Interpretation of the Lempel-Ziv Complexity Measure in the Context of Biomedical Signal Analysis

Mateo Aboy*, *Member, IEEE*, Roberto Hornero, *Member, IEEE*, Daniel Abásolo, *Member, IEEE*, and Daniel Álvarez

***Abstract*—Lempel-Ziv complexity (*LZ*) and derived LZ algorithms have been extensively used to solve information theoretic problems such as coding and lossless data compression. In recent years, *LZ* has been widely used in biomedical applications to estimate the complexity of discrete-time signals. Despite its popularity as a complexity measure for biosignal analysis, the question of *LZ* interpretability and its relationship to other signal parameters and to other metrics has not been previously addressed. We have carried out an investigation aimed at gaining a better understanding of the *LZ* complexity itself, especially regarding its interpretability as a biomedical signal analysis technique. Our results indicate that *LZ* is particularly useful as a scalar metric to estimate the bandwidth of random processes and the harmonic variability in quasi-periodic signals.**

***Index Terms*—Complex analysis, Lempel-Ziv complexity (LZ), nonlinear analysis.**

I. Introduction

THE metric of complexity proposed by Lempel and Ziv (*LZ*) to evaluate the randomness of finite sequences has been extensively used to solve information theoretic problems [1]–[9] and applications such as coding [10]–[12], data compression [13]–[19], and generation of test signals [20]–[22]. This complexity measure is related to the number of distinct substrings (i.e., patterns) and the rate of their occurrence along a given sequence [2].

In recent years, *LZ* has been applied extensively in biomedical signal analysis as a metric to estimate the complexity of discrete-time physiologic signals. For instance, *LZ* has been used for recognition of structural regularities [23], for complexity characterization of DNA sequences [24], to develop new methods for discovering patterns in DNA sequences by applying it to genomic sequences of Plasmodium falciparum [25], to characterize the responses of neurons of the primary visual cortex to different kinds of stimuli [26], and to estimate the entropy of neural discharges (spike trains) [27]. In a recent study, a new sequence distance measure for phylogenetic tree construction has been proposed based on the relative information between the sequences using *LZ* complexity [28]. *LZ* complexity has also been used to study brain function [29], brain information transmission [30], EEG complexity in patients with Alzheimer's disease [31], and epileptic seizures [32]. Other authors have used *LZ* complexity to study ECG dynamics [33], to detect ventricular tachycardia and fibrillation [34], [35], to predict movement in anaesthesia in animals [36], to estimate the depth of anaesthesia [37], [38] and to quantify oscillations in uterine electromyography [39].

Generally, previous work involving the application of *LZ* in the context of biomedical signal analysis has consisted of analyzing signals from a specific patient population or pathology, and identifying a *LZ* change associated with a specific condition of interest. However, despite its popularity as a tool for biomedical signal analysis, the question of *LZ* interpretability and its relationship to other signal parameters or metrics have not been previously addressed, and none of the previous studies carried out a thorough investigation aimed at gaining a better understanding of the *LZ* itself.

In this paper, we present the results of a study designed to enable researchers to interpret the *LZ* metric in terms of classical signal processing concepts such as frequency, number of harmonics, frequency variability of the harmonics, and signal bandwidth.

II. Methods: Lempel-Ziv Complexity

LZ complexity analysis is based on a coarse-graining of the measurements. Before calculating the *LZ* complexity measure $c(n)$, the signal must be transformed into a finite symbol sequence. In the context of biomedical signal analysis, typically the discrete-time biomedical signal $x(n)$ is converted into a binary sequence. By comparison with the threshold T_d, the signal data are converted into a 0–1 sequence P as follows [37]:

$$P = s(1), s(2), \ldots, s(n) \tag{1}$$

where

$$s(i) = \begin{cases} 0, & \text{if } x(i) < T_d \\ 1, & \text{otherwise} \end{cases}. \tag{2}$$

Usually the median is used as the threshold T_d because of its robustness to outliers [39].

Previous studies [29], [30], [32]–[34] have shown that 0–1 conversion is adequate to estimate the *LZ* complexity in biomedical signals. In order to compute *LZ* complexity, the sequence P is scanned from left to right and the complexity counter $c(n)$ is

Manuscript received October 30, 2005; revised April 30, 2006. This work was supported in part by through the Project of Junta de Castilla y León under Grant VA019-04. *Asterisk indicates corresponding author.*

*M. Aboy is with the Electronics Engineering Technology Department, Oregon Institute of Technology, Portland, OR 97006 USA (e-mail: mateoaboy@ieee.org).

R. Hornero, D. Abásolo, and D.Álvarez are with the Department of Signal Theory and Communications, ETSIT, University of Valladolid, Valladolid 47071 Spain, EU.

Digital Object Identifier 10.1109/TBME.2006.883696

increased by one unit every time a new subsequence of consecutive characters is encountered. The complexity measure can be estimated using the following algorithm [34], [37], [38].

1) Let S and Q denote two subsequences of P and SQ be the concatenation of S and Q, while sequence $SQ\pi$ is derived from SQ after its last character is deleted (π denotes the operation of deleting the last character in the sequence). Let $v(SQ\pi)$ denote the vocabulary of all different subsequences of $SQ\pi$. At the beginning, $c(n) = 1$, $S = s(1)$, $Q = s(2)$, therefore, $SQ\pi = s(1)$.
2) In general, $S = s(1), s(2), \ldots, s(r)$, $Q = s(r+1)$, then $SQ\pi = s(1), s(2), \ldots, s(r)$; if Q belongs to $v(SQ\pi)$, then Q is a subsequence of $SQ\pi$, not a new sequence.
3) Renew Q to be $s(r+1), s(r+2)$ and judge if Q belongs to $v(SQ\pi)$ or not.
4) Repeat the previous steps until Q does not belong to $v(SQ\pi)$. Now $Q = s(r+1), s(r+2), \ldots, s(r+i)$ is not a subsequence of $SQ\pi = s(1), s(2), \ldots, s(r+i-1)$, so increase $c(n)$ by one.
5) Thereafter, S is renewed to be $S = s(1), s(2), \ldots, s(r+i)$, and $Q = s(r+i+1)$.

The above procedure is repeated until Q is the last character. At this time the number of different subsequences in P—the measure of complexity—is $c(n)$. In order to obtain a complexity measure which is independent of the sequence length, $c(n)$ must be normalized. If the length of the sequence is n and the number of different symbols in the symbol set is α, it has been proved that the upper bound of $c(n)$ is given by [2]

$$c(n) < \frac{n}{(1-\epsilon_n)\log_\alpha(n)} \tag{3}$$

where ϵ_n is a small quantity and $\epsilon_n \to 0$ $(n \to \infty)$. In general, $n/\log_\alpha(n)$ is the upper bound of $c(n)$, where the base of the logarithm is α, i.e.,

$$\lim_{n\to\infty} c(n) = b(n) = \frac{n}{\log_\alpha(n)}. \tag{4}$$

For a 0–1 sequence, $\alpha = 2$, therefore

$$b(n) = \frac{n}{\log_2(n)} \tag{5}$$

and $c(n)$ can be normalized via $b(n)$

$$C(n) = \frac{c(n)}{b(n)} \tag{6}$$

where $C(n)$, the normalized *LZ* complexity, reflects the arising rate of new patterns in the sequence.

III. Simulation Study

A. Simulated Signals

In this section, we describe the signals used to study the *LZ* measure of complexity, and its interpretability in terms of classical signal processing concepts such as frequency, number of harmonics, frequency variability of signal harmonics, and signal bandwidth. These synthetic signals have been used to study the Approximate Entropy measure of irregularity, and have been described elsewhere [40]. We have included a summarized description of the signals in this paper for completeness.

Since most physiologic signals such as arterial blood pressure (ABP), intracranial pressure (ICP), electrocardiograms (ECG), pulse oximetry (SpO2), microelectrode recordings (MER), electroencephalograms (EEG), and magnetoencephalograms (MEG) are analog signals that are digitized for the purposes of analysis and storage, the digital signals we used in our study are assumed to be digital representations of analog signals. Consequently, the values of any of the time-series $x(n)$, $n = 1, \ldots, N$ we studied are assumed to correspond the measurements made on an analog physiologic signal $x(t)$ at times instances $t_n = nT_s$, where $T_s = (1/f_s)$ denotes the sampling period.

1) LZ Versus Frequency: The first test consisted in analyzing how changes in amplitude and frequency of sinusoidal signals affected the *LZ*. For this purpose we generated two synthetic signals. The first of these consisted in a constant amplitude chirp signal whose frequency was swept linearly from 0.5 Hz to 5 Hz in 40 s. The second signal was created by modulating the amplitude of the chirp signal by a pure sinusoid. *LZ* was applied to each of the two signals using a moving window of 10 s with 90% overlap with the objective of testing whether or not *LZ* is sensitive to frequency or amplitude changes. Fig. 1(a), and (b) shows the constant chirp signal, the amplitude modulated chirp, and the spectrogram of the amplitude modulated chirp, respectively.

2) LZ Versus Frequency Content: This test was designed to determine the relationship of *LZ* and the frequency content of periodic signals. We generated four periodic signals of 10 s in duration with 1, 2, 5, and 7 frequency components, respectively. The four signals were concatenated and the *LZ* metric was applied to the resulting vector using a moving window of 10 s with 90% overlap. Fig. 1(c) shows the composite signal used in this test.

3) LZ Versus Quasi-Periodic Signal Plus Noise: The objective of this test was to determine if *LZ* is sensitive to changes in noise power present in quasi-periodic signals. For this purpose we generated an amplitude-modulated harmonic quasi-periodic signal 40 s in duration, and added white Gaussian noise of various power to different portions of the signal. The noise power was increased every 10 s (1, 1.1, 1.3, and 1.5). The *LZ* metric was applied to the resulting vector using a moving window of 10 s with 90% overlap. Fig. 1(d) shows the sinusoid plus noise signal.

4) LZ Versus Noise Power: In this test, the *LZ* was applied using a moving window of 10 s with 90% overlap to a 40-s white Gaussian noise sequence with power increasing in steps each 10 s (0.1, 0.3, 0.5, and 0.7). The objective of this test was to determine if given the same underlying signal structure (white noise) the *LZ* is affected by the power of the noise. Fig. 1(e) shows this signal in the time domain.

5) LZ Versus Noise Bandwidth: This test consisted of determining the relationship between *LZ* and the noise bandwidth. The synthetic signal consisted on a 40-s time series composed

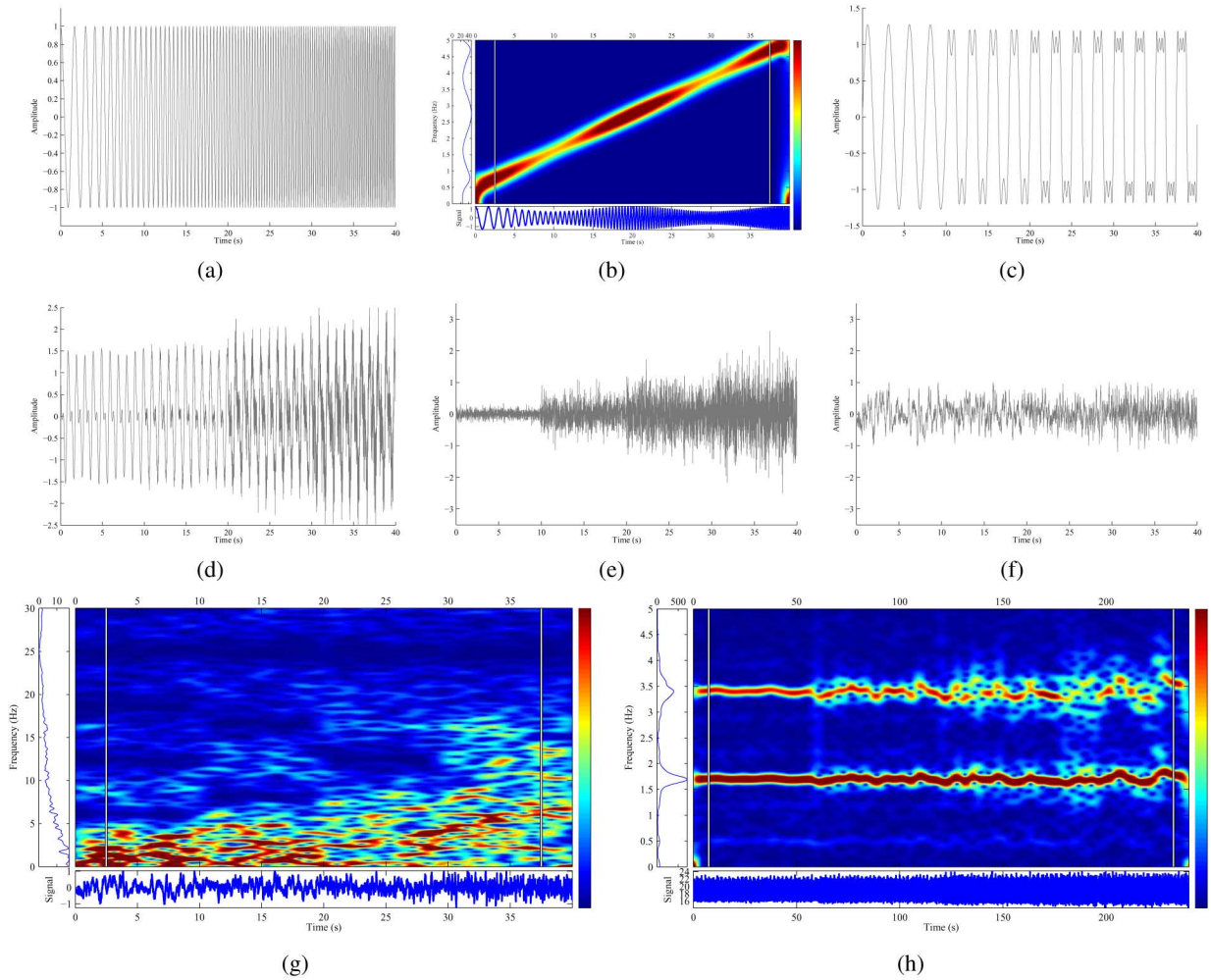

Fig. 1. Synthetic signals used in this study. (a) Chirp signal with constant amplitude ($f_1 = 0.5$, $f_2 = 5$ Hz), Test 1. (b) Amplitude modulated chirp signal. and its spectrogram. (c) Muti-tone signal used in Test 2. (d) Amplitude-modulated noisy signal with step increasing in noise power (Test 3). (e) White Gaussian Noise with step increases in noise power (Test 4). (f) Colored noise with step increases in noise bandwidth (Test 5). (g) Spectrogram of the signal shown in (f). (h) Spectrogram of the synthetic signal with step increases in stochastic variability (Test 6). (Color version available online at http://ieeexplore.ieee.org.)

of four segments of colored noise with increasing spectral bandwidth. The *LZ* metric was applied using a moving window of 10 s with 90% overlap. Fig. 1(f) and (g) show the signal in this time domain and its spectrogram, respectively.

6) LZ Versus Frequency Domain Stochastic Variability: In this test, we studied how the variability of signal harmonics affect the *LZ* complexity measure. We generated synthetic signals using a harmonic model with increasing stochastic variability of the signal harmonics. The fundamental frequency of the harmonics was modeled as an AR processes with adjustable bandwidth. This model is very general and can be used to model many biomedical signals including pressure signals, refer to [40] for details. We estimated the *LZ* of the synthetic signals using a moving window of 10 s with 90% overlap. Fig. 1(h) shows an spectrogram of a synthetic signal [41].

7) Standard Error of the LZ Versus Signal Bandwidth: In this test, we studied the standard error of *LZ* estimates as a function of the signal bandwidth. We generated 5000 independent realizations of four stochastic signals with different bandwidth: 1) gaussian white noise; 2) uniform white noise; 3) colored noise with bandwidth $B_1 = (f_s/4)$ (colored noise I); 4) colored noise with bandwidth $B_2 = (f_s/8)$ (colored noise II), where f_s denotes the sampling frequency ($f_s = 125$ in this case).

B. Example on Real Signals

As an example application on real signals we estimated the *LZ* complexity on intracranial pressure signals (ICP) containing acute intracranial hypertension (ICH) episodes. These signals were obtained from patients with brain injury admitted to the Intensive Care Unit at Doernbecher Children's Hospital, Oregon Health and Science University (Portland, OR). *LZ* was applied to detrended ICP signals (i.e., the ICP mean was removed) to study whether measures of complexity obtained from the pulse morphology alone correlate with mean ICP.

IV. Results and Discussion

The results of Test 1 show that for a sinusoidal signal with a quadratic angle $\theta = 2\pi\mu t + 2\pi f t + \phi$, *LZ* increases as the frequency of a sinusoid increases, and that amplitude modulation of a signal does not result in an increase of *LZ*, as it can be observed in Fig. 2(a) and (b), respectively.

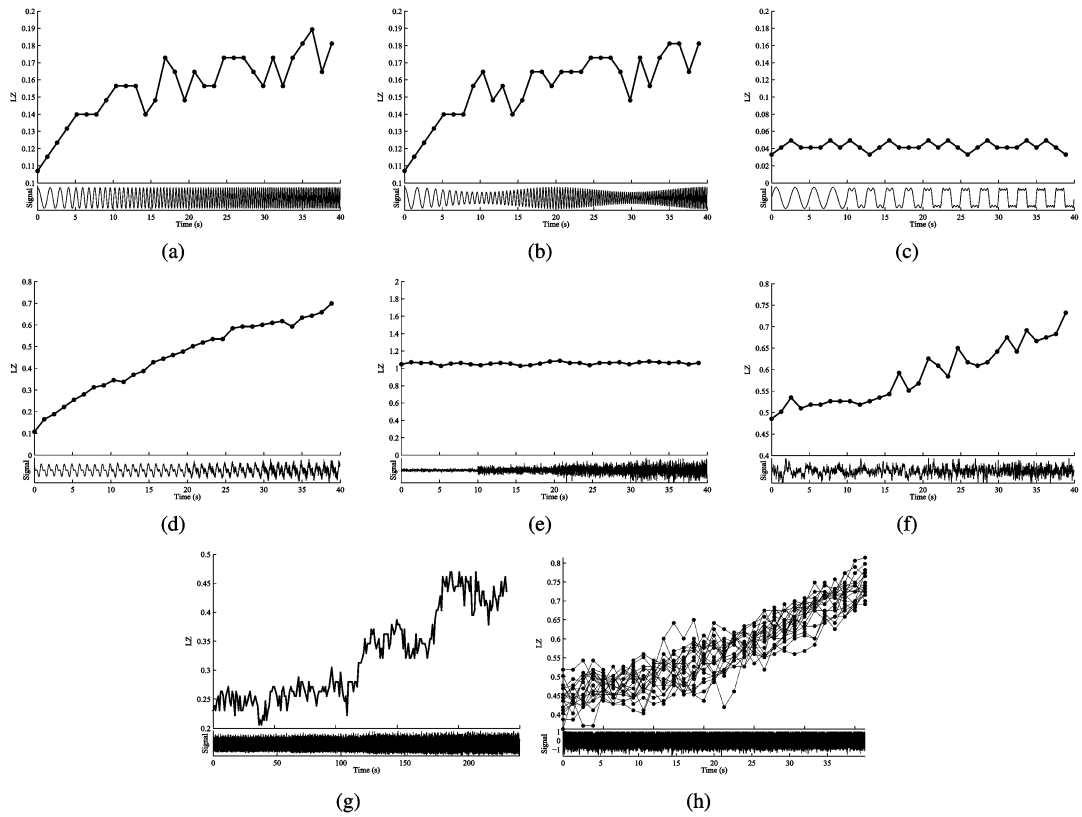

Fig. 2. Results of the tests performed to gain better understanding *LZ* and its interpretation. (a) Relationship between *LZ* and signal frequency, (b) *LZ* and amplitude modulation, (c) *LZ* versus number of harmonics, (d) *LZ* versus SNR, (e) *LZ* versus noise power, (f) *LZ* versus noise bandwidth, (g) *LZ* versus stochastic variability of signal harmonics, and (h) *LZ* versus Noise bandwidth variability.

Fig. 2(c) shows the results of Test 2. This test was designed to determine the relationship of *LZ* and the frequency content of periodic signals. Based on the results shown in Fig. 2(c) we conclude that *LZ* is not sensitive to the number of harmonics in periodic signals. Additionally, the *LZ* complexity in periodic signals with constant frequency is considerably less than in sinusoidal signals with variable frequency such as chirp signals where the frequency increases linearly with time as in Test 1 (e.g., $LZ = 0.04$ in periodic versus $LZ = 0.11 - 0.19$ in the chip signal we tested).

The results of Test 3 show that *LZ* is sensitive to changes in noise power present in quasi-periodic signals. *LZ* increases as the power of the noise increases. However, the results of Test 4 show that given the same underlying signal structure (white Gaussian noise) the *LZ* is not affected by the power of the noise as shown in Fig. 2(e). This indicates that there is a dependance between *LZ* and the noise power at high signal-to-noise ratio (SNR). *LZ* increases as the power of the noise increases until it reaches saturation, as it can be seen in Fig. 2(e). Note that in the case of white Gaussian noise, the *LZ* complexity measure is approximately equal to 1, independently of the noise power, which suggest that *LZ* complexity is bounded between 0 and 1 (approximately). This is a significant advantage of the *LZ* metric over other measures of complexity.

Fig. 2(f) shows the results of Test 5. This test consisted of determining the relationship between *LZ* and the noise bandwidth. The synthetic signal was composed of four segments of colored noise with increasing spectral bandwidth, as it can be seen in the spectrogram shown in Fig. 1(g). The results of this test demonstrate that *LZ* in sensitive to signal bandwidth changes. *LZ* increases as the signal bandwidth increases. This is significant in the context of biosignal analysis, since there are numerous applications where we are interested in estimating the bandwidth of a physiologic or residual signal. The *LZ* complexity measure may be used for this purpose.

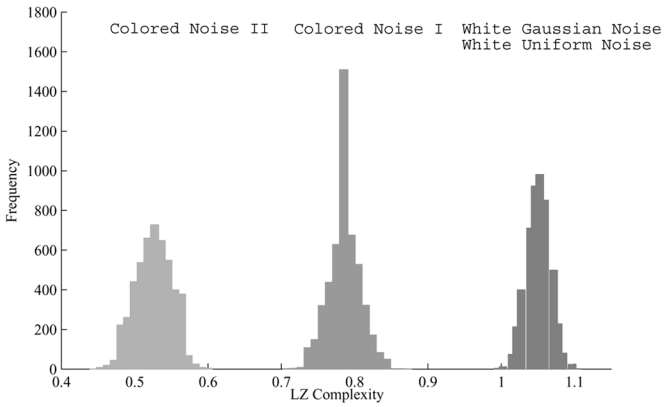

Fig. 3. Results of the test to study the standard error of the *LZ* as a function of the signal bandwidth. The standard error of the *LZ* estimates decreases as the bandwidth of the process increases.

In Test 6, we studied how the variability of signal harmonics affect the *LZ* complexity measure. The *LZ* metric was applied to synthetic signals generated using a harmonic model where the

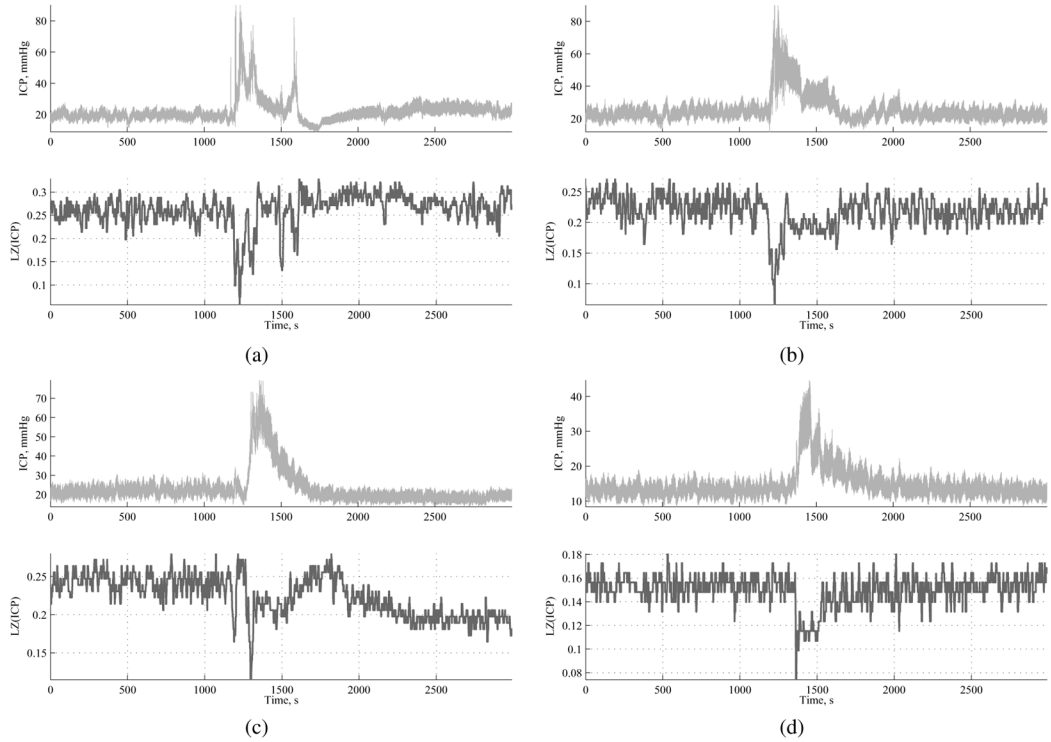

Fig. 4. Results showing the relation between the *LZ* applied to ICP versus the ICP signal. Note that *LZ* decreases during periods of ICH.

variability of the signal harmonics was increased, that is, the bandwidth of the process used to generate the signal harmonics was increased as a function of time. In this case, the signal was generated by concatenating four signals with different stochastic variability of the harmonics. The spectrogram shown in Fig. 1(h) shows how the variability of the harmonics increases as a function of time. Fig. 2(g) shows the results of this test. The test results demonstrate that *LZ* is sensitive to changes in the variability of the signal harmonics. It is also remarkable the fact that using the *LZ* metric we can determine the four abrupt changes in the stochastic variability of the signal harmonics, since these are not even detectable by visual examination of the spectrogram. The result of this test is particularly important in the context of biomedical signal processing and analysis, since many physiologic signals can be accurately modeled as a sum of sinusoids where the fundamental frequency is represented as an AR process (e.g., ABP, ICP, SpO2, and ECG). This result indicates that we can apply *LZ* complexity directly to a physiologic signal without any preprocessing and detect variability changes in the signal harmonics. For instance, if we analyze a pressure signal such as ABP or ICP by directly applying the *LZ* complexity measure to the time-series, we can detect heart rate variability (HRV) changes associated with specific patient populations or specific conditions, since the harmonics that compose pressure signals correspond to the cardiac components. More accurate HRV analysis can be performed by applying a bandpass filter centered at the heart rate frequency to the pressure signal, and analyzing the filtered signal using the *LZ* measure of complexity. This results also suggest that *LZ* complexity may be used to perform HRV analysis directly on ECG signals without the need to perform QRS detection (i.e., without the interbeat interval time-series). Thus, studies involving the application of *LZ* complexity to pressure signals or ECG signals should be interpreted in connection to HRV.

In Test 7, we studied the variability of the *LZ* estimates by creating replicas of the signal used in Test 4. Fig. 2(h) shows the variability of the *LZ* estimates as a function of time. Fig. 3 shows histograms for four signals with different bandwidths. These histograms can be used to estimate the standard error of the *LZ* estimates as a function of the signal bandwidth. Note that as the signal bandwidth increases the standard error of the *LZ* metric decreases. This result also demonstrates that the *LZ* complexity is the same for stochastic signals with different distributions (i.e., Gaussian versus Uniform) as long as they have the same bandwidth. Note that for processes with full bandwidth such as Gaussian white noise, or uniform white noise, the *LZ* measure of complexity is approximately equal to 1. The *LZ* corresponding to colored noise with bandwidth $B_1 = (f_s/4)$ is approximately 0.8, and the *LZ* corresponding to colored noise with bandwidth $B_1 = (f_s/8)$ is approximately 0.5. It is surprising that *LZ* can be used to characterize the bandwidth of a stochastic process, specifically considering that *LZ* is applied to a binary sequence generated from the original signal using a simple threshold.

The results of our study show that *LZ* is a metric that quantifies primarily the signal bandwidth and bandwidth of the signal harmonics that compose quasi-periodic signals. Consequently, *LZ* is correlated with second-order statistical metrics such as autocorrelation and power-spectral density. The fact that *LZ* quantifies primarily the signal bandwidth and bandwidth of the signal harmonics is relevant in biomedical signal analysis, since physiologic signals can often be modeled either as a quasi-periodic

signal, a colored noise signal, or a combination of the two. In the context of biosignal analysis, *LZ* measures derived from quasi-periodic physiologic signals should be interpreted as a harmonic variability metric. For instance, in the case of pressure signals such as ABP, ICP, and SpO2, *LZ* is affected primarily by the heart rate and the respiratory variability.

Fig. 4 shows four examples where *LZ* was calculated on ICP signals using a moving window of 10 s with 90% overlap. The ICP signal was highpass filtered with a cutoff frequency of 0.5 Hz to eliminate the mean (i.e., trend) prior to the computation of *LZ*. Note the *LZ* of ICP decreased as subjects progressed from a stable state of normal ICP to a state of acutely elevated ICP. This result provides further evidence to indicate that decreased complexity of ICP coincides with episodes of ICH in TBI. These results are consistent with the results of our previously published study involving the analysis of ICP during ICH using $ApEn$ [40], [42]. Based on our simulation studies involving *LZ*, in the case of pressure signals such as ICP, ABP, and SpO2, the *LZ* measure of complexity can be interpreted as a metric that quantifies the bandwidth of the signal harmonics. Thus, decreases in *LZ* complexity during ICH correspond to a decrease in the stochastic variability of the cardiac component, or a decrease in HRV. Additionally, based on the simulation study we conclude that values of *LZ* in pressure signals typically range from 0.05 to 0.3 which is consistent with the results obtained in this example. A *LZ* value close to 1 corresponds to a signal with full bandwidth (i.e., an uncorrelated stochastic process). Values of *LZ* around 0.05 correspond to periodic signals, and *LZ* values around 0.2 correspond to quasi-periodic signals with variable harmonics such as pressure signals.

V. Conclusion

In this paper, we studied the *LZ* measure of complexity and its interpretability in terms of classical signal processing concepts such as frequency, number of harmonics, frequency variability of signal harmonics, and signal bandwidth. Our results indicate that *LZ* is particularly useful as a scalar metric to estimate the bandwidth of random processes. Additionally, the *LZ* metric can be used to characterize the bandwidth of the harmonics in quasi-periodic signals. In the context of biosignal analysis, this result indicates that *LZ* measures derived from quasi-periodic physiologic signals should be interpreted as a harmonic variability metric.

This type of study is critical for the proper use and correct interpretation of the *LZ* metric in the context of biomedical signal analysis. Similar studies must also be carried out and published to aid the interpretability of some of the other nonlinear metrics currently used for biosignal analysis.

References

[1] A. Lempel, M. Cohn, and W. Eastman, "A class of balanced binary sequences with optimal autocorrelation properties," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 38–42, 1977.

[2] A. Lempel and J. Ziv, "On the complexity of finite sequences," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 1, pp. 75–81, 1976.

[3] A. Lempel, G. Seroussi, and J. Ziv, "On the power of straight-line computations in finite fields," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 6, pp. 875–880, 1982.

[4] J. Ziv and N. Merhav, "Estimating the number of states of a finite-state source," *IEEE Trans. Inf. Theory*, vol. 38, no. 1, pp. 61–65, 1992.

[5] M. Feder, N. Merhav, and M. Gutman, "Universal prediction of individual sequences," *IEEE Trans. Inf. Theory*, vol. 38, no. 4, pp. 1258–1270, Jul. 1992.

[6] E. Plotnik, M. Weinberger, and J. Ziv, "Upper bounds on the probability of sequences emitted by finite-state sources and on the redundancy of the Lempel-Ziv algorithm," *IEEE Trans. Inf. Theory*, vol. 38, no. 1, pp. 66–72, Jan. 1992.

[7] T. Kawabata, "Exact analysis of the Lempel-Ziv algorithm for i.i.d. sources," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 698–702, Mar. 1993.

[8] J. Ziv and J. Z. Merhav, "A measure of relative entropy between individual sequences with application to universal classification," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1270–1279, Jul. 1993.

[9] S. Savari, "Redundancy of the Lempel-Ziv incremental parsing rule," *IEEE Trans. Inf. Theory*, vol. 43, no. 1, pp. 9–21, Jan. 1997.

[10] P. Grassberger, "Estimating the information content of symbol sequences and efficient codes," *IEEE Trans. Inf. Theory*, vol. 35, no. 3, pp. 669–675, May 1989.

[11] A. Lapidoth and J. Ziv, "On the universality of the LZ-based decoding algorithm," *IEEE Trans. Inf. Theory*, vol. 44, no. 5, pp. 1746–1755, Sep. 1998.

[12] G. Louchard and W. Szpankowski, "On the average redundancy rate of the Lempel-Ziv code," *IEEE Trans. Inf. Theory*, vol. 43, no. 1, pp. 2–8, Jan. 1997.

[13] J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 5, pp. 530–536, Sep. 1978.

[14] A. Lempel and J. Ziv, "Compression of two-dimensional data," *IEEE Trans. Inf. Theory*, vol. 32, no. IT-1, pp. 2–8, Jan. 1986.

[15] A. Wyner and J. Ziv, "Some asymptotic properties of the entropy of a stationary ergodic data source with applications to data compression," *IEEE Trans. Inf. Theory*, vol. 35, no. 6, pp. 1250–1258, Nov. 1989.

[16] T. Luczak and W. Szpankowski, "A suboptimal lossy data compression based on approximate pattern matching," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1439–1451, Sep. 1997.

[17] D. Ornstein and B. Weiss, "Entropy and data compression schemes," *IEEE Trans. Inf. Theory*, vol. 39, no. 1, pp. 78–83, Jan. 1993.

[18] A. Gavish and A. Lempel, "Match-length functions for data compression," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1375–1380, Sep. 1996.

[19] E.-H. Yang and J. Kieffer, "Simple universal lossy data compression schemes derived from the Lempel-Ziv algorithm," *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 239–245, Jan. 1996.

[20] A. Lempel and M. Cohn, "Design of universal test sequences for VLSI," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 1, pp. 10–17, Jan. 1985.

[21] K. Visweswariah, S. Kulkarni, and S. Verdu, "Source codes as random number generators," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 462–471, 1998.

[22] T. Etzion and A. Lempel, "Algorithms for the generation of full-length shift-register sequences," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 3, pp. 480–484, May 1984.

[23] Y. L. Orlov and V. N. Potapov, "Complexity: An internet resource for analysis of DNA sequence complexity," in *Nucleic Acids Res.*, Jul. 2004, vol. 32, pp. W628–W633.

[24] V. D. Gusev, L. A. Nemytikova, and N. A. Chuzhanova, "On the complexity measures of genetic sequences," *Bioinformatics*, vol. 15, no. 12, pp. 994–999, Dec. 1999.

[25] L. Stern, L. Allison, R. L. Coppel, and T. I. Dix, "Discovering patterns in plasmodium falciparum genomic DNA," *Mol. Biochem. Parasitol.*, vol. 118, no. 2, pp. 175–186, Dec. 2001.

[26] J. Szczepaski, J. M. Amigó, E. Wajnryb, and M. V. Sanchez-Vives, "Application of Lempel-Ziv complexity to the analysis of neural discharges," *Network*, vol. 14, no. 2, pp. 335–350, May 2003.

[27] J. M. Amigó, J. Szczepaski, E. Wajnryb, and M. V. Sanchez-Vives, "Estimating the entropy rate of spike trains via Lempel-Ziv complexity," *Neural Comput*, vol. 16, no. 4, pp. 717–736, Apr. 2004.

[28] H. H. Otu and K. Sayood, "A new sequence distance measure for phylogenetic tree construction," *Bioinformatics*, vol. 19, no. 16, pp. 2122–2130, Nov. 2003.

[29] X. Wu and J. Xu, "Complexity and brain function," *Acta Biophysica Sinica*, vol. 7, pp. 103–106, 1991.

[30] J. Xu, Z. Liu, R. Liu, and Q. F. Yang, "Information transformation in human cerebral cortex," *Physica D*, vol. 106, pp. 363–374, 1997.

[31] 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," *Med. Eng. Phys.*, vol. 28, no. 4, pp. 315–322, May 2006.
[32] N. Radhakrishnan and B. Gangadhar, "Estimating regularity in epileptic seizure time-series data," *IEEE Eng. Med. Biol. Mag.*, vol. 17, no. 3, pp. 89–94, May-Jun. 1998.
[33] X. S. Zhang, Y. S. Zhu, and X. J. Zhang, "New approach to studies on ECG dynamics: extraction and analyses of QRS complex irregularity time series," *Med. Biol. Eng. Comput.*, vol. 35, no. 5, pp. 467–473, Sep. 1997.
[34] X. S. Zhang, Y. S. Zhu, N. V. Thakor, and Z. Z. Wang, "Detecting ventricular tachycardia and fibrillation by complexity measure," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 5, pp. 548–555, May 1999.
[35] H. X. Zhang, Y. S. Zhu, and Z. M. Wang, "Complexity measure and complexity rate information based detection of ventricular tachycardia and fibrillation," *Med. Biol. Eng. Comput.*, vol. 38, no. 5, pp. 553–557, Sep. 2000.
[36] X. S. Zhang and R. J. Roy, "Predicting movement during anaesthesia by complexity analysis of electroencephalograms," *Med. Biol. Eng. Comput.*, vol. 37, no. 3, pp. 327–334, May 1999.
[37] X. S. Zhang, R. J. Roy, and E. W. Jensen, "EEG complexity as a measure of depth of anesthesia for patients," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 12, pp. 1424–1433, Dec. 2001.
[38] X. S. Zhang and R. J. Roy, "Derived fuzzy knowledge model for estimating the depth of anesthesia," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 3, pp. 312–323, Mar. 2001.
[39] R. Nagarajan, "Quantifying physiological data with Lempel-Ziv complexity-certain issues," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 11, pp. 1371–1373, Nov. 2002.
[40] R. Hornero, M. Aboy, D. Abásolo, J. McNames, and B. Goldstein, "Interpretation of approximate entropy: Analysis of intracranial pressure approximate entropy during acute intracranial hypertension," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 10, pp. 1671–1680, Oct. 2005.
[41] M. Aboy, O. W. Márquez, J. McNames, R. Hornero, T. Trong, and B. Goldstein, "Adaptive modeling and spectral estimation of nonstationary biomedical signals based on Kalman filtering," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 8, pp. 1485–1489, Aug. 2005.
[42] R. Hornero, M. Aboy, D. Abasolo, J. McNames, W. Wakeland, and B. Goldstein, "Complex analysis of intracranial hypertension using approximate entropy," *Crit. Care Med.*, vol. 34, no. 1, pp. 87–95, Jan. 2006.

Mateo Aboy (M'98) was born in Pontevedra, Spain, in 1980. He received the B.S. degree (high honors) in electrical engineering and the B.S. degree (high honors) in physics from Portland State University (PSU), Portland, OR, in 2002. In 2004, he received the M.S. degree in electrical and computer engineering from PSU and the M.Phil (DEA) degree from ETSIT-Vigo, where he also received the Ph.D. degree from the Signal Theory and Communications Department.

Since September 2000, he has been a Research Member of the Biomedical Signal Processing Laboratory, PSU. He has been an Assistant Professor with the Electronics Engineering Technology Department at Oregon Institute of Technology, since 2005. His primary research interest is statistical signal processing.

Dr. Aboy is a full-member of The Scientific Research Society (Sigma XI), a lifetime honorary member of the Golden-Key Honor Society, a past Chapter President of HKN (International Electrical Engineering Honor Society), and past Corresponding Secretary of TBP (National Engineering Honor Society).

Roberto Hornero (M'04) was born in Plasencia, Spain, in 1972. He received the Engineer degree in telecommunication engineering and the Ph.D. degree from the University of Valladolid, Valladolid, Spain, in 1995 and 1998, respectively.

He is currently "Profesor Titular" in the Department of Signal Theory and Communications at the University of Valladolid. His main research interest is nonlinear analysis of biomedical signals to help physicians in the clinical diagnosis. He founded the Biomedical Engineering Group in 2004. The research interests of this group are connected with the field of nonlinear dynamics, chaotic theory, and wavelet transforms with applications in biomedical signal and image processing.

Dr. Hornero is a member of the Spanish Biomedical Engineering Society (SEIB).

Daniel Abásolo (M'04) was born in Valladolid, Spain, in 1976. He received the engineer degree in telecommunication engineering and the Ph.D. degree from the University of Valladolid in 2001 and 2006, respectively.

He is currently a Lecturer in the Department of Signal Theory and Communications at the University of Valladolid. His main research interest is nonlinear biomedical signal processing.

Dr. Abásolo is a member of the Biomedical Engineering Group and of the Spanish Biomedical Engineering Society (SEIB).

Daniel Álvarez was born in Bembibre, Spain, in 1978. He received the Engineer degree in telecommunications engineering from the University of Valladolid, Valladolid, Spain, in 2005. Currently, he is a scholarship holder in the Department of Signal Theory and Communications at the University of Valladolid, where he is working towards the Ph.D. degree.

He is a research member of the Biomedical Engineering Group. His research has been mainly focused in nonlinear processing of biomedical signals from overnight sleep studies.