, 2010.
- [56] J. C. McBride, X. Zhao, N. B. Munro et al., "Spectral and complexity analysis of scalp EEG characteristics for mild cognitive impairment and early Alzheimer's disease," *Computer Methods and Programs in Biomedicine*, vol. 114, no. 2, pp. 153–163, 2014.
- [57] J. Dauwels, K. Srinivasan, M. Ramasubba Reddy et al., "Slowing and loss of complexity in Alzheimer's EEG: Two sides of the same coin?" *International Journal of Alzheimer's Disease*, Article ID 539621, 2011.
- [58] P. Zhao, P. Van-Eetvelt, C. Goh, N. Hudson, S. Wimalaratna, and E. Ifeachor, "Characterization of EEGs in Alzheimer's disease using information theoretic methods," in *Proceedings of the 29th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society (EMBC '07)*, pp. 5127–5131, Lyon, France, August 2007.
- [59] B. Hamadicharef, C. Guan, N. Hudson, E. C. Ifeachor, and S. Wimalaratna, "Performance evaluation and fusion of methods for early detection of Alzheimer Disease," in *Proceedings of the 1st International Conference on BioMedical Engineering and Informatics (BMEI '08)*, pp. 347–351, Sanya, China, May 2008.
- [60] J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López, "Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy," *Physiological Measurement*, vol. 27, no. 11, article 004, p. 1091, 2006.
- [61] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, "Permutation entropy and its main biomedical and econophysics applications: a review," *Entropy*, vol. 14, no. 8, pp. 1553–1577, 2012.
- [62] T. J. De Bock, S. Das, M. Mohsin et al., "Early detection of Alzheimer's disease using nonlinear analysis of EEG via Tsallis entropy," in *Proceedings of the 2010 Biomedical Sciences and Engineering Conference (BSEC)*, pp. 1–4, Oak Ridge, TN, USA, 2010.
- [63] J. Dauwels, F. Vialatte, and A. Cichocki, "Diagnosis of Alzheimer's disease from EEG signals: where are we standing?" *Current Alzheimer Research*, vol. 7, no. 6, pp. 487–505, 2010.
- [64] C. Coronel, H. Garn, M. Waser et al., "Quantitative EEG markers of entropy and auto mutual information in relation to MMSE scores of probable Alzheimer's disease patients," *Entropy*, vol. 19, no. 3, article no. 130, 2017.
- [65] R. Sneddon, W. R. Shankle, J. Hara, A. Rodriguez, D. Hoffman, and U. Saha, "EEG Detection of Early Alzheimer's Disease Using Psychophysical Tasks," *Clinical EEG and Neuroscience*, vol. 36, no. 3, pp. 141–150, 2005.
- [66] H. Garn, M. Waser, M. Deistler et al., "Quantitative EEG markers relate to Alzheimer's disease severity in the Prospective Dementia Registry Austria (PRODEM)," *Clinical Neurophysiology*, vol. 126, no. 3, pp. 505–513, 2015.
- [67] D. Abásolo, S. Simons, R. Morgado da Silva, G. Tononi, and V. V. Vyazovskiy, "Lempel-Ziv complexity of cortical activity during sleep and waking in rats," *Journal of Neurophysiology*, vol. 113, no. 7, pp. 2742–2752, 2015.
- [68] M. W. Rivolta, M. Migliorini, M. Aktaruzzaman, R. Sassi, and A. M. Bianchi, "Effects of the series length on Lempel-Ziv Complexity during sleep," in *Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, pp. 693–696, USA, August 2014.
- [69] X.-S. Zhang, R. J. Roy, and E. W. Jensen, "EEG complexity as a measure of depth of anesthesia for patients," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 12, pp. 1424–1433, 2001.
- [70] M. Aboy, R. Hornero, D. Abásolo, and D. Álvarez, "Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 11, pp. 2282–2288, 2006.
- [71] S. Simons and D. Abásolo, "Distance-based Lempel-Ziv complexity for the analysis of electroencephalograms in patients with Alzheimer's disease," *Entropy*, vol. 19, no. 3, Article ID e19030129, 2017.
- [72] R. Hornero, D. Abásolo, J. Escudero, and C. Gómez, "Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease," *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 367, no. 1887, pp. 317–336, 2009.
- [73] R. Hornero, J. Escudero, A. Fernández, J. Poza, and C. Gómez, "Spectral and nonlinear analyses of MEG background activity in patients with Alzheimer's disease," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 6, pp. 1658–1665, 2008.
- [74] A. Fernandez, R. Hornero, C. Gómez et al., "Complexity analysis of spontaneous brain activity in alzheimer disease and mild cognitive impairment: An MEG study," *Alzheimer Disease & Associated Disorders*, vol. 24, no. 2, pp. 182–189, 2010.
- [75] A. Accardo, M. Affinito, M. Carrozzi, and F. Bouquet, "Use of the fractal dimension for the analysis of electroencephalographic time series," *Biological Cybernetics*, vol. 77, no. 5, pp. 339–350, 1997.
- [76] H. Preißl, W. Lutzenberger, F. Pulvermüller, and N. Birbaumer, "Fractal dimensions of short EEG time series in humans," *Neuroscience Letters*, vol. 225, no. 2, pp. 77–80, 1997.
- [77] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Physica D: Nonlinear Phenomena*, vol. 31, no. 2, pp. 277–283, 1988.
- [78] R. Esteller, G. Vachtsevanos, J. Echaz, and B. Litt, "A comparison of fractal dimension algorithms using synthetic and

- experimental data,” in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '99)*, vol. 3, pp. 199–202, IEEE, Orlando, Fla, USA, July 1999.
- [79] C. Gómez, Á. Mediavilla, R. Hornero, D. Abásolo, and A. Fernández, “Use of the Higuchi’s fractal dimension for the analysis of MEG recordings from Alzheimer’s disease patients,” *Medical Engineering & Physics*, vol. 31, no. 3, pp. 306–313, 2009.
- [80] F. M. Smits, C. Porcaro, C. Cottone, A. Cancelli, P. M. Rossini, and F. Tecchio, “Electroencephalographic fractal dimension in healthy ageing and Alzheimer’s disease,” *PLoS ONE*, vol. 11, no. 2, Article ID e0149587, 2016.
- [81] T. Staudinger and R. Polikar, “Analysis of complexity based EEG features for the diagnosis of Alzheimer’s disease,” in *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2011*, pp. 2033–2036, USA, September 2011.
- [82] L. Hirsch and R. Brenner, *Atlas of EEG in critical care*, John Wiley & Sons, 2011.
- [83] S. Sanei, *Adaptive processing of brain signals*, John Wiley & Sons, 2013.
- [84] M. F. Folstein, S. E. Folstein, and P. R. McHugh, ““Mini mental state”. A practical method for grading the cognitive state of patients for the clinician,” *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, 1975.
- [85] E. Strauss, E. M. S. Sherman, and O. Spreen, *A compendium of neuropsychological tests: Administration, norms, and commentary*, American Chemical Society, 2006.
- [86] A. L. Benton, *Revised Visual Retention Test: Clinical and Experimental Applications*, Psychological Corporation, New York, NY, USA, 1974.
- [87] G. A. Talland and M. Ekdahl, “Psychological studies of korszakoff’s psychosis: IV. The rate and mode of forgetting narrative material,” *The Journal of Nervous and Mental Disease*, vol. 129, no. 4, pp. 391–404, 1959.
- [88] J. Tuqan and P. P. Vaidyanathan, “Optimum low cost two channel IIR orthonormal filter bank,” in *Proceedings of the 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, pp. 2425–2428, April 1997.
- [89] M. Gell-Mann and C. Tsallis, *Nonextensive entropy: interdisciplinary applications*, Oxford University Press, 2004.
- [90] C. Gómez, D. Abásolo, J. Poza, A. Fernández, and R. Hornero, “MEG analysis in Alzheimer’s disease computing approximate entropy for different frequency bands,” in *Proceedings of the 2010 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC’10*, pp. 2379–2382, Argentina, September 2010.
- [91] J. Ramírez, J. M. Górriz, D. Salas-Gonzalez et al., “Computer-aided diagnosis of Alzheimer’s type dementia combining support vector machines and discriminant set of features,” *Information Sciences*, vol. 237, pp. 59–72, 2013.
- [92] R. Chaves, J. Ramírez, J. M. Górriz et al., “SVM-based computer-aided diagnosis of the Alzheimer’s disease using t-test NMSE feature selection with feature correlation weighting,” *Neuroscience Letters*, vol. 461, no. 3, pp. 293–297, 2009.
- [93] J. M. Górriz, J. Ramírez, A. Lassl et al., “Automatic computer aided diagnosis tool using component-based SVM,” in *Proceedings of the 2008 IEEE Nuclear Science Symposium Conference Record, NSS/MIC 2008*, pp. 4392–4395, Germany, October 2008.
- [94] A. So, D. Hooshyar, K. W. Park, and H. S. Lim, “Early diagnosis of dementia from clinical data by machine learning techniques,” *Applied Sciences (Switzerland)*, vol. 7, no. 7, article no. 651, 2017.
- [95] S. Raschka, *An Overview of General Performance Metrics of Binary Classifier Systems*, 2014, arXiv1410.5330.
- [96] D. M. W. Powers, “Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation, School of Informatics and Engineering, Flinders University, Adelaide, Australia,” *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.
- [97] S. Simons, D. Abasolo, and M. Hughes, “Investigation of Alzheimer’s disease EEG frequency components with Lempel-Ziv complexity,” in *Proceedings of the 6th European Conference of the International Federation for Medical and Biological Engineering, MBEC 2014*, pp. 46–49, Croatia, September 2014.
- [98] N. N. Kulkarni and V. K. Bairagi, “Extracting Salient Features for EEG-based Diagnosis of Alzheimer’s Disease Using Support Vector Machine Classifier,” *IETE Journal of Research*, vol. 63, no. 1, pp. 11–22, 2017.
- [99] D. V. Moretti, “Theta and alpha eeg frequency interplay in subjects with mild cognitive impairment: Evidence from EEG, MRI and spect brain modifications,” *Frontiers in Aging Neuroscience*, vol. 7, article no. 31, 2015.
- [100] C. Babiloni, R. Ferri, G. Binetti et al., “Directionality of EEG synchronization in Alzheimer’s disease subjects,” *Neurobiology of Aging*, vol. 30, no. 1, pp. 93–102, 2009.
- [101] U. A. Khan, L. Liu, F. A. Provenzano et al., “Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer’s disease,” *Nature Neuroscience*, vol. 17, no. 2, pp. 304–311, 2014.

Research Article

Multifractal-Multiscale Analysis of Cardiovascular Signals: A DFA-Based Characterization of Blood Pressure and Heart-Rate Complexity by Gender

Paolo Castiglioni ¹, Davide Lazzeroni,² Paolo Coruzzi,³ and Andrea Faini⁴

¹IRCCS Fondazione Don C. Gnocchi, Milan, Italy

²Fondazione Don C. Gnocchi, Parma, Italy

³Department of Medicine and Surgery, University of Parma, Parma, Italy

⁴Department of Cardiology, Istituto Auxologico Italiano, Milan, Italy

Correspondence should be addressed to Paolo Castiglioni; paolo.castiglioni@gmail.com

Received 14 September 2017; Revised 27 November 2017; Accepted 17 December 2017; Published 22 January 2018

Academic Editor: Hugo L. Rufiner

Copyright © 2018 Paolo Castiglioni et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Detrended Fluctuation Analysis (DFA) is a popular method for assessing the fractal characteristics of biosignals, recently adapted for evaluating the heart-rate multifractal and/or multiscale characteristics. However, the existing methods do not consider the beat-by-beat sampling of heart rate and have relatively low scale resolutions and were not applied to cardiovascular signals other than heart rate. Therefore, aim of this work is to present a DFA-based method for joint multifractal/multiscale analysis designed to address the above critical points and to provide the first description of the multifractal/multiscale structure of interbeat intervals (IBI), systolic blood pressure (SBP), and diastolic blood pressure (DBP) in male and female volunteers separately. The method optimizes data splitting in blocks to reduce the DFA estimation variance and to evaluate scale coefficients with Taylor's expansion formulas and maps the scales from beat domains to temporal domains. Applied to cardiovascular signals recorded in 42 female and 42 male volunteers, it showed that scale coefficients and degree of multifractality depend on the temporal scale, with marked differences between IBI, SBP, and DBP and with significant sex differences. Results may be interpreted considering the distinct physiological mechanisms regulating heart-rate and blood-pressure dynamics and the different autonomic profile of males and females.

1. Introduction

Beat-by-beat measures of cardiovascular variables show an intrinsic variability, even when the cardiovascular system is observed in steady-state conditions. These spontaneous changes may reflect the processes underlying the cardiovascular homeostasis. Components of this variability show a fractal nature and in the last two decades different authors suggested that, at least for the heart rate, such components may be the output of a complex system that generates self-similar signals [1, 2]. In fact, the cardiovascular system, like several complex dynamical systems, is composed of interacting subsystems embedded in a fractal structure. In particular, fractal networks of vessels and of nervous and humoral pathways connect and hierarchically regulate local

blood flows among several vascular beds. Accordingly, the cardiovascular system can be regarded as a dissipative system that preserves homeostasis evolving toward a self-organized state, not characterized by any intrinsic scale of time [3]. Properly assessing the fractal components of the spontaneous variability of cardiovascular signals is important, because it may help identifying early alterations in cardiovascular regulatory mechanisms and may contribute to stratifying more precisely the cardiovascular risk.

The first descriptions of cardiovascular self-similarity were based on modeling the heart rate as a time series belonging to the families of fractional Gaussian noises or of fractional Brownian motions and on estimating the corresponding Hurst exponent [4]. Successive studies recognized that such an approach oversimplifies a more complex phenomenon,

because the fractal characteristics of heart rate appear to depend on the scale of the observation [5]. This led some authors to propose multiscale approaches that quantify the cardiovascular complexity by a spectrum of self-similarity coefficients evaluated at different temporal scales [6–8]. Other authors provided evidence of the multifractal nature of heart rate [9–11], which means that the self-similar components of variability result from the superimposition of different fractal processes, interwoven at the same scales. This makes the inadequacy of methods based on monofractal models to describe the cardiovascular complexity even more apparent.

For these reasons, the more recent research in the field of heart-rate variability is aimed at proposing methods that take into account both the scale dependency of self-similarity and its multifractal nature [12, 13]. Following this line of research, the aim of the present study is to describe the multifractal and multiscale characteristics of cardiovascular signals in healthy subjects under controlled conditions. This is done by adapting previously proposed methods of multifractal and multiscale analysis and by comparing three cardiovascular signals frequently recorded in physiological and clinical studies: interbeat interval (IBI, inverse of heart rate), systolic blood pressure (SBP), and diastolic blood pressure (DBP). We expect different fractal dynamics for these three signals because they are influenced by different cardiovascular effectors: DBP is mainly modulated by changes in vascular resistances, SBP by changes in cardiac output, and IBI by changes in cardiac outflows of the autonomic nervous system. Since males and females are characterized by a different autonomic profile [14], the analysis also focuses on sex differences in the multifractal and multiscale dynamics.

2. Methods

2.1. Multifractal-Multiscale DFA. Our estimator of the multifractal-multiscale characteristics of beat-by-beat cardiovascular signals was based on detrended fluctuation analysis (DFA), a method originally proposed for calculating a scale exponent, α , strictly related to Hurst's exponent of monofractal time series [5]. DFA has been successively extended to analyze multifractal time series, obtaining distributions of α coefficients that describe the superposition of different fractal processes [15]. Moreover, DFA has been also extended to provide multiscale evaluations, that is, a spectrum of α coefficients function of the observation scale [6, 16]. Therefore, DFA is a versatile technique easily adaptable for multifractal or multiscale analysis. In this regard, recently Gieraltowski et al. combined both the approaches: they proposed a multifractal *and* multiscale method for the DFA of heart-rate variability, exploiting the possibility of adapting the multifractal DFA algorithm in order to provide estimates separately at different scales [12]. This method was recently applied for modeling heart-rate variability during sleep and blood-pressure variability [17, 18]. In the present study, we followed a similar approach, introducing, however, important variants. These took into account specific properties of the beat-to-beat cardiovascular dynamics that regard the way local slopes are derived and their proper mapping in the time domain.

Given a time series of the cardiovascular variable $x(j)$ with mean μ , evaluated over N consecutive heart beats ($1 \leq j \leq N$), its cumulative sum

$$y(i) = \sum_{j=1}^i (x(j) - \mu) \quad (1)$$

was calculated for $1 \leq i \leq N$. Then, fixing a block size n in number of beats, $y(i)$ was split into M blocks each containing a data segment of n beats. Often DFA is evaluated considering nonoverlapping consecutive blocks: in this case, $M = \text{int}(N/n)$ and a short segment of $N - nM$ data at the end of the series is not included in any of the M blocks if N is not multiple of n [19]. By contrast, we overlapped consecutive segments so that two successive blocks had $n - 1$ beats in common (maximal overlapping), M was equal to $(N - n + 1)$ and all the data were included in at least one block for any size n . As illustrated in Figure 1, maximal overlapping reduces the estimator variance substantially (this will allow evaluating the local slopes α with numerical differentiation formulas). Data were detrended in each of the M blocks with a least-square polynomial fitting of order 1. The standard deviation of the detrended data was calculated in each block k , $\sigma_n(k)$, for $1 \leq k \leq M$.

Data splitting was repeated for block sizes n between 6 and $N/4$ beats. Block sizes were selected as the closest integers to a distribution evenly spaced on a logarithmic scale, with density of about 13 samples per decade. For instance, for $N = 8400$ beats, corresponding to a 2-hour recording at the heart rate of 70 bpm, we considered 34 block sizes n between 6 and 1827 beats. Because of the low size of the smallest block, we set the order of the fitting polynomial equal to 1, to avoid overfitting the data with a too high order, which might remove not only the trend but also the significant components of variability.

According to the multifractal approach for DFA [15], a family of variability functions, $F_q(n)$, which depend on the multifractal parameter q , are calculated for each block size n , as

$$F_q(n) = \left(\frac{1}{M} \sum_{k=1}^M (\sigma_n^2(k))^{q/2} \right)^{1/q} \quad \text{for } q \neq 0 \quad (2)$$

$$F_q(n) = e^{(1/2M) \sum_{k=1}^M \ln(\sigma_n^2(k))} \quad \text{for } q = 0.$$

If $x(i)$ has a power law correlation—like fractional Gaussian noises or fractional Brownian motions—then $F_q(n)$ increases as a power of n for any choice of the parameter q : $F_q(n) \propto n^\alpha$. For monofractal time series, the exponent α is associated with Hurst's exponent H , being $\alpha = H$ for fractional Gaussian noises and $\alpha = H + 1$ for fractional Brownian motions. By contrast, if $x(i)$ is a multifractal series with fractal components of different amplitude, α reflects the superposition of different power law correlations. In particular, α mainly reflects the fractal components with larger amplitude if $q > 0$ and the fractal components with smaller amplitude if $q < 0$. Therefore, α coefficients that depend on q are sign of multifractal dynamics.

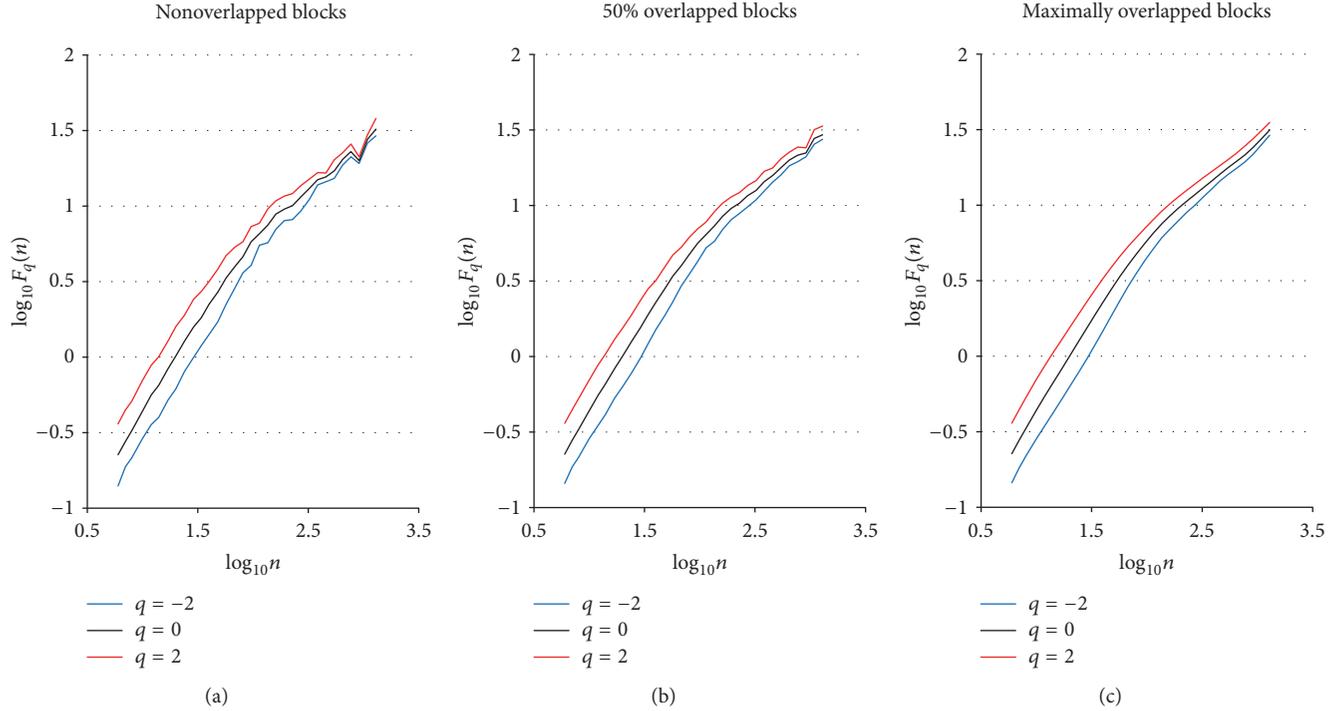


FIGURE 1: Examples of multifractal $F_q(n)$ functions plotted versus the block size n at different q values for nonoverlapped (a), 50% overlapped (b), and maximally overlapped blocks (c). Data from beat-to-beat IBI series of 2-hour duration recorded in a healthy volunteer sitting at rest.

The multifractal exponent $\alpha(q)$ can be estimated from $F_q(n)$ in (2) as the slope of a least-square linear regression between $\log F_q(n)$ and $\log n$ [20]. In this way, however, local deviations from the linear trend occurring at specific scales n cannot be detected. The assessment of local deviations may reflect changes in the sympathetic and vagal cardiac control not otherwise visible [7, 16], revealing subtle alterations in the overall autonomic regulation of the cardiovascular system [21] and characterizing pathological conditions [6, 22]. To evaluate a local slope, that is, α as function of n , methods with higher scale resolution are required. A simple way to obtain a multiscale representation is to calculate the least-square linear regression over a running window with constant width over the $\log n$ axis [12]. The estimated slope, $\alpha(q, n)$, is associated with the central scale n of the running window. However, the length of the running window influences the “smoothness” of the $\alpha(q, n)$ curves and limits the range of scales where α is estimated. Alternatively, $\alpha(q, n)$ could be estimated as derivative of $\log F_q(n)$ versus $\log n$, as proposed for monofractal DFA [16]. We followed this approach and since n was approximately spaced evenly on the logarithmic scale, we applied formula derived from Taylor’s expansion. Let us call $\{n_l\}$ with $1 \leq l \leq l_{\text{MAX}}$ the set of l_{MAX} block sizes, where we calculated $F_q(n)$ in (2). The 3-point expression of the derivative of $\log F_q(n)$ versus $\log n$ is

$$\alpha_B(q, n_l) = \frac{\log F_q(n_{l+1}) - \log F_q(n_{l-1})}{\log(n_{l+1}) - \log(n_{l-1})}. \quad (3)$$

In (3), the pedix B of α_B means that the scale coefficient is evaluated on the beat domain, n . For $l = 1$ and $l = l_{\text{MAX}}$ (3)

is not defined, and we used the expressions for right and left derivatives:

$$\begin{aligned} \alpha_B(q, n_l) &= \frac{-\log F_q(n_{l+2}) + 4 \log F_q(n_{l+1}) - 3 \log F_q(n_l)}{\log(n_{l+2}) - \log(n_l)} \\ &\quad \text{for } l = 1, \end{aligned} \quad (4)$$

$$\begin{aligned} \alpha_B(q, n_l) &= \frac{\log F_q(n_{l-2}) - 4 \log F_q(n_{l-1}) + 3 \log F_q(n_l)}{\log(n_l) - \log(n_{l-2})} \\ &\quad \text{for } l = l_{\text{MAX}}. \end{aligned}$$

Equations (3) and (4) approximate the first derivative of $\log F_q(n)$ versus $\log n$ with errors proportional to the amplitude of the derivatives of order higher than 2. A better approximation is provided by the formula on 5 points with errors proportional to derivatives of order higher than 3. Therefore, instead of (3), for $2 < l < l_{\text{MAX}} - 1$ we used the following equation:

$$\begin{aligned} \alpha_B(q, n_l) &= \frac{8(\log F_q(n_{l+1}) - \log F_q(n_{l-1})) - (\log F_q(n_{l+2}) - \log F_q(n_{l-2}))}{3(\log(n_{l+2}) - \log(n_{l-2}))}. \end{aligned} \quad (5)$$

Figure 2 illustrates how (3)–(5) derive $\alpha_B(q, n)$ in a real case.

Since α_B is a function of the scale expressed in number of beats, when series with different mean heart rate are

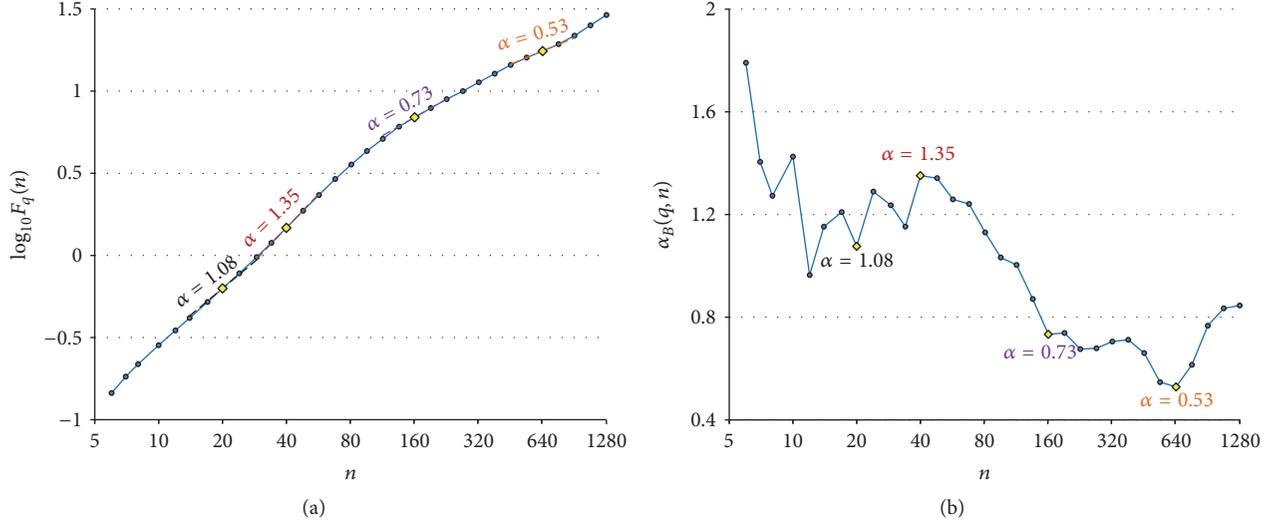


FIGURE 2: Example of local slopes estimation. (a) shows the same $F_q(n)$ function plotted for maximally overlapped blocks and $q = -2$ in Figure 1; the straight lines spanning over 5 points centered around the diamond symbols at $n = 20$, $n = 40$, $n = 161$, and $n = 646$ represent the local slopes $\alpha(n)$ evaluated by the 5-point derivative of (5). (b) shows the corresponding spectrum of local slopes.

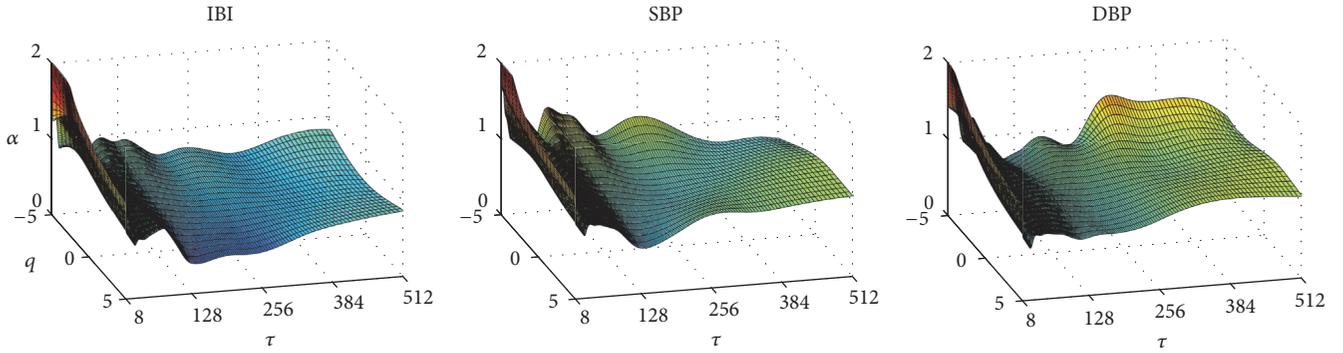


FIGURE 3: Example of multifractal and multiscale coefficients, $\alpha(q, \tau)$, for IBI, SBP, and DBP. Data recorded in a male, normotensive volunteer. In this example, $\alpha(q, \tau)$ of IBI decreases steeply with τ at scales shorter than 16 s, from values typical of fractional Brownian motions ($\alpha > 1$) to values typical of fractional Gaussian noises ($\alpha < 1$). Even if α is lower than 1 at larger scales, the $\alpha(q, \tau)$ surface is not flat and a relative maximum appears at $\tau = 32$ s, more pronounced for positive rather than for negative q values. SBP differs remarkably from IBI: α tends to be greater at scales > 64 s, and pronounced local maxima appear when $q < 0$. DBP differs from IBI and SBP: $\alpha(q, \tau)$ increases at $\tau > 128$ s for any q , reaching values greater than 1; and no local maxima appear at $\tau < 128$ s.

compared (e.g., a bradycardic versus a tachycardic subject or rest versus exercise conditions) the same scales n , in beats, correspond to different temporal scales, τ , in seconds. Therefore, to associate each scale coefficient with its temporal scale, we mapped the beat domain into the time domain [7]. Given the $\alpha_B(q, n_i)$ coefficients evaluated on the beat domain for the $\{n_i\}$ set of scales, the coefficients evaluated on the corresponding set of temporal scales, $\{\tau_i\}$, are

$$\alpha(q, \tau_i) = \alpha_B(q, n_i) \quad \text{for } \tau_i = n_i \times \mu_{\text{IBI}} \quad (6)$$

with μ_{IBI} being the mean IBI, in seconds.

The $\alpha(q, \tau_i)$ coefficients were interpolated over the τ axis with a spline function to obtain estimates at the same temporal scales for each recording. On the basis of the analysis of synthesized series with known self-similarity structure (see Appendix A), for our application on healthy volunteers (see Section 2.3) we interpolated 256 points evenly spaced over

the logarithmic τ axis, between $\tau = 8$ s and $\tau = 512$ s when $q > -3$ and between $\tau = 10$ s and $\tau = 512$ s when $q \leq -3$. The largest scale ($\tau = 512$ s) corresponds to less than 10% the average duration of the recordings. Estimates were obtained for q between -5 and $+5$, with incremental step of 0.5 . Figure 3 shows an example of estimated $\alpha(q, \tau)$ coefficients.

Finally, we defined a concise index of multifractality, function of the scale τ : $\text{MF}_I(\tau)$. For this purpose, fixing a parameter $q_r > 0$, we considered the range $-q_r \leq q \leq q_r$, symmetric around 0 with amplitude $2q_r$. For each temporal scale τ , we calculated the standard deviation of all $\alpha(q, \tau)$ values estimated over the $\pm q_r$ range, $\alpha_{\text{SD}}(\tau)$. The $\text{MF}_I(\tau)$ index is defined as the ratio between $\alpha_{\text{SD}}(\tau)$ and the range of corresponding q values:

$$\text{MF}_I(\tau) = \frac{\alpha_{\text{SD}}(\tau)}{2q_r}. \quad (7)$$

TABLE I: General characteristics of participants by sex.

	N	Age (years)	Body mass index (kg/m ²)	Prevalence of hypertension
Females	42	34.3 (9.7)	22.6 (2.8)	38.1%
Males	42	34.4 (10.1)	23.6 (2.4)	38.1%
p		0.96	0.09	

Since $\alpha_{SD}(\tau) \geq 0$, also $MF_I(\tau) \geq 0$, reaching values close to 0 if the series is monofractal at τ (Appendix A shows an example of $MF_I(\tau)$ for synthesized monofractal series). In this study, $MF_I(\tau)$ was calculated for τ between 10 s and 512 s, setting $q_r = 5$.

2.2. Power Spectral Analysis. The power spectrum of each cardiovascular series was also calculated. Beat-to-beat series were interpolated linearly at 10 Hz and resampled at 5 Hz. The Welch periodogram was estimated by splitting the resampled series in 50% overlapping Hann windows of 1638.4 s duration, by computing the FFT spectrum in each window and by averaging the spectra over all the windows. The final periodogram was smoothed with a broadband procedure [23].

2.3. Subjects and Experimental Protocol. We considered recordings previously collected in two studies aimed at evaluating the influence of sodium sensitivity on the cardiovascular control in normotensive [24] and hypertensive [25] healthy subjects. The original dataset in normotensive subjects consisted of recordings in 26 males and 45 females [24]. For the present analysis, we included all the 26 male participants and a subgroup of 26 female participants matched for age and body mass index. The original dataset in hypertensive subjects consisted of recordings in 30 males and 16 females, and, for the present analysis, we included all the 16 female participants and a subgroup of 16 male participants with age and body mass index matched with the female group. Table I summarizes by sex the general characteristics of the selected 84 participants.

Each participant was studied in a quiet environment in the morning, after 5 days of low-salt diet (30 mmol NaCl per day) to minimize the confounding effects of dietary sodium on cardiovascular variability. Continuous finger arterial blood pressure was recorded for about two hours, in sitting position at rest, by Portapres model-2 (Finapres Medical Systems B. V., Amsterdam, Netherlands). The finger cuff was placed on the mid finger of the left hand. SBP, DBP, and IBI (calculated as time interval between consecutive SBP values) were derived beat-by-beat for the whole duration of the recording. Brachial blood pressure was measured simultaneously with a cuff on the right arm every 15 minutes, and the SBP and DBP readings of the brachial device were used to calibrate beat-by-beat SBP and DBP values from the finger cuff.

2.4. Statistics. We described statistical patterns in $\alpha(q, \tau)$ estimates showing means and standard error of the means over the group, on the basis of previous observations reporting

that DFA coefficients follow a normal distribution [26] and in $MF_I(\tau)$ estimates showing median and standard error of the median, this latter estimated by bootstrapping using 100 bootstrap samples. Statistical inferences were performed with nonparametric tests for all the estimates not to make any assumption on the distribution of scale coefficients and multifractal indices at any τ . In particular, $\alpha(q, \tau)$ coefficients were compared between signals (IBI versus SBP, IBI versus DBP and SBP versus DBP) by the paired Wilcoxon test; $\alpha(q, \tau)$ and $MF_I(\tau)$ were compared between males and females by the unpaired Mann–Whitney test. Power spectra were compared between genders by unpaired t -test after log-transformation, to obtain normal distributions of power spectra [27]. The analyses were performed with “R: A Language and Environment for Statistical Computing” software package (R Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2017).

3. Results

Figure 4 shows α as a function of τ for specific q values over the whole group (for comparison, the traditional multiscale analysis corresponds to α values evaluated for $q = 2$ only, and the traditional multifractal analysis corresponds to the generalized Hurst exponents shown in Appendix B). Figure 4 confirms patterns suggested in the example of Figure 3. IBI coefficients decrease with τ from values greater than 1 (as for fractional Brownian motions) at the shorter scales to values lower than 1 (as for fractional Gaussian noises) at the larger scales, with a minimum at τ around 250 s. A relative maximum appears at $\tau \cong 30$ s for $q \geq 2$. Moreover, α increases as q decreases, at any τ . At the shorter scales also α of SBP decreases steeply with τ from values >1 , with greater α estimates at lower q values. However, unlike α of IBI, it remains stable around 1 (as for “ $1/f$ ” processes) when $\tau > 30$ s. Similarly to IBI and to SBP, also α of DBP decreases with τ at the shorter scales. However, unlike IBI and SBP, at larger scales it shows an increasing trend with τ .

Figure 5 compares scale coefficients among signals, at different q . Comparing IBI with SBP when $q = 0$ and $q = 4$, α is significantly greater for SBP almost at all the scales; however, this is not the case for scales τ between 16 and 35 s, where α of IBI shows a local maximum. When $q = -4$, differences between IBI and SBP are less significant, and their scale coefficients coincide at scale $\tau \leq 16$ s.

Similarly, comparing IBI and DBP, α is greater for DBP almost at all the scales when $q = 0$ and $q = 4$. In this case, however, at the scales where α of IBI displays a relative maximum, α of DBP shows an absolute minimum, becoming as a result significantly lower than the IBI scale coefficients.

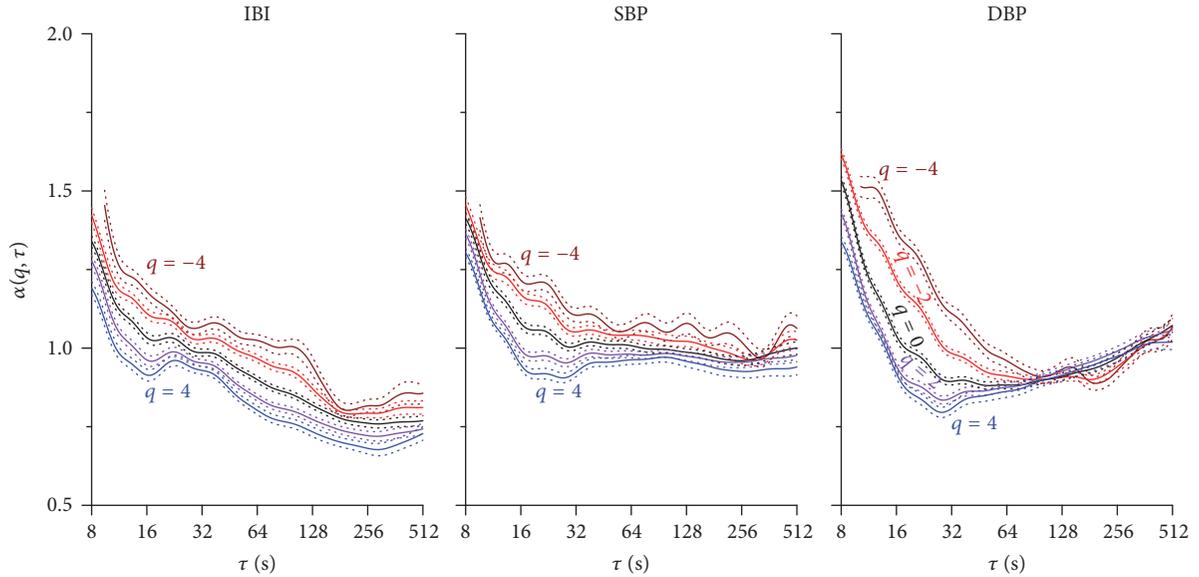


FIGURE 4: Multifractal and multiscale coefficients, $\alpha(q, \tau)$. Mean \pm standard error of the mean over the group ($N = 84$) as function of τ for five q values between -4 and $+4$. For clarity, the horizontal axis is plotted in a log scale.

Also SBP and DBP scale coefficients differ importantly. For $q = 0$ and $q = 4$, α of SBP is significantly greater within $20 < \tau < 100$ s. Similar differences appear for $q = -4$ but in a slightly higher band ($45 < \tau < 240$ s). At $\tau \leq 16$ s, α is greater for DBP when $q = -4$ while when $q = +4$, SBP and DBP scale coefficients coincide.

These results make it clear that the degree of multifractality is a function of τ and of the type of cardiovascular signal. This is summarized by Figure 6, which shows the index of multifractality, $MF_I(\tau)$, for IBI, SBP, and DBP separately. IBI reaches its highest degree of multifractality at $\tau = 10$ s; $MF_I(\tau)$ of IBI decreases at larger scales up to a minimum at $\tau = 30$ s. A different pattern characterizes $MF_I(\tau)$ of DBP and SBP: the highest degree of multifractality is not reached at the shortest scale ($\tau = 10$ s) as for IBI but between 16 and 32 s; and the lowest degree is reached at $\tau = 64$ s.

Gender Differences. Figure 7 compares $\alpha(q, \tau)$ for $q = -4$, $q = 0$, and $q = +4$, in males and females. When $q = +4$ or $q = 0$, α is greater in females at the larger scales (i.e., $\tau > 60$ s for IBI and SBP; $\tau > 20$ s for DBP). These differences vanish when $q = -4$. In addition, when $q = 0$ and $q = -4$, α is lower in females at $\tau \leq 10$ s. The fact that differences between males and females depend on q suggests sex-related differences also in the level of multifractality: these are actually highlighted by Figure 8, which shows a higher degree of multifractality in males, at scales shorter than 16 s for IBI and at scales centered around 20 s for SBP and 32 s for DBP.

Figure 9 compares IBI, SBP, and DBP power spectra by gender. The three signals have common spectral patterns, all showing a peak around 0.10 Hz and a “ $1/f$ ” component at frequencies lower than 0.03 Hz; moreover, IBI and SBP spectra also show a respiratory component at frequencies around 0.30 Hz. Although these patterns appear in both

sexes, spectra differ significantly between males and females. IBI spectral powers are greater in males between 0.008 and 0.13 Hz. Also SBP spectral components are greater in males, but in a larger band including all frequencies higher than 0.004 Hz. By contrast, DBP spectra coincide in men and women, with exclusion of the spectral peak around 0.10 Hz, higher in men.

4. Discussion

We presented a novel algorithm for quantifying cardiovascular complexity based on previous researches that in various ways adapted DFA for assessing multifractal or/and multiscale aspects of heart-rate variability. By applying our method to data collected in healthy volunteers, we provided the first detailed description of differences in multifractal and multiscale features among those cardiovascular time series more often recorded in clinical settings or in physiological studies: IBI, SBP, and DBP.

Three were the main results of our study. First, not only do self-similarity coefficients depend on the observational scale, but also the way α changes with τ depends on the cardiovascular series (IBI, SBP, or DBP). Second, the degree of multifractality also depends on τ and on the type of cardiovascular signal. Third, at the scales where the signals show a multifractal nature (e.g., at $\tau < 32$ s), increasing the multifractal index q from negative to positive values progressively decreases the estimate of α , thus indicating that fractal components with the lower α contribute more to the overall variability because positive q values emphasize components with larger amplitude. In addition, we also showed significant gender differences in these complex self-similarity structures. Although this study was designed to provide a solid description of the multifractal and multiscale features

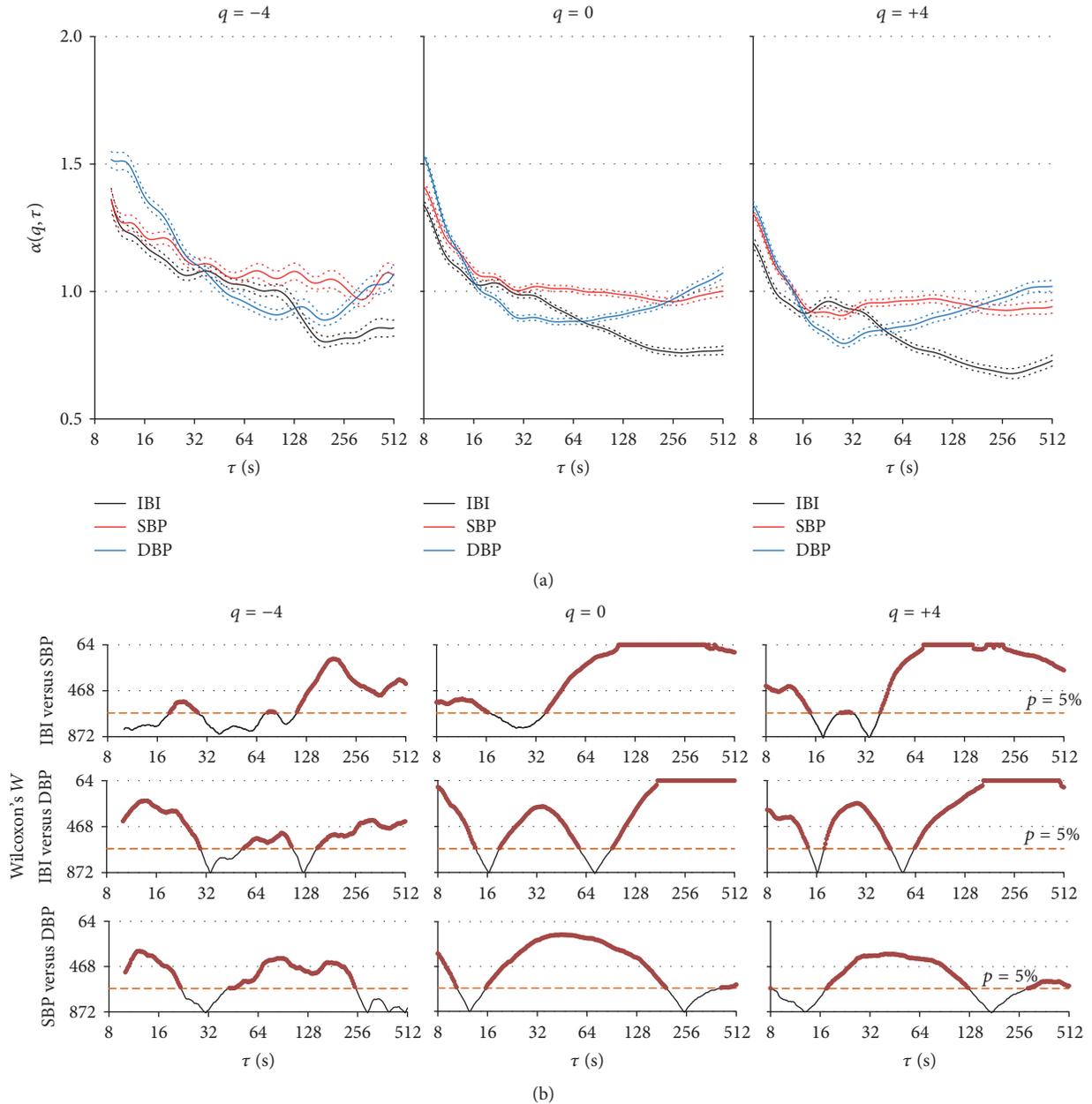


FIGURE 5: Comparison of $\alpha(q, \tau)$ among cardiovascular signals at different q values. (a) Mean \pm standard error of the mean over the group ($N = 84$) as function of τ . (b), from top to bottom: W statistics of Wilcoxon's tests at each scale τ for comparing IBI and SBP, IBI and DBP, and SBP and DBP. Dashed horizontal lines represent the 5% threshold of statistical significance; W values above the threshold are marked by a red dot and indicate statistically significant differences at $p < 5\%$.

of cardiovascular signals and not to find the mechanisms responsible for these features, in the following we may try to speculate on the possible origin and physiological meaning of our results.

We showed that $\alpha(q, \tau)$ of IBI decreases with τ in healthy individuals sitting at rest, from values typical for fractional Brownian motions to values typical for fractional Gaussian noises. A similar behavior has been previously observed in other studies that applied DFA with traditional monofractal approaches (i.e., with $q = 2$ only) to analyze the heart-rate

variability in sitting volunteers at rest [16, 28]. A possible explanation for this trend has been already proposed, based on the hypothesis that the heart-rate dynamics depend on the superposition of two fractal processes simultaneously modulating the heart rate [7, 29]. One process, with relatively homogeneous fractal characteristics at all scales, resembling fractional Gaussian noise, depends on the cardiac vagal outflow. The other process depends on the cardiac sympathetic outflow: it appears as a fractional Gaussian noise at the longer scales and as a Brownian motion at the shortest scales,

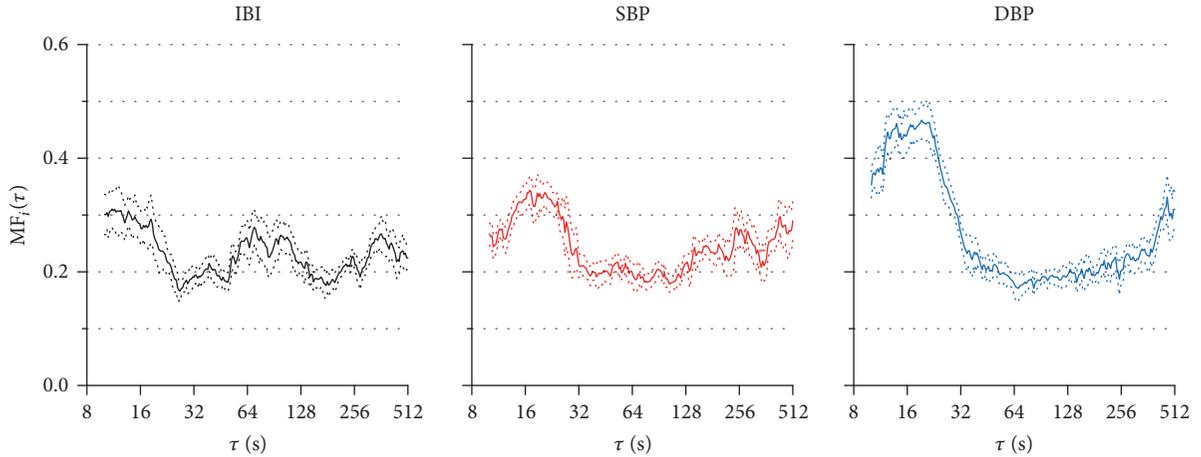


FIGURE 6: Multifractal index $MF_1(\tau)$ for IBI, SBP, and DBP. Median \pm standard error of the median over the group ($N = 84$).

due to a low-pass filter with time constant of 66 s caused by the time of removal of noradrenaline released by the sympathetic endings. The mixture of these two processes may therefore explain the α decreasing trend with τ from fractional Brownian motion to fractional Gaussian noise and may explain why α , when evaluated at the shortest scales only, can be considered index of sympathovagal balance [30]. However, if this hypothesis holds, it should predict some specific features of the multifractal dynamics of heart rate. First, we should expect that the heart rate is multifractal at the shortest scales, where it depends on two processes with different dynamics (fractional Gaussian noise for the vagal modulations, Brownian motion for the sympathetic modulations) and with different amplitude (greater for vagal than for sympathetic modulations). Second, we should expect that multifractality decreases at scales τ larger and closer to the time constant of the low-pass filter on the sympathetic outflow, because amplitude and fractal dynamics of sympathetic heart-rate modulations should become more similar to amplitude and fractal dynamics of vagal modulations of heart rate. Third, we should also expect that, at scales shorter than the time constant of the low-pass filter modeling the sympathetic outflow dynamics ($= 66$ s), α increases when q decreases. In fact, negative q exponents in (2) amplify the contribution of the fractal components with lower amplitude and reduce the contribution of the fractal components with higher amplitude in multifractal dynamics. Since at the shorter scales the sympathetic modulations of heart rate have lower amplitude than the vagal heart-rate modulations, negative q values should emphasize the Brownian motion contribution of the sympathetic outflow (with high α) rather than the fractal noise contribution of the vagal outflow (with low α). These three properties are actually demonstrated by our results on $\alpha(q, \tau)$ of IBI, which therefore support our interpretative hypothesis.

Interestingly, we found gender differences in $\alpha(q, \tau)$ of IBI at $\tau \leq 12$ s but only for $q = 0$ and $q = -4$. When $q = +4$, in fact, α coefficients are the same in males and females over these scales (Figure 7). The IBI power spectra

(Figure 9) indicate that males and females have the same amplitude for spectral components falling in the high-frequency band (>0.15 Hz) where only vagal modulations of heart rate are present and that males have higher amplitude of spectral components in the low-frequency band (around 0.1 Hz) where the heart rate is modulated importantly by both sympathetic and vagal outflows [31]. This means that, among our volunteers, males and females have a similar vagal tone, but males have a higher sympathovagal balance. In our interpretative hypothesis, α at the shorter scales represents mainly the vagal fractal modulation when q is high, and when q decreases the sympathetic contribution to α increases. Coherently, we observed that α at the shorter scales is greater in males, as expected for an index of sympathovagal balance, but only for low values of q . No differences are found for $q = +4$, when we expect α to represent mainly the vagal component of the fractal dynamics, in line with the observation that the high-frequency spectral power associated with pure vagal modulations of heart rate are the same in males and females. The presence of similar vagal modulations of heart rate in males and females and of a higher sympathetic tone in males would also explain the higher IBI multifractal index in males at the shorter scales (Figure 8).

We found marked differences between $\alpha(q, \tau)$ of IBI and DBP (Figure 5). While $\alpha(q, \tau)$ of IBI should reflect a mixture of fractal processes produced by the autonomic nervous system through vagal and sympathetic modulations of heart rate, $\alpha(q, \tau)$ of DBP is expected to mainly reflect modulations of total peripheral resistances, which the autonomic nervous system modulates through vascular sympathetic outflows only. This would explain why the highest degree of multifractality occurs at the shortest scale for IBI, where the two branches of the cardiac autonomic nervous system modulate the heart rate with different fractal dynamics. By contrast, DBP shows the highest degree of multifractality between 16 and 32 s (Figure 6). A possible explanation is that the total peripheral resistance results from the combined effect of individual vascular resistances of several vascular beds, each with its own local regulation hierarchically controlled at

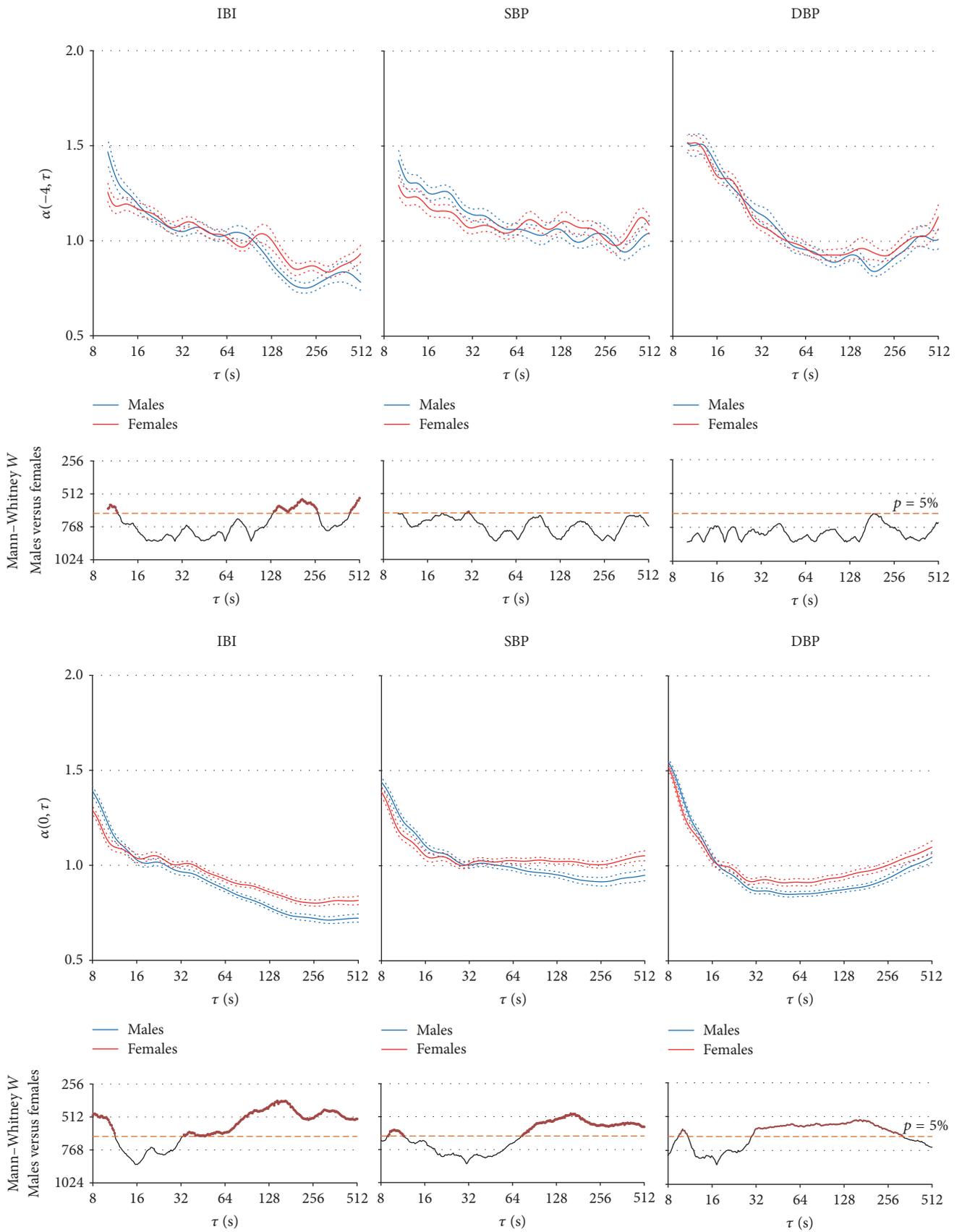


FIGURE 7: Continued.

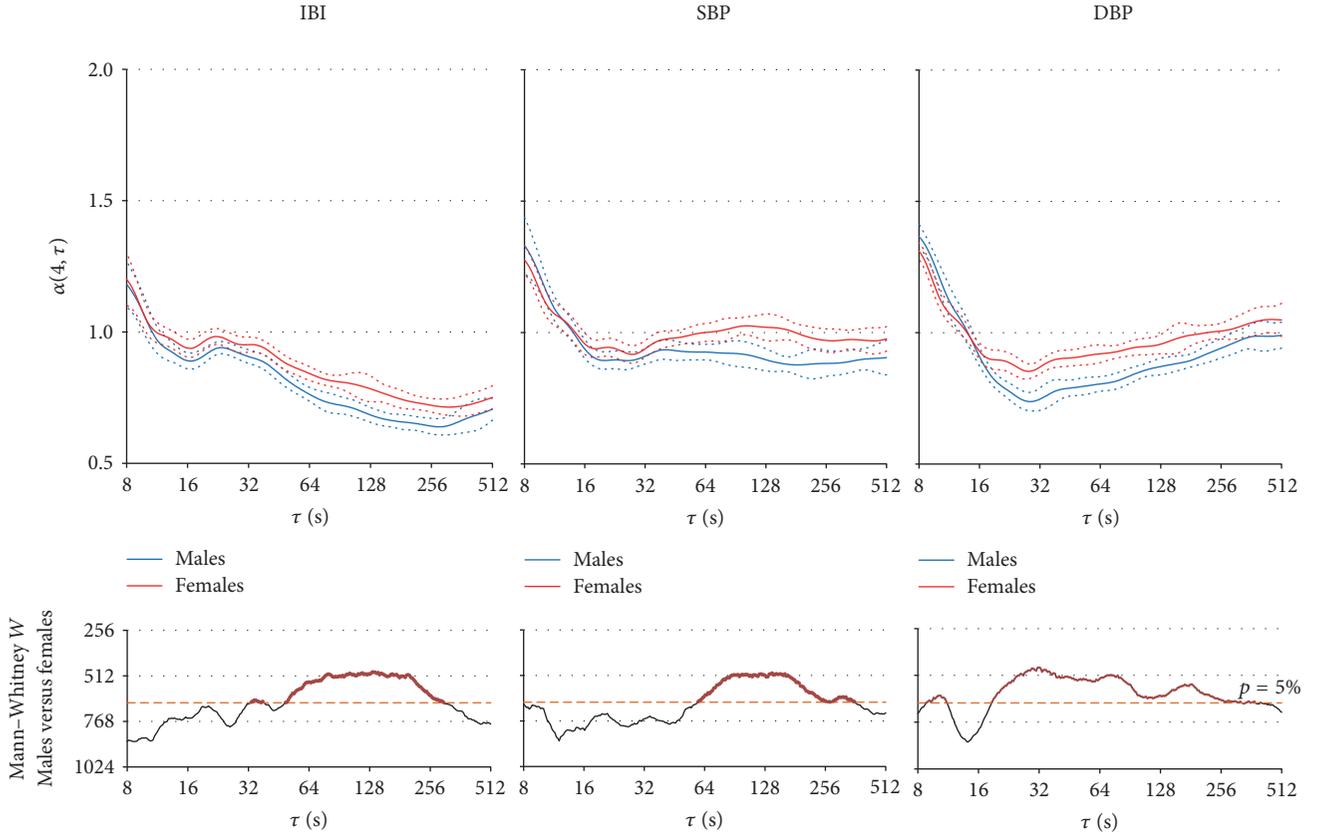


FIGURE 7: Multifractal-multiscale coefficients, $\alpha(q, \tau)$: comparison by sex. Mean \pm standard error of the mean (upper panels) and W statistics of Mann–Whitney tests (lower panels) for comparing males ($N = 42$) and females ($N = 42$), as function of τ . Dashed horizontal lines represent the 5% significance threshold; W values above the threshold (red dots) indicate differences significant at $p < 5\%$. From top to bottom: $q = -4$; $q = 0$; $q = +4$.

the central level through vascular sympathetic outflows. Therefore, at scales around 16 and 32 s the DBP multifractal dynamics would reflect the superimposition of various local regulations associated with different vascular districts.

A final consideration regards the different physiological information derivable from traditional spectral analyses and from a multifractal-multiscale analysis of cardiovascular signals. We found gender differences in $\alpha(q, \tau)$ but not in power spectra at the corresponding frequencies, and vice versa we found power spectral differences without significant differences in the fractal structure at the corresponding scales. For instance, $\alpha(q, \tau)$ of DBP is greater in females for τ between 32 and 256 s and $q = 0$ or $q = +4$, but at the corresponding frequencies (between 0.031 and 0.004 Hz) the DBP spectra are exactly the same in males and females. On the other hand, $\alpha(q, \tau)$ of SBP is very similar in males and females at $\tau < 64$ s, while the SBP spectra differ markedly by gender at frequencies higher than 0.01 Hz. This means that complexity methods and spectral methods are complementary approaches and that one method of analysis may reveal aspects of cardiovascular dynamics that go undetected with the other method. This is a promising perspective for designing new clinical tools based on multifractal-multiscale analysis that, used in

addition to traditional frequency-domain and time-domain methods, might substantially increase the risk stratification in the healthy population or improve the prediction of adverse events in cardiac patients.

Appendix

A. Multifractal-Multiscale DFA of Synthesized Time Series

To evaluate the range of scales where our algorithm provides reliable estimates of multifractal coefficients, we applied it on synthesized series with known self-similarity structure. For this aim, we generated 100 series each of $N = 8400$ samples to simulate a 2-hour beat-by-beat cardiovascular recording at the mean heart rate of 70 bpm, which corresponds to the average heart rate of the participants to our study. The series were generated by the MATLAB *pinknoise* function (version 1.6, made available at <http://goo.gl/PiiPw7> by H. Zhivomirov), which is expected to simulate a pure “ $1/f$ ” monofractal process with $\alpha = 1$. We selected a pink-noise generator because the “ $1/f$ ” process is considered the monofractal noise that better reproduces the fractal structure of cardiovascular signals. The $F_q(n)$ functions shown in Figure 10(a) were

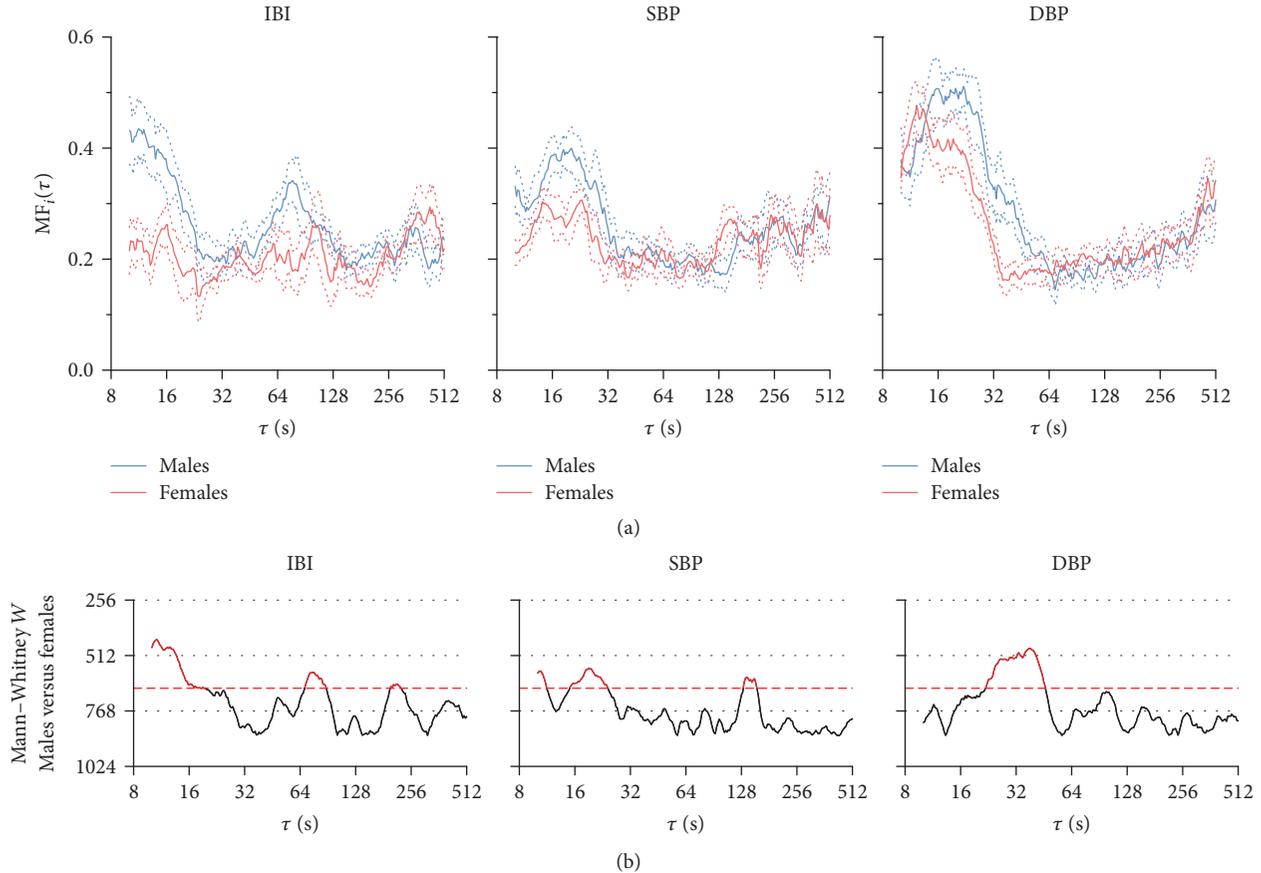


FIGURE 8: Multifractal index $MF_I(\tau)$: comparison by sex. (a) Median \pm standard error of the median; (b) W statistics of Mann-Whitney tests for comparing males ($N = 42$) and females ($N = 42$), as function of τ . Dashed horizontal lines represent the 5% significance threshold; W values above the threshold (red dots) indicate differences significant at $p < 5\%$.

calculated as described in methods for block sizes n between 6 and $N/4$ beats and q between -5 and $+5$. They appear as parallel straight lines with deviations from the constant slope at the shorter blocks for $q \leq -3$. The corresponding multifractal-multiscale coefficients $\alpha_B(q, n)$ calculated as in (3)–(5) are shown in Figure 10(b). The estimated coefficients are close to the $\alpha = 1$ theoretical value that characterizes “ $1/f$ ” processes almost over all the scales, with the exception of the estimates for $q \leq -3$ which deviate largely from the theoretical value at blocks shorter than $n = 12$ beats (corresponding to the time scale of 10 s at 70 bpm). On the basis of these results, we considered reliable estimates of $\alpha(q, \tau)$ for $8 \text{ s} \leq \tau \leq 512 \text{ s}$ when $q > -3$ and for $10 \text{ s} \leq \tau \leq 512 \text{ s}$ when $q \leq -3$.

The multifractal index of monofractal signals should, in theory, be equal to zero. In practice, however, we expect $MF_I(\tau)$ values greater than zero also for monofractal signals for two reasons. At the shorter block sizes, deviations from the constant slope α may occur for negative q values, as we observed in Figure 10. In addition, at all the block sizes α might not be the same when evaluated at different q values even for monofractal series because of the intrinsic variability of the estimate. The estimator variability may increase with the time scale because at the larger block sizes n the

number M of independent blocks for estimating (2) decreases. Figure 11 shows $MF_I(\tau)$ estimated for the synthesized series of Figure 10 and plotted for τ between 10 s and 512 s, where $\alpha(q, \tau)$ can be reliably estimated at all q . As expected, $MF_I(\tau)$ decreases from the shortest scale up to a minimum at $\tau = 16 \text{ s}$ and shows an increasing trend for $\tau > 16 \text{ s}$, but remaining lower than 0.2 over the whole range of scales considered in this study.

B. Traditional Multifractal Analysis

As reference, we also performed a traditional multifractal analysis for each cardiovascular series. For this purpose, we calculated the generalized Hurst exponents, $h(q)$, as slope of the regression line fitting $\log F_q(n)$ and $\log n$ over all the block sizes n for each q between -5 and $+5$, using the MATLAB implementation provided in [20]. Mean and 95% confidence intervals are shown in Figure 12. The analysis reveals greater generalized Hurst exponents for blood pressure than for heart rate at all q values. It also shows that the generalized Hurst exponents are greater at the lower q than at the higher q for all the cardiovascular signals, the relation between $h(q)$ and q decreasing almost linearly for IBI. These features, highlighted by the traditional multifractal analysis, are also

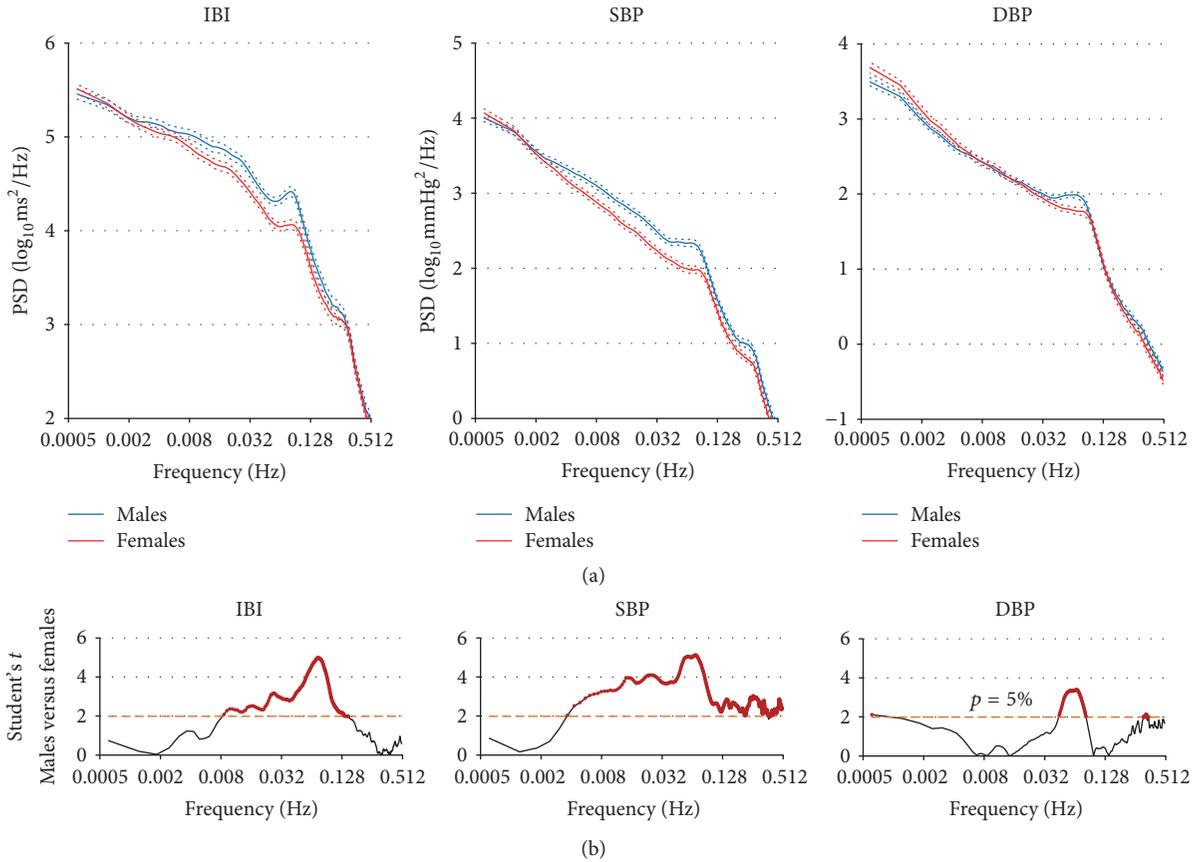


FIGURE 9: Power spectra of IBI, SBP, and DBP: comparison by sex. (a) Mean \pm standard error of the mean; (b) Student's t statistics at each frequency for comparing males ($N = 42$) versus females ($N = 42$). Dashed horizontal lines represent the 5% significance threshold; t values above the threshold (red dots) indicate differences significant at $p < 5\%$.

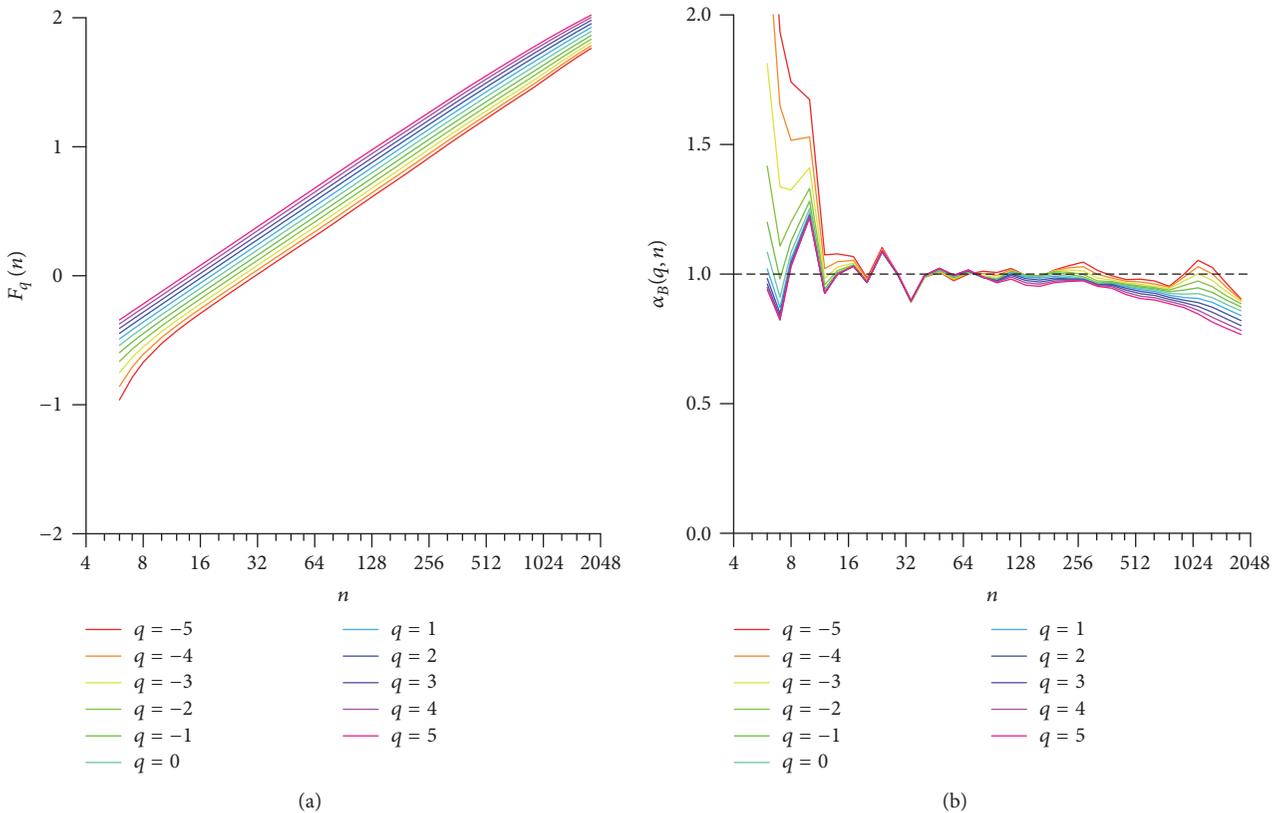


FIGURE 10: Multifractal-multiscale analysis of synthesized series. (a) Average of $F_q(n)$ functions evaluated for 100 synthesized series of pink noise. (b) Average of $\alpha_B(q, n)$ scale coefficients for the same 100 pink-noise signals; note the large deviations from $\alpha = 1$ when $q \leq -3$ at block sizes $n < 12$ beats.

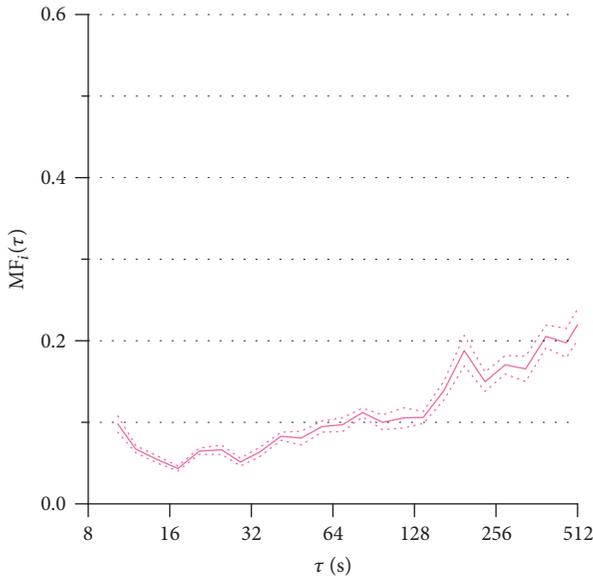


FIGURE 11: Multifractal index of pink noise. Median \pm standard error of the median for the 100 synthesized series of Figure 10.

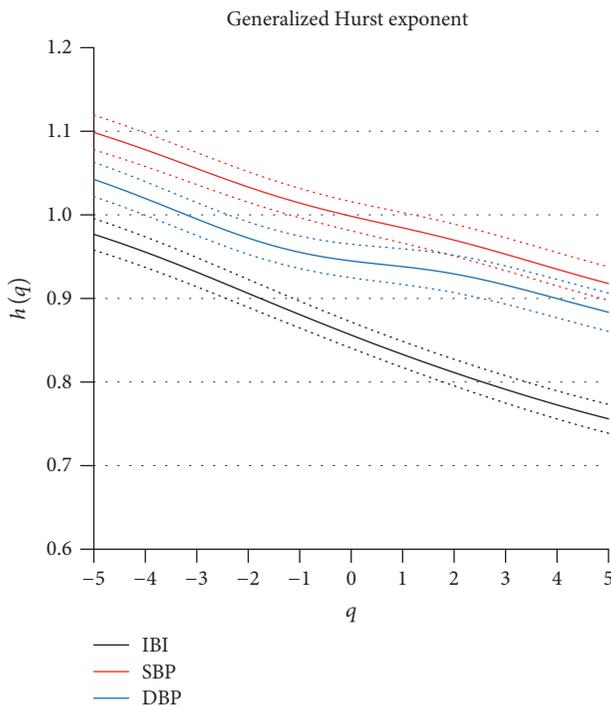


FIGURE 12: Generalized Hurst exponents, $h(q)$, for IBI, SBP, and DBP. Mean and 95% confidence intervals over the group of $N = 84$ participants.

described by the multifractal-multiscale approach proposed in this work. However, unlike the proposed multifractal-multiscale approach, traditional multifractal analysis cannot indicate that these features regard some scales more than others.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] A. Eke, P. Herman, L. Kocsis, and L. R. Kozak, "Fractal characterization of complexity in temporal physiological signals," *Physiological Measurement*, vol. 23, no. 1, pp. R1–R38, 2002.
- [2] R. Sassi, S. Cerutti, F. Lombardi et al., "Advances in heart rate variability signal analysis: joint position statement by the e-Cardiology ESC Working Group and the European Heart Rhythm Association co-endorsed by the Asia Pacific Heart Rhythm Society," *Europace*, vol. 17, no. 9, pp. 1341–1353, 2015.
- [3] P. Bak, C. Tang, and K. Wiesenfeld, "Self-organized criticality: an explanation of the $1/f$ noise," *Physical Review Letters*, vol. 59, no. 4, pp. 381–384, 1987.
- [4] A. Eke, P. Hermán, J. Bassingthwaighe et al., "Physiological time series: distinguishing fractal noises from motions," *Pflügers Archiv - European Journal of Physiology*, vol. 439, no. 4, pp. 403–415, 2000.
- [5] C.-K. Peng, S. Havlin, J. M. Hausdorff, J. E. Mietus, H. E. Stanley, and A. L. Goldberger, "Fractal mechanisms and heart rate dynamics: Long-range correlations and their breakdown with disease," *Journal of Electrocardiology*, vol. 28, pp. 59–64, 1995.
- [6] E. R. Bojorges-Valdez, J. C. Echeverría, R. Valdés-Cristerna, and M. A. Pêa, "Scaling patterns of heart rate variability data," *Physiological Measurement*, vol. 28, no. 6, article no. 010, pp. 721–730, 2007.
- [7] P. Castiglioni, G. Parati, M. Di Rienzo, R. Carabalona, A. Cividjian, and L. Quintin, "Scale exponents of blood pressure and heart rate during autonomic blockade as assessed by detrended fluctuation analysis," *The Journal of Physiology*, vol. 589, no. 2, pp. 355–369, 2011.
- [8] J. Xia, P. Shang, and J. Wang, "Estimation of local scale exponents for heartbeat time series based on DFA," *Nonlinear Dynamics*, vol. 74, no. 4, pp. 1183–1190, 2013.
- [9] R. Sassi, M. G. Signorini, and S. Cerutti, "Multifractality and heart rate variability," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 19, no. 2, Article ID 028507, 2009.
- [10] M. Meyer and O. Stiedl, "Self-affine fractal variability of human heartbeat interval dynamics in health and disease," *European Journal of Applied Physiology*, vol. 90, no. 3-4, pp. 305–316, 2003.
- [11] P. C. Ivanov, L. A. Nunes Amaral, A. L. Goldberger et al., "Multifractality in human heartbeat dynamics," *Nature*, vol. 399, pp. 461–465, 1999.
- [12] J. Gierałtowski, J. J. Zebrowski, and R. Baranowski, "Multiscale multifractal analysis of heart rate variability recordings with a large number of occurrences of arrhythmia," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 85, no. 2, Article ID 021915, 2012.
- [13] D. Makowiec, A. Dudkowska, R. Gałaska, and A. Rynkiewicz, "Multifractal estimates of monofractality in RR-heart series in power spectrum ranges," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 17, pp. 3486–3502, 2009.
- [14] J. Sztajzel, M. Jung, and A. Bayes De Luna, "Reproducibility and gender-related differences of heart rate variability during all-day activity in young men and women," *Annals of Noninvasive Electrocardiology*, vol. 13, no. 3, pp. 270–277, 2008.

- [15] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, "Multifractal detrended fluctuation analysis of nonstationary time series," *Physica A: Statistical Mechanics and its Applications*, vol. 316, no. 1–4, pp. 87–114, 2002.
- [16] P. Castiglioni, G. Parati, A. Civijian, L. Quintin, and M. D. Rienzo, "Local scale exponents of blood pressure and heart rate variability by detrended fluctuation analysis: Effects of posture, exercise, and aging," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 3, pp. 675–684, 2009.
- [17] M. Soliński, J. Gieraltowski, and J. Zebrowski, "Modeling heart rate variability including the effect of sleep stages," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 26, no. 2, Article ID 023101, 2016.
- [18] P. Castiglioni, D. Lazzeroni, V. Brambilla, P. Coruzzi, and A. Faini, "Multifractal multiscale dfa of cardiovascular time series: Differences in complex dynamics of systolic blood pressure, diastolic blood pressure and heart rate," in *Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3477–3480, Jeju Island, South Korea, July 2017.
- [19] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde, "Detecting long-range correlations with detrended fluctuation analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 295, no. 3–4, pp. 441–454, 2001.
- [20] E. A. F. Ihlen, "Introduction to multifractal detrended fluctuation analysis in Matlab," *Frontiers in Physiology*, vol. 3, Article ID Article 141, 2012.
- [21] P. Castiglioni and G. Merati, "Fractal analysis of heart rate variability reveals alterations of the integrative autonomic control of circulation in paraplegic individuals," *Physiological Measurement*, vol. 38, no. 5, pp. 774–786, 2017.
- [22] J. C. Echeverría, M. S. Woolfson, J. A. Crowe, B. R. Hayes-Gill, G. D. H. Croaker, and H. Vyas, "Interpretation of heart rate variability via detrended fluctuation analysis and $\alpha\beta$ filter," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 13, no. 2, pp. 467–475, 2003.
- [23] M. Di Rienzo, P. Castiglioni, G. Parati, G. Mancina, and A. Pedotti, "Effects of sino-aortic denervation on spectral characteristics of blood pressure and pulse interval variability: A wide-band approach," *Medical & Biological Engineering & Computing*, vol. 34, no. 2, pp. 133–141, 1996.
- [24] P. Castiglioni, G. Parati, D. Lazzeroni et al., "Hemodynamic and Autonomic Response to Different Salt Intakes in Normotensive Individuals," *Journal of the American Heart Association*, vol. 5, no. 8, p. e003736, 2016.
- [25] P. Castiglioni, G. Parati, L. Brambilla et al., "Detecting sodium-sensitivity in hypertensive patients: Information from 24-hour ambulatory blood pressure monitoring," *Hypertension*, vol. 57, no. 2, pp. 180–185, 2011.
- [26] P. Castiglioni, G. Parati, C. Lombardi, L. Quintin, and M. Di Rienzo, "Assessing the fractal structure of heart rate by the temporal spectrum of scale exponents: A new approach for detrended fluctuation analysis of heart rate variability," *Biomedizinische Technik. Biomedical Engineering*, vol. 56, no. 4, pp. 175–183, 2011.
- [27] P. Castiglioni, G. Parati, S. Omboni et al., "Broad band spectral analysis of 24 h continuous finger blood pressure: Comparison with intra-arterial recordings," *Clinical Science*, vol. 97, no. 2, pp. 129–139, 1999.
- [28] A. C. A. De Souza, J. R. Cisternas, L. C. De Abreu et al., "Fractal correlation property of heart rate variability in response to the postural change maneuver in healthy women," *International Archives of Medicine*, vol. 7, no. 1, article no. 25, 2014.
- [29] P. Castiglioni, "Self-similarity in Physiological Time Series: New Perspectives from the Temporal Spectrum of Scale Exponents," in *Computational Intelligence Methods for Bioinformatics and Biostatistics*, vol. 7548 of *Lecture Notes in Computer Science*, pp. 164–175, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
- [30] M. P. Tulppo, R. L. Hughson, T. H. Mäkikallio, K. E. Airaksinen, T. Seppänen, and H. V. Huikuri, "Effects of exercise and passive head-up tilt on fractal and complexity properties of heart rate dynamics," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 280, no. 3, pp. H1081–H1087, 2001.
- [31] M. Malik, "Heart rate variability," *Annals of Noninvasive Electrocardiology*, vol. 1, no. 2, pp. 151–181, 1996.

Research Article

Identification of Alcoholism Based on Wavelet Renyi Entropy and Three-Segment Encoded Jaya Algorithm

Shui-Hua Wang,^{1,2,3} Khan Muhammad ,⁴ Yiding Lv,⁵ Yuxiu Sui ,⁵
Liangxiu Han ,⁶ and Yu-Dong Zhang ^{1,2,3}

¹School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, Henan 454000, China

²Department of Informatics, University of Leicester, Leicester LE1 7RH, UK

³School of Computer Science and Technology, Nanjing Normal University, Nanjing, Jiangsu 210023, China

⁴Digital Contents Research Institute, Sejong University, Seoul, Republic of Korea

⁵Department of Psychiatry, Nanjing Brain Hospital, Nanjing Medical University, Nanjing, Jiangsu 210029, China

⁶School of Computing, Mathematics and Digital Technology (SCMDT), Manchester Metropolitan University, Manchester M156BH, UK

Correspondence should be addressed to Khan Muhammad; khanmuhammad@sju.ac.kr, Yuxiu Sui; suiyuxiu@aliyun.com, Liangxiu Han; l.han@mmu.ac.uk, and Yu-Dong Zhang; yudongzhang@ieee.org

Received 11 September 2017; Revised 22 November 2017; Accepted 12 December 2017; Published 17 January 2018

Academic Editor: Javier Escudero

Copyright © 2018 Shui-Hua Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The alcohol use disorder (AUD) is an important brain disease, which could cause the damage and alteration of brain structure. The current diagnosis of AUD is mainly done manually by radiologists. This study proposes a novel computer-vision-based method for automatic detection of AUD based on wavelet Renyi entropy and three-segment encoded Jaya algorithm from MRI scans. The wavelet Renyi entropy is proposed to provide multiresolution and multiscale analysis of features, describe the complexity of the brain structure, and extract the distinctive features. Grid search method was used to select the optimal wavelet decomposition level and Renyi order. The classifier was constructed based on feedforward neural network and a three-segment encoded (TSE) Jaya algorithm providing parameter-free training of the weights, biases, and number of hidden neurons. We have conducted the experimental evaluation on 235 subjects (114 are AUDs and 121 healthy). *k*-fold cross validation has been used to avoid overfitting and report out-of-sample errors. The results showed that the proposed method outperforms four state-of-the-art approaches in terms of accuracy. The proposed TSE-Jaya provides a better performance, compared to the conventional approaches including plain Jaya, multiobjective genetic algorithm, particle swarm optimization, bee colony optimization, modified ant colony system, and real-coded biogeography-based optimization.

1. Introduction

Alcohol use disorder (AUD) affected 208 million people worldwide in 2010. It can cause severe adverse effects to the brain, liver, heart, and pancreas. The long-term misuse can lead to increased tolerance to alcohol, making it difficult to control the consumption. The short-term misuse can lead to “blood alcohol concentration (BAC).” A BAC from 0.35% to 0.80% can cause fatal respiratory depression and life-threatening alcohol poisoning.

This paper studies the effect of long-term alcohol misuse on the brain. The alcohol misuse can have a damaging effect on the brain neurons; hence, patients with long-term

AUD have smaller volumes of white matter and gray matter than age-matched controls. Besides, alcohol causes adverse effect on the prefrontal cortex and cerebellum. The current diagnosis of AUD mainly relies on manual observation based on brain images. However, due to mild symptoms, the radiologists may miss the slight shrinkage of AUD brains and be unable to identify it at an early stage. It is necessary to create an efficient approach that can monitor the patient brain via magnetic resonance imaging (MRI) and provide automatic, early diagnosis.

Over the last decades, computer-vision-based techniques have been proposed for automatically detecting changes on brain structure for brain related disease diagnosis based on

TABLE 1: Demographic characteristics.

	Alcoholic		Nonalcoholic	
	Women (56)	Men (58)	Women (62)	Men (59)
Age (y)	59.0 + 8.0	56.5 + 8.9	56.9 + 8.4	55.3 + 7.9
Education (y)	9.9 + 1.9	9.3 + 1.6	9.6 + 2.4	9.3 + 2.5
DHD (y)	13.2 + 3.5	19.6 + 5.4	0.0 ± 0.0	0.0 ± 0.0
DDE (grams/d)	197.6 + 62.1	300.7 + 92.1	6.8 + 4.7	4.9 + 3.8
LOS (y)	9.8 + 5.1	6.7 + 3.4	-	-
AUDIT	25.1 + 4.2	25.4 + 4.9	1.5 + 2.0	1.6 + 2.2

(DDE, daily drinks of ethanol; DHD, duration of heavy drinking; AUDIT, alcohol use disorders identification test; LOS, length of sobriety).

MRI scans. Nayak et al. (2016) [1] presented a brain image classification algorithm based on random forest. Alweshah and Abdullah (2015) [2] hybridized firefly algorithm (FA) and probabilistic neural network (PNN). They used the proposed method for detecting changes in the brain. Lv and Hou [3] proposed an improved particle swarm optimization (IPSO) to detect alcoholism in MRI scanning. Monnig (2012) [4] suggested detecting white matter atrophy in neuroimaging of AUD. Yang (2017) [5] combined Hu moment invariant (HMI) and support vector machine (SVM) for pathological brain detection. Jiang and Zhu (2017) [6] explored the method using pseudo Zernike moment (PZM). Lv and Sui (2017) [7] used data augmentation technique for alcoholism detection.

Although several of the above methods were developed for pathological brain detection, they can be easily transferred and applied to alcoholism detection. Nevertheless, these methods suffer from several common problems: first, approaches that do not take into account expressions of complexity of the brain structure do not exhibit good performance in AUD. Second, the training algorithms of existing classifiers may fall into local optimal and it is difficult to optimize the hyperparameters (e.g., the number of hidden neurons in a feedforward neural network) of the classifiers [8].

To address the above problems, we propose in this study a novel identification method of alcoholic use disorder. Our contributions include the following: (1) a novel feature extraction method—wavelet Renyi entropy, which can describe the complexity of brain structure at multiple scales—and (2) an improved Jaya algorithm to train a feedforward neural network, which can optimize the weights, biases, and the number of hidden neurons simultaneously. Our training algorithm does not need to set algorithm-specific parameters.

The rest of this paper is organized as follows: Section 2 describes the subjects, scan protocol, and slice selection method. Section 3 presents the proposed feature extraction method—wavelet Renyi entropy. Section 4 describes the classifier construction and the proposed training algorithm: three-segment encoded Jaya algorithm. Section 5 provides the implementation procedure and the evaluation method. Besides, we show how to use grid search to optimize the parameters of wavelet Renyi entropy. The results and discussions are presented in Section 6. Section 7 concludes the work.

2. Materials

2.1. Subjects. The subjects went through a medical history interview to guarantee they met the inclusion criteria. Those qualified applicants received the computerized diagnostic interview schedule version IV, which ascertains the presence or absence of major psychiatric disorders. Applicants were excluded if mandarin was not their first language, if they were left-handed, or if they had HIV, epilepsy, and stroke; Wernicke–Korsakoff syndrome; bipolar disorder; cirrhosis or liver failure or seizures unrelated to alcoholism, head injury with loss of consciousness more than 15 minutes unrelated to alcoholism, depression, schizophrenia, and other psychotic disorders.

Finally, we enrolled 114 abstinent long-term chronic alcoholic participants (58 men and 56 women) and 121 nonalcoholic control participants (59 men and 62 women). Participants were enrolled through flyers posted in Jiangsu Province Hospital, Nanjing Children’s Hospital, and Nanjing Brain Hospital, as well as the Internet-based advertisements. The data collection lasted for a total of three years. The research was approved by the Institutional Review Board of the participating hospitals. Informed consent was obtained from each participant.

The 235 participants were tested by the “Alcohol Use Disorder Identification Test (AUDIT)” [9]. The unit “ounce” was transformed to “gram,” since the former is not widely identified in China. Their demographic characteristics are shown in Table 1. In this study, we only focus on the structural imaging data.

2.2. Scan Protocol. All 235 subjects lied down as still as possible, with their eyes closed and remaining conscious. Scanning was implemented by a Siemens Verio Tim 3.0T MR scanner (Siemens Medical Solutions, Erlangen, Germany). In total, 216 sagittal slices covering the whole brain were acquired, using an MP-RAGE sequence. The imaging parameters were listed as follows: slice thickness = 0.8 mm, TE = 2.50 ms, TR = 2000 ms, TI = 900 ms, FA = 9°, matrix = 256 × 256, and FOV = 256 mm × 256 mm. The acquired image was 16-bit gray level depth, and we reduced it to 8-bit gray level depth, since the alcoholism alters the structure of healthy brain and it does not change the gray level of brain images. Besides, 8-bit gray level provides enough information, so it is unnecessary to use 16-bit gray level images.

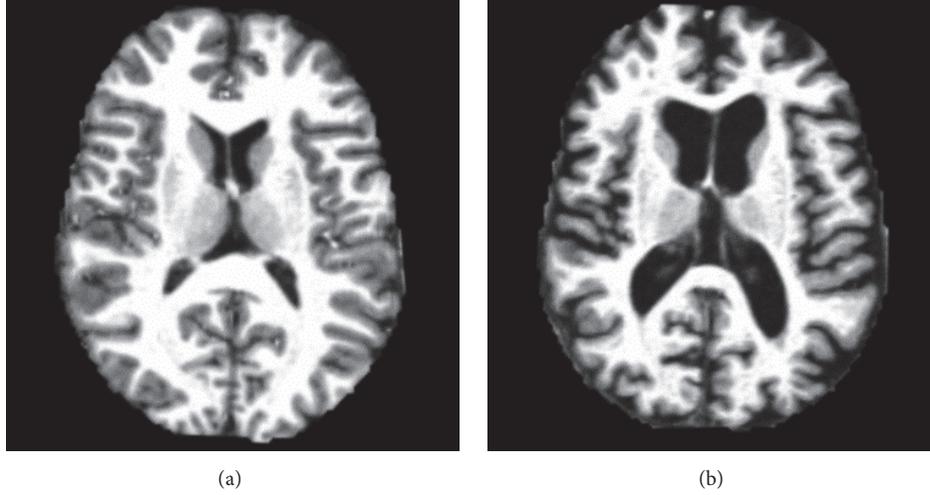


FIGURE 1: Slice examples between (a) a nonalcoholic brain and (b) an alcoholic brain.

2.3. Slice Selection. We used FMRIB software library (FSL) v5.0 software [16, 17] to extract brain and remove skulls for each scanned 3D image. All the volumetric images were normalized to a standard MNI template. Afterwards, we resampled each image to 2 mm isotropic voxel. The slice at $Z = 80$ (8 mm) at MNI_152 coordinate, which is an average of 152 T1-weighted MRI scans linearly transformed to Talairach space, was chosen for each patient. The reason for selecting the 80th slice is that it contains the two distinguishing features of alcoholic patients: (i) the enlarged ventricle and (ii) the shrunken gray matter, for example, the precentral gyrus [18], inferior frontal gyrus [19], and middle temporal gyrus [20]. Figure 1 shows the clear difference between the alcoholic and healthy samples.

Afterwards, the background was cropped, leaving a rectangle matrix with size of 176×176 for the subsequent classifier training. The datasets used in this study are available upon request.

3. The Proposed Feature Extraction Method

To extract distinctive features, this study proposed a new wavelet Renyi entropy (WRE), which combines discrete wavelet transform and Renyi entropy in order to describe the complexity of the brain structure. The wavelet decomposition provides multiresolution and multiscale analysis, while the Renyi entropy provides the complexity description of the wavelet subbands of brain structure.

3.1. Wavelet Decomposition. For a specific signal/image, the discrete wavelet transform (DWT) transforms the signal/image to the wavelet domain. It performs the transformation at multiple levels, by delivering the previous approximation subband to the quadrature mirror filters (abbreviated as QMF) [21]. Compared to traditional Fourier transform, DWT has the key advantage of temporal/spatial resolution.

Let $q(t)$ be a given one-dimensional signal, and the continuous wavelet transform of $q(t)$ is depicted as

$$O_{\psi}(f_s, f_t) = \int_{-\infty}^{\infty} q(t) \times \psi(t | f_s, f_t) dt, \quad (1)$$

where O represents the coefficients and $\psi(t)$ the mother wavelet. $\psi(t | f_s, f_t)$ is defined as

$$\psi(t | f_s, f_t) = \frac{1}{\sqrt{f_s}} \psi\left(\frac{t - f_t}{f_s}\right). \quad (2)$$

Here, f_s represents the scale factor and f_t the translation factor (both f_s and $f_t > 0$). Formula (1) can be discretized by replacing f_s and f_t to discrete variables a and b .

$$\begin{aligned} f_s &= 2^a \\ f_t &= b \times 2^a, \end{aligned} \quad (3)$$

where the parameters a and b represent the values of scale and translation factors, respectively. By this means, we can produce the DWT as

$$\begin{aligned} L(n | a, b) &= \downarrow \left(\sum_n q(n) \times l_a^*(n - 2^a b) \right) \\ G(n | a, b) &= \downarrow \left(\sum_n q(n) \times g_a^*(n - 2^a b) \right). \end{aligned} \quad (4)$$

Here, n means the discrete version of variable t . \downarrow means the downsampling. The functions $l(n)$ and $g(n)$ represent the low-pass filter and high-pass filter, respectively. L and G represent the approximation subband and the detail subband, respectively.

For a two-dimensional DWT (abbreviated as 2D-DWT) [22], suppose the image is symbolized as $q(m, n)$, and there are four subbands in all after each decomposition (A , H , V , and D), shown in Figure 2. The subband A is the approximation component of original image. Subbands H , V , and D represent horizontal, vertical, and diagonal position, respectively. A will be decomposed into four new subbands at a higher level, to produce corresponding higher-level subbands.

```

Step 1 Input a given brain image  $X$ .
Step 2 Perform a  $k$ -level wavelet decomposition, and obtained  $(3k + 1)$  subbands.
Step 3 For  $i = 1: 3k + 1$ 
    Get the 256-bin histogram of the  $i$ th subband wavelet coefficients;
    Obtain Renyi entropy with order of  $\alpha$  over the histogram.
    End
Step 4 Output the catenation of all Renyi entropy values of all subbands.

```

PSEUDOCODE 1: Pseudocode of wavelet Renyi entropy.

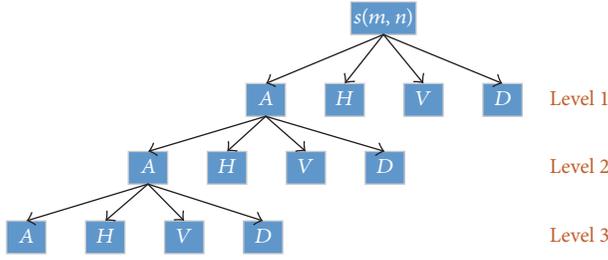


FIGURE 2: Pipeline of 2D-DWT.

3.2. Renyi Entropy. Each subband of wavelet decomposition can be regarded as a discrete variable X . Suppose X has possible outcomes as

$$X = \{x_1, x_2, \dots, x_n\}. \quad (5)$$

Suppose the corresponding probability is defined as

$$p_i = \Pr(X = x_i). \quad (6)$$

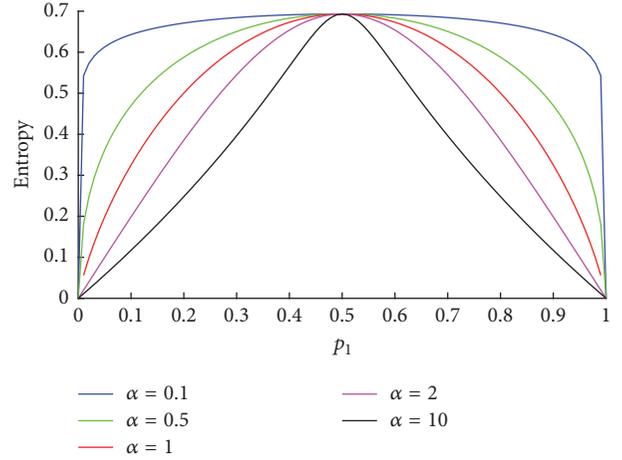
The α -order Renyi entropy is defined as [23]

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left(\sum_{i=1}^n p_i^\alpha \right). \quad (7)$$

The Renyi entropy is Schur-concave and it is a nonincreasing function in α . In some special cases, the Renyi entropy will turn to other types of entropies. For instance, $H_0(X)$ is called Hartley entropy, $H_1(X)$ is Shannon entropy, and $H_\infty(X)$ is the min-entropy [24].

Suppose we have a binary random variable X with $P = [p_1, p_2]$, where $p_2 = 1 - p_1$. The Renyi entropies with different α -values against p_1 are plotted in Figure 3. The concaveness and the nonincreasing against α properties are obvious from this picture. Zero values are included, since this does not affect the calculation of Renyi entropy.

3.3. Wavelet Renyi Entropy. In the past, scholars have proposed the so-called ‘‘wavelet Renyi entropy.’’ Nevertheless, our proposed WREs are different from traditional WREs. First, traditional WREs are mainly for one-dimensional signal, while ours are for two-dimensional image. Second, traditional WREs calculate entropies over the approximation subbands, while our method calculates entropies over both approximation and detail subbands of wavelet coefficients.

FIGURE 3: Renyi entropy with different α -values for a binary random variable.

The pseudocode of WRE is depicted in Pseudocode 1. The bin number of wavelet coefficient histogram is set to 256 in this study.

For a given image, our proposed WRE produced a $(3k + 1)$ -element feature vector. Here, we choose the optimal values of α and k by grid searching approach. The detailed implementation is explained in Section 5.3.

4. The Classifier Construction Based on a Feedforward Neural Network (FNN) and Three-Segment Jaya

To train the classifier, we have proposed using a feedforward neural network (FNN) and three-segment Jaya algorithm. Scholars have used various classifiers in medical brain image analysis, such as decision tree, support vector machine [25], and naive Bayesian classifier. Nevertheless, the feedforward neural network (FNN) won remarkable success, because of the universal approximation theorem [26], which says the following.

Suppose η is a bounded and nonconstant continuous function. Given any function f and any small number $\varepsilon > 0$, there exist an integer N , real vectors ω_i , and real constants c_i and b_i , such that we have

$$F(x) = \sum_{i=1}^N c_i \eta(\omega_i^T x + b_i), \quad (8)$$

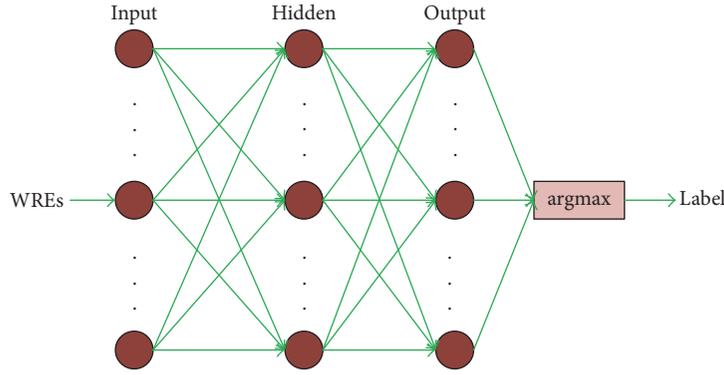


FIGURE 4: Diagram of an FNN.

where $F(x)$ can be used as an approximation realization of function f , which satisfies

$$|F(x) - f(x)| < \varepsilon. \quad (9)$$

However, the traditional FNN training algorithm is a backpropagation (BP) gradient descent algorithm. The BP and its variants often converge to local optimal points. Three-segment encoded Jaya is introduced to address the problem. In the following sections, we will detail each of the methods.

4.1. Structure of FNN. Structurally, the FNN include three layers: (i) an input layer accepted the features; (ii) a hidden layer contains hidden neurons; (iii) an output layer outputs the scores of each class. Finally, the “argmax” function predicts the class associated with the largest score. Figure 4 presents the diagram of FNN. The number of input neurons is the same as the number of features extracted from brain images, the number of output neurons is the same as the number of classes, and the number of hidden neurons is commonly obtained by hyperparameter optimization.

4.2. Jaya Algorithm. As mentioned earlier, the traditional FNN training has an issue with global optimization; to address this problem, a massive number of global optimization algorithms were proposed and employed to train FNN, particularly in the field of brain image classification. For example, Hajimani et al. (2017) [11] designed a multiobjective genetic algorithm (MOGA) to detect cerebral vascular accidents. Chen et al. (2017) [12] used particle swarm optimization (PSO) to classify MRI brain tissues. Subramaniam and Radhakrishnan (2016) [13] used bee colony optimization (BCO) to classify brain cancer image. Raghtate and Salankar (2015) [14] proposed a modified ant colony system (MACS) to realize automatic brain MRI classification. Chen and Du (2017) [15] proposed a real-coded biogeography-based optimization (RCBBO) method for pathological brain detection.

Those algorithms make the classifier more robust than BP; nevertheless, their own parameters need to be fine-tuned, which causes the hyperparameter optimization problem. To overcome the limitation of existing optimization approaches, Jaya as a powerful global optimization approach has been introduced by Rao (2016) [27] as a benchmark function for constrained and unconstrained problems. It is

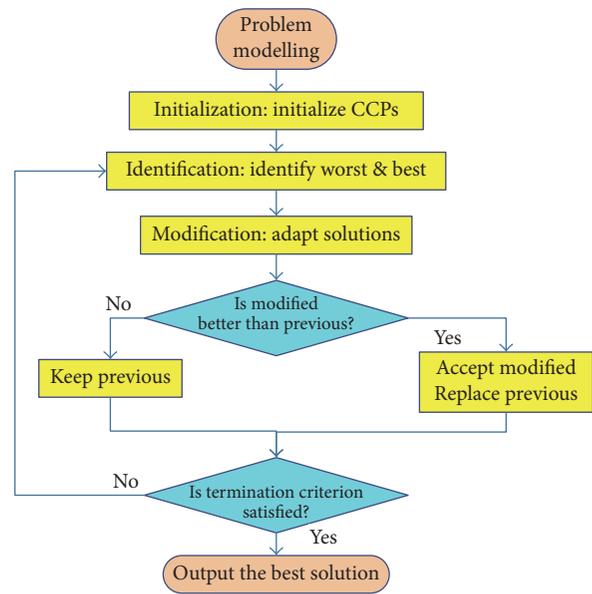


FIGURE 5: Diagram of Jaya algorithm.

an algorithm-specific parameter-free approach which has been proven to be superior to state-of-the-art optimization algorithms and has been successfully applied in thermal performance optimization [28], photovoltaic model identification [29], cooling tower design [30], sensing period adaptation [31], heat change optimization [32], and so forth.

Figure 5 shows the diagram of Jaya algorithm. Assume x , y , and z are the index of iteration, variable, and candidate. Assume $V(x, y, z)$ means the y th variable of z th solution candidate at x th step. Assume $R(x, y, 1)$ and $R(x, y, 2)$ are two positive numbers in the range of $[0, 1]$ and generated at random. The modified candidate $Q(x, y, z)$ is defined as

$$\begin{aligned} Q(x, y, z) = & V(x, y, z) + R(x, y, 1) \\ & \times [V(x, y, b) - |V(x, y, z)|] \\ & - R(x, y, 2) \\ & \times [V(x, y, a) - |V(x, y, z)|], \end{aligned} \quad (10)$$

where a and b denote the index of worst and best candidate within the population:

$$\begin{aligned} b &= \arg \min_z (V(x, y, z)) \\ a &= \arg \max_z (V(x, y, z)). \end{aligned} \quad (11)$$

Hence, $V(x, y, a)$ and $V(x, y, b)$ denote the worst and best value of y th variable at x th iteration.

The 2nd term " $R(x, y, 1) \times [V(x, y, b) - |V(x, y, z)|]$ " in (10) represents that the candidate needs to move closer to the best one. The 3rd term " $-R(x, y, 2) \times [V(x, y, a) - |V(x, y, z)|]$ " in (10) represents that the candidate needs to move away from the worst candidate, noting the "-" symbol before $R(x, y, 2)$ [33]. The updated candidate at iteration $(x + 1)$ is written as

$$\begin{aligned} V(x+1, y, z) &= \begin{cases} Q(x, y, z) \\ V(x, y, z), \end{cases} \\ \text{if } \begin{cases} h[N(x, y, z)] < h[V(x, y, z)] \\ h[N(x, y, z)] \geq h[V(x, y, z)], \end{cases} \end{aligned} \quad (12)$$

where h represents the fitness function.

Equation (12) indicates that $V(x + 1, y, z)$ is assigned with $Q(x, y, z)$ if the modified candidate $Q(x, y, z)$ is better in terms of fitness than $V(x, y, z)$; otherwise it is assigned with $V(x, y, z)$. The algorithm iterates until the termination criterion is satisfied. We set the termination criterion as follows: either the algorithm reaches maximum iteration epoch or the error does not reduce for five epochs.

4.3. Three-Segment Encoded Jaya. The existing Jaya algorithm is mainly used to train weights and biases of FNN as described in Phillips (2017) [10]. However, we believe the number of hidden neurons is also an important hyperparameter that influences the classification performance of FNN. Hence, we proposed a three-segment encoded Jaya algorithm (TSE-Jaya), which aims to optimize the weights, biases, and number of hidden neurons simultaneously. The candidate $V(x, y, z)$ now contains three parts as

$$V(x, y, z) = [V_1(x, y, z) \ V_2(x, y, z) \ V_3(x, y, z)], \quad (13)$$

where V_1 , V_2 , and V_3 represent extracting the first part, second part, and third part of the solution candidate representation. The first part encodes the weights, the second part encodes the biases, and the third part encodes the number of hidden neurons (NHN). The modified candidate is defined consequently as

$$Q(x, y, z) = [Q_1(x, y, z) \ Q_2(x, y, z) \ Q_3(x, y, z)]. \quad (14)$$

The modification rule does not obey (10), and the new modification rule is three-fold as follows:

$$\begin{aligned} Q_1(x, y, z) &= V_1(x, y, z) + R(x, y, 1) \\ &\quad \times [V_1(x, y, b) - |V_1(x, y, z)|] \\ &\quad - R(x, y, 2) \\ &\quad \times [V_1(x, y, a) - |V_1(x, y, z)|] \end{aligned}$$

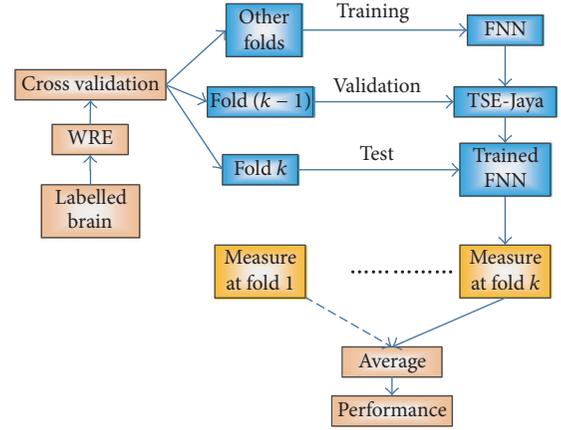


FIGURE 6: The flowchart of the proposed method.

$$\begin{aligned} & - R(x, y, 2) \\ & \times [V_1(x, y, a) - |V_1(x, y, z)|] \\ Q_2(x, y, z) &= V_2(x, y, z) + S(x, y, 1) \\ & \times [V_2(x, y, b) - |V_2(x, y, z)|] \\ & - S(x, y, 2) \\ & \times [V_2(x, y, a) - |V_2(x, y, z)|] \\ Q_3(x, y, z) &= V_3(x, y, z) + T(x, y, 1) \\ & \times [V_3(x, y, b) - |V_3(x, y, z)|] \\ & - T(x, y, 2) \\ & \times [V_3(x, y, a) - |V_3(x, y, z)|], \end{aligned} \quad (15)$$

where S and T are two random positive numbers, similar to variable R . Other procedures are the same as those in Jaya algorithm.

5. Implementation and Evaluation

5.1. Cross Validation Based Implementation. Figure 6 presents the flowchart of our method. Here, the k -fold cross validation method [34] was used in order to avoid overfitting and report out-of-sample errors. We divide the whole dataset 10-fold. In k th trial, $(k - 1)$ th fold is used as validation, k th fold is used as test, and other folds are used as training. The training iterates until the accuracy over validation (a_v) set increases for five continuous epochs. For a clear understanding, we plotted a toy example in Figure 7. Here, at epoch 6, the validation error reaches the minimum. Then, from the 6th to 11th epoch, we can observe the validation increases although the training error decreases, which indicates an overfitting occurs. Hence, we should select the weights corresponding to the 6th epoch. The goal of k -fold cross validation in this study is to avoid overfitting.

TABLE 2: Grid searching of hyperparameters.

Grid searching	Implementation
Wavelet decomposition level k	Search [1, 2, 3, 4, 5] and obtain the best level k^*
Order α of Renyi entropy	<i>Step 1.</i> Search [0, 1, 2, ..., 6] and obtain the best candidate c <i>Step 2.</i> Search [$c - 0.5, c - 0.4, \dots, c + 0.4, c + 0.5$] and obtain the best order α^*

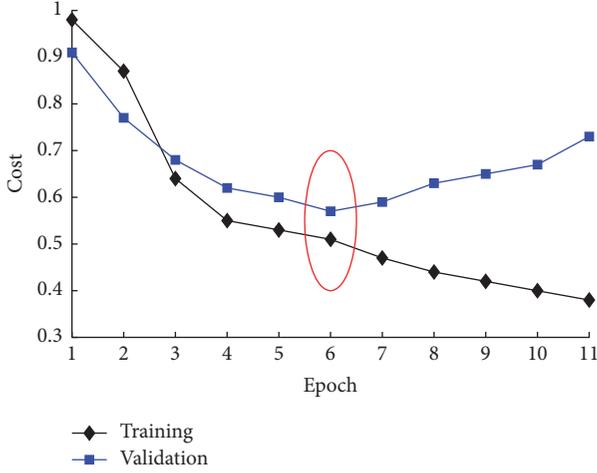


FIGURE 7: A toy example explaining the effect of validation set.

After the training terminates, the measure over the test set is recorded. Finally, all the measures over all test sets of all trials are averaged, and the final classification performance was calculated. The ideal confusion matrix of one time of 10-fold cross validation is

$$I(f = 10, t = 1) = \begin{bmatrix} 114 & 0 \\ 0 & 121 \end{bmatrix}, \quad (16)$$

where I means the ideal confusion matrix, f the number of folds, and t the number of repetitions. In this study, we run the 10-fold cross validation 10 times, and the ideal confusion matrix is

$$I(f = 10, t = 10) = \begin{bmatrix} 1140 & 0 \\ 0 & 1210 \end{bmatrix}. \quad (17)$$

5.2. Evaluation. The evaluation was performed on the realistic confusion matrix of 10×10 -fold cross validation. Suppose the positive class is alcoholism, and the negative class is the control. We can define true positive (TP) as alcoholism correctly identified, true negative (TN) as control correctly identified, false positive (FP) as control mispredicted as alcoholism, and false negative (FN) as alcoholism mispredicted as control. Finally, we define three measures: sensitivity (Sen), specificity (Spc), and accuracy (Acc).

$$\begin{aligned} \text{Sen} &= \frac{\text{TP}}{\text{TP} + \text{FN}} \\ \text{Spc} &= \frac{\text{TN}}{\text{TN} + \text{FP}} \\ \text{Acc} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \end{aligned} \quad (18)$$

5.3. Grid Searching. In the grid searching, the criterion uses the “accuracy (Acc)” measure defined above. The implementation is explained in Table 2. For wavelet decomposition level k , a simple grid searching from 1 to 5 with an increase of 1 was used, since k should be an integer.

$$k^* = \arg \max_{[1,2,\dots,5]} (\text{Acc}). \quad (19)$$

For the order α of Renyi entropy, a coarse-to-fine searching strategy was used. First, a coarse grid was set from 0 to 6 with an increase of 1, and we obtained the coarse candidate c .

$$c = \arg \max_{[0,1,\dots,6]} (\text{Acc}). \quad (20)$$

Then, a fine grid was set from $c - 0.5$ to $c + 0.5$ with an increase of 0.1, and the optimal order α^* is obtained as

$$\alpha^* = \arg \max_{[c-0.5, c-0.4, \dots, c+0.5]} (\text{Acc}). \quad (21)$$

6. Results and Discussions

Our programs were developed in-house. The experiment ran on the platform of Dell laptop with 2.20 GHz Intel Core i7-4702HQ CPU and 16 GB RAM. The operating system was Windows 10. MATLAB 2017a is the programming development environment.

6.1. Statistical Analysis. The cross validation divides the dataset into 10 sets. In each run, the number of sets resulting from the division is different. We set the wavelet decomposition level as 4 and Renyi α -value as 1.2. The maximum iterative epoch is set as 1000, and the population in Jaya algorithm is set to 20. The sensitivity, specificity, and accuracy of our method are listed in Tables 3, 4, and 5, respectively. We can observe our proposed method achieved a sensitivity of $93.60 \pm 1.55\%$, a specificity of $93.72 \pm 1.42\%$, and an accuracy of $93.66 \pm 1.23\%$.

6.2. Comparison to State-of-the-Art Approaches. We compared this proposed method “WMI + FNN + TSE-Jaya” with four state-of-the-art approaches: FA + PNN [2], IPSO method [3], HMI + SVM [5], and PZM [6]. All the algorithms were run over a 10×10 -fold cross validation over our dataset. The results of 10×10 -fold cross validation of four state-of-the-art methods are shown in Table 6. Finally, the comparison result is presented in Table 7.

6.3. Optimal Wavelet Decomposition Level. In this experiment, we fixed the Renyi order α to 1.2 and let the wavelet decomposition level change from 1 to 5 with increase of 1. The

TABLE 3: Sensitivity of the proposed method (unit: %).

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Total
R1	100.00	90.91	100.00	81.82	100.00	83.33	100.00	100.00	100.00	90.91	94.74
R2	100.00	100.00	100.00	91.67	100.00	72.73	72.73	100.00	100.00	100.00	93.86
R3	100.00	100.00	90.91	100.00	100.00	100.00	91.67	100.00	81.82	91.67	95.61
R4	100.00	90.91	81.82	90.91	75.00	100.00	100.00	100.00	100.00	90.91	92.98
R5	100.00	81.82	100.00	100.00	100.00	100.00	63.64	90.91	90.91	91.67	92.11
R6	100.00	100.00	100.00	100.00	100.00	90.91	72.73	66.67	100.00	91.67	92.11
R7	100.00	83.33	90.91	90.91	90.91	81.82	91.67	100.00	100.00	100.00	92.98
R8	91.67	100.00	91.67	100.00	90.91	90.91	100.00	90.91	100.00	90.91	94.74
R9	90.91	100.00	100.00	100.00	100.00	90.91	90.91	91.67	100.00	91.67	95.61
R10	100.00	90.91	91.67	81.82	100.00	90.91	90.91	83.33	90.91	91.67	91.23
Avr											93.60 ± 1.55

(F, fold; R, run).

TABLE 4: Specificity of the proposed method (unit: %).

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Total
R1	100.00	100.00	100.00	100.00	100.00	100.00	83.33	91.67	91.67	91.67	95.87
R2	100.00	91.67	100.00	83.33	100.00	100.00	91.67	84.62	100.00	91.67	94.21
R3	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	75.00	66.67	94.21
R4	91.67	91.67	100.00	92.31	83.33	100.00	100.00	91.67	66.67	100.00	91.74
R5	91.67	83.33	100.00	91.67	91.67	92.31	100.00	83.33	100.00	91.67	92.56
R6	100.00	66.67	84.62	100.00	100.00	100.00	100.00	83.33	100.00	100.00	93.39
R7	91.67	100.00	83.33	100.00	83.33	100.00	100.00	100.00	100.00	91.67	95.04
R8	100.00	91.67	83.33	100.00	75.00	100.00	100.00	100.00	75.00	91.67	91.74
R9	100.00	100.00	100.00	100.00	91.67	100.00	83.33	100.00	100.00	75.00	95.04
R10	83.33	100.00	91.67	83.33	100.00	91.67	100.00	84.62	100.00	100.00	93.39
Avr											93.72 ± 1.42

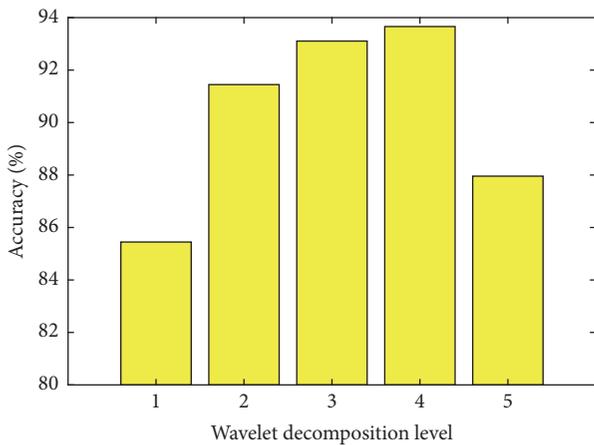


FIGURE 8: Choosing the optimal wavelet decomposition level.

corresponding accuracy varied as shown in Figure 8. Here, the accuracy is 85.45%, 91.45%, 93.11%, 93.66%, and 87.96% when decomposition level is 1, 2, 3, 4, and 5, respectively. Obviously, the 4th-level decomposition yields the greatest accuracy; hence, we chose the optimal decomposition level as 4. The Renyi entropy was then calculated over all the subbands of this 4-level wavelet decomposition.

Note that 3D-DWT is more straightforward than 2D-DWT over a one particular slice. Nevertheless, our aim in this study is to select a distinguishing slice, which is related to brain regions affected by alcoholism, in order to reduce the computation burden. In the future, we shall test the results of 3D-DWT.

6.4. Optimal Renyi Order. In this experiment, we shall illustrate why we set the Renyi order α as 1.2 by a coarse-to-fine grid search. First, we search the coarse grid from 0 to 6 with an increase of 1. The result was shown in Figure 9(a), and the value of 1 was selected as the initial point for fine grid search. Second, the fine grid from 0.5 to 1.5 with an increase of 0.1 was established, and the result was shown in Figure 9(b). We can observe that $\alpha = 1.2$ can yield the greatest accuracy.

6.5. Effectiveness of Three-Segment Encoding. Our proposed TSE-Jaya can train the weights, biases, and number of hidden neurons (NHN) simultaneously. In this experiment, we compare TSE-Jaya to plain Jaya algorithm [10], which can only train the weights and biases of feedforward neural network [10]; hence, we have to fix the number of hidden neurons by experience. Here, we set NHN as 10 for plain Jaya algorithm. The settings of other parameters were the same as previous experiments. The results of 10×10 -fold cross validation of plain Jaya [10] are shown in Table 8, and the

TABLE 5: Accuracy of the proposed method (unit: %).

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Total
R1	100.00	95.65	100.00	91.30	100.00	91.67	91.30	95.83	95.83	91.30	95.32
R2	100.00	95.83	100.00	87.50	100.00	86.96	82.61	91.67	100.00	95.83	94.04
R3	100.00	100.00	95.65	100.00	100.00	100.00	95.83	100.00	78.26	79.17	94.89
R4	95.83	91.30	91.30	91.67	79.17	100.00	100.00	95.65	83.33	95.65	92.34
R5	95.65	82.61	100.00	95.83	95.83	95.83	82.61	86.96	95.65	91.67	92.34
R6	100.00	83.33	92.00	100.00	100.00	95.65	86.96	75.00	100.00	95.83	92.77
R7	95.65	91.67	86.96	95.65	86.96	91.30	96.00	100.00	100.00	95.65	94.04
R8	96.00	95.83	87.50	100.00	82.61	95.65	100.00	95.65	86.96	91.30	93.19
R9	95.65	100.00	100.00	100.00	95.65	95.65	86.96	95.83	100.00	83.33	95.32
R10	91.67	95.65	91.67	82.61	100.00	91.30	95.65	84.00	95.65	95.83	92.34
Avr											93.66 ± 1.23

TABLE 6: Results of 10 × 10-fold cross validation of four state-of-the-art methods.

FA + PNN [2]	Sen	Spc	Acc	IPSO [3]	Sen	Spc	Acc
R1	86.84	86.78	86.81	R1	90.35	87.60	88.94
R2	83.33	90.08	86.81	R2	92.11	85.95	88.94
R3	84.21	87.60	85.96	R3	89.47	90.08	89.79
R4	90.35	85.12	87.66	R4	88.60	90.08	89.36
R5	87.72	84.30	85.96	R5	88.60	88.43	88.51
R6	84.21	86.78	85.53	R6	88.60	90.08	89.36
R7	81.58	91.74	86.81	R7	89.47	90.91	90.21
R8	86.84	87.60	87.23	R8	85.96	93.39	89.79
R9	85.09	86.78	85.96	R9	87.72	90.91	89.36
R10	87.72	85.95	86.81	R10	87.72	86.78	87.23
Avr	85.79 ± 2.58	87.27 ± 2.21	86.55 ± 0.67	Avr	88.86 ± 1.66	89.42 ± 2.23	89.15 ± 0.83
HMI + SVM [5]	Sen	Spc	Acc	PZM [6]	Sen	Spc	Acc
R1	82.46	86.78	84.68	R1	87.72	86.78	87.23
R2	85.96	83.47	84.68	R2	87.72	87.60	87.66
R3	85.09	86.78	85.96	R3	85.09	90.91	88.09
R4	87.72	85.12	86.38	R4	85.96	86.78	86.38
R5	85.09	87.60	86.38	R5	85.96	85.95	85.96
R6	83.33	86.78	85.11	R6	89.47	87.60	88.51
R7	85.96	85.95	85.96	R7	85.09	83.47	84.26
R8	79.82	89.26	84.68	R8	85.96	88.43	87.23
R9	80.70	82.64	81.70	R9	85.96	85.95	85.96
R10	89.47	82.64	85.96	R10	83.33	86.78	85.11
Avr	84.56 ± 3.02	85.70 ± 2.21	85.15 ± 1.40	Avr	86.23 ± 1.71	87.02 ± 1.91	86.64 ± 1.35

comparison between Jaya [10] and our proposed TSE-Jaya is shown in Table 9.

The superiority of proposed TSE-Jaya to plain Jaya [10] is clear. This demonstrates the importance of choosing the optimal number of hidden neurons, that is, that the variable number of hidden neurons gives a better performance than fixed number of hidden neurons, which is also validated by Carleo and Troyer (2017) [35].

6.6. Training Algorithm Comparison. To demonstrate the efficiency of the proposed algorithm, we have compared the TSE-Jaya with a several global optimization algorithms including MOGA [11], PSO [12], BCO [13], MACS [14],

and RCBBO [15]. All the settings of common controlling parameters are the same: the maximum iterative epoch is set as 1000, and the population in all algorithms is set to 20. The algorithm-specific parameters of those four comparison algorithms are assigned by experiences. The results of those four training algorithms over 10 × 10-fold cross validation are shown in Table 10, and the final comparison with our proposed TSE-Jaya was shown in Table 11.

Table 11 shows that the proposed TSE-Jaya performed the best among all six algorithms. The PSO [12] and RCBBO [15] ranked the second and the third, with their accuracies over than 90%. The BCO [13] ranked the fourth, the MACS [14] ranked the fifth, and MOGA [11] performed the worst. The

TABLE 7: Method comparison.

Alcoholism identification method	Sensitivity	Specificity	Accuracy
FA + PNN [2]	85.79 \pm 2.58	87.27 \pm 2.21	86.55 \pm 0.67
IPSO [3]	88.86 \pm 1.66	89.42 \pm 2.23	89.15 \pm 0.83
HMI + SVM [5]	84.56 \pm 3.02	85.70 \pm 2.21	85.15 \pm 1.40
PZM [6]	86.23 \pm 1.71	87.02 \pm 1.91	86.64 \pm 1.35
WMI + FNN + TSE-Jaya (ours)	93.60 \pm 1.55	93.72 \pm 1.42	93.66 \pm 1.23

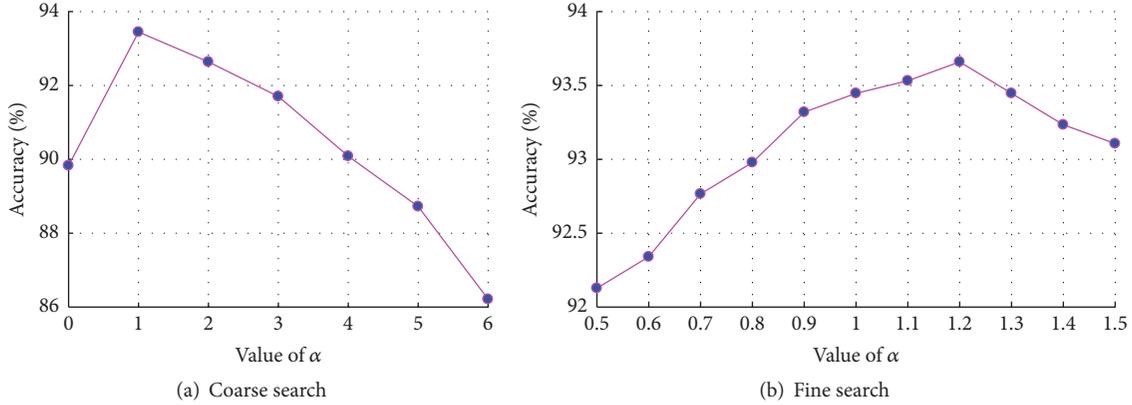


FIGURE 9: Choosing the optimal Renyi order by coarse-to-fine grid search.

TABLE 8: Results of 10×10 -fold cross validation of plain Jaya algorithm [10].

Jaya [10]	Sen (%)	Spcc (%)	Acc (%)
R1	92.11	92.56	92.34
R2	94.74	95.04	94.89
R3	87.72	93.39	90.64
R4	92.11	91.74	91.91
R5	94.74	91.74	93.19
R6	96.49	94.21	95.32
R7	89.47	95.04	92.34
R8	92.11	92.56	92.34
R9	92.11	90.91	91.49
R10	95.61	92.56	94.04

TABLE 9: Plain Jaya algorithm versus the proposed TSE-Jaya.

Method	Sensitivity	Specificity	Accuracy
Jaya [10]	92.72 \pm 2.74	92.98 \pm 1.42	92.85 \pm 1.50
TSE-Jaya (ours)	93.60 \pm 1.55	93.72 \pm 1.42	93.66 \pm 1.23

reasons behind efficiency of the proposed approach can be explained from two aspects: (i) The Jaya does not need to set the algorithm-specific parameters, making it more reliable than other algorithms. (ii) The TSE guarantees the variable number of hidden neurons at each run.

6.7. Validation of the Selected Slice. In this experiment, we validated in terms of the classification performance the

selection of the 80th slice. We set a range of Z increasing from 30 to 150 with an increment of 10 as shown in Figure 10.

Other settings were the same as the previous experiments. Again, 10 repetitions of 10-fold cross validation were utilized. The curve of accuracy is drawn in Figure 11. It is observed that the 80th slice gives the highest accuracy among all candidate slices. The reason is that this slice contains the enlarged ventricle and the shrunk gray matter caused by alcoholism. On the contrary, hippocampus [36] and striatum [37] are also related to alcoholism. Nevertheless, their altered volumes are relatively small and hence do not provide an excellent performance in this task.

In this case, the optimal slice could be in a position that is vertical to x or y -axes, or it can be even an oblique plane to all three axes. Here, we choose a slice vertical to z -axis, which is for the convenience of radiologists, since they usually read the axial slices. In the future, we shall develop techniques to handle multislices, and we may develop surface analysis techniques [38].

7. Conclusions

In this study, we proposed a novel alcoholism identification method from healthy controls based on a computer-vision approach. Our method was based on three components: the proposed wavelet Renyi entropy, feedforward neural network, and the proposed three-segment encoded Jaya algorithm. The experiments showed that our method achieved a sensitivity of 93.60 \pm 1.55%, a specificity of 93.72 \pm 1.42, and an accuracy of 93.66 \pm 1.23 over a 10×10 -fold cross validation. The performance is superior to four state-of-the-art alcoholism algorithms. We validated the optimal wavelet

TABLE 10: Results of 10 × 10-fold cross validation of four training algorithms.

MOGA [11]	Sen	Spc	Acc	PSO [12]	Sen	Spc	Acc
R1	78.07	75.21	76.60	R1	90.35	90.91	90.64
R2	76.32	79.34	77.87	R2	87.72	87.60	87.66
R3	72.81	77.69	75.32	R3	87.72	90.08	88.94
R4	75.44	77.69	76.60	R4	92.98	93.39	93.19
R5	76.32	76.86	76.60	R5	88.60	90.08	89.36
R6	75.44	75.21	75.32	R6	89.47	87.60	88.51
R7	78.95	80.99	80.00	R7	88.60	87.60	88.09
R8	76.32	73.55	74.89	R8	93.86	90.91	92.34
R9	73.68	81.82	77.87	R9	90.35	92.56	91.49
R10	79.82	80.17	80.00	R10	92.98	94.21	93.62
Avr	76.32 ± 2.19	77.85 ± 2.72	77.11 ± 1.83	Avr	90.26 ± 2.28	90.50 ± 2.41	90.38 ± 2.18
BCO [13]	Sen	Spc	Acc	MACS [14]	Sen	Spc	Acc
R1	87.72	85.12	86.38	R1	84.21	85.95	85.11
R2	92.11	90.91	91.49	R2	86.84	89.26	88.09
R3	87.72	88.43	88.09	R3	81.58	85.95	83.83
R4	85.96	90.91	88.51	R4	84.21	79.34	81.70
R5	84.21	86.78	85.53	R5	80.70	83.47	82.13
R6	87.72	87.60	87.66	R6	82.46	85.12	83.83
R7	91.23	90.08	90.64	R7	83.33	84.30	83.83
R8	91.23	85.12	88.09	R8	84.21	86.78	85.53
R9	84.21	85.12	84.68	R9	86.84	86.78	86.81
R10	84.21	91.74	88.09	R10	88.60	86.78	87.66
Avr	87.63 ± 3.05	88.18 ± 2.61	87.91 ± 2.09	Avr	84.30 ± 2.50	85.37 ± 2.64	84.85 ± 2.19
RCBBO [15]	Sen	Spc	Acc				
R1	84.21	90.08	87.23				
R2	88.60	90.91	89.79				
R3	88.60	90.08	89.36				
R4	93.86	93.39	93.62				
R5	91.23	90.91	91.06				
R6	92.11	93.39	92.77				
R7	90.35	88.43	89.36				
R8	91.23	90.08	90.64				
R9	90.35	89.26	89.79				
R10	88.60	87.60	88.09				
Avr	89.91 ± 2.62	90.41 ± 1.88	90.17 ± 1.95				

TABLE 11: FNN training algorithm comparison.

FNN training algorithm	Sensitivity	Specificity	Accuracy	Rank
MOGA [11]	76.32 ± 2.19	77.85 ± 2.72	77.11 ± 1.83	6
PSO [12]	90.26 ± 2.28	90.50 ± 2.41	90.38 ± 2.18	2
BCO [13]	87.63 ± 3.05	88.18 ± 2.61	87.91 ± 2.09	4
MACS [14]	84.30 ± 2.50	85.37 ± 2.64	84.85 ± 2.19	5
RCBBO [15]	89.91 ± 2.62	90.41 ± 1.88	90.17 ± 1.95	3
TSE-Jaya (ours)	93.60 ± 1.55	93.72 ± 1.42	93.66 ± 1.23	1

decomposition to be 4, and the optimal Renyi order was 1.2. Besides, comparing to the existing global optimization approaches, the proposed three-segment encoded Jaya is proven to provide a better performance than other methods such as plain Jaya and another five training algorithms.

Finally, we validated the reason why we chose the 80th slice.

The shortcomings of our method lie in two aspects. First, our method needs to scan the whole brain and select the 80th slice at z -axis. Second, the wavelet Renyi entropy was

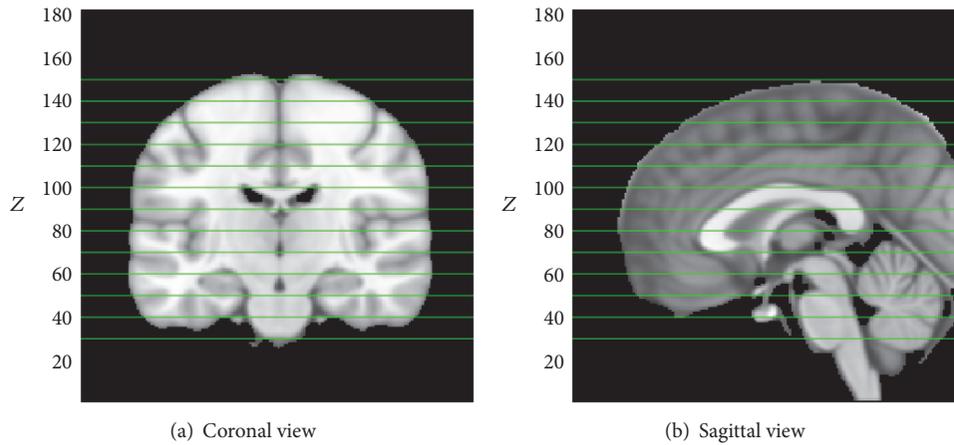


FIGURE 10: Position of Z-slice.

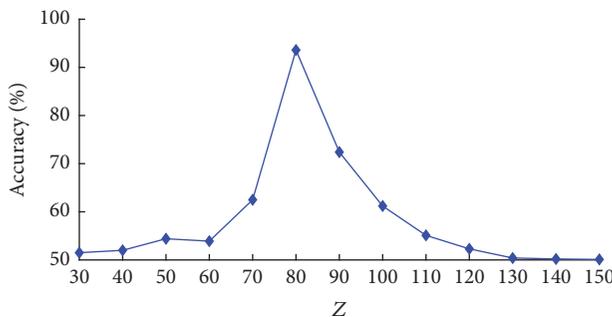


FIGURE 11: Classification performance changes with slice index Z.

extracted, but in the future we shall try to find more efficient features.

In the future work, we will develop classifiers based on multisource data, such as facial image, EEG, and spectrum data. We shall explore the changes of functional connectivity of alcoholism brain. Second, other image features, such as histogram of oriented gradient, will be tested. Third, deep learning methods will be tested.

Conflicts of Interest

The authors have no conflicts of interest to disclose with regard to the subject matter of this paper.

Acknowledgments

This study was supported by National Natural Science Foundation of China (61602250), Natural Science Foundation of Jiangsu Province (BK20150983), Project of Science and Technology of Henan Province (172102210272), Program of Natural Science Research of Jiangsu Higher Education Institutions (16KJB520025), and Open Fund of Key Laboratory of Guangxi High Schools Complex System and Computational Intelligence (2016CSCI01).

References

- [1] D. R. Nayak, R. Dash, and B. Majhi, "Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests," *Neurocomputing*, vol. 177, pp. 188–197, 2016.
- [2] M. Alweshah and S. Abdullah, "Hybridizing firefly algorithms with a probabilistic neural network for solving classification problems," *Applied Soft Computing*, vol. 35, pp. 513–524, 2015.
- [3] Y.-D. Lv and X.-X. Hou, "Alcoholism detection by medical robots based on Hu moment invariants and predator-prey adaptive-inertia chaotic particle swarm optimization," *Computers & Electrical Engineering*, vol. 63, pp. 126–138, 2017.
- [4] M. A. Monnig, "Observed power and projected sample sizes to detect white matter atrophy in neuroimaging of alcohol use disorders," *Alcoholism-Clinical And Experimental Research*, vol. 36, article 272A, 2012.
- [5] J. Yang, "Pathological brain detection in MRI scanning via Hu moment invariants and machine learning," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 29, no. 2, pp. 299–312, 2017.
- [6] Y. Jiang and W. Zhu, "Exploring a smart pathological brain detection method on pseudo Zernike moment," *Multimedia Tools and Applications*, pp. 1–16, 2017.
- [7] Y.-D. Lv and Y. Sui, "Alcoholism detection by data augmentation and convolutional neural network with stochastic pooling," *Journal of Medical Systems*, vol. 42, no. 1, article 2, 2017.
- [8] R. M. Stephenson, G. R. Naik, and R. Chai, "A system for accelerometer-based gesture classification using artificial neural networks," in *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '17)*, pp. 4187–4190, Seogwipo, South Korea, July 2017.
- [9] A. Barik, R. K. Rai, and A. Chowdhury, "Alcohol use-related problems among a rural Indian population of West Bengal: an application of the Alcohol Use Disorders Identification Test (AUDIT)," *Alcohol and Alcoholism*, vol. 51, no. 2, pp. 215–223, 2016.
- [10] P. Phillips, "Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm," *Neurocomputing*, vol. 272, pp. 668–676, 2017.
- [11] E. Hajimani, M. Ruano, and A. Ruano, "An intelligent support system for automatic detection of cerebral vascular accidents

- from brain CT images,” *Computer Methods and Programs in Biomedicine*, vol. 146, pp. 109–123, 2017.
- [12] R. M. Chen, S. C. Yang, and C. M. Wang, “MRI brain tissue classification using unsupervised optimized extenics-based methods,” *Computers & Electrical Engineering*, vol. 58, pp. 489–501, 2017.
- [13] S. Subramaniam and M. Radhakrishnan, “Neural network with bee colony optimization for MRI brain cancer image classification,” *International Arab Journal of Information Technology*, vol. 13, no. 1, pp. 118–124, 2016.
- [14] G. S. Raghtate and S. S. Salankar, “Automatic brain MRI classification using modified ant colony system and neural network classifier,” in *Proceedings of the 7th International Conference on Computational Intelligence and Communication Networks (CICN '15)*, pp. 1241–1246, Jabalpur, India, December 2015.
- [15] P. Chen and S. Du, “Pathological brain detection via wavelet packet tsallis entropy and real-coded biogeography-based optimization,” *Fundamenta Informaticae*, vol. 151, no. 1-4, pp. 275–291, 2017.
- [16] M. W. Woolrich, S. Jbabdi, B. Patenaude et al., “Bayesian analysis of neuroimaging data in FSL,” *NeuroImage*, vol. 45, no. 1, pp. S173–S186, 2009.
- [17] S. M. Smith, M. Jenkinson, M. W. Woolrich et al., “Advances in functional and structural MR image analysis and implementation as FSL,” *NeuroImage*, vol. 23, supplement 1, pp. S208–S219, 2004.
- [18] T. Hatchard, O. Mioduszewski, C. Fall, A. Byron-Alhassan, P. Fried, and A. M. Smith, “Neural impact of low-level alcohol use on response inhibition: an fMRI investigation in young adults,” *Behavioural Brain Research*, vol. 329, pp. 12–19, 2017.
- [19] T. Schmidt, P. Roser, O. Ze, G. Juckel, B. Suchan, and P. Thoma, “Cortical thickness and trait empathy in patients and people at high risk for alcohol use disorders,” *Psychopharmacology*, vol. 234, no. 23-24, pp. 3521–3533, 2017.
- [20] P. Bach, E. Zois, S. Vollstädt-Klein et al., “Association of the alcohol dehydrogenase gene polymorphism rs1789891 with gray matter brain volume, alcohol consumption, alcohol craving and relapse risk,” *Addiction Biology*, 2017.
- [21] T. Koza and N. Karaboga, “Quadrature mirror filter bank design for mitral valve doppler signal using artificial bee colony algorithm,” *Elektronika ir Elektrotechnika*, vol. 23, no. 1, pp. 57–62, 2017.
- [22] G. Savić, M. Prokin, V. Rajović, and D. Prokin, “High-performance 1-D and 2-D inverse DWT 5/3 filter architectures for efficient hardware implementation,” *Circuits, Systems and Signal Processing*, vol. 36, no. 9, pp. 3674–3701, 2017.
- [23] N. Masi, “Rényi entropy for particle systems as an instrument to enlarge the Boltzmannian concept of entropy: some holographic perspectives,” *International Journal of Modern Physics D: Gravitation, Astrophysics, Cosmology*, vol. 26, no. 10, Article ID 1750105, 11 pages, 2017.
- [24] C. Y. Kee, S. G. Ponnambalam, and C. K. Loo, “Binary and multi-class motor imagery using Renyi entropy for feature extraction,” *Neural Computing & Applications*, vol. 28, no. 8, pp. 2051–2062, 2017.
- [25] J. M. Gorriz and J. Ramírez, “Wavelet entropy and directed acyclic graph support vector machine for detection of patients with unilateral hearing loss,” *Frontiers in Computational Neuroscience*, vol. 10, article 160, 2016.
- [26] D. A. Winkler and T. C. Le, “Performance of deep and shallow neural networks, the universal approximation theorem activity cliffs, and QSAR,” *Molecular Informatics*, vol. 36, no. 11, Article ID 1781141, 2017.
- [27] R. V. Rao, “Jaya: a simple and new optimization algorithm for solving constrained and unconstrained optimization problems,” *International Journal of Industrial Engineering Computations*, vol. 7, no. 1, pp. 19–34, 2016.
- [28] P. Ochoń, P. Cisek, M. Rerak et al., “Thermal performance optimization of the underground power cable system by using a modified Jaya algorithm,” *International Journal of Thermal Sciences*, vol. 123, pp. 162–180, 2018.
- [29] K. Yu, J. Liang, B. Qu, X. Chen, and H. Wang, “Parameters identification of photovoltaic models using an improved JAYA optimization algorithm,” *Energy Conversion and Management*, vol. 150, pp. 742–753, 2017.
- [30] R. Rao and K. More, “Optimal design and analysis of mechanical draft cooling tower using improved Jaya algorithm,” *International Journal of Refrigeration*, vol. 82, pp. 312–324, 2017.
- [31] A. Kaur, S. Sharma, and A. Mishra, “Sensing period adaptation for multiobjective optimisation in cognitive radio using Jaya algorithm,” *IEEE Electronics Letters*, vol. 53, no. 19, pp. 1335–1336, 2017.
- [32] R. V. Rao and A. Saroj, “Constrained economic optimization of shell-and-tube heat exchangers using elitist-Jaya algorithm,” *Energy*, vol. 128, pp. 785–800, 2017.
- [33] K. Abhishek, V. R. Kumar, S. Datta, and S. S. Mahapatra, “Application of JAYA algorithm for the optimization of machining performance characteristics during the turning of CFRP (epoxy) composites: comparison with TLBO, GA, and ICA,” *Engineering with Computers*, vol. 33, no. 3, pp. 457–475, 2017.
- [34] H. M. Lu, “Facial emotion recognition based on biorthogonal wavelet entropy, fuzzy support vector machine, and stratified cross validation,” *IEEE Access*, vol. 4, pp. 8375–8385, 2016.
- [35] G. Carleo and M. Troyer, “Solving the quantum many-body problem with artificial neural networks,” *Science*, vol. 355, no. 6325, pp. 602–606, 2017.
- [36] T. Murano, H. Koshimizu, H. Hagihara, and T. Miyakawa, “Transcriptomic immaturity of the hippocampus and prefrontal cortex in patients with alcoholism,” *Scientific Reports*, vol. 7, Article ID 44531, 2017.
- [37] M. Spoelder, P. Hesselting, M. Styles et al., “Dopaminergic neurotransmission in ventral and dorsal striatum differentially modulates alcohol reinforcement,” *European Journal of Neuroscience*, vol. 45, no. 1, pp. 147–158, 2017.
- [38] B. M. Craver, M. M. Acharya, B. D. Allen et al., “3D surface analysis of hippocampal microvasculature in the irradiated brain,” *Environmental and Molecular Mutagenesis*, vol. 57, no. 5, pp. 341–349, 2016.

Research Article

Efficient Computation of Multiscale Entropy over Short Biomedical Time Series Based on Linear State-Space Models

Luca Faes,^{1,2} Alberto Porta,^{3,4} Michal Javorka,^{5,6} and Giandomenico Nollo^{1,7}

¹BIOTech, Department of Industrial Engineering, University of Trento, Trento, Italy

²Dipartimento di Energia, Ingegneria dell'Informazione e Modelli Matematici (DEIM), University of Palermo, Palermo, Italy

³Department of Biomedical Sciences for Health, University of Milan, Milan, Italy

⁴Department of Cardiothoracic-Vascular Anesthesia and Intensive Care, IRCCS Policlinico San Donato, San Donato Milanese, Milan, Italy

⁵Department of Physiology, Jessenius Faculty of Medicine, Comenius University in Bratislava, Mala Hora 4C, 03601 Martin, Slovakia

⁶Biomedical Center Martin, Jessenius Faculty of Medicine, Comenius University in Bratislava, Mala Hora 4C, 03601 Martin, Slovakia

⁷Bruno Kessler Foundation, Trento, Italy

Correspondence should be addressed to Luca Faes; faes.luca@gmail.com

Received 18 September 2017; Accepted 13 November 2017; Published 7 December 2017

Academic Editor: Anne Humeau-Heurtier

Copyright © 2017 Luca Faes et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The most common approach to assess the dynamical complexity of a time series across multiple temporal scales makes use of the multiscale entropy (MSE) and refined MSE (RMSE) measures. In spite of their popularity, MSE and RMSE lack an analytical framework allowing their calculation for known dynamic processes and cannot be reliably computed over short time series. To overcome these limitations, we propose a method to assess RMSE for autoregressive (AR) stochastic processes. The method makes use of linear state-space (SS) models to provide the multiscale parametric representation of an AR process observed at different time scales and exploits the SS parameters to quantify analytically the complexity of the process. The resulting linear MSE (LMSE) measure is first tested in simulations, both theoretically to relate the multiscale complexity of AR processes to their dynamical properties and over short process realizations to assess its computational reliability in comparison with RMSE. Then, it is applied to the time series of heart period, arterial pressure, and respiration measured for healthy subjects monitored in resting conditions and during physiological stress. This application to short-term cardiovascular variability documents that LMSE can describe better than RMSE the activity of physiological mechanisms producing biological oscillations at different temporal scales.

1. Introduction

An intrinsic feature of almost all physiological systems, which is clearly visible in the time course of the variables measured from these systems, is their dynamical complexity. It is indeed widely acknowledged that physiological systems such as the brain, the cardiovascular system, and the muscular system produce output signals that exhibit highly complex dynamics, resulting from the combined activity of several mechanisms of physiological regulation which are coupled with each other through structural and functional pathways [1–4]. Since these multiple and simultaneously active mechanisms typically operate across multiple temporal scales, a surge of interest has emerged in the last two decades in computational methods

able to connect the complexity of a dynamic oscillation to its temporal scale. The research in this context has been driven by the work of Costa et al. [5], who introduced multiscale entropy (MSE) as a measure of the complexity of a time series evaluated as a function of the time scale at which the series is observed. Since its introduction, MSE has been successfully employed in several fields of science [6], becoming a prevailing method to quantify the complexity of biomedical time series [7, 8] and gaining particular popularity in the analysis of brain [9, 10] and cardiovascular variability [11, 12] signals.

In spite of its acknowledged usefulness, MSE has been shown to present some shortcomings which have led many authors to propose improvements and modifications of the original algorithm [6]. The main problems identified in the

original MSE formulation are related to both its defining algorithmic steps, the rescaling procedure that changes the temporal scale of the observed series progressively filtering out the shorter scales and keeping the longer ones, and the computation of the entropy rate of the rescaled time series performed by means of the sample entropy (SampEn) metric [13]. Specifically, it has been reported [11, 14] that the rescaling procedure makes use of a suboptimal low-pass filter (i.e., an averaging filter) which cannot prevent aliasing and that the progressive application of SampEn employs a badly designed coarse-graining step which makes MSE artificially decrease as a function of the time scale. These problems were overcome by the introduction of the so-called refined MSE (RMSE) [11], which exploits well-designed procedures for low-pass filtering and coarse graining.

Even though RMSE and other refinements and extensions [6] have improved the performance and widened the applicability of this methodology, some limitations still remain which hamper a full exploitation of the concept underlying MSE. The main limitation is in the fact that MSE requires long time series to be reliably computed: as any nonparametric entropy estimator, the SampEn is highly data-demanding, and the issue is exacerbated at increasing time scales where the rescaled signals get progressively shorter. This issue is only partly addressed by versions of MSE that intensify the number of patterns over which SampEn is calculated [15, 16], because the problem ultimately stands in the lower number of cycles of the slowest oscillations that are available for a given data length. Another limitation is in the fact that the length of the temporal scales which can be explored is usually expressed in number of samples and is restricted to integer values, thus limiting the possibility of fine-tuning the elimination of the fast temporal scales. These issues strongly limit the applicability of MSE or RMSE to short biomedical time series, such as those typically considered in short-term cardiovascular variability, where the dynamics of interest are deployed over a few minutes (~ 300 samples) [17]. Moreover, in this applicative context fine-tuning of the filtering process is fundamental to elicit changes in complexity related to the oscillatory rhythms typically observed in cardiovascular variability, that is, low-frequency (LF, from 0.04 Hz to 0.15 Hz) and high-frequency (HF, from 0.15 Hz to 0.5 Hz) oscillations [18, 19]. A further issue with existing MSE techniques is the lack of theoretical procedures to derive exact values of the multiscale complexity for known dynamical processes which one can study analytically and can fully control; though being theoretical, this issue has practical implications as analytical methods would allow one to assess the estimation bias of existing MSE measures.

The present study introduces a new approach for the evaluation of multiscale complexity that is explicitly designed to address the limitations of existing MSE methods described above. The approach builds on recent theoretical advances providing exact techniques for the analytical computation of information-theoretic measures, including entropy rates, for linear autoregressive (AR) stochastic processes [20–23]. The key steps of the derivation of the new measure of multiscale complexity, which we denote as linear MSE (LMSE), are the closed-form representation of filtered and downsampled AR

processes as state-space (SS) processes and the analytical quantification of complexity from SS parameters. Following these steps, we devise a procedure which allows an exact computation of the refined version of MSE for linear processes, starting from their AR parameters and from the desired scale factor. Additionally, a technique to set the scale factor at any rational number instead of an integer number is devised, thus allowing the fine-tuning of the filtering step of MSE computation. The proposed approach is tested on simulations of linear stochastic processes, first in a theoretical analysis aimed at relating multiscale complexity to the dynamical properties of the process without the unavoidable errors connected with estimation procedures and then in realizations of these simulated processes generated to assess the computational reliability of the LMSE compared with the traditional RMSE estimator. Moreover, the comparative ability of LMSE and RMSE in assessing the multiscale complexity of short-term cardiovascular variability is assessed on the beat-to-beat time series of the heart period, arterial pressure, and respiration measured in a large group of healthy subjects resting in a relaxed state and during two commonly studied conditions of physiological stress, that is, postural stress and mental stress.

2. Approaches to Multiscale Entropy Analysis

2.1. Multiscale Entropy. Multiscale entropy (MSE), originally proposed by Costa et al. [5], is a measure assessing the complexity of a process across multiple temporal scales. Its computation relies first on rescaling the observed process (i.e., focusing on a specific range of temporal scales) and then on assessing the dynamical complexity of the rescaled process through the computation of its rate of entropy production.

Specifically, let us consider a discrete-time, stationary stochastic process x with zero mean and variance σ_x^2 . Let us further define $x(n)$ as the variable obtained sampling the process at the discrete time n and set an integer scale factor τ . The rescaling step is aimed at eliminating the fast temporal scales and is performed applying the following transformation to the original process x , which averages τ consecutive samples to yield the rescaled process x_d :

$$x_d(n) = \frac{1}{\tau} \sum_{i=0}^{\tau-1} x(n\tau - i); \quad (1)$$

the rescaled process is denoted here as x_d because it is a downsampled version of the original process x ; in fact, in (1) one sample of the process x_d is obtained by taking one sample every τ values of the averaged version of x , where averaging is performed over τ consecutive samples. The second part of the procedure is implemented by calculating the conditional entropy of the present state of the rescaled process, $x_d(n)$, given its past states, $x_d^-(n) = [x_d(n-1), x_d(n-2), \dots]$. In MSE, this is accomplished approximating the past history of the rescaled process with the finite-length pattern $x_d^m(n) = [x_d(n-1), \dots, x_d(n-m)]$ and then on estimating the conditional entropy by means of the Sample Entropy (SampEn) metric [13]. SampEn is an estimate of the probability that patterns which are detected as similar in the m -dimensional

space remain similar also in the $(m + 1)$ -dimensional space when they are incremented with their future values, that is, the probability that $x_d^{m+1}(n_1 + 1)$ and $x_d^{m+1}(n_2 + 1)$ are similar given that $x_d^m(n_1)$ and $x_d^m(n_2)$ are similar; similarity is assessed through a coarse-graining procedure which applies the Heaviside step function with parameter r on the distance between patterns (i.e., r is a threshold such that $x_d^m(n_1)$ and $x_d^m(n_2)$ are similar if their distance in the m -dimensional space is lower than r , where the distance is assessed using the maximum norm).

The MSE measure resulting from the above procedure, which we denote as $C_x(\tau)$, quantifies the dynamical complexity of the original process x observed at scale τ . The free parameters of the MSE estimator are the length m of the patterns used to approximate the past of the process and the threshold r that sets the similarity between patterns. In the application of MSE, the pattern length is limited to a few samples (typically, $m = 2$), while the threshold distance is a fraction of the standard deviation of the original process (typically, $r = 0.2 \cdot \sigma_x$).

2.2. Refined Multiscale Entropy. The original MSE formulation suffers from two main limitations: the suboptimal procedure for elimination of the fast temporal scales implemented by (1), which tends to introduce spurious oscillations in the rescaled time series, and the fact that the threshold parameter of the coarse graining is kept at a constant value for all time scales, which brings about an artificial reduction of the MSE values with increasing scales. Valencia et al. [11] proposed a refined MSE (RMSE) measure that is aimed at overcoming these limitations. The solution to the first limitation is based on the rationale that the rescaling procedure actually consists of two steps: a filtering step which eliminates the fast temporal scales from the original process and a downsampling step that eliminates the redundancy introduced by the first step. Accordingly, the filtering step was designed to limit as much as possible the aliasing that can occur with the following downsampling step.

In fact, the change of scale of a process x is performed first applying a low-pass filter to obtain the filtered process $x_f(n)$ and then reducing the sampling rate of the filtered process to obtain the rescaled (downsampled) process $x_d(n)$. The two steps yield, respectively, the processes

$$x_f(n) = \sum_{i=0}^q b(i) x(n-i) - \sum_{j=1}^r c(j) x_f(n-j), \quad (2a)$$

$$x_d(n) = x_f(n\tau), \quad (2b)$$

where $b(i)$ and $c(j)$ are the filter coefficients and q and r are the filter orders. As it averages τ consecutive samples, the original MSE formulation [5] implicitly applies a finite impulse response (FIR) low-pass filter [24] of orders $q = \tau - 1$ and $r = 0$ (the filter is nonrecursive) and with coefficients $b(i) = 1/\tau$ for each $i = 0, 1, \dots, q = \tau - 1$; the cutoff frequency of the filter is constrained to the value $f_\tau = 1/(2\tau)$. To improve elimination of the fast temporal scales and satisfy the Shannon theorem in the subsequent downsampling

step, the RMSE approach implements an infinite impulse response (IIR) avoiding ripples in the stop band; specifically, a Butterworth filter [24] of order q (with $r = q - 1$) is implemented in which the coefficients $b(i)$ and $c(j)$ are set to realize a low-pass implementation with cutoff frequency f_τ [11]. This choice brings also the advantage that since the cutoff frequency f_τ can take any real value, in RMSE the scale factor $\tau = 1/(2f_\tau)$ is not constrained to integer numbers as in MSE; this allows filtering out the oscillatory components of the original process with a better resolution before computing the complexity of the rescaled process.

Besides the type of filter, another crucial difference exists between the original and refined formulations of MSE. Whereas in MSE the parameter r is constant for all scale factors, as it is set at a fraction of the standard deviation of the original process (e.g., $r = 0.2 \cdot \sigma_x$), in RMSE this parameter is set at a fraction of the standard deviation of the rescaled process (e.g., $r = 0.2 \cdot \sigma_{x_d}$). This choice is meant to reduce the dependence of the estimated conditional entropy on the decrease of variance due to filtering. As a consequence of this refinement, RMSE does not exhibit a tendency to decline with the time scale as a consequence of the inherent reduction of the variance of the filtered process: for example, for a white noise process MSE decreases with the time scale while RMSE is constant.

2.3. Linear Multiscale Entropy. In this section we present a method to assess the multiscale complexity of linear Gaussian stochastic processes. The method is based on the fact that if the variables obtained sampling the considered process x have a joint Gaussian distribution, all the variability that defines the entropy rate of the process is fully captured by a linear autoregressive (AR) model [25]. Accordingly, let us consider the linear AR representation of the process x , defined as

$$x(n) = \sum_{k=1}^p a(k) x(n-k) + e(n), \quad (3)$$

where p is the order of the process, $a(k)$, $k = 1, \dots, p$, are linear regression coefficients describing the interactions between the process variables as a function of the lag k , and $e(n)$ is an innovation process with variance σ_e^2 . The process x is fully described by the AR model (3) when the innovation process e is formed by uncorrelated Gaussian variables.

Given the AR representation, the variance of the process and of the innovations can be used to derive an information-theoretic description of the statistical structure of the process [26]. Specifically, the entropy of the process is related to its variance by [27]

$$H_x = \frac{1}{2} \ln 2\pi e \sigma_x^2, \quad (4)$$

and the entropy rate of the process, that is, the conditional entropy of the present $x(n)$ given the past $x^-(n) = [x(n-1), x(n-2), \dots]$, is related to the variance of the innovations by [28]

$$H_{x|x^-} = \frac{1}{2} \ln 2\pi e \sigma_e^2. \quad (5)$$

Then, (4) and (5) can be combined to provide a version of the conditional entropy which quantifies the dynamical complexity of the normalized process as

$$C_x = H_{x|x^-} - H_x + \frac{1}{2} \ln 2\pi e = \frac{1}{2} \ln 2\pi e \frac{\sigma_e^2}{\sigma_x^2}; \quad (6)$$

note that applying (6) corresponds to computing the conditional entropy of the original process after normalizing it to unit variance.

Now we turn to show how to compute analytically the variance of the AR innovations after rescaling the original process, in a way such that this variance can be used as in (6) to assess multiscale complexity. First, we perform a virtual upsampling of the original process x , by setting an integer upsampling factor s and by defining the process

$$x_u(n) = \sum_{k=1}^{p_s} a_u(k) x_u(n-k) + e_u(n), \quad (7)$$

where $p_s = p \cdot s$ is the order of the upsampled process and the coefficients a_u are set to be a zero-padded version of the original AR coefficients; that is, we set $a_u(ls) = a(l)$ for each $l = 1, \dots, p$ and $a_u(k) = 0$ for each $k = 1, \dots, p_s, k \neq ls$. Note that the innovations for the original and upsampled processes are formally the same, that is, $\sigma_{e_u}^2 = \sigma_e^2$. Then, we filter the upsampled process using a linear FIR filter, that is, a filter implemented as in (2a) with $c(j) = 0$ for each j ; with some algebraic manipulation we find that, substituting (7) in (2a) applied with $r = 0$ and with x_u in place of x , the filtered representation of the upsampled process takes the form

$$x_f(n) = \sum_{k=1}^{p_s} a_u(k) x_f(n-k) + \sum_{i=0}^q b(i) e_u(n-i). \quad (8)$$

Equation (8) shows that the filtering step introduces a moving average (MA) component in the original AR process, transforming it into an ARMA process [21]. Then, we exploit the connection between ARMA processes and state-space (SS) processes [29] to evidence that the ARMA process (8) can be expressed in an SS form as

$$\begin{aligned} Z_f(n+1) &= \mathbf{A}_f \cdot Z_f(n) + K_f e_f(n), \\ x_f(n) &= C_f \cdot Z_f(n) + e_f(n), \end{aligned} \quad (9)$$

where $Z_f(n) = [x_f(n-1) \cdots x_f(n-p_s) u(n-1) \cdots u(n-q)]^T$ is a $(p_s + q)$ -dimensional state process, $e_f(n)$ is the scalar

SS innovation process defined as $e_f(n) = b(0)e_u(n)$, and the vectors C_f and K_f and the matrix \mathbf{A}_f are defined as

$$\begin{aligned} \mathbf{A}_f &= \begin{bmatrix} a_u(1) & \cdots & a_u(p_s-1) & a_u(p_s) & b(1) & \cdots & b(q-1) & b(q) \\ 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \\ K_f &= \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ b(0) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \\ C_f &= [a_u(1) \cdots a_u(p_s-1) \ a_u(p_s) \ b(1) \cdots b(q-1) \ b(q)]. \end{aligned} \quad (10)$$

Such a connection between ARMA and SS models allows finding, through (10) and the relation $\sigma_{e_f}^2 = b(0)^2 \sigma_{e_u}^2$, the parameters of the SS process descriptive of the process x_f . Such a process is a filtered version of the upsampled process x_u , obtained passing x_u through a FIR low-pass filter with cutoff frequency equal to $f_\tau = 1/(2\tau)$; therefore, since the sampling frequency of the upsampled process x_u is s times higher than that of the original process x , the cutoff frequency of the low-pass filter applied to x for obtaining x_f becomes $f_{\tau_s} = s/(2\tau)$.

The next step of the procedure is to provide a representation for the downsampled process x_d , where downsampling is applied to the filtered process x_f . To this end, we exploit recent theoretical findings leading to an analytical derivation of the parameters of the SS model which describes the downsampled process [20, 21, 30]. According to these findings, the downsampled process $x_d(n) = x_f(n\tau)$ can be represented in SS form involving a vector state process Y as

$$\begin{aligned} Y(n+1) &= \mathbf{A}_d \cdot Y(n) + W(n) \\ x_d(n) &= C_d \cdot Y(n) + v(n), \end{aligned} \quad (11)$$

where the parameters of the SS model (11) are the matrix \mathbf{A}_d , the vector C_d , the variance of the scalar noise process $v(n)$, σ_v^2 , the covariance matrix of the vector state noise process $W(n)$, Σ_W , and the cross-covariance between $v(n)$ and $W(n)$,

Σ_{Wv} . These parameters can be computed at scale τ , in terms of parameters previously obtained, as follows [20, 30]:

$$\begin{aligned} \mathbf{A}_d &= (\mathbf{A}_f)^T; \\ C_d &= C_f; \\ \sigma_v^2 &= \sigma_{e_f}^2 \\ \Sigma_{Wv} &= (\mathbf{A}_f)^{T-1} K_f \sigma_{e_f}^2 \\ \Sigma_W(\tau) &= \mathbf{A}_f \Sigma_W(\tau-1) \mathbf{A}_f^T + \sigma_{e_f}^2 K_f K_f^T, \quad \tau \geq 2 \\ \Sigma_W(\tau) &= \sigma_{e_f}^2 K_f K_f^T, \quad \tau = 1. \end{aligned} \quad (12)$$

Then, the SS model (11) can be transformed in a form similar to that of (9) which evidences the innovations:

$$\begin{aligned} Z_d(n+1) &= \mathbf{A}_d \cdot Z_d(n) + K_d e_d(n), \\ x_d(n) &= C_d \cdot Z_d(n) + e_d(n), \end{aligned} \quad (13)$$

by solving a so-called discrete algebraic Riccati equation [20, 30]

$$\begin{aligned} \mathbf{P} &= \mathbf{A}_d \mathbf{P} \mathbf{A}_d^T + \Sigma_W - (\mathbf{A}_d \mathbf{P} C_d + \Sigma_{Wv}) (C_d \mathbf{P} C_d^T + \sigma_v^2)^{-1} \\ &\cdot (C_d \mathbf{P} \mathbf{A}_d^T + \Sigma_{Wv}^T) \end{aligned} \quad (14)$$

applied to the parameters of the SS model (11), to complete the derivation of the parameters of (13) as

$$\begin{aligned} \sigma_{e_d}^2 &= C_d \mathbf{P} C_d^T + \sigma_v^2 \\ K_d &= \frac{\mathbf{A}_d \mathbf{P} C_d^T + \Sigma_{Wv}}{\sigma_v^2}. \end{aligned} \quad (15)$$

Finally, the variance of the downsampled process can be computed analytically solving a discrete-time Lyapunov equation [20] as

$$\begin{aligned} \mathbf{\Omega} &= \mathbf{A}_d \mathbf{\Omega} \mathbf{A}_d^T + \sigma_v^2 K_d K_d^T \\ \sigma_{x_d}^2 &= C_d \mathbf{\Omega} C_d^T + \sigma_v^2. \end{aligned} \quad (16)$$

The derivations above allow computing analytically all the parameters of the state-space process (see (13)) which describes the filtered (see (9)) and downsampled (see (13)) versions of the upsampled process (see (7)), which in turn can be analytically described starting from the original AR process x given by (3). Among these parameters, the relevant ones for our purpose are the variance of the downsampled process x_d and that of the relevant innovations e_d , which are used to compute the complexity of x_d in analogy to (6), thus obtaining our MSE measure:

$$C_x(\tau_s) = C_{x_d} = \frac{1}{2} \ln 2\pi e \frac{\sigma_{e_d}^2}{\sigma_{x_d}^2}. \quad (17)$$

Note that MSE measure defined in (17), which we denote as linear MSE (LMSE), is relevant to a time scale given by the rational number $\tau_s = \tau/s$, obtained setting integer values for the scale factor τ and the upsampling factor s . As this corresponds to applying to the original process a low-pass filter with cutoff frequency $f_{\tau_s} = 1/(2\tau_s)$, such cutoff frequency can be arranged according to the value of τ_s in order to provide fine-tuning of the filtering process. Moreover we stress that since we filter the original process with a FIR filter that prevents aliasing and we compute at each time the complexity of the normalized process, our state-space MSE measure follows the philosophy of the RMSE method rather than that of the original MSE.

2.4. Practical Implementation and Applicability of LMSE and RMSE. The proposed approach for multiscale complexity analysis is implemented in the LMSE MATLAB® toolbox, which includes the algorithms for computing LMSE and RMSE for the simulated processes and exemplificative realizations of the cardiovascular data studied in this paper. The toolbox is uploaded as supplementary material to this article and is freely available for download from <http://www.lucafaes.net/LMSE.html>. The codes allow also nonexpert users to implement multiscale complexity analysis without the need to go deep into the mathematical details presented in the previous subsections. Complexity is assessed at a given time scale, after removing the faster temporal scales through low-pass filtering, as the unpredictability of future values of the time series given past observations. Unpredictability is quantified either using a model-free estimation of the conditional entropy (RMSE) or using a linear model (LMSE). As we will see in the next two sections, the analytical expressions underlying the computation of LMSE make this measure much more reliable than RMSE in the presence of short time series and long temporal scales. On the other hand, one should bear in mind that the context of application for LMSE is that of time series which are adequately represented by linear stationary stochastic processes. Linear stochastic processes are nondeterministic dynamic processes for which the relation between samples can be described by linear equations. The range of applicability of this simple generative model for time series data is very broad, including time series collected from many physical, biological, and social systems. Nevertheless, if the observed system is supposed to generate time series with strong nonlinear dynamics or significant departures from the null hypothesis of Gaussianity of the distribution are proven, the LMSE measure may capture only partly the complexity of these time series; in such a case the RMSE measure is a more appropriate choice. Furthermore, the hypothesis of stationarity implies that the statistical properties of the observed process (e.g., mean and variance) do not vary across time. For time series with evident nonstationary behaviors, the potential applicability of the method here proposed to short time series allows its easy adaptation to time-varying formulations.

3. Simulation Study

To investigate the theoretical profiles of the dynamical complexity of a stochastic process as a function of the parameters

determining its dynamics, as well as to assess the computational reliability of the proposed estimator of the refined MSE, in this section we consider a set of linear processes simulated with varying oscillatory components, for which we compute exact values of MSE and compare LMSE and RMSE estimates obtained from short process realizations.

3.1. Simulation Design. Simulations are designed considering AR processes described by l pairs of complex-conjugate poles with assigned modulus ρ_i and phase $\varphi_i = 2\pi f_i$, $i = 1, \dots, l$ (the order of the process is $p = 2l$). Type 1 simulations are designed with one pair of complex-conjugate poles ($l = 1$, $p = 2$) according to two configurations: (a) fixing the pole phase to $\varphi_1 = \pi/5$ (frequency $f_1 = 0.1$ Hz) and varying the modulus in the range $\rho_1 = \{0, 0.5, 0.8, 0.9\}$; (b) fixing the pole modulus to $\rho_1 = 0.8$ and varying the frequency in the range $f_1 = \{0.1, 0.2, 0.3, 0.4\}$. Type 2 simulations are designed with two pairs of complex-conjugate poles ($l = 2$, $p = 4$) according to two configurations both with a fixed low-frequency pole obtained setting $\rho_1 = 0.8$ and $f_1 = 0.1$; (c) fixing the phase of the second pole to $\varphi_2 = \pi/2$ (frequency $f_2 = 0.25$ Hz) and varying the modulus in the range $\rho_2 = \{0, 0.3, 0.6, 0.8\}$; (d) fixing the modulus of the second pole to $\rho_2 = 0.8$ and varying the frequency in the range $f_2 = \{0.15, 0.25, 0.3\}$. Type 1 simulations (configurations (a, b)) are set to study how the complexity of a process with a single stochastic oscillatory component varies with the amplitude and the frequency of such component. Type 2 simulations (configurations (c, d)) are set to study how the complexity of a process with two stochastic oscillatory components varies with the mismatch in the amplitude and in the frequency of these two components.

Given the configurations described above, first we determine the theoretical values of the MSE using the procedure described in Section 2.3. To this end, the exact values of the AR coefficients are determined as the coefficients of the polynomial with roots given by the poles set in the complex plane as $z_i = \rho_i \cdot (\cos \varphi_i \pm j \sin \varphi_i)$; these coefficients, together with the innovation variance $\sigma_e^2 = 1$, are set as parameters of the AR model of (3). From these parameters, MSE values are computed for specific values of the upsampling factor s and the scale factor τ , that is, $(s, \tau) = \{(1, 1), (8, 9), (4, 5), (7, 10), (3, 5), (5, 9), (1, 2), (8, 18), (2, 5), (7, 20), (3, 10), (1, 4), (1, 5), (3, 20), (1, 10), (1, 20)\}$, chosen in order to yield an approximately uniform range of values for the cutoff frequency of the rescaling low-pass filter, that is, $f_{\tau_s} = s/(2\tau) = \{0.5, 0.444, 0.4, 0.35, 0.3, 0.278, 0.25, 0.222, 0.2, 0.175, 0.15, 0.125, 0.1, 0.075, 0.05, 0.025\}$; the filter order is set to $q = 48$.

After theoretical analysis, practical estimation of MSE was performed choosing two representative cases of the parameter setting for Type 1 simulation ($\rho_1 = 0.8$, $f_1 = 0.1$) and Type 2 simulation ($\rho_1 = 0.8$, $f_1 = 0.1$; $\rho_2 = 0.8$, $f_2 = 0.2$), generating for each case 100 realizations of the simulation by feeding the model of (3) with realizations of 300 white noise samples taken from the Gaussian distribution with zero mean and unit variance and computing LMSE and RMSE estimates. For each realization, LMSE was obtained through the procedure described in Section 2.3, identifying an AR

model through the standard least squares method, setting the model order according to the Bayesian Information Criterion [31], and using a FIR low-pass filter of order $q = 48$ [24]; RMSE was obtained through the procedure described in Section 2.2 and using the parameters of [11], that is, filtering the data with a Butterworth low-pass filter of order $q = 6$ [24] and calculating SampEn with embedding and tolerance parameters $m = 2$ and $r = 0.2 \cdot \hat{\sigma}_{x_d}$ ($\hat{\sigma}_{x_d}$ is the variance of the downsampled signal).

3.2. Theoretical Analysis. Results of the theoretical analysis are reported in Figure 1. As a general result, the complexity of the simulated AR processes tends to increase at increasing the time scale $\tau_s = \tau/s$ (i.e., at decreasing the cutoff frequency of the low-pass filter applied to the original process, $f_{\tau_s} = 1/(2\tau_s) = s/(2\tau)$). This result is expected, because filtering removes the oscillatory components that bring regular dynamics to the signal, in a way such that when all stochastic oscillations are removed the signal is left with no regular dynamics and the maximum complexity level is achieved (i.e., $C_x = 0.5 \ln 2\pi e$, corresponding to uncorrelated white noise).

Figures 1(a) and 1(b) report the exact values of MSE computed for simulated stochastic processes featuring a single oscillatory component. Generating oscillations with fixed frequency and varying amplitude (Figure 1(a)), the MSE of the process increases at decreasing the pole modulus, documenting the higher complexity of stochastic oscillations associated with poles with higher distance from the unit circle in the complex plane [32]. Looking at the multiscale behavior, MSE reaches its maximum value when the regular component, oscillating at $f_1 = 0.1$ Hz in this example, is completely removed by the filtering procedure; the slope of the rise in complexity decreases with the pole modulus, becoming zero for $\rho_1 = 0$ when the process is a Gaussian white noise without regular dynamics. The behavior of MSE for oscillations with fixed amplitude ($\rho_1 = 0.8$) and varying frequency is more complicated (Figure 1(b)). At scale one ($f_{\tau_s} = 0.5$), the MSE is the same if the frequency of the oscillatory dynamics, f_1 , has the same distance from half the Nyquist frequency of 0.25 Hz and decreases with such a distance; the same symmetric behavior, with maximum complexity at half the Nyquist frequency, was observed in [33]. Then, the multiscale behavior of the complexity measure is related to the frequency of the stochastic oscillation in a way such that faster oscillations are removed more easily than slower oscillations, and thus MSE reaches its higher values for higher values of f_{τ_s} (see the trends of $f_1 = 0.4$ versus $f_1 = 0.1$ and $f_1 = 0.3$ versus $f_1 = 0.2$, in Figure 1(b)).

Figures 1(c) and 1(d) show the exact values of MSE computed for simulated stochastic processes featuring two oscillatory components. Simulating the rise of a high-frequency component in the presence of a stable low-frequency stochastic oscillation (Figure 1(c), where $\rho_1 = 0.8$, $f_1 = 0.1$, and ρ_2 increases with fixed $f_2 = 0.25$) determines an increase of MSE that is revealed across multiple temporal scales. Simulating the progressive separation of stochastic oscillations within a process (Figure 1(d), where $\rho_1 = 0.8$, $f_1 = 0.1$, and f_2 increases with fixed $\rho_2 = 0.8$) determines again an increase of MSE across multiple scales. These results suggest that the

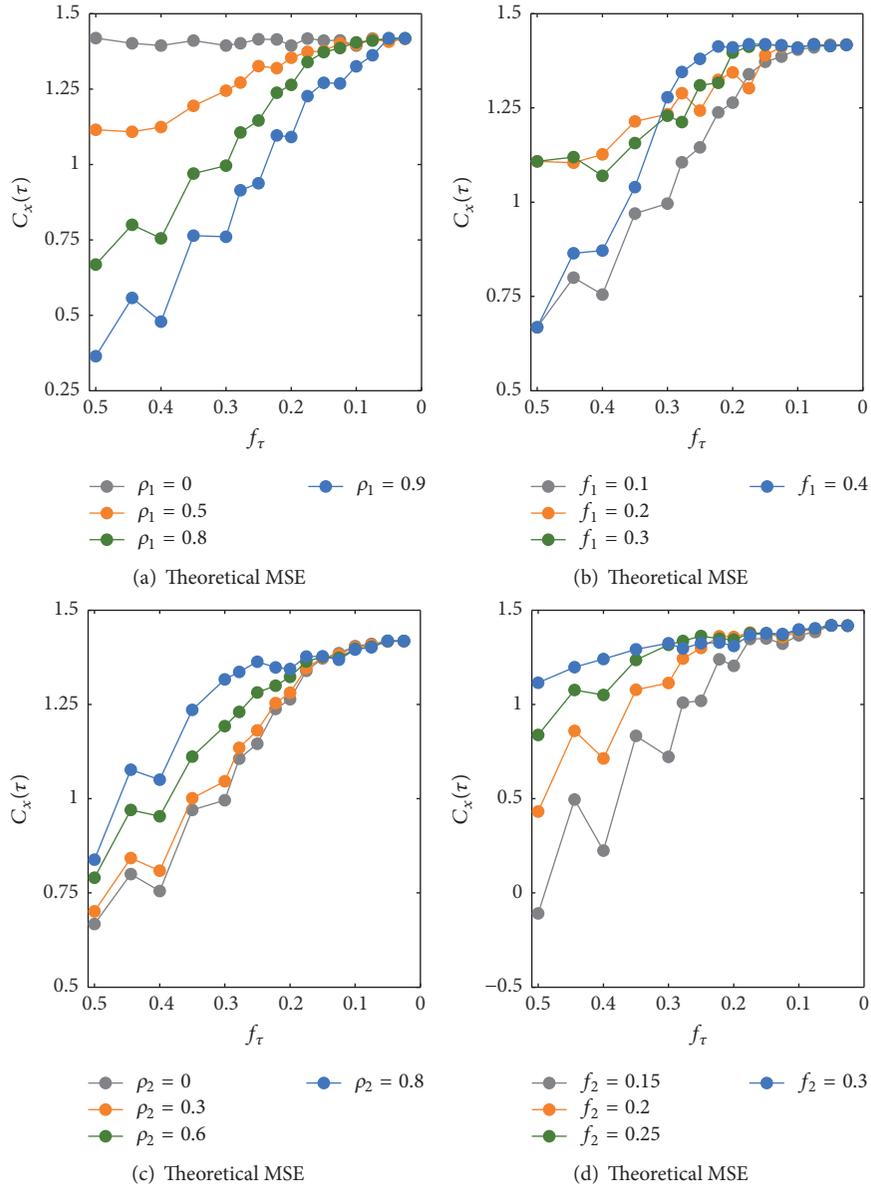


FIGURE 1: Theoretical profiles of the multiscale entropy (MSE) computed applying the proposed approach to the true parameters of simulated AR processes. Plots depict the exact values of MSE ($C_x(\tau_s)$) computed as a function of the cutoff frequency of the low-pass filter used to eliminate the fast temporal scales of the process (f_τ) for different values of the modulus ρ_1 (a) and of the frequency f_1 (b) of the pole set in Type 1 simulations, and for different values of the modulus ρ_2 (c) and frequency f_2 (d) of the second pole set in Type 2 simulations. The multiscale complexity of AR processes increases with decreasing the amplitude (a) or moving the frequency of a single stochastic oscillation closer to half the Nyquist frequency (b), as well as adding a second stochastic oscillation with increasing amplitude (c) or with increasing frequency mismatch compared to the first one (d).

simultaneous presence of multiple oscillatory mechanisms tends to produce more complex dynamics than in the case of single mechanisms, with a complexity degree that increases with the strength of the stochastic oscillations and with the mismatch of their frequency. These findings are supported by the results of recent studies [32, 33] and are observed over a range of time scales that comprises the characteristic periods of all oscillations.

3.3. Estimation Performance. Figure 2 reports the results of the practical estimation of complexity over short realizations

of the simulations, performed using the refined approach and the linear approach to MSE computation. For both the parameter settings chosen to simulate an individual stochastic oscillation (Figures 2(a) and 2(b); $\rho_1 = 0.8$, $f_1 = 0.1$) and a pair of stochastic oscillations (Figures 2(c) and 2(d); $\rho_1 = 0.8$, $f_1 = 0.1$; $\rho_2 = 0.8$, $f_2 = 0.2$), it is evident that the LMSE estimator proposed in this study outperforms the RMSE estimator. In fact, LMSE estimates are substantially unbiased, since the median estimated profile of multiscale complexity overlaps with the true profile, and exhibit a low variability around their median value, particularly at

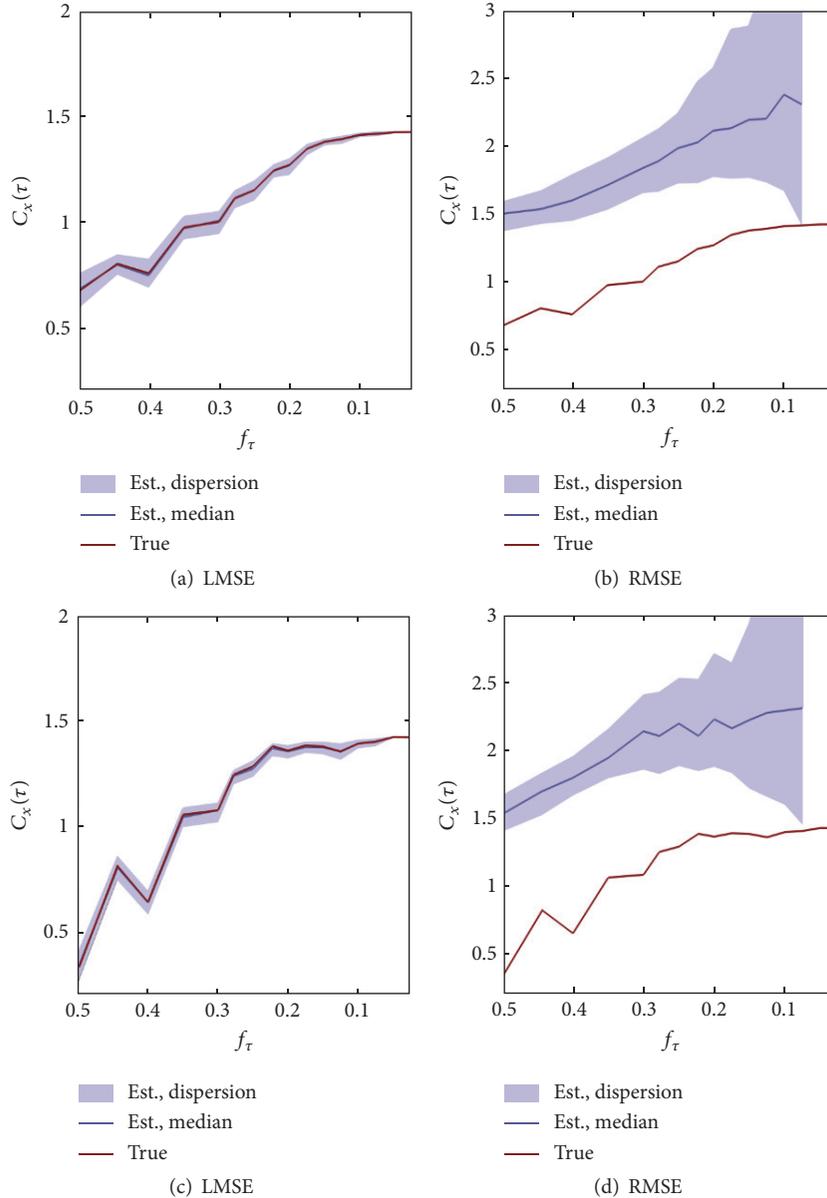


FIGURE 2: Estimation of multiscale entropy (MSE) over finite-length realizations of simulated AR processes. Plots depict the exact values (red lines) and the distributions (median and 10th–90th percentiles) of the MSE estimates ($C_x(\tau_s)$) computed as a function of the cutoff frequency of the low-pass filter used to eliminate the fast temporal scales of the process (f_τ) for representative parameter settings of Type 1 simulation (a, b) and of Type 2 simulation (c, d). Estimates are obtained using the linear MSE (LMSE) method proposed in this study (a, c) and using the refined MSE (RMSE) method proposed in [11] (b, d). In simulated AR processes, RMSE estimates exhibit high bias and a variance increasing with the time scale, while LMSE estimates display high computational reliability at all time scales.

high time scales where they converge to the expected value $0.5 \ln 2\pi e$ (Figures 2(a) and 2(c)). On the contrary, RMSE estimates are strongly biased at all time scales and also exhibit a substantial variance that increases dramatically with the time scale (Figures 2(b) and 2(d)); note that RMSE could not be computed at the higher time scales ($f_\tau = 0.05$, $f_\tau = 0.025$) due to the limited number of data points. The poor performance of RMSE confirms the known difficulty of yielding accurate model-free complexity estimates using the SampEn estimator applied to short time series [33, 34].

4. Application to Short-Term Cardiovascular Variability Series

To illustrate the application of the proposed approach for the computation of MSE over short biomedical time series, this section reports the analysis of cardiovascular and respiratory variability series. Specifically, we compare the abilities of LMSE and RMSE in detecting the multiscale complexity of heart period (HP), systolic arterial pressure (SAP), and respiration (RESP) time series measured from a large group of

healthy subjects in a resting state condition as well as during two types of physiological stress, that is, postural stress and mental stress [35].

4.1. Experimental Protocol and Data Analysis. The analyzed time series belong to a database collected to assess the dynamics of HP, SAP, and RESP during two types of physiological stress commonly studied in cardiovascular variability, that is, postural stress induced by head-up tilt (HUT) and mental stress induced by mental arithmetics (MA); we refer to [35, 36] for a detailed description of the population and the experimental setup. Briefly, a group of 61 young healthy subjects (37 females, 17.5 ± 2.4 years old) were monitored in a relaxed state in the resting supine position (SU), in the upright position during HUT, and in the supine position during MA, measuring the surface electrocardiogram (ECG), the finger arterial blood pressure collected noninvasively by the photoplethysmographic method, and the respiration signal measured through respiratory inductive plethysmography. From these signals, the beat-to-beat time series of HP, SAP, and RESP were measured, respectively, as the sequences of the temporal distances between consecutive R peaks of the ECG, the maximum values of the arterial pressure waveform measured inside the consecutively detected heart periods, and the values of the respiratory signal sampled at the onset of the consecutively detected heart periods.

The analysis was performed on segments of 300 consecutive points, free of artifacts and satisfying stationarity requirements, extracted from the three time series for each subject and condition. The preprocessing steps consisted in removing the linear trend from each sequence and in reducing the series to zero mean. Then, for each individual time series, a linear AR model was identified using the standard least squares method and using the Bayesian Information Criterion to set the model order within the range $\{1, \dots, 12\}$ [37]. From the estimated AR parameters, LMSE was computed by means of the procedure described in Section 2.3, using a low-pass FIR filter of order $q = 48$. The computation of RMSE was performed as described in Section 2.2, using a sixth-order Butterworth low-pass filter before resampling and computing SampEn with parameters $m = 2$ and r equal to 20% of the variance of the resampled signal. As for simulations, in the computation of both LMSE and RMSE, the parameters s and τ determining the time scale were set to obtain the following values for the cutoff frequency of the resampling filter: $f_{\tau_s} = 1/(2\tau_s) = s/(2\tau) = \{0.5, 0.444, 0.4, 0.35, 0.3, 0.278, 0.25, 0.222, 0.2, 0.175, 0.15, 0.125, 0.1, 0.075, 0.05, 0.025\}$.

Statistically significant differences among the MSE profiles obtained in the three conditions (i.e., SU, HUT, and MA) were first assessed by means of the multivariate ANOVA. Then, if the null hypothesis that the means of MSE computed across time scales for each condition are the same multivariate vector was rejected, the univariate ANOVA was applied to the three distributions of MSE obtained during SU, HUT, and MA at any assigned time scale. Furthermore, if at a given time scale the null hypothesis that the means of MSE computed in the three conditions are the same number was rejected, a post hoc pairwise test (i.e., the Student t -test for paired data) was performed to assess the statistical significance

of the differences between rest and stress conditions (i.e., HUT versus. REST, or MA versus. REST). A p value < 0.05 was always assumed as statistically significant; both in the univariate ANOVA and in the pairwise tests, a Bonferroni-Holm correction for multiple comparisons was employed.

4.2. Results and Discussion. The results of multiscale complexity analysis of HP, SAP, and RESP are depicted, respectively, in Figures 3, 4, and 5; results are presented as median and interquartile range of the distributions of LMSE and RMSE computed for the 61 subjects during the three considered experimental conditions (SU, HUT, and MA). As a general result, we find that LMSE estimates exhibit lower intersubject variability than RMSE estimates, especially when increasing the scale factor (i.e., when reducing the cutoff frequency of the low-pass filter). This result holds for all time series and experimental conditions, likely reflecting the problems associated with the computation of RMSE for short time series and at long time scales. On the contrary, the variability of LMSE estimates decreases when increasing the scale factor, allowing eliciting statistically significant differences that in RMSE are masked by the larger intersubject variability.

The analysis of LMSE and RMSE computed for the HP time series, reported in Figure 3, documents that the complexity of HP is significantly lower during HUT than during REST, while no significant differences are observed between REST and MA. The lower complexity during HUT is detected up to $f_{\tau_s} = 0.225$ Hz using RMSE (Figure 3(b)) and up to $f_{\tau_s} = 0.1$ Hz using LMSE (Figure 3(a)). At scale 1 ($f_{\tau_s} = 0.5$ Hz), these results confirm a number of previous investigations documenting the decreased complexity of heart rate variability during postural stress and its unaltered complexity during mental stress [36, 38–40]. Here we extend these findings showing that during MA the complexity of HP is left unchanged at any time scale and during HUT it is kept at lower levels until both the high-frequency (HF, >0.15 Hz) and the low-frequency (LF, ~ 0.1 Hz) oscillations are removed by filtering. This latter result, which is documented only using the LSME estimator, is in agreement with a recent study showing that the reduction in complexity induced by HUT is a result of the presence of more regular HP oscillations in the LF band rather than in the HF band and is thus associated more to an enriched sympathetic modulation than to a vagal withdrawal [32].

The MSE analysis performed for the SAP time series, reported in Figure 4, evidences a clear increase in the MSE computed during MA. This result, which is again more evident using LMSE than RMSE, has been previously observed at scale 1 [36] and is possibly related to complex patterns of autonomic activation following cognitive load [41, 42]; our results, showing the persistence of higher MSE values up to long time scales ($f_{\tau_s} = 0.15$ Hz), point to an involvement of LF oscillations in the generation of higher complexity during mental stress. The multiscale pattern of complexity induced by postural stress is more complicated, indicating lower MSE during HUT than during REST for short time scales ($f_{\tau_s} = 0.5$ Hz), no significant alterations for intermediate time scales (up to $f_{\tau_s} = 0.275$ Hz), and higher MSE during HUT than during REST for longer time scales (f_{τ_s} between 0.275 Hz and

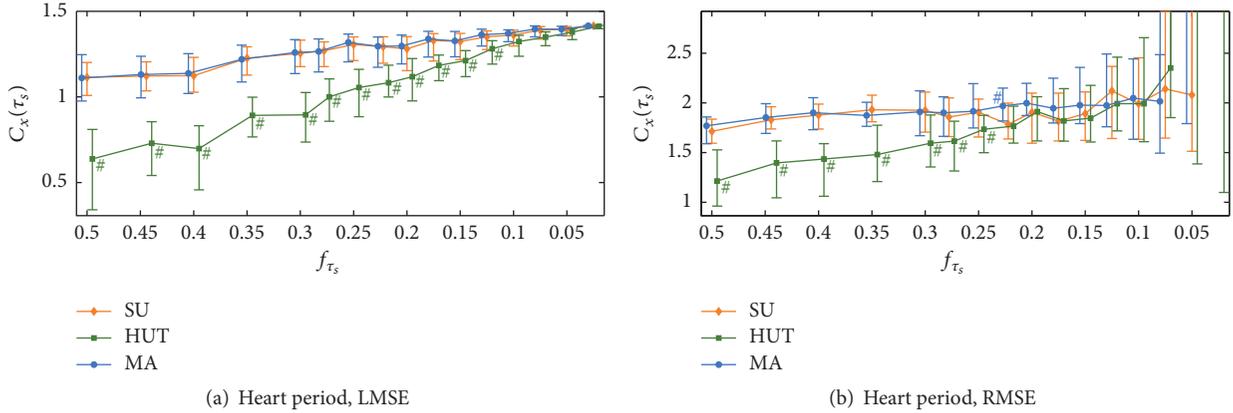


FIGURE 3: Estimation of multiscale entropy (MSE) for the time series of the heart period. Plots depict the distributions (median and 25th–75th percentiles) of the MSE estimates ($C_x(\tau_s)$) computed as a function of the cutoff frequency of the low-pass filter used to eliminate the fast temporal scales of the process (f_{τ_s}) in the resting supine position (SU, diamonds), during postural stress induced by head-up tilt (HUT, squares), and during mental stress induced by mental arithmetics (MA, circles). While MA does not induce changes, HUT evokes a significant reduction of the complexity of heart period variability, which is observed using LMSE across a wide range of time scales including both low- and high-frequency oscillations. #Statistically significant difference, HUT versus SU or MA versus SU.

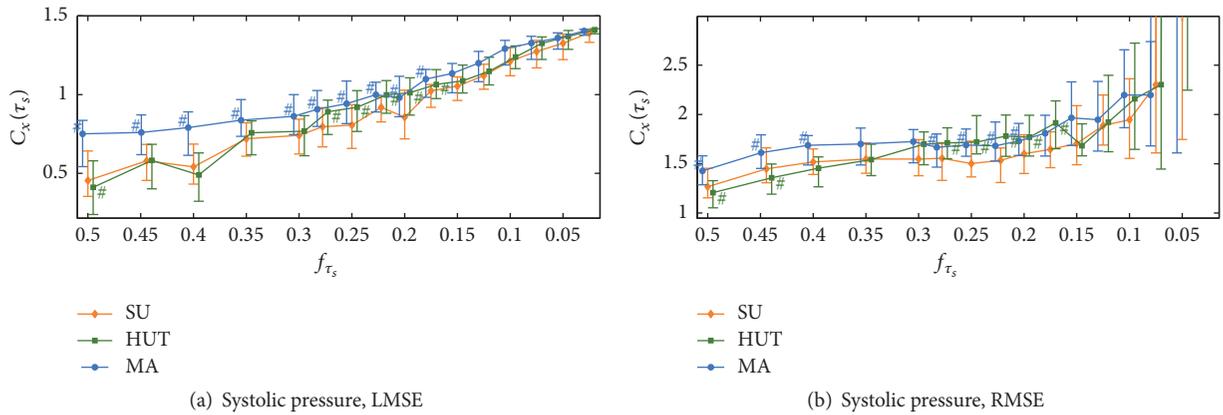


FIGURE 4: Estimation of multiscale entropy (MSE) for the time series of the systolic arterial pressure. Plots depict the distributions (median and 25th–75th percentiles) of the MSE estimates ($C_x(\tau_s)$) computed as a function of the cutoff frequency of the low-pass filter used to eliminate the fast temporal scales of the process (f_{τ_s}) in the resting supine position (SU, diamonds), during postural stress induced by head-up tilt (HUT, squares), and during mental stress induced by mental arithmetics (MA, circles). The multiscale complexity of systolic pressure dynamics increases consistently across multiple time scales during MA and increases for time scales associated with low-frequency oscillations during HUT. #Statistically significant difference, HUT versus SU or MA versus SU.

0.15 Hz). The emergence of a significantly higher complexity of SAP during HUT when HF fluctuations are filtered out is a novel finding, suggesting that the postural stress decreases the regularity of LF arterial pressure oscillations, while HF oscillations display similar or even increased regularity.

The multiscale complexity analysis of respiration variability, reported in Figure 5, documents that MSE is significantly decreased during HUT and significantly increased during MA. These variations are consistently observed across multiple scales, going up to $f_{\tau_s} = 0.3$ Hz for the lower LMSE and RMSE during HUT and up to $f_{\tau_s} = 0.15$ Hz for the higher LMSE during MA. The higher regularity of the respiratory dynamics during postural stress, which seems to be confined to the HF band where respiration usually occurs, may be related to an increased tidal volume during this condition.

On the other hand, the lower regularity of respiration during mental stress might be explained by the appearance of long pauses or sighs in the respiration pattern while performing mental arithmetics [43], making it more erratic and thus complex. The existence of more erratic patterns, which likely span a broad band of the frequency spectrum, may also explain the fact that we observe higher complexity of respiration up to long time scales, going beyond the common frequency bands at which respiratory activity is observed.

5. Conclusions and Future Developments

The present study introduces for the first time a multiscale entropy measure that is based on theoretical rather than empirical grounds and can thus be analytically computed

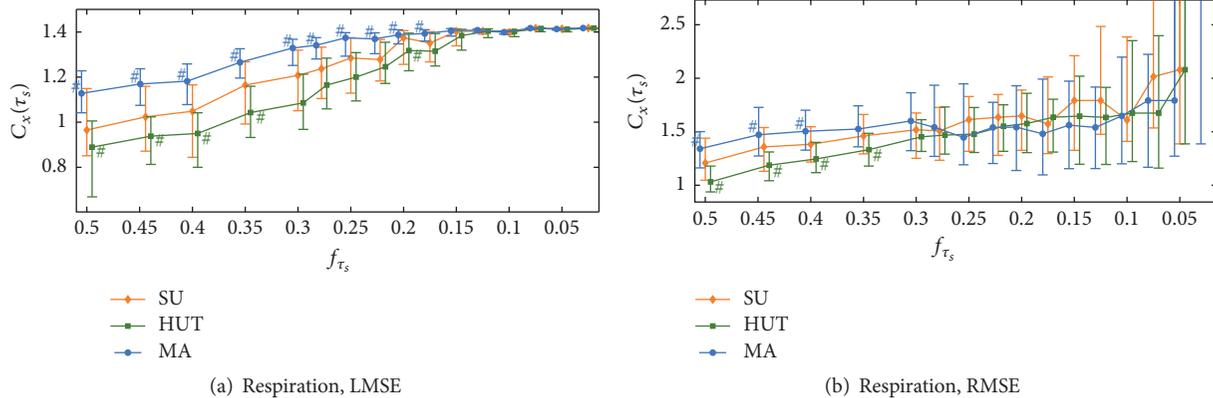


FIGURE 5: Estimation of multiscale entropy (MSE) for the time series of respiration. Plots depict the distributions (median and 25th–75th percentiles) of the MSE estimates ($C_x(\tau_s)$) computed as a function of the cutoff frequency of the low-pass filter used to eliminate the fast temporal scales of the process (f_{τ_s}) in the resting supine position (SU, diamonds), during postural stress induced by head-up tilt (HUT, squares), and during mental stress induced by mental arithmetics (MA, circles). The multiscale complexity of respiratory dynamics increases consistently across multiple time scales during MA and decreases for time scales associated with high-frequency oscillations during HUT. #Statistically significant difference, HUT versus SU or MA versus SU.

from the parametric representation of an observed stochastic process. As a matter of fact, the proposed LMSE method is highly data-efficient because it stems from simple linear parametric modeling and is thus much more reliable than MSE or its modifications [6], including RMSE, in assessing the complexity of short time series at long time scales. Compared with another recently proposed method to assess multiscale complexity from short data sequences [32], LMSE shares the philosophy of being derived for linear AR processes, but differs in the fact that it is designed closely following the two algorithmic steps of MSE computation, that is, low-pass filtering and downsampling. The high computational reliability of LMSE comes from the fact that filtering and downsampling are not actually implemented on the measured time series, but result from the analytical quantification of the impact that they have on the state-space parameters of the observed AR process.

Our approach can be fruitfully exploited, as done in the present study, to relate the multiscale complexity of a stochastic process to the parameters that establish its dynamical features, or to estimate patterns of multiscale complexity from short process realizations. In this work, we have formalized the dependence of the multiscale complexity of an AR process on the amplitude and frequency of its stochastic oscillatory components and have assessed multiscale patterns of short-term cardiovascular complexity which cannot be fully retrieved using standard MSE methods. Our results emphasize the role of the sympathetic control in driving the increased regularity of low-frequency heart rate oscillations and the increased complexity of low-frequency arterial pressure oscillations during postural stress. LMSE analysis stresses also the importance of dynamics occurring within the low-frequency band in determining the increased complexity of arterial pressure, and even that of respiration, during mental stress.

The main strength of the proposed approach, that is, the linear parametric formulation, constitutes also one of

its major limitations. In fact, the computation of LMSE holds exactly only if the observed process has a Gaussian distribution; in such a case, the linear AR description fully captures all of the variability in the process that determines the measured entropy rates, and model-free formulations as the one implemented by SampEn have no additional utility [25]. On the contrary, departures from linearity leading to non-Gaussian distributions may generate dynamics which are captured only partially by LMSE, and thus important features of multiscale complexity may be missed in this case. The suitability of LMSE for the applicative context of this study is supported by the fact that linear methods are ubiquitously exploited in short-term cardiovascular variability studies (e.g., performing parametric spectral analysis [17, 44] or parametric coupling and causality analysis [37, 45]). Nevertheless, future studies should deepen the comparison of LMSE with RMSE or other nonparametric MSE techniques in order to clarify the importance of accounting for nonlinear dynamics in the multiscale analysis of biomedical time series. On the other hand, keeping the linear analysis framework, a desirable extension of the present study would be to integrate the standard AR representation with fractionally integrated (FI) innovation modeling [46]; this would allow the computation of multiscale complexity for ARFI processes, thus incorporating in the approach the ability to describe long-range correlations phenomena which are typically assessed in long-term multiscale analyses [47].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Luca Faes and Giandomenico Nollo were partly supported by funding from the Healthcare Research and Implementation

Program (IRCS) of the Autonomous Province of Trento (PAT), Italy. Michal Javorka was supported by Grants APVV-0235-12, VEGA 1/0117/17, and VEGA 1/0202/16 and project “Biomedical Center Martin,” ITMS code 26220220187, the project cofinanced from EU sources.

Supplementary Materials

The LMSE MATLAB toolbox is available as supplementary material to this article. The package contains functions for the computation of multiscale entropy based on the linear state-space approach and on the refined approach proposed in [11], as well as scripts that allow performing multiscale complexity analysis for the simulated and real data considered in the present study. (*Supplementary Materials*)

References

- [1] W. W. Burggren and M. G. Monticino, “Assessing physiological complexity,” *Journal of Experimental Biology*, vol. 208, no. 17, pp. 3221–3232, 2005.
- [2] A. Bashan, R. P. Bartsch, J. W. Kantelhardt, S. Havlin, and P. C. Ivanov, “Network physiology reveals relations between network topology and physiological function,” *Nature Communications*, vol. 3, article no. 702, 2012.
- [3] M. A. Cohen and J. A. Taylor, “Short-term cardiovascular oscillations in man: Measuring and modelling the physiologies,” *The Journal of Physiology*, vol. 542, no. 3, pp. 669–683, 2002.
- [4] A. Porta, M. D. Rienzo, N. Wessel, and J. Kurths, “Addressing the complexity of cardiovascular regulation,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 367, no. 1892, pp. 1215–1218, 2009.
- [5] M. Costa, A. L. Goldberger, and C.-K. Peng, “Multiscale entropy analysis of complex physiologic time series,” *Physical Review Letters*, vol. 89, no. 6, Article ID 068102, p. 68102, 2002.
- [6] A. Humeau-Heurtier, “The multiscale entropy algorithm and its variants: A review,” *Entropy*, vol. 17, no. 5, pp. 3110–3123, 2015.
- [7] A. Humeau-Heurtier, G. Mahé, S. Durand, and P. Abraham, “Multiscale entropy study of medical laser speckle contrast images,” *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 3, pp. 872–879, 2013.
- [8] M. Costa, A. L. Goldberger, and C.-K. Peng, “Multiscale entropy analysis of biological signals,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 71, no. 2, Article ID 021906, 2005.
- [9] J. Courtiol, D. Perdikis, S. Petkoski et al., “The multiscale entropy: Guidelines for use and interpretation in brain signal analysis,” *Journal of Neuroscience Methods*, vol. 273, pp. 175–190, 2016.
- [10] J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López, “Analysis of electroencephalograms in Alzheimer’s disease patients with multiscale entropy,” *Physiological Measurement*, vol. 27, no. 11, article 004, p. 1091, 2006.
- [11] J. F. Valencia, A. Porta, M. Vallverdú et al., “Refined multiscale entropy: application to 24-h holter recordings of heart period variability in healthy and aortic stenosis subjects,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 9, pp. 2202–2213, 2009.
- [12] L. Angelini, R. Maestri, D. Marinazzo et al., “Multiscale analysis of short term heart beat interval, arterial blood pressure, and instantaneous lung volume time series,” *Artificial Intelligence in Medicine*, vol. 41, no. 3, pp. 237–250, 2007.
- [13] 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, no. 6, pp. H2039–H2049, 2000.
- [14] V. V. Nikulin and T. Brismar, “Comment on Multiscale Entropy Analysis of Complex Physiologic Time Series,” *Physical Review Letters*, vol. 92, no. 8, p. 89803, 2004.
- [15] Y.-C. Chang, H.-T. Wu, H.-R. Chen et al., “Application of a modified entropy computational method in assessing the complexity of pulse wave velocity signals in healthy and diabetic subjects,” *Entropy*, vol. 16, no. 7, pp. 4032–4043, 2014.
- [16] 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, no. 3, pp. 1069–1084, 2013.
- [17] M. Malik, J. Bigger, A. Camm, and R. Kleiger, “Heart rate variability. standards of measurement, physiological interpretation, and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology,” *European Heart Journal*, vol. 17, pp. 354–381, 1996.
- [18] B. Pomeranz et al., “Assessment of autonomic function in humans by heart rate spectral analysis,” *American Journal of Physiology*, vol. 248, pp. H151–H153, 1985.
- [19] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, A. C. Berger, and R. J. Cohen, “Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control,” *Science*, vol. 213, no. 4504, pp. 220–222, 1981.
- [20] L. Barnett and A. K. Seth, “Granger causality for state-space models,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 91, no. 4, Article ID 040101, 2015.
- [21] L. Faes, D. Marinazzo, and S. Stramaglia, “Multiscale information decomposition: exact computation for multivariate gaussian processes,” *Entropy*, vol. 19, p. 408, 2017.
- [22] A. Porta, B. De Maria, V. Bari, A. Marchi, and L. Faes, “Are nonlinear model-free conditional entropy approaches for the assessment of cardiac control complexity superior to the linear model-based one?” *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 6, pp. 1287–1296, 2017.
- [23] L. Faes, G. Nollo, S. Stramaglia, and D. Marinazzo, “Multiscale Granger causality,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 96, p. 42150, 2017.
- [24] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete Time Signal Processing*, vol. 1999, 1999.
- [25] A. B. Barrett, L. Barnett, and A. K. Seth, “Multivariate Granger causality and generalized variance,” *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 81, no. 4, Article ID 041907, 2010.
- [26] L. Faes, A. Porta, and G. Nollo, “Information decomposition in bivariate systems: Theory and application to cardiorespiratory dynamics,” *Entropy*, vol. 17, no. 1, pp. 277–303, 2015.
- [27] J. Kieffer, “Elements of Information Theory (Thomas M. Cover and Joy A. Thomas),” *Society for Industrial and Applied Mathematics Review*, vol. 36, no. 3, pp. 509–511, 1994.
- [28] L. Barnett, A. B. Barrett, and A. K. Seth, “Granger causality and transfer entropy are equivalent for gaussian variables,” *Physical Review Letters*, vol. 103, no. 23, Article ID 238701, 2009.
- [29] M. Aoki and A. Havenner, “State space modeling of multiple time series,” *Econometric Reviews*, vol. 10, no. 1, pp. 1–99, 1991.

- [30] V. Solo, "State-space analysis of Granger-Geweke causality measures with application to fMRI," *Neural Computation*, vol. 28, Article ID NECO.a.00828, 2016.
- [31] A. A. Neath and J. E. Cavanaugh, "The Bayesian information criterion: Background, derivation, and applications," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 2, pp. 199–203, 2012.
- [32] A. Porta, V. Bari, G. Ranuzzi, B. De Maria, and G. Baselli, "Assessing multiscale complexity of short heart rate variability series through a model-based linear approach," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 27, no. 9, p. 93901, 2017.
- [33] W. Xiong, L. Faes, and P. C. Ivanov, "Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 95, no. 6, p. 62114, 2017.
- [34] J. M. Yentes, N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath, and N. Stergiou, "The appropriate use of approximate entropy and sample entropy with short data sets," *Annals of Biomedical Engineering*, vol. 41, no. 2, pp. 349–365, 2013.
- [35] M. Javorka, J. Krohova, B. Czipelova et al., "Basic cardiovascular variability signals: mutual directed interactions explored in the information domain," *Physiological Measurement*, vol. 38, no. 5, Article ID aa5b77, pp. 877–894, 2017.
- [36] L. Faes, A. Porta, G. Nollo, and M. Javorka, "Information decomposition in multivariate systems: Definitions, implementation and application to cardiovascular networks," *Entropy*, vol. 19, no. 1, article no. 5, 2017.
- [37] L. Faes, S. Erla, and G. Nollo, "Measuring connectivity in linear multivariate processes: Definitions, interpretation, and practical analysis," *Computational and Mathematical Methods in Medicine*, vol. 2012, Article ID 140513, 2012.
- [38] A. Porta, T. Gnechi-Ruscione, E. Tobaldini, S. Guzzetti, R. Furlan, and N. Montano, "Progressive decrease of heart period variability entropy-based complexity during graded head-up tilt," *Journal of Applied Physiology*, vol. 103, no. 4, pp. 1143–1149, 2007.
- [39] A. Porta, E. Tobaldini, S. Guzzetti, R. Furlan, N. Montano, and T. Gnechi-Ruscione, "Assessment of cardiac autonomic modulation during graded head-up tilt by symbolic analysis of heart rate variability," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 293, no. 1, pp. H702–H708, 2007.
- [40] Z. Visnovcova, M. Mestanik, M. Javorka et al., "Complexity and time asymmetry of heart rate variability are altered in acute mental stress," *Physiological Measurement*, vol. 35, no. 7, pp. 1319–1334, 2014.
- [41] N. T. Kuipers, C. L. Sauder, J. R. Carter, and C. A. Ray, "Neurovascular responses to mental stress in the supine and upright postures," *Journal of Applied Physiology*, vol. 104, no. 4, pp. 1129–1136, 2008.
- [42] H. K. Lackner, I. Papousek, J. J. Batzel, A. Roessler, H. Scharfetter, and H. Hinghofer-Szalkay, "Phase synchronization of hemodynamic variables and respiration during mental challenge," *International Journal of Psychophysiology*, vol. 79, no. 3, pp. 401–409, 2011.
- [43] M. Grassmann, E. Vlemincx, A. Von Leupoldt, J. M. Mittelstädt, and O. Van Den Bergh, "Respiratory changes in response to cognitive load: A systematic review," *Neural Plasticity*, vol. 2016, Article ID 8146809, 2016.
- [44] A. Malliani, "The pattern of sympathovagal balance explored in the frequency domain," *News in Physiological Sciences*, vol. 14, no. 3, pp. 111–117, 1999.
- [45] S. Schulz et al., "Cardiovascular and cardiorespiratory coupling analyses: a review," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 371, Article ID 20120191, 2013.
- [46] S. Ling and W. K. Li, "On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity," *Journal of the American Statistical Association*, vol. 92, no. 439, pp. 1184–1194, 1997.
- [47] H. E. Stanley, S. V. Buldyrev, A. L. Goldberger et al., "Fractal landscapes in biological systems: Long-range correlations in DNA and interbeat heart intervals," *Physica A: Statistical Mechanics and its Applications*, vol. 191, no. 1-4, pp. 1–12, 1992.