2D Fourier Transform Based Analysis Comparing the DFA with the DMA

Bastien Berthelot*†, Éric Grivel‡, Pierrick Legrand†, Marc Donias‡, 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*—Even if they can be outperformed by other methods, the detrended fluctuation analysis (DFA) and the detrended moving average (DMA) are widely used to estimate the Hurst exponent because they are based on basic notions of signal processing. For the last years, a great deal of interest has been paid to compare them and to better understand their behaviors from a mathematical point of view. In this paper, our contribution is the following: we first propose to express the square of the so-called fluctuation function as a 2D Fourier transform (2D-FT) of the product of two matrices. The first one is defined from the instantaneous correlations of the signal while the second, called the weighting matrix, is representative of each method. Therefore, the 2D-FT of the weighting matrix is analyzed in each case. In this study, differences between the DFA and the DMA are pointed out when the approaches are applied on non-stationary processes.**

***Index Terms*—frequency analysis, Hurst, DFA, DMA.**

I. Introduction

In various applications, signal features can be extracted for classification. Model parameters, powers in some frequency bands but also zero crossing are often considered. In the field of biomedical, the long-range dependence (LRD) based on the Hurst exponent is often used.

Two main families of approaches exist to estimate the Hurst exponent. The first one consists of frequency-domain estimators including the local Whittle method, the periodogram method, the approach based on the empirical mode decomposition (EMD) [17] or the fractional Fourier transform [18], the wavelet-based method [1] and the semi-parametric method [11]. The second one gathers the time-domain estimators such as the so-called rescaled range analysis, the aggregated variance method, the absolute-value method and the variance-of-residuals method. See [20] and [19] for more details.

To estimate the Hurst exponent of a pure mono-fractal time series, Peng *et al.* proposed the fluctuation analysis (FA) [14] and then the detrended fluctuation analysis (DFA) [15]. The first step of the DFA was to define the trend of the integrated signal. The authors suggested defining it by concatenating discontinuous local trends modeled by straight lines of length N_{DFA}. Then, the fluctuation function, corresponding to the square root of the power of the difference between the integrated signal and its trend, is shown to be proportional to N_{DFA}^{α}, where α is the scaling exponent. The Hurst exponent is then deduced as it is equal to $\alpha-1$. Since, various ways have been proposed to obtain the global trend of a signal [8]. This led to several variants of the DFA such as the higher-order DFA where the local trends are approximated by polynomials of order higher than 1, the regularized DFA which is based on a regularized least-squares (LS) criterion to obtain the trend [21] and the adaptive fractal analysis (AFA) [16] correcting the discontinuities *a posteriori*. In the following, we focus our attention on the detrended moving average (DMA), also known as "moving average filtering" in economics. It consists in low-pass filtering the integrated signal to deduce the trend. In its standard version, the filter has a finite impulse response of length N_{DMA} [2] [13].

The DFA and the DMA belong to the nonlinear dynamical system analysis techniques which also include Lyapunov exponents [12]. In addition to the approximate entropy and the sample entropy, they are widely used especially in the field of biomedical, gait analysis, stock market prediction and meteorology. Unlike the wavelet-based approach or the local Whittle method, the DFA and the DMA have the advantage of *a priori* not requiring advanced skills in statistical signal processing since they are based on simple concepts like regression and linear filtering. This is probably one of the reasons of their popularity even if they may be outperformed by other approaches. It is a compromise between performance, computational cost and simplicity of implementation and use.

For the last years, efforts have been made to provide extensions of the algorithms [3], to develop fast versions of these approaches [22] [23], and to propose mathematical analysis in order to better understand their behaviors [4]–[7] [9] [10].

In [4], the authors aimed at finding a relation between the square of the fluctuation function and an estimation of the normalized covariance function of the signal, by assuming that the process was wide-sense stationary (w.s.s.) and ergodic and by making some approximations. Kiyono *et al.* analyzed the single-frequency responses of both the DFA and the higher-order DFA [9] and the centered DMA [10], but did not take into account the centering step. They concluded that, for stochastic processes whose power spectral density (PSD) is a function of the frequency f of the form $f^{-\beta}$, the higher-order DFA is convenient to estimate α as long as $\alpha = \frac{\beta+1}{2}$.

Our work is complementary to the above recent studies. It

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

is based on a uniform matrix way to express the square of the fluctuation function. For the DFA and the DMA, we first propose analytic expressions of the trend vectors, the residuals and the powers of the residuals. Then, we give an interpretation of these powers as the 2D-FT of a matrix defined from the instantaneous correlations of the signal multiplied by a weighting matrix. Therefore, analyzing the 2D-FT of the weighting matrix is a way to compare the DFA with the DMA. The remainder of this paper is organized as follows: In section II, the main steps of the DFA and the DMA are briefly recalled. Section III deals with the uniform presentation and the interpretation. Conclusions and perspectives are then given.

II. Steps of the DFA and the DMA

Let us consider M consecutive samples $\{y(m)\}_{m=1,...,M}$ of the signal. The DFA and the DMA are defined by the following four steps [15] [2]:

1. The profile $y_{int}(m) = \sum_{i=1}^{m}(y(i) - \mu_y)$ is first computed, where $\mu_y = \frac{1}{M}\sum_{m=1}^{M} y(m)$ is the mean of y.

2. The trend of the profile is then estimated.

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

$$x_l(n) = a_{l,1}[(l-1)N_{DFA} + n] + a_{l,0} \quad (1)$$

Then, $\forall l \in [\![1; L]\!]$, the parameter vector $\theta_l = \begin{bmatrix} a_{l,0} & a_{l,1} \end{bmatrix}^T$ is estimated in the LS sense from the profile.

- With the DMA, known as "simple moving average" [24], the profile is low-pass filtered. The impulse response of the filter is given by $h_{DMA}(n) = \frac{1}{N_{DMA}}$ for $n = 0, ..., N_{DMA}-1$. Due to the impulse-response symmetry, it is a linear-phase filter with a constant group delay equal to $\frac{N_{DMA}-1}{2}T_s$, with T_s the sampling period. Therefore, N_{DMA} is chosen odd.

In the following, although N_{DFA} and N_{DMA} do not have the same meaning, $N_{DFA} = N_{DMA} = N$.

3. The resulting trend is subtracted to the profile. The fluctuation function, *i.e.* the square root of the residual power, $F_\bullet(N)$, is computed. $_{(\bullet)}$ denotes the method (DFA or DMA).

4. Steps 2 and 3 are repeated for different values of N. At this stage, as $F_\bullet(N) \propto N^\alpha$ [14], $\log(F_\bullet(N))$ is plotted as a linear function of $\log(N)$. The final step is to search a straight line fitting the log-log representation. The quantity α is hence estimated in the LS sense.

III. Uniform presentation of the DFA and the DMA based on a matrix form

A. Notations

Let us first introduce some notations that will be useful in the remainder of the paper. 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 $[.]$. Thus, $diag(\mathbb{1}_{1\times N-1}, 1)$ is the square matrix of size N whose 1^{st} sub-diagonal above the main one has its elements equal to 1. X_l is a $N \times 1$ vector storing the values of $x_l(n)$. Y and Y_{int} are two column vectors storing the samples $\{y(n)\}_{n=1,...,M}$ and $\{y_{int}(n)\}_{n=1,...,M}$ respectively. Using the above notations, this leads to:

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

with $H_M = \sum_{r=0}^{M-1} diag(\mathbb{1}_{1\times M-r}, -r)$ a low triangular matrix filled with ones. Finally, let us introduce the matrix of size (j, M) defined as follows:

$$C_{j,k} = [\mathbb{0}_{j\times k} \quad I_j \quad \mathbb{0}_{j\times(M-(j+k))}] \quad (3)$$

In this case, one can easily express the first LN elements of the vector Y_{int} as follows:

$$\begin{aligned} Y_{int}(1:LN) &= [y_{int}(1), y_{int}(2), ..., y_{int}(LN)]^T \quad (4) \\ &= C_{LN,0}Y_{int} \underset{(2)}{=} C_{LN,0}\mathrm{H}_M J_M Y \end{aligned}$$

Finally, for the sake of simplicity, let us define $N' = \frac{N-1}{2}$.

B. Matrix form of the power of the residual with DFA

Using a vector form of (1), the local trend vector satisfies:

$$X_l = A_l \theta_l \; \forall l \in [\![1; L]\!] \quad (5)$$

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 $\{(l-1)N + n\}_{n=1,...,N}$. By introducing the parameter vector $\Theta_{DFA} = \begin{bmatrix}\theta_1 \dots \theta_L\end{bmatrix}^T$ of size $2L\times 1$, and the $(LN\times 2L)$ matrix A_{DFA} which is block diagonal matrix defined from the set of matrices $\{A_l\}_{l=1,...,L}$, the parameters defining the local trends satisfy:

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

This leads to:

$$\hat{\Theta}_{DFA} = (A_{DFA}^T A_{DFA})^{-1} A_{DFA}^T C_{LN,0} Y_{int} \quad (7)$$

The trend vector $T_{DFA} = A_{DFA}\hat{\Theta}_{DFA}$ can then be deduced:

$$T_{DFA} = A_{DFA}(A_{DFA}^T A_{DFA})^{-1} A_{DFA}^T C_{LN,0} Y_{int} \quad (8)$$

Therefore, the residual $R_{DFA} = C_{LN,0}Y_{int} - T_{DFA}$ of the projection of $C_{LN,0}Y_{int}$ onto the space spanned by the columns of A_{DFA} is given by:

$$R_{DFA} \underset{(8)}{=} \left[I_{LN} - A_{DFA}(A_{DFA}^T A_{DFA})^{-1} A_{DFA}^T\right] C_{LN,0} Y_{int} \quad (9)$$

Defining B_{DFA} as $\left[I_{LN} - A_{DFA}(A_{DFA}^T A_{DFA})^{-1} A_{DFA}^T\right]$ and by combining (2) and (9), one has:

$$R_{DFA} = B_{DFA} C_{LN,0} \mathrm{H}_M J_M Y = \overline{B}_{DFA} Y \quad (10)$$

Given $\Gamma_{DFA} = \frac{1}{LN}\overline{B}_{DFA}^T \overline{B}_{DFA}$ of size $M \times M$ and using the properties of the trace of a matrix, the power of the residual, also called the generalized variance [22], can be expressed as follows[1]:

$$F_{DFA}^2(N) = \frac{1}{LN} R_{DFA}^T R_{DFA} = Tr(\Gamma_{DFA} Y Y^T) \quad (11)$$

[1] F_{DFA}^2 is explicitly written as a function of N. For the sake of simplicity, this is omitted for the previous quantities (B_{DFA}, R_{DFA}, T_{DFA}, etc.).

Comments on Γ_{DFA}: Let us rewrite Γ_{DFA} :

$$\Gamma_{DFA} = \frac{1}{LN} J_M^T \mathrm{H}_M^T C_{LN,0}^T \times (I_{LN} - A_{DFA}(A_{DFA}^T A_{DFA})^{-1} A_{DFA}^T) C_{LN,0} \mathrm{H}_M J_M \quad (12)$$

In (12), the matrix $I_{LN} - A_{DFA}(A_{DFA}^T A_{DFA})^{-1} A_{DFA}^T$ is block diagonal. Each block is defined by $I_N - A_l(A_l^T A_l)^{-1} A_l^T$ with $l = 1, ..., L$. They are all equal. Indeed, $A_l(A_l^T A_l)^{-1} A_l^T$ corresponds the orthogonal projector onto the space spanned by the columns of A_l. However, the 1^{st} column of A_l is equal to the 1^{st} column of A_1 while the 2^{nd} column of A_l is a linear combination of the 1^{st} and 2^{nd} columns of A_1. So, for $l = 1, ..., L$, one has:

$$\mathbf{A_1} = I_N - A_l(A_l^T A_l)^{-1} A_l^T = I_N - A_1(A_1^T A_1)^{-1} A_1^T \quad (13)$$

The above matrix $\mathbf{A_1}$ has three main properties that will be used in the following: 1/ $\mathbf{A_1} = \mathbf{A_1}^T$. 2/ the space spanned by the columns of $\mathbb{1}_{N\times N}$ corresponds to the space spanned by the first column of A_1. Therefore, $\mathbf{A_1}\mathbb{1}_{N\times N}$ is the null matrix. 3/ the space spanned by the columns of $\mathrm{H}_N \mathbb{1}_{N\times N}$ corresponds to the space spanned by the two columns of A_1. Therefore, $\mathbf{A_1}\mathrm{H}_N \mathbb{1}_{N\times N}$ is also the null matrix. By combining all the properties, it can be shown that $\mathbb{1}_{N\times N}\mathbf{A_1}$ and $\mathbb{1}_{N\times N}\mathrm{H}_N^T \mathbf{A_1}$ are null matrices. So, taking into account the above comments on $\mathbf{A_1}$, Γ_{DFA} can be expressed as follows:

$$\Gamma_{DFA} = \frac{1}{LN} \mathrm{H}_M^T C_{LN,0}^T (I_L \otimes \mathbf{A_1}) C_{LN,0} \mathrm{H}_M \quad (14)$$

where $\otimes$ denotes the Kronecker product. This matrix is a block diagonal matrix, where each block is the same.

C. Matrix form of the power of the residual with DMA

With the DMA, the M samples of the profile are filtered. Instead of using a convolution at each time step, let us express the vector storing the filter output samples. This can be done by premultiplying Y_{int} by the filtering matrix M_{filt}:

$$M_{filt} = \frac{1}{N} \sum_{r=0}^{N-1} diag(\mathbb{1}_{1\times M-r}, -r) \quad (15)$$

In addition, the group delay of the filter corresponding to N' samples has to be compensated. This can be done by adding a pre-multiplication by the following $M \times M$ matrix:

$$M_{comp} = diag\left(\mathbb{1}_{1\times(M-N')}, N'\right) \quad (16)$$

The resulting trend vector is equal to $M_{comp} M_{filt} Y_{int}$. However, the last N' elements of this vector are equal to 0. In addition, due to the transient behavior of the filtering which corresponds to the first $N-1$ samples and the delay compensation introduced above, the first N' elements of the current trend vector should not be taken into account. For the above reasons, only a vector of size $M-N+1$ should be considered. This amounts to adding another pre-multiplication by the matrix $C_{M-N+1,N'}$. Therefore, the trend vector satisfies:

$$\begin{aligned} T_{DMA} &= C_{M-N+1,N'} M_{comp} M_{filt} Y_{int} \quad (17) \\ &\underset{(2)}{=} C_{M-N+1,N'} M_{comp} M_{filt} \mathrm{H}_M J_M Y \end{aligned}$$

Therefore, the following residual is considered:

$$\begin{aligned} R_{DMA} &= C_{M-N+1,N'} \left(I_M - M_{comp} M_{filt}\right) \mathrm{H}_M J_M Y \quad (18) \\ &= B_{DMA} \mathrm{H}_M J_M Y = \overline{B}_{DMA} Y \end{aligned}$$

Then, using trace properties, $F^2_{DMA}(N)$ can be deduced:

$$F^2_{DMA}(N) = \frac{1}{M-N+1} R_{DMA}^T R_{DMA} = Tr(\Gamma_{DMA} Y Y^T) \quad (19)$$

Comments on $\Gamma_{DMA} = \frac{1}{M-N+1}\overline{B}_{DMA}^T \overline{B}_{DMA}$: It can be shown that B_{DMA} is a rectangular Toeplitz matrix of size $(M-N+1)\times M$ whose first column is $[-\frac{1}{N} \quad \mathbb{0}_{1\times(M-N)}]^T$ and first row is $[\underbrace{-\frac{1}{N} \quad \cdots \quad -\frac{1}{N}}_{N'} \quad 1 - \frac{1}{N} \quad \underbrace{-\frac{1}{N} \quad \cdots \quad -\frac{1}{N}}_{N'} \quad \mathbb{0}_{1\times(M-N)}]$. Therefore, $B_{DMA}\mathrm{H}_M$ is also a rectangular Toeplitz matrix of size $(M-N+1)\times M$ whose first column is $\mathbb{0}_{(M-N+1)\times 1}$ and first row is $[0 \ \ \frac{1}{N} \cdots \frac{N'}{N} \ \ -\frac{N'}{N} \cdots -\frac{1}{N} \ \ \mathbb{0}_{1\times(M-N)}]$. Finally, $B_{DMA}\mathrm{H}_M 1_{M\times M}$ is the null matrix. Therefore, Γ_{DMA} is a symmetric matrix of size $M \times M$ whose first row and first column are null. The rest of the matrix is a null matrix except the main diagonal and the first $N'+1$ sub-diagonals above and below the main diagonal.

D. Towards a uniform expression of the residual power

Given the above two subsections, the power of the residual $F^2_\bullet(N)$ with $\bullet =_{DFA}$ or $_{DMA}$ can be expressed as:

$$F^2_\bullet(N) = Tr(\Gamma_\bullet Y Y^T) \quad (20)$$

It should be noted that as α is computed from the log of $F^2_\bullet(N)$ for multiple values of N, it is a non-linear function of Y -as we could expect-. As M is the size of the square matrix $\Gamma_\bullet$, (20) becomes:

$$\begin{aligned} F^2_\bullet(N) =& \sum_{k=1}^{M} \Gamma_\bullet(k,k) y^2(k) \quad (21) \\ &+ \sum_{r=1}^{M-1} \sum_{k=1}^{M-r} [\Gamma_\bullet(k,k+r) + \Gamma_\bullet(k+r,k)] y(k) y(k+r) \end{aligned}$$

Thus, $F^2_\bullet(N)$ can be expressed from the instantaneous correlations $\{y(k)y(k+r)\}_{r=-M+1,...,M-1}$ which are then weighted. Given the structure of $\Gamma_\bullet$ provided in sections III-B and III-C, $F^2_{DFA}(N)$ and $F^2_{DMA}(N)$ do not depend on the instantaneous correlation of the signal y whose lag is strictly larger than $N-1$ and $N'+1$ respectively.

Let us now introduce the following two matrices of size $(2M-1 \times 2M)$. Y_{corr} has its k^{th} column storing the instantaneous correlations of the signal y at time k for the lags varying from $M-1$ to $1-M$:

$$Y_{corr} = \begin{bmatrix} y(1)y(1-(M-1)) & \dots & y(M)y(M-(M-1)) \\ y(1)y(1-(M-2)) & \dots & y(M)y(M-(M-2)) \\ \vdots & \vdots & \vdots \\ y(1)y(1) & \dots & y(M)y(M) \\ \vdots & \vdots & \vdots \\ y(1)y(1+(M-2)) & \dots & y(M)y(M+(M-2)) \\ y(1)y(1+(M-1)) & \dots & y(M)y(M+(M-1)) \end{bmatrix} \quad (22)$$

The weighting matrix $W_\bullet$ is filled with zeros. Its i-th anti-diagonal[2], with $i = 1, ..., M$ is of length M and equal to $[\Gamma_\bullet(1,i), \Gamma_\bullet(2,i), \ldots, \Gamma_\bullet(M,i)]$. Thus, for $M = 3$, one has:

$$W_\bullet = \begin{bmatrix} 0 & 0 & \Gamma_\bullet(3,1) \\ 0 & \Gamma_\bullet(2,1) & \Gamma_\bullet(3,2) \\ \Gamma_\bullet(1,1) & \Gamma_\bullet(2,2) & \Gamma_\bullet(3,3) \\ \Gamma_\bullet(1,2) & \Gamma_\bullet(2,3) & 0 \\ \Gamma_\bullet(1,3) & 0 & 0 \end{bmatrix} \quad (23)$$

Let us give two examples of $W_\bullet$ in Fig. 1 with $N = 9$ and $N = 21$. The colormap was chosen to distinguish the null values from the negative ones and the positive ones. Due to the expression and the properties of Γ_{DFA}, W_{DFA} corresponds to the copy of a pattern $\lfloor \frac{M}{N} \rfloor$ times. Additionally, the details given about the structure of $\Gamma_\bullet$ in subsections III-B and III-C make it possible to explain the sparse matrices we get.

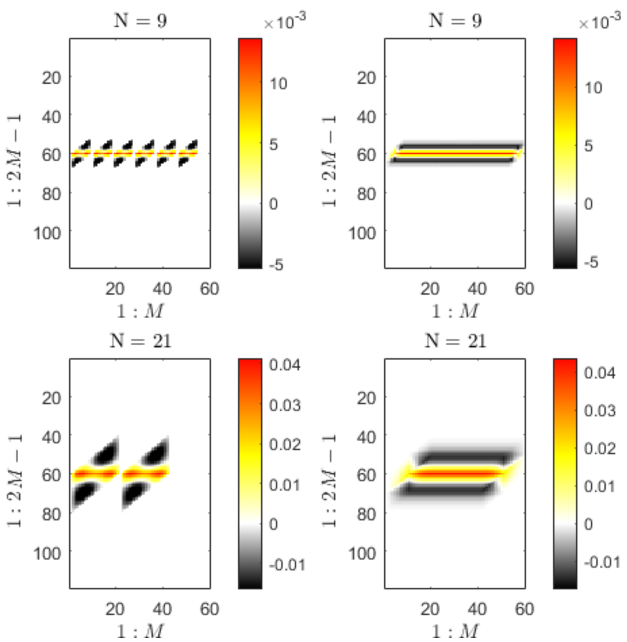

Figure 1. Weighting matrices W_{DFA} (left) and W_{DMA} (right) for $M = 60$.

Given (22) and (23), (21) can be seen as the two-dimensional Fourier transform (2D-FT) of the element-wise multiplication of Y_{corr} and $W_\bullet$ for the set of the spatial frequencies $(u, v) = (0, 0)$. Given the properties of the Fourier transform, $F_\bullet^2(N)$ can also be expressed as the convolution between the 2D-FTs of Y_{corr} and $W_\bullet$:

$$\begin{aligned} F_\bullet^2(N) &= \mathscr{F}(Y_{corr} W_\bullet)|_{(u=0,v=0)} \qquad (24) \\ &= \Big(\mathscr{F}(Y_{corr}) \circledast \mathscr{F}(W_\bullet)\Big)|_{(u=0,v=0)} \end{aligned}$$

where u and v are the normalized spatial frequencies, $\mathscr{F}$ denotes the 2D-FT and $\circledast$ the convolution.
Therefore, one way to compare the DMA with the DFA is to compare the properties of the 2D-FT of $W_\bullet$.

[2]The i^{th} anti-diagonal of the matrix $W_\bullet$ corresponds to the set of elements located at the $(2M+1-j-i)^{th}$ row and the j^{th} column, with $j = 1, ..., M$.

E. Analysis of $\mathscr{F}(W_\bullet)$ and its influence on $F_\bullet^2(N)$

1) About the 2D-FT of the weighting matrix: Let us first give some comments on the 2D-FT of the weighting matrix $W_\bullet$ for the DFA and the DMA.

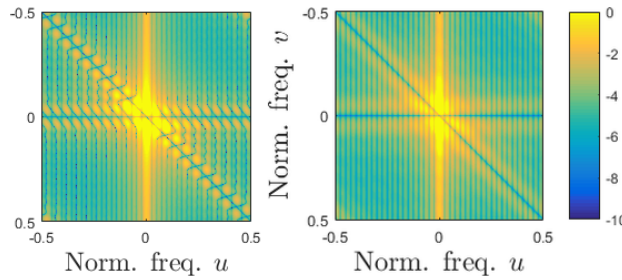

Figure 2. $\log(|\mathscr{F}(W_\bullet)|)$ with $\bullet = DFA$ (left) and $\bullet = DMA$ (right) for $M = 60$ and $N = 21$.

In Fig. 2, taking advantage of the properties of the 2D-FT[3], the large values of $\log(|\mathscr{F}(W_{DFA})|)$ appearing when u is a multiple of $1/N$ are representative of two features of the weighting matrix W_{DFA}: the periodicity of W_{DFA} along the x-axis induces the large values of $\log(|\mathscr{F}(W_{DFA})|)$ on both sides of the axis $v = 0$. The non-null anti-diagonals of W_{DFA} induce the large values located on both sides of the axis $v = -u$. For the DMA, $\log(|\mathscr{F}(W_{DMA})|)$ is mainly characterized by frequencies around $u = 0$ and $v = -u$.
Let us study how $\log(|\mathscr{F}(W_\bullet)|)$ evolves when N increases. As shown in Fig. 3, for both the DFA and the DMA, the contribution of $(u = 0, v = 0)$ remains close to 0. When N increases, $W_\bullet$ is more and more composed of frequencies $\sqrt{u^2 + v^2}$ that become smaller and smaller. For each approach, $|\log(F(W_\bullet)|$ has two main lobes located at $(0, v_{N,\bullet})$ and $(0, -v_{N,\bullet})$ where $v_{N,\bullet}$ becomes smaller when N increases.
Using our comments on the 2D-FT of $W_\bullet$, let us now deduce how it influences the power of the residual $F_\bullet^2$.

2) Impact on the power of the residual: Let us first look at the DMA case. Given (24) and the fact that the modulus of the 2D-FT of W_{DMA} exhibits two main peaks at $\pm v_{N,DMA}$, the frequency components of Y_{corr} at $u = 0$ and $v = \mp v_{N,DMA}$ are amplified and contribute to $F_{DMA}^2(N)$.
Let us analyze $E[F_\bullet^2(N)]$ when y is w.s.s. In this case, $E[Y_{corr}]$ is a matrix where each row contains the same element, which is the correlation function of y for a specific lag. In this case, by neglecting the windowing influence along the x-axis, the modulus of the 2D-FT of $E[Y_{corr}]$ exhibits frequencies only at $u = 0$. In addition, the modulus of the 2D-FT of $E[Y_{corr}]$ is null for every couple (u, v), except for $u = 0$, where it corresponds to the PSD of the process y at the frequency v, by neglecting the windowing influence along the y-axis. Thus, with the DMA, convolving $\mathscr{F}(W_{DMA})$

[3]Let us recall the interpretation of a 2D-FT analysis. Let two Dirac pulses be located at (u, v) and $(-u, -v)$ in the frequency domain. This leads to a vector of coordinates $(2u, 2v)$ which defines a direction. In the spatial domain, this corresponds to a sinusoid of frequency $\sqrt{(u^2 + v^2)}$ along this direction, and to a constant when looking perpendicular to the direction. When dealing with images, this corresponds to equally-spaced bands along the direction.

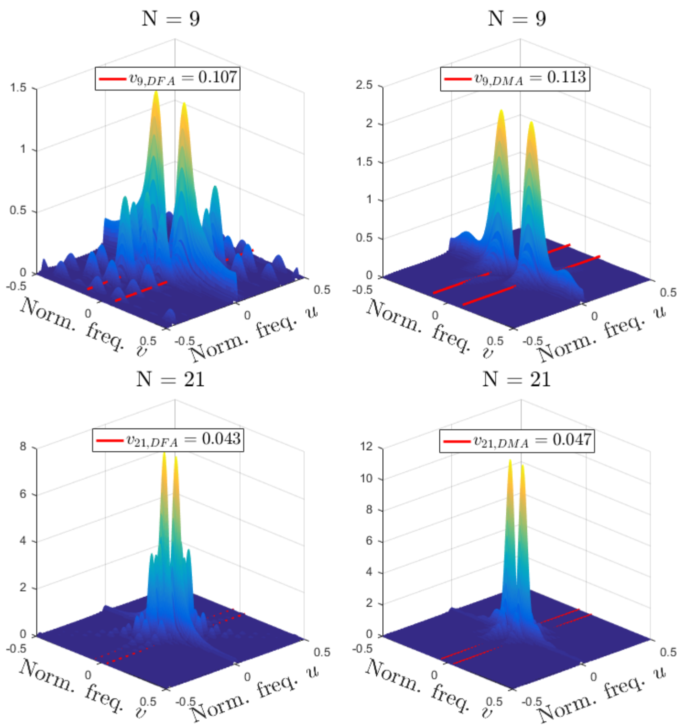

Figure 3. $|\mathscr{F}(W_\bullet)|$ with $\bullet = DFA$ (left) and $\bullet = DMA$ (right) vs. normalized frequencies u and v, for $M = 60$

with $E[Y_{corr}]$ amounts to filtering the process y with a band-pass filter whose central frequency decreases with N. It is confirmed that DMA acts as an *ad hoc* wavelet based analysis. The same comments can be done for the DFA.

Let us now analyze $F_\bullet^2(N)$ in the non-stationary case. The modulus of the 2D-FT of Y_{corr} no longer exhibits frequency only at $u = 0$, but also at other values of the normalized spatial frequency u. Thus, given the above-mentioned properties of $\mathscr{F}(W_{DMA})$, the convolution in (24) amounts to emphasizing the frequency of Y_{corr} located at $v = \pm v_{N,DMA}$. When the DFA is used, there is not only the above phenomenon but others that are due to the various lobes of $\log(|\mathscr{F}(W_{DFA})|)$ appearing at each normalized spatial frequency u multiple of $\frac{1}{N}$.

IV. Conclusions and Perspectives

Our purpose is to help the user have a better understanding of the DFA and the DMA. To this end, we suggest expressing the square of the fluctuation function in a matrix form to show it depends on the instantaneous correlation function of the signal. In the stationary case, our analysis confirms that the methods can be seen as an *ad hoc* wavelet based approach. In the non-stationary case, additional terms appear for the DFA.

References

[1] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch. Self-similarity and long-range dependence through the wavelet lens. In *Theory and Applications of Long-range Dependence*, pages 527–556. Birkhäuser, 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 - Condensed Matter and Complex Systems*, 27, 2:197–200, 2002.

[3] S. Arianos, A. Carbone, and C. Türk. Self-similarity of higher-order moving averages. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 84 4 Pt 2:046113, 2011.

[4] M. Höll and H. Kantz. The relationship between the detrended fluctuation analysis and the autocorrelation function of a signal. *The European Physical Journal B*, 88:1–7, 2015.

[5] M. Höll, H. Kantz, and Y. Zhou. Detrended fluctuation analysis and the difference between external drifts and intrinsic diffusionlike nonstationarity. *Physical review. E*, 94 4-1:042201, 2016.

[6] 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.

[7] 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.

[8] S.-J. Kim, K. Koh, S. Boyd, and D. Gorinevsky. l_1 trend filtering. *SIAM Review*, 51, (2):339–360, 2009.

[9] K. Kiyono. Establishing a direct connection between detrended fluctuation analysis and fourier analysis. *Physical Review E*, 92 (4):042925, 2015.

[10] K. Kiyono. Theory and applications of detrending-operation-based fractal-scaling analysis. *International Conference on Noise and Fluctuations (ICNF)*, pages 1–4, 2017.

[11] 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.

[12] 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*, 6920420, 2018.

[13] D. Osborne. Moving average detrending and the analysis of business cycles. *Oxford Bull. Econom. Statist.*, 57:547–558, 1995.

[14] 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.

[15] 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.

[16] 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.

[17] G. Rilling, P. Flandrin, and P. Gonçalves. Empirical mode decomposition, fractional gaussian noise and hurst exponent estimation. *ICASSP*, pages 489–492, 2005.

[18] R. Sun. Fractional order signal processing: techniques and applications. *Thesis of Master of science in electrical Engineering, Utah state University*, 2007.

[19] 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.

[20] M. S. Taqqu, V. Teverovsky, and W. Willinger. Estimators for long range dependence: an empirical study. *Fractals*, 3, (4):785–788, 1995.

[21] M. P. Tarvainen, P. O. Ranta-aho, and P. A. Karjalainen. An advanced detrending method with application to hrv analysis. *IEEE Transactions on Biomedical Engineering*, 49, 2:172–175, 2002.

[22] 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), 2016.

[23] Y. Tsujimoto, Y. Miki, E. Watanabe, J. Hayano, Y. Yamamoto, T. Nomura, and K. Kiyono. Fast algorithm of long-range cross-correlation analysis using savitzky-golay detrending filter and its application to biosignal analysis. *International Conference on Noise and Fluctuations (ICNF)*, 2017.

[24] L. Xu, P. Ch. Ivanov, K. Hu, Z. Chen, A. Carbone, and H. E. Stanley. Quantifying signals with power-law correlations: A comparative study of detrended fluctuation analysis and detrended moving average techniques. *Phys. Rev. E covering statistical, nonlinear, biological, and soft matter physics*, 71, 5:051101, 2005.