A new distribution-based test of self-similarity

Bianchi, Sergio

University of Cassino (Italy)

2004

Fractals, Vol. 12, No. 3 (2004) 331–346

World Scientific
www.worldscientific.com

A NEW DISTRIBUTION-BASED TEST OF SELF-SIMILARITY

SERGIO BIANCHI
Faculty of Economics, University of Cassino
Via Mazzaroppi — 03043 Cassino (FR), Italy
sbianchi@eco.unicas.it

Received October 28, 2003
Accepted January 20, 2004

Abstract

In studying the scale invariance of an empirical time series a twofold problem arises: it is necessary to test the series for self-similarity and, once passed such a test, the goal becomes to estimate the parameter H_0 of self-similarity. The estimation is therefore correct only if the sequence is truly self-similar but in general this is just assumed and not tested in advance. In this paper we suggest a solution for this problem. Given the process $\{X(t), t \in T\}$, we propose a new test based on the diameter δ of the space of the rescaled probability distribution functions of $X(t)$. Two necessary conditions are deduced which contribute to discriminate self-similar processes and a closed formula is provided for the diameter of the fractional Brownian motion (fBm). Furthermore, by properly choosing the distance function, we reduce the measure of self-similarity to the Smirnov statistics when the one-dimensional distributions of $X(t)$ are considered. This permits the application of the well-known two-sided test due to Kolmogorov and Smirnov in order to evaluate the statistical significance of the diameter δ, even in the case of strongly dependent sequences. As a consequence, our approach both tests the series for self-similarity and provides an estimate of the self-similarity parameter.

Keywords: Distance; Fractional Brownian Motion; Kolmogorov-Smirnov Test; Self-similarity.

1. INTRODUCTION

Since the pioneering works by Mandelbrot in the early 1960s,[1–3] stochastic processes exhibiting statistical self-similarity, i.e. distributional invariance with respect to a proper scaling rule, have been suggested as models in many fields such as finance, geophysics or network traffic modeling (see e.g. Refs. 4–9 or 10–12 for a comprehensive bibliography). Particularly, in finance an increasing attention is being paid to the scaling laws of stock returns because the assumption of self-similarity in strong sense is implicit both in the standard financial theory (stating that the self-similarity parameter H_0 equals $\frac{1}{2}$) and in less consolidate frameworks such as the fractal Gaussian models (for whom $H_0 \in (0, 1]$) or the α-stable models (see, e.g. Ref. 13). The scaling structure of actual data is usually deduced by analyzing the sample moments, but this approach could be misleading because of several reasons, the most "embarassing" being perhaps the assumption of existence of the analyzed moments; just as an example it suffices to think of the non-degenerate α-stable processes, having finite moments of order lower than α.[14] Moreover, even when variance does exist, the best one can do using the large class of variance-based estimators is to discriminate second order self-similar sequences but not truly self-similar processes.

In order to overcome this problem, in this work we will focus on distributions: the standard condition of self-similarity based on the equality of the finite-dimensional distributions will be given using the notion of "diameter." Roughly speaking, this is the maximum of the set whose elements are the distances between the distribution functions generated by the scaling rule which defines the notion of self-similarity. Being the value of the scaling parameter H_0 *a priori* unknown, the diameter — which momentarily we denote by δ — will depend on the variable H.

We first state that a stochastic process is self-similar with exponent H_0 if and only if δ equals zero when $H = H_0$ (Proposition 1). Hence, the basic idea is to estimate the self-similarity parameter by solving the equation $\delta = 0$ but concretely, when the empirical distributions are considered, even the diameter of a truly self-similar process will be larger than zero for $H = H_0$. Therefore, it becomes necessary to evaluate the statistical significance of δ, but in order to do this a candidate H_0, the diameter of which must have been calculated in advance, is needed. The problem is solved by a nice property of the diameter, which for an H_0 self-similar process is proved to be non-increasing for $H \leq H_0$ and non-decreasing for $H \geq H_0$ (Proposition 2). This permits the application of a statistical test with respect to $\min(\delta)$, which for an H_0 self-similar process is unique.

A further necessary condition of self-similarity is exploited in Proposition 3. In addition, when the process is the celebrated fractional Brownian motion (fBm) a closed formula is provided for the diameter of the space of the rescaled probability distributions (Proposition 4).

Finally, just a few words on the statistical test used to evaluate the significance of $\min(\delta)$: a proper choice of the norm which defines the diameter allows to apply the Kolmogorov-Smirnov goodness of fit test, even in the case of strongly dependent data (Proposition 5).

So, the technique we propose both tests scale invariance and estimates the self-similarity parameter. The remainder of the paper is organized as follows: in Sec. 2 some definitions are given and, once characterized the behavior of the metric we will introduce, in Sec. 3 the use of the Kolmogorov-Smirnov test is motivated. Section 4 provides a comparative application on simulated data coming from samples of self-similar and non self-similar processes. Appendix A contains the proofs of the Propositions.

2. A METRIC FOR SELF-SIMILARITY

Before introducing the new approach we suggest, let us recall the definition of self-similarity.

Definition 1. The continuous time, real-valued process $\{X(t),\, t \in T\}$, with $X(0) = 0$, is self-similar with index $H_0 > 0$ (concisely, H_0-ss) if, for any $a \in \mathbb{R}^+$ and any integer k such that $t_1, \ldots, t_k \in T$, the following equality holds for its finite-dimensional distributions

$$\{X(at_1),\, X(at_2), \ldots, X(at_k)\} \stackrel{d}{=} \{a^{H_0}X(t_1),\, a^{H_0}X(t_2), \ldots, a^{H_0}X(t_k)\}\,. \quad (1)$$

Recall also that the second-order stationary, real-valued stochastic process $X(t)$ is said H_0-second order self-similar if — denoted by $Y(t, a)$ its a-lagged increments, namely $Y(t,\, a) = X(t + a) - X(t)$, and by $\overline{Y}(t,\, m) = m^{-1}\sum_{\tau=(t-1)m+1}^{tm} Y(\tau,\, 1)$, $m,\, t \in \{1, 2, \ldots\}$ the averaged (over blocks of length m) sequence — it holds

$$\mathrm{Var}(\overline{Y}(t,\, m)) = m^{2(H_0-1)}\mathrm{Var}(Y(t,\, 1))\,.$$

Finally, $X(t)$ is said H_0-second order asymptotically self-similar if

$$\mathrm{Var}(\overline{Y}(t,\, km)) \sim k^{2(H_0-1)}\mathrm{Var}(\overline{Y}(t,\, m)) \quad \text{as } m \to \infty, k \in \{1, 2, \ldots\}\,.$$

Since equality (1) implies

$$\mathbb{E}(|X(t)|^q) = t^{H_0 q}\mathbb{E}(|X(1)|^q) \tag{2}$$

self-similarity is usually tested via the scaling behavior of the sample moments of $X(t)$ but this approach has some drawbacks. First, there is no equivalence between moments and distribution, therefore (2) does not imply (1). Second, the moment-based techniques study only particular moments, usually absolute moments or variance (see e.g. 15 and 16 for a survey of several methods); as a consequence, no conclusion can be drawn about the process' self-similarity but only about some weak form of self-similarity, such as, for example, the second order or the asymptotical self-similarity. Finally, the analyses are generally lacking from an inferential (possibly non-parametric) perspective.

A different perspective is to test directly Eq. (1). To do this, let $\mathcal{A}$ be any bounded subset of $\mathbb{R}^+$, $\mathfrak{a} = \min(\mathcal{A})$ and $\mathfrak{A} = \max(\mathcal{A}) < \infty$. For any $a \in \mathcal{A}$, consider the set $\{X(at)\}$ of the a-lagged rescaled process and — denoted by Φ the k-dimensional distribution of X — rewrite equality (1) as

$$\Phi_{\mathbb{X}(a)}(\mathbf{x}) = \Phi_{a^{H_0}\mathbb{X}(1)}(\mathbf{x}) \tag{3}$$

where, with concise notation, we have set $\mathbb{X}(a) = (X(at_1), \ldots, X(at_k))$ and $\mathbf{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k$. A trivially equivalent but insightful way of writing Eq. (3) uses the *variable* H as follows. From (3), one has

$$\begin{aligned}
&\Phi_{a^{-H}\mathbb{X}(a)}(\mathbf{x})\\
&\quad = \Pr(a^{-H}X(at_1) < x_1, \ldots, a^{-H}X(at_k) < x_k)\\
&\quad = \text{by self-similarity}\\
&\quad = \Pr(a^{H_0-H}X(t_1) < x_1, \ldots, a^{H_0-H}X(t_k) < x_k)\\
&\quad = \Phi_{a^{H_0-H}\mathbb{X}(1)}(\mathbf{x})\\
&\quad = \Pr(X(t_1) < a^{H-H_0}x_1, \ldots, X(t_k) < a^{H-H_0}x_k)\\
&\quad = \Phi_{\mathbb{X}(1)}(a^{H-H_0}\mathbf{x})\,.
\end{aligned} \tag{4}$$

Denoted by $\Psi_H = \{\Phi_{a^{-H}\mathbb{X}(a)}(\mathbf{x}),\ a \in \mathcal{A}\}$ the set of the absolutely continuous k-dimensional probability distribution functions of $\{a^{-H}X(at)\}$ and considered as distance function ρ the one induced by the sup-norm $\|\cdot\|_\infty$ with respect to the set $\mathcal{A}$, we assume the diameter[a] of the metric space (Ψ_H, ρ)

$$\delta^k(\Psi_H) = \sup_{\mathbf{x}\in\mathbb{R}^k}\sup_{a_i,a_j\in\mathcal{A}} |\Phi_{a_i^{-H}\mathbb{X}(a_i)}(\mathbf{x}) - \Phi_{a_j^{-H}\mathbb{X}(a_j)}(\mathbf{x})| \tag{5}$$

as a measure of *discrepancy* among the rescaled distributions.

Given the above assumptions, the following proposition holds

Proposition 1. *$\{X(t),\ t \in T\}$ is H_0-ss if and only if, for any bounded $\mathcal{A} \subset \mathbb{R}^+$ and any integer k, $\delta^k(\Psi_{H_0}) = 0$.*

(Proof in Appendix A).

Remark. Notice that combining Proposition 1 and the uniqueness of the self-similarity parameter H_0,[17] *for an H_0-ss process $\delta^k(\Psi_H) > 0$* for $H \neq H_0$. □

Whenever the process X is H_0-ss, by (4) one has

$$\begin{aligned}
&\sup_{a_i,a_j\in\mathcal{A}} |\Phi_{a_i^{-H}\mathbb{X}(a_i)}(\mathbf{x}) - \Phi_{a_j^{-H}\mathbb{X}(a_j)}(\mathbf{x})|\\
&\quad = \sup_{a_i,a_j\in\mathcal{A}} |\Phi_{\mathbb{X}(1)}(a_i^{H-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(a_j^{H-H_0}\mathbf{x})|
\end{aligned}$$

and, Φ being a cumulative distribution function and thus monotonous, the supremum is reached by maximizing the dilation of the vector $|a_i^{H-H_0} - a_j^{H-H_0}|\mathbf{x}$, that is the distance between the two terms $a_i^{H-H_0}$ and $a_j^{H-H_0}$, namely

$$\sup_{a_i,a_j\in\mathcal{A}} |(a_i^{H-H_0} - a_j^{H-H_0})\mathbf{x}| = |(\mathfrak{A}^{H-H_0} - \mathfrak{a}^{H-H_0})\mathbf{x}|\,. \tag{6}$$

By (6), the diameter of an H_0-ss process becomes

$$\delta^k(\Psi_H) = \sup_{\mathbf{x}\in\mathbb{R}^k} |\Phi_{\mathfrak{A}^{-H}\mathbb{X}(\mathfrak{A})}(\mathbf{x}) - \Phi_{\mathfrak{a}^{-H}\mathbb{X}(\mathfrak{a})}(\mathbf{x})| \tag{7}$$

and relationship (7) characterizes the behavior of $X(t)$ as follows

Proposition 2. *Let $X(t)$ be H_0-ss, $\mathfrak{A} > \mathfrak{a}$ and $\mathbf{x} \geqq \mathbf{0}$ or $\mathbf{x} \leqq \mathbf{0}$. Then $\delta^k(\Psi_H)$ as function of H is non-increasing for $H \leq H_0$ and non-decreasing for $H \geq H_0$.* (Proof in Appendix A.)

Remark. Proposition 2 is very useful for at least two reasons: first, being a necessary condition, it provides a preliminary filter in the analysis of self-similarity, which can be rejected if the monotonicity does not occur; second, as will be seen afterwards,

[a]We recall that, given the metric space (X, ρ), the number $\delta = \sup_{x,y\in X}\rho(x, y)$ is called *diameter of the set (space) X*.

it gives the value of the diameter for which an inferential test can be performed in order to evaluate the statistical significance of $\min(\delta)$. □

Proposition 3. *Let $X(t)$ be H_0-ss, $\mathbf{x} \geqq \mathbf{0}$ or $\mathbf{x} \leqq \mathbf{0}$ and let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be a sequence of sets such that — denoted by $\mathfrak{a}_i = \min(\mathcal{A}_i)$ and by $\mathfrak{A}_i = \max(\mathcal{A}_i) < \infty$, $i = 1, \ldots, n$ — it is $\mathfrak{a}_i \leq \mathfrak{a}_j$ and $\mathfrak{A}_i \geq \mathfrak{A}_j$ for $i > j$. Then with respect to the sequence $\{\mathcal{A}_n\}$ the diameter δ^k is:* (i) *non-decreasing, provided that $H \neq H_0$;* (ii) *an identically zero function if $H = H_0$.* (Proof in Appendix A)

Remark. When $k = 1$, inequalities $\mathbf{x} \geqq \mathbf{0}$ or $\mathbf{x} \leqq \mathbf{0}$ of Proposition 2 and 3 can be relaxed as can be easily seen by the proofs. □

Proposition 4. *Let $X(t)$ be a fractional Brownian motion with parameter H_0. Then*

$$\delta^1(\Psi_H) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} \left| \displaystyle\int_{-\hat{x}\mathfrak{a}^{H_0-H}}^{-\hat{x}\mathfrak{A}^{H_0-H}} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u \right|, & H \neq H_0 \\ 0, & H = H_0 \,. \end{cases} \tag{8}$$

where $\hat{x} = \sqrt{\frac{2(H_0-H)}{\mathfrak{A}^{2(H_0-H)} - \mathfrak{a}^{2(H_0-H)}} \ln \frac{\mathfrak{A}}{\mathfrak{a}}}$. (Proof in Appendix A).

3. EVALUATING H_0-ss BY THE KOLMOGOROV-SMIRNOV GOODNESS-OF-FIT TEST

Proposition 2 and relationship (7) are the core of this work: the former discriminates if an empirical realization *can* come from a self-similar process and, once checked that the necessary condition is not violated, the latter provides the way to assess the statistical significance of the discrepancy between the distributions. In particular, self-similarity of $X(t)$ implies that $\delta^k(\Psi_{H_0})$ is zero for any real bounded $\mathcal{A}$ but this theoretical result must be stated precisely when the sample distributions are considered. To be rigorous, consider the case of the one-dimensional distributions, and momentarily assume to know the scaling parameter H_0. Denoting by $\tilde{\mathbb{X}}(a) = (a^{-H_0}\tilde{X}(at_k))$, $k = 1, \ldots, n(a)$, the array whose components represent the a-lagged, H_0-rescaled sample path, the step functions

$$\begin{aligned} \tilde{\Phi}_{n(\mathfrak{A}),\tilde{\mathbb{X}}(\mathfrak{A})}(x) &= \frac{1}{n(\mathfrak{A})} \#\{1 \leq k \leq n(\mathfrak{A}) : \mathfrak{A}^{-H_0}\tilde{X}(\mathfrak{A}t_k) \leq x\} \\ \tilde{\Phi}_{n(\mathfrak{a}),\tilde{\mathbb{X}}(\mathfrak{a})}(x) &= \frac{1}{n(\mathfrak{a})} \#\{1 \leq k \leq n(\mathfrak{a}) : \mathfrak{a}^{-H_0}\tilde{X}(\mathfrak{a}t_k) \leq x\} \end{aligned} \tag{9}$$

represent the empirical distribution functions corresponding to the maximum and the minimum of $\mathcal{A}$. Owing to empirical errors, even for a true H_0-ss process the sample diameter will be a non-negative random variable. Therefore the crucial point becomes to establish if there exists an H_0 such that the corresponding $\tilde{\delta}^1(\tilde{\Psi}_{H_0})$ is statistically negligible (here $\tilde{\Psi}_{H_0}$ denotes the set of the empirical distributions and $\tilde{\delta}$ the sample diameter). The choice $\rho = \|\cdot\|_\infty$ is motivated just by this aim: setting $k = 1$, the problem of evaluating the significance of $\tilde{\delta}^1(\tilde{\Psi}_H)$ for some H can be reduced by Proposition 2 to the two-sided Kolmogorov-Smirnov goodness-of-fit test (shortly, KS test), specified as follows: there exists a value H_0 such that, for any $a_i, a_j \in \mathcal{A}$ and any bounded $\mathcal{A} \subset \mathbb{R}^+$

Null hypothesis[b] $\quad \tilde{\Phi}_{n(a_i),\tilde{\mathbb{X}}(a_i)}(x) = \tilde{\Phi}_{n(a_j),\tilde{\mathbb{X}}(a_j)}(x)$ for each x

against $\quad \tilde{\Phi}_{n(a_i),\tilde{\mathbb{X}}(a_i)}(x) \neq \tilde{\Phi}_{(a_j),\tilde{\mathbb{X}}(a_j)}(x)$ for at least one x.

[b] As will be better observed in the following, if the distribution of $X(t)$ is completely specified and known, denoted by $n(a)$ the sample size, the alternatives become (null hypothesis) $\tilde{\Phi}_{n(a),\tilde{\mathbb{X}}(a)}(x) = \Phi(x)$ for each x against $\tilde{\Phi}_{n(a),\tilde{\mathbb{X}}(a)}(x) \neq \Phi(x)$ for at least one x and the statistics to be used is due to Kolmogorov.

The good news is that, when one deals with ergodic sequences, the assumption of independence needed by the KS test to ensure that the empirical distributions really approximate the unconditional one can be relaxed. In fact, once the sequences are rescaled, the equality of the sample distributions is ensured by the *self-similarity* itself and one has the following.

Proposition 5. *Let:*

(i) $X(t)$ *be* H_0*-ss with stationary increments* $Y(t, a) = X(t+a) - X(t)$;
(ii) $\tilde{X}(t_k)_{k=1,\dots,N}$, $t_k \in \mathbb{R}^+$, *be one of the trajectories of* $X(t)$;
(iii) $\tilde{\mathbb{X}}(a) = (a^{-H_0}\tilde{X}(at_k))$ *be the a-lagged,* H_0*-rescaled sample path of* $X(t)$;
(iv) $\tilde{\mathbb{Y}}(a) = (a^{-H_0}\tilde{Y}(t_k, a))$ *be the a-lagged,* H_0*-rescaled sample path designed by the empirical increments* $\tilde{Y}$.

Then, denoted by $\mathcal{L}(z)$ *the distribution derived by Smirnov,*[18]

$$\Pr\left(\sup_{1\le x\le n}\left|\tilde{\Phi}_{n,\tilde{\mathbb{X}}(1)}(x) - \tilde{\Phi}_{n,\tilde{\mathbb{X}}(a)}(x)\right| < \frac{z}{\sqrt{n}}\right) \xrightarrow[n\longrightarrow\infty]{} \mathcal{L}(z)\,. \tag{10}$$

(Proof in Appendix A).

We now come to consider another problem: the KS statistic is a function of one variable (x) whereas the diameter $\tilde{\delta}^1$ is a function of x, but also of H and $\mathcal{A}$; hence, the number of variables must be reduced.

The choice of the set $\mathcal{A}$ is in some ways determined by exogenous considerations about the nature of the analyzed process. In the case of financial stocks, $\mathcal{A}$ is a set of trading horizons and therefore it is constituted by those horizons for which it makes sense to consider a trading activity. The negligible relevance of the set $\mathcal{A}$ can be motivated in the light of Proposition 3: when $X(t)$ is a *true* H_0-ss the sequence $\{\mathcal{A}_n\}_{n\in\mathbb{N}}$ only affects the values $\tilde{\delta}^1(\tilde{\Psi}_{H\neq H_0})$, since — by definition of self-similarity and by Proposition 1 — $\tilde{\delta}^1(\tilde{\Psi}_{H_0})$ is a non-negative stationary random function with respect to any sequence $\mathcal{A}_n \subset \mathbb{R}^+$, with empirical distribution given by (10). Hence, the problem becomes to determine the value H_0 the KS statistic must be calculated for and this can be done by virtue of the diameter's monotonicity.

Restricting here the interest to non-degenerate H_0-ss processes with at least the first moment finite and to Gaussian processes, in the following it will be assumed $H \in (0, 1]$ (for the details about the range of H in these cases, see Ref. 14, pp. 316–317). In this interval H can take any real value and the problem to choose the right value is addressed by Proposition 2. When X is H_0-ss we expect the diameter to be non-increasing for $H \le H_0$ and non-decreasing for $H \ge H_0$; as a consequence, fixed the set $\mathcal{A}$, we will first calculate $\tilde{\delta}_{\min} = \min_{H\in(0,1]}\{\tilde{\delta}^1(\tilde{\Psi}_H)\}$ and then the significance will be evaluated for the value $\tilde{\delta}_{\min}$.

At this point two cases should be distinguished:

Case 1. The distribution of the process is known and equals $\Phi(x)$. This is, for example, the case of the fractional Brownian motion $B_{H_0}(t)$ (fBm), for which $\Phi(x)$ is given by (21) in Appendix A. In this case the statistic is the one originally defined by Kolmogorov

$$D_n = \sup_{x\in\mathbb{R}}|\tilde{\Phi}_{n,a^{-H}\mathbb{X}(a)}(x) - \Phi(x)|\,, \quad \text{for any fixed positive real } a\,.$$

As is well known, the exact computation of the distribution of D_n is possible but very toilsome; hence Kolmogoro[19] first and Smirnov[18] later provided the following asymptotical distribution

$$\Pr\left(D_n \le \frac{z}{\sqrt{n}}\right) = \mathcal{L}(z) = 1 - 2\sum_{j=1}^{\infty}(-1)^{j-1}\exp(-2j^2z^2) \tag{11}$$

an approximation of which, obtained using only the first term of the series, is satisfactory for $n \ge 35$ and leads to

$$\frac{z}{\sqrt{n}} = z_{\alpha,n} \simeq \sqrt{-\frac{1}{2n}\ln\frac{\alpha}{2}}$$

where as usual α denotes the significance level.

Case 2. The distribution of the process is unknown. In this case we can only observe the discrepancy between the sample distributions. So, the statistics is the one argued by Smirnov

$$D_{n_1,n_2} = \sup_{x\in\mathbb{R}}|\tilde{\Phi}_{n_1,a_i^{-H}\mathbb{X}(a_i)}(x) - \tilde{\Phi}_{n_2,a_j^{-H}\mathbb{X}(a_j)}(x)|\,, \quad \text{for any fixed } a_i, a_j \in \mathbb{R}^+\,.$$

Also in this case the asymptotical distribution (11) holds and the approximation becomes

$$z_{a,n_1,n_2} \simeq \sqrt{-\frac{1}{2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\ln\frac{\alpha}{2}}.$$

Once calculated $\tilde{\delta}_{\min}$, we set $D_n = \tilde{\delta}_{\min}$ (or $D_{n_1,n_2} = \tilde{\delta}_{\min}$, depending on the case) and compare the value with $z_{\alpha,n}$ (or z_{α,n_1,n_2}); self-similarity (with parameter $\tilde{H}_0 = \tilde{\delta}^{-1}_{\min}$) will be rejected at a significance level α if $D_n > z_{\alpha,n}$ ($D_{n_1,n_2} > z_{\alpha,n_1,n_2}$).

Although in this paper the continuity of the distribution function has been assumed, the Kolmogorov test can be used even when the continuity does not hold (as in the case of α-stable processes). In this case in fact $\Pr(D_n \leq z_{\alpha,n}) \geq \sqrt{-\frac{1}{2n}\ln\frac{\alpha}{2}}$, so the test can be used conservatively.[20]

4. NUMERICAL SIMULATION

Self-similarity has been tested using as term of comparison the celebrated fractional Brownian motion (fBm), which is the only Gaussian self-similar random process with stationary increments. fBm is widely used in many fields such as finance, insurance, signal processing or geophysics. Furthermore, in order to show the improvement achieved by our estimator, examples have been designed where self-similarity does not hold and the estimates obtained by using different methods have been compared. To do this we have considered two further sets of simulations: first we have generated some samples of multifractional Brownian motion (mBm), which is known to be not self-similar, and finally we have taken into consideration the case of ergodic sequences based on uniformly distributed random variables.

4.1. Fractional Brownian Motion (fBm)

The analysis has been performed through the following steps:

Step 1. Two hundred independent samples of fBm of length $N = 2000$ have been generated for each of the following parameters: $H_0 = 0.3, 0.4, 0.5, 0.6, 0.7$. The simulations have been carried out with the circulant matrix method introduced by Wood and Chan.[21] The algorithm, known to be as one of the most stable, provides in a fast way an excellent approximation of fBm (see Refs. 22, 23 or 24 for a detailed discussion of several simulating methods).

Step 2. Once fixed the set of lags $\mathcal{A} = \{1, 2, \ldots, 100\}$, for $a \in \mathcal{A}$ and for H in the range (0, 1] with step $\Delta H = 0.005$, the a-lagged, H-rescaled sequences have been calculated for each simulated series.

Step 3. Once calculated the empirical distribution function of each sequence, the distance $\tilde{\delta}(\tilde{\Psi}_H) = \sup_x |\tilde{\Phi}_{n_1,\tilde{\mathbb{X}}(1)}(x) - \tilde{\Phi}_{n_a,\tilde{\mathbb{X}}(a)}(x)|$ has been computed (for sake of simplicity, we dropped from the notation the index 1 which reminds that the one-dimensional diameter is being considered).

Step 4. The significance of $\tilde{\delta}_{\min} = \min_H\{\tilde{\delta}(\tilde{\Psi}_H)\}$ has been evaluated by comparing it with the KS statistic.

Step 5. Whenever the process is H_0-ss, Proposition 2 ensures that the map $\tilde{\delta}_{\min}$ is invertible. Hence, the parameter of self-similarity is estimated as $\tilde{H}_0 = \tilde{\delta}^{-1}_{\min}$.

The software needed for the analysis has been developed in S-PLUS® 6.1 for Windows.

When the analysis is performed on samples of fBm there would be no need to check the monotonicity of $\tilde{\delta}$ since the diameter is expected to behave as stated by Proposition 4, where the integral function (8) is monotonic at left and at right with respect to H_0. Just as an example, the behavior of $\tilde{\delta}$ is reproduced for one set of realizations in Fig. 1; the samples are relative to 50 simulated fBm's with nominal parameter $H_0 = 0.6$. It is quite apparent that all the samples: (1) have minima corresponding to abscissae close (centered) to the value 0.6, which is indeed the nominal generation parameter of the simulated series; and (2) behave as stated by Proposition 2, being decreasing for $H \leq H_0$ and increasing for $H \geq H_0$. Figure 2 shows the diameter's trend for increasing values of the maximum lag; as claimed by Proposition 3, $\tilde{\delta}$ is non-decreasing when H is different from H_0 and substantially constant when $H = H_0$ (really, $\tilde{\delta}(\tilde{\Psi}_{0.605}) = 0.0374$ for $\mathcal{A} = \{1, \ldots, 25\}$, $\tilde{\delta}(\tilde{\Psi}_{0.60}) = 0.0396$ for $\mathcal{A} = \{1, \ldots, 50\}$, $\tilde{\delta}(\tilde{\Psi}_{0.60}) = 0.0.0429$ for $\mathcal{A} = \{1, \ldots, 100\}$ and $\tilde{\delta}(\tilde{\Psi}_{0.61}) = 0.0431$ for $\mathcal{A} = \{1, \ldots, 200\}$, but in all cases the diameter is statistically negligible at the level $\alpha = 5\%$).

Figure 3 shows the comparison between the (averaged) sample and the theoretical diameter for the same 50 traces of fBm. The latter, calculated by Proposition 4, has been shifted upward of the quantity $z_{0.05,1900} \simeq 0.019854$.

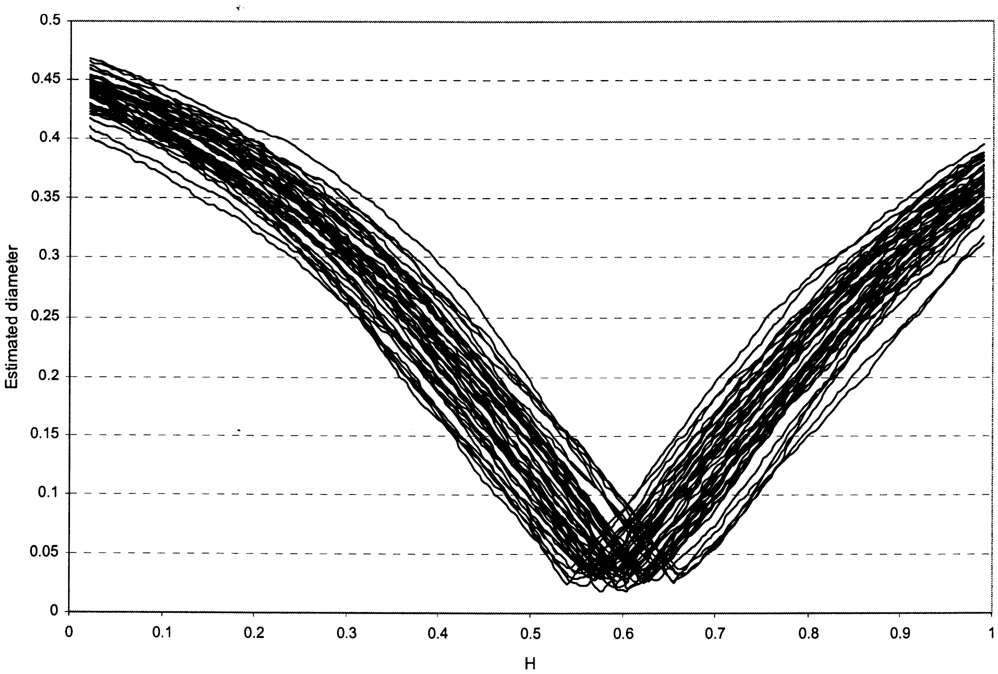

Fig. 1 Diameter for a set of 50 simulated fBm with nominal parameter $H_0 = 0.6$.

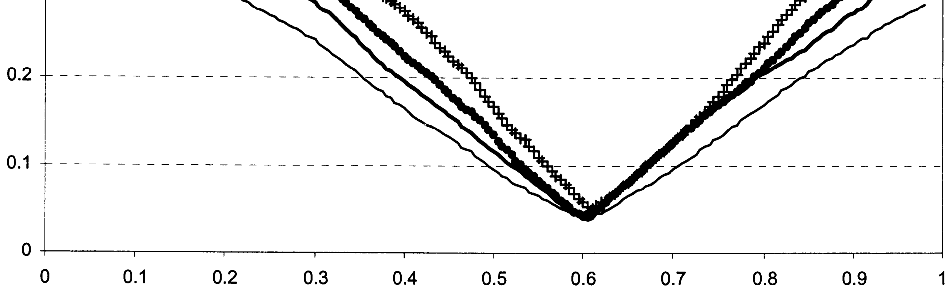

Fig. 2 Behavior of the estimated diameter of a sampled $B_{0.6}(t)$ for increasing maximum lag.

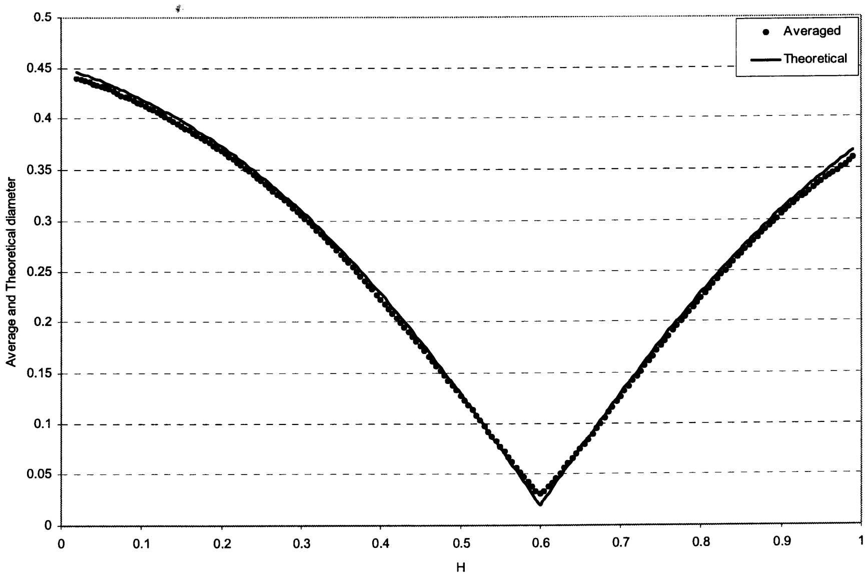

Fig. 3 Sample (average) and theoretical diameter for fBm with parameter $H_0 = 0.6$.

Table 1 **Self-similarity analysis for fBm $N = 2000$, $\mathfrak{a} = 1$, $\mathfrak{A} = 100$, 200 samples.**

(*a*) Nominal Value	(*b*) $\tilde{H}_0^\dagger$	(*c*) St. Dev.	(*d*) $\tilde{\delta}^\dagger_{\min}$	(*e*) St. Dev.	(*f*) 95%	(*g*) 99%
$H_0 = 0.30$	0.297	0.02493	0.02748*	0.00718	99.5	100.0
$H_0 = 0.40$	0.395	0.02918	0.02756*	0.00781	95.0	99.5
$H_0 = 0.50$	0.500	0.03130	0.03137*	0.00667	95.0	100.0
$H_0 = 0.60$	0.603	0.03392	0.03177*	0.00632	95.0	100.0
$H_0 = 0.70$	0.697	0.03689	0.03304*	0.00595	96.5	100.0

*Not significant at level $\alpha = 0.05$ of Type 1 error.
†Average over 200 samples.

The analysis concerning the statistical significance of the diameter is summarized in Table 1 and in Figs. 4 and 5. In detail, for each set of sampled fBm of nominal parameter displayed in column (*a*), Table 1 reproduces the parameter $\tilde{H}_0$ estimated by averaging 200 estimates [column (*b*)], the standard deviations of the estimates [column (*c*)], the sample diameters $\tilde{\delta}_{\min}$ for which the significance is given at 95% [column (*d*)] and the respective standard deviations [column (*e*)]. Finally, columns (*f*) and (*g*) display the empirical frequencies of acceptance of $\tilde{\delta}_{\min}$, once fixed the nominal level of the test (for $\alpha = 5\%$ and $\alpha = 1\%$). For all the samples, we infer that — as expected — self-similarity cannot be rejected for fBm at a significance level of 5% (or more). This confirms the robustness of the KS statistic also in the case of strongly dependent data, of course provided that the process satisfies the conditions of Proposition 5.

Figures 4 and 5 respectively display the values of the biases $\tilde{H}_0 - H_0$ and the distribution of $\tilde{\delta}_{\min}$ for the different nominal self-similarity parameters. As usual, the two boxplots show lines at the lower quartile (quantile 25%), median and upper quartile (quantile 75%) values. The whiskers are square brackets extending from each end of the box to show

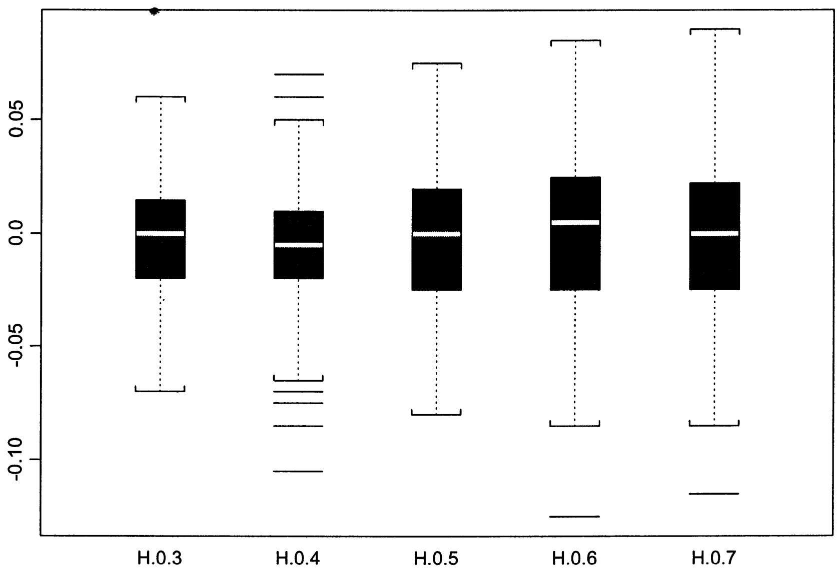

Fig. 4 Boxplot of the bias of the estimated H_0's.

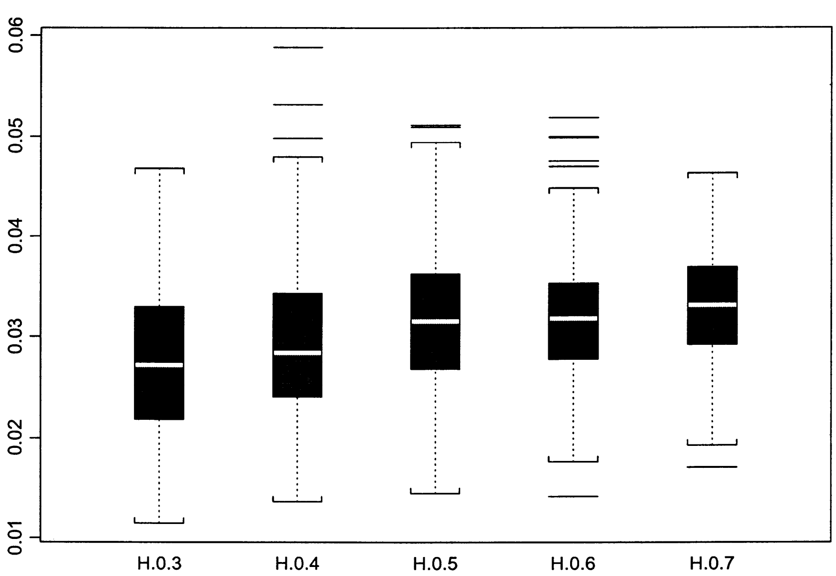

Fig. 5 Boxplot of the estimated diameters for the different H_0's.

the extent of the rest of the values. Outliers are marked by bars.

4.2. Non Self-similar Processes

In this section, we want to compare the estimates of the self-similarity parameter of non self-similar processes obtained with different methods. Even if this is a clear conceptual contradiction, it should be considered that — in studying the scale invariance of an empirical time series — self-similarity is usually assumed but almost never tested in advance. As a consequence, the estimates are correct only if the assumption is correct. Hence, it is a useful exercise

Table 2 Self-similarity analysis for some traces of mBm.

	$H_{\text{min}} = 0.30, H_{\text{max}} = 0.70$				$H_{\text{min}} = 0.35, H_{\text{max}} = 0.65$			
	Function				Function			
	arctg	linear	step	sin	arctg	linear	step	sin
$\delta(50)$	0.490 0.08770*	0.530 0.08269*	0.495 0.07700*	0.475 0.08316*	0.515 0.06640*	0.565 0.08211*	0.535 0.09273*	0.520 0.05685*
$\delta(100)$	0.540 0.11269*	0.580 0.14252*	0.510 0.09215*	0.495 0.10171*	0.550 0.11212*	0.545 0.09704*	0.540 0.10062*	0.515 0.07167*
LR	0.389	0.349	0.332	0.368	0.381	0.432	0.421	0.484
W	0.387	0.392	0.351	0.388	0.406	0.452	0.431	0.456
NC	0.576	0.510	0.529	0.552	0.514	0.545	0.559	0.505
OV	0.403	0.370	0.345	0.382	0.404	0.443	0.434	0.474
GV	0.408	0.370	0.347	0.387	0.407	0.444	0.434	0.471

	$H_{\text{min}} = 0.40, H_{\text{max}} = 0.60$				$H_{\text{min}} = 0.45, H_{\text{max}} = 0.55$			
	Function				Function			
	arctg	linear	step	sin	arctg	linear	step	sin
$\delta(50)$	0.460 0.06209*	0.635 0.16105*	0.570 0.11521*	0.520 0.06842*	0.535 0.04262*	0.500 0.09040*	0.450 0.09216*	0.470 0.11054*
$\delta(100)$	0.465 0.06821*	0.720 0.20011*	0.505 0.15443*	0.515 0.10196*	0.555 0.05930*	0.525 0.09988*	0.460 0.11602*	0.450 0.16229*
LR	0.393	0.458	0.410	0.434	0.515	0.486	0.428	0.433
W	0.430	0.484	0.440	0.457	0.518	0.501	0.456	0.461
NC	0.524	0.512	0.515	0.513	0.535	0.515	0.535	0.546
OV	0.409	0.461	0.420	0.460	0.527	0.481	0.455	0.458
GV	0.414	0.464	0.425	0.459	0.524	0.484	0.453	0.457

*Significant at level $\alpha = 0.05$ of Type 1 error.

to check what happens with different estimators when one tries to calculate the self-similar parameter for a non self-similar process. As examples of non self-similar sequences, we have first considered several realizations of the multifractional Brownian motion (mBm), as defined in Péltier and Lévy Véhel.[25] MBm generalizes fBm to the case where the self-similarity parameter is no longer a constant, but a function of the time index of the process. This extension, resulting in the loss of self-similarity, allows to model non-stationary continuous processes and is particularly useful in the study of financial time series. Samples of mBm of length $n = 2000$ have been simulated with arctangent, linear, step and sinusoidal functional parameter. In order to appreciate the sensitiveness of the estimator, for each simulation the functional parameter has been allowed to vary within the four ranges [0.30, 0.70], [0.35, 0.65], [0.40, 0.60] and [0.45, 0.55].

For the sequences so generated, the self-similarity parameter has been estimated using the following methods, based on different approaches (see, e.g. Refs. 22 and 23 for their description):

- distributional approach: our diameter-based estimator, with maximum lag $\mathfrak{A} = 50$ and $\mathfrak{A} = 100$;
- spectral approach: the periodogram estimation in the variant of Lobato and Robinson[26] (LR);
- maximum likelihood approach: Whittle's estimator[27] (W); and
- temporal approach: number of level crossings[28] (NC) and discrete variations, both with ordinary (OV) and generalized least squares (GV).[29]

Table 2 shows the results of the analysis.

Table 3 Self-similarity analysis of S_n.

	$\delta(50)$	LR	W	NC	OV	GV
$\tilde{H}_0^{\dagger}$	0.497 0.0864*	0.499	0.499	0.519	0.500	0.500
St. dev.	0.0247	0.0209	0.0140	0.0151	0.0165	0.0158

*All the sequences are significant at level $\alpha = 0.01$ (or less) of Type 1 error.
†Average over 100 samples.

In all cases a (senseless) self-similarity parameter is estimated but there is a conclusive improvement using the diameter: it is always statistically significant at level $\alpha = 5\%$ (the only exception being the $\delta(50)$ estimated for the arctangent functional parameter with $H_{\min} = 0.45$ and $H_{\max} = 0.55$) and therefore at 95% we can reject self-similarity, even if a parameter has been calculated in some ways. It is clear that this is a mere exercise, since applying the above estimators to mBm is conceptually incorrect. More, in the case of the diameter the time series is not ergodic, so no justification can be given for the use of KS statistic.

More persuasive is therefore the same analysis performed with respect to some ergodic, non self-similar stochastic process, such as the sequences of independent and identically distributed (neither Gaussian nor stable) random variables. To do this, we have generated 100 samples of (2000 each) independent, uniformly distributed random variables and built the process

$$(S_n)_{n=1,\dots,N} = \sum_{j=1}^{n} \varsigma_j \quad S_0 = 0,\, \varsigma_j \sim U\left(-\frac{1}{2}, \frac{1}{2}\right).$$

Clearly the sequence S_n is not self-similar since S_1 (and $S_n - S_{n-1}$) is uniform, S_2 (and $S_n - S_{n-2}$) is triangular and, in general, S_k (and $S_n - S_{n-k}$) tends to the normal distribution as k tends to infinity. In spite of this, if self-similarity is tested only with regard to some moments of the distribution then one concludes that the sequence is self-similar with parameter $H_0 = \frac{1}{2}$ (see Table 3) and — by virtue of the central limit theorem — this result becomes the more likely the larger the lags are taken. This is a consequence of the fact that the moment-based estimators are generally unable to distinguish self-similarity from second order self-similarity (or from asymptotic self-similarity). The diameter $\delta(50)$ is on the contrary statistically significant for all the generated sequences, even at the probability level of 1%: the mean diameter is in fact 0.0864 (the standard deviation is 0.0245), whereas the threshold is $z_{0.01} = 0.0518$.

5. CONCLUSIONS

In this paper a new method has been proposed that both tests scale invariance and estimates the self-similarity parameter. Our method has two main remarkable features: being distribution-based but distribution-free, it is robust with respect to infinite-moments self-similar stochastic processes and, when the one-dimensional distributions are considered, the significance of its estimates can be evaluated using a standard inferential tool such as the Kolmogorov-Smirnov goodness-of-fit test, even in the case of strongly dependent ergodic sequences.

ACKNOWLEDGMENTS

The author thanks the anonymous referees for their comments and suggestions which helped improving the presentation of this paper.

REFERENCES

1. B. B. Mandelbrot, Une classe de processus stochastiques homothétiques à soi; application à la loi climatologique de H.E. Hurst, *Comptes Rendus de l'Académie des Sciences de Paris* **260** (1965) 3274–3277.
2. B. B. Mandelbrot, Sporadic random functions and conditional spectral analysis: self-similar examples and limits, *Proc. 5th Berkeley Symp. Math. Stat. Prob.* **3** (1967) 155–179.
3. B. B. Mandelbrot and J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, *SIAM Rev.* **10** (1968) 422–437.
4. K. Berkner and C. J. G. Evertsz, Large deviations and self-similarity analysis of graphs: DAX stock prices, *Chaos Solitons Fractals* **6** (1995) 121–130.

5. C. J. G. Evertsz, Self-similarity of high-frequency USD-DEM exchange rates, in *Proceedings of the First International Conference on High Frequency Data in Finance* (Zurich, Switzerland, 1995).
6. Y. Kumagai, Fractal structure of financial high frequency data, *Fractals* **10**(1) (2002) 13–18.
7. U. A. Müller, M. M. Dacorogna, R. B. Olsen, O. V. Pictet, M. Schwarz and C. Morgenegg, Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis, *J. Bank Fin.* **14** (1990) 1189–1208.
8. U. A. Müller, M. M. Dacorogna, R. D. Davé, R. B. Olsen, O. V. Pictet and J. R. Ward, Fractals and intrinsic time; a challenge to econometricians, in *The 34th International Conference of the Applied Econometrics Association*, Luxembourg, 1993.
9. A. Razdan, Scaling in the Bombay Stock Exchange Index, *Pramana* **58**(3) (2002) 537–544.
10. B. B. Mandelbrot, *Gaussian Self-Affinity and Fractals* (Springer, New York, 2002).
11. W. Willinger, M. S. Taqqu and A. Erramilli, A bibliographical guide to self-similar traffic and performance modeling for high-speed networks, in *Stochastic Networks*, eds. F. P. Kelly, S. Zachary and I. Ziedins (Oxford University Press, Oxford, 1996) pp. 339–366.
12. P. Doukhan, G. Oppenheim and M. S. Taqqu (eds.), *Long-Range Dependence: Theory and Applications* (Birkhauser, 2002).
13. B. B. Mandelbrot, *Fractals and Scaling in Finance* (Springer, New York, 1997).
14. G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes* (Chapman & Hall, New York, 1994).
15. P. Embrechts and M. Maejima, *Self-similar Processes* (Princeton University Press, Princeton, 2002).
16. M. S. Taqqu, V. Teverovsky and W. Willinger, Estimators for long-range dependence: an empirical study, *Fractals* **3**(4) (1995) 785–798.
17. J. W. Lamperti, Semi-stable stochastic processes, *Trans. Amer. Math. Soc.* **104** (1962) 62–78.
18. N. V. Smirnov, On the estimation of the discrepancy between empirical curves of distribution for two independent samples, *Bull. Math. Univ. Moscow* **2** (1939) 3–14.
19. A. N. Kolmogorov, *Foundations of the Theory of Probability* (Chelsea Publishing Co., New York, 1933).
20. D. A. Darling, The Kolmogorov-Smirnov, Cramér-Von Mises tests, *Ann. Math. Stat.* **28** (1957) 823–838.
21. A. Wood and G. Chan, Simulation of stationary Gaussian processes in $[0, 1]^d$, *J. Comput. Graph. Stat.* **3**(4) (1994) 409–432.
22. J. Bardet, G. Lang, G. Oppenheim, A. Philippe and M. S. Taqqu, Generators of long-range dependent processes: a survey, in *Theory and Application of Long Range Dependence*, eds. M. S. Taqqu, G. Oppenheim and P. Doukhan (Birkhauser, Boston, 2002).
23. J. F. Coeurjolly, *Simulation and Identification of the Fractional Brownian Motion: A Bibliographical and Comparative Study*, available at http://www.jstatsoft.org/v05/i07/simest.pdf (2000).
24. R. Jennane, R. Harba and G. Jacquet, Estimation de la qualité des méthodes de synthèse du mouvement Brownien fractionnaire, *Traitement du Signal* **13**(4) (1996) 289–302.
25. R. F. Péltier and J. Lévy Véhel, *Multifractional Brownian Motion: Definition and Preliminary Results* (Rapport de Recherche INRIA n. 2645, 1995).
26. I. Lobato and P. M. Robinson, Averaged periodogram estimation of long memory, *J. Econom.* **73** (1996) 303–324.
27. J. Beran, *Statistics for Long Memory Processes* (Chapman and Hall, London, 1994).
28. A. Feuerverger, P. Hall and T. A. Wood, Estimation of fractal index and fractal dimension of a Gaussian process by counting the number of level crossings, *J. Times Ser. Anal.* **15**(6) (1994) 587–606.
29. J. F. Coeurjolly, *Identification du Mouvement Brownien Fractionnaire par Variations Discrètes* (Rapport de Recherche 1016-M, IMAG, Grenoble, 1999).
30. I. P. Cornfeld, Y. G. Sinai and S. V. Fomin, *Ergodic Theory* (Springer Verlag, Berlin, 1982)
31. W. Feller, *An Introduction to Probability Theory and its Applications*, 2nd edn. (John Wiley & Sons, New York, 1971).
32. B. V. Gnedenko and V. C. Koroljuk, On the maximal deviation between two empirical distributions, *Doklady Akad. Nauk SSSR* **80** (1951) 525–528.

APPENDIX A: PROOFS

Proof of Proposition 1 (necessity). Using notation of (3) for the H_0-ss process $X(t)$, the following equalities hold

$$\Phi_{a_i^{-H_0}\mathbb{X}(a_i)}(\mathbf{x}) = \Phi_{\mathbb{X}(1)}(\mathbf{x}) \tag{12a}$$

$$\Phi_{a_j^{-H_0}\mathbb{X}(a_j)}(\mathbf{x}) = \Phi_{\mathbb{X}(1)}(\mathbf{x}) \tag{12b}$$

for any $\mathbf{x} \in \mathbb{R}^k$, any a_i, $a_j \in \mathcal{A}$ and any $\mathcal{A} \subset \mathbb{R}^+$.

From (12a) and (12b), it follows that the definition of self-similarity can be equivalently written as

$$\Phi_{a_i^{-H_0}\mathbb{X}(a_i)}(\mathbf{x}) = \Phi_{a_j^{-H_0}\mathbb{X}(a_j)}(\mathbf{x})\,. \tag{13}$$

Hence, by (13) and by the definition of diameter it is

$$\delta^k(\Psi_{H_0}) = \sup_{\mathbf{x}\in\mathbb{R}^k}\sup_{a_i,a_j\in\mathcal{A}} |\Phi_{a_i^{-H_0}\mathbb{X}(a_i)}(\mathbf{x}) - \Phi_{a_j^{-H_0}\mathbb{X}(a_j)}(\mathbf{x})| = 0\,. \tag{14}$$

(sufficiency). The proof directly follows from the condition of self-similarity [equivalent to (1)] given in the form of (13). In fact, $\delta^k(\Psi_{H_0}) = 0$ implies $\Phi_{a_i^{-H_0}\mathbb{X}(a_i)}(\mathbf{x}) - \Phi_{a_j^{-H_0}\mathbb{X}(a_j)}(\mathbf{x}) = 0$ for any $\mathbf{x} \in \mathbb{R}^k$, any a_i, $a_j \in \mathcal{A}$ and any $\mathcal{A} \subset \mathbb{R}^+$, that is (13). □

Proof of Proposition 2. First assume without loss of generality that $\mathfrak{a} = 1$. This is possible because $\mathcal{A}$ is any subset of $\mathbb{R}^+$ such that $\mathfrak{a} = \min(\mathcal{A})$ and $\mathfrak{A} = \max(\mathcal{A})$. Therefore, it is always possible to define the new set $\mathcal{A}^* = \{\frac{a_i}{\mathfrak{a}}\}$ with $a_i \in \mathcal{A}$ and, obviously, $\min(\mathcal{A}^*) = 1$.

Case (a). For each $H' < H'' \le H_0$ we have to prove that $\delta^k(\Psi_{H'}) \ge \delta^k(\Psi_{H''})$ or, similarly,

$$\sup_{\mathbf{x}\in\mathbb{R}^k} |\Phi_{\mathfrak{A}^{-H'}\mathbb{X}(\mathfrak{A})}(\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x})| \ge \sup_{\mathbf{x}\in\mathbb{R}^k} |\Phi_{\mathfrak{A}^{-H''}\mathbb{X}(\mathfrak{A})}(\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x})|\,.$$

Since the process is assumed to be self-similar, equality (4) holds; therefore we can write

$$\sup_{\mathbf{x}\in\mathbb{R}^k} |\Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H'-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x})| \ge \sup_{\mathbf{x}\in\mathbb{R}^k} |\Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H''-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x})|\,. \tag{15}$$

Notice that, when $H < H_0$ the quantity $\Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x})$ is non-positive when $\mathbf{x} \geqq \mathbf{0}$ and non-negative when $\mathbf{x} \leqq \mathbf{0}$. Therefore inequality (15) holds because from $H' < H'' < H_0$ it follows that $\mathfrak{A}^{H'-H_0}\mathbf{x} \leqq \mathfrak{A}^{H''-H_0}\mathbf{x}$ for $\mathbf{x} \geqq \mathbf{0}$ and $\mathfrak{A}^{H'-H_0}\mathbf{x} \geqq \mathfrak{A}^{H''-H_0}\mathbf{x}$ for $\mathbf{x} \leqq \mathbf{0}$, that is — Φ being a distribution law —

$$\begin{cases} \Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H'-H_0}\mathbf{x}) \le \Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H''-H_0}\mathbf{x})\,, & \mathbf{x} \geqq \mathbf{0} \\ \Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H'-H_0}\mathbf{x}) \ge \Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H''-H_0}\mathbf{x})\,, & \mathbf{x} \leqq \mathbf{0} \end{cases} \Longrightarrow$$

$$\begin{cases} \Phi_{\mathbb{X}(1)}(\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H'-H_0}\mathbf{x}) \ge \Phi_{\mathbb{X}(1)}(\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H''-H_0}\mathbf{x})\,, & \mathbf{x} \geqq \mathbf{0} \\ \Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H'-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x}) \ge \Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H''-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x})\,, & \mathbf{x} \leqq \mathbf{0}\,. \end{cases}$$

Observe that both the left-hand and right-hand sides of the two inequalities are positive and hence the two cases can be parsimoniously gathered using the modulus as follows

$$|\Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H'-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x}) \ge |\Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H''-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x})|\,. \tag{16}$$

Inequality (16) holds for any $\mathbf{x} \geqq \mathbf{0}$ or $\mathbf{x} \leqq \mathbf{0}$ and therefore it holds for the supremum, which completes the proof of Case (a).

In the same way we discuss Case (b), for which we assume $H' > H'' \ge H_0$. In this case it will be $\delta^k(\Psi_{H'}) \le \delta^k(\Psi_{H''})$; in fact using the same arguments of Case (a) and recalling that $H' - H_0 > H'' - H_0 \ge 0$, it is easily seen that

$$\sup_{\mathbf{x}\in\mathbb{R}^k} |\Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H'-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x})| \le \sup_{\mathbf{x}\in\mathbb{R}^k} |\Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H''-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathbf{x})|$$

□

Proof of Proposition 3. We prove only case (i) of the Proposition, case (ii) being a trivial consequence of the former. Firstly observe that $X(t)$ being self-similar both (4) and (7) hold; therefore one has

$$\delta^k(\Psi_H) = \sup_{\mathbf{x}\in\mathbb{R}^k} |\Phi_{\mathbb{X}(1)}(\mathfrak{a}^{H-H_0}\mathbf{x})\Phi_{\mathbb{X}(1)}(\mathfrak{A}^{H-H_0}\mathbf{x})|\,.$$

Case (a): $H_0 > H$. Since $\mathfrak{a}_1 \ge \cdots \ge \mathfrak{a}_n$ and $\mathfrak{A}_1 \le \cdots \le \mathfrak{A}_n$, for the infimum it will be

$$\begin{cases} \mathfrak{a}_1^{H-H_0}\mathbf{x} \leqq \cdots \leqq \mathfrak{a}_n^{H-H_0}\mathbf{x}\,, & \mathbf{x} \geqq \mathbf{0} \\ \mathfrak{a}_1^{H-H_0}\mathbf{x} \geqq \cdots \geqq \mathfrak{a}_n^{H-H_0}\mathbf{x}\,, & \mathbf{x} \leqq \mathbf{0} \end{cases} \Longrightarrow$$

$$\begin{cases} \Phi_{\mathbb{X}(1)}(\mathfrak{a}_1^{H-H_0}\mathbf{x}) \leq \cdots \leq \Phi_{\mathbb{X}(1)}(\mathfrak{a}_n^{H-H_0}\mathbf{x}), & \mathbf{x} \geqq \mathbf{0} \\ -\Phi_{\mathbb{X}(1)}(\mathfrak{a}_1^{H-H_0}\mathbf{x}) \leq \cdots \leq -\Phi_{\mathbb{X}(1)}(\mathfrak{a}_n^{H-H_0}\mathbf{x}), & \mathbf{x} \leqq \mathbf{0} \end{cases} \tag{17}$$

and for the supremum it will be

$$\begin{cases} \mathfrak{A}_1^{H-H_0}\mathbf{x} \geqq \cdots \leqq \mathfrak{A}_n^{H-H_0}\mathbf{x}, & \mathbf{x} \geqq \mathbf{0} \\ \mathfrak{A}_1^{H-H_0}\mathbf{x} \leqq \cdots \leqq \mathfrak{A}_n^{H-H_0}\mathbf{x}, & \mathbf{x} \leqq \mathbf{0} \end{cases} \Longrightarrow$$

$$\begin{cases} -\Phi_{\mathbb{X}(1)}(\mathfrak{A}_1^{H-H_0}\mathbf{x}) \leq \cdots \leq -\Phi_{\mathbb{X}(1)}(\mathfrak{A}_n^{H-H_0}\mathbf{x}), & \mathbf{x} \geqq \mathbf{0} \\ \Phi_{\mathbb{X}(1)}(\mathfrak{A}_1^{H-H_0}\mathbf{x}) \leq \cdots \leq \Phi_{\mathbb{X}(1)}(\mathfrak{A}_n^{H-H_0}\mathbf{x}), & \mathbf{x} \leqq \mathbf{0}. \end{cases} \tag{18}$$

Adding up (17) and (18) and using the compact notation of modulus

$$|\Phi_{\mathbb{X}(1)}(\mathfrak{a}_1^{H-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathfrak{A}_1^{H-H_0}\mathbf{x})| \leq \cdots \leq |\Phi_{\mathbb{X}(1)}(\mathfrak{a}_n^{H-H_0}\mathbf{x}) - \Phi_{\mathbb{X}(1)}(\mathfrak{A}_n^{H-H_0}\mathbf{x})|. \tag{19}$$

Since (19) holds for each non-negative (non-positive) $\mathbf{x}$ it holds for the supremum too, which completes the proof of Case (a).

Case (b): $H_0 < H$. The proof is omitted since it uses the same argument of Case (a). □

Proof of Proposition 4. When $X(t)$ is the fractional Brownian motion (fBm) of parameter H_0, in notation $B_{H_0}(t)$, one has

$$\{a^{-H}(B_{H_0}(t+a) - B_{H_0}(t))\} \stackrel{d}{=} \mathcal{N}(0, a^{2(H_0-H)}\sigma^2) \tag{20}$$

where $\sigma^2 = \mathbb{E}((B_{H_0}(t+1) - B_{H_0}(t))^2)$. Assuming $k = 1$, due to the stationarity of the increments, it is $B_{H_0}(t+a) - B_{H_0}(t) \stackrel{d}{=} B_{H_0}(a)$ and hence, by the same argument used for the general case in Step 1 of the proof of Proposition 5, (4) can be written in terms of the increments of the fBm, i.e.

$$\begin{aligned} \Phi_{a^{-H}B_{H_0}(a)}(x) &= \Pr(a^{-H}B_{H_0}(a) < x) \\ &= \Pr(a^{-H}(B_{H_0}(t+a) - B_{H_0}(t)) < x) \\ &= \text{combining with (20)} \\ &= \frac{1}{a^{H_0-H}\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{u^2}{2a^{2(H_0-H)}\sigma^2}\right) du \\ &= \text{(by self-similarity)} \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{xa^{H-H_0}} \exp\left(-\frac{u^2}{2\sigma^2}\right) du. \end{aligned} \tag{21}$$

Written as (21), the diameter becomes

$$\delta^1(\Psi_H) = \sup_{x\in\mathbb{R}} \sup_{a_i,a_j\in\mathcal{A}} \left| \int_{\frac{x}{a_j^{(H_o-H)}}}^{\frac{x}{a_i^{(H_o-H)}}} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{u^2}{2\sigma^2}\right) du \right|.$$

Since the integrand is a positive function, one has the supremum with respect to a taking the maximum interval of integration, that is, for $a_i = \mathfrak{a}$ ($a_i = \mathfrak{A}$) and $a_j = \mathfrak{A}$ ($a_j = \mathfrak{a}$). Hence

$$\delta^1(\Psi_H) = \sup_{x\in\mathbb{R}} \left| \int_{\frac{x}{\mathfrak{A}^{H_o-H}}}^{\frac{x}{\mathfrak{a}^{H_o-H}}} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{u^2}{2\sigma^2}\right) du \right|. \tag{22}$$

Maximizing the integral function (22), one has

$$\frac{1}{\mathfrak{a}^{(H_o-H)}\sigma\sqrt{2\pi}}\exp\left(-\frac{x^2}{2\mathfrak{a}^{2(H_o-H)}\sigma^2}\right)-\frac{1}{\mathfrak{A}^{(H_o-H)}\sigma\sqrt{2\pi}}\exp\left(-\frac{x^2}{2\mathfrak{A}^{2(H_o-H)}\sigma^2}\right)=0$$

from which trivially follows

$$x=\pm\sqrt{\frac{2(\mathfrak{a}\mathfrak{A})^{2(H_o-H)}\sigma^2}{\mathfrak{A}^{2(H_o-H)}-\mathfrak{a}^{2(H_o-H)}}(H_0-H)\ln\left(\frac{\mathfrak{A}}{\mathfrak{a}}\right)}\,.$$

The theoretical one-dimensional diameter for a fractional Brownian motion with parameter H_0 is therefore independent on σ^2 and equals

$$\delta^1(\Psi_H)=\begin{cases}\dfrac{1}{\sqrt{2\pi}}\left|\displaystyle\int_{-\hat{x}\mathfrak{a}^{H_o-H}}^{-\hat{x}\mathfrak{A}^{H_o-H}}\exp\left(-\frac{u^2}{2}\right)du\right|, & H\neq H_0\\ 0\,, & H=H_0\,.\end{cases}$$

where $\hat{x}=\sqrt{\frac{2(H_0-H)}{\mathfrak{A}^{2(H_o-H)}-\mathfrak{a}^{2(H_o-H)}}\ln\frac{\mathfrak{A}}{\mathfrak{a}}}$. □

Proof of Proposition 5. The proof ensues combining the following three lemmas and steps. In particular, Step 1 reduces the analysis of the distribution of $\tilde{\mathbb{X}}(a)$ to the distribution of $\tilde{\mathbb{Y}}(a)$. The extension of the KS statistic to ergodic (dependent) random variables as deduced in Step 2 is a direct consequence of Lemmas 2 and 3. The convergence of the diameter of a self-similar process to $\mathcal{L}(z)$ is justified in Step 3. □

Lemma 1 (Smirnov). Let $Y_1, Y_2, \ldots, Y_n$ be n mutually independent continuous random variables with common distribution function $F_{Y_j}(x)=F(x)$ for all j and let $\hat{F}_n(x)$ be the sample cumulative distribution function drawn by the Y_j's. Then

$$\lim_{n\to\infty}\Pr\left(\sup_x|\hat{F}_n(x)-F(x)|\le\frac{z}{\sqrt{n}}\right)=\mathcal{L}(z):=1-2\sum_{j=1}^{\infty}(-1)^{j-1}\exp(-2j^2z^2)\,.\quad\square$$

Lemma 2 (Glivenko-Cantelli). Let (Y_k) be a stationary ergodic sequence defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ with common distribution function $F_{Y_j}(x)=F(x)$ and empirical distribution function $\hat{F}_n(x)$ at stage n. Then, with probability one $\sup_{x\in\mathbb{R}}|\hat{F}_n(x)-F(x)|\to 0$. □

Lemma 3. The sequence of the increments $B_{H_0}(t+1)-B_{H_0}(t)$ of an fBm is ergodic (see e.g. Ref. 30, Theorem 14.2.1). □

Proof of Step 1. Notice that

$$|\tilde{\Phi}_{n,\tilde{\mathbb{X}}(1)}(x)-\tilde{\Phi}_{n,\tilde{\mathbb{X}}(a)}(x)|\overset{d}{=}|\tilde{\Phi}_{n,\tilde{\mathbb{Y}}(1)}(x)-\tilde{\Phi}_{n,\tilde{\mathbb{Y}}(a)}(x)|\,.$$

In fact, self-similarity of $X(t)$ ensues that

$$(\tilde{X}(at_1),\,\tilde{X}(at_2),\ldots,\tilde{X}(at_n))\overset{d}{=}(a^{H_0}\tilde{X}(t_1),a^{H_0}\tilde{X}(t_2),\ldots,a^{H_0}\tilde{X}(t_n))\,.\quad(23)$$

Since, for each positive real t_k, $\tilde{X}(at_k)\overset{d}{=}t_k^{H_0}\tilde{X}(a)$, (23) can be written as

$$(t_1^{H_0}\tilde{X}(a),\,t_2^{H_0}\tilde{X}(a),\ldots,t_n^{H_0}\tilde{X}(a))\overset{d}{=}((at_1)^{H_0}\tilde{X}(1),(at_2)^{H_0}\tilde{X}(1),\ldots,(at_n)^{H_0}\tilde{X}(1))$$

or equivalently as

$$(a^{-H_0}\tilde{X}(a),\,a^{-H_0}\tilde{X}(a),\ldots,a^{-H_0}\tilde{X}(a))\overset{d}{=}(\tilde{X}(1),\,\tilde{X}(1),\ldots,\tilde{X}(1))\,.\quad(24)$$

Due to the stationarity of the increments, it is

$$\tilde{Y}(t_k,a):=\tilde{X}(t_k+a)-\tilde{X}(t_k)\overset{d}{=}\tilde{X}(a)$$

and hence, written in terms of increments, (24) becomes

$$\tilde{\mathbb{Y}}(a)\overset{d}{=}\tilde{\mathbb{Y}}(1)\,.$$

Finally, we conclude that

$$\tilde{\Phi}_{n,\tilde{\mathbb{X}}(1)}(x)=\tilde{\Phi}_{n,\tilde{\mathbb{X}}(a)}(x)\Leftrightarrow\tilde{\Phi}_{n,\tilde{\mathbb{Y}}(1)}(x)=\tilde{\Phi}_{n,\tilde{\mathbb{Y}}(a)}(x)\,.\quad\square$$

Proof of Step 2. Independence of the random variables is required in Lemma 1 to guarantee that $\hat{F}_n(x)\to F(x)$ as n tends to infinity as in classical Bernoulli trials; otherwise, assuming some form of dependence among the Y_j's, $\hat{F}_n(x)$ would approximate a (certain) conditional distribution rather than the unconditional one. In spite of this, when $Y_1, \ldots, Y_n$ forms a stationary ergodic sequence Lemma 2 holds, which generalizes the case of i.i.d. random variables. In this way the convergence to $F(x)$ is ensured as well and this is, for example, the case of fBm, whose increments are stationary and ergodic under time shifts, despite the strong correlations shown when $H > \frac{1}{2}$ (Lemma 3). □

Proof of Step 3. To prove now that the diameter calculated for a (dependent or independent) H_0 self-similar process distributes for $H = H_0$ as the KS statistic, we will refer to the argument given in 1951 by Gnedenko and Koroljuk.[31] The scheme leads to the following conclusion: $\Pr(\sup_{1\le x\le n} |\tilde{\Phi}_{n,\tilde{\mathbb{X}}(1)}(x) - \tilde{\Phi}_{n,\tilde{\mathbb{X}}(a)}(x)| < \frac{z}{\sqrt{n}})$ equals the probability that a symmetric random walk of length $2n$ starting and terminating at the origin does not reach the points $\pm r$, namely the well-known probability of a random walk with absorbing barriers, i.e. $\mathcal{L}(z)$ (see e.g. Ref. 32). To show this, it suffices to consider integral r.

(1) The $2n$ random variables $\tilde{Y}(t_1, 1), \ldots, \tilde{Y}(t_n, 1), a^{-H_0}\tilde{Y}(t_1, a), \ldots, a^{-H_0}\tilde{Y}(t_n, a)$ are sorted in order of increasing magnitude and placed in the vector $\mathbf{y} = (y_s)_{s=1,\ldots,2n}$.

(2) Define the vector $\mathbf{r} = (r_s)_{s=1,\ldots,2n}$ where $r_s = \mathbb{I}_{\tilde{\mathbb{Y}}(1)}(y_s) - \mathbb{I}_{\tilde{\mathbb{Y}}(a)}(y_s)$. So, $r_s = 1$ if the sth element of $\mathbf{y}$ comes from $\tilde{\mathbb{Y}}(1)$ and $r_s = -1$ if the sth element of $\mathbf{y}$ comes from $\tilde{\mathbb{Y}}(a)$.

(3) This is the core of the proof. The vector $\mathbf{r}$ contains n plus ones and n minus ones and — what is here of importance — all $\binom{2n}{n}$ orderings are equally likely. In fact — by definition of self-similarity — since $a^{-H_0}\tilde{Y}(t_k, a) \overset{d}{=} \tilde{Y}(t_k, 1)$ there is no reason for a particular displacement to hold when the elements of $\mathbf{y}$ are ordered. This property is trivially true when conditions of Lemma 1 hold but continues to be true if self-similarity occurs, whether independence is assumed or not.

(4) The elements of $\mathbf{r}$ are therefore in a one-to-one correspondence with the paths of length $2n$ starting and terminating at the origin; this means that if $\sum_{s=1}^{j} r_s = q$ in the first j positions there are $\frac{j+q}{2}$ variables of $\tilde{\mathbb{Y}}(1)$ and $\frac{j-q}{2}$ variables of $\tilde{\mathbb{Y}}(a)$. Hence, there exists a point x such that $\tilde{\Phi}_{\tilde{\mathbb{Y}}(1)}(x) = \frac{j+q}{2n}$ and $\tilde{\Phi}_{\tilde{\mathbb{Y}}(a)}(x) = \frac{j-q}{2n}$ and then $|\tilde{\Phi}_{n,\tilde{\mathbb{Y}}(1)}(x) - \tilde{\Phi}_{n,\tilde{\mathbb{Y}}(a)}(x)| \overset{d}{=} |\tilde{\Phi}_{n,\tilde{\mathbb{X}}(1)}(x) - \tilde{\Phi}_{n,\tilde{\mathbb{X}}(a)}(x)| = \frac{|q|}{n}$. It follows that $\sup_{1\le x\le n} |\tilde{\Phi}_{n,\tilde{\mathbb{X}}(1)}(x) - \tilde{\Phi}_{n,\tilde{\mathbb{X}}(a)}(x)| \ge \frac{|q|}{n}$.

(5) The argument in reverse completes the proof. □