Combining the global trends of DFA or CDFA of different orders

Eric Grivel, Gaetan Colin, Jeremie Soetens, Pierrick Legrand

Version of Record: https://www.sciencedirect.com/science/article/pii/S1051200423000015
Manuscript_7b728d98ed63bc2ad5092946d18eb980

Combining the global trends of DFA or CDFA of different orders

Eric Grivel[a], Gaetan Colin[b], Jérémie Soetens[b], Pierrick Legrand[c]

[a] *Bordeaux University - INP Bordeaux - IMS - UMR CNRS 5218, FRANCE*
[b] *Bordeaux INP, ENSEIRB-MATMECA, France*
[c] *Bordeaux University - IMB UMR CNRS 5251 - ASTRAL team, INRIA, FRANCE*

Abstract

When dealing with the detrended fluctuation analysis (DFA) and its variants such as the higher-order DFA and the continuous DFA (CDFA), removing the mean, integrating the signal and detrending it amount to applying an equivalent linear filter characterized by its frequency response. The different variants can hence be compared and their performances can be better understood by looking at the differences between the frequency responses. In this paper, our contribution is threefold: 1/ the higher-order CDFA is derived using two types of approaches. 2/ we show that the frequency responses of two DFAs of different orders can be expressed from each other. Moreover the frequency responses of the CDFA and the DFA are also related to each other. 3/ we propose to combine the global trends obtained with the DFA or the CDFA of different orders. Our purpose is to show how the frequency responses are modified. Illustrations and comments on ARFIMA processes and Weierstrass functions to evaluate their long range dependence are also given.

Keywords: DFA, CDFA, frequency response, Hurst coefficient.

A great deal of interest has been paid to the detrended fluctuation analysis (DFA) initially proposed by Peng in the 90ies [1] and whose purpose was to characterize the long range dependence (LRD). This type of feature can be useful in many applications, especially in the field of biomedical applications [2–9], but not only. Indeed, many papers were published in the field of econometrics [10, 11], meteorology [12, 13], or geophysics [14].

The DFA has been designed to estimate the LRD of signals exhibiting fluctuations around a trend varying over time. Two main steps characterize this approach: 1/ The estimation of the profile trend, *i.e.* the estimation of the trend of the integrated centered signal. 2/ The estimation of the variance of the detrended profile, decimated by a factor equal to N. This corresponds to the definition of the square of the so-called fluctuation function. By selecting several values for N, the estimated variances are combined to deduce the LRD of the signal.

For the last years, several research works were conducted by various authors having different skills in mathematics, statistics and signal processing. Sometimes, some of them have addressed the same type of problems independently at the same period. First of all, various contributions were related to variants of the DFA. Indeed, as one of the first steps of the DFA consists in estimating the global trend of the profile, different authors have proposed an alternative to the discontinuous piecewise linear function used in the standard DFA. Thus, when dealing with the higher-order DFA, polynomials of degree larger than 1 are used to model the local trends. However, the latter may still be discontinuous. To avoid or to weaken the above feature, the adaptive fractal analysis (AFA) [15], the regularized DFA (RDFA) [16] and the continuous DFA (CDFA) [17] were developed. All these methods can hence belong to a first family of methods characterized by an *a priori* model of the local trends. A second family of approaches has been proposed and consists in using a low-pass linear filter. This is the case of the detrending moving average (DMA) [18] if the impulse response is finite but there are also some other methods like the backward DMA (BDMA), the forward DMA (FDMA), the weighted DMA of order l

(WDMA-l) or the centered weighted DMA (CWDMA) [19]. A third family including the higher-order DMA [20], or equivalently the method based on the scatterplot smoothing (LOESS), and the one based on the locally weighted scatterplot smoothing (LOWESS) [21] makes it possible to bridge the gap between the above two families. See Figure 1.

It should be noted that proposing new variants of the DFA is not the only aspect that has been addressed. Thus, the way to extend the approach to the analysis of multifractal time series has been for instance studied in [22]. As the computational cost may be of real interest in some applications, fast versions have been developed [23]. Even if it is not easy to provide an exhaustive state of the art on a method proposed thirty years ago, the reader can also refer to various mathematical analyzes that were also led to better understand the behavior of the DFA and its variants. Thus, the properties of the square of the fluctuation function has been addressed by several authors in various papers [24–37]. In [38], the distribution of a detrended integrated Gaussian white noise is studied. When the DMA is used, the residue distribution is Gaussian whereas it is not the case when the DFA is applied. Finally, in [39–41], complementary analysis for short-memory processes were conducted.

In [16], when dealing with wide-sense-stationary processes, the first steps of the DFA, from the integration step to the detrending one, can be modeled by an "equivalent" linear filter whose impulse response (and consequently the frequency response) depend on the way to estimate the trend. Better understanding the properties and the performance of the variants of the DFA can be done by looking at the frequency response of the equivalent filter. Thus, for the DFA, the DMA and the CDFA, the null frequency is always rejected, which is consistent with the purpose of detrending. Moreover, we showed that when $N = 3$, the filters associated with the DFA and the DMA are high-pass whereas they are band-pass for larger values of N. Regarding the CDFA, the filter is always band-pass. For the three methods, we noticed that the larger N, the smaller the -3dB bandwidth of the filter and the spikier the resonances of the frequency responses. The resonances also move to low frequencies when

N increases. For any value of N, the CDFA always provides the spikiest and lowest resonance. In this new paper, our contribution is threefold. Firstly, as done for the DFA a few years ago, the higher-order CDFA is derived. Two types of approaches are proposed: the first one is based on a similar reasoning as the one presented in [16], except that the degree of the polynomial is now larger than or equal to 1. Then, an alternative way is presented based on Lagrange multipliers. Secondly, we aim at searching if there is a link between two DFAs of different orders. We analyze how the frequency responses of the equivalent filters can be expressed from each other. Moreover we analyze if the frequency responses of the CDFA and the DFA are also related to each other. Finally, instead of *a priori* selecting the degree of the polynomial of the local trends and estimating the corresponding model parameters in the LS sense, we propose to *a posteriori* combine two global trends obtained with DFAs or CDFAs of different orders. To this end, we analyze how the frequency responses of the equivalent filter are modified. Illustrations on autoregressive fractionally integrated moving average (ARFIMA) processes and Weierstrass functions as well as comments are also given. A presentation of the variants of the DFA and a summary of our contributions are given Figure 1.

The remainder of this paper is organized as follows: Section 1 deals with the main steps of the DFA using a matrix form. Then, links between two DFAs of different orders is addressed. In Section 2, the higher-order CDFA is studied. Two methods are proposed to derive the expression of the global trend. More particularly, we reformulate the estimation of the parameters of the continuous trend used in the CDFA in terms of optimisation under constraints using Lagrange multipliers. We will see that this makes it possible to create a direct link between the CDFA and the DFA. The links between two CDFAs of different orders is also presented. Then, in Section 3, combining the approaches is analyzed. Finally, A comparison on a set of predefined signals is presented in Section 4.

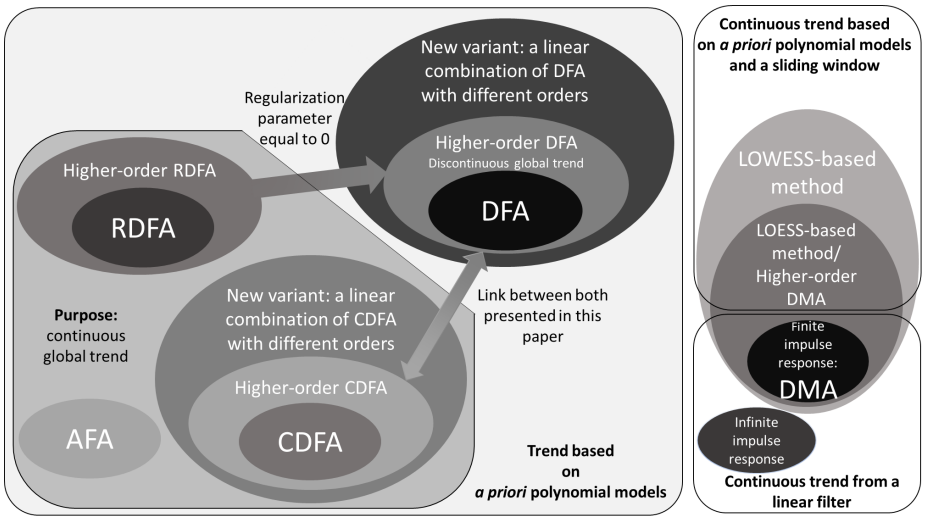


Figure 1: Variants of the DFA and our contribution in this paper

1. Links between two DFAs of different orders

As the FA was sensitive to signals exhibiting fluctuations around a trend varying over time, the detrended fluctuation analysis (DFA) was developed in [1]. Its purpose is to estimate the trend of the integrated centered process, and then to analyze the LRD of the process when the trend has been removed. In the standard version of the DFA, the trend is defined as a piecewise linear approximation of the integrated centered process while it is a polynomial of degree d in the higher-order version.

1.1. Notations

- $\mathbb{1}_{j\times k}$ and $\mathbf{0}_{j\times k}$ are matrices of size $j \times k$ filled with 1s and 0s, respectively.
- $diag([.], l)$ is a matrix whose elements are null except the l^{th} diagonal which is equal to [.].
- $J_j = I_j - \frac{1}{j}\mathbb{1}_{j\times j}$ with I_j the identity matrix of size j.
- T_l is a $N \times 1$ vector storing the values of the l^{th} local trend $t_l(n)$.
- Y and Y_{int} are two column vectors storing respectively the samples $\{y(n)\}_{n=1,\ldots,M}$

and $\{y_{int}(n)\}_{n=1,...,M}$ that are related as follows:

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

with $H_M = \sum_{r=0}^{M-1} diag(\mathbb{1}_{1\times M-r}, -r)$ a lower triangular matrix filled with 1s.

• $A(i:j, k:l)$ is the part of the matrix A corresponding to the elements belonging to the rows i to j and to the columns k to l.

• Given the following matrix of size $j \times M$:

$$C_{j,k} = [\mathbf{0}_{j\times k} \quad I_j \quad \mathbf{0}_{j\times(M-(j+k))}] \tag{2}$$

the first LN elements of the vector Y_{int} can be expressed as follows:

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

1.2. Steps of the DFA and its higher-order variant with some comments

The DFA and its higher-order variant operate with following steps when M consecutive samples $\{y(m)\}_{m=1,...,M}$ are available.

Note that when the subscript $_{\mathbf{1},d}$ is used in a notation, this means that the quantity is related to the d^{th}-order DFA method, also called DFAd.

- **Step 1, preprocessing:** The profile y_{int} is obtained by centering and integrating y:

$$y_{int}(m) = \sum_{i=1}^{m} (y(i) - \mu_y) \tag{4}$$

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

- **Step 2, profile trend estimation:** The trend of the profile is estimated by splitting the profile into L non-overlapping segments of length N, whose samples are $\{y_{int,l}(n)\}_{l=1,...,L}$ with $n \in [\![1; N]\!]$. Only the first LN samples of the sequences are kept. Using a vector form, the d^{th}-degree polynomial trend of the l^{th} segment can be expressed $\forall l \in [\![1; L]\!]$ by the following product:

$$\mathrm{T}_{\mathbf{1},d,l} = \mathsf{A}_{\mathbf{1},d,l}\theta_{d,l} \tag{5}$$

where the parameter vector is $\theta_{d,l} = \left[a_{0,l}\ a_{1,l}\ ...\ a_{d,l}\right]^T$. When $d = 1$, the local trend is a linear function, $a_{1,l}$ being the local slope and $a_{0,l}$ the vertical intercept. In addition, $\mathsf{A}_{\mathbf{1},d,l}$ is a $N \times (d+1)$ matrix defined as follows:

$$\mathsf{A}_{\mathbf{1},d,l} = \begin{bmatrix} 1 & \cdots & ((l-1)N+1)^d \\ \vdots & & \vdots \\ 1 & \cdots & (lN)^d \end{bmatrix} = \begin{bmatrix} \mathsf{C}_{0,l} & \cdots & \mathsf{C}_{d,l} \end{bmatrix} \tag{6}$$

where $\mathsf{C}_{c,l}$ is the c^{th} column of the matrix $\mathsf{A}_{\mathbf{1},d,l}$, with $c = 0, ..., d$.

Given the parameter vector $\Theta_{\mathbf{1},d} = \left[\theta_{d,1}^T \dots \theta_{d,L}^T\right]^T$ of size $(d+1)L \times 1$, and the following block diagonal matrix of size $LN \times (d+1)L$:

$$A_{\mathbf{1},d} = diag\begin{bmatrix} \mathsf{A}_{1,d,1} & \dots & \mathsf{A}_{\mathbf{1},d,L} \end{bmatrix} \tag{7}$$

the parameters of the local trends can be deduced by minimizing the least-square (LS) criterion $\left|\left|C_{LN,0}Y_{int} - A_{\mathbf{1},d}\Theta_{\mathbf{1},d}\right|\right|^2$. For $d \geq 1$, the parameter vector estimate has the following expression:

$$\hat{\Theta}_{\mathbf{1},d} = (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} A_{\mathbf{1},d}^T C_{LN,0} Y_{int} \tag{8}$$

Therefore, combining (1) and (8), the global trend vector $T_{\mathbf{1},d}$ obtained with the d^{th}-order DFA is given by:

$$\begin{aligned} T_{\mathbf{1},d} &= A_{\mathbf{1},d}\hat{\Theta}_{\mathbf{1},d} = A_{\mathbf{1},d}(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} A_{\mathbf{1},d}^T C_{LN,0} H_M J_M Y \\ &= P_{\mathbf{1},d} C_{LN,0} H_M J_M Y \end{aligned} \tag{9}$$

In the equation above, the trend vector $T_{\mathbf{1},d}$ is deduced from the orthogonal projector:

$$P_{\mathbf{1},d} = \text{diag}\left(\left[\mathsf{P}_{\mathbf{1},d,1}, \cdots, \mathsf{P}_{\mathbf{1},d,L}\right]\right) \tag{10}$$

with $\{\mathsf{P}_{\mathbf{1},d,l}\}_{l=1,\dots,L}$ the orthogonal projectors onto the space spanned by the column vectors of $\{\mathsf{A}_{\mathbf{1},d,l}\}_{l=1,\dots,L}$ defined as follows:

$$\mathsf{P}_{\mathbf{1},d,l} = \mathsf{A}_{\mathbf{1},d,l}\left(\mathsf{A}_{\mathbf{1},d,l}{}^T \mathsf{A}_{\mathbf{1},d,l}\right)^{-1} \mathsf{A}_{\mathbf{1},d,l}{}^T \tag{11}$$

Note that in a previous paper [16], we showed that $\forall l \in [\![1; L]\!]$:

$$\mathsf{P}_{\mathbf{1},d,l} = \mathsf{P}_{\mathbf{1},d,1} \tag{12}$$

- **Step 3, computing the fluctuation function:** Given (9), the residue vector $R_{\mathbf{1},d}$, *i.e.* the expression of the difference between the profile and its trend storing discontinuous polynomials of degree d, can be expressed from the signal vector Y as follows:

$$R_{\mathbf{1},d} = C_{LN,0}\mathrm{H}_M J_M \mathrm{Y} - T_{\mathbf{1},d} = B_{\mathbf{1},d} Y \tag{13}$$

where:

$$\begin{aligned} B_{\mathbf{1},d} &= \big(I_{LN} - A_{\mathbf{1},d}(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} A_{\mathbf{1},d}^T\big) C_{LN,0} H_M J_M \\ &= \big(I_{LN} - P_{\mathbf{1},d}\big) C_{LN,0} H_M J_M \end{aligned} \tag{14}$$

The next step is to define the square of the fluctuation function $F^2_{\mathbf{1},d}(N)$ which can be also interpreted as the power of the residue. By introducing the matrix[1] of size $M \times M$:

$$\Gamma_{\mathbf{1},d} = \frac{1}{LN} B_{\mathbf{1},d}^T B_{\mathbf{1},d} \tag{15}$$

$F^2_{\mathbf{1},d}(N)$ can be written this way:

$$F^2_{\mathbf{1},d}(N) = Tr\left(\Gamma_{\mathbf{1},d} Y Y^T\right) \tag{16}$$

- **Step 4, iterating over different values of N:** Steps 2 and 3 are repeated for different values of N, *e.g.* N_1, N_2, ..., N_{max} usually selected in the interval $[3; \lfloor M/2 \rfloor]$, where $\lfloor\rfloor$ is the floor function. It should be noted that each pair $(\log(F_{\mathbf{1},d}(N)), \log(N))$ satisfies :

$$\log(F_{\mathbf{1},d}(N)) = b + \alpha \log(N) + \epsilon(N), \quad \text{for } N = N_1, ..., N_{max} \tag{17}$$

with b the vertical intercept and $\epsilon(N)$. the error variable.

- **Step 5, estimation of the Hurst exponent H:** The final step aims at using the pairs $(log\left(F_{\mathbf{1},d}(N)\right), log(N))$ obtained in the previous step and searching the straight line that fits them. Its slope α, which is no longer equal to H but to $H + 1$ due to the integration, is estimated in the LS sense.

[1]Note that the expression of $\Gamma_{\mathbf{1},d}$ could be simplified since $(I_{LN} - P_{\mathbf{1},d})^T (I_{LN} - P_{\mathbf{1},d}) = (I_{LN} - P_{\mathbf{1},d})$ due to the property of the orthogonal projector. However, we do not change this expression as it will be useful to define of the equivalent filter in the rest of the document.

Given the above description of the DFA and its higher-order variant, we proposed in [16] to define the equivalent filter of the processing chain from the step 1 to step 3 when dealing with wide-sense stationary processes (w.s.s.). The latter can be defined by its frequency response[2] $\Psi_{\mathbf{1},d}(f_n)$, that can be deduced by computing the Fourier transform of the sequence $\{Tr(\Gamma_{\mathbf{1},d}, r)\}_{r=-LN+1,\ldots,LN-1}$, where $Tr(.,r)$ denotes the sum of the elements of the r^{th} sub-diagonal of the matrix. Indeed, by taking the expectation of (16), one obtains:

$$\begin{aligned} E[F^2_{\mathbf{1},d}(N)] &= \mathscr{F}\mathscr{T}^{-1}\Big(\big(\sum_{r=-LN+1}^{LN-1} Tr(\Gamma_{\mathbf{1},d}, r)e^{-j2\pi f_n r}\big)S_{yy}(f_n)\Big)\big|_{\tau=0} \\ &= \mathscr{F}\mathscr{T}^{-1}\Big(\Psi_{\mathbf{1},d}(f_n)S_{yy}(f_n)\Big)\big|_{\tau=0} \end{aligned} \tag{18}$$

where $\mathscr{F}\mathscr{T}^{-1}$ is the inverse Fourier transform and $S_{yy}(f_n)$ is the power spectral density (PSD) of the signal under study at the normalized frequency f_n.

This frequency response is of interest as it makes it possible to compare the different variants of the DFA. We also analyzed the differences between our approach and the works done in [30] where the square of the fluctuation function is approximated by a weighted sum of the estimates of the correlation function of the signal and [29]. Although Kiyono's approach [29] led to frequency responses the closest to ours for most of the values of N, the frequency responses obtained by our colleagues are different from ours, especially in high frequency and for small values of N. This is due to the fact that some approximations are made by the authors in their mathematical developments while the expectation is used in our case. When the process is no longer w.s.s., we suggested expressing the fluctuation function as a 2D Fourier transform of the product of two matrices: the first one defined from the instantaneous correlations of the signal $\{y(k)y(k+r)\}_{r=-LN+1,\ldots,LN-1}$ and the second defined from the element of $\Gamma_{\mathbf{3},d}$ [16]. In that case, the comparison between the different methods can be done by analyzing the properties of the 2D-FT of the second matrix.

In the next section, we propose to compare the frequency response of the equivalent filters related to DFAs of different orders. To help the reader, we propose

[2] f_n denotes the normalized frequency.

to present the different results by using a proposition-proof structure.

1.3. Link between two frequency responses of the equivalent filters related to DFAs of different orders

In this section, our purpose is to find a link between the DFA of order d_1 and the DFA of order $d_2 > d_1$. More particularly, our goal is to find a relation between the frequency responses of the equivalent filters related to the DFAs of orders d_1 and d_2. To this end, we first show that the orthogonal projector related to the DFA of order d_2 is the sum of two projectors, one of which is the orthogonal projector related to the DFA of order d_1. Then, we deduce a relation between the matrices $\Gamma_{\mathbf{1},d_2}$ and $\Gamma_{\mathbf{1},d_1}$. At that stage, we will be able to find a link between $\Psi_{\mathbf{1},d_2}(f_n)$ and $\Psi_{\mathbf{1},d_1}(f_n)$.

1.3.1. Relation between the orthogonal projectors

In the following, we propose to find a link between the orthogonal projector $P_{\mathbf{1},d_2}$ and the orthogonal projector $P_{\mathbf{1},d_1}$ whose definition is given in (10). More particularly, we propose the following theorem:

Theorem 1. *The orthogonal projector $P_{\mathbf{1},d_2}$ can be expressed as the sum of two terms: the orthogonal projector $P_{\mathbf{1},d_1}$ and a second term denoted as $P_{\mathbf{1},d_1\Rightarrow d_2}$, where the subscript $_{d_1\Rightarrow d_2}$ is used to show the transition from the order d_1 to the order d_2.*

Proof. As a preamble, let us consider the following block diagonal matrices $\{A_{1,d}\}_{d=d_1,d_2}$ whose definition is given in (7). They are diagonal matrices respectively defined by the set of matrices $\{\mathsf{A}_{\mathbf{1},d_1,l}\}_{l=1,\dots,L}$ and $\{\mathsf{A}_{\mathbf{1},d_2,l}\}_{l=1,\dots,L}$. More particularly, $\forall l \in [\![1; L]\!]$, one has:

$$\mathsf{A}_{\mathbf{1},d_2,l} = \begin{bmatrix} \mathsf{A}_{\mathbf{1},d_1,l} & \mathsf{C}_{d_1+1,l} & \cdots & \mathsf{C}_{d_2,l} \end{bmatrix} = \begin{bmatrix} \mathsf{A}_{\mathbf{1},d_1,l} & \mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} \end{bmatrix} \tag{19}$$

where $\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}$ is the matrix storing the columns $\mathsf{C}_{d_1+1,l}$ to $\mathsf{C}_{d_2,l}$.
Then, given the definition of the orthogonal projector (11), let us rewrite the product $\mathsf{A}^T_{\mathbf{1},d_2,l}\mathsf{A}_{\mathbf{1},d_2,l}$ and give an expression of its inverse. Finally, the result will

be pre- and post-multiplied by $\mathsf{A}_{\mathbf{1},d_2,l}$ and $\mathsf{A}^T_{\mathbf{1},d_2,l}$ respectively.
Thus, given (19), one has $\forall l \in [\![1; L]\!]$:

$$\begin{aligned}\mathsf{A}^T_{\mathbf{1},d_2,l}\mathsf{A}_{\mathbf{1},d_2,l} &= \begin{bmatrix} \mathsf{A}^T_{\mathbf{1},d_1,l} \\ \mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l} \end{bmatrix} \begin{bmatrix} \mathsf{A}_{\mathbf{1},d_1,l} & \mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} \end{bmatrix} \\ &= \begin{bmatrix} \mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l} & \mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} \\ \mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\mathsf{A}_{\mathbf{1},d_1,l} & \mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} \end{bmatrix}\end{aligned} \tag{20}$$

Then, let us take advantage of one result of the block matrix inversion[3] one gets:

$$\begin{aligned}&\left(\mathsf{A}^T_{\mathbf{1},d_2,l}\mathsf{A}_{\mathbf{1},d_2,l}\right)^{-1} = \\ &\begin{bmatrix} \left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1} & \mathbf{0}_{(d_1+1)\times(d_2-d_1)} \\ \mathbf{0}_{(d_2-d_1)\times(d_1+1)} & \mathbf{0}_{(d_2-d_1)\times(d_2-d_1)} \end{bmatrix} + \begin{bmatrix} -\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1}\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} \\ I_{d_2-d_1} \end{bmatrix} \\ &\times \left[\mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} - \mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\mathsf{A}_{\mathbf{1},d_1,l}\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1}\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}\right]^{-1} \\ &\times \left[-\mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\mathsf{A}_{\mathbf{1},d_1,l}\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1} \quad I_{d_2-d_1}\right]\end{aligned} \tag{23}$$

[3]Let us consider a matrix partitioned into four blocks denoted as A, B, C and D. In this case, its inverse can be expressed as:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix} \tag{21}$$

where $(D-CA^{-1}B)^{-1}$ is the Schur complement of the block A. The inverse can be rewrittten as follows:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix} \left[D - CA^{-1}B\right]^{-1} \begin{bmatrix} -CA^{-1} & I \end{bmatrix} \tag{22}$$

The projection matrix $\mathsf{P}_{\mathbf{1},d_2,l}$ is given $\forall l \in [\![1;L]\!]$ by:

$$\mathsf{P}_{\mathbf{1},d_2,l} = \begin{bmatrix}\mathsf{A}_{\mathbf{1},d_1,l} & \mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}\end{bmatrix} \begin{bmatrix}\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1} & \mathbf{0}_{(d_1+1)\times(d_2-d_1)} \\ \mathbf{0}_{(d_2-1)\times(d_1+1)} & \mathbf{0}_{(d_2-d_1)\times(d_2-d_1)}\end{bmatrix} \begin{bmatrix}\mathsf{A}^T_{\mathbf{1},d_1,l} \\ \mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\end{bmatrix} \tag{24}$$

$$+ \begin{bmatrix}\mathsf{A}_{\mathbf{1},d_1,l} & \mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}\end{bmatrix} \begin{bmatrix}-\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1}\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} \\ I_{d_2-d_1}\end{bmatrix}$$

$$\times \left[\mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} - \mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\mathsf{A}_{\mathbf{1},d_1,l}\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1}\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}\right]^{-1}$$

$$\times \begin{bmatrix}-\mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\mathsf{A}_{\mathbf{1},d_1,l}\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1} & I_{d_2-d_1}\end{bmatrix} \begin{bmatrix}\mathsf{A}^T_{\mathbf{1},d_1,l} \\ \mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\end{bmatrix}$$

The above equality (24) can be simplified as follows:

$$\begin{aligned}\mathsf{P}_{\mathbf{1},d_2,l} =& \mathsf{A}_{\mathbf{1},d_1,l}\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1}\mathsf{A}^T_{\mathbf{1},d_1,l} \\ &+ \left(-\mathsf{A}_{\mathbf{1},d_1,l}\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1}\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} + \mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}\right) \\ &\quad\times \left(\mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\left(I_N - \mathsf{A}_{\mathbf{1},d_1,l}\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1}\mathsf{A}^T_{\mathbf{1},d_1,l}\right)\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}\right)^{-1} \\ &\quad\times \left(-\mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\mathsf{A}_{\mathbf{1},d_1,l}\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1}\mathsf{A}^T_{\mathbf{1},d_1,l} + \mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\right)\end{aligned} \tag{25}$$

As $\mathsf{P}_{\mathbf{1},d_1,l} = \mathsf{A}_{\mathbf{1},d_1,l}\left(\mathsf{A}^T_{\mathbf{1},d_1,l}\mathsf{A}_{\mathbf{1},d_1,l}\right)^{-1}\mathsf{A}^T_{\mathbf{1},d_1,l}$, (25) becomes:

$$\begin{aligned}\mathsf{P}_{\mathbf{1},d_2,l} =& \mathsf{P}_{\mathbf{1},d_1,l} + \left(-\mathsf{P}_{\mathbf{1},d_1,l}\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} + \mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}\right) \\ &\times \left(\mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\left(I_N - \mathsf{P}_{\mathbf{1},d_1,l}\right)\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}\right)^{-1}\left(-\mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\mathsf{P}_{\mathbf{1},d_1,l} + \mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\right)\end{aligned} \tag{26}$$

or equivalently:

$$\begin{aligned}\mathsf{P}_{\mathbf{1},d_2,l} =& \mathsf{P}_{\mathbf{1},d_1,l} + \left(I_N - \mathsf{P}_{\mathbf{1},d_1,l}\right)\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l} \\ &\times \left(\mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\left(I_N - \mathsf{P}_{\mathbf{1},d_1,l}\right)^T\left(I_N - \mathsf{P}_{\mathbf{1},d_1,l}\right)\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}\right)^{-1}\mathsf{D}^T_{\mathbf{1},d_1\Rightarrow d_2,l}\left(I_N - \mathsf{P}_{\mathbf{1},d_1,l}\right)^T\end{aligned} \tag{27}$$

since $\mathsf{P}_{\mathbf{1},d_1,l} = \mathsf{P}^T_{\mathbf{1},d_1,l}$ and $\left(I_N - \mathsf{P}_{\mathbf{1},d_1,l}\right)^T\left(I_N - \mathsf{P}_{\mathbf{1},d_1,l}\right) = \left(I_N - \mathsf{P}_{\mathbf{1},d_1,l}\right)$ due to the property of the orthogonal projector $\mathsf{P}_{\mathbf{1},d_1,l}$.

Therefore, the projector $\mathsf{P}_{\mathbf{1},d_2,l}$ can be expressed as the sum of two terms: the first one is the projector $\mathsf{P}_{\mathbf{1},d_1,l}$ while $\mathsf{P}_{\mathbf{1},d_1\Rightarrow d_2,l}$ is the orthogonal projection onto the space spanned by the columns of $\left(I_N - \mathsf{P}_{\mathbf{1},d_1,l}\right)\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}$. Hence (27) becomes $\forall l \in [\![1;L]\!]$:

$$\mathsf{P}_{\mathbf{1},d_2,l} = \mathsf{P}_{\mathbf{1},d_1,l} + \mathsf{P}_{\mathbf{1},d_1\Rightarrow d_2,l} \tag{28}$$

Let us now show that $\forall l = 1, ..., L$:

$$\mathsf{P}_{\mathbf{1},d_1\Rightarrow d_2,l} = \mathsf{P}_{\mathbf{1},d_1\Rightarrow d_2,1} \tag{29}$$

On the one hand, $(I_N - \mathsf{P}_{\mathbf{1},d_1,l})\,\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}$ is a matrix equal to $\begin{bmatrix}(I_N - \mathsf{P}_{\mathbf{1},d_1,l})\,\mathsf{C}_{d_1+1,l} & \cdots & (I_N - \mathsf{P}_{\mathbf{1},d_1,l})\,\mathsf{C}_{d_2,l}\end{bmatrix}$.
On the other hand, for $k = d_1 + 1, ..., d_2$, one has:

$$\mathsf{C}_{k,l} = \sum_{q=0}^{k} \binom{k}{q} ((l-1)N)^{k-q} \mathsf{C}_{q,1} \tag{30}$$

Therefore, as $\forall q \leqslant d_1, \mathsf{C}_{q,1} \in \mathrm{Im}(\mathsf{P}_{\mathbf{1},d_1,l})$, one can deduce that:

$$(I_N - \mathsf{P}_{\mathbf{1},d_1,l})\,\mathsf{C}_{k,l} = \sum_{q=d_1+1}^{k} \binom{k}{q} ((l-1)N)^{k-q} (I_N - \mathsf{P}_{\mathbf{1},d_1,1})\,\mathsf{C}_{q,1} \tag{31}$$

This means that the columns of $(I_N - \mathsf{P}_{\mathbf{1},d_1,l})\,\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}$ can be expressed as a linear combination of the columns of $(I_N - \mathsf{P}_{\mathbf{1},d_1,l})\,\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,1}$. As a consequence, the orthogonal projection onto the columns of the matrix $(I_N - \mathsf{P}_{\mathbf{1},d_1,1})\,\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,l}$ is the same as the orthogonal projection onto the columns of the matrix $(I_N - \mathsf{P}_{\mathbf{1},d_1,1})\,\mathsf{D}_{\mathbf{1},d_1\Rightarrow d_2,1}$.
Introducing the block diagonal matrix:

$$P_{1,d_1\Rightarrow d_2} = diag\left([\mathsf{P}_{\mathbf{1},d_1\Rightarrow d_2,1}, \cdots, \mathsf{P}_{\mathbf{1},d_1\Rightarrow d_2,L}]\right) \tag{32}$$

and given (10) and (12), one obtains the following equality:

$$P_{\mathbf{1},d_2} = P_{\mathbf{1},d_1} + P_{\mathbf{1},d_1\Rightarrow d_2} \tag{33}$$

This means that the orthogonal projector related to the DFA of order d_2 is the sum of two projectors, one of which is the orthorgonal projector related to the DFA of order d_1. □

1.3.2. Relation between the residue vectors and the frequency responses of the equivalent filters

In the following, we propose to find a link between the frequency response of the equivalent filter $\Psi_{\mathbf{1},d_2}(f_n)$ and the frequency response of the equivalent filter $\Psi_{\mathbf{1},d_1}(f_n)$. More particularly, we propose the following theorem:

Theorem 2. *$\Psi_{\mathbf{1},d_2}(f_n)$ which is the frequency response of the equivalent filter related to the DFA of order d_2 can be expressed as the difference between two terms: $\Psi_{\mathbf{1},d_1}(f_n)$ which is the frequency response of the equivalent filter related to the DFA of order d_1 and a second term $\Psi_{\mathbf{1},d_1\Rightarrow d_2}(f_n)$ which is defined from $P_{\mathbf{1},d_1\Rightarrow d_2}$.*

Proof. Given the final result (33) of the previous Section 1.3.1, let us express the residue vector $B_{\mathbf{1},d_2}$ as a function of the residue vector $B_{\mathbf{1},d_1}$:

$$\begin{aligned} B_{\mathbf{1},d_2} &= (I_N - P_{\mathbf{1},d_2})C_{LN,0}H_M J_M \\ &= (I_N - P_{\mathbf{1},d_2} - P_{\mathbf{1},d_1\Rightarrow d_2})C_{LN,0}H_M J_M \\ &= (I_N - P_{\mathbf{1},d_1})C_{LN,0}H_M J_M - P_{1,d_1\Rightarrow d_2}C_{LN,0}H_M J_M \\ &= B_{\mathbf{1},d_1} - B_{\mathbf{1},d_1\Rightarrow d_2} \end{aligned} \tag{34}$$

where $B_{\mathbf{1},d_1\Rightarrow d_2} = P_{\mathbf{1},d_1\Rightarrow d_2}C_{LN,0}H_M J_M$.
The next step consists in deducing the expression of the matrix $\Gamma_{\mathbf{1},d_2}$. We will see that it can be rewritten as a function of $\Gamma_{\mathbf{1},d_1}$. Thus, taking advantage of the idempotent property of the orthogonal projector and using (34), one gets:

$$\Gamma_{\mathbf{1},d_2} = \frac{1}{LN}B_{\mathbf{1},d_2}^T B_{\mathbf{1},d_2} \tag{35}$$

$$\begin{aligned} &= \frac{1}{LN}J_M^T H_M^T C_{LN,0}^T (I_{LN} - P_{\mathbf{1},d_2})C_{LN,0}H_M J_M \\ &= \frac{1}{LN}J_M^T H_M^T C_{LN,0}^T (I_{LN} - P_{\mathbf{1},d_1})C_{LN,0}H_M J_M \\ &- \frac{1}{LN}J_M^T H_M^T C_{LN,0}^T P_{\mathbf{1},d_1\Rightarrow d_2}C_{LN,0}H_M J_M \end{aligned} \tag{36}$$

By introducing the matrix $\Gamma_{\mathbf{1},d_1\Rightarrow d_2}$ equal to $\frac{1}{LN}J_M^T H_M^T C_{LN,0}^T P_{\mathbf{1},d_1\Rightarrow d_2}C_{LN,0}H_M J_M$, one obtains:

$$\Gamma_{\mathbf{1},d_2} = \Gamma_{\mathbf{1},d_1} - \Gamma_{\mathbf{1},d_1\Rightarrow d_2} \tag{37}$$

Given the definition of $\Psi_{\mathbf{1},d_2}(f_n)$ based on the Fourier transform of the traces of the sub-diagonals of the matrix $\Gamma_{\mathbf{1},d_2}$, one gets:

$$\Psi_{\mathbf{1},d_2}(f_n) = \Psi_{\mathbf{1},d_1}(f_n) - \Psi_{\mathbf{1},d_1\Rightarrow d_2}(f_n) \tag{38}$$

where $\Psi_{\mathbf{1},d_1\Rightarrow d_2}(f_n)$ is similarly deduced from $\Gamma_{\mathbf{1},d_1\Rightarrow d_2}$ as $\Psi_{\mathbf{1},d_2}(f_n)$ and $\Psi_{\mathbf{1},d_1}(f_n)$ are respectively obtained from $\Gamma_{\mathbf{1},d_2}$ and $\Gamma_{\mathbf{1},d_1}$. □

1.3.3. Comments on the frequency responses of DFA of different orders

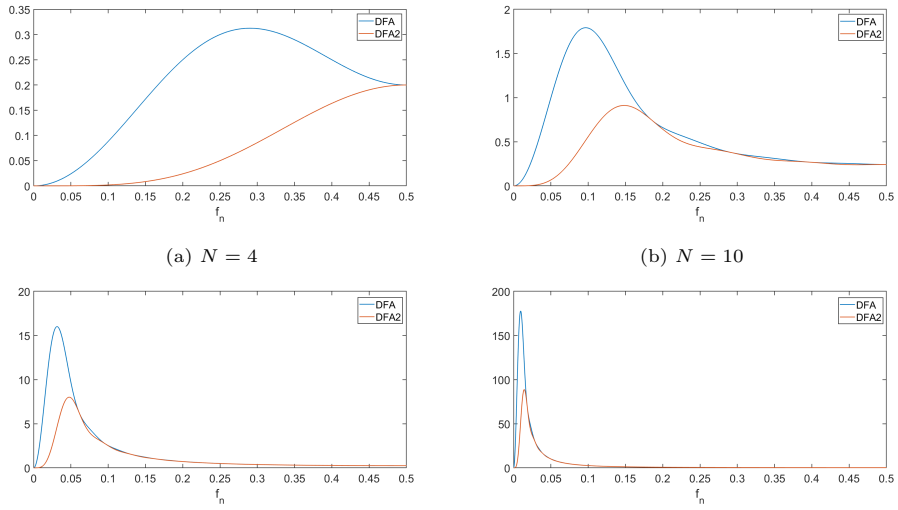
In the Figure 2, the frequency responses of the standard DFA and the 2^{nd}-order DFA for $N = 4$, $N = 10$, $N = 30$ and $N = 100$ are given.

As $\Gamma_{\mathbf{1},d_1\Rightarrow d_2}$ is the result of a product by its transpose, $\Psi_{\mathbf{1},d_1\Rightarrow d_2}(f_n)$ is always positive or null. So, one has $\forall f_n \in \left[0, \frac{1}{2}\right]$:

$$\Psi_{\mathbf{1},d_1}(f_n) \geqslant \Psi_{\mathbf{1},d_2}(f_n) \tag{39}$$

Consequently, the larger the order of the DFA, the smaller the gain for a given normalized frequency f_n. This also means that the larger the order of the DFA is, the smaller the square of the fluctuation function will be for a given value of N. This result is not surprising. Indeed, when the order of the DFA increases, the local trends are modeled by polynomials whose degree increases, necessarily leading to a residue power that is smaller.

In addition, there may be frequencies for which the frequency responses are equal. They can be easily obtained by searching the values of the normalized frequency for which $\Psi_{\mathbf{1},d_1\Rightarrow d_2}(f_n)$ is null.



(a) $N = 4$ (b) $N = 10$

(c) $N = 30$ (d) $N = 100$

Figure 2: Frequency responses of the DFA and the 2^{nd}-order DFA (DFA2) for $N = 4$, $N = 10$, $N = 30$ and $N = 100$

Finally, we can notice that the frequency response of the equivalent filter is always null for the null frequency whatever the order of the DFA. This result is consistent with the purpose of detrending.

For $N > 4$, for a given order, the equivalent filter is band-pass and its frequency response exhibits a resonance. This phenomenon can be explained by the fact that the equivalent filter is the combination of the integration step whose frequency response is low-pass (and more and more low-pass when N increases) and the detrending step which tends to reject more and more the low frequencies when the order increases.

Consequently, for a given order, when N increases, the resonance frequency becomes smaller and smaller and spikier and spikier. When comparing the standard DFA and the 2^{nd}-order DFA, one can see that the resonance frequency of the filter associated with the DFA of higher order is always higher than the resonance frequency of the filter associated with the DFA of lower order.

Let us now start a short discussion when applying these approaches on autoregressive fractionally integrated moving average (ARFIMA) process of order[4] (p, δ, q). This process is a generalization of the ARIMA process where δ is no longer an integer but can take a real value. It can be seen as the output of a linear filter whose input is a fractionally integrated (FI) white noise. Its main properties are the following: If $W(z)$ denotes the z-transform of the FI white noise w_n and $U(z)$ the z-transform of the driving process u_n, assumed to be a zero-mean white noise with variance σ_u^2, one has:

$$W(z)(1 - z^{-1})^{\delta} = U(z) \tag{40}$$

Thus, given (40), the PSD $S_{fi}(f_n)$ of the FI white noise can be expressed as follows :

$$S_{fi}(f_n) = \sigma_u^2 |2sin(\pi f_n)|^{-2\delta} \tag{41}$$

Therefore, the PSD tends to infinity (resp. 0) when f_n tends to 0 if $\delta > 0$ (resp. $\delta < 0$). For any δ, when $f_n = \frac{1}{\pi} asin(\frac{1}{2})$, $S_{fi}(f_n) = \sigma_u^2$. In addition, when f_n

[4]In this paper δ is the differencing order. A more usual notation is d but this is already used for the degree of the polynomial used in the DFA, in this paper.

tends to $\pm\frac{1}{2}$, the PSD of the FI white noise tends to $\frac{\sigma_u^2}{2^{2\delta}}$. For a given normalized frequency $0 < f_n < \frac{1}{\pi}asin(\frac{1}{2})$ (resp. $\frac{1}{\pi}asin(\frac{1}{2}) < f_n < \frac{1}{2}$), the larger $\delta > 0$, the larger (resp. the smaller) $S_{fi}(f_n)$. In Figure 3, a time representation of a FI white noise, its power spectrum (in blue) as well as its PSD (in red) are given where $\delta = 0.2$.

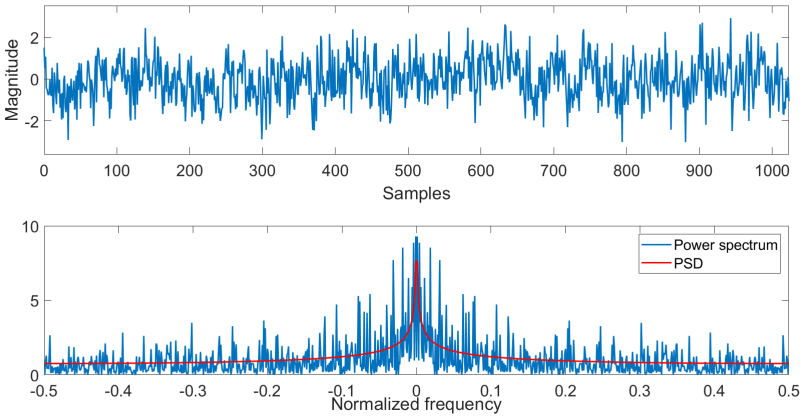


Figure 3: Time representation of a FI white noise, its power spectrum and the corresponding PSD

Using the inverse Fourier transform of (41), the correlation function of the FI white noise satisfies [42] when $-\frac{1}{2} < \delta < \frac{1}{2}$:

$$R_{fi,\tau} = \frac{\Gamma(\tau+\delta)\Gamma(1-2\delta)}{\Gamma(\tau+1-\delta)\Gamma(1-\delta)\Gamma(\delta)}\sigma_u^2 \tag{42}$$

where Γ is the Gamma function.

Consequently, $\frac{R_{fi,\tau}}{R_{fi,0}} = \frac{\Gamma(1-\delta)}{\Gamma(\delta)}\frac{\Gamma(\tau+\delta)}{\Gamma(\tau+1-\delta)} \approx \frac{\Gamma(1-\delta)}{\Gamma(\delta)}\tau^{2\delta-1}$ for large values of τ by taking advantage of Stirling formula. As it decays at a hyperbolic rate for $0 < \delta < \frac{1}{2}$, it is a long-memory process. When $-\frac{1}{2} < \delta < 0$, the process is anti–persistent with $R_{fi,\tau} < 0$ for $\tau > 0$.

For this type process, there is a theoretical relationship between the differencing order δ and the Hurst exponent H:

$$H = \delta + 0.5 \tag{43}$$

Taking into account the above properties and introducing the sets of poles $\{p_i\}_{i=1,\dots,p}$ whose modulus are smaller than 1 and the sets of zeros $\{z_i\}_{i=1,\dots,q}$ of the transfer function of the linear filter whose input is the FI white noise and the output is the ARFIMA process, the PSD of the ARFIMA process satisfies:

$$\begin{aligned} S_{xx}(f_n) &= S_{fi}(f_n)\Big|\frac{\prod_{i=1}^{q}(1-z_i z^{-1})}{\prod_{i=1}^{p}(1-p_i z^{-1})}\Big|^2_{z=exp(j2\pi f_n)} \\ &= \sigma_u^2|2sin(\pi f_n)|^{-2\delta}\Big|\frac{\prod_{i=1}^{q}(1-z_i z^{-1})}{\prod_{i=1}^{p}(1-p_i z^{-1})}\Big|^2_{z=exp(j2\pi f_n)} \end{aligned} \tag{44}$$

In Figure 4, a time representation of a realization of an ARFIMA process, its power spectrum (in blue) as well as its PSD (in red) are given. This 1024 sample-realization of a real ARFIMA$(2, 0.2, 2)$ process, whose corresponding transfer function is defined by the two complex conjugate non-unit poles equal to $0.9e^{\pm\frac{2\pi}{3}}$ and by the two complex-conjugate zeros equal to $0.7e^{\pm\frac{\pi}{3}}$, is deduced from the previous FI white noise.

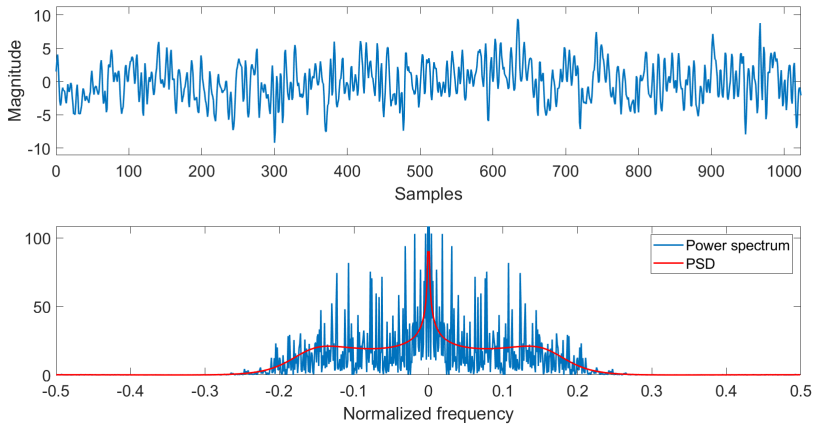


Figure 4: Time representation of a ARFI MA process, its power spectrum and the corresponding PSD

Finally, if the ARFIMA process evolves around a trend that is independent from the driving process u of the FI noise, the PSD of the resulting data is

the sum of $S_{xx}(f_n)$ and the spectral contribution of the trend. In Figure 5, a time representation of a realization of an ARFIMA process evolving around its exponential trend and its spectrogram are given.

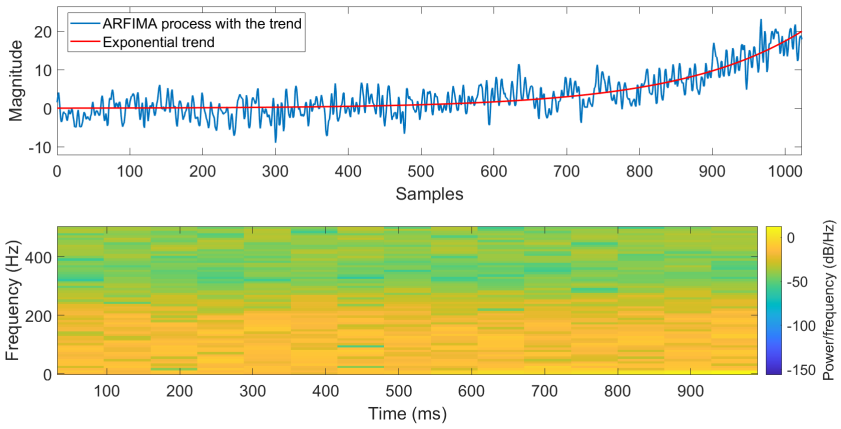


Figure 5: Time representation of an ARFIMA process evolving around its additive exponential trend and its spectrogram

According to (18), $E[F^2_{\mathbf{1},d}(N)]$ depends on the product between $S_{xx}(f_n)$ and $\Psi_{\mathbf{1},d}(f_n)$.

When $\delta > 0$ becomes larger and larger, $S_{fi}(f_n)$ exhibits a resonance around the null frequency that becomes spikier and spikier. In addition, from a filtering point of view, the zeros $\{z_i\}_{i=1,..,q}$ tend to create rejections. If one of the zeros has its modulus equal to 1 and its argument equal to $2\pi f_{n,z}$, the frequency component $f_{n,z}$ of the FI noise is totally rejected. Regarding the poles $\{p_i\}_{i=1,..,q}$, they tend to create "resonances". More particularly, if one of the poles has its modulus close to 1 and its argument equal to $2\pi f_{n,p}$, the frequency components of the FI white noise around $f_{n,p}$ are amplified by the filter whose transfer function is $\frac{\prod_{i=1}^{q}(1-z_i z^{-1})}{\prod_{i=1}^{p}(1-p_i z^{-1})}$, and so all the more as the modulus is close to 1. Consequently, the power of the ARFIMA process is rather located in low-frequencies when δ increases. When N is small, as shown in Figure 2, the frequency responses $\Psi_{\mathbf{1},1}(f_n)$ and $\Psi_{\mathbf{1},2}(f_n)$ are rather band-pass

in the medium and high-frequencies. Consequently, for these small values of N, there may be a high risk that the product $\Psi_{\mathbf{1},1}(f_n)S_{xx}(f_n)$ or $\Psi_{\mathbf{1},2}(f_n)S_{xx}(f_n)$ can take small values whatever the normalized frequency f_n. We have also to take into account the influence of the signal trend on the value taken by the fluctuation function. Even if the profile trend is removed in the DFA or its higher-order variant, $E[F^2_{\mathbf{1},d}(N)]$ also depends on the spectral properties of the trend around which the ARFIMA process evolves. If the trend was a constant or slowly changes overtime, its power would be mainly located in the very-low frequencies. Therefore, depending on the shape of the signal trend spectrum, $E[F^2_{\mathbf{1},d}(N)]$ can be very small for some small values of N. In practice, the fluctuation function $F_{\mathbf{1},d}(N)$ should be small and we can guess that the result cannot be reliable. This problem was also pointed out by [26] where the authors suggested introducing a correction function, but not justified by considering the frequency response of the equivalent filter. Given Figure 2, this unwanted phenomenon is probably more noticeable with the 2^{nd}-order DFA due to the shape of $\Psi_{\mathbf{1},2}(f_n)$. When N increases, due to the evolution of the shapes of $\Psi_{\mathbf{1},1}(f_n)$ and $\Psi_{\mathbf{1},2}(f_n)$ which are more and more low-frequency, the influence of $S_{xx}(f_n)$ becomes more and more important in the definition of the fluctuation function. Due to the difference between $\Psi_{\mathbf{1},1}(f_n)$ and $\Psi_{\mathbf{1},2}(f_n)$ shown in Figure 2, we may guess that larger values of N will be necessary.

The discussion will go further in Section 4. In the next section, let us address a variant where there is no discontinuity between local trends.

2. Higher-order CDFA: two types of derivations, links between two CDFAs of different orders and link with the DFA

In this section, after briefly recalling what the CDFA is all about, our purpose is to find connections between the CDFA and the DFA in its standard version, but also for a higher order. As the higher-order CDFA has not been presented yet, we propose to derive it by first following the reasoning we initially proposed in [16] and then by using Lagrange multipliers. These derivations will be useful

to deduce some links between two frequency responses of the equivalent filters related to CDFA of different orders and then between the CDFA and the DFA of different orders. Note that subscript $_{\mathbf{2},d}$ will be used to refer to the d-order CDFA, also called CDFAd.

2.1. Recalling what CDFA is all about

In [16], the CDFA was presented by first expressing the criterion to be minimized and the constraints to be followed. It was shown that the joint estimation of the $2L$ parameters defining the global trend, *i.e.* the slope and the vertical intercept of each local trend, reduced to the joint estimation of $L+1$ parameters due to the $L-1$ constraints of continuity between the local trends. Thus, one has either:

$$t_{l+1}(1) = a_{1,l+1}(lN+1) + a_{0,l+1} = t_l(N+1) = a_{1,l}(lN+1) + a_{0,l} \tag{45}$$

or

$$t_{l+1}(0) = a_{1,l+1}(lN) + a_{0,l+1} = t_l(N) = a_{1,l}(lN) + a_{0,l} \tag{46}$$

where $t_l(n)$ is the n^{th} sample of the l^{th} local trend.
A matrix form was then deduced. This led to the following criterion:

$$\begin{aligned} J(a_{1,1},..,a_{L,1},a_{1,0}) &= \left|\left|C_{LN,0}Y_{int} - A_{\mathbf{2},1}\Theta_{\mathbf{2},1}\right|\right|^2 \\ &= [C_{LN,0}Y_{int} - A_{\mathbf{2},1}\Theta_{\mathbf{2},1}]^T[C_{LN,0}Y_{int} - A_{\mathbf{2},1}\Theta_{\mathbf{2},1}] \end{aligned} \tag{47}$$

where the $(L+1)\times 1$ column parameter vector $\Theta_{\mathbf{2},1}$ is defined by $[a_{1,1},..,a_{L,1},a_{1,0}]^T$ and the $LN \times (L+1)$ matrix $A_{\mathbf{2},1}$ is given by:

$$A_{\mathbf{2},1} = \begin{bmatrix} B_{1,1} & \mathbb{1}_{N\times 1} \\ \vdots & \vdots \\ B_{L,1} & \mathbb{1}_{N\times 1} \end{bmatrix} \tag{48}$$

where:

$$B_{1,1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 2 & & \vdots & \\ \vdots & & \vdots & \\ N & \underbrace{0 \quad \cdots \quad 0}_{L-1} & & \end{bmatrix} \tag{49}$$

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

$$B_{l,1} = \begin{bmatrix} \beta(1) & N & \cdots & N & (l-1)N+1-\beta(l-1) & 0 & \cdots & 0 \\ \beta(1) & N & \cdots & N & (l-1)N+2-\beta(l-1) & & \vdots & \\ \vdots & & \vdots & & \vdots & & \vdots & \\ \beta(1) & \underbrace{N \quad \cdots \quad N}_{l-2} & & & lN-\beta(l-1) & \underbrace{0 \quad \cdots \quad 0}_{L-l} & & \end{bmatrix} \tag{50}$$

with $\beta(l) = lN + 1$ (resp. lN) if the first constraint (45) (resp. the second constraint (46)) is taken into account. Note that the first and the second value in the subscript $_{l,1}$ respectively refer to the the local trend (but also the position of the matrix in the block matrix $A_{\mathbf{2},1}$ along the rows) and the degree here equal to 1. This notation will be useful when extending the CDFA to an order larger than 1.

Consequently, the trend vector $T_{\mathbf{2},1} = A_{\mathbf{2},1}\hat{\Theta}_{\mathbf{2},1}$ was expressed this way:

$$\begin{aligned} T_{\mathbf{2},1} = A_{\mathbf{2},1}[A_{\mathbf{2},1}^T A_{\mathbf{2},1}]^{-1} A_{\mathbf{2},1}^T C_{LN,0} Y_{int} \quad &= A_{\mathbf{2},1}[A_{\mathbf{2},1}^T A_{\mathbf{2},1}]^{-1} A_{\mathbf{2},1}^T C_{LN,0} H_M J_M Y \\ = P_{\mathbf{2},1} C_{LN,0} H_M J_M Y & \end{aligned} \tag{51}$$

The trend vector was hence seen as the orthogonal projection of $C_{LN,0}H_M J_M Y$ onto the space spanned by the columns of $A_{\mathbf{2},1}$. It is given by $P_{\mathbf{2},1}$ in the equation above.

In the next subsection, the standard CDFA is extended to the higher-order and then we address the derivation of the trend differently using the Lagrange multipliers.

2.2. Derivations of the higher-order CDFA

2.2.1. Formulation of the higher-order CDFA following previous contributions

In this section, the formulation of the higher-order CDFA following the reasoning proposed in [16] is presented. Similarly to the standard CDFA, the higher-order CDFA consists in building a continuous global from local trends modeled by polynomials of degree d. The continuity between locals trend is ensured if one of the following equality holds $\forall l \in [\![0; L-1]\!]$:

$$t_{d,l+1}(1) = \sum_{k=0}^{d} a_{k,l+1}(lN+1)^k = t_{d,l}(N+1) = \sum_{k=0}^{d} a_{k,l}(lN+1)^k \tag{52}$$

or

$$t_{d,l+1}(0) = \sum_{k=0}^{d} a_{k,l+1}(lN)^k = t_{d,l}(N) = \sum_{k=0}^{d} a_{k,l}(lN)^k \tag{53}$$

where $t_{d,l}(n)$ is the n^{th} sample of the l^{th} local trend modeled by a d^{th}-degree polynomial.

One can rewrite the constraint as follows:

$$a_{0,l+1} = a_{0,l} + \sum_{k=1}^{d} (a_{k,l} - a_{k,l+1})\beta^k(l) \tag{54}$$

with $\beta(l) = lN + 1$ (resp. $\beta(l) = lN$) using the first equality (52) (resp. the second equality (53)). Then, by induction, one can show that:

$$a_{0,l+1} = a_{0,1} + \sum_{j=1}^{l} \sum_{k=1}^{d} (a_{k,j} - a_{k,j+1})\beta^k(j) \tag{55}$$

and by interchanging both sums, this leads to:

$$a_{0,l+1} = a_{0,1} + \sum_{k=1}^{d} \left[\beta^k(1)a_{k,1} + \sum_{j=2}^{l} \underbrace{(\beta^k(j) - \beta^k(j-1))}_{:=\Delta\beta^k(j)} a_{k,j} - \beta^k(l)a_{k,l+1} \right] \tag{56}$$

The joint estimation of the $(d+1)L$ parameters, $\{a_{k,l}\}_{l=1,\ldots,L \text{ and } k=0,\ldots,d}$ consists in minimizing a LS criterion under the $L-1$ constraints (55). It can hence be rewritten as follows:

$$\begin{aligned} J = & \sum_{n=1}^{N} (y_{int}(n) - \sum_{k=0}^{d} a_{k,1} n^k)^2 \\ & + \sum_{l=2}^{L} \sum_{n=1}^{N} \left(y_{int}((l-1)N + n) - \sum_{k=1}^{d} a_{k,l}((l-1)N + n)^k \right. \\ & \left. - a_{0,1} - \sum_{k=1}^{d} \left[\beta^k(1)a_{k,1} - \sum_{j=2}^{l-1} \Delta\beta^k(j)a_{k,j} + \beta^k(l-1)a_{k,l} \right] \right)^2 \end{aligned} \tag{57}$$

Let us express this criterion in a matrix form. To this end, let us introduce the set of matrices $\{B_{l,k}\}$ of size $N \times L$. For $l \geq 2$ and $k \geq 1$, one has:

$$B_{l,k} = \begin{bmatrix} \beta^k(1) & \Delta\beta^k(2) & \ldots & \Delta\beta^k(l-1) & ((l-1)N+1)^k - \beta^k(l-1) & 0 & \cdots & 0 \\ \beta^k(1) & & \vdots & & ((l-1)N+2)^k - \beta^k(l-1) & & \vdots & \\ \vdots & & \vdots & & \vdots & & \vdots & \\ \beta^k(1) & \underbrace{\Delta\beta^k(2) \quad \ldots \quad \Delta\beta^k(l-1)}_{l-2} & & & (lN)^k - \beta^k(l-1) & \underbrace{0 \quad \cdots \quad 0}_{L-l} & & \end{bmatrix} \tag{58}$$

Note that $\Delta\beta(j) = N\ \forall j \in [2; l-1]$. In addition, one special case is given by:

$$B_{1,k} = \begin{bmatrix} 1^k & 0 & \cdots & 0 \\ 2^k & & \vdots & \\ \vdots & & \vdots & \\ N^k & \underbrace{0 \quad \cdots \quad 0}_{L-1} & & \end{bmatrix} \tag{59}$$

Given $\Theta_{\mathbf{2},d}$ the parameter vector of size $dL+1 \times 1$ defined by:

$$\Theta_{\mathbf{2},d} = [a_{d,1}, .., a_{d,L}, a_{d-1,1}, .., a_{d-1,L}, ..., a_{1,1}, .., a_{1,L}, a_{0,1}]^T \tag{60}$$

$$T_{\mathbf{2},d,l} = \begin{bmatrix} B_{l,d} & B_{l,d-1} & \ldots & B_{l,1} & \mathbb{1}_{N\times 1} \end{bmatrix} \Theta_{\mathbf{2},d} \tag{61}$$

By defining the block matrix $\boldsymbol{A}_{2,d}$ of size $LN \times (dL+1)$ by:

$$A_{\mathbf{2},d} = \begin{bmatrix} B_{1,d} & B_{1,d-1} & \ldots & B_{1,1} & \mathbb{1}_{N\times 1} \\ \vdots & \vdots & & \vdots & \vdots \\ B_{L,d} & B_{L,d-1} & \ldots & B_{L,1} & \mathbb{1}_{N\times 1} \end{bmatrix} \tag{62}$$

the optimization problem in its matrix form can be expressed this way:

$$\arg\min_{\Theta} ||C_{LN,0}Y_{int} - A_{\mathbf{2},d}\Theta_{\mathbf{2},d}||^2 \tag{63}$$

This minimization gives the parameter vector $\hat{\Theta}_{\mathbf{2},d}$:

$$\hat{\Theta}_{\mathbf{2},d} = (A_{\mathbf{2},d}^T A_{\mathbf{2},d})^{-1} A_{\mathbf{2},d}^T C_{LN,0} Y_{int} \tag{64}$$

The global trend $T_{\mathbf{2},d} = A_{\mathbf{2},d}\hat{\Theta}_{\mathbf{2},d}$ can be deduced as follows:

$$T_{\mathbf{2},d} = A_{\mathbf{2},d}(A_{\mathbf{2},d}^T A_{\mathbf{2},d})^{-1} A_{\mathbf{2},d}^T C_{LN,0} Y_{int} = P_{\mathbf{2},d} C_{LN,0} Y_{int}. \tag{65}$$

where $P_{\mathbf{2},d}$ is the orthogonal projector on the space generated by the column of $A_{\mathbf{2},d}$. This is hence the generalization of (51) to the degree d.

In the Figure 6, examples are given of trends and residues obtained using the 2^{nd}-order DFA and the 2^{nd}-order CDFA.

Let us give some comments on Figure 7: once again, the frequency response of the equivalent filter is always null for the null frequency whatever the order of the CDFA. In addition, the shape of the frequency response of the equivalent filter associated with the CDFA always exhibits a resonance. When N

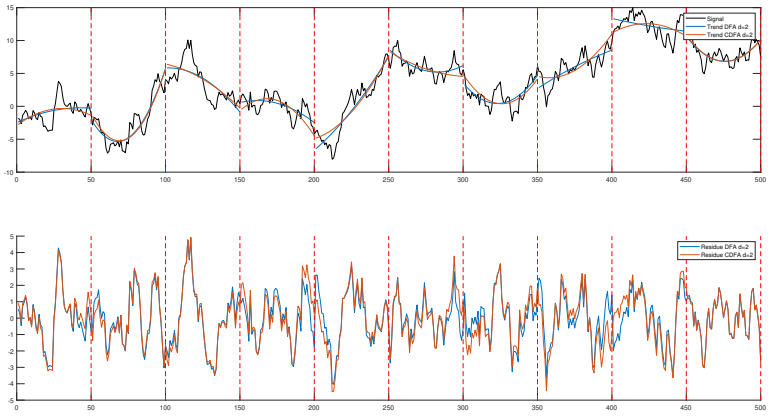


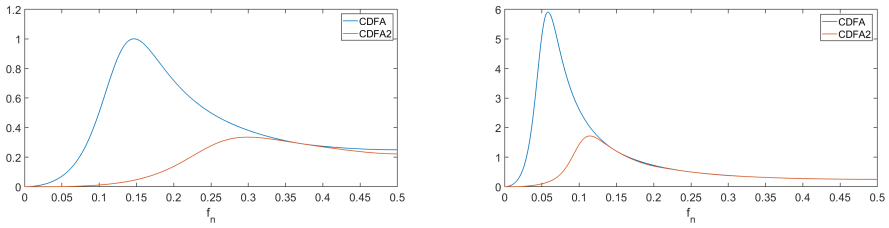
Figure 6: Estimated trends using the 2^{nd}-order DFA and 2^{nd}-order CDFA (CDFA2) and the corresponding residues

increases, the resonance frequency becomes smaller and smaller and spikier and spikier. Finally, when comparing Figure 2 and Figure 7, one can notice that the frequency response of the 2^{nd}-order CDFA is more "low-frequency" than the frequency response of the 2^{nd}-order DFA.
In the next subsection, let us present an alternative way to obtain this trend using Lagrange multipliers.

2.2.2. Formulation of the higher-order CDFA with Lagrange multipliers

Let us now introduce the formulation of the higher-order CDFA by using the Lagrange multipliers. This will enable us to find a link between the trend vector obtained with the higher-order DFA and the one deduced from the higher-order CDFA.

Theorem 3. *The trend vector $T_{\mathbf{2},d}$ deduced from the d^{th}-order CDFA can be expressed as the difference between two terms: The trend vector $T_{\mathbf{1},d}$ deduced from the d^{th}-order DFA and a corrective term denoted as $T_{\mathbf{1}\Rightarrow\mathbf{2},d}$, which can be interpreted as the result of an orthogonal projection.*



(a) $N = 4$ (b) $N = 10$

(c) $N = 30$ (d) $N = 100$

Figure 7: Frequency responses of the CDFA and the 2^{nd}-order CDFA for $N = 4$, $N = 10$, $N = 30$ and $N = 100$

Proof. Given $\underline{H}_{d,l} = \begin{bmatrix} 1 & \beta(l) & \dots & \beta^d(l) \end{bmatrix}$, the constraints (52) and (53) can be written in a matrix form as follows:

$$\begin{bmatrix} 1 & \beta(l) & \dots & \beta^d(l) \end{bmatrix} \begin{bmatrix} a_{0,l+1} \\ a_{1,l+1} \\ \vdots \\ a_{d,l+1} \end{bmatrix} = \underline{H}_{d,l}\theta_{d,l+1} \tag{66}$$

$$= \underline{H}_{d,l}\theta_{d,l} = \begin{bmatrix} 1 & \beta(l) & \dots & \beta^d(l) \end{bmatrix} \begin{bmatrix} a_{0,l} \\ a_{1,l} \\ \vdots \\ a_{d,l+1} \end{bmatrix}$$

The above equality (66) can be rewritten $\forall l \in [\![1; L-1]\!]$ this way:

$$\underline{H}_{d,l}(\theta_{d,l+1} - \theta_{d,l}) = 0 \tag{67}$$

Therefore, the LS criterion to be considered is the following:

$$||C_{LN,0}Y_{int} - A_{\mathbf{1},d}\Theta||^2 \tag{68}$$

in which $\Theta = \begin{bmatrix} \theta_{d,1}^T & \dots & \theta_{d,L}^T \end{bmatrix}^T$ and the $L-1$ constraints (67) are going to be included by considering $L-1$ Lagrange multipliers $\{\lambda_i\}_{i=1,\dots,L-1}$ stored in one vector $\underline{\lambda}$. This can be rewritten as follows:

$$||C_{LN,0}Y_{int} - A_{\mathbf{1},d}\Theta||^2 + \underline{\lambda}K_d^T\Theta \tag{69}$$

where K_d^T is a block matrix defined by:

$$K_d^T = \begin{bmatrix} -\underline{H}_{d,1} & \underline{H}_{d,1} & \mathbf{0}_{1\times(d+1)} & \dots & \dots & \mathbf{0}_{1\times(d+1)} \\ \mathbf{0}_{1\times(d+1)} & -\underline{H}_{d,2} & \underline{H}_{d,2} & \mathbf{0}_{1\times(d+1)} & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \mathbf{0}_{1\times(d+1)} & -\underline{H}_{d,L-1} & \underline{H}_{d,L-1} & \mathbf{0}_{1\times(d+1)} \\ \mathbf{0}_{1\times(d+1)} & \dots & \dots & \mathbf{0}_{1\times(d+1)} & -\underline{H}_{d,L} & \underline{H}_{d,L} \end{bmatrix} \tag{70}$$

After expressing the derivative of (69) with respect to Θ, one obtains:

$$-2A_{\mathbf{1},d}^T C_{LN,0}Y_{int} + 2A_{\mathbf{1},d}^T A_{\mathbf{1},d}\Theta + K_d\underline{\lambda}^T = 0 \tag{71}$$

which can be rewritten as:

$$\Theta = (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}A_{\mathbf{1},d}^T C_{LN,0}Y_{int} - \frac{1}{2}(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}K_d\underline{\lambda}^T \tag{72}$$

Let us now take into account the constraint (67) $\forall l \in [\![0; L-1]\!]$ which, by construction of the matrix K_d^T, is equivalent to:

$$K_d^T\Theta = \mathbf{0}_{(L-1)\times 1} \tag{73}$$

Then using (72), this amounts to writing:

$$\begin{aligned} K_d^T\Theta &= K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}A_{\mathbf{1},d}^T C_{LN,0}Y_{int} - \frac{1}{2}K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}K_d\underline{\lambda}^T \\ &= \mathbf{0}_{(L-1)\times 1} \end{aligned} \tag{74}$$

or equivalently:

$$K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}K_d\underline{\lambda}^T = 2K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}A_{\mathbf{1},d}^T C_{LN,0}Y_{int} \tag{75}$$

As $K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}K_d$ is invertible, one has:

$$\underline{\lambda}^T = 2\Big(K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}K_d\Big)^{-1}K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}A_{\mathbf{1},d}^T C_{LN,0}Y_{int} \tag{76}$$

Combining (72) and (76) leads to:

$$\begin{aligned}\Theta &= (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} A_{\mathbf{1},d}^T C_{LN,0} Y_{int} \\ &- (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} K_d \Big(K_d^T (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} K_d \Big)^{-1} K_d^T (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} A_{\mathbf{1},d}^T C_{LN,0} Y_{int}\end{aligned} \tag{77}$$

Consequently, the trend satisfies:

$$\begin{aligned}T_{\mathbf{2},d} &= A_{\mathbf{1},d} (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} A_{\mathbf{1},d}^T C_{LN,0} Y_{int} \\ &- A_{\mathbf{1},d} (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} K_d \Big(K_d^T (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} K_d \Big)^{-1} K_d^T (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} A_{\mathbf{1},d}^T C_{LN,0} Y_{int}\end{aligned} \tag{78}$$

By introducing the following matrix:

$$W_d^T = K_d^T (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} A_{\mathbf{1},d}^T \tag{79}$$

and as one can notice that:

$$\begin{aligned}W_d^T W_d &= K_d^T (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} A_{\mathbf{1},d}^T A_{\mathbf{1},d} (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} K_d \\ &= K_d^T (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} K_d\end{aligned} \tag{80}$$

one can rewrite the trend given in (78) as follows:

$$\begin{aligned}T_{\mathbf{2},d} &= A_{\mathbf{1},d} (A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} A_{\mathbf{1},d}^T C_{LN,0} Y_{int} - W_d (W_d^T W_d)^{-1} W_d^T C_{LN,0} Y_{int} \\ &= P_{\mathbf{1},d} C_{LN,0} Y_{int} - P_{\mathbf{1}\Rightarrow\mathbf{2},d} C_{LN,0} Y_{int} \\ &= T_{\mathbf{1},d} - P_{\mathbf{1}\Rightarrow\mathbf{2},d} C_{LN,0} Y_{int} = T_{\mathbf{1},d} - T_{\mathbf{1}\Rightarrow\mathbf{2},d}\end{aligned} \tag{81}$$

where $T_{\mathbf{1}\Rightarrow\mathbf{2},d}$ is a correction term.

Therefore, the global trend obtained with the d^{th}-order DFA is the result of two orthogonal projections of $C_{LN,0}Y_{int}$ on two different subspaces: one spanned by the columns of $A_{\mathbf{2},d}$ and the other spanned by the columns of $W_d = A_{\mathbf{1},d}(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1} K_d$. It is the sum of the global trend obtained with the d^{th}-order CDFA and a corrective term. □

2.3. Links between two approaches

2.3.1. Link between the frequency responses of two equivalent filters: the first one related to a higher-order CDFA and the second related to the higher-order DFA

Theorem 4. *$\Psi_{\mathbf{2},d}(f_n)$ which is the frequency response of the equivalent filter related to the d^{th}-order CDFA can be expressed as the sum of two terms: $\Psi_{\mathbf{1},d}(f_n)$ which is the frequency response of the equivalent filter related to the d^{th}-order DFA and a second term denoted as $\Psi_{\mathbf{1}\Rightarrow\mathbf{2},d}(f_n)$ which depends on the projector $P_{\mathbf{1}\Rightarrow\mathbf{2},d}$.*

Proof. One can express the matrix $\Gamma_{\mathbf{2},d}$ from the matrix $\Gamma_{\mathbf{1},d}$. To this end, one just has to define the residue and the matrix $B_{\mathbf{2},d}$ from $B_{\mathbf{1},d}$.

$$\begin{aligned} B_{\mathbf{2},d} &= C_{LN,0}H_M J_M - T_{\mathbf{2},d} \\ &= (I - (P_{\mathbf{1},d} - P_{\mathbf{1}\Rightarrow\mathbf{2},d}))C_{LN,0}H_M J_M \\ &= B_{\mathbf{1},d} + \underbrace{P_{\mathbf{1}\Rightarrow\mathbf{2},d}C_{LN,0}H_M J_M}_{=B_{\mathbf{1}\Rightarrow\mathbf{2},d}} \end{aligned} \tag{82}$$

Then one has:

$$\begin{aligned} \Gamma_{\mathbf{2},d} &= \frac{1}{LN}B_{\mathbf{2},d}^T B_{\mathbf{2},d} = \frac{1}{LN}(B_{\mathbf{1},d} + B_{\mathbf{1}\Rightarrow\mathbf{2},d})^T(B_{\mathbf{1},d} + B_{\mathbf{1}\Rightarrow\mathbf{2},d}) \\ &= \Gamma_{\mathbf{1},d} + \frac{1}{LN}B_{\mathbf{1},d}^T B_{\mathbf{1}\Rightarrow\mathbf{2},d} + \frac{1}{LN}B_{\mathbf{1}\Rightarrow\mathbf{2},d}^T B_{\mathbf{1},d} + \frac{1}{LN}B_{\mathbf{1}\Rightarrow\mathbf{2},d}^T B_{\mathbf{1}\Rightarrow\mathbf{2},d} \end{aligned} \tag{83}$$

By developing the third term $B_{\mathbf{1}\Rightarrow\mathbf{2},d}^T B_{\mathbf{1},d}$ in (83) and as $P_{\mathbf{1}\Rightarrow\mathbf{2},d}^T = P_{\mathbf{1}\Rightarrow\mathbf{2},d}$, one gets:

$$\begin{aligned} B_{\mathbf{1}\Rightarrow\mathbf{2},d}^T B_{\mathbf{1},d} &= (P_{\mathbf{1}\Rightarrow\mathbf{2},d}C_{LN,0}H_M J_M)^T(I - P_{\mathbf{1},d})C_{LN,0}H_M J_M \\ &= J_M^T H_M^T C_{LN,0}^T P_{\mathbf{1}\Rightarrow\mathbf{2},d}^T(I - P_{\mathbf{1},d})C_{LN,0}H_M J_M \\ &= J_M^T H_M^T C_{LN,0}^T(P_{\mathbf{1}\Rightarrow\mathbf{2},d} - P_{\mathbf{1}\Rightarrow\mathbf{2},d}P_{\mathbf{1},d})C_{LN,0}H_M J_M \end{aligned} \tag{84}$$

However, given the expressions of the two projectors, one can check that:

$$\begin{aligned} &P_{\mathbf{1}\Rightarrow\mathbf{2},d}P_{\mathbf{1},d} = A_{\mathbf{1},d}(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}K_d \\ &\times \Big(K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}K_d\Big)^{-1}K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}A_{\mathbf{1},d}^T A_{\mathbf{1},d}(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}A_{\mathbf{1},d}^T \\ &= A_{\mathbf{1},d}(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}K_d\Big(K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}K_d\Big)^{-1}K_d^T(A_{\mathbf{1},d}^T A_{\mathbf{1},d})^{-1}A_{\mathbf{1},d}^T \\ &= P_{\mathbf{1}\Rightarrow\mathbf{2},d} \end{aligned} \tag{85}$$

Finally, (84) becomes:

$$B_{\mathbf{1}\Rightarrow\mathbf{2},d}^T B_{\mathbf{1},d} = B_{\mathbf{1},d}^T B_{\mathbf{1}\Rightarrow\mathbf{2},d} = \mathbf{0}_{M\times M} \tag{86}$$

Consequently, (83) can be rewritten:

$$\begin{aligned}\Gamma_{\mathbf{2},d} &= \Gamma_{\mathbf{1},d} + \frac{1}{LN} B^T_{\mathbf{1}\Rightarrow\mathbf{2},d} B_{\mathbf{1}\Rightarrow\mathbf{2},d} \\ &= \Gamma_{\mathbf{1},d} + \Gamma_{\mathbf{1}\Rightarrow\mathbf{2},d}\end{aligned} \tag{87}$$

By linearity of the Fourier transform, one can deduce the link between the frequency responses of the equivalent filters related to the d^{th}-order CDFA of order and the d^{th}-order DFA:

$$\Psi_{\mathbf{2},d}(f_n) = \Psi_{\mathbf{1},d}(f_n) + \Psi_{\mathbf{1}\Rightarrow\mathbf{2},d}(f_n) \tag{88}$$

The CDFA was introduced to correct the discontinuity between each local trend by introducing a constraint in the minimization of the first-order DFA. We have seen that this method can be extended to a higher order and that the orthogonal projector related to the CDFA can be expressed from the orthogonal projector related to the DFA. □

2.3.2. Link between the frequency responses of two equivalent filters related to CDFA of different orders

Theorem 5. *$\Psi_{\mathbf{2},d_2}(f_n)$ which is the frequency response of the equivalent filter related to the CDFA of order d_2 can be expressed as the difference between two terms: $\Psi_{\mathbf{2},d_1}(f_n)$ which is the frequency response of the equivalent filter related to the CDFA of order d_1 and a second term $\Psi_{\mathbf{2},d_1\Rightarrow d_2}(f_n)$ which is defined from $P_{\mathbf{2},d_1\Rightarrow d_2}$.*

First of all, given the definition (62), one can rewrite $A_{\mathbf{2},d_2}$ as a function of $A_{\mathbf{2},d_1}$ as follows:

$$A_{\mathbf{2},d_2} = \begin{bmatrix} D_{\mathbf{2},d_1\Rightarrow d_2} & A_{\mathbf{2},d_1} \end{bmatrix} \tag{89}$$

As we did in a previous section, let us express the orthogonal projector related

to the CDFA of order d_2. To this end, let us start by the following product:

$$A_{2,d_2}^T A_{2,d_2} = \begin{bmatrix} D_{\mathbf{2},d_1\Rightarrow d_2}^T \\ A_{2,d_1}^T \end{bmatrix} \begin{bmatrix} D_{\mathbf{2},d_1\Rightarrow d_2} & A_{2,d_1} \end{bmatrix} \tag{90}$$
$$= \begin{bmatrix} D_{\mathbf{2},d_1\Rightarrow d_2}^T D_{\mathbf{2},d_1\Rightarrow d_2} & D_{\mathbf{2},d_1\Rightarrow d_2}^T A_{2,d_1} \\ A_{2,d_1}^T D_{\mathbf{2},d_1\Rightarrow d_2} & A_{2,d_1}^T A_{2,d_1} \end{bmatrix}$$

Using the same type of inverse of block matrix based on the Schur complement[5] of the block located located at the bottom right of the matrix, one gets:

$$(A_{\mathbf{2},d_2}^T A_{\mathbf{2},d_2})^{-1} = \begin{bmatrix} \mathbf{0}_{L(d_2-d_1)\times L(d_2-d_1)} & \mathbf{0}_{L(d_2-d_1)\times(Ld_1+1)} \\ \mathbf{0}_{(Ld_1+1)\times L(d_2-d_1)} & (A_{\mathbf{2},d_1}^T A_{\mathbf{2},d_1})^{-1} \end{bmatrix} + \begin{bmatrix} I_{L(d_2-d_1)} \\ -(A_{\mathbf{2},d_1}^T A_{\mathbf{2},i})^{-1} A_{\mathbf{2},d_1}^T D_{\mathbf{2},d_1\Rightarrow d_2} \end{bmatrix} \tag{92}$$

$$\times [D_{\mathbf{2},d_1\Rightarrow d_2}^T D_{\mathbf{2},d_1\Rightarrow d_2} - D_{\mathbf{2},d_1\Rightarrow d_2}^T P_{2,d_1} D_{\mathbf{2},d_1\Rightarrow d_2}]^{-1} \begin{bmatrix} I_{L(j-i)} & -D_{\mathbf{2},d_1\Rightarrow d_2} A_{\mathbf{2},d_1} (A_{2,d_1}^T A_{\mathbf{2},d_1})^{-1} \end{bmatrix}$$

After pre and post-multiplying by $A_{\mathbf{2},d_2}$and $A_{\mathbf{2},d_2}^T$respectively and some simplification, one can deduce that:

$$\begin{aligned} P_{\mathbf{2},d_2} &= P_{\mathbf{2},d_1} + (I_{LN} - P_{\mathbf{2},d_1}) D_{\mathbf{2},d_1\Rightarrow d_2} \\ &\times (D_{\mathbf{2},d_1\Rightarrow d_2}^T (I_{LN} - P_{\mathbf{2},d_1})^T (I_{LN} - P_{\mathbf{2},d_1}) D_{\mathbf{2},d_1\Rightarrow d_2})^{-1} \\ &\times D_{\mathbf{2},d_1\Rightarrow d_2}^T (I_{LN} - P_{\mathbf{2},d_1})^T \end{aligned} \tag{93}$$

or equivalently:

$$P_{\mathbf{2},d_2} = P_{\mathbf{2},d_1} + P_{\mathbf{2},d_1\Rightarrow d_2} \tag{94}$$

where $P_{\mathbf{2},d_1\Rightarrow d_2}$ is the orthogonal projection onto the space spanned by the columns of $(I_{LN} - P_{\mathbf{2},d_1}) D_{\mathbf{2},d_1\Rightarrow d_2}$. Then, following the same reasoning as the one presented for the DFA for different orders, one obtains:

$$\Psi_{\mathbf{2},d_2}(f_n) = \Psi_{\mathbf{2},d_1}(f_n) - \Psi_{\mathbf{2},d_1\Rightarrow d_2}(f_n) \tag{95}$$

where $\Psi_{\mathbf{2},d_1\Rightarrow d_2}$ is the Fourier transform of the sequence defined by the traces of the sub-diagonal of the matrix $\Gamma_{\mathbf{2},d_1\Rightarrow d_2}$ that is deduced from $P_{\mathbf{2},d_1\Rightarrow d_2}$.

---

[5] One has now:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & D^{-1} \end{bmatrix} + \begin{bmatrix} I \\ -D^{-1}C \end{bmatrix} (A - BD^{-1}C)^{-1} \begin{bmatrix} I & -BD^{-1} \end{bmatrix} \tag{91}$$

2.4. Summary

The higher-order CDFA has been derived using two different methods. Both methods make it possible to define the link between the frequency responses of the equivalent filters related to the CDFA of different orders but also those related to the CDFA and the DFA of different orders.
By now, when dealing with the DFA or its variants, the global trend consisted in assuming that each sub-part of length N was modeled by a polynomial of a predefined degree. Then, depending on the variant -DFA or CDFA-, there was a continuity between the local trends or not. In the next section, we propose to *a posteriori* combine two global trends obtained with different orders and to analyze the equivalent filtering by looking at its frequency response. The results derived in Sections 1 and 2 will be useful.

3. *A posteriori* combining the global trends of different variants

In many fields where the development of statistical signal processing based on *a priori* models is required, using a single model is not necessarily relevant Despite the efforts made to select an appropriate model based on different criteria, the most popular of which are the Akaike information criterion (AIC) or the Bayesian information criterion (BIC), this type of approach is most of the time reliable when the number of samples is large enough. Moreover, as mentioned by Yang in [43], the model selection is not necessarily robust to a small perturbation in the data. When the regression function estimation is one of the main goal, *a priori* selecting a single model is one solution, but taking advantage of different models can be an alternative. Thus, when dealing with time-series forecasting, both strategies are used. We speak of forecasts model selection and forecast model averaging [44]. In the field of target tracking, when dealing with the estimation of position from radar measurements, Kalman filtering based on a *a priori* assumption of a motion model can be considered. However, as there can be maneuvering targets and as some model parameters can be unknown, interactive multiple model (IMM) approaches based on different models [45] can

be considered. In what follows, we propose to take into account this idea of combining different models in the framework of the DFA. More particularly, we suggest *a posteriori* combining two trends obtained with two DFA of different orders.

3.1. Linear combination of DFA of different orders (LC-DFA)

In this section, we suggest *a posteriori* combining the global trends of the DFA of orders d_1 et $d_2 > d_1$. Thus, by introducing the weight $\mu \in [0, 1]$, we propose to analyze the following projection matrix which is a linear combination of the orthogonal projector related to the DFA of order d_1 and d_2:

$$P_{\mathbf{1},equ}(\mu) = (1-\mu)P_{\mathbf{1},d_1} + \mu P_{\mathbf{1},d_2} \tag{96}$$

Thus, the resulting trend satisfies:

$$\begin{aligned} T_{\mathbf{1},equ}(\mu) &= P_{\mathbf{1},equ}(\mu) C_{LN,0} H_M J_M Y \\ &= \left((1-\mu)P_{\mathbf{1},d_1} + \mu P_{\mathbf{1},d_2}\right) C_{LN,0} H_M J_M Y \\ &= (1-\mu)T_{\mathbf{1},d_1} + \mu T_{\mathbf{1},d_2} \end{aligned} \tag{97}$$

Our purpose is to characterize the equivalent filter by means of its frequency response denoted as $\Psi_{\mathbf{1},equ,\mu}(f_n)$. To this end, one must first find the expression of $B_{\mathbf{1},equ}(\mu)$ similarly defined as $B_{\mathbf{1},d_1}$ and $B_{\mathbf{1},d_2}$:

$$\begin{aligned} B_{\mathbf{1},equ}(\mu) &= (I_{LN} - (1-\mu)P_{\mathbf{1},d_1} - \mu P_{\mathbf{1},d_2}) C_{LN,0} H_M J_M \\ &= (1-\mu)B_{\mathbf{1},d_1} + \mu B_{\mathbf{1},d_2} \end{aligned} \tag{98}$$

Then, given (98), the matrix $\Gamma_{\mathbf{1},equ}(\mu)$ can be deduced this way:

$$\begin{aligned} &\Gamma_{\mathbf{1},equ}(\mu) = \frac{1}{LN} B^T_{\mathbf{1},equ}(\mu) B_{\mathbf{1},equ}(\mu) \\ &= \frac{1}{LN}((1-\mu)B_{\mathbf{1},d_1} + \mu B_{\mathbf{1},d_2})^T((1-\mu)B_{\mathbf{1},d_1} + \mu B_{\mathbf{1},d_2}) \\ &= \frac{1}{LN}\Big((1-\mu)^2 B^T_{\mathbf{1},d_1} B_{\mathbf{1},d_1} + (1-\mu)\mu(B^T_{\mathbf{1},d_1} B_{\mathbf{1},d_2} + B^T_{\mathbf{1},d_2} B_{\mathbf{1},d_1}) + \mu^2 B^T_{\mathbf{1},d_2} B_{\mathbf{1},d_2}\Big) \end{aligned} \tag{99}$$

As $P_{\mathbf{1},d_1}P_{\mathbf{1},d_2} = P_{\mathbf{1},d_2}P_{\mathbf{1},d_1} = P_{\mathbf{1},d_1}$ for $d_1 \leqslant d_2$, the matrix product $B^T_{\mathbf{1},d_1}B_{\mathbf{1},d_2}$ can be rewritten as follows:

$$\begin{aligned} B^T_{\mathbf{1},d_1}B_{\mathbf{1},d_2} &= J_M^T H_M^T C^T_{LN,0}(I_{LN} - P_{\mathbf{1},d_1})(I_{LN} - P_{\mathbf{1},d_2})C_{LN,0}H_M J_M \\ &= J_M^T H_M^T C^T_{LN,0}(I_{LN} - P_{\mathbf{1},d_2} - P_{\mathbf{1},d_1} + P_{\mathbf{1},d_1}P_{\mathbf{1},d_2})C_{LN,0}H_M J_M \\ &= J_M^T H_M^T C^T_{LN,0}(I_{LN} - P_{\mathbf{1},d_2})C_{LN,0}H_M J_M \\ &= B^T_{\mathbf{1},d_2}B_{\mathbf{1},d_2} \end{aligned} \tag{100}$$

Thus, combining (99) and (100), $\Gamma_{\mathbf{1},equ}(\mu)$ can be written as a linear combination of $\Gamma_{\mathbf{1},d_1}$ and $\Gamma_{\mathbf{1},d_2}$:

$$\begin{aligned} \Gamma_{\mathbf{1},equ}(\mu) &= (1-\mu)^2\Gamma_{\mathbf{1},d_1} + (2\mu - \mu^2)\Gamma_{\mathbf{1},d_2} \\ &= (1-\mu)^2\Gamma_{\mathbf{1},d_1} + (1 - (1-\mu)^2)\Gamma_{\mathbf{1},d_2} \end{aligned} \tag{101}$$

Consequently, the frequency responses satisfy:

$$\begin{aligned} \Psi_{\mathbf{1},equ,\mu}(f_n) &= (1-\mu)^2\Psi_{\mathbf{1},d_1}(f_n) + (2\mu - \mu^2)\Psi_{\mathbf{1},d_2}(f_n) \\ &= (1-\mu)^2\Psi_{\mathbf{1},d_1}(f_n) + (1 - (1-\mu)^2)\Psi_{\mathbf{1},d_2}(f_n) \end{aligned} \tag{102}$$

However, in Section 2, we found links between the frequency responses of DFA of different orders. Therefore, using (38), one obtains:

$$\begin{aligned} \Psi_{\mathbf{1},equ,\mu}(f_n) &= (1-\mu)^2\Psi_{\mathbf{1},d_1}(f_n) + (2\mu - \mu^2)(\Psi_{\mathbf{1},d_1}(f_n) - \Psi_{\mathbf{1},d_1\Rightarrow d_2})(f_n) \\ &= \Psi_{\mathbf{1},d_1}(f_n) - (1 - (1-\mu)^2)\Psi_{\mathbf{1},d_1\Rightarrow d_2}(f_n) \end{aligned} \tag{103}$$

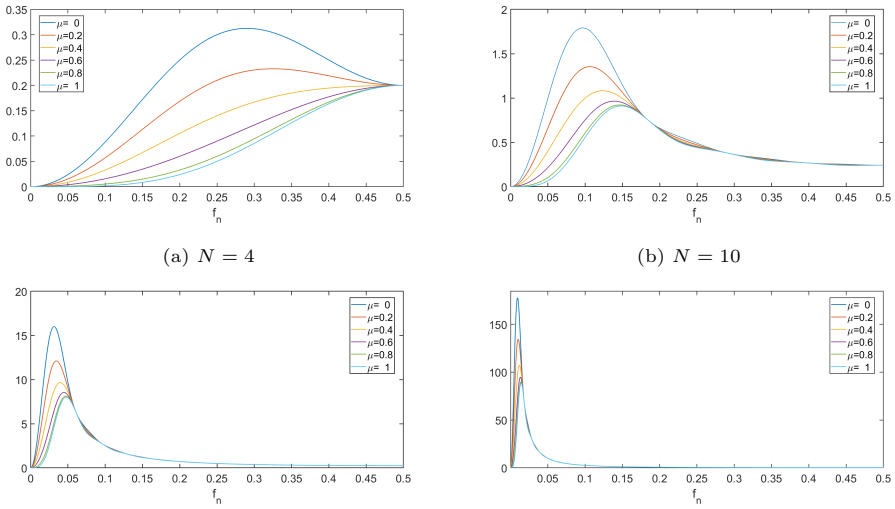
It should be noted that when $\mu = 0$, $\Psi_{\mathbf{1},equ,\mu}(f_n) = \Psi_{\mathbf{1},d_1}(f_n)$ whereas when $\mu = 1$, $\Psi_{\mathbf{1},equ,\mu}(f_n) = \Psi_{\mathbf{1},d_1}(f_n) - \Psi_{\mathbf{1},d_1\Rightarrow d_2}(f_n) = \Psi_{\mathbf{1},d_2}(f_n)$.

In addition, taken into account the fact that the function $x \mapsto 2x - x^2$ increases when $x \in [0,1]$, and that $\forall f_n, \Psi_{\mathbf{1},d_1\Rightarrow d_2}(f_n) \geqslant 0$, one has if $\mu_1 > \mu_2$:

$$\Psi_{\mathbf{1},equ,\mu_1}(f_n) \geqslant \Psi_{\mathbf{1},equ,\mu_2}(f_n) \tag{104}$$

In Figure 8, six frequency responses are provided to show how the frequency responses evolve when the weight μ varies, for $N = 4$, $N = 10$, $N = 30$ and $N = 100$.

The frequency response of the equivalent filter of the LC-DFA with orders equal to 1 and 2 is upper-bounded by the frequency response of the DFA and lower-bounded by the frequency response of the 2^{nd}-order DFA.



(a) $N = 4$ (b) $N = 10$

(c) $N = 30$ (d) $N = 100$

Figure 8: Frequency responses of the LC-DFA with orders 1 and 2 for different values of the weight, for $N = 4$, $N = 10$, $N = 30$ and $N = 100$

In the remainder, let us address the linear combination of trends obtained with CDFA of different orders.

3.2. Linear combination of CDFA of different order (LC-CDFA)

We propose to follow a similar reasoning for the combination of CDFA of different orders as the one presented in the previous section for the LC-DFA. Thus, using the same type of notation, one has:

$$P_{\mathbf{2},equ} = (1-\mu)P_{\mathbf{2},d_1} + \mu P_{\mathbf{2},d_2} \tag{105}$$

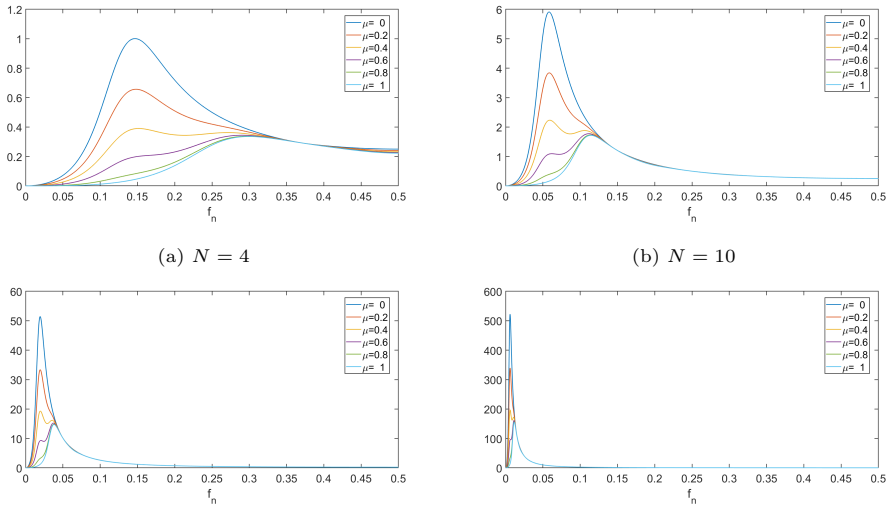
The trend hence satisfies:

$$T_{\mathbf{2},equ} = (1-\mu)T_{\mathbf{2},d_1} + \mu T_{\mathbf{2},d_2} \tag{106}$$

and following the same reasoning and mathematical development as the one presented for the DFA, the frequency response is given by:

$$\Psi_{\mathbf{2},equ,\mu}(f_n) = \Psi_{\mathbf{2},d_1}(f_n) - (1-(1-\mu)^2)\Psi_{\mathbf{2},d_1 \Rightarrow d_2}(f_n) \tag{107}$$

with $\Psi_{\mathbf{2},d_1\Rightarrow d_2}(f_n)$ similarly defined $\Psi_{\mathbf{1},d_1\Rightarrow d_2}(f_n)$. In addition, the same type of comments as the ones for the DFA made can be made. In Figure 9, six frequency responses are provided to show how the frequency responses evolve when the weight μ varies, for $N = 4$, $N = 10$, $N = 30$ and $N = 100$.



(a) $N = 4$ (b) $N = 10$

(c) $N = 30$ (d) $N = 100$

Figure 9: Frequency responses of the LC-CDFA with orders 1 and 2 for different values of the weight, for $N = 4$, $N = 10$, $N = 30$ and $N = 100$

4. Illustrations

In this section, we propose to first analyze the relevance of the linear combination on the ARFIMA process we introduced above. The purpose of this part will be to better understand the type of results one can obtain especially in the Step 5 of the DFA and its variants, which is dedicated to the estimation of the Hurst exponent. Then, we propose to study other signals, namely Weierstrass signals.

4.1. Comparison on an ARFIMA process

As described in Section 1.2, the last step of the DFA or one of its variants is to represent the logarithm of the fluctuation function as a function of logarithm

of N and to estimate the slope of the regression line. As our analysis is based on the expectation of the square of the fluctuation function, we propose to operate as follows: we suggest using a sliding rectangular window of length 800 on the realization of the ARFIMA process varying along an exponential trend in order to create a set of different data. Each resulting "frame" is processed by using the DFA, the 2^{nd}-order DFA, the LC-DFA, the CDFA, the 2^{nd}-order DFA, the LC-CDFA. Then, the fluctuation functions are averaged. See Figure 10. We follow the same type of approaches as the Welch periodogram that is used to estimate the power spectral density of a random process. Note that some authors operate differently by first analyzing the first LN samples and then the last LN samples.

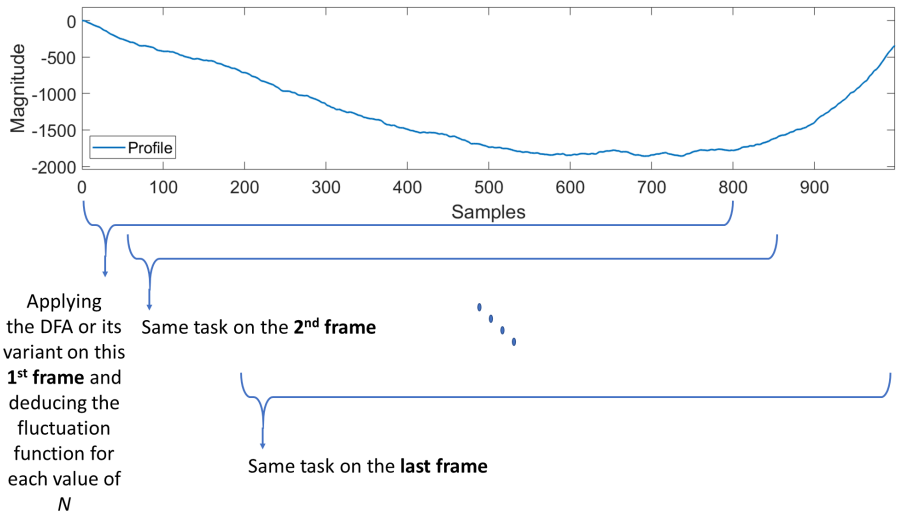
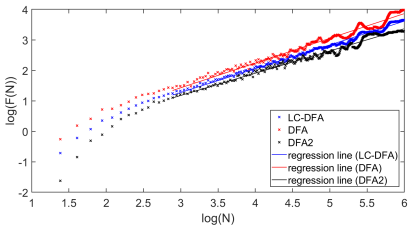
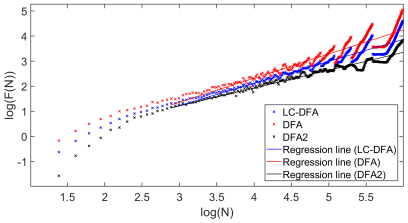


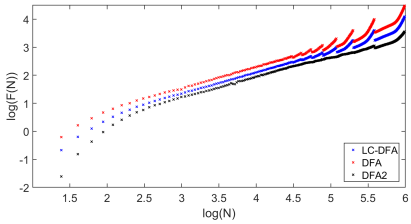
Figure 10: Proposed process based on a windowing of the profile to get an estimation of the expectation of the square of the fluctuation function



(a) When dealing with the first frame



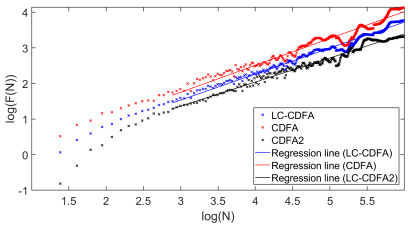
(b) When dealing with the last frame



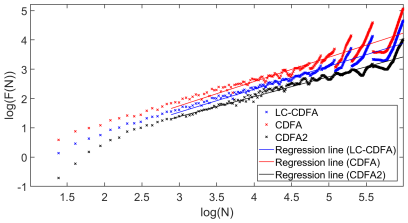
(c) When averaging

Figure 11: Estimation of the LRD with the DFA, the 2^{nd}-order DFA and the LC-DFA with $\mu = 0.4$

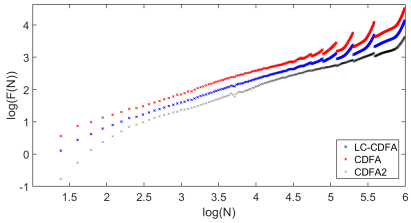
When looking at Figures 11 and 12, averaging makes it possible to smooth the curves and reduce the data spread. In addition, it is no coincidence that if the data associated with the DFA (resp. CDFA) are located above the data associated with the 2^{nd}-order DFA (2^{nd}-order CDFA). This phenomenon is the consequence of the shapes of the frequency responses of the equivalent filters (See properties (39) and (104)). When N is small, it is confirmed that we cannot trust the results obtained when the DFA is used. Consequently these values do not have to be considered to deduce the regression line. This is less the case of the CDFA as the frequency response of the latter is rather band-pass in the low frequency even for N small. Regarding the large values of N, one can notice that depending on the frame, the behavior is not necessarily the same. This is due to the exponential trend of the signal under study. As shown in Figure 13, the resulting profile can be poorly approximated by using a linear-piecewise trend, which leads to these large variations of the logarithm of the square of the fluctuation function when N is large in Figures 11b and 12b. This is really different when increasing the degree of the polynomial. Moreover, in



(a) When dealing with the first frame



(b) When dealing with the last frame



(c) When averaging

Figure 12: Estimation of the LRD with the DFA, the 2^{nd}-order CDFA and the LC-CDFA with $\mu = 0.4$

this example, the DFA tends to overestimate the Hurst exponent H (between 0.7 to 0.85) whereas the DFA tends to underestimate it (between 0.6 to 0.7). Therefore, the LC-DFA with $\mu = 0.4$ can provide an interesting compromise leading to an estimation close to the true value, i.e. 0.7.

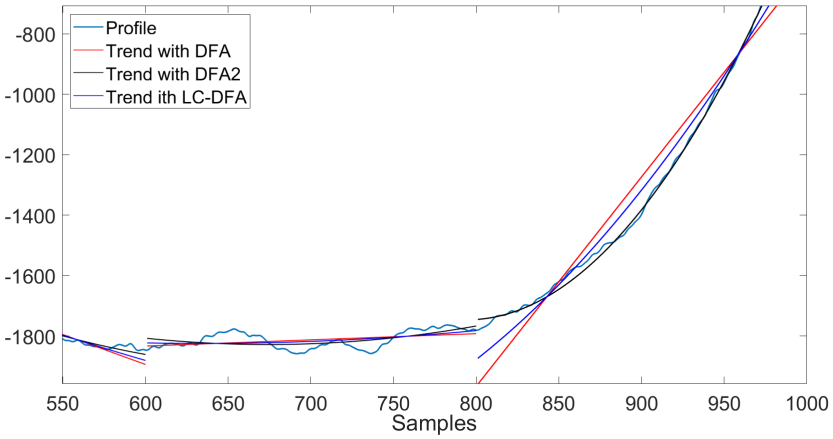


Figure 13: Zoom on the profile and the estimation of the profile trend using the DFA, the 2^{nd}-order DFA and the LC-DFA

4.2. Comparison on a set of Weierstrass signals

In this section, we propose to analyze if the combination of the trends of different orders can be of interest to analyze the LRD. To this end, a set of Weierstrass functions[6] of length 500 whose Hurst coefficient H is predefined is considered. H varies from 0.2 to 0.8 with a step equal to 0.2. The estimation of H is based on the selection of the values of N. For this reason, two different strategies have been considered. The details will be given in Section 4.2.1. Finally, the linear combination of the trends of different orders is defined by the weight μ. When $\mu = 0$, this corresponds to the first method whereas when $\mu = 1$, the combination of the two methods reduces to the second method.
The rest of this section is organized as follows. First, the linear combination of DFA of different orders is addressed. Then, the results for a linear combination of CDFA of different orders are presented.

[6]Weierstrass (WEI) functions are continuous functions that can be derived from nowhere [46]. Each WEI is constructed as a sum of damped sinusoids with increasing frequencies.

4.2.1. Estimation of the Hurst exponent with a linear combination of trends obtained with DFA of different orders: LC-DFA

Two strategies have been considered for the estimation of H.

- Strategy 1: Consider all the values $N \in [\![d+2, M/2]\!] = [\![d+2, 250]\!]$ with d the order of the method.
- Strategy 2: Consider a sub-sample of the values of N depending of the total size of the signal. Here, $N \in [\![M/8, M/2]\!] = [\![62, 250]\!]$.

The purpose of the second strategy is to avoid selecting too small values of N since the fluctuation function is *a priori* known to be underestimated in this case. This underestimation can tend to overestimate the value of the slope α and consequently the Hurst exponent.

Given Tables 1 and 2, one can first notice that the estimation of the Hurst exponent is slightly overestimated whatever the method chosen. For this set of signals and these simulations, if we only look at the results of the DFA and its higher-order variants, the DFA is the most accurate. If one compares these results with the ones obtained for the combinations of methods, one can see that the larger H is, the larger the weight μ must be when combining the standard DFA and the 2^{nd}-order DFA in order to get the smaller error mean. This combination seems to be relevant for both strategies.

4.2.2. Estimation of the Hurst exponent with a linear combination of trends obtained with CDFA of different orders: LC-CDFA

The two strategies to select N that are presented in the DFA are still considered for the combination of the trends obtained with CDFA of different orders. Given Tables 3 and 4, and looking at the CDFA or its higher-order variants only, the 2^{nd}-order CDFA is always the most accurate approach. However, when choosing the first strategy, this method can be outperformed by the combination of CDFA of different orders. More particularly, the combination of CDFA of order 1 and 2 provides the most accurate results when $H = 0.4$, $H = 0.6$ and $H = 0.8$. For $H = 0.2$, even if the combination of CDFA of order 2 and 3 provides

Weight μ and methods	$H = 0.2$	$H = 0.4$	$H = 0.6$	$H = 0.8$
0.2 (DFA of orders 1 and 2)	**-0.010**	-0.005	-0.007	-0.030
0.2 (DFA of orders 1 and 3)	-0.013	-0.009	-0.012	-0.037
0.2 (DFA of orders 2 and 3)	-0.034	-0.030	-0.025	-0.020
0.4 (DFA of orders 1 and 2)	**-0.010**	**-0.003**	-0.003	-0.025
0.4 (DFA of orders 1 and 3)	-0.014	-0.009	-0.011	-0.036
0.4 (DFA of orders 2 and 3)	-0.032	-0.029	-0.024	-0.018
0.6 (DFA of orders 1 and 2)	-0.013	-0.004	**-0.001**	-0.016
0.6 (DFA of orders 1 and 3)	-0.017	-0.010	-0.010	-0.033
0.6 (DFA of orders 2 and 3)	-0.032	-0.030	-0.026	-0.020
0.8 (DFA of orders 1 and 2)	-0.021	-0.012	-0.003	**-0.005**
0.8 (DFA of order 1 and 3)	-0.025	-0.019	-0.013	-0.027
0.8 (DFA of order 2 and 3)	-0.035	-0.035	-0.033	-0.030
DFA of order 1	-0.012	-0.008	-0.011	-0.034
DFA of order 2	-0.028	-0.025	-0.019	-0.013
DFA of order 3	-0.038	-0.041	-0.043	-0.043

Table 1: Mean of the difference between the true Hurst coefficient and the estimated one for different values of the weight μ. Strategy 1 for the selection of N. Bold values correspond to the most accurate results

the most accurate results, the combination of CDFA of order 1 and 2 remains among the most accurate solutions.

When opting for the second strategy, it is more difficult to draw some conclusions. The combination of CDFA of order 1 and 2 or of order 2 and 3 are of interest. For a predefined value H, when combining CDFA of order 1 and 2 or 2 and 3, the error decreases when the weight increases.

Depending on the signal under study and the selection of the values of N, we can see that it is always easy to draw some definitive conclusions. In practice, one has to pay attention to the spectral properties of the signal, the overall temporal representation of the signal, but also the log/log plot.

Weight μ and methods	$H = 0.2$	$H = 0.4$	$H = 0.6$	$H = 0.8$
0.2 (DFA of orders 1 and 2)	-0.011	-0.011	-0.025	-0.060
0.2 (DFA of orders 1 and 3)	-0.018	-0.018	-0.032	-0.066
0.2 (DFA of orders 2 and 3)	-0.037	-0.027	-0.018	**-0.010**
0.4 (DFA of orders 1 and 2)	**-0.009**	-0.005	-0.016	-0.051
0.4 (DFA of orders 1 and 3)	-0.022	-0.020	-0.032	-0.067
0.4 (DFA of orders 2 and 3)	-0.042	-0.032	-0.022	-0.013
0.6 (DFA of orders 1 and 2)	-0.011	**-0.002**	-0.006	-0.035
0.6 (DFA of orders 1 and 3)	-0.033	-0.028	-0.035	-0.067
0.6 (DFA of orders 2 and 3)	-0.051	-0.044	-0.034	-0.025
0.8 (DFA of orders 1 and 2)	-0.022	-0.009	**0.001**	-0.011
0.8 (DFA of order 1 and 3)	-0.054	-0.048	-0.047	-0.067
0.8 (DFA of order 2 and 3)	-0.066	-0.064	-0.060	-0.054
DFA of order 1	-0.016	-0.017	-0.032	-0.066
DFA of order 2	-0.035	-0.026	-0.017	-0.011
DFA of order 3	-0.076	-0.081	-0.086	-0.091

Table 2: Mean of the difference between the true Hurst coefficient and the estimated one for different values of the weight μ. Strategy 2 for the selection of N.

5. Conclusions and perspectives

The purpose of this paper was to analyze a new class of variants of DFA. Indeed, we suggest using a linear combination of the trends obtained with the DFA and the CDFA in their standard formulations with a higher-order variant. To this end, our main concern was to analyze how the frequency response of the equivalent filter representing the centering, integration and detrending steps could evolve. Illustrations on an ARFIMA process show that the proposed combination can lead to a good compromise. The *a priori* choice of the weight remains an open topic and will be probably addressed in future works. This will probably depend on the type of profile trend and the number of samples available.

Weight μ and methods	$H = 0.2$	$H = 0.4$	$H = 0.6$	$H = 0.8$
0.2 (CDFA of orders 1 and 2)	-0.089	-0.086	-0.081	-0.085
0.2 (CDFA of orders 1 and 3)	-0.092	-0.090	-0.085	-0.088
0.2 (CDFA of orders 2 and 3)	-0.074	-0.073	-0.071	-0.069
0.4 (CDFA of orders 1 and 2)	-0.083	-0.081	-0.078	-0.076
0.4 (CDFA of orders 1 and 3)	-0.089	-0.088	-0.083	-0.087
0.4 (CDFA of orders 2 and 3)	**-0.072**	-0.072	-0.070	-0.069
0.6 (CDFA of orders 1 and 2)	-0.077	-0.074	-0.069	-0.072
0.6 (CDFA of orders 1 and 3)	-0.085	-0.085	-0.081	-0.085
0.6 (CDFA of orders 2 and 3)	**-0.072**	-0.073	-0.073	-0.071
0.8 (CDFA of orders 1 and 2)	-0.073	**-0.069**	**-0.063**	**-0.061**
0.8 (CDFA of order 1 and 3)	-0.078	-0.080	-0.078	-0.080
0.8 (CDFA of order 2 and 3)	-0.074	-0.079	-0.081	-0.083
CDFA of order 1	-0.093	-0.090	-0.085	-0.088
CDFA of order 2	-0.074	-0.074	-0.072	-0.071
CDFA of order 3	-0.076	-0.084	-0.091	-0.098

Table 3: Mean of the difference between the true Hurst coefficient and the estimated one for different values of the weight μ. Strategy 1 for the selection of N.

References

[1] 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, vol. 49, (2), pp. 1685–1689, 1994.

[2] 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, pp. 206–210, 2015.

[3] A. A. Pranata, G. W. Adhane, and D. S. Kim, "Detrended fluctuation analysis on ECG device for home environment," Consumer Communications and Networking Conference, pp. 4233–4236, 2017.

[4] A. G. Ravelo-Garcia, U. Casanova-Blancas, S. Martin-González, E. Hernández-Pérez, I. Guerra-Moreno, P. Quintana-Morales, N. Wessel,

Weight μ and methods	$H = 0.2$	$H = 0.4$	$H = 0.6$	$H = 0.8$
0.2 (CDFA of orders 1 and 2)	-0.124	-0.124	-0.120	-0.122
0.2 (CDFA of orders 1 and 3)	-0.133	-0.132	-0.127	-0.127
0.2 (CDFA of orders 2 and 3)	**-0.073**	-0.074	-0.072	-0.070
0.4 (CDFA of orders 1 and 2)	-0.110	-0.109	-0.107	-0.113
0.4 (CDFA of orders 1 and 3)	-0.128	-0.128	-0.125	-0.127
0.4 (CDFA of orders 2 and 3)	-0.077	-0.072	-0.066	**-0.065**
0.6 (CDFA of orders 1 and 2)	-0.094	-0.091	-0.089	-0.097
0.6 (CDFA of orders 1 and 3)	-0.120	-0.123	-0.121	-0.125
0.6 (CDFA of orders 2 and 3)	-0.081	-0.075	-0.069	-0.075
0.8 (CDFA of orders 1 and 2)	-0.080	-0.072	**-0.067**	-0.068
0.8 (CDFA of orders 1 and 3)	-0.115	-0.117	-0.116	-0.121
0.8 (CDFA of orders 2 and 3)	-0.097	-0.100	-0.101	-0.101
CDFA of order 1	-0.135	-0.134	-0.128	-0.128
CDFA of order 2	-0.076	**-0.071**	**-0.067**	-0.066
CDFA of order 3	-0.105	-0.114	-0.123	-0.134

Table 4: Mean of the difference between the true Hurst coefficient and the estimated one for different values of the weight μ. Strategy 2 for the selection of N.

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, pp. 905–908, 2014.

[5] R.U Acharya, C.M Lim, and P Joseph, "Heart rate variability analysis using correlation dimension and detrended fluctuation analysis," ITBM-RBM, vol. 23, no. 6, pp. 333–339, 2002.

[6] R. Kumagai and M. Uchida, "Detrended fluctuation analysis of repetitive handwriting," in 2017 International Conference on Noise and Fluctuations (ICNF), 2017, pp. 1–4.

[7] Wajid Mumtaz, Aamir Malik, Syed Saad Ali, mohd azhar mohd yasin, and Hafeez Ullah Amin, "Detrended fluctuation analysis for major depressive disorder," in Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 08 2015, pp. 4162–4165.

[8] A. Adda and H. Benoudnine, "Detrended fluctuation analysis of eeg recordings for epileptic seizure detection," in 2016 International Conference on Bio-engineering for Smart Technologies (BioSMART), 2016, pp. 1–4.

[9] A. Kitlas Golińska, "Detrended fluctuation analysis (DFA) in biomedical signal processing: Selected examples," Studies in Logic, Grammar and Rhetoric, vol. 29, pp. 107–115, 01 2012.

[10] Aviral Tiwari, Claudiu Albulescu, and Seong-Min Yoon, "A multifractal detrended fluctuation analysis of financial market efficiency: Comparison using dow jones sector etf indices," Physica A: Statistical Mechanics and its Applications, vol. 483, pp. 182–192, 05 2017.

[11] Apostolos Serletis, Olga Yu Uritskaya, and Vadim M. Uritsky, "Detrended fluctuation analysis of the us stock market," International Journal of Bifurcation and Chaos, vol. 18, no. 02, pp. 599–603, 2008.

[12] Peter Talkner and Rudolf Weber, "Power spectrum and detrended fluctuation analysis: Application to daily temperatures," Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics, vol. 62, pp. 150–60, 08 2000.

[13] A. Király and I. M. Jánosi, "Detrended fluctuation analysis of daily temperature records: Geographic dependence over australia," Meteorology and Atmospheric Physics, vol. 88, no. 3-4, pp. 119–128, 2005.

[14] Wen-wen Tung, Jianbo Gao, Jing Hu, and Lei Yang, "Detecting chaos in heavy-noise environments," Phys. Rev. E, vol. 83, pp. 046210, Apr 2011.

[15] M. A. Riley, S. Bonnette, N. Kuznetsov, S. Wallot, and J. Gao, "A tutorial introduction to adaptive fractal analysis," Frontiers in Physiology, vol. 3, pp. 371, 2012.

[16] B. Berthelot, E. Grivel, P. Legrand, J.-M. André, and P. Mazoyer, "Regularized DFA to study the gaze position of an airline pilot," EUSIPCO, pp. 2403–2407, 2020.

[17] B. Berthelot, E. Grivel, P. Legrand, J.-M. André, and P. Mazoyer, "Alternative ways to compare the detendred fluctuation analysis and its variants. application to visual tunneling detection," Digital Signal Processing, vol. 108, pp. 102865, 2020.

[18] E. Alessio, A. Carbone, G. Castelli, and V. Frappietro, "Second-order moving average and scaling of stochastic time series," The European Physical Journal B, vol. 27, 2, pp. 197–200, 2002.

[19] 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, vol. 71, 5, pp. 051101, 2005.

[20] S. Arianos, A. Carbone, and C. Türk, "Self-similarity of higher-order moving averages," Physical Review E, vol. 84(4 Pt 2), pp. 046113, 2011.

[21] B. Berthelot, E. Grivel, and P. Legrand, "New Variants of DFA based on LOESS and LOWESS methods: generalization of the detrended moving average," ICASSP, 2021.

[22] 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, vol. 316, (1-4), pp. 87–114, 2002.

[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, pp. 1–4, 2017.

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

[25] 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, pp. 177–217, 1996.

[26] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde, "Detecting long-range correlations with detrended fluctuation analysis," Physica A: Statistical Mechanics and its Applications, vol. 295, (3-4), pp. 441–454, 2001.

[27] J.-M. Bardet and Imen Kammoun, "Asymptotic properties of the detrended fluctuation analysis of long range dependence processes," IEEE Transactions on Information theory, vol. 54, (5), pp. 2041–2052, 2008.

[28] N. Crato, R. R. Linhares, and S. R. C. Lopes, "Statistical properties of detrended fluctuation analysis," Journal of Statistical Computation and Simulation, vol. 80, (6), pp. 625–641, 2010.

[29] K. Kiyono, "Establishing a direct connection between detrended fluctuation analysis and fourier analysis," Physical Review E, vol. 92, pp. 042925, 2015.

[30] M. Höll and H. Kantz, "The relationship between the detrendend fluctuation analysis and the autocorrelation function of a signal," The European Physical Journal B, vol. 88, pp. 327, 2015.

[31] M. Höll, H. Kantz, and Y. Zhou, "Detrended fluctuation analysis and the difference between external drifts and intrinsic diffusion-like non-stationarity," Physical Review E, vol. 94, pp. 042201, 2016.

[32] K. Kiyono, "Theory and applications of detrending-operation-based fractal-scaling analysis," International Conference on Noise and Fluctuations, pp. 1–4, 2017.

[33] O. Løvsletten, "Consistency of detrended fluctuation analysis," Physical Review E, vol. 96, pp. 012141, 2017.

[34] M. Höll, K. Kiyono, and H. Kantz, "Theoretical foundation of detrending methods for fluctuation analysis such as detrended fluctuation analysis and detrending moving average," Physical Review E, vol. 99, pp. 033305, 2019.

[35] B. Berthelot, E. Grivel, P. Legrand, M. Donias, J.-M. André, P. Mazoyer, and T. Ferreira, "2D Fourier Transform Based Analysis Comparing the DFA with the DMA," EUSIPCO, pp. 1–5, 2019.

[36] G. Sikora, M. Höll, J. Gajda, H. Kantz, A. Chechkin, and A. Wyloman-ska, "Probabilistic properties of detrended fluctuation analysis for gaussian processes," Physical Review E, vol. 101, pp. 032114, 2020.

[37] E. Grivel, B. Berthelot, P. Legrand, and A. Giremus, "Dfa-based abacuses providing the hurst exponent estimate for short-memory processes," Digital Signal Processing, vol. 116, pp. 103102, 2021.

[38] Ken Kiyono and Yutaka Tsujimoto, "Nonlinear filtering properties of detrended fluctuation analysis," Physica A: Statistical Mechanics and its Applications, vol. 462, pp. 807–815, 2016.

[39] D. Maraun, H. W. Rust, and J. Timmer, "Tempting long-memory on the interpretation of dfa results," Nonlinear Processes in Geophysics, pp. 495–503, 2004.

[40] V. V. Morariu, L. Buimaga-Iarinca, C. Vamoş, and S. M. Soltuz, "Detrended fluctuation analysis of autoregressive processes," submitted to Fluctuation and Noise Letters, p. arXiv:0707.1437, 2007.

[41] M. Höll and H. Kantz, "The fluctuation function of the detrended fluctuation analysis — investigation on the ar(1) process," The European Physical Journal B, vol. 88, pp. 126, 2015.

[42] W. Palma, Long-Memory Time Series, Theory and Methods, Wiley series in probability and statistics, 2007.

[43] Y. Yang, "Adaptive regression by mixing," Journal of the American Statistical Association, vol. 96, Issue 54, pp. 574–588, 2001.

[44] P. Montero-Manso, G. Athanasopoulos, R. Hyndman, and T. Talagala, "Fforma: Feature-based forecast model averaging," International Journal of Forecasting, vol. 36, pp. 86–92, 2020.

[45] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, Estimation With Applications To Tracking and Navigation, Wiley, Hoboken, NJ, USA, 2001.

[46] G.H. Hardy, "On weierstrass' non-differentiable functions," Trans. of the American Mathematical Society, vol. 17, (3), pp. 301–325, 1916.