Consistency of detrended fluctuation analysis

O. Løvsletten
Department of Mathematics and Statistics, UiT Artic University of Norway, Norway.

The scaling function \( F(s) \) in detrended fluctuation analysis (DFA) scales as \( F(s) \sim s^{2H} \) for stochastic processes with Hurst exponents \( H \). We prove this scaling law for both stationary stochastic processes with \( 0 < H < 1 \), and non-stationary stochastic processes with \( 1 < H < 2 \). For \( H < 0.5 \) we observe that using the asymptotic (power-law) auto-correlation function (ACF) yield \( F(s) \sim s^{1/2} \).

We also show that the fluctuation function in DFA is equal in expectation to: i) A weighted sum of the ACF ii) A weighted sum of the second order structure function. These results enable us to compute the exact finite-size bias for signals that are scaling, as well as studying DFA for signals that do not have power-law statistics. We illustrate this with examples, where we find that a previous suggested modified DFA will increase the bias for signals with Hurst exponents \( H > 1 \).

As a final application of the new theory, we present an estimator \( \hat{F}(s) \) that can handle missing data in regularly sampled time series without the need for interpolation schemes. Under mild regularity conditions, \( \hat{F}(s) \) is equally in expectation to the fluctuation function \( F(s) \) in the gap-free case.

I. INTRODUCTION

Detrended fluctuation analysis (DFA) was introduced in a study of long-range dependence in DNA sequences [1]. It has later been applied in a wide range of scientific disciplines [2]. Some recent examples are found in scientific studies of climate [3], finance [4] and medicine [5]. The most common usage of DFA is to estimate the Hurst exponent. The assumption is then that the input signal has second-moments that are scaling. If this assumption is satisfied the DFA fluctuation function \( F(s) \) scales as a power-law

\[
\mathbb{E} F^2(s) \sim s^{2H},
\]

where \( \mathbb{E} \) denotes the mathematical expectation operator.

Important examples of stochastic processes \( X(t) \) with scaling properties are self-similar and multifractal models, see e.g. [6]. For this large class of models the existing \( q \)-moments satisfy \( \mathbb{E} |X(t + t_0) - X(t_0)|^q \propto t^{q(2)} \). In particular, if the variance is finite, the second moments are scaling and we define the Hurst exponent \( H \) by the relation \( \zeta(2) = 2H - 2 \). The power-law of the DFA fluctuation function in this case \( (1 < H < 2) \) has been established empirically. A mathematical proof has not been published prior to this paper, except for random walks \( (H = 1.5) \) [7].

For stationary stochastic processes \( X(t) \) with second-moments that are scaling, the Hurst exponent is in the range \( 0 < H < 1 \). For \( H = 0.5 \), \( X(t) \) is white noise, while \( H \neq 1/2 \) implies an auto-correlation function (ACF) \( \rho(\tau) \) on the form

\[
\rho(\tau) \sim H(2H - 1)\tau^{2H-2}.
\]

For \( H < 1/2 \) the ACF is negative for all time lags \( \tau \neq 0 \), while for \( H > 1/2 \) the ACF is positive. Moreover, in the persistent case \( (H > 1/2) \), the ACF decays so slowly that the series \( \sum_{\tau=-\infty}^{\infty} \rho(\tau) \) diverges.

In the case of a stationary input signal \( X(t) \), with Hurst exponent \( 0 < H < 1 \), Eq. (1) has been partly proved. Taqqu et al. [9] constructed a proof for DFA1. DFA\( m \), or DFA of order \( m \), means that a \( m \)'th order polynomial is applied in the DFA algorithm (Section IIIA).

For Hurst exponents restricted to the range \( 0.5 < H < 1 \), the proof has been extended to include higher order polynomials \( m \geq 1 \) [8]. We make the new observation that, for \( 0 < H < 0.5 \), in order for Eq. (1) to be satisfied, it is the exact auto-covariance function (acvf) that must be used. If one instead apply the asymptotic acvf, then \( \mathbb{E} F^2(s) \sim s^{1/2} \).

For stationary signals [8] showed that the squared DFA fluctuation function is equal, in expectation, to a weighted sum of the acvf \( \gamma(\cdot) \):

\[
\mathbb{E} F^2(s) = \gamma(0)G(0,s)s^{-1} + 2s^{-1} \sum_{j=1}^{s-1} G(j,s)\gamma(j),
\]

where the weight function \( G(j,s) \) will be defined in Section III. We present the more general result

\[
\mathbb{E} F^2(s) = -\frac{1}{s} \sum_{j=1}^{s-1} G(j,s)S(j),
\]

where \( S(t) = \mathbb{E} [X(t + t_0) - X(t_0)]^2 \), which also holds for non-stationary stochastic processes with stationary increments. The quantity \( S(t) \) is known as the variogram. We note that the relationship between DFA and the power spectral density was derived, partly based on numerical calculations, in [10].

Eqs. (3) and (4) have applications beyond proving Eq. (1). For instance, one can compute the exact finite-size bias for signals that are scaling, as well as studying DFA for signals that are not scaling. In [2] the bias of DFA for stochastic processes with Hurst exponents in the range \( 0.5 < H < 1 \) was found by means of Monte Carlo.

An analytical study of the behaviour of DFA for auto-regressive processes of order one (AR(1)) was investigated in [8]. We demonstrate the use of Eqs. (3) and (4) by simple extensions of the aforementioned examples.

As a final application of the new theory presented in this paper, we propose estimators (modifications of the
DFA fluctuation function) that can handle missing data in regularly sampled time series. One simple way of handling missing data is to apply linear interpolation, random re-sampling or mean filling. However, this will typically destroy, or add artificial, correlations to the time series under study. The effect on DFA using these three gap-filling techniques was examined in [11] for signals with Hurst exponents 0 < H < 1. It was found that these interpolation schemes introduced significant deviation from the expected scaling. In contrast, the modified fluctuation functions we propose have the property of equality in expectation to the fluctuation function in the gap-free case. For the wavelet variance, estimators that can handle missing data in a proper statistical way was presented in [12]. These wavelet variances are similar in construction to the DFA estimators we present.

This paper is organised as follows. In Section II we review the definition of Hurst exponent adopted in this paper. Examples of stochastic processes with well-defined Hurst exponents are given. In Section III we present the relationship between DFA and acvf/variogram, and the proof of Eq. (1). Examples of applications are given in Section IV. Bias for scaling signals, DFA of Ornstein-Uhlenbeck processes and modification of the DFA fluctuation function to handle missing data.

II. HURST EXPONENT

A. Definition and properties

Let X(t) be a stochastic process with mean EX(t) = 0. If

i) X(t) is non-stationary with stationary increments and

\[ \mathbb{E}[X(t + t_0) - X(t_0)]^2 \propto t^{2H-2}, \]

or

ii) X(t) is stationary and

\[ \mathbb{E}[Y(t + t_0) - Y(t_0)]^2 \propto t^{2H}, \quad Y(t) = \sum_{k=1}^{t} X(k), \]

holds, then we define H to be the Hurst exponent of the process X(t). The Hurst exponent determines the correlation at all time scales. Assume that X(t) has Hurst exponent 1 < H < 2, i.e., X(t) is non-stationary. We have

\[ 2X(t)X(s) = X(t)^2 + X(s)^2 - \{X(t) - X(s)\}^2. \]

By stationary increments

\[ \mathbb{E}\{X(t) - X(s)\}^2 = \mathbb{E}\{|t - s|^2\}. \]

It follows that

\[ \mathbb{E}X(t)X(s) = \frac{\sigma^2}{2} \{ |s|^{2h} + |t|^{2h} - |t - s|^{2h} \}, \quad (5) \]

with \( \mathbb{E}X(1)^2 = \sigma^2 \) and \( h = H - 1 \). The increments \( \Delta X(t) = X(t) - X(t - 1) \) have Hurst exponent h. The acvf \( \gamma(\tau) \) of the increments follows from (5), and is given by

\[ \gamma(\tau) = \frac{\sigma^2}{2} (|\tau + 1|^{2h} - 2|\tau|^{2h} + |\tau - 1|^{2h}). \quad (6) \]

For \( h = 1/2 \) the increments are white noise, while for \( h \neq 1/2 \) the acvf is asymptotically a power-law

\[ \gamma(\tau) \sim \frac{\sigma^2}{2} \frac{d^2}{d\tau^2} \tau^{2h} = \sigma^2 h(2h - 1)\tau^{2h-2}, \]

as \( \tau \to \infty \). Thus, \( h \neq 1/2 \) implies dependent increments. Choosing \( 0 < h < 1/2 \) results in negatively correlated increments, while for \( h > 1/2 \) the increments are persistent. Moreover, in the persistent case, the acvf decays so slowly that the series \( \sum_{\tau=-\infty}^{\infty} \gamma(\tau) \) diverges.

B. Examples

If, in addition to Hurst exponents in the range 1 < H < 2, we require that X(t) is Gaussian, then this defines the class of fractional Brownian motions (fBm’s) whose increments are known as fractional Gaussian noises (fGn’s) [13]. An fBm is an example of a self-similar process. By definition self-similar processes X(t), with self-similar exponent h, satisfy the self-similarity

\[ X(at) \overset{d}{=} M(a)X(t), \quad (7) \]

with \( M(a) = a^h \) [14], and where \( \overset{d}{=} \) denotes equality in finite-dimensional distributions. The class of log-ininitely divisible multifractal processes [15, 16] also satisfy Eq. (7), but now M(a) is random variable with an arbitrarily log-ininitely divisible distribution. The scaling law Eq. (7) implies \( \mathbb{E}[X(t + t_0) - X(t_0)]^q \propto t^{\zeta(q)} \). Thus, if the second moments exist the Hurst exponent H is given by the relation \( \zeta(2) = 2H - 2 \). These examples are summarised in Table I. We emphasise that neither multifractality or self-similarity is needed to have a process with well-defined Hurst exponent. An example is the class of smoothly truncated Lévy flights (STLF’s) [17]. For STLF’s all moments exist, and the property of stationary and independent increments implies a Hurst exponent \( H = 1.5 \). The STLF behaves like a Lévy flight on small time scales, while on long time scales, the statistics are close to Brownian motion [18]. Thus, it is clearly not self-similar nor multifractal, which was proven in [19].
TABLE I. Examples of stochastic processes with well-defined Hurst-exponents $H$. Finite variance is assumed in all examples.

| Stochastic process               | Hurst exponent $H$ |
|----------------------------------|--------------------|
| White noise                      | $H = 1/2$          |
| Random walks                     | $H = 3/2$          |
| fractional Gaussian noise        | $0 < H < 1$        |
| fractional Brownian motion       | $1 < H < 2$        |
| $h$-selfsimilar processes        | $H = h + 1$        |
| Scaling function $\zeta(q)$      | $H = \zeta(2)/2 + 1$ |

III. DETRENDED FLUCTUATION ANALYSIS

A. DFA algorithm

Let $X(1), X(2), \ldots, X(n)$ be the input to DFA. The first step in DFA is to construct the profile

$$Y(t) = \sum_{k=1}^{t} X(k).$$

For a given scale $s$ one considers windows of length $s$. In each window a polynomial of degree $m$ is fitted to the profile. Subtracting the fitted polynomial from the profile gives a set of residuals. From these residuals the variance is calculated. We denote by $F^2_t(s)$ the residual variance. The squared fluctuation function $F^2$ is the average of $F^2_t$. To express the residual variance mathematically, we introduce some notation. Define the vector $Y(t) = [Y(t+1), Y(t+2), \ldots, Y(t+s)]^T$. Let $B$ be the $(m+1) \times s$ design matrix in the ordinary least square (OLS) regression. That is, row $k$ of $B$ is the vector $(1^{k-1}, 2^{k-1}, \ldots, s^{k-1})$. Define

$$Q = B^T (BB^T)^{-1} B,$$

which is known as the hat matrix in statistics. The residual variance is given by

$$F^2_t(s) = \frac{1}{s} Y(t)^T (I - Q) Y(t),$$

where $I$ is the $(s \times s)$ identity matrix.

B. Relation between DFA and variogram/acvf

It is convenient to express the squared fluctuation function explicitly in terms of the input series. Let $X(t) = [X(t+1), X(t+2), \ldots, X(t+s)]^T$. We define the $s \times s$ matrix $D$ by letting element $(i, j)$ of $D$ be equal to one if $i \geq j$ and zero otherwise. Left-multiplying $D$ with $X(t)$ gives the vector of cumulative sums $(X(t+1), X(t+1) + X(t+2), \ldots, \sum_{k=1}^{s} X(t+k))$. Define

$$A = D^T (I - Q) D,$$

and let $a_{k,j}$ be element $(k, j)$ of the matrix $A$. The fluctuation function can be written

$$F^2_t(s) = \frac{1}{s} X(t)^T A X(t)$$

$$= \frac{1}{s} \sum_{k=1}^{s} \sum_{j=1}^{s} a_{k,j} X(t+k) X(t+j).$$

In the definition of DFA the profile is constructed for the whole time series prior to windowing. Eq. (11) states that constructing the profile in each window gives the same residual variance (squared fluctuation function). A proof is found in Appendix A. In the sequel we make the assumption that $\mathbb{E} X(t) = 0$. Denote by $\gamma(t, s)$ the auto-covariance function (acvf) of $X(t)$. Applying the expectation operator to Eq. (10), it is seen that

$$\mathbb{E} F^2_t(s) = \frac{1}{s} \sum_{k=1}^{s} \sum_{j=1}^{s} a_{k,j} \gamma(t+k, t+j).$$

If we add the further restriction of stationarity of the process $X(t)$, Eq. (11) simplifies to Eq. (3), with $\gamma(t) = \gamma(0, t)$ and

$$G(j, s) = \sum_{k=1}^{s-j} a_{k,k+j}.$$
Let us note that the weight functions $G(j,s)$ can be computed exactly. In this work this has been done using Mathematica. The weight function for DFA1 and DFA2 are listed in Table II while the map $j \mapsto G(j,100)$ for DFA2 and DFA5 are shown in Fig. [1].

While Eq. (11) seemingly is time-dependent when $X(t)$ is non-stationary with stationary increments, this is not the case. To establish that $\mathbb{E} F^2_t(s)$ does not depend on the window $t$, we use the following form of the residual variance

$$F^2_t(s) = \frac{1}{2s} \sum_{k=1, j=1}^s a_{k,j} [X(t+k) - X(t+j)]^2,$$  \hspace{1cm} (13)

with proof found in Appendix A. Applying the expectation operator to Eq. (13) results in Eq. (4), with $G(j,s)$ the weight function defined in Eq. (12). We note that Eq. (4) also holds when $X(t)$ is stationary. Since $\mathbb{E} F^2_t(s)$ does not depend on the window $t$, it follows that $\mathbb{E} F^2_t(s) = \mathbb{E} F^2(s)$.

$\text{C. Proof of DFA scaling}$

We are now in a position to prove

$$\mathbb{E} F^2(s) \sim \lambda_m H s^{2H},$$  \hspace{1cm} (14)

for input signals $X(t)$ with Hurst exponents $H \in \{0, 1 \cup (1,2)\}$. We assume $\mathbb{E} X(t)^2 = 1$. In the appendix we derive the asymptotic weight function $G_{\text{asym}}(j,s) \sim G(j,s)$, which takes the form

$$G_{\text{asym}}(j,s) = \left\{ \begin{array}{ll} \sum_{q=0}^{2m+3} s^{2-q} j^{q} d_q & \text{if } j > 0, \\ d_0 s^2 & \text{if } j = 0. \end{array} \right.$$

Expressions for the coefficients $\{d_q\}$ can be found in Appendix B. The values of $\{d_q\}$ for orders $m \leq 6$ are listed in Table III.

For $H \geq 1$, using Eq. (4) and the asymptotic weight function, yield the asserted scaling:

$$\mathbb{E} F(s)^2 \sim -s^{2H} \sum_{q=0}^{2m+3} s^{-1-q} d_q \sum_{j=1}^{s-1} j^{q+2H-2} \sim -s^{2H} \sum_{q=0}^{2m+3} \frac{d_q}{q + 2H - 1}.$$  \hspace{1cm} (15)

In the stationary case $0 < H < 1$, using Eq. (3), we find

$$\mathbb{E} F^2(s) \sim s d_0 \gamma(0) + 2 \sum_{q=0}^{2m+3} s^{-1-q} d_q \sum_{j=1}^{s-1} j^{q} \gamma(j).$$

White noise ($H = 1/2$) is trivial:

$$\mathbb{E} F(s)^2 \sim s d_0 \gamma(0).$$

For $1/2 < H < 1$, in Eq. (15) the second term dominates the first term. Coupled with the asymptotic form of the acvf we have

$$\mathbb{E} F(s)^2 \sim 2H(2H-1) \sum_{q=0}^{2m+3} s^{-1-q} d_q \sum_{j=1}^{s-1} j^{q+2H-2} \sim s^{2H} 2H(2H-1) \sum_{q=0}^{2m+3} \frac{d_q}{q + 2H - 1}.$$  \hspace{1cm} (16)

For $0 < H < 1/2$ the two terms in Eq. (15) scale as $s$. However, the linear term cancels and we end up with the expected scaling: Denote by $\rho(\tau)$ the auto-correlation function (ACF), i.e., $\rho(\tau) = \gamma(\tau)/\gamma(0)$. It is well-known that $\sum_{\tau=-\infty}^{\infty} \rho(\tau) = 0$ (e.g., [20]). Since $\rho(0) = 1$, and the ACF is a symmetric function, we have $-\gamma(0)/2 = \sum_{\tau=1}^{\infty} \gamma(\tau)$. Thus

$$\mathbb{E} F(s)^2 \sim -2d_0 s \sum_{j=s}^{\infty} \gamma(j) + 2 \sum_{j=1}^{2m+3} s^{-1-q} d_q \sum_{j=1}^{s-1} j^{q} \gamma(j) \sim -2d_0 s H(2H-1) \sum_{j=s}^{\infty} j^{2H-2} + 2H(2H-1) \sum_{q=1}^{2m+3} s^{-1-q} d_q \sum_{j=1}^{s-1} j^{q+2H-2} \sim s^{2H} 2H(2H-1) \sum_{q=0}^{2m+3} \frac{d_q}{q + 2H - 1}.$$  \hspace{1cm} (17)

| $q$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $d_q$ | $\frac{1}{15}$ | $-\frac{1}{3}$ | $\frac{1}{7}$ | $\frac{1}{9}$ | $\frac{1}{2}$ | $\frac{1}{10}$ | $\frac{1}{12}$ | $\frac{1}{15}$ | $\frac{1}{18}$ | $\frac{1}{21}$ | $\frac{1}{24}$ | $\frac{1}{27}$ | $\frac{1}{30}$ | $\frac{1}{33}$ | $\frac{1}{36}$ |

TABLE III: The coefficients $\{d_q\}$ for DFA of order $m = 1,2, \ldots, 6$. 
IV. APPLICATION

A. Bias for scaling signals

In [2] the bias (of the DFA fluctuation function) for Hurst exponents $H = 0.5, 0.65, 0.9$ was found by means of Monte Carlo. From this bias they proposed the modified DFA fluctuation function

$$F^2_{\text{mod}}(s) = \frac{F^2(s)}{K^2(s)},$$

(16)

with

$$K^2(s) = \frac{\mathbb{E}F^2(s)\tau^{2H}}{\mathbb{E}F^2(\tau)s^{2H}}.$$  

If we assume $\tau$ is large such that $\mathbb{E}F^2(\tau) = \lambda_{m,H}^{\tau^{2H}}$ holds (approximately), then

$$K^2(s) = \frac{\mathbb{E}F^2(s)}{\lambda_{m,H}s^{2H}},$$

(17)

which implies $\mathbb{E}F^2_{\text{mod}}(s) = \lambda_{m,H}s^{2H}$.

We can use Eq. (3) to calculate the bias for signals with Hurst exponents $0 < H < 1$. An example is shown in Fig. 2, where we have used $H = 0.9$. Of course, this give similar result as in [2] (see their Fig. 2a). Using Eq. (4) we can also compute the bias for signals with Hurst exponents $1 < H < 2$. The bias for $H = 1.1$ is shown in Fig. 2b.

The correction functions Eq. (17) for $H = 0.9$ and $H = 1.1$ are shown in Fig. 3. A practical problem is that $K(s)$ depends on the (unknown) Hurst exponent. In [2] this dependence was found to be weak for $H = 0.5, 0.65, 0.9$. Based on this finding the authors suggested to use Eq. (16) with the correction function for $H = 0.5$. While using this modified DFA will improve the scaling for $H < 1$, it will actually increase the bias for signals with Hurst exponents $H > 1$. For $H = 0.9$ and $H = 1.1$, this can be seen from Figs. 2 and 3, where we observe that the bias has different signs.

B. Ornstein Uhlenbeck

Another application is to study the behaviour of DFA for signals that are not scaling. Here we consider the class of Ornstein-Uhlenbeck (OU) processes. An OU is the solution to the Langevin equation

$$dX(t) = -\frac{1}{\tau}X(t)dt + \sigma dB(t),$$

(18)

where $B(t)$ is a standard ($\mathbb{E}B(1)^2 = 1$) Brownian motion, $\sigma > 0$ is a scale parameter and $\tau > 0$ is the characteristic correlation time. We choose initial condition such that $X(t)$ is stationary. This imply that the auto-covariance function takes the form

$$\mathbb{E}X(t)X(s) = \frac{\exp(-|t-s|/\tau)}{2\sigma^2}$$

(19)

Again, we can use Eq. (3) to calculate the expected value of the squared DFA fluctuation. An example is shown in Fig. 4.

While OU processes do not have well-defined Hurst exponents as defined in Section [1] the second moments scales asymptotically: On long time scales ($\tau \to \infty$) $X(t)$ is white noise, while on short time scales ($\tau \to 0$) $X(t)$ converges to a Brownian motion. Thus, for the DFA fluctuation function we should expect a scaling exponent close to $H = 0.5$ on long time scales. It is seen in Fig. 4 that this holds. On small time scales we need to keep in mind that there is a bias in DFA for signals that are scaling. Relevant here is the bias for random walks ($H = 1.5$). In Fig. 4 it is seen that the OU DFA fluctuation function, with $\tau = 20$, is consistent with random walks on small time scales.

We note that an AR(1) is an discretised OU process, and more results on the AR(1) DFA fluctuation function can be found in [3].

FIG. 2. Detrended fluctuation analysis for input signals with Hurst exponent (a) $H = 0.9$ and (b) $H = 1.1$. In both figures (a,b) the graphs from bottom to top corresponds to DFA of increasing order $m$, from $m = 1$ (bottom) to $m = 6$ (top). Dashed lines are the asymptotic scaling $\lambda_{m,H}^s/\tau^H$ (see text). The squared fluctuation functions have been shifted by factors $10^m$.  

Again, we can use Eq. (3) to calculate the expected value of the squared DFA fluctuation. An example is shown in Fig. 4.

While OU processes do not have well-defined Hurst exponents as defined in Section [1] the second moments scales asymptotically: On long time scales ($\tau \to \infty$) $X(t)$ is white noise, while on short time scales ($\tau \to 0$) $X(t)$ converges to a Brownian motion. Thus, for the DFA fluctuation function we should expect a scaling exponent close to $H = 0.5$ on long time scales. It is seen in Fig. 4 that this holds. On small time scales we need to keep in mind that there is a bias in DFA for signals that are scaling. Relevant here is the bias for random walks ($H = 1.5$). In Fig. 4 it is seen that the OU DFA fluctuation function, with $\tau = 20$, is consistent with random walks on small time scales.

We note that an AR(1) is an discretised OU process, and more results on the AR(1) DFA fluctuation function can be found in [3].
C. Missing data

Based on Eq. \([13]\) we can modify DFA to handle missing data. Define \(\delta(t)\) to be zero if \(X(t)\) is missing and one otherwise. We make the assumption that at least one \(X(t+k)X(t+j)\) is non-missing. A sufficient, but not necessary, condition for this to hold is that at least one window contain no gaps. Let

\[
p_{k,j} = \frac{\# \text{ of windows}}{\# \text{ of non-missing } X(t+k)X(t+j)}.
\]

We propose the estimator

\[
\hat{F}_t^2(s) = -\frac{1}{2s} \sum_{k=1,j=1}^{s} p_{k,j} a_{k,j} \times \]

\[
[X(t+k) - X(t+j)]^2 \times \delta(t+k)\delta(t+j),
\]

(20)

We define \(\hat{F}_t^2(s)\) to be the average of \(\hat{F}_t^2(s)\) (averaging over the different windows \(t\) used). Without missing data the fluctuation function \(\hat{F}(s)\) is the same as fluctuation function \(F(s)\) in the gap-free case. For a time series with gaps, \(\hat{F}(s)\) is equal in expectation to \(F(s)\). The equality \(\mathbb{EF}_t^2(s) = \mathbb{EF}^2(s)\) holds if the input signal is stationary or non-stationary with stationary increments: Applying the expectation operator on Eq. \([20]\) we have

\[
\mathbb{EF}_t^2(s) = -\frac{1}{2s} \sum_{k=1,j=1}^{s} p_{k,j} a_{k,j} S(|k-j|) \times \]

\[
\delta(t+k)\delta(t+j),
\]

and since at least one \(\delta(t+k)\delta(t+j)\) is assumed non-zero, the equality \(\mathbb{EF}_t^2(s) = \mathbb{EF}^2(s)\) follows.

An alternative estimator, based on Eq. \([10]\), is

\[
\tilde{F}_t^2(s) = \frac{1}{s} \sum_{k=1,j=1}^{s} p_{k,j} a_{k,j} \times \]

\[
X(t+k)X(t+j) \times \delta(t+k)\delta(t+j),
\]

(21)

and \(\tilde{F}_t^2(s)\) defined as the average of \(\hat{F}_t^2(s)\). It is straightforward to verify that \(\mathbb{EF}_t^2(s) = \mathbb{EF}^2(s)\) for stationary input signals. However, for input signals that are non-stationary with stationary increments, \(\hat{F}\) does not have the desirable property of equality in expectation to the fluctuation function \(F(s)\) in the gap-free case. As an example, consider an input signal with Hurst exponent \(1 < H < 2\). In the gap-free case, the time-dependent part of the expected squared fluctuation vanish, i.e.,

\[
\mathbb{EF}_t^2(s) = \frac{1}{s} \sum_{k=1}^{s} \sum_{j=1}^{s} a_{k,j} (|t+k|^{2h} + |t+j|^{2h}) = 0,
\]

see Eq. \([A1]\) in the appendix. For \(\hat{F}_t^2(s)\) the time-dependent part will not, in general, vanish. This is due to the additional multiplicative factors \(p_{k,j}\delta(t+k)\delta(t+j)\) in Eq. \([21]\)
V. CONCLUDING REMARKS

Several new propositions, with proofs, for DFA have been presented. This include the relationship between the DFA fluctuation function and acvf/variograms, which was derived from the sample forms Eqs. (10) and (13). We have proven the power-law scaling of the DFA fluctuation function for stochastic processes with Hurst exponents $H \in \{(0,1) \cup (1,2)\}$. These results was derived under the assumption that the input signal in DFA has mean zero. However, we can replace the assumption of mean zero with a polynomial trend of order $q$, provided $q < m$. See Appendix A.1 for a proof. Results on trends not accounted for in DFA can be found in [2] [21].

We have demonstrated that the new theory of DFA has applications. For this purpose we used the weight functions and asymptotic weight functions. The Mathematica code for these functions are found as supplementary material to this article.

Some of the theory presented in this paper is probably a suitable starting point to prove the correctness of the multifractal DFA introduced in [22], as well as the variance and limiting distribution of the DFA fluctuation function.

ACKNOWLEDGMENTS

This work has received support from the Norwegian Research Council under Contract 229754/E10.

Appendix A: Simple proofs

Recall the definition of the weight matrix

$$A = D^T(I - Q)D$$

and hat matrix

$$Q = B^T (BB^T)^{-1} B.$$  

Since $Q$ is a projection matrix, vectors $v$ that are in the row-space of $B$ will be mapped to itself, i.e. $Qv = v$, and thus $(I - Q)v = 0$.

(A9) $\iff$ (A10): Let 1 be a $(s \times 1)$ vector of ones. For $t > 0$:

$$Y(t) = DX(t) + 1Y(t).$$

The proof is completed by noting that 1 is in the row-space of $B$.

(A10) $\iff$ (A13): This equality holds if

$$\sum_{k=1,j=1}^{s} a_{k,j}(X(t+k)^2 + X(t+j)^2) = 0$$

Fix $k$ and consider the sum

$$\sum_{j=1}^{s} a_{k,j}X(t+k)^2,$$

which is element $k$ of the vector $A1X(t+k)^2$. We have $D1 = (1,2, \ldots, n)^T$, which is in the row-space of $B$. Thus $A1X(t+k)^2 = 0$. Since this holds for all $k = 1, 2, \ldots, s$, we can conclude that

$$\sum_{k=1,j=1}^{s} a_{k,j}X(t+k)^2 = 0.$$

Since $A$ is a symmetric matrix we also have

$$\sum_{k=1,j=1}^{s} a_{k,j}X(t+j)^2 = 0.$$

1. Trends

We want to show that the relationship between the DFA fluctuation function and acvf/variograms, Eqs. (11) and (4), and the power-law scaling of the DFA fluctuation function Eq. (14), remains valid when certain trends is superposed the signal. Let $Z(t)$ be a stochastic process with mean zero and acvf $\gamma(t,s)$. It is assumed that $Z(t)$ is either stationary or non-stationary with stationary increments. Define

$$T(t) = \beta_0 + \beta_1 t + \ldots + \beta_q t^q, \quad t = 1, \ldots, n,$$

where $q$ is an integer in the range $0 \leq q \leq m - 1$. Let

$$X(t) = T(t) + Z(t).$$

By Eq. (10) we have

$$F_t^2(s) = \frac{1}{s} \sum_{k=1,j=1}^{s} a_{k,j}Z(t+k)Z(t+j) + \frac{1}{s} \sum_{k=1,j=1}^{s} a_{k,j}Z(t+k)T(t+j) + \frac{1}{s} \sum_{k=1,j=1}^{s} a_{k,j}T(t+k)Z(t+j) + \frac{1}{s} \sum_{k=1,j=1}^{s} a_{k,j}T(t+k)T(t+j)$$

(A2)

Since $EZ(t) = 0$, the middle terms vanish in expectation, and thus

$$\mathbb{E}F_t^2(s) = \frac{1}{s} \sum_{k=1,j=1}^{s} a_{k,j}\gamma(t+k, t+j) + \frac{1}{s} \sum_{k=1,j=1}^{s} a_{k,j}T(t+k)T(t+j)$$

(A3)
We have
\[ \sum_{k=1, j=1}^{s} a_{k,j}T(t + k)T(t + j) = T(t)^T \mathbf{T}(t), \quad (A4) \]
where \( \mathbf{T}(t) = [T(t + 1), \ldots, T(t + s)]^T \). One can use the formulas for sums of powers, e.g. [23], to verify that \( DT(t) \) is in the row-space of \( B \). Hence
\[
\sum_{k=1, j=1}^{s} a_{k,j}T(t + k)T(t + j) = 0,
\]
and thus
\[ EF^2_t(s) = \frac{1}{s} \sum_{k=1, j=1}^{s} a_{k,j} \gamma(t + k, t + j). \]

Appendix B: Asymptotic weight function

The weight matrix can be written
\[ A = D^T D - D^T Q D, \]
where \( Q \) is the hat matrix defined in Eq. (8). Element \((i, j)\) of \( D \) is one if \( i \geq j \) and zero otherwise. Thus, element \((i, j)\) in the first matrix product is
\[ (D^T D)_{i, j} = s + 1 - \max\{i, j\}. \]
Summing the \( j \)th (sub)-diagonal yield
\[
\sum_{k=1}^{s-j} (D^T D)_{k,k+j} = \sum_{k=1}^{s-j} (s + 1 - (k + j)) \\
= s^2/2 - sj + s/2 + j^2/2 - j/2 \\
\sim s^2/2 - sj + j^2/2. \quad (B1)
\]
Calculating the term \( D^T Q D \) is more tedious, but straightforward. Let us start with the hat matrix \( Q \). Denote by \((BB^T)_{i,j}\) element \((i, j)\) of the inverse of \( BB^T \). By observing that column \( j \) of \( B \) is \((j^0, j^1, \ldots, j^m)\), we see that
\[ Q_{p,q} = \sum_{d=1, l=1}^{m+1} p^{d-1} q^{l-1} (BB^T)^{-1}_{d,l}. \]
Using the asymptotic expression of \( BB^T \),
\[ (BB^T)_{i,j} = \sum_{t=1}^{s} t^{i+j} s^{i+j-1} \quad \frac{1}{i+j-1}, \]
one can use the definition of the inverse matrix to verify that
\[ (BB^T)^{-1}_{d,l} \sim \hat{c}_{d,l}/s^{d+l-1} \quad (B3) \]
Inserting Eq. (B3) into Eq. (B2) yields
\[ [D^T Q D]_{i_1,i_2} \sim \sum_{d=1, l=1}^{m+1} (s^d - q^d) (s^l - q^l) c_{d,l} / s^{d+l-1}, \]
where we have defined \( c_{d,l} = \hat{c}_{d,l}/(dl) \). Summing the \( j \)th (sub)-diagonal yield
\[ \sum_{k=1}^{s-j} (D^T Q D)_{k,k+j} \sim \sum_{k=1}^{s-j} (s^d - q^d) (s^d - (k + j)^d) c_{d,l} / s^{d+l-1} \sim \sum_{k=1}^{s-j} c_{d,l} \left( s^d - sj - \frac{s^2}{l+1} + \frac{s^{l-j+1}}{l+1} \right) \]
\[ - \sum_{d=1, l=1}^{m+1} c_{d,l} \frac{s^{-d+1} (s - j)^{d+1}}{d+1} \]
\[ + \sum_{d=1, l=1}^{m+1} c_{d,l} \sum_{r=0}^{l} \left( \frac{1}{r} \right) \frac{(s - j)^{d+l+1-r} j^r}{(d+l+1-r)s^{d+l-1}} \]
\[ = \sum_{q=0}^{2m+3} s^2 - q^2 \hat{b}_q. \quad (B7) \]
The terms [B4]-[B6] can be written
\[ \sum_{q=0}^{2m+3} s^2 - q^2 \hat{b}_q^{(k)}, \quad k = 1, 2, 3, \]
respectively. This implies the equality Eq. (B7), with
\[ b_q = b_q^{(1)} + b_q^{(2)} + b_q^{(3)}. \]
The coefficients \( b_q^{(k)} \), found by re-organising terms, are given by:
\[
\begin{align*}
b_q^{(1)} &= \left\{ \begin{array}{ll}
\sum_{d=1, l=1}^{m+1} c_{d,l} - c_{d,l+1}/(d) & \text{if } q = 0, \\
- \sum_{d=1, l=1}^{m+1} c_{d,l} & \text{if } q = 1, \\
\frac{1}{q} \sum_{d=1}^{m+1} c_{d,q-1} - 0 & \text{if } 2 \leq q \leq m + 2, \\
0 & \text{if } q > m + 2.
\end{array} \right.
\end{align*}
\]
Using Eq. (B1) and (B7), the coefficients follow:

\[ b_q^{(2)} = \begin{cases} 
- \sum_{d=1, l=1}^{m+1} \frac{c_{d,l}}{d+1} & \text{if } q = 0, \\
\sum_{d=1, l=1}^{m+1} \frac{c_{d,l}}{d+1} (d+1) & \text{if } q = 1, \\
(-1)^{q-1} \sum_{d=q, l=1}^{m+1} \frac{c_{d,l}}{d+1} (d+1-q) & \text{if } 2 \leq q \leq m+2, \\
0 & \text{if } q > m+2. 
\end{cases} \]

\[ b_q^{(3)} = \sum_{d+l+q-1, d \geq 1, l \geq 1}^{m+1} a_q^{(d,l)} c_{d,l}, \]

\[ a_k^{(d,l)} = \min \{l,k\} \sum_{r=0}^\infty \binom{l}{r} \frac{(-1)^{k-r}}{d+l+1-r} \frac{(d+l+1-r)}{(d+l+1-k)}, \]

\[ d_q = \begin{cases} 
1/2 - b_0 & \text{if } q = 0, \\
-1 - b_1 & \text{if } q = 1, \\
1/2 - b_2 & \text{if } q = 2, \\
-b_q & \text{if } q > 2. 
\end{cases} \]

[1] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Phys. Rev. E 49, 1685 (1994).
[2] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde, Physica A 295, 441 (2001).
[3] O. Løvsletten and M. Rypdal, J. Clim. 29, 4057 (2016).
[4] S. Lahmiri, Physica A 437, 130 (2015).
[5] J.-Y. Chiang, J.-W. Huang, L.-Y. Lin, C.-H. Chang, F.-Y. Chu, Y.-H. Lin, C.-K. Wu, J.-K. Lee, J.-J. Hwang, J.-L. Lin, and F.-T. Chiang, PloS one 11, e0147282 (2016).
[6] O. Løvsletten and M. Rypdal, Phys. Rev. E 85, 046705 (2012).
[7] M. Höll and H. Kantz, Eur. Phys. J. B 88, 327 (2015).
[8] M. Höll and H. Kantz, Eur. Phys. J. B 88, 126 (2015).
[9] M. S. Taqqu, V. Teverovsky, and W. Willinger, Fractals 03, 785 (1995).
[10] C. Heneghan and G. McDarby, Phys. Rev. E 62, 6103 (2000).
[11] P. S. Wilson, A. C. Tomsett, and R. Touni, Phys. Rev. E 68, 017103 (2003).
[12] D. Mondal and D. B. Percival, Ann. Inst. Stat. Math. 62, 943 (2008).
[13] B. B. Mandelbrot and J. W. Van Ness, SIAM Rev. 10, 422 (1968).
[14] P. Embrechts and M. Maejima, *Selfsimilar Processes, Princeton Series in Applied Mathematics* (Princeton University Press, 2002).
[15] J.-F. Muzy and E. Bacry, Phys. Rev. E 66, 056121 (2002).
[16] E. Bacry and J. F. Muzy, Commun. Math. Phys. 236, 449 (2003).
[17] I. Koponen, Phys. Rev. E 52, 1197 (1995).
[18] G. Terdik, W. A. Woyczynski, and A. Piryatinska, Phy. Lett. A 348, 94 (2006).
[19] M. Rypdal and K. Rypdal, Earth Syst. Dyn. 7, 281 (2016).
[20] J. Beran, *Statistics for long-memory processes*, Vol. 61 (CRC press, 1994).
[21] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. E. Stanley, Phys. Rev. E 64, 011114 (2001).
[22] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, Physica A 316, 87 (2002).
[23] E. W. Weisstein, “Power sum,” From MathWorld–A Wolfram Web Resource. [PowerSum.html](http://mathworld.wolfram.com/PowerSum.html).