



A Fast DFA Algorithm for Multifractal Multiscale Analysis of Physiological Time Series

Paolo Castiglioni^{1*} and Andrea Faini²

¹ IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy, ² Department of Cardiovascular Neural and Metabolic Sciences, Istituto Auxologico Italiano, IRCCS, S.Luca Hospital, Milan, Italy

OPEN ACCESS

Edited by:

Zbigniew R. Struzik,
The University of Tokyo, Japan

Reviewed by:

Piotr Suffczynski,
University of Warsaw, Poland
Stanislaw Drozd,
Institute of Nuclear Physics (PAN),
Poland

*Correspondence:

Paolo Castiglioni
pcastiglioni@dongnocchi.it
orcid.org/0000-0002-8775-2605

Specialty section:

This article was submitted to
Computational Physiology and
Medicine,
a section of the journal
Frontiers in Physiology

Received: 08 October 2018

Accepted: 30 January 2019

Published: 01 March 2019

Citation:

Castiglioni P and Faini A (2019) A Fast
DFA Algorithm for Multifractal
Multiscale Analysis of Physiological
Time Series. *Front. Physiol.* 10:115.
doi: 10.3389/fphys.2019.00115

Detrended fluctuation analysis (DFA) is a popular tool in physiological and medical studies for estimating the self-similarity coefficient, α , of time series. Recent researches extended its use for evaluating multifractality (where α is a function of the multifractal parameter q) at different scales n . In this way, the multifractal-multiscale DFA provides a bidimensional surface $\alpha(q, n)$ to quantify the level of multifractality at each scale separately. We recently showed that scale resolution and estimation variability of $\alpha(q, n)$ can be improved at each scale n by splitting the series into maximally overlapped blocks. This, however, increases the computational load making DFA estimations unfeasible in most applications. Our aim is to provide a DFA algorithm sufficiently fast to evaluate the multifractal DFA with maximally overlapped blocks even on long time series, as usually recorded in physiological or clinical settings, therefore improving the quality of the $\alpha(q, n)$ estimate. For this aim, we revise the analytic formulas for multifractal DFA with first- and second-order detrending polynomials (i.e., DFA₁ and DFA₂) and propose a faster algorithm than the currently available codes. Applying it on synthesized fractal/multifractal series we demonstrate its numerical stability and a computational time about 1% that required by traditional codes. Analyzing long physiological signals (heart-rate tachograms from a 24-h Holter recording and electroencephalographic traces from a sleep study), we illustrate its capability to provide high-resolution $\alpha(q, n)$ surfaces that better describe the multifractal/multiscale properties of time series in physiology. The proposed fast algorithm might, therefore, make it easier deriving richer information on the complex dynamics of clinical signals, possibly improving risk stratification or the assessment of medical interventions and rehabilitation protocols.

Keywords: hurst exponent, multiscale analysis, multifractality, HRV, EEG

INTRODUCTION

In the mid of the 90s, the algorithm of detrended fluctuation analysis (DFA) gave a great boost to the study of fractal physiology providing an easy-to-calculate method for evaluating the Hurst's exponent of physiological times series (Castiglioni et al., 2017). DFA estimates the slope α of the log-log plot of a fluctuations function, $F(n)$, over a range of time scales n . The fluctuations function is the root mean square, or second-order moment, of the deviations of blocks of n data from a polynomial trend. Due to the intrinsic variability of the $F(n)$ estimate, α is calculated as the slope of the least-square regression line fitting $\log F(n)$ on $\log n$.

DFA became quickly popular in the field of heart rate variability analysis and since its first applications the $F(n)$ of heart rate was described by two slopes, one at the shorter ($n < 16$) and one at the longer scales (Peng et al., 1995). Then, similar multi-slope approaches were proposed for electroencephalogram (EEG) recordings (Hwa and Ferree, 2002; Jospin et al., 2007). The idea that α may change with the scale n and that the resulting $\alpha(n)$ profile may provide information on physiological or pathophysiological mechanisms was further investigated by different groups in the following years. Continuous $\alpha(n)$ profiles from noisy $F(n)$ estimates were obtained by applying a simplified Kalman filter whose coefficients, however, were defined empirically, making the procedure somehow arbitrary (Echeverria et al., 2003). Other authors extended the two-slope approach fitting the regression line over a short running window (Gieraltowski et al., 2012; Xia et al., 2013): in this way a continuous $\alpha(n)$ profile was obtained as the central point of the window, n , moved from the shortest to the longest scale. The running-window approach was also applied to study multifractality. Multifractal series are composed by interwoven fractal processes and the multifractal DFA approach extends the calculation of the second-order fluctuations function to a range of positive and negative moments that includes $q = 2$ (the second-order moment): in fact, positive q moments amplify the contribution of fractal components with larger amplitude and negative q moments the contribution of fractal components with smaller amplitude (Kantelhardt et al., 2002). This led to the calculation of the local slope of the q^{th} -order fluctuation function, $\alpha(q, n)$, as the slope of the regression lines fitting $F_q(n)$ over a running window at each q separately (Gieraltowski et al., 2012). It should be considered that the scale resolution achievable with the regression line method is limited because the window width cannot be too short to make α insensitive to the estimation variability.

A third approach is based on estimating the slope as the first derivative of $\log F(n)$ (Castiglioni et al., 2009, 2011a). Assuring higher $\alpha(n)$ resolution, it allows obtaining a more detailed description of the deviations of $\log F(n)$ from the straight line, which resulted useful for describing the autonomic integrative control (Castiglioni et al., 2011a) and its alterations (Castiglioni and Merati, 2017). However, this approach requires minimizing the variability of the $F(n)$ estimator because of its sensitivity to noise. Traditional DFA estimators calculate $F(n)$ splitting the series into consecutive non-overlapped blocks. If the series is split into maximally overlapped blocks, which means that consecutive blocks of size n have $n-1$ samples in common, the variability of the $F(n)$ estimator decreases substantially, allowing the evaluation of the slope as the first derivative (Castiglioni et al., 2018). A drawback of this approach is the much higher computational load that may make unfeasible the analysis of long series of data, as often occur in heart-rate variability or EEG studies. The computational load is even greater for calculating $F_q(n)$ and the multifractal-multiscale spectrum, $\alpha(q, n)$.

Strictly related to the determination of the local slope of $F_q(n)$ is the identification of number and position of crossover

scales where the fractal properties change. Crossovers may be due to interferences affecting the measures (Ludescher et al., 2011) or may reflect specific aspects of the physiological system generating the series, like short-range correlations (Höll and Kantz, 2015). Statistical models for identifying crossover points in $F_q(n)$ are based on fitting piecewise regressions and on iterative hypothesis testing (Ge and Leung, 2013). Other approaches identify the scales where the $F_q(n)$ curves change their slope by introducing the focus-based multifractal formalism (Mukli et al., 2015): accordingly, if one assumes a bimodal scale dynamics, the multifractal fluctuations function is directly modeled by one or two fan-like structures in which log-log straight lines, each corresponding to an order q , may converge to a focus at the largest scale. Multifractal crossovers are then decomposed by an iterative process that minimizes the residual errors of least-square fittings (Nagy et al., 2017; Mukli et al., 2018). Clearly, the estimation variability of the $F_q(n)$ curves negatively affects the identification of the focuses and decreases the overall statistical power of these methods. Therefore, also these methods can take advantage of the use of maximally overlapped blocks for improving the $F_q(n)$ estimates and for better modeling the multifractal dynamics.

The examples of **Figure 1** illustrate the reduction of estimation variability and the related computational costs when using maximum overlapping. In the example of a first-order autoregressive process generated by white Gaussian noise (**Figure 1**, Upper), the variability of the $F_q(n)$ estimate without overlapping might suggest the presence of a focus at the larger scales. A focus should not be present, because at scales larger than the cut-off frequency the dynamics is determined by white noise. Actually, when the $F_q(n)$ curves are estimated with maximally overlapped blocks, they run in parallel (see inset) with slope α close to 0.5 at the largest scale. In the example of a random series with multifractal Cauchy distribution (**Figure 1**, Lower), the large variability of the estimate without overlapping makes it difficult to identify the scale where the focus occurs and may completely hide possible crossover points between $n = 10^3$ and $n = 10^4$ for positive q . These drawbacks are largely mitigated by using maximum overlapping but at the cost of an increased computational load that may lead to unacceptable calculation times for long series.

To overcome this limit, we designed a fast algorithm for $F_q(n)$ calculation, particularly efficient in case of maximally overlapped blocks. Therefore, the aim of the present work is (1) to describe the theoretical aspects on which we based our fast algorithm; (2) to illustrate its performance by comparison with the traditional DFA algorithm; (3) to show potential applications in the fields of heart rate variability analysis and of EEG signals analysis; and (4) to provide the source code for its implementation. The algorithm optimizes the computation time of $F_q(n)$ and readers can then estimate $\alpha(q, n)$ with any of the methods proposed in literature: Kalman filtering, least-square regression over a running window or first derivative of $\log F_q(n)$ vs. $\log n$. In this work, we also include the code for slope estimation with the latter method because our fast

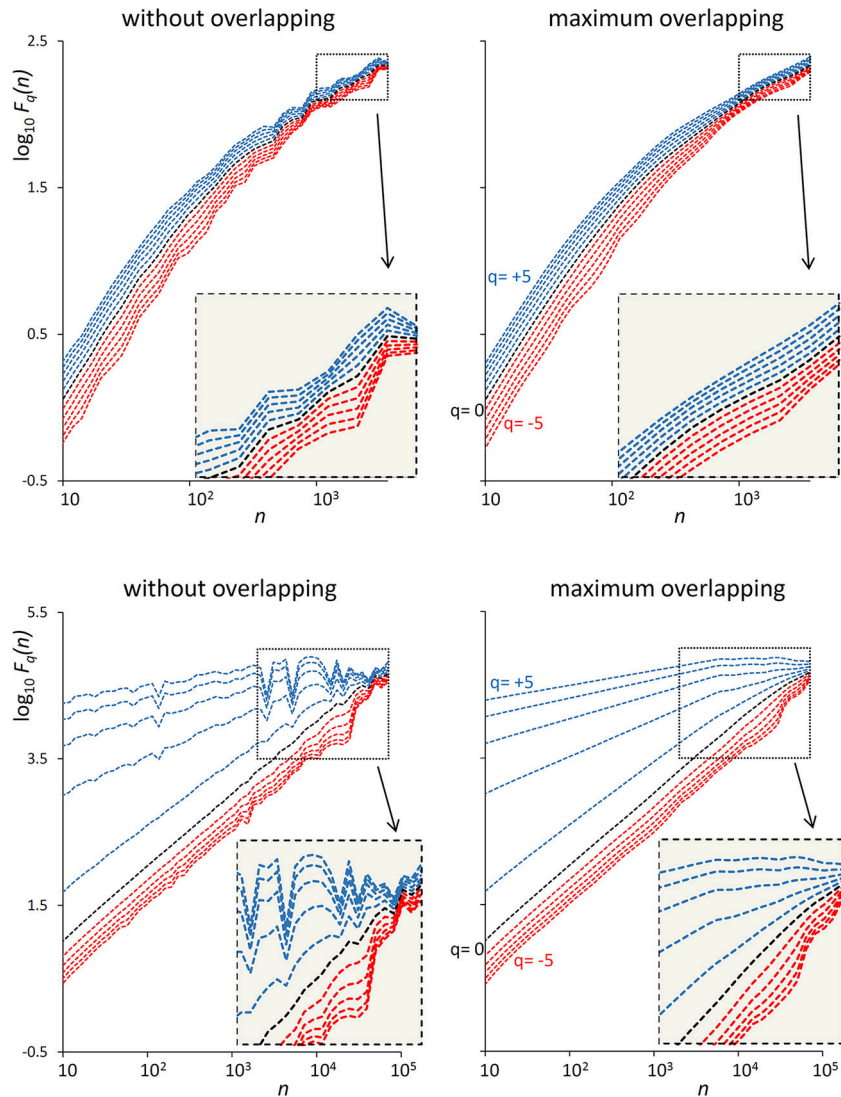


FIGURE 1 | Effect of blocks overlapping on variability and computation time of $F_q(n)$ estimates. **(Upper)** DFA₁ fluctuation functions for an autoregressive series $\{x_i\}$ of $N = 16,384$ samples generated as $x_i = ax_{i-1} + wn_i$ where wn_i is white Gaussian noise with zero mean and unit variance and $a = 0.9391014$ as in Kiyono (2015), corresponding to a low-pass filter with cut-off frequency $f_{CO} = 0.01$, or a cross-over scale $n_{CO} = 100$ (Höll and Kantz, 2015); calculation time for estimating $F_q(n)$ over 38 block-sizes n was 4 s without overlapping and 3 min with maximum overlapping on a personal computer. **(Lower)** DFA₁ fluctuation functions for $N = 300,000$ random samples with Cauchy distribution (location 0, scale 3); calculation time for estimating $F_q(n)$ over 55 block-sizes n was 58 s without overlapping, 5 h and 22 min with maximum overlapping.

algorithm was specifically designed for evaluating $\alpha(q, n)$ with the derivative approach.

FAST $F_q(n)$ CALCULATION

Given a time series of N samples, S_j , with $j = 1, \dots, N$, mean μ_S and standard deviation σ_S , the mean is removed and the series normalized to unit standard deviation:

$$s_j = \frac{S_j - \mu_S}{\sigma_S} \tag{1}$$

The normalization in Equation (1), not required by DFA, is useful for recognizing data blocks with very low dynamics (see section *Calculation precision and small blocks overfitting*). The cumulative sum y_i , with $i = 1, \dots, N$, is:

$$y_i = \sum_{j=1}^i s_j \tag{2}$$

To evaluate F_q , the y_i series is split into consecutive blocks of size n . At a given scale n , the number of blocks, M , depends on the length of the series N , on the block size n and on the

number of overlapped samples between consecutive blocks, L , with $0 \leq L < n$, as:

$$M = \left\lfloor \frac{N - n}{n - L} \right\rfloor + 1 \tag{3}$$

If $N-n$ is not multiple of $n-L$, a short segment of $N-(M-1)(n-L)-n$ samples at the end of the series are excluded from the analysis. In the case of maximum overlapping, when $L = n-1$, all the N samples contribute to the estimation of $F_q(n)$ and the number of blocks reaches the highest value: $M = N-n+1$. Each of the M blocks is detrended by least-square fitting a polynomial $p(i)$. The usual notation to indicate a DFA employing a polynomial of order O is DFA_O . The variance of the residuals in the k -th detrended block is

$$\sigma_n^2(k) = \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} (y_i - p(i))^2 \tag{4a}$$

with $I_k = (n-L) \times (k-1) + 1$ index of the first sample falling into the k -th block. According to Kantelhardt et al. (2002), the variability functions for DFA multifractal analysis are:

$$\begin{cases} F_q(n) = \left(\frac{1}{M} \sum_{k=1}^M (\sigma_n^2(k))^{q/2} \right)^{1/q} & \text{for } q \neq 0 \\ F_q(n) = e^{\frac{1}{2M} \sum_{k=1}^M \ln(\sigma_n^2(k))} & \text{for } q = 0 \end{cases} \tag{5}$$

Because of the normalization in Equation (1), the variability function of the original series is $F_q(n) \times \sigma_S$.

Most of the computational load for evaluating $F_q(n)$ is due to the calculation of Equation (4a), which requires (1) a least-square polynomial fitting and (2) the evaluation of the variance of the residuals, over n points. These steps should be repeated for each of the M blocks and M reaches very large values, close to N , for maximally overlapped blocks (Equation 3). In the following, we describe how to make these calculation steps faster for detrending polynomials of order 1 or 2, i.e., for DFA_1 and DFA_2 . We did not consider polynomials of a higher order because all the DFA biomedical applications we are aware of did not employ detrending polynomials of order greater than 2. However, the same optimization strategy we present here can be extended to higher orders.

Least Square Polynomial Fitting

Freely available DFA codes, as accessible in Ihlen (2012) or described in Peng et al. (1995) and Gieraltowski et al. (2012) and downloadable in Goldberger et al. (2000), calculate the polynomial $p(i)$ of order O with algorithms for least-square fitting a data-set of n points with coordinates (i, y_i) . This means solving n linear equations operating with a $n \times (O+1)$ Vandermonde matrix. The fitting polynomials of first and second order are:

$$y = b_1 i + a_1 \tag{6a}$$

$$y = c_2 i^2 + b_2 i + a_2 \tag{6b}$$

Coefficients of Equation (6a) are:

$$b_1 = \frac{S_{iy}}{S_{ii}} \tag{7a}$$

$$a_1 = \bar{y} - b_1 \bar{i} \tag{7b}$$

with

$$\bar{i} = \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} i \tag{8a}$$

$$\bar{y} = \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} y_i \tag{8b}$$

$$S_{iy} = \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} i y_i - \bar{i} \bar{y} \tag{8c}$$

$$S_{ii} = \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} i^2 - \bar{i}^2 \tag{8d}$$

Coefficients of Equation (6b) are:

$$c_2 = \frac{S_{i^2y} S_{ii} - S_{iy} S_{ii^2}}{S_{ii} S_{i^2i^2} - (S_{ii^2})^2} \tag{9a}$$

$$b_2 = \frac{S_{iy} S_{i^2i^2} - S_{i^2y} S_{ii^2}}{S_{ii} S_{i^2i^2} - (S_{ii^2})^2} \tag{9b}$$

$$a_2 = \bar{y} - b_2 \bar{i} - c_2 \bar{i}^2 \tag{9c}$$

with, in addition to Equation (8a-d):

$$\bar{i}^2 = \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} i^2 \tag{10a}$$

$$S_{ii^2} = \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} i^3 - \bar{i} \bar{i}^2 \tag{10b}$$

$$S_{i^2i^2} = \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} i^4 - \bar{i}^2 \bar{i}^2 \tag{10c}$$

$$S_{i^2y} = \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} i^2 y_i - \bar{i}^2 \bar{y} \tag{10d}$$

Equations (8a,d) and Equations (10a-c) require summing powers “ V ” of i , with V an integer between 1 and 4 and i consecutive natural numbers. Remembering the Faulhaber’s formula:

$$\sum_{i=1}^n i^V = \frac{1}{V+1} \sum_{i=1}^{V+1} (-1)^{\delta_{iV}} \binom{V+1}{i} B_{V+1-i} n^i \tag{11}$$

where B_i is the i -th Bernoulli's number, $\binom{V}{i}$ is the binomial coefficient, and δ_{iV} is the Kronecker's delta (Weisstein, 1999), we have:

$$\sum_{i=1}^n i = \frac{n^2}{2} + \frac{n}{2} \tag{12a}$$

$$\sum_{i=1}^n i^2 = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6} \tag{12b}$$

$$\sum_{i=1}^n i^3 = \frac{n^4}{4} + \frac{n^3}{2} + \frac{n^2}{4} \tag{12c}$$

$$\sum_{i=1}^n i^4 = \frac{n^5}{5} + \frac{n^4}{2} + \frac{n^3}{3} - \frac{n}{30} \tag{12d}$$

Since

$$\sum_{i=I_k}^{I_k+n-1} i^V = \sum_{i=1}^{I_k+n-1} i^V - \sum_{j=1}^{I_k-1} j^V \tag{13}$$

the sum of powers of n consecutive integers does not require the actual summation of n terms and can be obtained through the much faster expressions of Equations (12). However, Equations (8b,c, 10d) include sums of the product between $i^V y_i$, with V integer between 0 and 2. There are no analytic expressions for these sums but a fast way to calculate them is to define the arrays $A_{V,W}(i)$, $i = 0, \dots, N$, as:

$$\begin{cases} A_{V,W}(i) = 0 & \text{for } i = 0 \\ A_{V,W}(i) = A_{V,W}(i-1) + i^V y_i^W & \text{for } i > 0 \end{cases} \tag{14}$$

Once the arrays have been initialized, the sum of n consecutive $i^V y_i^W$ products is calculated by the following difference:

$$\sum_{i=I_k}^{I_k+n-1} i^V y_i^W = A_{V,W}(I_k+n-1) - A_{V,W}(I_k-1) \tag{15}$$

Variance of the Residuals

To avoid another summation over n points, let's writing the squared expression in Equation (4a) for the linear fitting (Equation 6a) as:

$$\begin{aligned} \sigma_n^2(k) &= \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} [y_i - (b_1 i + a_1)]^2 = \\ &= a_1^2 + \frac{1}{n} \left[\sum_{i=I_k}^{I_k+n-1} y_i^2 + b_1^2 \sum_{i=I_k}^{I_k+n-1} i^2 \right. \\ &\quad \left. + 2a_1 b_1 \sum_{i=I_k}^{I_k+n-1} i - 2b_1 \sum_{i=I_k}^{I_k+n-1} i y_i - 2a_1 \sum_{i=I_k}^{I_k+n-1} y_i \right] \end{aligned} \tag{16a}$$

and for quadratic fitting (Equation 6b) as:

$$\begin{aligned} \sigma_n^2(k) &= \frac{1}{n} \sum_{i=I_k}^{I_k+n-1} [y_i - (c_2 i^2 + b_2 i + a_2)]^2 = \\ &= a_2^2 + \frac{1}{n} \left[c_2^2 \sum_{i=I_k}^{I_k+n-1} i^4 + 2c_2 b_2 \sum_{i=I_k}^{I_k+n-1} i^3 + (b_2^2 + 2a_2 c_2) \sum_{i=I_k}^{I_k+n-1} i^2 \right. \\ &\quad \left. + 2b_2 a_2 \sum_{i=I_k}^{I_k+n-1} i \right] + \\ &\quad + \frac{1}{n} \left[\sum_{i=I_k}^{I_k+n-1} y_i^2 - 2c_2 \sum_{i=I_k}^{I_k+n-1} i^2 y_i - 2b_2 \sum_{i=I_k}^{I_k+n-1} i y_i - 2a_2 \sum_{i=I_k}^{I_k+n-1} y_i \right] \end{aligned} \tag{16b}$$

The sums in Equations (16a,b) are obtained analytically with Equations (12a-d, 13), and directly from Equation (15) with $V = 0$ and $W = 1$ or 2, and with $W = 1$ and $V = 1$ or 2.

CALCULATION PRECISION

To avoid subtracting terms with very different magnitude, which may introduce errors due to the finite precision of numbers representation, the starting point of the k -th block, I_k , is shifted to 1 rewriting Equation (4a) as:

$$\sigma_n^2(k) = \frac{1}{n} \sum_{j=1}^n (y_{j+I_k-1} - p'(j))^2 \tag{4b}$$

The difference in Equation (13) is no more required and Equation (15) is calculated as in **Appendix**. However, numerical errors may affect $A_{V,W}(i)$ in Equation (14). In fact, the array initialization requires adding the term $i^V y_i^W$ to a running summation $A_{V,W}(i-1)$ which may reach very high values for long series when i is close to N , propagating precision errors from the lower to the higher indices i . To avoid this problem, the running summation is split into a sequence of shorter sums over segments of 128 samples. This is done defining the matrix $A_{V,W}(r,c)$ of 129 rows ($0 \leq r \leq 128$) and $Q = \lfloor (N-1)/128 \rfloor + 1$ columns ($1 \leq c \leq Q$), as:

$$\begin{cases} A_{V,W}(r,c) = 0 & \text{for } r = 0 \\ A_{V,W}(r,c) = A_{V,W}(r-1,c) \\ \quad + [r + (c-1) \times 128]^V y_{r+(c-1) \times 128}^W & \text{for } r > 0 \end{cases} \tag{17}$$

Any index $i > 0$ corresponds to the column $c = \lfloor (i-1)/128 \rfloor + 1$ and to the row $r = i - c \times 128$, and $A_{V,W}(i)$ is calculated as

$$\begin{cases} A_{V,W}(i) = A_{V,W}(r,c) & \text{for } c = 1 \\ A_{V,W}(i) = A_{V,W}(r,c) + \sum_{k=1}^{c-1} A_{V,W}(128,k) & \text{for } c > 1 \end{cases} \tag{18}$$

avoiding most of the sums that include terms with very different magnitude.

For extremely long series, even Equation (18) cannot exclude errors due to the finite precision of numbers representation.

To control further these errors, the variance of the residuals in Equation (4b) is compared with a threshold. The idea is that when the variance of the residuals is extremely low, the relative weight of errors due to the finite precision of numbers representation might have a detectable influence on the estimate. In this case, the variance is recalculated *directly* as the summation of the n consecutive powers $i^V y_i^W$ and not through (Equation 15). Because of the normalization in Equation (1), the threshold does not depend on the standard deviation of the time series, but only on its length N , on the precision of number representation and on the detrending order. In our implementation the default value of the DFA₁ threshold, Th_1 , is

$$\begin{cases} Th_1 = 10^{-3} & \text{for } N \leq 10^5 \\ Th_1 = \frac{N}{10^8} & \text{for } N > 10^5 \end{cases} \quad (19a)$$

and of the DFA₂ threshold, Th_2 , is

$$\begin{cases} Th_2 = 10^{-2} & \text{for } N \leq 10^2 \\ Th_2 = \frac{N}{10^4} & \text{for } N > 10^2 \end{cases} \quad (19b)$$

A numerical problem of different nature may require comparing the variance of the residuals in Equation (4b) with another

threshold EPS , as described in Ihlen (2012). A detrending polynomial of too high order might overfit the blocks of smaller size n , or pre-processing procedures, as low-pass filtering, might remove fractal components at the shortest scales. In these cases, the variance of the residuals could be so low to severely distort $\log F_q(n)$ at negative q . The solution proposed in Ihlen (2012) is to discard variances lower than an EPS threshold set “to the precision of the measurement device that is recording the biomedical time series.” Since we normalized the series to unit variance, our approach is to set EPS as a fraction of the dynamics of the original series (for instance, $EPS=10^{-4}$).

ALGORITHM PERFORMANCE

To assess numerical stability and speed of our algorithm (see Matlab source code in the **Supplemental File** *FMDFDA.m*), we applied it to synthesized time series. We synthesized three series of $N = 10^6$ samples: a white Gaussian noise with zero mean and unit variance, $\{wn_i\}$; a Brownian motion as the integral of a zero-mean white noise with variance = 0.01986918, $\{Bm_i\}$; and

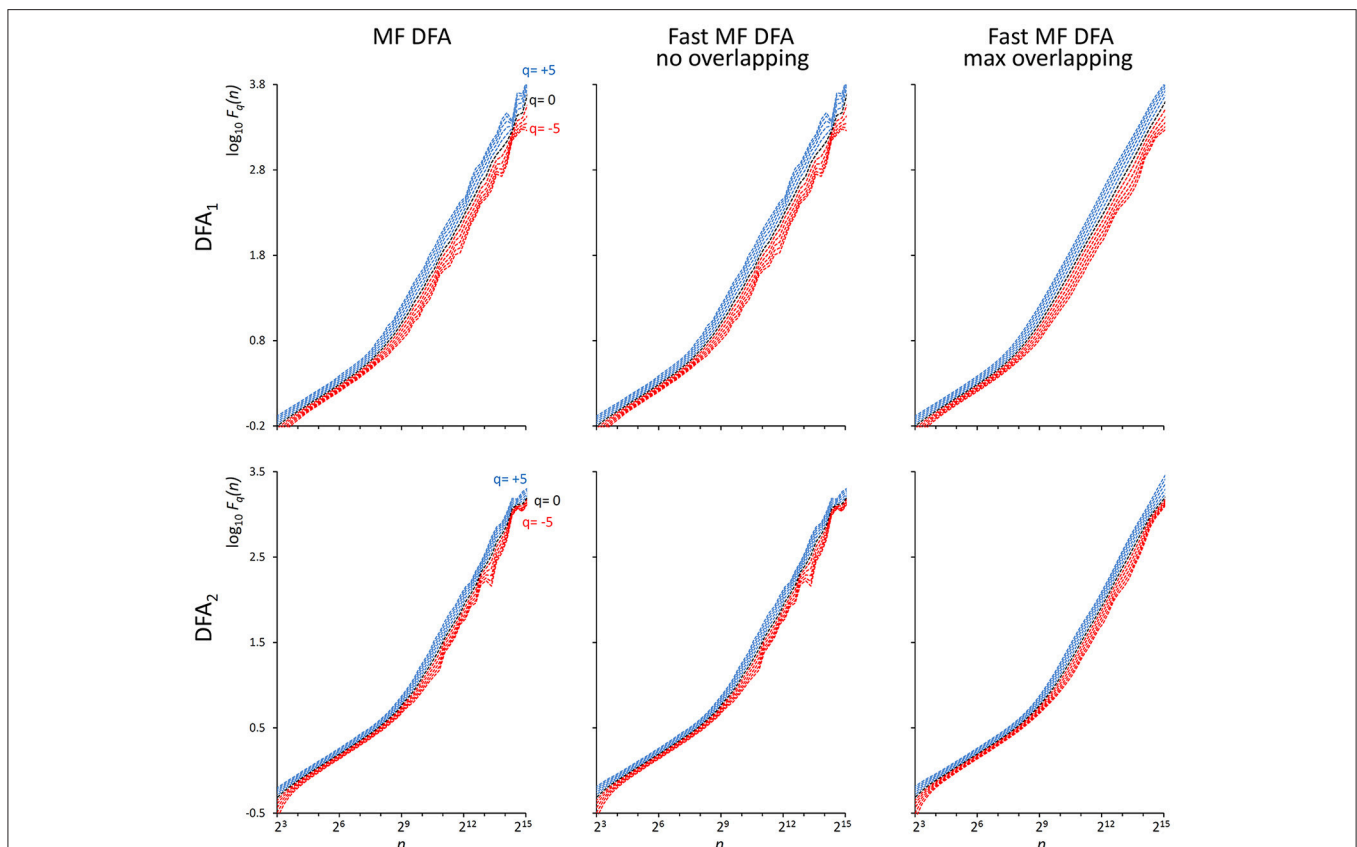


FIGURE 2 | $F_q(n)$ functions by the traditional (MF DFA) and the proposed fast algorithm (Fast MF DFA) for the $\{wb_i\}$ series. The series of $N = 200,000$ samples is the superposition of white noise and Brownian motion (see text); estimates without overlapping by the traditional algorithm (**Left**) and by our fast algorithm (**Center**) and with maximum overlapping by our fast algorithm (**Right**) were performed over 56 block sizes exponentially distributed over the n axis, with thresholds Th_1 and Th_2 as in Equations (19) and $EPS = 0$. On a laptop computer, the traditional algorithm employed $T = 57$ s for calculating DFA₁ and $T = 146$ s for calculating DFA₂ without overlapping, the fast algorithm employed 2.2 s without overlapping and 65 s with maximum overlapping for calculating both the DFA₁ and DFA₂ estimates simultaneously.

the series $\{wb_i\}$ obtained as their superposition, i.e.:

$$wb_i = wn_i + Bm_i \quad (20)$$

As indicated in Kiyono (2015), the power spectrum of $\{wb_i\}$ has a crossover frequency separating the spectral profiles of white

TABLE 1 | Largest relative error in $F_q(n)$ estimates induced by calculating the sum of n products with the fast Equation (15).

Time series					
$\{wn_i\}; \{Bm_i\}; \{wb_i\};$			$\{bcn_i\};$		
Length	DFA ₁	DFA ₂	Length	DFA ₁	DFA ₂
$N = 10^2$	1.4E-11	1.9E-11	$N = 2^8$	3.1E-11	2.1E-11
$N = 10^3$	1.0E-09	2.3E-08	$N = 2^{10}$	4.8E-10	1.2E-10
$N = 10^4$	1.0E-07	5.0E-07	$N = 2^{12}$	2.6E-09	3.9E-09
$N = 10^5$	1.9E-05	6.1E-06	$N = 2^{14}$	1.1E-06	4.0E-07
$N = 10^6$	6.0E-03	4.4E-04	$N = 2^{16}$	4.6E-06	2.2E-06

$\{wn_i\}$, white noise; $\{Bm_i\}$, Brownian motion; $\{wb_i\}$, sum of $\{wn_i\}$ and $\{Bm_i\}$; $\{bcn_i\}$, stochastic binomial cascade embedded in noise (see text for details).

and brown noises at $f_{co} = 10^{-2.5}$, corresponding to a crossover scale $n_{co} = 316$ samples. **Figure 2** shows an example of DFA₁ and DFA₂ estimates on $N = 200,000$ samples of $\{wb_i\}$ (we analyze EEG series of the same length in the next paragraph “Application on real biomedical time series”).

To evaluate how the finite precision of number representation may affect the sum of n consecutive products when obtained by Equation (15), we extracted segments of length $N = 10^{k+1}$, with k integer between 1 and 5, from the three synthesized series. For each segment, we estimated $F_q(n)$ for all integers q between -5 and $+5$ and for scales $n_b = 10^b$ with b an integer between 1 and $k-1$ (e.g., for a segment of length $N_4 = 10,000$ samples we considered the scales $n_1 = 10$, $n_2 = 100$ and $n_3 = 1,000$ samples). Estimates were performed with the fast algorithm with maximally overlapped blocks, setting Th_1 and Th_2 as in Equations (19) and $EPS = 0$. These “fast” estimates were compared with estimates from a reference algorithm in which the difference on the right side of Equation (15) is replaced by the slower but more precise summation of n products on the left side of the equation. The relative error between the fast estimate, $F_q(n_b)^F$, and the reference estimate, $F_q(n_b)^R$, was computed for each q and each block

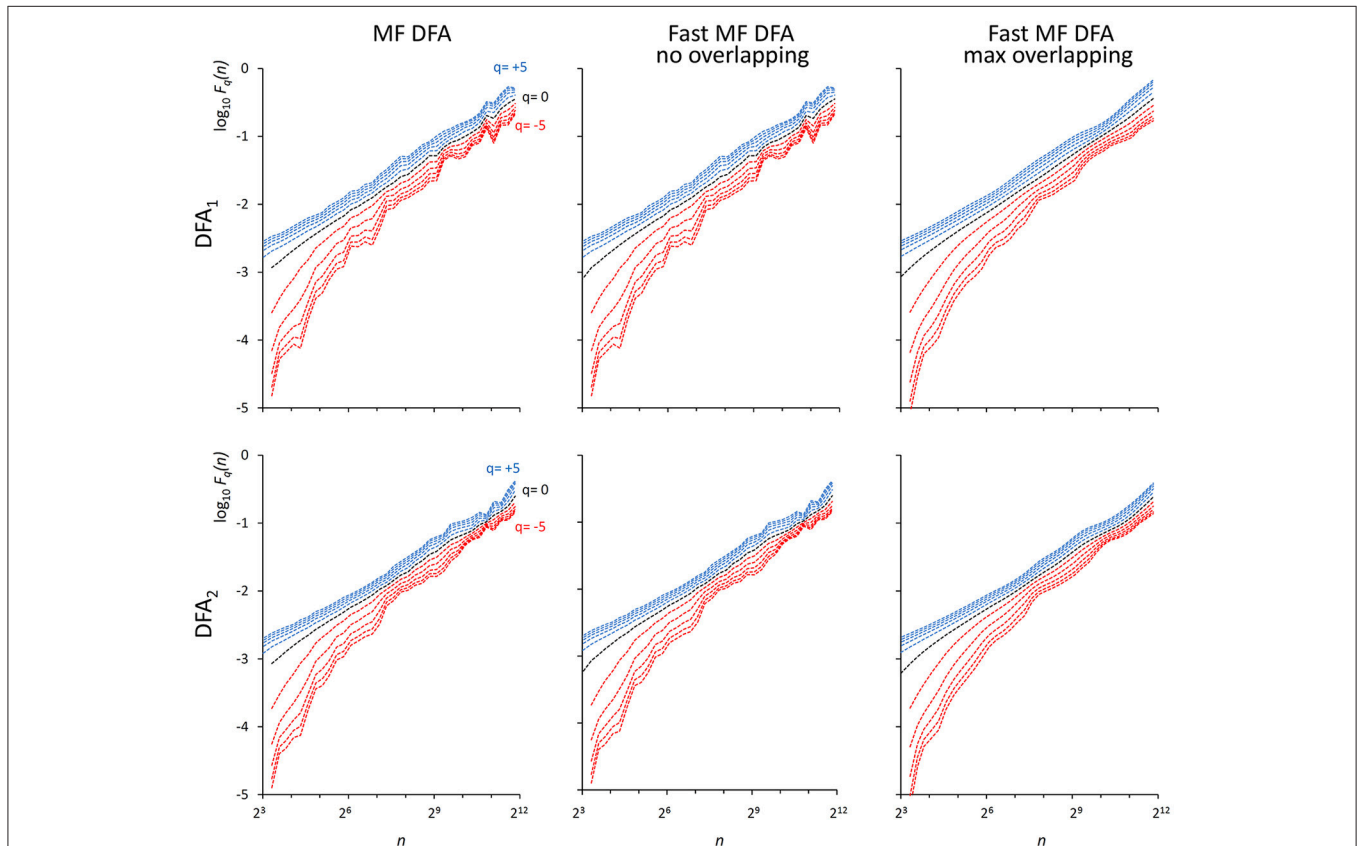


FIGURE 3 | $F_q(n)$ functions for a multifractal binomial cascade embedded in noise of $N = 65,536$ samples. Estimates without overlapping by the traditional (Left) and fast (Center) algorithm and with maximum overlapping by the fast algorithm (Right) were performed over 38 block sizes, with thresholds as in Equations (19) and $EPS = 0$. On a laptop computer, the traditional algorithm employed $T = 4.3$ s for DFA₁ and $T = 9.5$ s for DFA₂ without overlapping while the fast algorithm employed 0.12 s without overlapping and 3.6 s with maximum overlapping for calculating DFA₁ and DFA₂ estimates simultaneously.

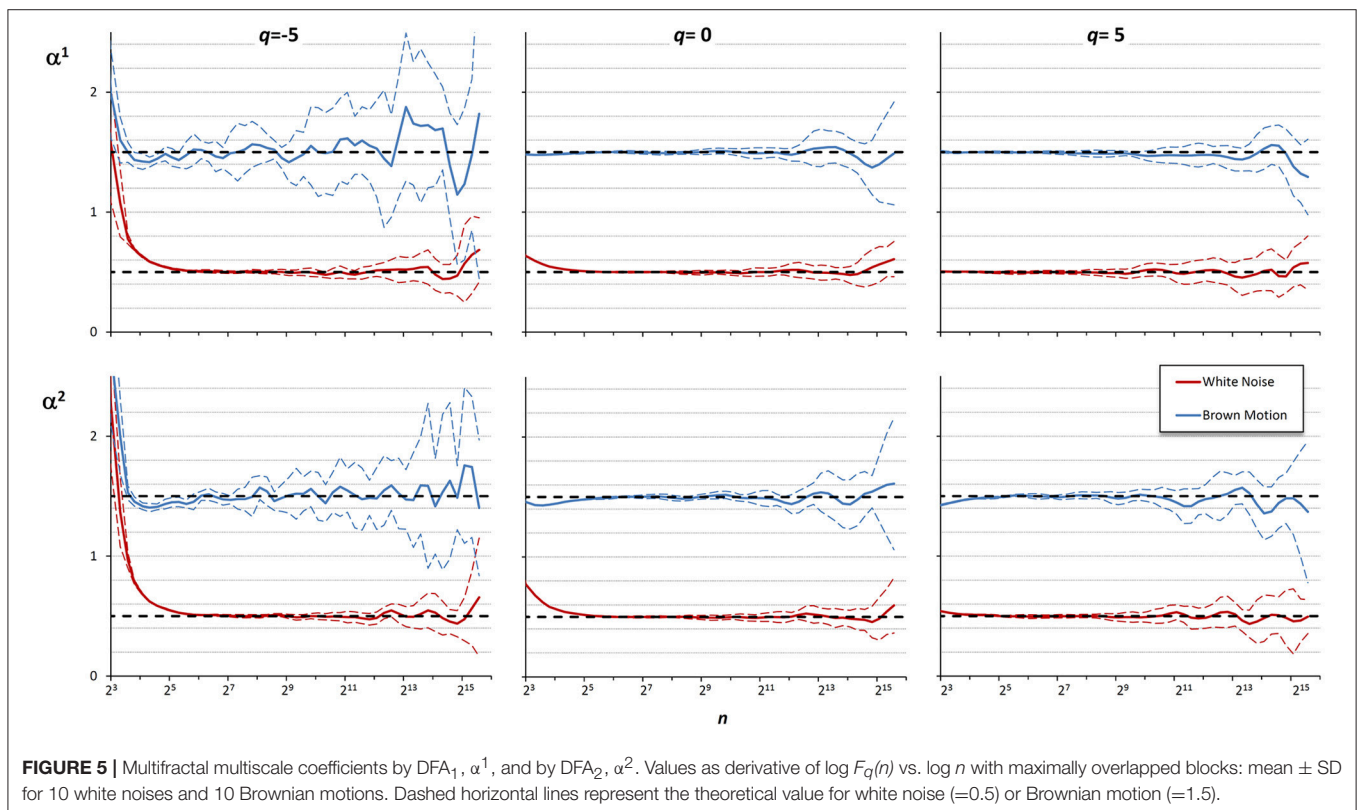
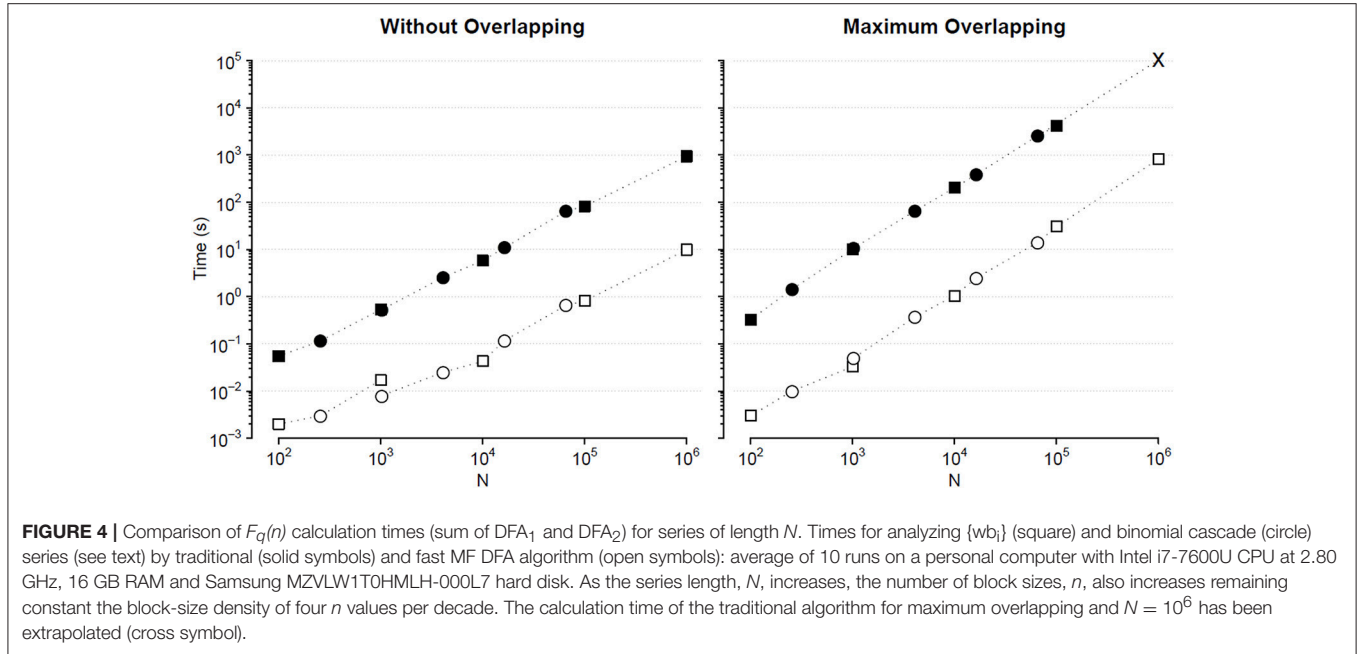
size n_b as:

$$\varepsilon(q, n_b) = |F_q(n_b)^F - F_q(n_b)^R| / F_q(n_b)^R \quad (21)$$

The largest of these errors among all the n_b block sizes, among all the q values and among the three series ($\{wb_i\}$, $\{wn_i\}$ and $\{Bm_i\}$) was taken as a global measure of the precision of the

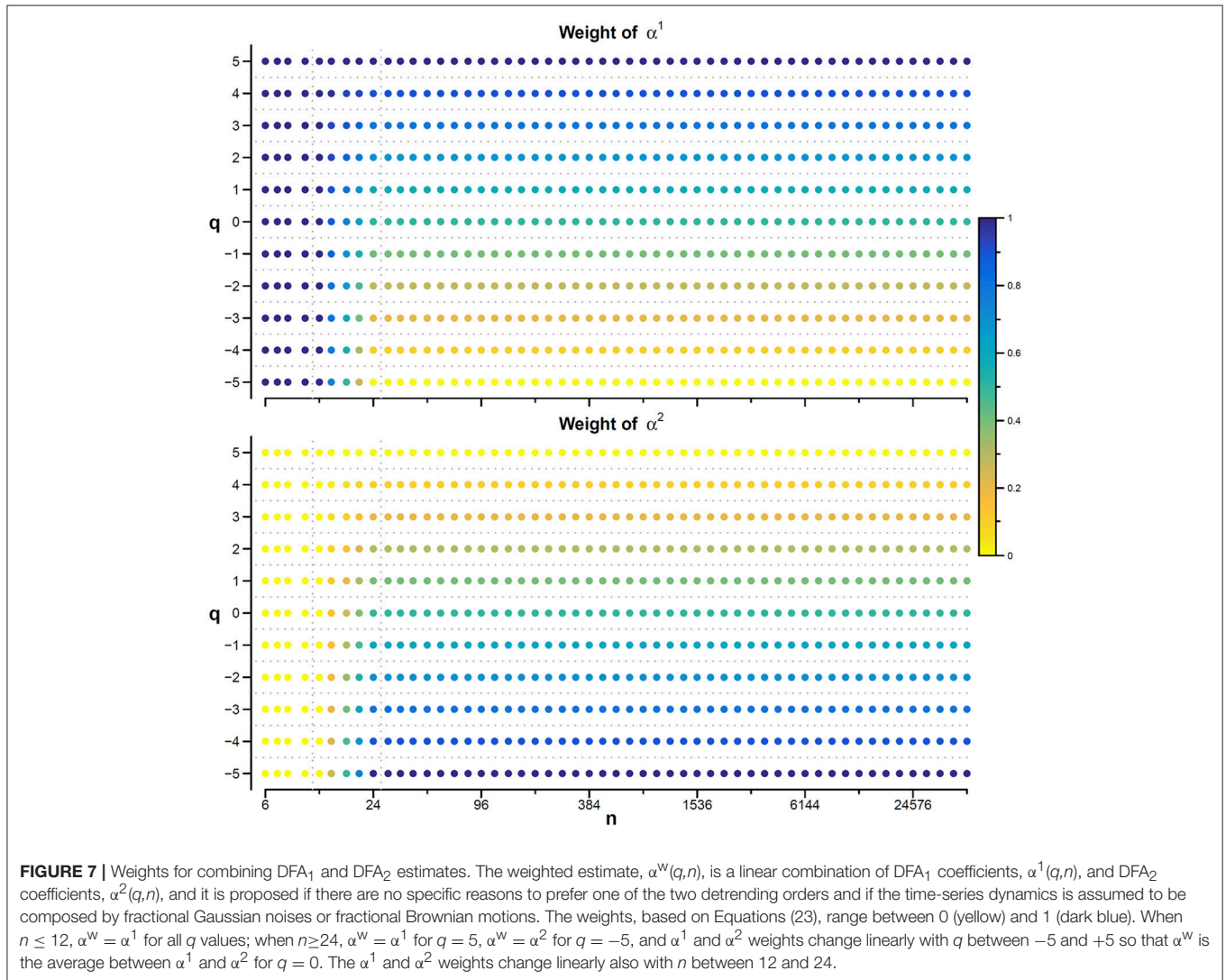
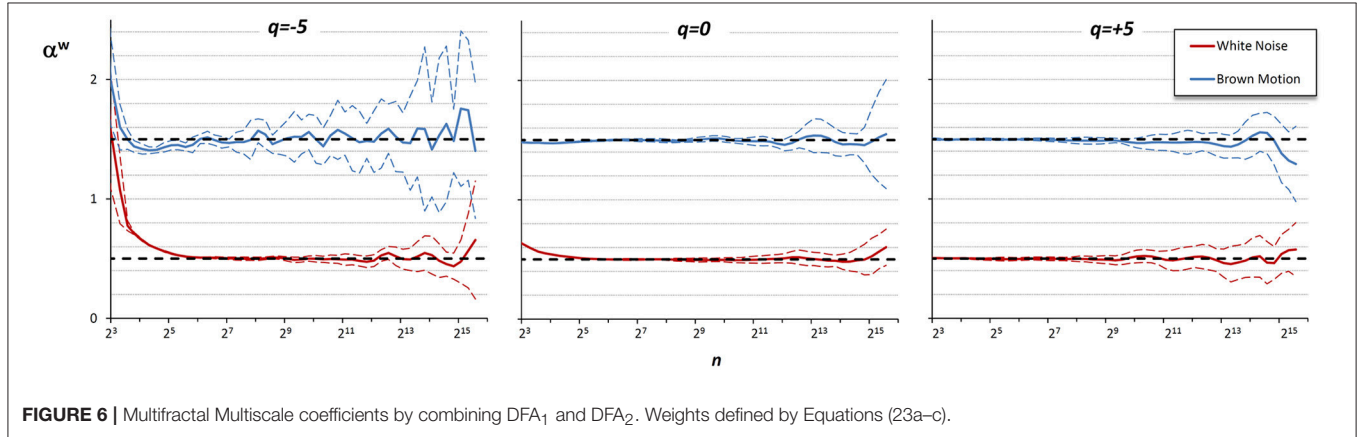
fast algorithm for any given length N_k . Relative errors are reported in **Table 1**: even in the worst case (corresponding to the length $N = 10^6$ samples), the highest relative error is less than 1%.

We also generated a stochastic binomial cascade embedded in noise, $\{bcn_i\}$, with a procedure similar to that described in Gieraltowski et al. (2012). We started from a series of $N = 2^{14}$



samples equal to 1, we split it into two segments of the same length, and multiplied one segment by the weight 0.25 and one by the weight (1-0.25), randomly selecting the weights of the first and second segment. We repeated the splitting/weighting procedure for each segment of the previous step, up to 14 steps.

To include random noise, we substituted all values lower than 10^{-6} with random samples uniformly distributed between 0 and 0.01. From this cascade, we subtracted its reversed duplicate to obtain a symmetric distribution. **Figure 3** shows examples of $F(n)$ estimates for the $\{bcn_i\}$ series (we analyzed heart-rate



series of similar length in the paragraph “*Application on real biomedical time series*”). We generated and concatenated four of these binomial cascades to obtain a series of length $N = 2^{16}$. Relative errors were assessed from fast, $F_q(n_b)^F$, and reference, $F_q(n_b)^R$, estimates (calculated with maximally overlapped blocks, thresholds as in Equation 19 and $EPS = 0$) for segments of length $N_k = 2^{2k+6}$ with $1 \leq k \leq 5$, and block sizes $n_b = 2^{2b+1}$ with $1 \leq b \leq k$ (e.g., for the length $N_4 = 16,384$ we considered the scales $n_1 = 16$, $n_2 = 64$, $n_3 = 256$ and $n_4 = 1,024$). Results in **Table 1** show negligible errors also for this multifractal series.

We compared the calculation times employed by the traditional and the fast MF DFA algorithm to analyze $\{wb_i\}$ series

of length $N_k = 10^{k+1}$ and binomial cascade series of length $N_k = 2^{2k+6}$, with $1 \leq k \leq 5$. Results in **Figure 4** indicate that the fast algorithm is about two orders of magnitude faster than the traditional one.

LOCAL SLOPES OF $F_q(n)$

Once that $F_q(n)$ has been estimated with maximally overlapped blocks, the multifractal DFA coefficients $\alpha(q, n)$ may be obtained as the first derivative of $\log F_q(n)$ vs. $\log n$. To have equispaced $\log n$ values we interpolate the $\log F_q(n)$ samples by a spline function, obtaining H values $F_q(n_h)$

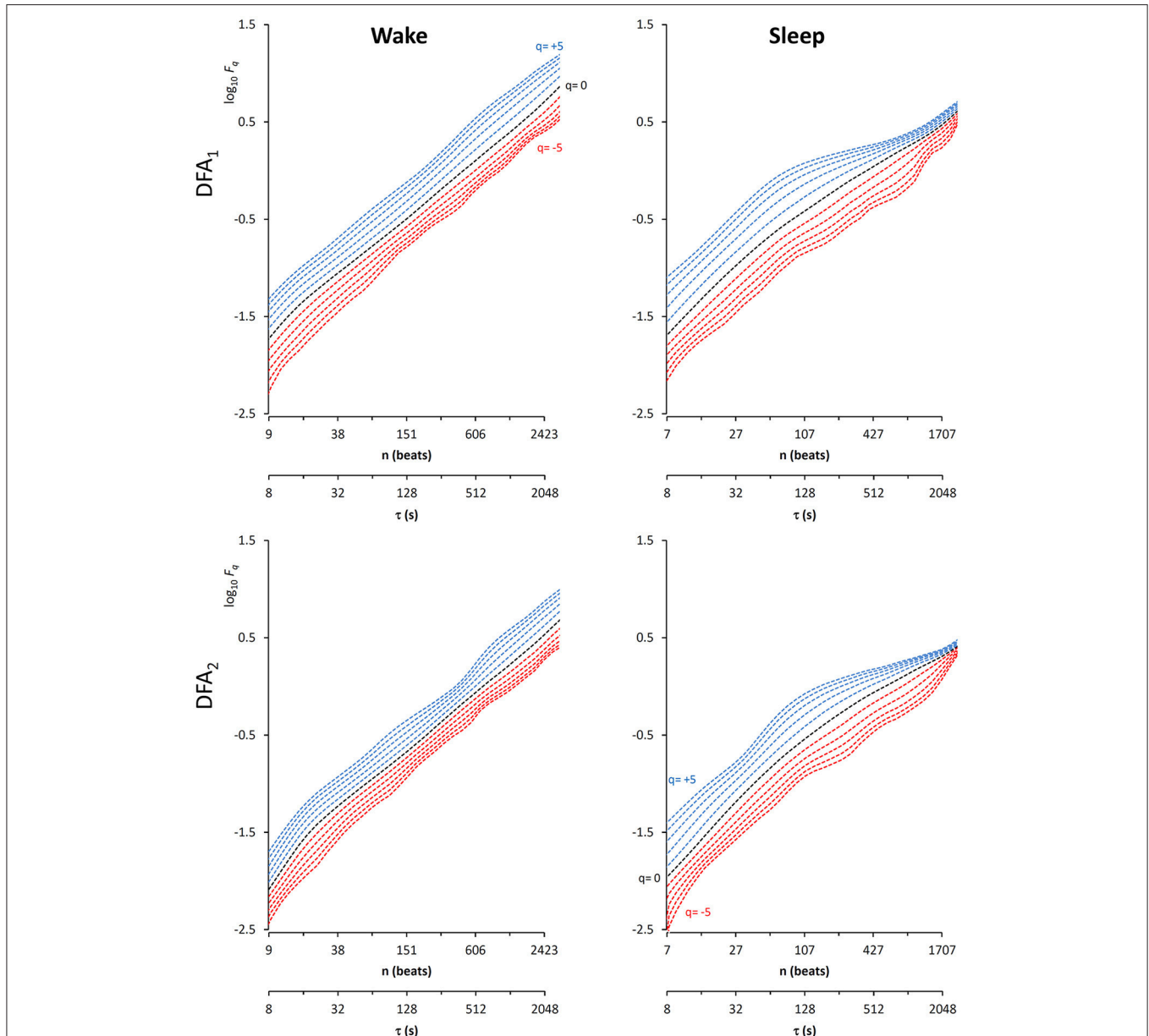


FIGURE 8 | Multifractal fluctuation functions of heart-rate variability during Wake and Sleep with maximally overlapped blocks. The original series are two 4-h segments of beat-by-beat R-R intervals from a 24-h ECG recording in a healthy volunteer. The first horizontal axis is the box size n , in number of beats; the second horizontal axis is the temporal scale τ , in seconds, obtained multiplying n by the mean R-R interval, which is shorter in *Wake* than in *Sleep*. Thresholds Th_1 and Th_2 as in Equations (19) and $EPS = 0$.

with n_h real numbers spaced exponentially on the scale axis. The slope at n_h is the derivative formula on 5 points (Castiglioni et al., 2018):

$$\alpha(q, n_h) = \frac{8(\log F_q(n_{h+1}) - \log F_q(n_{h-1})) - (\log F_q(n_{h+2}) - \log F_q(n_{h-2}))}{3(\log(n_{h+2}) - \log(n_{h-2}))}$$

for $2 < h < H - 1$ (22a)

or on 3 points at the extremes of the scale axis:

$$\alpha(q, n_h) = \frac{-\log F_q(n_{h+2}) + 4\log F_q(n_{h+1}) - 3\log F_q(n_h)}{\log(n_{h+2}) - \log(n_h)} \text{ for } h = 1 \quad (22b)$$

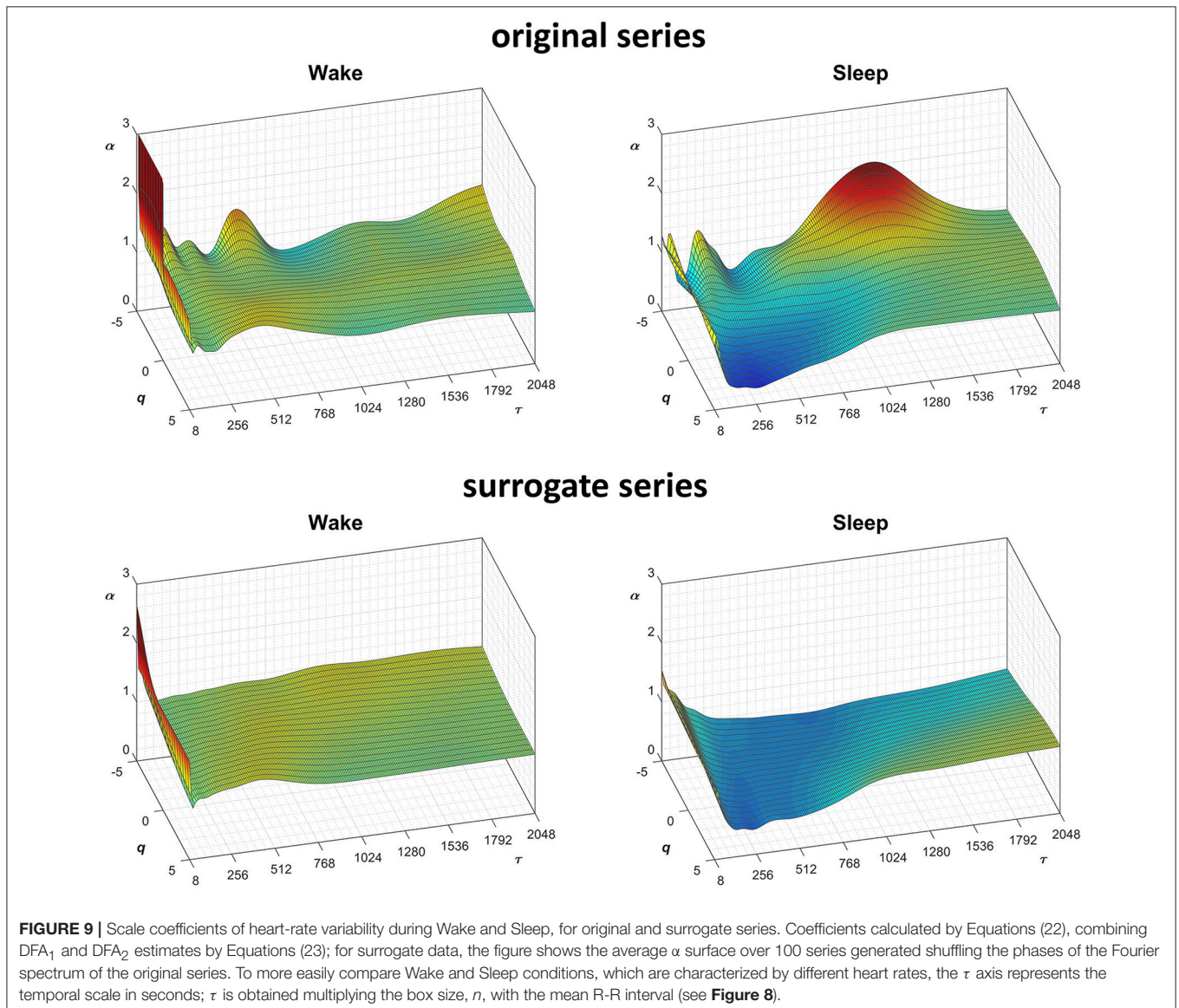
$$\alpha(q, n_h) = \frac{\log F_q(n_{h+1}) - \log F_q(n_{h-1})}{\log(n_{h+1}) - \log(n_{h-1})} \text{ for } h = 2 \text{ and } h = H - 1 \quad (22c)$$

$$\alpha(q, n_h) = \frac{\log F_q(n_{h-2}) - 4\log F_q(n_{h-1}) + 3\log F_q(n_h)}{\log(n_h) - \log(n_{h-2})} \text{ for } h = H \quad (22d)$$

The code for calculating Equations (22) is available as the **Supplemental Matlab File** *slpMFMSDFA.m*.

COMBINING INFORMATION FROM DFA₁ AND DFA₂

Our algorithm provides both DFA₁ and DFA₂ estimates in a single run. In some cases, there are reasons to prefer one of the two estimates. For instance, DFA₁ may identify the scale where possible crossovers occur better than DFA₂ (Höll and Kantz, 2015; Kiyono, 2015); however, if the power spectrum of the time series depends on the frequency f proportionally to $1/f^\beta$, then DFA₂ estimates the correct α up to $\beta < 5$, while DFA₁ can estimate it only up to $\beta < 3$ (Kiyono, 2015). Furthermore, it is possible that the nature of the physiological system under analysis may make preferable one of the two detrending orders. For instance, in the field of heart-rate



variability analysis, some studies highlighted that the slope α at $q = 2$ reflects the sympatho/vagal balance over scales shorter than 11 beats and allows stratifying the cardiovascular risk (Perkiomaki, 2011). Since high-order detrending polynomials overfit the shortest blocks (Kantelhardt et al., 2001), most of the clinical studies estimated this index by DFA₁, and a first-order detrending should be used to get an autonomic index comparable with the medical literature. On the other hand, a crossover at the larger heart-rate scales was found for DFA₁ but not for DFA₂ during light sleep, suggesting that in this sleep phase DFA₂ can better remove long-term trends (Bunde et al., 2000). This would make DFA₂ preferable to DFA₁ when comparing heart rate variability by phases of sleep over the larger scales.

These examples suggest that the best detrending order may differ if the focus of the analysis is on the shorter rather than on the larger scales. For a multifractal series, it is even possible that the best detrending order may depend also on q because components with different amplitude may have different fractal structures. Therefore, it is conceivable that the assessment of $\alpha(q, n)$ could be improved by properly combining the results of DFA₁ and DFA₂ at each scale. To investigate this possibility, we considered 10 white-noise and 10 Brownian-motion series, each of $N = 200,000$ samples. Often the fractal dynamics of physiological time series is described as belonging to the family of fractional Gaussian noises and fractional Brownian motions (Eke et al., 2000), and the behavior at the extremes of this family of random processes (i.e., white noise and Brownian motion) may be assumed to represent a large class of physiological series. Then we calculated $\alpha(q, n)$ by DFA₁ and by DFA₂ with $q = -5, q = 0$, and $q = +5$, for each series. **Figure 5** illustrates the deviations of the scale exponents from the theoretical value showing mean and

standard deviation of the estimates separately over the groups of white noise and Brownian motion. The figure shows that at the shorter scales the estimation bias is greater for DFA₂ while, at the larger scales, DFA₂ appears more stable than DFA₁ for $q = -5$, less stable than DFA₁ for $q = +5$ and of quality similar to DFA₁ for $q = 0$.

These findings suggest that if the time series is composed by fractional Gaussian noises and/or fractional Brownian motions, DFA₁ is preferable to DFA₂ at the shorter scales, while at the larger scales the choice may depend on the sign of q , and if $q = 0$ averaging the DFA₁ and DFA₂ coefficients might improve the estimate. Accordingly, we combined the estimates by DFA₁, $\alpha^1(q, n)$, and by DFA₂, $\alpha^2(q, n)$, defining a weighted average, $\alpha^w(q, n)$, when $q = -5$, as:

$$\begin{cases} \alpha^w(-5, n) = \alpha^1(-5, n) & \text{for } n < 12 \\ \alpha^w(-5, n) = \frac{24-n}{12}\alpha^1(-5, n) + \frac{n-12}{12}\alpha^2(-5, n) & \text{for } 12 \leq n \leq 24 \\ \alpha^w(-5, n) = \alpha^2(-5, n) & \text{for } n > 24 \end{cases} \quad (23a)$$

when $q = 0$, as:

$$\begin{cases} \alpha^w(0, n) = \alpha^1(0, n) & \text{for } n < 12 \\ \alpha^w(0, n) = \frac{36-n}{24}\alpha^1(0, n) + \frac{n-12}{24}\alpha^2(0, n) & \text{for } 12 \leq n \leq 24 \\ \alpha^w(0, n) = \frac{\alpha^1(0, n) + \alpha^2(0, n)}{2} & \text{for } n > 24 \end{cases} \quad (23b)$$

and when $q = +5$, as:

$$\alpha^w(5, n) = \alpha^1(5, n) \quad (23c)$$

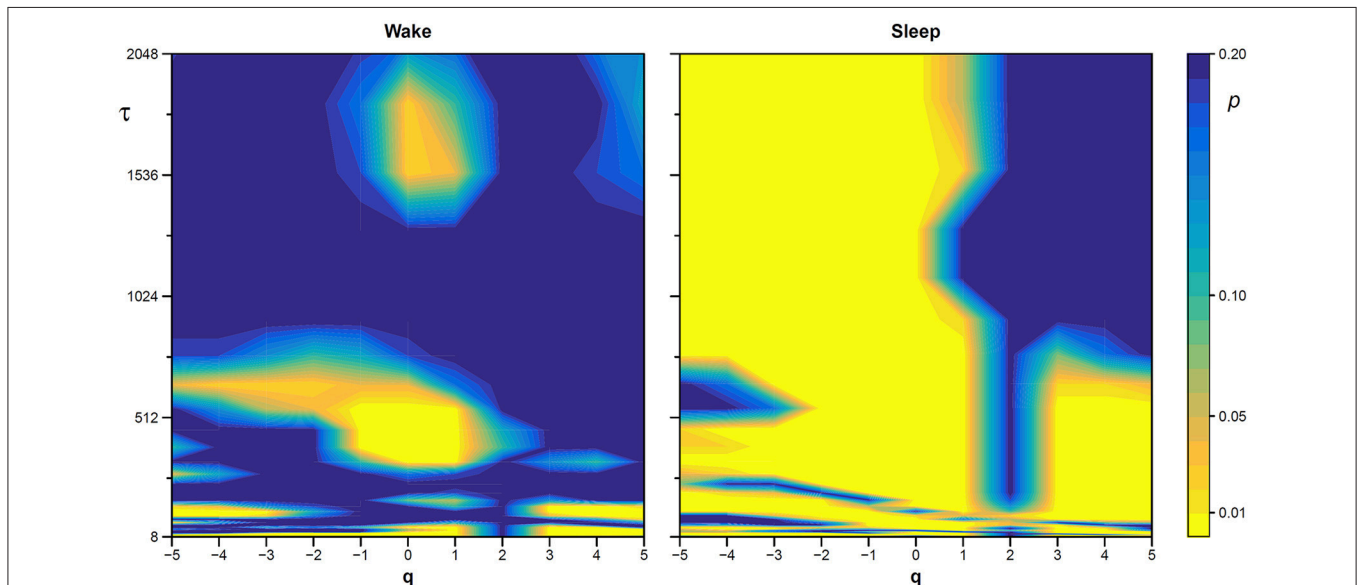


FIGURE 10 | Statistical significance of the nonlinearity test on scale coefficients of heart rate. Each point represents the two-sided, type-I error probability to reject the null hypothesis comparing scale coefficients of the original series and of Fourier-phase shuffled surrogate series at each scale τ and at each moment-order q ; differences significant at $p < 0.01$ are indicated in yellow. The temporal scale τ , in seconds, is obtained multiplying the box size, n , by the mean R-R interval (see **Figure 8**).

Figure 6 shows the α^w coefficients derived from α^1 and α^2 of Figure 5.

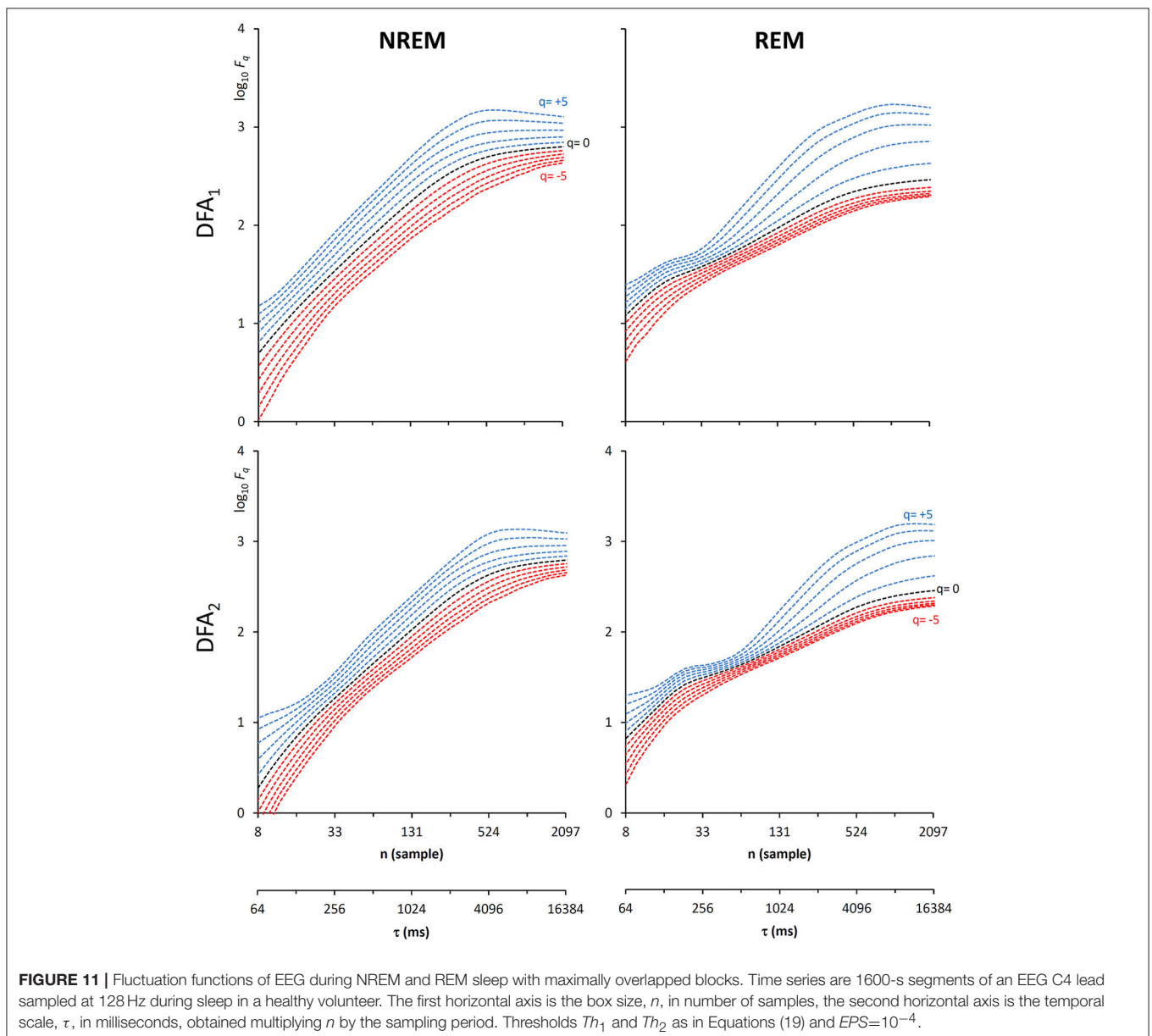
Values of α^w for q between -5 and $+5$ are calculated linearly interpolating the weights for $q = -5$ and $q = +5$. Figure 7 illustrates how the weights for the mixed coefficients are defined.

APPLICATIONS ON REAL BIOMEDICAL TIME SERIES

This section presents examples of applications of our algorithm on physiological time series collected in volunteers that gave their written informed consent to participate to the study, which was approved by the Ethics Committee of Istituto Auxologico Italiano, IRCCS, Milan, Italy.

Heart Rate

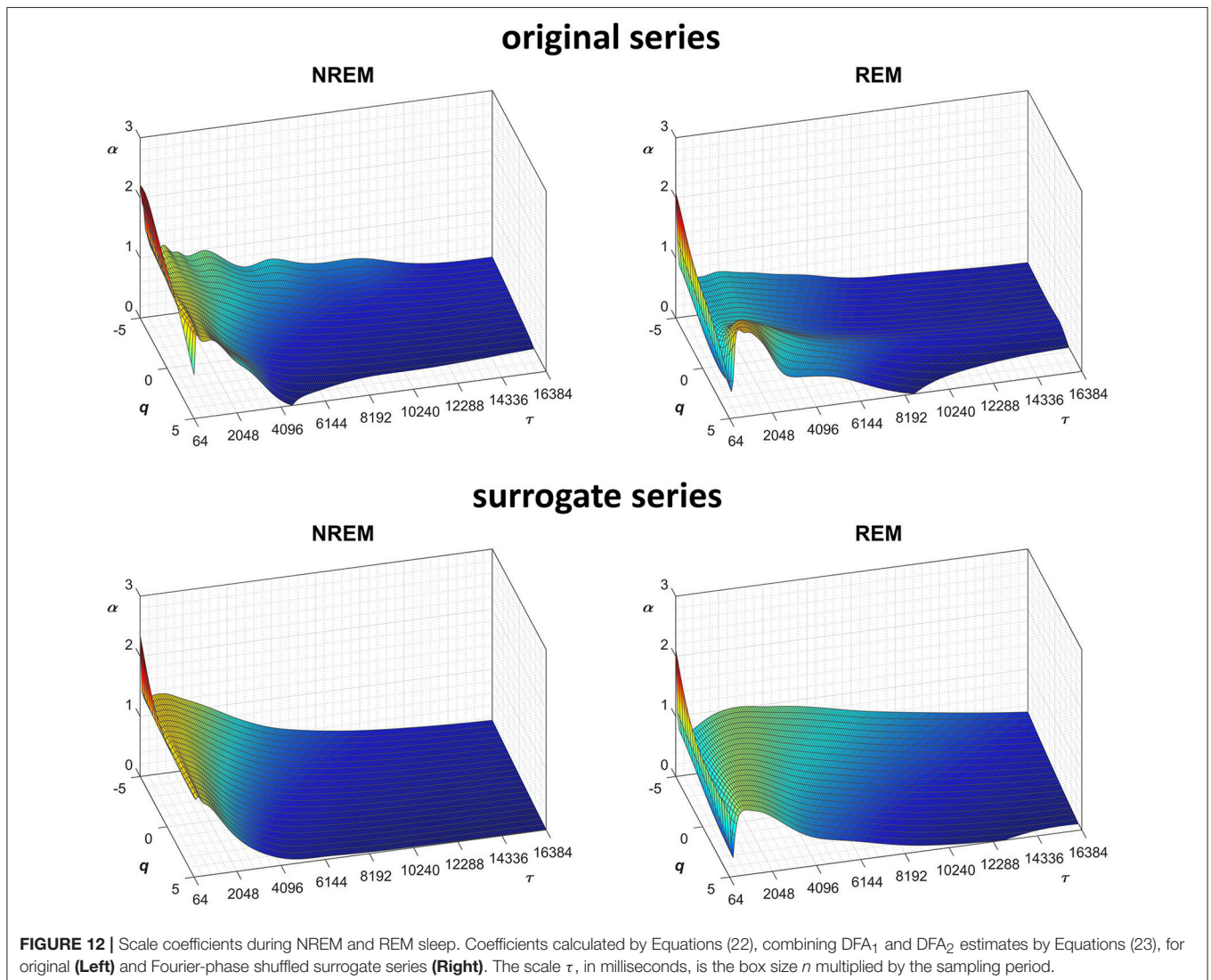
Most of DFA applications in physiology regard the analysis of heart rate variability. In this field, the evaluation of the self-similarity exponents as a continuous function of the scale, $\alpha(n)$, proved useful for associating a short-term crossover to the dynamics of removal of noradrenaline released by the sympathetic nerve endings (Castiglioni, 2011), for quantifying alterations during sleep at high-altitude (Castiglioni et al., 2011b) and for evaluating clinical conditions like congestive heart failure (Bojorges-Valdez et al., 2007) or spinal lesions (Castiglioni and Merati, 2017). When the self-similarity exponents were estimated as a multifractal multiscale surface of the moment q and scale n , $\alpha(q, n)$, they provided information on the autonomic development from fetal heart rate recordings (Gieraltowski et al., 2013), on differences between the dynamics of heart



rate and other cardiovascular variables (Castiglioni et al., 2018) and helped modeling the heart rate dynamics during sleep (Solinski et al., 2016).

In this context, our fast algorithm is expected to improve existing methods by calculating $\alpha(q, n)$ surfaces with higher scale resolution and lower estimation variability, thanks to its ability to provide estimates with maximally overlapped blocks in short time. As an example in the field of heart rate variability, we compare multiscale multifractality during nighttime sleep and during daytime activities in a healthy male volunteer. We derived R-R interval series from a 24-h electrocardiogram sampled at 250 Hz. We compared $F_q(n)$ functions calculated on two segments of 4 h duration, one selected during daytime, from 2 p.m. to 6 p.m. (Wake condition), one during nighttime after sleep onset, from 2 a.m. to 6 a.m. (Sleep condition). The mean heart rate was 50 bpm during Sleep ($N = 12,000$ beats) and 71 bpm during Wake ($N = 17,000$ beats). Four premature beats were identified visually and removed, and $F_q(n)$ functions were derived by DFA₁ and DFA₂ with maximally overlapped

blocks. Since the mean heart rate is lower in Sleep than in Wake condition, the same block of n beats corresponds to different time scales, in seconds, during Sleep and during Wake. Therefore, the fluctuation functions were plotted vs. the corresponding temporal scale τ , in seconds, by multiplying n by the mean R-R interval. **Figure 8** shows that Wake and Sleep have different fluctuations functions, increasing with τ as parallel lines in Wake and converging to a focus at the longest scales in Sleep. This suggests a multifractal behavior (i.e., different slopes for different q), more evident at some scales (i.e., a multiscale dynamics), during sleep. The $F_q(\tau)$ functions by DFA₁ and by DFA₂ appear similar: this confirms a previous observation on 24-h heart rate recordings (Gieraltowski et al., 2012) and does not suggest the presence of long-term trends that make one of the two detrending orders preferable. Therefore, we derived the surface of multifractal multiscale coefficients by applying the weighted approach of Equations (23). **Figure 9** illustrates the resulting $\alpha(q, \tau)$ surfaces. During daytime, the surface represents a relatively monofractal process



with multiscale dynamics, i.e., α depends more on τ than on q . By contrast, during sleep, the surface also shows a strong dependence on q at some scales, like at $\tau = 1250$ s, where it ranges between $\alpha = 2.05$ at $q = -5$ and $\alpha = 0.9$ at $q = +5$.

Fractal structures in physiological systems may arise from nonlinear chaotic dynamics or may be due to linear spectral features. We tested whether the $\alpha(q, n)$ surfaces we estimated for the heart-rate series actually reflect an underlying nonlinear dynamics with the method of surrogate data analysis (Theiler et al., 1992). For this purpose, we generated a dataset of 100 Fourier-phase shuffled time series (Schreiber and Schmitz, 2000) from each of the two recordings. The surrogate series have the same power spectrum of the original recording but a random phase. We calculated $\alpha(q, n)$ for the “Wake” and “Sleep” surrogate datasets. The average of each dataset is plotted in **Figure 9** for comparison with the original data. Multifractal structures evident in the original series are absent in the surrogate data.

The statistical significance of the difference between the original and surrogate series was assessed evaluating in which percentile of the surrogate dataset the $\alpha(q, \tau)$ coefficients of the original series fall. **Figure 10** plots the significance p at each scale and moment order. The figure demonstrates the presence of a multifractal dynamics that does not depend on the linear characteristics of the power spectrum, and which characterizes Sleep more than Wake. The whole surrogate analysis of the 200 phase-shuffled series took less than 10 min with our algorithm, while the traditional algorithm would take more than 18 h according to **Figure 4**.

Electroencephalogram

Another important field of DFA applications in physiology regards the study of EEG recordings (Hardstone et al., 2012). Traditionally, EEG is described by two (Ferree and Hwa, 2003;

Abasolo et al., 2008) or three DFA coefficients (Jospin et al., 2007), but $\alpha(n)$ represented as a continuous function of the scale n suggests more complex structures (Castiglioni, 2011). Therefore, we considered an overnight EEG recording (C4 lead) during sleep in a healthy adult male as an additional example. The EEG signal was band-pass filtered (cut-off frequencies between 0.3 and 35 Hz) before sampling at 128 Hz, and two segments of 1,600 s duration ($N = 204,800$ samples) were selected, one during NREM and one during REM sleep. Due to the 128 Hz sampling rate, the scale n , expressed in number of samples, corresponds to the temporal scale $\tau = n \times (1000/128)$ milliseconds. The fluctuation functions (**Figure 11**) have an almost flat profile at scales larger than 4 s, likely an effect of the filter. Furthermore, they show remarkable differences between sleep phases at shorter scales, running in parallel during NREM sleep and following a sigmoidal pattern in REM sleep. Since DFA_1 and DFA_2 estimates are similar, excluding the shorter scales where DFA_1 provides a more linear increase of $\log F_q(\tau)$ with $\log \tau$, we estimated the scale coefficients with the weighted approach of Equations (23). The $\alpha(q, \tau)$ surfaces (**Figure 12**) show a more evident multifractal structure in REM than in NREM sleep.

We also generated a surrogate dataset of 100 Fourier-phase shuffled time series for each recording, one during NREM and one during REM sleep. The average $\alpha(q, \tau)$ surface of each surrogate dataset is plotted in **Figure 12** for comparison with the original series. **Figure 13**, shows the significance of the difference between original and phase-shuffled series and indicates that the null hypothesis is rejected and nonlinearity detected almost at all scales and moment orders. The whole analysis on 200 surrogate series took 3 h and 22 min: this calculation time should be compared with the calculation time of 27 days and 8 h that the traditional algorithm would take, as from **Figure 4**.

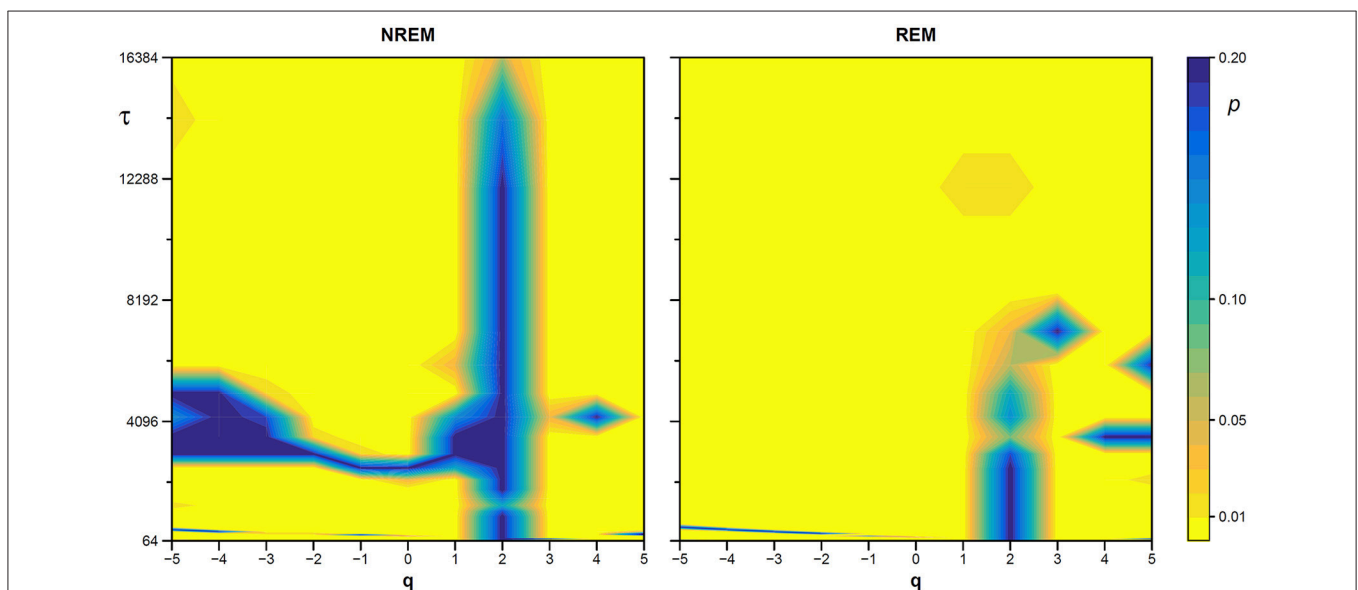


FIGURE 13 | Statistical significance of nonlinearity test on EEG scale coefficients. Each point represents the two-sided, type-I error probability to reject the null hypothesis comparing scale coefficients of the original series and of Fourier-phase shuffled surrogate series at each scale τ and moment-order q ; differences significant at $p < 0.01$ are indicated in yellow. The scale, τ , in milliseconds, is the box size n multiplied by the sampling period.

SUMMARY OF ADVANTAGES AND LIMITATIONS, AND CONCLUSIONS

Since the DFA introduction, hundredths of works in physiological or clinical settings quantified the fractal dynamics by α . Almost all of them provided a monofractal description, most considering only one or two ranges of scales (Perkiomaki, 2011; Sassi et al., 2015). Even if the results were promising, particularly in the field of heart rate variability, it is questioned whether a monofractal DFA estimating one or two coefficients may improve health assessment and risk stratification compared to nonfractal methods (Sassi et al., 2015). Therefore, the more advanced research in this field is investigating how DFA can be enhanced by jointly depicting the multifractal and multiscale nature of the time series in order to model pathophysiological mechanisms more completely (Zebrowski et al., 2015; Solinski et al., 2016).

In this regard, the proposed algorithm is aimed at improving scale resolution and at reducing estimation variability of the multifractal fluctuations function without the cost of an increase in calculation time that may make unfeasible the analysis of relatively long series. Therefore, we redesigned the multifractal DFA algorithm to provide estimates in a relatively short time even by using maximally overlapped blocks, with first and second-order detrending polynomials simultaneously. This allows comparing DFA₁ and DFA₂ estimates easily, which may be useful for evaluating the consistency of the estimates and for selecting the best detrending order. Simultaneously obtaining the DFA₁ and DFA₂ scale coefficients, which have a certain degree of statistical independence, also suggests new ways to improve the estimate, as in the weighted approach we proposed. The examples on real biomedical series point out that the algorithm is sufficiently fast to allow using challenging bootstrapping procedures on surrogate data that may separate linear and nonlinear components of the multiscale/multifractal dynamics. The possibility to apply such tests in clinical studies may make

it easier deriving richer information to improve diagnosis or risk stratification from DFA.

Some limitations, however, should be considered. The algorithm employs detrending polynomials of first and second order only. Although these orders are used in most (if not all) DFA applications on real biosignal, the algorithm cannot be considered as general as other DFA codes. If a higher order “o” is required, the analytic expression of the polynomial regression should be derived as done in Equations (7–10) for DFA₁ and DFA₂, powers summations should be calculated by extending the array definition of Equation (14) and, to maintain calculation errors negligible also for DFA₀, a new threshold Th_o should be defined in Equations (19), which will decrease the speed of the algorithm.

AUTHOR CONTRIBUTIONS

PC conceived the study and wrote the manuscript. AF wrote the code and created the figures. PC and AF developed the theoretical formulas on which the algorithm is based, designed the algorithm validation with synthesized series, made the acquisition and the analysis of real time series, interpreted the results and critically revised the manuscript.

ACKNOWLEDGMENTS

Matlab source code for generating surrogate data was provided by Temu Gautama in Matlab central, downloadable from <https://it.mathworks.com/matlabcentral/fileexchange/4612-surrogate-data>.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2019.00115/full#supplementary-material>

REFERENCES

- Abasolo, D., Hornero, R., Escudero, J., and Espino, P. (2008). A study on the possible usefulness of detrended fluctuation analysis of the electroencephalogram background activity in Alzheimer's disease. *IEEE Trans. Biomed. Eng.* 55, 2171–2179. doi: 10.1109/TBME.2008.923145
- Bojorges-Valdez, E. R., Echeverria, J. C., Valdes-Cristerna, R., and Pena, M. A. (2007). Scaling patterns of heart rate variability data. *Physiol. Meas.* 28, 721–730. doi: 10.1088/0967-3334/28/6/010
- Bunde, A., Havlin, S., Kantelhardt, J. W., Penzel, T., Peter, J.-H., and Voigt, K. (2000). Correlated and Uncorrelated Regions in Heart-Rate Fluctuations during Sleep. *Phys. Rev. Lett.* 85, 3736–3739. doi: 10.1103/PhysRevLett.85.3736
- Castiglioni, P. (2011). “Self-similarity in physiological time series: new perspectives from the temporal spectrum of scale exponents,” in *Computational Intelligence Methods for Bioinformatics and Biostatistics: 8th International Meeting*, eds E. Biganzoli, A. Vellido, F. Ambrogi, R. Tagliaferri CIBB (Gargnano del Garda; Berlin, Heidelberg: Springer), 164–175.
- Castiglioni, P., Di Rienzo, M., and Faini, A. (2017). “Self-similarity and detrended fluctuation analysis of cardiovascular signals,” in *Complexity and Nonlinearity in Cardiovascular Signals*, eds R. Barbieri, E. P. Scilingo, and G. Valenza (Cham: Springer International Publishing), 197–232. doi: 10.1007/978-3-319-58709-7_7
- Castiglioni, P., Lazzeroni, D., Coruzzi, P., and Faini, A. (2018). Multifractal-multiscale analysis of cardiovascular signals: a DFA-based characterization of blood pressure and heart-rate complexity by gender. *Complexity* 2018, 1–14. doi: 10.1155/2018/4801924
- Castiglioni, P., and Merati, G. (2017). Fractal analysis of heart rate variability reveals alterations of the integrative autonomic control of circulation in paraplegic individuals. *Physiol. Measur.* 38, 774–786. doi: 10.1088/1361-6579/aa5b7e
- Castiglioni, P., Parati, G., Civijian, A., Quintin, L., and Di Rienzo, M. (2009). Local scale exponents of blood pressure and heart rate variability by detrended fluctuation analysis: effects of posture, exercise, and aging. *IEEE Trans. Biomed. Eng.* 56, 675–684. doi: 10.1109/TBME.2008.2005949
- Castiglioni, P., Parati, G., Di Rienzo, M., Caraballona, R., Cividjian, A., and Quintin, L. (2011a). Scale exponents of blood pressure and heart rate during autonomic blockade as assessed by detrended fluctuation analysis. *J. Physiol.* 589, 355–369. doi: 10.1113/jphysiol.2010.196428
- Castiglioni, P., Parati, G., Lombardi, C., Quintin, L., and Di Rienzo, M. (2011b). Assessing the fractal structure of heart rate by the temporal spectrum of scale

- exponents: a new approach for detrended fluctuation analysis of heart rate variability. *Biomed. Tech.* 56, 175–183. doi: 10.1515/bmt.2011.010
- Echeverria, J. C., Woolfson, M. S., Crowe, J. A., Hayes-Gill, B. R., Croaker, G. D., and Vyas, H. (2003). Interpretation of heart rate variability via detrended fluctuation analysis and alpha-beta filter. *Chaos*. 13, 467–475. doi: 10.1063/1.1562051
- Eke, A., Herman, P., Bassingthwaite, J. B., Raymond, G. M., Percival, D. B., Cannon, M., et al. (2000). Physiological time series: distinguishing fractal noises from motions. *Pflug. Arch.* 439, 403–415. doi: 10.1007/s004249900135
- Ferree, T. C., and Hwa, R. C. (2003). Power-law scaling in human EEG: relation to Fourier power spectrum. *Neurocomputing* 52–54, 755–761. doi: 10.1016/S0925-2312(02)00760-9
- Ge, E., and Leung, Y. (2013). Detection of crossover time scales in multifractal detrended fluctuation analysis. *J. Geograph. Syst.* 15, 115–147. doi: 10.1007/s10109-012-0169-9
- Gieraltowski, J., Hoyer, D., Tetschke, F., Nowack, S., Schneider, U., and Zebrowski, J. (2013). Development of multiscale complexity and multifractality of fetal heart rate variability. *Auton. Neurosci.* 178, 29–36. doi: 10.1016/j.autneu.2013.01.009
- Gieraltowski, J., Zebrowski, J. J., and Baranowski, R. (2012). Multiscale multifractal analysis of heart rate variability recordings with a large number of occurrences of arrhythmia. *Phys. Rev. E Stat. Nonlin. Soft. Matter Phys.* 85:021915. doi: 10.1103/PhysRevE.85.021915
- Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101, E215–E220. doi: 10.1161/01.CIR.101.23.e215
- Hardstone, R., Poil, S.-S., Schiavone, G., Jansen, R., Nikulin, V. V., Mansvelder, H. D., et al. (2012). Detrended fluctuation analysis: a scale-free view on neuronal oscillations. *Front. Physiol.* 3:450. doi: 10.3389/fphys.2012.00450
- Höll, M., and Kantz, H. (2015). The relationship between the detrended fluctuation analysis and the autocorrelation function of a signal. *Eur. Phys. J. B.* 88:327. doi: 10.1140/epjb/e2015-60721-1
- Hwa, R. C., and Ferree, T. C. (2002). Scaling properties of fluctuations in the human electroencephalogram. *Phys. Rev. E Stat. Nonlin. Soft. Matter Phys.* 66:021901. doi: 10.1103/PhysRevE.66.021901
- Ihlen, E. A. F. (2012). Introduction to multifractal wavelet and detrended fluctuation analyses. *Front. Physiol.* 3:141. doi: 10.3389/fphys.2012.00141
- Jospin, M., Caminal, P., Jensen, E. W., Litvan, H., Vallverdu, M., Struys, M. M., et al. (2007). Detrended fluctuation analysis of EEG as a measure of depth of anesthesia. *IEEE Trans. Biomed. Eng.* 54, 840–846. doi: 10.1109/TBME.2007.893453
- Kantelhardt, J. W., Koscielny-Bunde, E., Rego, H. H. A., Havlin, S., and Bunde, A. (2001). Detecting long-range correlations with detrended fluctuation analysis. *Phys. A* 295, 441–454. doi: 10.1016/S0378-4371(01)00144-3
- Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., and Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Phys. A* 316, 87–114. doi: 10.1016/S0378-4371(02)01383-3
- Kiyono, K. (2015). Establishing a direct connection between detrended fluctuation analysis and Fourier analysis. *Phys. Rev. E* 92:042925. doi: 10.1103/PhysRevE.92.042925
- Ludescher, J., Bogachev, M. I., Kantelhardt, J. W., Schumann, A. Y., and Bunde, A. (2011). On spurious and corrupted multifractality: the effects of additive noise, short-term memory and periodic trends. *Phys. A* 390, 2480–2490. doi: 10.1016/j.physa.2011.03.008
- Mukli, P., Nagy, Z., and Eke, A. (2015). Multifractal formalism by enforcing the universal behavior of scaling functions. *Phys. A* 417, 150–167. doi: 10.1016/j.physa.2014.09.002
- Mukli, P., Nagy, Z., Racz, F. S., Herman, P., and Eke, A. (2018). Impact of healthy aging on multifractal hemodynamic fluctuations in the human prefrontal cortex. *Front. Physiol.* 9:1072. doi: 10.3389/fphys.2018.01072
- Nagy, Z., Mukli, P., Herman, P., and Eke, A. (2017). Decomposing multifractal crossovers. *Front. Physiol.* 8:533. doi: 10.3389/fphys.2017.00533
- Peng, C. K., Havlin, S., Stanley, H. E., and Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* 5, 82–87. doi: 10.1063/1.166141
- Perkiomaki, J. S. (2011). Heart rate variability and non-linear dynamics in risk stratification. *Front. Physiol.* 2:81. doi: 10.3389/fphys.2011.00081
- Sassi, R., Cerutti, S., Lombardi, F., Malik, M., Huikuri, H. V., Peng, C. K., et al. (2015). Advances in heart rate variability signal analysis: joint position statement by the e-Cardiology ESC Working Group and the European Heart Rhythm Association co-endorsed by the Asia Pacific Heart Rhythm Society. *Europace* 17, 1341–1353. doi: 10.1093/europace/euv015
- Schreiber, T., and Schmitz, A. (2000). Surrogate time series. *Phys. D* 142, 346–382. doi: 10.1016/S0167-2789(00)00043-9
- Solinski, M., Gieraltowski, J., and Zebrowski, J. (2016). Modeling heart rate variability including the effect of sleep stages. *Chaos* 26:023101. doi: 10.1063/1.4940762
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Doynne Farmer, J. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Phys. D* 58, 77–94. doi: 10.1016/0167-2789(92)90102-S
- Weisstein, E. W. (1999). *CRC Concise Encyclopedia of Mathematics*. Boca Raton, Fla: CRC Press.
- Xia, J., Shang, P., and Wang, J. (2013). Estimation of local scale exponents for heartbeat time series based on DFA. *Nonlin. Dyn.* 74, 1183–1190. doi: 10.1007/s11071-013-1033-2
- Zebrowski, J. J., Kowalik, I., Orłowska-Baranowska, E., Andrzejewska, M., Baranowski, R., and Gieraltowski, J. (2015). On the risk of aortic valve replacement surgery assessed by heart rate variability parameters. *Physiol. Meas.* 36, 163–175. doi: 10.1088/0967-3334/36/1/163

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Castiglioni and Faini. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

Given the time series y_i with index i from 1 to N , let's shift i by the integer quantity J so that each sample y_i is associated with the new index $j = i - J$ running from $1 - J$ to $N - J$.

The average $\bar{i} = \frac{1}{n} \sum_{i=J+1}^{J+n} i$ can be rewritten as

$$\begin{aligned}\bar{i} &= \frac{1}{n} \sum_{j=1}^n (j + J) = \\ &= \frac{1}{n} \sum_{j=1}^n j + J = \\ &= \bar{j} + J\end{aligned}\quad (\text{A1})$$

Similarly, the summation $s_{iy} = \frac{1}{n} \sum_{i=J+1}^{J+n} iy_i$ in Equation (8c) can be rewritten as

$$\begin{aligned}s_{iy} &= \frac{1}{n} \sum_{j=1}^n (j + J) y_i = \\ &= \frac{1}{n} \sum_{j=1}^n jy_i + J \times \bar{y} = \\ &= s_{jy} + J \times \bar{y}\end{aligned}\quad (\text{A2})$$

and $s_{i^2y} = \frac{1}{n} \sum_{i=J+1}^{J+n} i^2 y_i$ in Equation (10d) as

$$\begin{aligned}s_{i^2y} &= \frac{1}{n} \sum_{j=1}^n (j + J)^2 y_i = \\ &= \frac{1}{n} \sum_{j=1}^n j^2 y_i + 2J \frac{1}{n} \sum_{j=1}^n jy_i + J^2 \frac{1}{n} \sum_{j=1}^n jy_i \\ &= s_{j^2y} + (2J + J^2) \times s_{jy}\end{aligned}\quad (\text{A3})$$

Through Equations (A1-3), s_{jy} and s_{j^2y} can be obtained from s_{iy} and s_{i^2y} as calculated in Equation (15).