Filtering-based Analysis Comparing the DFA with the CDFA for Wide Sense Stationary Processes

Bastien Berthelot*†, Éric Grivel‡, Pierrick Legrand†,
Jean-Marc André‡, Patrick Mazoyer* and Thierry Ferreira*

* THALES AVS France, Campus THALES Bordeaux/THALES Toulouse
† Bordeaux University - IMB UMR CNRS 5251 - INRIA, FRANCE
‡ Bordeaux University - INP Bordeaux - IMS - UMR CNRS 5218, FRANCE

***Abstract*—The detrended fluctuation analysis (DFA) is widely used to estimate the Hurst exponent. Although it can be outperformed by wavelet based approaches, it remains popular because it does not require a strong expertise in signal processing. Recently, some studies were dedicated to its theoretical analysis and its limits. More particularly, some authors focused on the so-called fluctuation function by searching a relation with an estimation of the normalized covariance function under some assumptions. This paper is complementary to these works. We first show that the square of the fluctuation function can be expressed in a similar matrix form for the DFA and the variant we propose, called Continuous-DFA (CDFA), where the global trend is constrained to be continuous. Then, using the above representation for wide-sense-stationary processes, the statistical mean of the square of the fluctuation function can be expressed from the correlation function of the signal and consequently from its power spectral density, without any approximation. The differences between both methods can be highlighted. It also confirms that they can be seen as *ad hoc* wavelet based techniques.**

***Index Terms*—filter, interpretation, Hurst, DFA, CDFA.**

I. Introduction

In many applications such as speech processing, autoregressive moving average (ARMA) processes are often used to model the data. In this case, the correlation function, r_τ, with τ the lag, decays exponentially to zero. As $\sum_\tau r_\tau$ is absolutely summable, these processes are short-memory. However, in statistics, econometrics and finance [5], the correlation function may decay slower than exponentially. In these applications, the ARMA model is no longer well-suited. Long-memory models must be considered: the autoregressive fractionally integrated moving average (ARFIMA) models [7] [9] can be used as well as the fractional Gaussian noise which is a kind of $1/f$ noise.

One of the earliest studies mentioning that time series may exhibit long-range dependence (LRD) is based on the Hurst exponent, denoted as H [17]. Thus, a process is said to have LRD if $0.5 < H < 1$ whereas $0 < H < 0.5$ corresponds to anti-persistent processes. For a Brownian noise, a pink noise and a white noise, H is equal to 0.5, 0 and -0.5 respectively. Two main families of approaches exist to estimate it:

1. Frequency-domain estimators can be used and aim at analyzing the power spectral density (PSD) of the time series [27]. This is for instance the case of the local Whittle method, the periodogram method, the wavelet-based method [1] and the semi-parametric method [4] [18]. More recently, authors have proposed solutions based on the empirical mode decomposition (EMD) [25] or the fractional Fourier transform [27]. Some comparative studies such as [6] have been also led.
2. Time-domain estimators can be used. They include the rescaled range analysis, the aggregated variance method, the absolute-value method and the variance-of-residuals method. The reader may refer to [28] for instance.

In 1992, Peng *et al.* suggested using the fluctuation analysis (FA) to estimate the Hurst exponent of a pure mono-fractal time series [21]. Then, the detrended fluctuation analysis (DFA) [22] has shown good performance. Its first step is to define the trend of the integrated signal. This latter consists of discontinuous local trends modeled by straight lines of length N. However, as recalled in [14], there are many other ways to obtain the global trend of a signal. This is the reason why several variants of the DFA exist. To name a few, the adaptive fractal analysis (AFA) [24] *a posteriori* corrects the discontinuities. Tarvainen's method [29] is based on a regularized least-squares (LS) criterion to obtain the local trends. Note that the trend extraction is similar to the so-called Hodrick-Prescott filtering, widely used in econometrics [8]. Finally, the detrended moving average (DMA) is based on a low-pass filtering of a signal in order to obtain the trend. In its standard version, the filter has a causal finite-impulse response of length N_{DMA} [2]. Some variants called centered moving average and weighted moving average of order l [31] have been proposed and are respectively based on a non-causal impulse response or an infinite-impulse response. All these methods provide the so-called scaling exponent, denoted α, which is related to the Hurst exponent, as explained below.

A great deal of interest has been paid to the DFA and the DMA, especially in the field of meteorology, stock market prediction, biomedical to analyze heart-rate variability [23], breathing pattern [19], voice pathology [3] and EEG analysis [26]. It should be noted that the practitioner using these methods also considers other nonlinear dynamical system analysis techniques [20] as well as the sample entropy and the multi-scale entropy to characterize the recorded signals or time series. Even if the DFA and the DMA can suffer some drawbacks and can be outperformed by other approaches based

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

on wavelets or the local Whittle method, they can be used by people who do not have advanced skills in signal processing and statistics because the methods are only based on regression and linear filtering. This probably explains their popularity and the fact that there is still an active research on them. Thus, fast versions have been developed [30]. Multifractal aspects [13] as well as theoretical studies on the DFA and DMA [10]–[12] [15] [16] have been done for the last years. More particularly, in [10], by assuming that the number of segments is large enough, by supposing that the signal is wide-sense stationary (w.s.s.) and ergodic and by making some approximations such as replacing infinite temporal summations by finite sums, Höll *et al.* aimed at expressing the square of the so-called fluctuation function as a function of the normalized covariance function of the signal. Kiyono *et al.* analyzed the single-frequency responses of the DFA and the higher-order DFA [15] as well as the centered DMA [16]. They concluded that for stochastic processes whose PSD is a function of the frequency f of the form $f^{-\lambda}$, the higher-order DFA is convenient to estimate α as long as $\alpha = \frac{\lambda+1}{2}$.

In this paper, we propose a matrix formulation of the square of the fluctuation function for both the DFA and its variant, called Continuous-DFA (CDFA), where the global trend is constrained to be continuous. Provided that the signal under study is w.s.s., its statistical mean is then expressed from the correlation function without any approximation. At this stage, both methods can be compared from a filtering point of view. Therefore, we can highlight the differences between them with respect to the selection of N.

The remainder of this paper is organized as follows: in section II, the main steps of the DFA and the CDFA are briefly described. In section III, a comparative analysis is done.

In the following, I_j is the identity matrix of size j. $\mathbb{1}_{j\times k}$ and $\mathbb{0}_{j\times k}$ are matrices of size $j \times k$ filled with ones and zeros respectively. $J_j = I_j - \frac{1}{j}\mathbb{1}_{j\times j}$, $diag([.], j)$ is a matrix whose j^{th} diagonal is equal to $[.]$. $diag(\mathbb{1}_{1\times N-1}, 1)$ is hence the square matrix of size N whose 1^{st} sub-diagonal above the main one has its elements equal to 1. Finally, $C_{j,k}$ is a matrix of size (j, M) so that $C_{j,k} = [\mathbb{0}_{j\times k} \quad I_j \quad \mathbb{0}_{j\times(M-(j+k))}]$.

II. DFA VS CDFA

Let us describe the main steps of the DFA and the CDFA when the M samples $\{y(m)\}_{m=1,...,M}$ of the signal are available.

A. *Computation of the profile*

The profile, *i.e.* the integrated signal, is computed as follows:

$$y_{int}(m) = \sum_{i=1}^{m}(y(i) - \mu_y) \tag{1}$$

with $\mu_y = \frac{1}{M}\sum_{m=1}^{M} y(m)$ the mean of the signal y.

Let Y and Y_{int} be two column vectors storing the samples $\{y(n)\}_{n=1,...,M}$ and $\{y_{int}(n)\}_{n=1,...,M}$ respectively. Given $H_M = \sum_{r=0}^{M-1} diag(\mathbb{1}_{1\times M-r}, -r)$ a low triangular matrix filled with ones, one has :

$$Y_{int} = [y_{int}(1), ..., y_{int}(M)]^T = \mathrm{H}_M J_M Y \tag{2}$$

Therefore, the first LN elements Y_{int} can be expressed as:

$$\begin{aligned} Y_{int}(1:LN) &= [y_{int}(1), ..., y_{int}(LN)]^T \\ &= C_{LN,0}Y_{int} \underset{(2)}{=} C_{LN,0}\mathrm{H}_M J_M Y \end{aligned} \tag{3}$$

B. *Estimation of the trend of the profile*

1) With the DFA [22]: the profile is split into L non-overlapping segments of length N, denoted as $\{y_{int,l}(n)\}_{l=1,...,L}$ with $n \in [\![1; N]\!]$. As M is not necessarily a multiple of N, the last $M - LN$ samples of the profile are not taken into account. The l^{th} local trend, which is the trend x_l of the l^{th} segment $y_{int,l}$, is modeled as a straight line $\forall l \in [\![1; L]\!]$ and $\forall n \in [\![1; N]\!]$:

$$x_l(n) = a_{l,1}[(l-1)N + n] + a_{l,0} \tag{4}$$

By respectively denoting X_l and $\theta_l = \left[a_{l,0}\; a_{l,1}\right]^T$ the $N \times 1$ vector storing the values of $x_l(n)$ and the parameter vector $\forall l \in [\![1; L]\!]$, one has:

$$X_l = A_l\theta_l \tag{5}$$

where A_l is a $N \times 2$ matrix whose first column corresponds to a vector of 1 and whose second column is defined by the set of values $\{(l-1)N + n\}_{n=1,...,N}$.

Let $\Theta_{DFA} = \left[\theta_1 \dots \theta_L\right]^T$ be the parameter vector of size $2L \times 1$, and A_{DFA} the $(LN \times 2L)$ block diagonal matrix defined from the set of matrices $\{A_l\}_{l=1,...,L}$. In this case, the parameters defining the local trends satisfy:

$$\underset{\Theta_{DFA}}{\arg\min} \left|\left|C_{LN,0}Y_{int} - A_{DFA}\Theta_{DFA}\right|\right|^2 \tag{6}$$

Therefore, the estimated parameter vector and the trend vector $T_{DFA} = A_{DFA}\hat{\Theta}_{DFA}$ satisfy:

$$\begin{cases} \hat{\Theta}_{DFA} = (A_{DFA}^T A_{DFA})^{-1} A_{DFA}^T C_{LN,0} Y_{int} \\ T_{DFA} = A_{DFA}(A_{DFA}^T A_{DFA})^{-1} A_{DFA}^T C_{LN,0} Y_{int} \end{cases} \tag{7}$$

2) With the CDFA: Instead of *a posteriori* correcting the discontinuities as done in the AFA, we suggest *a priori* introducing a constraint of continuity between local trends by presenting the so-called CDFA. For the L segments under study, our purpose is to ensure continuity between the consecutive local trends $\forall l \in [1; L-1]$. Therefore, there are two possibilities: $x_{l+1}(1) = x_l(N+1)$ or $x_{l+1}(0) = x_l(N)$. Given (4), defining the constraints amounts to minimizing the following criterion:

$$\begin{aligned} J(a_{1,1}, .., a_{L,1}, a_{1,0}) &= \sum_{n=1}^{N}(y_{int}(n) - a_{1,1}n - a_{1,0})^2 \\ &+ \sum_{l=2}^{L}\sum_{n=1}^{N}[y_{int}((l-1)N + n) - a_{l,1}[(l-1)N + n] \\ &- a_{1,0} - \sum_{j=1}^{l-1}\beta(j)(a_{j,1} - a_{j+1,1})]^2 \end{aligned} \tag{8}$$

with $\beta(l) = lN + 1$. To rewrite it in a matrix form, let the vector of parameters be defined as follows:

$$\Theta_{CDFA} = [a_{1,1}, .., a_{L,1}, a_{1,0}]^T \tag{9}$$

In addition, let us introduce A_{CDFA} of size $LN \times (L+1)$ whose first N rows, $A_{CDFA}(1:N, 1:L+1)$, are given by:

$$A_{CDFA}(1:N,1:L+1) = \begin{bmatrix} 1 & 0 & \cdots & 0 & 1 \\ 2 & & \vdots & & 1 \\ \vdots & & \vdots & & \vdots \\ N & & & & 1 \end{bmatrix} \quad (10)$$

and $\forall l \in [2; L-1]$:

$$A_{CDFA}((l-1)\times N+1 : lN, 1 : L+1) = \quad (11)$$
$$\begin{bmatrix} \beta(1) & N & \cdots & N & lN+1-\beta(l) & 0 & \cdots & 0 & 1 \\ \beta(1) & N & \cdots & N & lN+2-\beta(l) & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \beta(1) & \underbrace{N}_{} & \underbrace{\cdots}_{l-2} & \underbrace{N}_{} & (l+1)N-\beta(l) & \underbrace{0}_{} & \underbrace{\cdots}_{L-l} & \underbrace{0}_{} & 1 \end{bmatrix}$$

The criterion (8) becomes:

$$J(a_{1,1},..,a_{L,1},a_{1,0}) = \left|\left|C_{LN,0}Y_{int} - A_{CDFA}\Theta_{CDFA}\right|\right|^2 \quad (12)$$

Therefore, the vector of the estimated parameters and the trend vector are equal to:

$$\begin{cases} \hat{\Theta}_{CDFA} = [A_{CDFA}^T A_{CDFA}]^{-1} A_{CDFA}^T C_{LN,0} Y_{int} \\ T_{CDFA} = A_{CDFA}[A_{CDFA}^T A_{CDFA}]^{-1} A_{CDFA}^T C_{LN,0} Y_{int} \end{cases} \quad (13)$$

C. Computation of the residual

In the remainder, the subscript $\bullet$ denotes the method that is considered, *i.e.* DFA or CDFA. The residual vector $R_\bullet = C_{LN,0}Y_{int} - T_\bullet$ of the projection of $C_{LN,0}Y_{int}$ onto the space spanned by the columns of $A_\bullet$ can be expressed as follows:

$$R_\bullet = \left[I_{LN} - A_\bullet(A_\bullet^T A_\bullet)^{-1}A_\bullet^T\right]C_{LN,0}Y_{int} \quad (14)$$

Let $B_\bullet$ be equal to $\left[I_{LN} - A_\bullet(A_\bullet^T A_\bullet)^{-1}A_\bullet^T\right]$. By combining (2) and (14), this leads to:

$$R_\bullet = B_\bullet C_{LN,0}\mathrm{H}_M J_M Y \quad (15)$$

D. Computation of the square of the fluctuation function

Let us now define the following $M \times M$ matrix:

$$\Gamma_\bullet = \frac{1}{LN} J_M^T \mathrm{H}_M^T C_{LN,0}^T B_\bullet^T B_\bullet C_{LN,0} \mathrm{H}_M J_M \quad (16)$$

Using the properties of the trace of a matrix, the power of the residual $F_\bullet^2(N)$ can be expressed as:

$$F_\bullet^2(N) = Tr(\Gamma_\bullet Y Y^T) \quad (17)$$

E. Estimation of α

As $F_\bullet(N) \propto N^\alpha$ [21], $\log(F_\bullet(N))$ is plotted as a linear function of $\log(N)$. The goal is to search a straight line fitting the log-log representation. The quantity α is its slope and is estimated in the LS sense. Then, $H = \alpha - 1$, since the integration adds 1 in the estimation of α.

In the following, using this matrix presentation, let us first express the power of the residual from the correlation function of the process under study and consequently from its PSD, before comparing both methods.

III. Comparative Study

A. Link between the power of the residual and the PSD of the process under study

Let us rewrite (17) as follows:

$$F_\bullet^2(N) = \sum_{k=1}^{M} \Gamma_\bullet(k,k) y^2(k) \quad (18)$$
$$+ \sum_{r=1}^{M-1}\sum_{k=1}^{M-r} [\Gamma_\bullet(k,k+r) + \Gamma_\bullet(k+r,k)] y(k) y(k+r)$$

Assuming that y is w.s.s and taking the statistical mean of (18), one has:

$$E[F_\bullet^2(N)] = \sum_{r=-M+1}^{M-1} Tr(\Gamma_\bullet, r) R_{y,y}(r) \quad (19)$$

where $R_{y,y}(r)$ is the correlation function of the process y and $Tr(\Gamma_\bullet, r)$ denotes the r^{th} diagonal of the matrix $\Gamma_\bullet$. As the correlation function for real signals is symmetric and by denoting $g_{\Gamma_\bullet}(r) = Tr(\Gamma_\bullet, r)$, the above equation can be expressed as result of a convolution as follows:

$$E[F_\bullet^2(N)] = g_{\Gamma_\bullet} * R_{y,y}(\tau)|_{\tau=0} \quad (20)$$

Given the Wiener-Khintchine theorem and the properties of the inverse Fourier transform (TF^{-1}), $E[F_\bullet^2(N)]$ can be expressed from the PSD of y, denoted as $S_{yy}(f)$:

$$E[F_\bullet^2(N)] = TF^{-1}\Big(\big(\sum_{r=-M+1}^{M-1} Tr(\Gamma_\bullet, r) e^{-j2\pi f_n r}\big) S_{yy}(f)\Big)\Big|_{\tau=0}$$
$$= TF^{-1}\Big(\Psi_\bullet(f) S_{yy}(f)\Big)\Big|_{\tau=0} \quad (21)$$

In (21), $\Psi_\bullet(f) = \sum_{r=-M+1}^{M-1} Tr(\Gamma_\bullet, r) e^{-j2\pi f_n r}$ corresponds to the Fourier transform of the sequence $\{Tr(\Gamma_\bullet, r)\}_{r=-M+1,\ldots,M-1}$. Let us look at the properties of the latter: first of all, as it is real and even, $\Psi_\bullet(f)$ is necessarily real and even. Moreover, as $\Gamma_\bullet$ is a Gramian matrix since it is the product between $\frac{1}{\sqrt{LN}} B_\bullet C_{LN,0}\mathrm{H}_M J_M$ and its transpose, the element $\Gamma_\bullet(i,j)$ located at the i^{th} row and the j^{th} column of $\Gamma_\bullet$ corresponds to the scalar product between the i^{th} and the j^{th} rows of $\frac{1}{\sqrt{LN}} B_\bullet C_{LN,0}\mathrm{H}_M J_M$. Given the properties of the scalar product, one has:

$$|\Gamma_\bullet(i,j)| \leq |\Gamma_\bullet(i,i)| \quad (22)$$

As a corollary, using the inequality (22), one has:

$$|Tr(\Gamma_\bullet, r)| \leq \sum_{k=1}^{M-r} |\Gamma_\bullet(k,k+r)| \leq \sum_{k=1}^{M-r} \Gamma_\bullet(k,k)$$
$$\leq \sum_{k=1}^{M-1} \Gamma_\bullet(k,k) = Tr(\Gamma_\bullet, 0) = Tr(\Gamma_\bullet)$$

In the above, note that $Tr(\Gamma_\bullet)$ corresponds to the square of the Froebenius norm of the matrix $\Gamma_\bullet$. It is necessarily positive. As a consequence, the sequence can be seen as the convolution of a vector with its flipped version and its Fourier transform

$\Psi_{\bullet}(f)$ is necessarily positive. Therefore it can be seen as the PSD of the signal y filtered by a filter whose transfer function $H_{filter}(z)$ satisfies: $\Psi_{\bullet}(f) = |H_{filter}(z)|^2_{z=exp(j\theta)}$, with $\theta = 2\pi f/f_s$ the normalized angular frequency. Consequently, we can conclude that $E[F_{\bullet}^2(N)]$ corresponds to the correlation function of the filter output calculated for the lag equal to 0, *i.e.* the power of the filter output.

B. Illustrations and comments

Let us study the influence of N on $\Psi_{DFA}(f)$ and $\Psi_{CDFA}(f)$. Using (21) and the expressions of Γ_{DFA} and Γ_{CDFA}, we noticed:

1. The null frequency is always rejected, which is consistent with the purpose of detrending as shown in Fig. 1. According to the simulations we carried out, the orders of magnitude of $\Psi_{DFA}(0)$ and $\Psi_{CDFA}(0)$ are 10^{-16} and 10^{-15}. When $N = 3$, the DFA acts as a high-pass filter whereas it exhibits two resonances at $\pm f_{N,DFA}$ for larger values of N. The CDFA always acts as a band-pass filter characterized by the resonance frequencies $\pm f_{N,CDFA}$. In the following, let $bw_{\bullet}$ be the -3 dB bandwidth[1] of the filter associated to $\Psi_{\bullet}$.

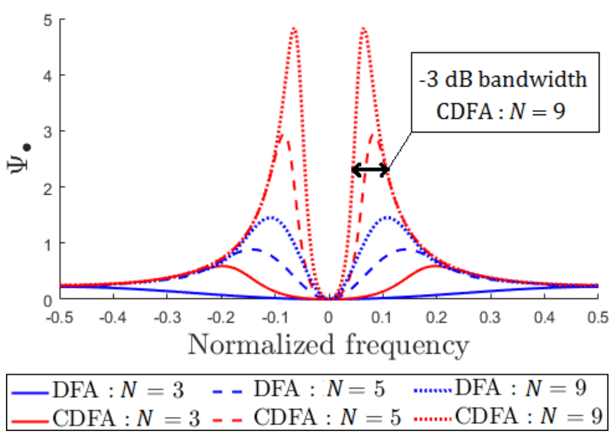

Figure 1. Comparison of the filtering induced by the DFA and the CDFA for $N = 3$, $N = 5$ and $N = 9$.

2. As depicted in Fig. 2, the larger N, the spikier the resonances of the frequency responses. The -3 dB bandwidth decreases when N increases. In addition, the resonance frequencies also move to low frequencies when N increases. The bandwidths and resonance frequencies of the filters corresponding to the DFA and the CDFA are clearly different for small values of N, but they tend to be the same as N increases.

3. For any N, $\Psi_{CDFA}(f)$ is spikier and larger than $\Psi_{DFA}(f)$, for most of the frequencies. See Fig. 1 where three values of N are presented for the sake of clarity. In addition, the log-spectral distance between $\Psi_{DFA}(f)$ and $\Psi_{CDFA}(f)$ based on FFT with zero padding, depicted in Fig. 3, decreases when N increases. This means that the DFA and the CDFA tend to have the same behaviour when N increases.

[1]It corresponds to the frequencies for which $10\log\frac{\Psi_{\bullet}(f)}{\Psi_{\bullet}(f_{N,\bullet})} > -3$

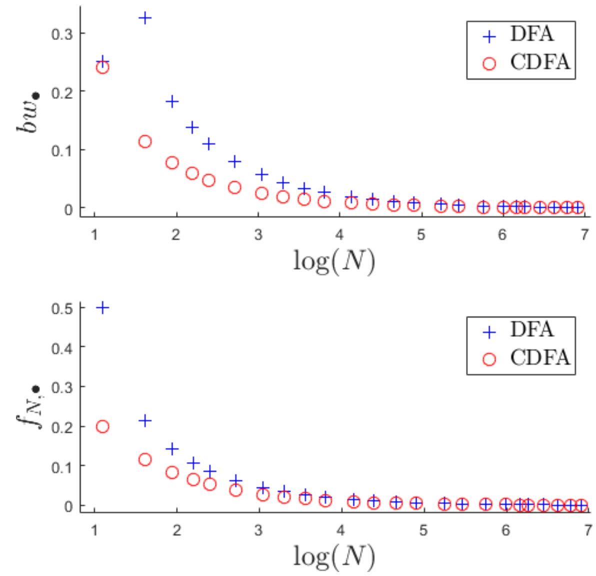

Figure 2. Evolution of $bw_{\bullet}$ (top) and $f_{N,\bullet}$ (bottom) of both filters as a function of $\log(N)$.

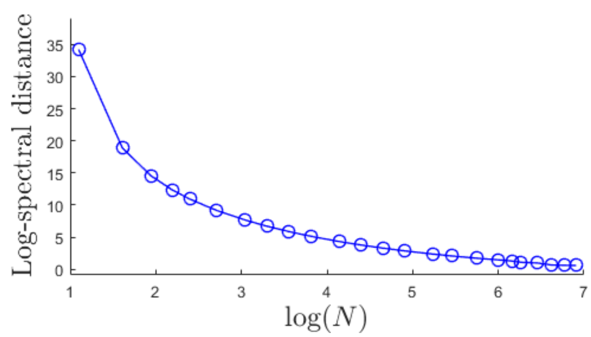

Figure 3. Evolution of the Log-spectral distance (LSD) between $\Psi_{DFA}(f)$ and $\Psi_{CDFA}(f)$ as a function of $\log(N)$.

C. Comparing the DFA and the CDFA with a toy example

Let us estimate the Hurst coefficient of a w.s.s. zero-mean white noise to illustrate the differences and the similarities between the DFA and the CDFA. In Fig. 4, $\log(F(N))$ is represented as a function of $\log(N)$, as done in step II-E of both methods. The difference between the values of $\log(F_{\bullet}(N))$ obtained with the DFA and the CDFA decreases as N increases. It is coherent with the analysis we did in the previous section, especially with Fig. 3. Then, the slope α -and consequently H- depends on the values of N that are considered. They tend to be the same if large values of N are used. In Fig 4, an example is given for one realization, where α is computed with the DFA or the CDFA and with small or large values of N. Then, given Table I, the CDFA provides more accurate estimations of α for small values of N.

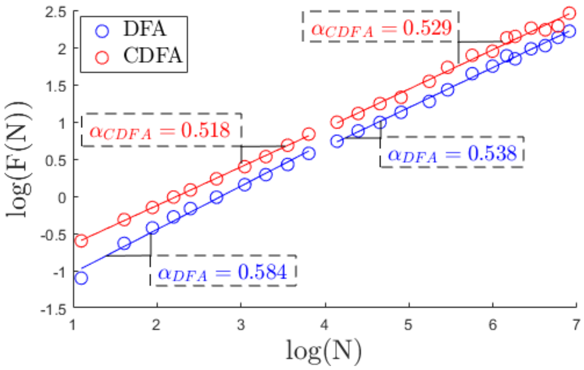

Figure 4. Evolution of $\log(F(N))$ as a function of $\log(N)$ for both the DFA and the CDFA, in the case of a white noise.

Table I
COMPARISON OF THE MEAN AND VARIANCE VALUES OF α WITH EACH APPROACH, ESTIMATED ON 500 WHITE NOISES FOR SMALL VALUES OF N.

	Mean	Variance	% err.
DFA	0.592	3.29×10^{-4}	18.4
CDFA	0.507	5.16×10^{-4}	1.40

IV. CONCLUSIONS AND PERSPECTIVES

In this paper, our purpose was to analyze the difference between the DFA in which the global trend is constructed from discontinuous local trends and its version when *a priori* constraints are added to guarantee the continuity of the global trend. To this end, we suggested analyzing both methods by comparing the squares of their fluctuation functions using a filtering-based interpretation. Both can be seen as *ad hoc* wavelet based methods, but their main difference stands when the length of the local trends are small. We currently prepare a global comparative study including the DMA, the AFA and the regularized DFA and analyzing a large set of processes.

REFERENCES

[1] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch. Self-similarity and long-range dependence through the wavelet lens. *Theory and Applications of Long-range Dependence*, pages 527–556, 2003.
[2] E. Alessio, A. Carbone, G. Castelli, and V. Frappietro. Second-order moving average and scaling of stochastic time series. *The European Physical Journal B*, 27, 2:197–200, 2002.
[3] P. N. Balljekar and H. A. Patil. A comparison of waveform fractal dimension techniques for voice pathology classification. *International Conference on Acoustics, Speech and Signal Processing*, pages 4461–4464, 2012.
[4] J.-M. Bardet. Testing for the presence of self-similarity of gaussian time series having stationary increments. *Journal of Time Series Analysis*, 21:497–515, 2000.
[5] C. L. Chang and M. McAleer. Econometric analysis of financial derivatives: an overview. *Journal of Econometrics*, 187, (2):403–407, 2015.
[6] F. Esposti and M. G. Signorini. Evaluation of a blind method for the estimation of Hurst's exponent in time series. *European Signal Processing Conference*, pages 1–5, 2006.
[7] C. W. J. Granger and R. Joyeux. An introduction to long-memory time series models and frafgaoctional differencing. *Journal of Times Series Analysis*, 1, (1):15 – 29, 1980.
[8] R. J. Hodrick and E. C. Prescott. Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29, (1):1–16, 1997.
[9] J. R. M. Hosking. Fractional differencing. *Biometrika*, 68, (1):165–176, 1981.
[10] M. Höll and H. Kantz. The relationship between the detrendend fluctuation analysis and the autocorrelation function of a signal. *The European Physical Journal B*, 88:327, 2015.
[11] M. Höll, H. Kantz, and Y. Zhou. Detrended fluctuation analysis and the difference between external drifts and intrinsic diffusionlike nonstationarity. *Physical Review E*, 94:042201, 2016.
[12] 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*, 295, (3-4):441–454, 2001.
[13] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, A. Bunde, S. Havlin, and H. Eugene Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, 316, (1-4):87–114, 2002.
[14] S.-J. Kim, K. Koh, S. Boyd, and D. Gorinevsky. l_1 trend filtering. *SIAM Review*, 51, (2):339–360, 2009.
[15] K. Kiyono. Establishing a direct connection between detrended fluctuation analysis and fourier analysis. *Physical Review E*, 92:042925, 2015.
[16] K. Kiyono. Theory and applications of detrending -operation -based fractal-scaling analysis. *International Conference on Noise and Fluctuations (ICNF)*, pages 1–4, 2017.
[17] B. B. Mandelbrot and J. W. Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422–437, 1968.
[18] E. Moulines, F. Roueff, and M. S. Taqqu. Central limit theorem for the log-regression wavelet estimation of the memory parameter in the gaussian semi-parametric context. *Fractals*, 15:301–313, 2007.
[19] X. Navarro, A. Beuchée, F. Porée, and G. Carrault. Performance analysis of Hurst's exponent estimators in higly immature breathing patterns of preterm infants. *International Conference on Acoustics, Speech and Signal Processing*, pages 701–704, 2011.
[20] S. K. Nayak, A. Bit, A. Dey, B. Mohapatra, and K. Pal. A review on the nonlinear dynamical system analysis of electrocardiogram signal. *Journal of Healthcare Engineering*, (2):1–19, 2018.
[21] C. K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley. Long-range correlations in nucleotide sequences. *Nature*, 356:168–170, 1992.
[22] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. Mosaic organization of DNA nucleotides. *Physical Review E*, 49, (2):1685–1689, 1994.
[23] A. G. Ravelo-Garcia, U. Casanova-Blancas, S. Martin-González, E. Hernández-Pérez, I. Guerra-Moreno, P. Quintana-Morales, N. Wessel, and J. L. Navarro-Mesa. An approach to the enhancement of sleep apnea detection by means of detrended fluctuation analysis of RR intervals. *Computing in Cardiology*, pages 905–908, 2014.
[24] M. A. Riley, S. Bonnette, N. Kuznetsov, S. Wallot, and J. Gao. A tutorial introduction to adaptive fractal analysis. *Frontiers in Physiology*, 3:371, 2012.
[25] G. Rilling, P. Flandrin, and P. Gonçalves. Empirical mode decomposition, fractional gaussian noise and hurst exponent estimation. *International Conference on Acoustics, Speech and Signal Processing*, pages 489–492, 2005.
[26] S. Sanyal, A. Banerjee, R. Pratihar, A. Kumar Maity, S. Dey, V. Agrawal, R. Sengupta, and D. Ghosh. Detrended fluctuation and power spectral analysis of alpha and delta EEG brain rhythms to study music elicited emotion. *International Conference on Signal Processing, Computing and Control*, pages 206–210, 2015.
[27] R. Sun. Fractional order signal processing: techniques and applications. *Thesis of Master of science in electrical Engineering, Utah State University*, 2007.
[28] M. S. Taqqu and V. Teverovsky. On estimating the intensity of long range dependence in finite and infinite variance time series. *A practical guide to heavy tails: statistical techniques and applications*, pages 177–217, 1996.
[29] M. P. Tarvainen, P. O. Ranta-aho, and Pasi A. Karjalainen. An advanced detrending method with application to hrv analysis. *IEEE Trans. on Biomedical Engineering*, 49, 2:172–175, 2002.
[30] Y. Tsujimoto, Y. Miki, S. Shimatani, and K. Kiyono. Fast algorithm for scaling analysis with higher-order detrending moving average method. *Physical Review E*, 93 (5):053304, 2016.
[31] L. Xu, P. Ch. Ivanov, K. Hu, Z. Chen, A. Carbone, and H. E. Stanley1. Quantifying signals with power-law correlations: A comparative study of detrended fluctuation analysis and detrended moving average techniques. *Physical Review E*, 71, 5:051101, 2005.