DETRENDED FLUCTUATION ANALYSIS FOR EMPIRICAL MODE DECOMPOSITION BASED DENOISING

Ahmet Mert *

Dept. of Mechanical Engineering
Piri Reis University
34940, Tuzla, Istanbul, Turkey
amert@pirireis.edu.tr

Aydin Akan

Dept. of Electrical and Electronics Eng.
Istanbul University
34320 Avcilar, Istanbul, Turkey
akan@istanbul.edu.tr

ABSTRACT

Empirical mode decomposition (EMD) is a recently proposed method to analyze non-linear and non-stationary time series by decomposing them into intrinsic mode functions (IMFs). One of the most popular application of such a method is noise elimination. EMD based denoising methods require a robust threshold to determine which IMFs are noise related components. In this study, detrended fluctuation analysis (DFA) is suggested to obtain such a threshold. The scaling exponential obtained by the root mean squared fluctuation is capable of distinguishing uncorrelated white Gaussian noise and anti-correlated signals. Therefore, in our method the slope of the scaling exponent is used as the threshold for EMD based denoising. IMFs with lower slope than the threshold are assumed to be noisy oscillations and excluded in the reconstruction step. The proposed method is tested on various signal to noise ratios (SNR) to show its denoising performance and reliability compared to several other methods.

***Index Terms*—** Empirical mode decomposition, Detrended fluctuation analysis, Denoising, Thresholding

1. INTRODUCTION

A common approach in denoising is decomposing the noisy signal into its components and applying a thresholding scheme [1]. Wavelet transform based thresholding is a well-known and popular denoising method where the performance depends on the thresholding approaches and estimators [2, 3], as well as the trial and error method of the selection of orthogonal basis.

The empirical mode decomposition (EMD) is an alternative method to analyze non-linear and non-stationary signals [4]. After applying EMD, a finite number of amplitude and frequency modulated (AM/FM) zero-mean oscillations called intrinsic mode functions (IMFs) are obtained. IMFs are extracted as the signal dependent semi-orthogonal basis functions via an iterative algorithm called sifting. On the other hand, it is a challenging study to explain the meaning of each IMF, or determine which IMF refers to noise, that is the main obstacle in EMD based denoising.

Detrended fluctuation analysis (DFA) [5] can be considered as an enhanced version of Hurst exponent to analyze non-stationary time series [6, 7]. Slope of DFA, α more clearly describes the statistical properties of a process than the Hurst exponent. The special cases $\alpha = 0.5$, $\alpha = 1$ and $\alpha = 1.5$ correspond to completely uncorrelated white noise, pink noise, and Brownian noise, respectively. Moreover, the slope α can also be considered as an indicator of roughness [8]: the larger value, the smoother time series or slower fluctuations. Here, we present a method where we use α as a quantitative measure to select the noisy IMFs.

Denoising performance of any thresholding based method using wavelet transform or EMD depends on the used threshold estimator. The most popular threshold for wavelet denoising $T = \hat{\sigma}\sqrt{2\ln N}$ with the robust variance estimator, $\hat{\sigma} = median\,(\mid c_i \mid : \, i = 1, \ldots, N)\,/0.6745$ (c_i states the wavelet coefficients of the noisy signal in the ortogonal basis) [2, 3], it is clear that the value of threshold is determined by the distribution of coefficients belonging to both noise free signal and the noise. Therefore, threshold value depends on the SNR and the properties of the noise, which may cause inaccurate threshold estimation. For EMD based applications, white Gaussian noise (WGn) and fractional Gaussian noise (fGn) referenced models [9, 10] or information theoretical methods including mutual information [11, 12] are applied to analyze the discrepancy of the scores in order to determine the irrelevant or noisy IMFs. However, these are all reference and comparison based approaches. In this study, we propose a DFA thresholding for EMD based denoising as a robust metric to determine IMFs of the noise. The excluded oscillations are detected and accepted as noise by comparing their statistical properties instead of comparing with the others. Noisy IMF detection performance of the algorithm is tested on synthetic and real signals at low SNR levels, compared to wavelet based soft and hard thresholding. While the mentioned methods above are based on the comparison of referenced or mod-

*This work was partially supported by The Research Fund of The University of Istanbul. Project number:35830.

eled systems, our approach uses DFA score without comparing with other results which is the main advantage of the proposed DFA-EMD based denoising.

2. THE EMD ALGORITHM

The EMD has been introduced as a tool of a data driven decomposition method for multicomponent signal so that sum of the IMFs is equal to the original signal [4] . IMFs satisfy two criteria: First, the number of the extrema and the number of zero crossings must be equal or must differ by one at most. Second, the mean of the envelopes determined by the local maxima and minima called upper and lower envelope should be zero. Therefore, instantaneous frequency (IF) fluctuations of IMFs can be decreased [13]. The most important process of the EMD algorithm to extract IMFs is called *Sifting*, which is composed of the following steps [14]:

(i) Find local maxima,M_i, $i = 1, 2, \ldots,$ and minima m_k ,$k = 1, 2, \ldots,$ in $x(n)$.

(ii) Compute the interpolating signals $M(n) = f_M(M_i, n)$, and $m(n) = f_m(m_k, n)$ using cubic spline, which are the upper and lower envelopes of the signal.

(iii) Compute mean of the envelopes, $e(n) = [M(n) + m(n)]/2$.

(iv) If $e(n)$ satisfies the IMF requirements, keep it as an IMF, and substract $e(n)$ from the signal; $x(n) = x(n) - e(n)$.

(v) Return to step (i) and stop, after $x(n)$ remains nearly unchanged.

(vi) After obtaining an IMF,$\varphi_i(n)$, substract IMF from the signal $x(n) = x(n) - \varphi_i(n)$ and return to (i) if $x(n)$ is not constant or trend, $r(n)$.

Hence, the original signal can be reconstructed by the sum of IMFs described by $x(n) = \sum_{i=1}^{L} \varphi_i(n) + r(n)$, where L is the number of IMFs. EMD has been an alternative approach for signal analysis such as instantaneous frequency, autoregressive parameter estimation, classification, audio coding [15–18], and denoising which is the most popular application of EMD.

3. DETRENDED FLUCTUATION ANALYSIS

Hurst exponent [19] is defined as the index of long-range dependence, mild or wild randomness. In case of non-stationarities, it is not a suitable method which causes spurious score [5]. Hence, DFA is a recently proposed method to obtain more reliable scaling exponent for signals having non-stationary properties especially different trends with unknown duration. The DFA score is also adopted in a similar way with the Hurst exponent in log-log scale. The basic principle of the DFA is to compute how the average root mean square (RMS) fluctuation of the time series around the local trend in the box size varies as a function of the time scale n. The first step is to find integrated time series $y(k)$ after removing the mean by

$$y(k) = \sum_{i=1}^{k} [x(i) - \bar{x}], \; 1 \leq k \leq N \tag{1}$$

where $\bar{x}$ states the average of the time series in the range $[1, N]$. $y(k)$ is divided into n sample long segments called box size for DFA. For each box, the estimated local trend $y_n(k)$ is found by using least square linear fitting. Finally the RMS fluctuation $F(n)$ is computed by subtracting $y_n(k)$ from the integrated series $y(k)$ as:

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} [y(k) - y_n(k)]^2} \tag{2}$$

After calculating the $F(n)$ for various n, the slope, α of the curve $\log(F(n)) / \log(n)$ is called the scaling exponent which demonstrates power-law behaviour as $F(n) \propto n^{\alpha}$.

The reason that α gives a reliable metric is it yields $\alpha = 0.5$ for completely uncorrelated data such as white noise, $\alpha = 1$ for pink noise and $\alpha = 1.5$ for Brownian noise. When $0 < \alpha < 0.5$, the signal is called “anti-correlated” such that large fluctuations are followed by small ones and vice versa. When it is in the range between 0.5 and 1.0, temporal correlations are available. If $\alpha > 1$, the correlations do not exhibit power-law behaviour versus time. The slope is also considered an indicator of the roughness. The larger the value, the smoother the time series. In other words, small value indicates more rapid fluctuations [8]. However, the linear region for the log-log plot used to find α is another parameter of the DFA related applications. The selection of the box size n is signal dependent, but the range $4 \leq n \leq 16$ is the most popular and reliable linear region, which is successfully applied in biomedical signal processing such as electroencephalography (EEG), [8, 20] and electrocardiography (ECG) [21].

4. PROPOSED DFA THRESHOLDING FOR EMD BASED DENOSING

EMD based denoising methods focus on determining irrelevant and information free IMFs. After decomposition of a corrupted noisy signal $x(n)$, few IMFs may be the oscillations of the noise free original signal $\bar{x}(n)$, and the others correspond to noise $\eta(n)$. From this point of view, a reliable metric to determine noisy IMFs is the main step of a denoising algorithm.

Our proposed method is based on using the DFA slope, α as a threshold. The method is independent of any estimators depending on other components, is based on excluding the IMFs with lower α score than the threshold θ. The threshold is determined by the static score of the DFA slope $\alpha = 0.5$

for white noise. However, the EMD algorithm has a drawback called "mode-mixing". In other words, an IMF is not a mono-component, and it is mixed with the other oscillations in the signal. Therefore, the threshold is determined with 0.2 confidential interval. The detailed steps of the proposed method are described as:

a) Let $x(n) = \bar{x}(n) + \sigma\eta(n)$ be the observed noisy signal where $\bar{x}(n)$ and $\eta(n)$ state the noise free signal and the AWGN with unknown variance σ respectively.

b) Apply DFA to decomposed IMFs, $\varphi_i(n)$ of $x(n)$ to compute α_i of each IMF. $i = 1, 2, \ldots, L$.

c) Determine $\theta = 0.7$ as the sum of white noise slope 0.5 with 0.2 confidential interval.

d) Then reconstruct a denoised signal estimate using the IMFs with higher slope score α_i than the threshold as: $\tilde{x}(n) = \sum_j \varphi_j(n),\ j = \{i \mid \alpha_i > \theta\}$.

Therefore, irrelevant IMFs which contain mostly white noise can be determined and removed from the observed signal.

5. RESULTS

The proposed DFA thresholding for EMD based denosing (DFA-EMD) method is evaluated using signals with various SNRs. IMFs with DFA scores, and the denoised versions are given in the rest of the section, comparing the performance with wavelet denosing and another EMD-based denosing method given in [9].

1. The well-studied piecewise regular signal with 0 dB SNR additive white Gaussian noise, and 2048 samples, shown in Fig. 1 is applied to the proposed method.

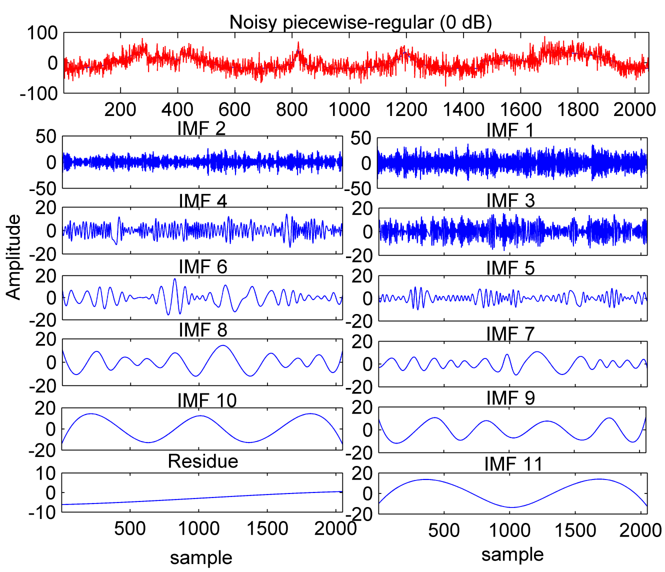


Fig. 1. The decomposition of piecewise regular signal with 0 dB SNR.

After obtaining DFA slopes of the IMFs in Fig. 1, the noisy components should have lower slope α than the threshold θ, and the results are shown in Fig. 4a. The IMFs with $\alpha_i \leq 0.7$ are determined as the oscillations of the AWGN. IMFs above the threshold are combined to estimate the denoised signal, $\tilde{x}(n) = \sum_{j=3}^{12} \varphi_j(n)$. We compare the proposed approach with wavelet thresholding, and the results are given in Fig. 2.

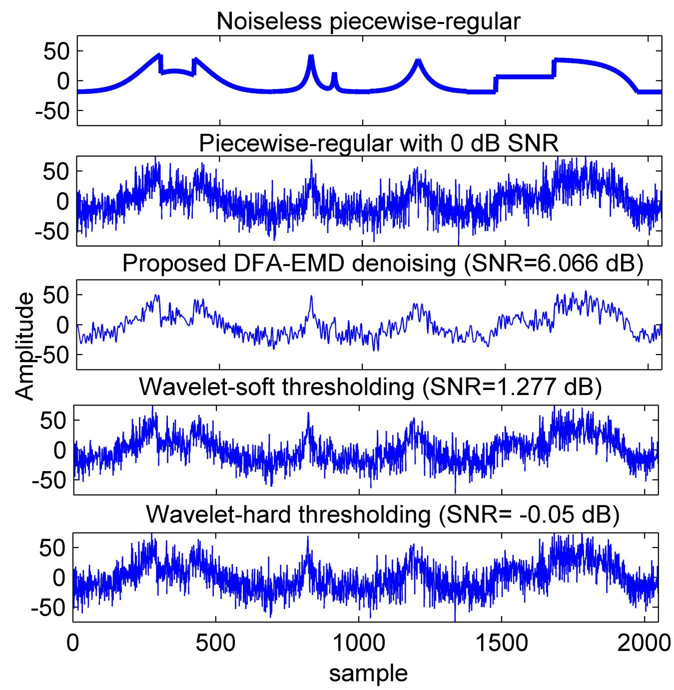


Fig. 2. Comparison of the proposed DFA-EMD based denoising at 0 dB SNR.

The proposed denoising method results in 6.066 dB SNR, while hard and soft threshold wavelet denosing with universal threshold estimator at level 3 by Symlet8 result in -0.05, and 1.277 dB, respectively. In our simulations, the mother wavelet Symlet8 at level 3 shows more denoising performance when compared to symlet3 and Symlet20 at level 7 and Daubechies2 at level 3 decomposition.

2. We applied the above procedure to the same signal with 20 dB SNR, and the decomposed IMFs are shown in Fig. 3.

As opposed to the results in Fig. 2 with 11 IMFs, 12 IMFs are obtained in Fig. 3. Since there are two IMFs lower than the threshold in the previous example, only IMF 1 has lower slope ($\alpha_1 = 0.282$), as such it is not included in the reconstruction. The slopes of the exponents of the IMFs are presented in Fig. 4b.

The DFA-EMD method can successfully denoise 20 dB piecewise-regular signal, and the SNR becomes 22.373 dB. However, the wavelet thresholding methods with resulting SNRs of 26.508 and 24.901 dB give better results, which is

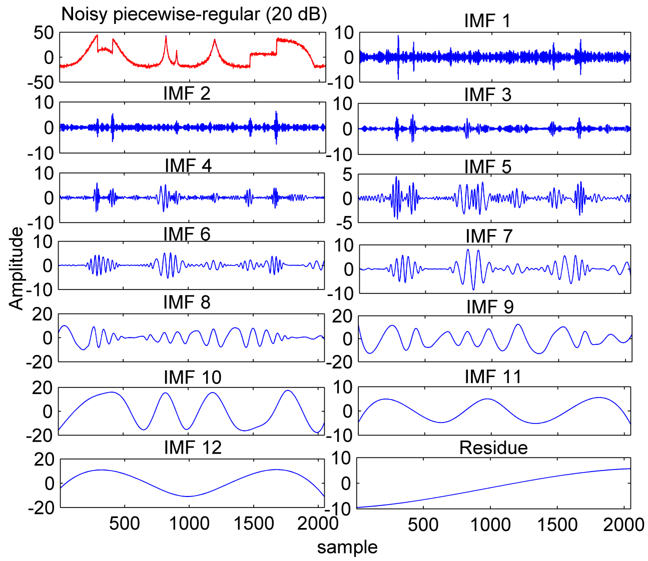
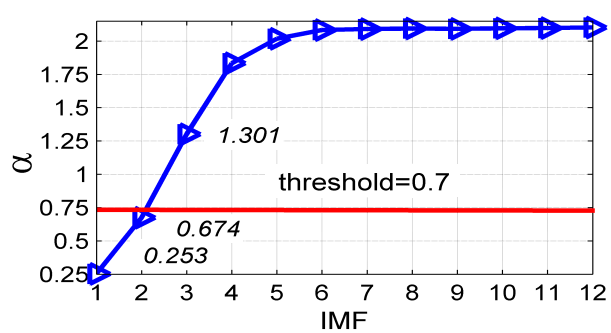
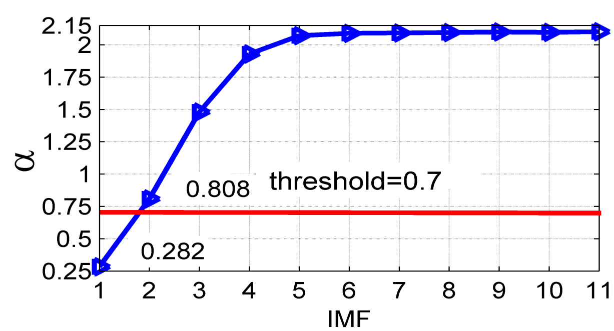


Fig. 3. Decomposition of piecewise-regular signal with 20 dB SNR.



(a) 0 dB



(b) 20 dB

Fig. 4. DFA slopes of the decomposed IMFs for piecewise-regular signal.

caused by the nature of any EMD algorithm namely mode-mixing problem. In other words, some components of white noise cannot be separated, and it is distributed to other IMFs.

3. The proposed method is also compared with a previous EMD-based denoising called noise-only model [9], and the

Table 1. The comparison of the DFA-EMD denoising with wavelet denoising at various SNRs

Signal	SNR (dB)	DFA-EMD (dB)	Soft(dB)	Hard(dB)
Piecewise-Regular	0	6.066	1.277	-0.050
	10	15.586	14.120	10.215
	20	22.373	26.508	24.901
Epileptic EEG	0	5.873	-0.043	-0.061
	10	14.706	9.991	9.938
	20	22.259	20.083	19.939
Normal EEG	0	5.940	0.111	-0.062
	10	15.484	10.354	9.938
	20	22.710	21.231	19.955
Bumps	0	5.900	7.522	7.521
	10	12.941	11.825	11.825
	20	19.412	12.631	12.631

results are given in Fig. 5. We observe that the fifth and the

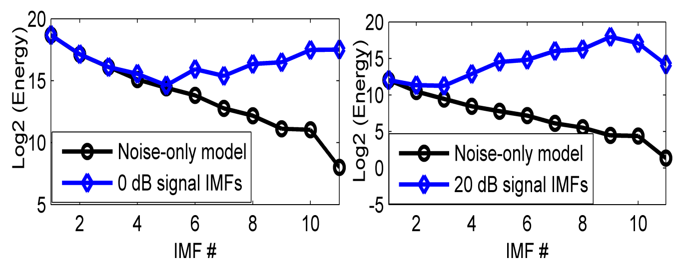


Fig. 5. The WGn model [9] to identify noise free IMFs.

second IMF diverge from the theoretical noise-only model. Thus, $\sum_{i=1}^{5}\varphi_i$ for 0 dB and $\sum_{i=1}^{2}\varphi_i$ for 20 dB are assumed as WGn, and then the SNRs of the EMD denoising employing the model above result in 5.46 dB and 22.320 dB SNR for piecewise-regular signal. However, our proposed DFA-EMD has denoising performance of 6.066 dB and 22.273 dB for piecewise-regular signal.

4. The proposed method is also tested on EEG signals, and Bumps signal to illustrate its performance. The results are given in Table 1.

For piecewise-regular signal denoising, the proposed EMD based method performs better than wavelet denoising, when the SNRs are 0 and 10 dB. However, in the case of 20 dB, it performs poorer than the wavelet methods. In contrast, the resulting SNRs after denoising epileptic and normal EEG signals have different trend; the DFA-EMD is more successful for 0, 10 and 20 dB SNR, when compared to wavelet soft and hard thresholding. Briefly, EMD-based method suppresses noise for lower SNR, better than wavelet methods.

6. CONCLUSION

We propose a detrended fluctuation analysis (DFA) thresholding for empirical mode decomposition (EMD) based (DFA-EMD) denoising. A few intrinsic mode functions (IMFs) of a noisy signal may be components of the noiseless signal, and the others may belong to the noise. DFA is deployed as a metric to determine which IMFs are noisy oscillations and to be excluded in the reconstruction step. The proposed DFA-EMD based approach is independent of the SNR and signal properties, which makes it advantageous over other threshold estimators in wavelet or EMD denoising. Computer simulations show that the denoising performance of the proposed DFA-EMD method is better than wavelet soft and hard-thresholding and a previous EMD-based denoising approach in [9] at especially low SNR values.

REFERENCES

[1] A. Chkeir, C. Marque, J. Terrien, and B. Karlsson. De-noising electrohysterogram via empirical mode decomposition. In *Proc. of PSSNIP Biosig. and Biorobot. Conf.*, pages 32–35, 2010.

[2] D.L. Donoho and I.M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrica*, 81:525–455, 1994.

[3] D.L. Donoho. De-noising by soft-thresholding. *IEEE Trans. Inform. Theory*, 41:613–627, 1995.

[4] N.E. Huang, Z. Shen, and S.R. Long. The empirical mode decomposition and the hilbert spectrum for non-linear and non-stationary time series analysis. *Proceeding of Royal Society A*, 454:903–995, 1998.

[5] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, and A.L. Goldberger. Mosaic organization of dna nucleotides. *Phys. Rev. E*, 49:1685–1689, 1994.

[6] D. Horvatic, H.E. Stanley, and B. Podobnik. Detrended cross-correlation analysis for non-stationary time series with periodic trends. *Europhys. Lett.*, 94:1–7, 2011.

[7] J.-M. Lee, D.-J. Kim, I.-Y. Kim, K.-S. Park, and S.I. Kim. Detrended fluctuation analysis of eeg in sleep apnea using mit/bih polysomnograpghy data. *Comput. Biol. Med.*, 32:37–47, 2002.

[8] S. Leistedt, M. Dumont, J.-P. Lanquart, F. Jurysta, and P. Linkowski. Characterization of the sleep eeg in acutely depressed men using detrended fluctuation analysis. *Clin. Neurophysiol.*, 118:940–950, 2007.

[9] Z. Wu and N.E. Huang. A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. Roy. Soc. London. A*, 460:1597–1611, 2004.

[10] P. Flandrin, G. Rilling, and P. Goncalves. Empirical mode decomposition as a filter bank. *IEEE Signal Process. Lett.*, 11:112–114, 2004.

[11] A. Ayenu-Prah and N. Attoh-Okine. A criterion for selecting relevant intrinsic mode functions in empirical mode decomposition. *Adv. in Adapt. Data Analy.*, 2:1–24, 2010.

[12] C.Y. Tseng and H.C. Lee. Entropic interpretation of empirical mode decomposition and its applications in signal processing. *Adv. in Adapt. Data Analy.*, 2:429–449, 2010.

[13] G. Rilling and P. Flandrin. One or two frequencies: The empirical mode decomposition answers. *IEEE Trans. Signal Proces*, 56:85–95, 2008.

[14] R.T. Rato, M.D. Ortigueira, and A.G. Batista. On the hht, its problems, and some solutions. *Mech. Syst. Signal Pr.*, 22:1374–1384, 2008.

[15] R. Gupta, A. Kumar, and R. Bahl. Estimation of instantaneous frequencies using iterative empirical mode decomposition. *Signal Image Video Proces*, 2012.

[16] Md.K. Hasan, M.S. Apu, and Md.K.I. Molla. A robust method for parameter estimation of ar systems using empirical mode decomposition. *Signal Image Video Proces.*, 2009.

[17] L. Ai, J. Wang, and R. Yao. Classification of parkinsonian and essential tremo using empirical mode decomposition and support vector meachines. *Digit. Signal Process.*, 21:543–550, 2011.

[18] K. Khaldi, A.-O. Boudraa, B. Torresani, and T. Chonavel. Hht- based audio coding. *Signal Image Video Proces.*, 2013.

[19] H .E. Hurst. Long-term storage of reservoirs. *T. Am Soc Civ Eng*, 116:770–808, 1951.

[20] M. Jospin, P. Caminal, E.W. Jensen, H. Litvan, M. Vallverdu, M.M.R.F. Struys, H.E.M. Vereeckle, and D.T. Kaplan. Detrended fluctation analysis of eeg as a measure of depth of anesthesia. *IEEE Trans. Biomed. Eng.*, 54:840–846, 2007.

[21] F.S Leitet, J. L. A. Carvalho, and A. F. da Rocha. Matlab software for detrended fluctuation analysis of heart rate variability. In *Proc. of International Conf. on Bio-inspired Sys. and Signal Proc.*, pages 225–229, 2010.