Robustness of the R/S statistic for fractional stable noises *†‡§

Florin Avram
Heriot Watt University

Murad S. Taqqu
Boston University

November 16, 2000

Abstract

The R/S statistic is used to detect long-range dependence in a time series and to estimate its intensity. One of its virtues is robustness against different distributions. We show here that the R/S statistic continues to be robust if the time series is a moving average with long-range dependence with innovations that are in the domain of attraction of an infinite variance stable process.

1 Introduction

The R/S statistic provides one of the oldest techniques for detecting long-range dependence and measuring its intensity. It was discussed by the British physicist H.E. Hurst [10] who studied the levels of the Nile River. It was later used by Mandelbrot and Wallis [15, 17, 16] to analyze long-range dependence in geophysical data through a graphical technique which is also described in Mandelbrot and Taqqu [14] and in the books of Beran [6] and Shiryaev [20]. Its effectiveness is analyzed in Taqqu, Teverovsky and Willinger [23] and Taqqu and Teverovsky [22]. A modified and non-graphical version of the R/S statistic, introduced by Lo [11], is discussed in Teverovsky, Taqqu and Willinger [24] and Willinger, Taqqu and Teverovsky [26]. Bootstrap techniques are presented in Davies and Harte [8] and Hall, Härdle, Kleinow and Schmidt [9].

While the R/S statistic is biased (see Taqqu, Teverovsky and Willinger [23] and Taqqu and Teverovsky [22]), it is often used in practice because it appears robust. We show in this paper that the robustness of the R/S statistic extends to infinite variance moving averages with long-range dependence. The results presented here are part of the first author's Ph.D. thesis (Avram [3]), which was written under the supervision of the second author. Because of numerous requests, we have decided to publish these results.

The R/S statistic applied to a strictly stationary time series $X_i,\ i = 1, 2, \dots$ is the ratio $R(n)/S(n)$, where

$$R(n) = \max_{0\le u\le n}\left[\sum_{i=1}^{u} X_i - \frac{u}{n}\sum_{i=1}^{n} X_i\right] - \min_{0\le u\le n}\left[\sum_{i=1}^{u} X_i - \frac{u}{n}\sum_{i=1}^{n} X_i\right]$$

*The second author was partially supported by the NSF Grants DMS-9404093 and ANI-9805623 at Boston University.

†*AMS Mathematics Subject Classifications (1991).* Primary: 60G18, 62F12; Secondary: 60E07, 60F17.

‡*Keywords and phrases:* self-similar processes, functional limit theorems, stable distributions, weak convergence, Skorohod topology.

§*Running head:* Robustness of the R/S statistic is the "adjusted range" of X_i and where

$$S(n) = \left[1/n \sum_{i=1}^{n} X_i^2 - \left(1/n \sum_{i=1}^{n} X_i \right) \right]^{1/2}$$

is the sample standard deviation. The presence of $S(n)$ in the denominator narrows down the fluctuation of $R(n)$ because $S(n)$ tends to be large when $R(n)$ is large. It also contributes to the "robustness" of the statistic R/S (see below). The objective of the R/S analysis of time series is to determine whether there exists $0 \leq J \leq 1$ and a slowly varying function L so that the ratio

$$\frac{R([nt])}{n^J L(n) S(n)}, \quad 0 \leq t \leq 1,$$

converges weakly as n converges to infinity, and also to estimate J. The exponent J is called the *R/S exponent* or Hurst coefficient after the physicist H.E. Hurst.

The statistic R/S is said to be *robust* if J depends on the exponents characterizing long-range dependence but does not depend otherwise on the underlying distribution of the time series. Formally R/S is robust if

$$J = d + 1/2$$

where $d \in (0, 1/2)$ is a measure of the long-range dependence of the time series X_i. The parameter d is defined below.

We will show that R/S is robust when the X_i are fractional stable noises, i.e. moving averages which are attracted to the fractional stable process with self-similarity parameter $1/\alpha < H < 1$. In that case,

$$d = H - 1/\alpha$$

and

$$J = d + 1/2 = H - 1/\alpha + 1/2. \tag{1.1}$$

This includes the large family of FARIMA time series (see Samorodnitsky and Taqqu [19]). We show robustness also in the case when the moving averages are attracted to the Lévy stable process. In that case $d = 0$ and $J = 1/2$.

Both results have been conjectured by Mandelbrot in an unpublished note. Mandelbrot [13] has also shown that R/S is robust in the finite variance case $\alpha = 2$ when X_i are attracted to any of the following processes $X(t)$: Brownian motion ($H = 1/2,\ d = 1/2 - 1/2 = 0$), fractional Brownian motion, for which $d = H - 1/2$, and also the Hermite processes $X(t)$ which are defined as multiple Wiener-Itô integrals with respect to Brownian motion. The R/S statistic is also robust when the X_i are i.i.d. and attracted to the Lévy stable motion ($H = 1/\alpha,\ d = H - 1/\alpha = 0$). These facts are based on the following lemma:

Lemma 1.1 (Mandelbrot [13]) *Let X_i be a stationary sequence, such that:*

$$\left(\frac{1}{n^{H'} L'(n)} \sum_{i=1}^{[nt]} X_i,\ \frac{1}{n^{H''} L''(n)} \sum_{i=1}^{[nt]} X_i^2 \right) \stackrel{*}{\to} (U(t),\ V(t)) \tag{1.2}$$

where L' and L'' are slowly varying functions and where $\stackrel{*}{\to}$ denotes convergence in some functional sense, strong enough to imply convergence of the $\sup_{0 \leq t \leq T}$ and $\inf_{0 \leq t \leq T}$ functionals.

Then the properly normalized R/S statistic $\frac{R([nt])}{n^J L(n) S(n)}$, $\quad 0 \le t \le 1$, *converges weakly as* $n \to \infty$, *with an R/S exponent given by the formula:*

$$J = H' - (1/2)H'' + 1/2. \tag{1.3}$$

Our main results, stated in Section 2 and proved in the subsequent sections, show that Mandelbrot's condition (1.2) holds for some types of moving average processes with infinite variance innovations. We also identify the limit.

2 Main results

We now turn to the precise formulation of our results. Consider the moving averages:

$$Y_k = \sum_{i=-\infty}^{\infty} c_{k-i} X_i \tag{2.1}$$

where X_i are i.i.d. and belong to $\mathcal{D}(\alpha)$, $0 < \alpha < 2$, that is,

$$P[|X_i| > x] = x^{-\alpha} H(x), \quad \lim_{x\to\infty} \frac{P[X_i > x]}{P[X_i < -x]} = \frac{a_+}{a_-}, \tag{2.2}$$

with $a_\pm \ge 0, a_+ + a_- > 0$, and $H(x)$ is a slowly varying function (s.v.f.), and $EX_i = 0$ (when it exists).

The moving average Y_k is well-defined if there exists ν, $0 < \nu < \alpha$, such that

$$\sum_{i=-\infty}^{\infty} |c_i|^\nu < \infty, \tag{2.3}$$

since by the Bahr-Esseen inequality (see Lemma 4.1 below), $E|Y_k|^\nu \le 2E|X_0|^\nu \sum_{i=-\infty}^{\infty} |c_i|^\nu$, and thus, when (2.3) holds, Y_k is well defined in the L_ν sense.

Our first result is that (1.2) holds under either one of the two following conditions (A) and (B). Condition (A) involves $d > 0$ in (1.1), and hence, the resulting exponent J satisfies $J > 1/2$. Condition (B) corresponds to $d = 0$ and hence to $J = 1/2$.

Condition (A)

Suppose $1 < \alpha < 2$ and

$$c_n = \begin{cases} c_+ n^{d-1} L(n) & \text{if } n \ge 0, \\ c_- |n|^{d-1} L(|n|) & \text{if } n < 0, \end{cases} \tag{2.4}$$

where

$$0 < d < 1 - 1/\alpha, \tag{2.5}$$

$L(n)$ is a s.v.f., and c_+ and c_- are nonnegative and not both equal to 0. The exponent d is a measure of "long-range dependence."

Observe that in this case, (2.3) is satisfied. Indeed, note that the condition (2.5) implies $(d-1)\alpha < -1$; but then we may choose ν strictly less than α but sufficiently close to it that it satisfies $(d-1)\nu < -1$, which is precisely the condition for the sequence c_n^ν to be summable.

Condition (B)

i) If $0 < \alpha \leq 1$, suppose that $c_i \geq 0$, and that $\sum_{i=-\infty}^{\infty} c_i^\nu < \infty$ for some $\nu < \alpha$.

ii) If $1 < \alpha < 2$, suppose that $c_i \geq 0$, $\{c_i,\ i \geq 0\}$ and $\{c_i,\ i < 0\}$ are monotone sequences, and that $\sum_{i=-\infty}^{\infty} c_i^\nu < \infty$ for some $\nu < \alpha$ and $\nu < 1$.

Instead of supposing that $\{c_i,\ i \geq 0\}$ and $\{c_i,\ i < 0\}$ are monotone sequences in ii), one can suppose more generally that the technical condition (T.C.) in Avram and Taqqu [5], page 486, is satisfied.

Even though there is no parameter d in Condition (B), it is customary in the statistical literature to say that it corresponds to the case $d = 0$ (no long-range dependence).

Before stating our results, we introduce the normalization factor $a_n = \inf\left[y\colon P\{|X_i| > y\} \leq n^{-1}\right]$. It is well-known that one gets from (2.2) that $a_n = n^{1/\alpha} H_\alpha^{1/\alpha}(n)$, where H_α denotes the unique s.v.f. that satisfies as $n \to \infty$, $H\left[n^{1/\alpha} H_\alpha^{1/\alpha}(n)\right] \sim H(n)$. By the classical central limit theorem, $a_n^{-1} \sum_{i=1}^{[nt]} X_i$ converges weakly in the J_1-Skorohod topology to a non-trivial α-stable Lévy process. We denote this process by $X_\alpha(t)$. Let $[X_\alpha](t)$ denote its square variation, that is

$$[X_\alpha](t) = \lim_{n\to\infty} \sum_{k=0}^{[nt]} \left[X_\alpha\left[\frac{k+1}{n}\right] - X_\alpha\left[\frac{k}{n}\right]\right]^2,$$

where the convergence is in the sense of the finite-dimensional distributions. $[X_\alpha](t)$ is a stable subordinator of order $\alpha/2$.

The following theorem and subsequent one involve weak convergence in various topologies. These are defined in Section 3.

Theorem 2.1 *(Case A) Let Y_k be given by* (2.1)*, with X_i zero mean random variables in $\mathcal{D}(\alpha)$, $1 < \alpha < 2$, and c_i given by* (2.4)*, with $0 < d < 1 - 1/\alpha$, and $L(n)$ a s.v.f. Then,*

$$\left(\frac{1}{b_n}\sum_{k=1}^{[nt]} Y_k,\ \frac{1}{a_n^2}\sum_{k=1}^{[nt]} Y_k^2\right) \overset{M_1}{\to} \left(X_{\alpha,d}(t),\ (\sum_{i=-\infty}^{\infty} c_i^2)[X_\alpha](t)\right) \tag{2.6}$$

where $\overset{M_1}{\to}$ denotes the weak limit with respect to the M_1 two-dimensional Skorohod topology. Here

$$b_n = n^d L(n) a_n, \tag{2.7}$$

and

$$X_{\alpha,d}(t) = \int_{-\infty}^{\infty} k_t(s)\, dX_\alpha(s). \tag{2.8}$$

The kernel $k_t(s)$ is given by

$$k_t(s) = \int_0^t q(u-s)\,du \quad \text{with} \quad q(x) = c_+(x_+)^{d-1} + c_-(x_-)^{d-1}.$$

Remark. The convergence of the first coordinate has been established by Astrauskas [2] in the case $c_- = 0$, and by Maejima [12] in the case $c_+ = c_- = 1$. The finite-dimensional distribution (f.d.d.) convergence of the second coordinate has been established by Astrauskas [1]. The process $X_{\alpha,d}$ is the linear fractional stable motion with self-similarity parameter $H = d + 1/\alpha$ (see Samorodnitsky and Taqqu [19]).

We consider now case (B) in which Y_k are attracted to a Lévy stable motion and we have

Theorem 2.2 *(Case B) Let Y_k be given by (2.1), (2.2) and assume $c_i \geq 0$, $\sum_{i=-\infty}^{\infty} c_i^{\nu} < \infty$, $\nu < \alpha$, and $\nu < 1$. Then,*

$$\left(\frac{1}{a_n}\sum_{k=1}^{[nt]} Y_k,\ \frac{1}{a_n^2}\sum_{k=1}^{[nt]} Y_k^2\right) \overset{M_1^*}{\to} \left((\sum_{i=-\infty}^{\infty} c_i)X_\alpha(t),\ (\sum_{i=-\infty}^{\infty} c_i^2)[X_\alpha](t)\right) \tag{2.9}$$

where $\overset{M_1^*}{\to}$ denotes weak convergence with respect to the product of the one-dimensional M_1 topologies.

Note that here both components of the limiting process are discontinuous, in contrast to the situation considered in Theorem 2.1 and in Proposition 4.3 below. We were not able to extend in this case the f.d.d. convergence to weak convergence in any of the two-dimensional Skorohod topologies. In fact, we conjecture that the convergence in (2.9) does not hold in any of the two-dimensional Skorohod topologies.

Theorems 2.1 and 2.2 can now be used to obtain the asymptotic behavior of the R/S statistic. In the setup of Theorem 2.1, where $d > 0$, we obtain:

Theorem 2.3 *When Y_k are as in Theorem 2.1, and $R(n), S(n)$ are given by:*

$$R(n) = \max_{0\leq t\leq 1}\left[\sum_{i=1}^{[nt]} Y_i - t\sum_{i=1}^{n} Y_i\right] - \min_{0\leq t\leq 1}\left[\sum_{i=1}^{[nt]} Y_i - t\sum_{i=1}^{n} Y_i\right]$$

$$S(n) = \left[\frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^2\right]^{1/2}$$

then

$$\lim_{n\to\infty}\frac{R(n)}{n^J L(n) S(n)} = \frac{\max_{0\leq t\leq 1}[X_{\alpha,d}(t) - tX_{\alpha,d}(1)] - \min_{0\leq t\leq 1}[X_{\alpha,d}(t) - tX_{\alpha,d}(1)]}{\left[(\sum_{i=-\infty}^{\infty} c_i^2)[X_\alpha](1)\right]^{1/2}}$$

where

$$J = d + 1/2.$$

When the Y_k are attracted to a Lévy stable motion, we get:

Theorem 2.4 *Under the conditions of Theorem 2.2,*

$$\lim_{n\to\infty}\frac{R(n)}{n^J L(n) S(n)} \overset{d}{=} \frac{\max_{0\leq t\leq 1}[X_\alpha(t) - tX_\alpha(t)] - \min_{0\leq t\leq 1}[X_\alpha(t) - tX_\alpha(1)]}{\sqrt{[X_\alpha](1)}} \cdot \frac{\sum_{i=-\infty}^{\infty} c_i}{\sqrt{\sum_{i=-\infty}^{\infty} c_i^2}},$$

with

$$J = 1/2.$$

3 The Skorohod topologies

Consider the space $D_{[0,1]}$ of functions in $\mathbb{R}^m$, $m \geq 1$, on $[0,1]$, with left limits and continuous to the right (and to the left at $t=1$). Skorohod introduced four topologies in $D_{[0,1]}$, denoted J_1, J_2, M_1 and M_2 in addition to the uniform topology U. Although J_1 is the commonly used Skorohod topology, there are many situations where J_1 must be replaced by a weaker topology such as M_1. Avram and Taqqu [5], for example, show that normalized sums of moving averages of order $1 < q < \infty$ that converge in finite-dimensional distributions to a stable-Lévy process, do not converge in the J_1 sense; however, they can converge in the M_1 sense. For other instances of M_1 convergence but no J_1 convergence, see Whitt [25]. Having M_1 convergence is useful, because the M_1 topology is strong enough to make continuous the widely used functions $\max_{0 \leq t \leq 1}$ and $\min_{0 \leq t \leq 1}$.

The Skorohod topologies differ in the way convergent sequences of deterministic functions A_n approach their limit A in the neighborhood jump of A. In the case of the J_1 topology, A_n must have a single jump around a jump of A close to the jump of A in location and magnitude. In the case of the M_2 topology, several jumps are allowed, but the extended graph of A_n (graph + vertical segments) must be close to that of A. In the case of the M_1 topology, several jumps are allowed, but the graph of A_n must be, within ϵ, a "monotone staircase," which gets "compressed" into a single jump of A as $n \to \infty$. For illustration see the figure in Avram and Taqqu [5], page 484.

To define these topologies, consider 3 points x_1, x_2 and x in $\mathbb{R}^m$ and let

$$\begin{aligned} J(x_1, x, x_2) &= \text{the smallest distance from } x \text{ to the two-point set } \{x_1, x_2\} \\ &= \min(|x - x_1|,\ |x - x_2|), \end{aligned}$$

where $|\ |$ denotes the Euclidian norm in $\mathbb{R}^m$, and let

$$\begin{aligned} M(x_1, x, x_2) &= \text{the distance from } x \text{ to } [x_1, x_2] \\ &= \inf_{0 \leq \lambda \leq 1} |\lambda x_1 + (1-\lambda) x_2 - x|. \end{aligned} \tag{3.1}$$

Given a function $\{A(t),\ 0 \leq t \leq 1\}$ in $D_{[0,1]}$, define its corresponding M_1-oscillation

$$\omega_\delta^{M_1}(A) = \sup_{\substack{0 \leq t - t_1 \leq \delta \\ 0 \leq t_2 - t \leq \delta}} M[A(t_1), A(t), A(t_2)] \tag{3.2}$$

and its U-oscillation as

$$\omega_\delta^{U}(A) = \sup_{0 \leq |t-s| \leq \delta} |A(t) - A(s)|$$

The $\omega_\delta^{J_1}$ oscillation is obtained by replacing M by J in (3.2). The J_2 and M_2 oscillations are obtained by taking the sup over $\delta/2 \leq t - t_1 \leq \delta$ and $\delta/2 \leq t_2 - t \leq \delta$.

Let $\{A_n(t),\ 0 \leq t \leq 1\}$ be a sequence of functions in $D_{[0,1]}$ and suppose there exists a function A in $D_{[0,1]}$ and a sequence t_i, dense in $[0,1]$ and including 0 and 1, such that $\lim_{n\to\infty} A_n(t_i) = A(t_i)$ for all i. Let T stand for either U, M_1 (or J_1, J_2, M_2) topologies. We say that A_n converges to A in the T topology if

$$\lim_{\delta \to 0} \overline{\lim_{n \to \infty}} \omega_\delta^T(A_n) = 0.$$

Noting that $M(x_1, x, x_2) \leq J(x_1, x, x_2)$, that $J(x_1, x, x_1) = |x - x_1|$ and that $\omega_\delta^{M_2} \leq \omega_\delta^{M_1}$, $\quad \omega_\delta^{J_2} \leq \omega_\delta^{J_1}$, we find that

$$\omega_\delta^{M_2} \leq \omega_\delta^{M_1} \leq \omega_\delta^{J_1} \leq \omega_\delta^{U}, \quad \omega_\delta^{M_2} \leq \omega_\delta^{J_2} \leq \omega_\delta^{J_1} \leq \omega_\delta^{U}.$$

Hence, the topologies are related as follows:

$$U \Rightarrow J_1 \Rightarrow J_2 \Rightarrow M_2 \quad \text{and} \quad U \Rightarrow J_1 \Rightarrow M_1 \Rightarrow M_2.$$

We use the notation M_1^* to denote the *product topology* on $D_{[0,1]}^2 = D_{[0,1]} \times D_{[0,1]}$ induced by the M_1-topologies on $D_{[0,1]}$.

4 Proof of Theorem 2.1

The proof of Theorem 2.1 will be divided into a number of propositions. We will use U-lim, M_1-lim and M_1^*-lim to denote the weak convergence of random functions in the U, M_1 and M_1^* topology respectively.

Proposition 4.1 *Theorem 2.1 holds if, in its conclusion, we replace weak convergence in the M_1 sense by the convergence of the finite-dimensional distributions.*

Proposition 4.2 *If the conditions of Theorem 2.1 hold, then:*

a)
$$U - \lim_{n\to\infty} b_n^{-1} \sum_{i=1}^{[nt]} Y_i = X_{\alpha,d}(t)$$

b)
$$M_1 - \lim_{n\to\infty} a_n^{-2} \sum_{i=1}^{[nt]} Y_i^2 = (\sum_i c_i^2)[X_\alpha](t)$$

Proposition 4.3 *Let $A_n(t)$, $B_n(t)$ be a pair of processes in $D_{[0,1]}$, such that (i)* f.d.d. $-\lim_{n\to\infty}(A_n(t), B_n(t)) = (A(t), B(t))$, *(ii)* $M_1 - \lim_{n\to\infty} A_n = A$, *(iii)* $U - \lim_{n\to\infty} B_n = B$. *Then* $M_1 - \lim_{n\to\infty}(A_n, B_n) = (A, B)$.

Hence, Theorem 2.1 follows from Proposition 4.1, 4.2 and 4.3, applied to:

$$A_n(t) = a_n^{-2} \sum_{i=1}^{[nt]} Y_i^2 \quad \text{and} \quad B_n(t) = b_n^{-1} \sum_{i=1}^{[nt]} Y_i.$$

4.1 Proof of Proposition 4.1

To obtain the result, which is stated without proof in Astrauskas [1], we decompose Y_k^2 as $Y_k^2 = W_k + Z_k$, where

(4.1)
$$W_k = \sum_i c_{k-i}^2 X_i^2$$

is a moving average, and

(4.2)
$$Z_k = \sum_{i \neq j} c_{k-i} c_{k-j} X_i X_j$$

is a *quadratic* moving average.

Note that W_k are well defined as moving averages of random variables in $\mathcal{D}(\alpha/2)$, since if we let $\nu' = \nu/2$, we have $\nu' < \alpha/2$ and $\sum_i |c_i^2|^{\nu'} < \infty$. The Z_k's are also well defined (see Lemma 4.2 below).

We will show below that the order of magnitude of the sum of the quadratic moving averages Z_k is negligible when compared to the sum of W_k. We shall need a moment inequality of the Bahr-Esseen type, namely

Lemma 4.1 *If $1 < \nu \leq 2$ and W_i, $i = 1, \ldots, m$ satisfies $E[W_k | W_1 + \ldots + W_{k-1}] = 0$ for $k = 2, \ldots, m$, or if $0 < \nu \leq 1$ and $\{W_i\}$ is any sequence of random variables, then $E|\sum_{i=1}^m W_i|^\nu \leq 2\sum_{i=1}^m E|W_i|^\nu$.*

This lemma follows from Chatterji [7], Lemma 1. We need here the following multilinear extension:

Lemma 4.2 *a) Let X_i be an i.i.d. sequence in $\mathcal{D}(\alpha)$, $0 < \alpha < 2$, where, if $\alpha > 1$, we assume $EX_i = 0$. Let $A_n^{(k)} = \{\mathbf{i} \colon \mathbf{i} = (i_1, \ldots, i_k),\ 1 \leq i_j \leq n,\ \forall j\}$ and let $f_{\mathbf{i}}$ denote an "off-diagonal" sequence, i.e. $f_{\mathbf{i}} = 0$ whenever $\mathbf{i}$ has two equal components. Then,*

$$E\Big|\sum_{\mathbf{i} \in A_n^{(k)}} f_{\mathbf{i}} X_{i_1} \ldots X_{i_k}\Big|^\nu \leq M(E|X_i|^\nu)^k \sum_{\mathbf{i} \in A_n^{(k)}} |f_{\mathbf{i}}|^\nu, \tag{4.3}$$

where $M = 2^k (k!)^{\nu - 1}$.

b) In the case $n = \infty$, if $\sum_{\mathbf{i} \in A_\infty^{(k)}} |f_{\mathbf{i}}|^\nu < \infty$ and $E|X_0|^\nu < \infty$, then $\sum_{\mathbf{i} \in A_\infty^{(k)}} f_{\mathbf{i}} X_{i_1} \ldots X_{i_k}$ is well defined in L_ν and (4.3) still holds.

Proof of Lemma 4.2. It is enough to prove Part a). The case when $0 < \nu \leq 1$ follows from the elementary inequality $|\sum_i a_i|^\nu \leq \sum_i |a_i|^\nu$ and the independence of the X_i's. Let now $1 < \nu \leq 2$. Note that is enough to prove:

$$E\Big|\sum_{i_1 < \ldots < i_k} f_{i_1, \ldots, i_k} X_{i_1} \ldots X_{i_k}\Big|^\nu \leq M' \sum_{i_1 < \ldots < i_k} |f_{i_1, \ldots, i_k}|^\nu \prod_{j=1}^k E|X_{i_j}|^\nu, \tag{4.4}$$

where M' depends on k only, since then (4.3) follows, with $M = M'(k!)^{\nu-1}$, by Jensen's inequality.

We use induction on k to prove that (4.4) holds, with:

$$M' = 2^k \tag{4.5}$$

When $k = 1$, this follows from Lemma 4.1. Suppose now that (4.4), (4.5) hold up to $k - 1$, and introduce the random variables

$$A_m = \sum_{i_1 < \ldots < i_k = m} f_{i_1, \ldots, i_{k-1}, m} X_{i_1} \ldots X_m = X_m B_{m-1},$$

where

$$B_{m-1} = \sum_{i_1 < \ldots < i_{k-1} \leq m-1} f_{i_1, \ldots, i_{k-1}, m} X_{i_1} \ldots X_{i_{k-1}}.$$

Since $E\left[A_m | X_1, \ldots, X_{m-1}\right] = B_{m-1} E X_m = 0$, and $A_1, \ldots, A_{m-1}$ are functions of $X_1, \ldots, X_{m-1}$, then $E\left[A_m | A_1 + \ldots A_{m-1}\right] = 0$ and we can apply Lemma 4.1 to get $E|\sum_{m=1}^n A_m|^\nu \leq 2\sum_{m=1}^n E|X_m|^\nu E|B_{m-1}|^\nu$. Since, by the induction hypothesis,

$$E|B_{m-1}|^\nu \leq 2^{k-1} \sum_{i_1 < \ldots < i_{k-1} \leq m-1} |f_{i_1, \ldots, i_{k-1}, m}|^\nu \prod_{j=1}^{k-1} E|X_{i_j}|^\nu,$$

it follows that (4.4) and (4.5) hold for k also. □

We now establish upper bounds for the rate of increase of sums of quadratic moving averages.

Lemma 4.3 *Let Z_k be given by (4.2), with X_i in $\mathcal{D}(\alpha)$, $0 < \alpha < 2$, and $EX_1 = 0$ when $\alpha > 1$. Assume there exists $\nu < \alpha$ such that $\sum_i |c_i|^\nu < \infty$. Then, for every ϵ,*

$$P\left\{ \sup_{0\le t\le 1} \left| \frac{1}{n^{(\overline{\nu}+\eta)/\nu}} \sum_{k=1}^{[nt]} Z_k \right| > \epsilon \right\} \to 0, \tag{4.6}$$

where $\overline{\nu} = \max(\nu, 1)$, and $\eta > 0$ is arbitrarily small.

Proof: Applying Lemma 4.2, we get $E|\sum_{k=1}^m Z_k|^\nu = E|\sum_{i\neq j} X_i X_j \sum_{k=1}^m c_{k-i}c_{k-j}|^\nu \le 2^\nu [E|X_i|^\nu]^2 \sum_{i\neq j} |\sum_{k=1}^m c_{k-i}c_{k-j}|^\nu \le 2^\nu [E|X_i|^\nu]^2 m^{\overline{\nu}} [\sum_i |c_i|^\nu]^2$. The last step is obvious when $\nu \le 1$, and when $\nu > 1$, we have by Jensen's inequality, $|\sum_{k=1}^m c_{k-i}c_{k-j}|^\nu \le m^{\nu-1} \sum_{k=1}^m |c_{k-i}|^\nu |c_{k-j}|^\nu$, and thus, $\sum_{i\neq j} |\sum_{k=1}^m c_{k-i}c_{k-j}|^\nu \le m^\nu [\sum_i |c_i|^\nu]^2$.

Applying now the results of Moricz [18] (Theorem 1, in the case $\nu > 1$, and Theorem 2 in the case $\nu \le 1$), we get $E[\sup_{1\le m\le n} |\sum_{k=1}^m Z_k|]^\nu \le An^{\overline{\nu}}$, where A is a constant independent of n. It follows then, by Chebychev's inequality, that $P\{n^{-(\frac{\overline{\nu}+\eta}{\nu})} \sup_{1\le m\le n} |\sum_{k=1}^m Z_k| > \epsilon\} \le \epsilon^{-\nu} n^{-(\overline{\nu}+\eta)} An^{\overline{\nu}}$ and thus the result holds. □

Remark. Consider the power of n in (4.6). If $\alpha > 1$, ν is close to α, $\bar{\nu} = \nu$, and hence this power can be taken to be $1 + \eta$. If $\alpha \le 1$, then ν is close to α, $\bar{\nu} = 1$ and hence this power can be taken to be $\frac{1}{\alpha} + \eta$. Here $\eta > 0$ is arbitrarily small.

We are now in position to prove Proposition 4.1.

Proof of Proposition 4.1. The normalization for the partial sums of $Y_k^2 = W_k + Z_k$ in (2.6) is a_n^2. Since $\alpha > 1$, we can take the power of n in (4.6) to be $1 + \eta$, and since $a_n = n^{1/\alpha} H_\alpha^{1/\alpha}(n)$ implies $n^{1+\eta} = o(a_n^2)$ for some $\eta > 0$, we can ignore the Z_k term. It is then enough to show that

$$\Big(\frac{1}{b_n} \sum_{k=1}^{[nt]} Y_k,\ \frac{1}{a_n^2} \sum_{k=1}^{[nt]} W_k \Big) \overset{f.d.d.}{\to} \Big(X_{\alpha,d}(t), \Big(\sum_{i=-\infty}^{\infty} c_i^2 \Big) [X_\alpha](t) \Big). \tag{4.7}$$

Let

$$f_t^{(n)}(u) = n^{-d} L^{-1}(n) \sum_{k=1}^{[nt]} c_{k-[nu]} \quad \text{and} \quad g_t^{(n)}(u) = \sum_{k=1}^{[nt]} c_{k-[nu]}^2.$$

If

$$\|f\|_{\alpha,\eta} = \max\left\{ \int_{-\infty}^{\infty} |f(x)|^{\alpha-\eta} dx, \Big(\int_{-\infty}^{\infty} |f(x)|^{\alpha+\eta} dx \Big)^{\frac{\alpha-\eta}{\alpha+\eta}} \right\},$$

then one can show that

$$\|f_t^{(n)}(u) - k_t(u)\|_{\alpha,\eta} \to 0 \quad \text{and} \quad \|g_t^{(n)}(u) - \Big(\sum_{i=-\infty}^{\infty} c_i^2 \Big) 1_{[0,t]}^{(u)}\|_{\alpha/2,\eta} \to 0,$$

so that (4.7) follows from the Scheme of discrete stochastic integrals (see Avram and Taqqu [4], Theorem 4). □

4.2 Proof of Proposition 4.2

a) The f.d.d. convergence (from Proposition 4.1) and a standard argument show that weak convergence in the uniform topology U holds (see for example Maejima [12]).

b) The f.d.d. convergence holds of course by Proposition 4.1, and as in the proof of that proposition, we can ignore the Z_k term in $Y_k^2 = W_k + Z_k$, so that it is enough to show

$$M_1 - \lim_{n\to\infty} a_n^{-2} \sum_{k=1}^{[nt]} W_k = (\sum_{i=-\infty}^{\infty} c_i^2)[X_\alpha](t). \tag{4.8}$$

The convergence in the M_1-Skorohod topology holds trivially since $W_k \geq 0$. $\square$

4.3 Proof of Proposition 4.3

Let $A_n(t)$ and $B_n(t)$ be processes in $D_{[0,1]}$. We first show that

$$\left[\omega_\delta^M(A_n, B_n) > 2\epsilon\right] \subset \left[\omega_\delta^M(A_n) > \epsilon\right] \cup \left[\omega_\delta^M(B_n) > \epsilon\right] \tag{4.9}$$

Indeed, if

$$\sup_{\substack{0\leq t-t_1\leq\delta \\ 0\leq t_2-t\leq\delta}} M\left[A_n(t_1), A_n(t),\ A_n(t_2)\right] < \epsilon, \tag{4.10}$$

and

$$\sup_{0\leq t-s\leq\delta} |B_n(t) - B_n(s)| < \epsilon \tag{4.11}$$

then, for every $t_1 < t < t_2$, with $t - t_1 \leq \delta,\ t_2 - t_1 \leq \delta$, we have:

$$\begin{aligned}
&M[(A_n(t_1), B_n(t_1)),\ (A_n(t), B_n(t)),\ (A_n(t_2), B_n(t_2))] \\
&= \inf_{0\leq\lambda\leq 1} \|\lambda(A_n, B_n)(t_1) + (1-\lambda)(A_n, B_n)(t) - (A_n, B_n)(t)\| \\
&\leq \inf_{0\leq\lambda\leq 1} [|\lambda A_n(t_1) + (1-\lambda)A_n(t_2) - A_n(t)| + |\lambda B_n(t_1) + (1-\lambda)B_n(t_2) - B_n(t)|] \\
&\leq \inf_{0\leq\lambda\leq 1} |\lambda A_n(t_1) + (1-\lambda)A_n(t_2) - A_n(t)| + \lambda\epsilon + (1-\lambda)\epsilon \quad (\text{by}(4.11)) \\
&= M\left[A_n(t_1), A_n(t),\ A_n(t_2)\right] + \epsilon \leq 2\epsilon \quad (\text{by}(4.10)),
\end{aligned}$$

where $\|\ \|$ denotes the Euclidian norm (we used the inequality $\|(x,y)\| \leq |x| + |y|$). So (4.9) holds. It follows now that for every ϵ,

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\left[\omega_\delta^{M_1}(A_n, B_n) > 2\epsilon\right] \leq \lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\left[\omega_\delta^{M_1}(A_n) > \epsilon\right] + \lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\left[\omega_\delta^{U}(B_n) > \epsilon\right] = 0,$$

and thus the result follows, by Theorem 3.2.1 of Skorohod [21]. $\square$

5 Proof of the other theorems

Proof of Theorem 2.3. By Theorem 2.1,

$$M_1 - \lim_{n\to\infty} \Big(\frac{1}{n^d L(n)} \sum_{i=1}^{[nt]} \frac{Y_i}{a_n},\ \sum_{i=1}^{[nt]} \Big[\frac{Y_i}{a_n}\Big]^2\Big) = (X_{\alpha,d}(t),\ s_2[X_\alpha](t)) \tag{5.1}$$

where $s_2 = \sum_{i=-\infty}^{\infty} c_i^2$. Since $L^2(n)/n^{1-2d} \to 0$ as $n \to \infty$, we have

$$U - \lim_{n\to\infty} \Big(\frac{t}{n^d L(n)} \sum_{i=1}^{n} \frac{Y_i}{a_n},\ \frac{L^2(n)}{n^{1-2d}} \Big[\frac{1}{n^d L(n)} \sum_{i=1}^{[nt]} \frac{Y_i}{a_n}\Big]^2\Big) = (tX_{\alpha,d}(1), 0). \tag{5.2}$$

Hence, by (5.1) and (5.2),

$$\begin{aligned} M_1 - \lim_{n\to\infty} \Big(\frac{1}{n^d L(n)} \Big[\sum_{i=1}^{[nt]} \frac{Y_i}{a_n} - t\sum_{i=1}^{n} \frac{Y_i}{a_n}\Big],\ \sum_{i=1}^{[nt]} \frac{Y_i^2}{a_n^2} - \frac{L^2(n)}{n^{1-2d}} \Big[\frac{1}{n^d L(n)} \sum_{i=1}^{[nt]} \frac{Y_i}{a_n}\Big]^2\Big) \\ = (X_{\alpha,d}(t) - tX_{\alpha,d}(1),\ s_2[X_\alpha](t)). \end{aligned} \tag{5.3}$$

The limiting process induces a measure μ on $D^2_{[0,1]}$. Note now that the functional

$$\kappa\colon\ D^2_{[0,1]} \to \mathbb{R}^3, \quad \kappa(f,g) = (\max_{0\le t\le 1} f(t), \min_{0\le t\le 1} f(t), g(1))$$

is continuous a.s. μ, i.e. each component of the $\mathbb{R}^3$ vector $\kappa(f,g)$ is a.s. continuous. Indeed, the continuity of max and min with respect to the one-dimensional M_1 topology was established in Skorohod [21] and the projection at 1 is continuous, since the probability that the limit process has a jump at $t = 1$ is 0.

Furthermore, the function ϕ defined by

$$\phi\colon\ \mathbb{R}^3 \to \mathbb{R}, \quad \phi(x,y,z) = \frac{x-y}{z}$$

is continuous a.s. with respect to the distribution of the $\mathbb{R}^3$-random vector $\kappa(X_{\alpha,d}(t) - tX_{\alpha,d}(1),\ s_2[X_\alpha](t))$, since $P([X_\alpha](1) = 0) = 0$.

Hence, applying first κ and then ϕ to the $D^2_{[0,1]}$-process $(X_{\alpha,d}(t) - tX_{\alpha,d}(1), s_2[X_\alpha](t))$, we get

$$\frac{1}{n^d L(n) a_n} R(n) / \frac{\sqrt{n}}{a_n} S_n \to \frac{\max_{0\le t\le 1}[X_{\alpha,d}(t) - tX_{\alpha,d}(1)] - \min_{0\le t\le 1}[X_{\alpha,d}(t) - tX_{\alpha,d}(1)]}{[s_2[X_\alpha](1)]^{1/2}} \tag{5.4}$$

Thus the normalizing constants in (5.4) are $n^{d+1/2}L(n)$, yielding the result. □

Proof of Theorem 2.2. Weak convergence with respect to the product of the one-dimensional M_1 topologies is equivalent to the f.d.d. convergence, this follows from Theorem 1 of Avram and Taqqu [4]. The M_1 weak convergence of the first component follows from Theorem 2 of Avram and Taqqu [5] (note that the technical condition (T.C.) there is satisfied by Proposition 1(i) of that same paper). Finally, the M_1 weak convergence of the second component may be established just as in case (A). □

Proof of Theorem 2.4. The proof is similar to that of Theorem 2.3 but uses Theorem 2.2. Note that the κ function used in the previous proof is still continuous on $D^2_{[0,1]}$ endowed with the product of the one-dimensional M_1 topologies. □

References

[1] A. Astrauskas. Limit theorems for quadratic forms of linear processes. *Lithuanian Mathematical Journal*, 23(4):355–361, 1983.

[2] A. Astrauskas. Limit theorems for sums of linearly generated random variables. *Lithuanian Mathematical Journal*, 23(2):127–134, 1983.

[3] F. Avram. *Some limit theorems for stationary sequences with infinite or finite variance*. PhD thesis, Cornell University, Ithaca, New York, 1986.

[4] F. Avram and M. S. Taqqu. Weak convergence of moving averages with infinite variance. *Dependence in Probability and Statistics*, pages 399–415, 1986. E. Eberlein and M. S. Taqqu, eds., Boston: Birkhäuser.

[5] F. Avram and M. S. Taqqu. Weak convergence of sums of moving averages in the α-stable domain of attraction. *The Annals of Probability*, 20:483–503, 1992.

[6] J. Beran. *Statistics for Long-Memory Processes*. Chapman & Hall, New York, 1994.

[7] S. D. Chatterji. An L^p-convergence theorem. *The Annals of Mathematical Statistics*, 40:1068–1070, 1969.

[8] R. B. Davies and D. S. Harte. Tests for Hurst effect. *Biometrika*, 74:95–101, 1987.

[9] P. Hall, W. Härdle, T. Kleinow, and P. Schmidt. Semiparametric bookstrap approach to hypothesis tests and confidence intervals for the Hurst coefficient. Preprint, 1999.

[10] H. E. Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–808, 1951.

[11] A. W. Lo. Long-term memory in stock market prices. *Econometrica*, 59:1279–1313, 1991.

[12] M. Maejima. On a class of self-similar processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 62:235–245, 1983.

[13] B. B. Mandelbrot. Limit theorems on the self-normalized range for weakly and strongly dependent processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 31:271–285, 1975.

[14] B. B. Mandelbrot and M. S. Taqqu. Robust R/S analysis of long-run serial correlation. In *Proceedings of the 42nd Session of the International Statistical Institute*, Manila, 1979. Bulletin of the International Statistical Institute. Vol.48, Book 2, pp. 69-104.

[15] B .B. Mandelbrot and J. R. Wallis. Computer experiments with fractional Gaussian noises, Parts 1,2,3. *Water Resources Research*, 5:228–267, 1969.

[16] B. B. Mandelbrot and J. R. Wallis. Robustness of the rescaled range R/S in the measurement of noncyclic long-run statistical dependence. *Water Resources Research*, 5:967–988, 1969.

[17] B. B. Mandelbrot and J. R. Wallis. Some long-run properties of geophysical records. *Water Resources Research*, 5:321–340, 1969.

[18] F. Moricz. Moment inequalities and the strong laws of large numbers. *Zeitschrift für Wahrscheinlichskeitstheorie und verwandte Gebiete*, 35:299–314, 1976.

[19] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Processes: Stochastic Models with Infinite Variance*. Chapman and Hall, New York, London, 1994.

[20] A. N. Shiryaev. *Essentials of Stochastic Finance: Facts, Models, Theory*. World Scientific, Singapore, 1999.

[21] A. V. Skorohod. Limit theorems for stochastic processes. *Theory of Probability and Its Applications*, 1:261–290, 1956.

[22] M. S. Taqqu and V. Teverovsky. On estimating the intensity of long-range dependence in finite and infinite variance series. In R. Adler, R. Feldman, and M. S. Taqqu, editors, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, pages 177–217, Boston, 1998. Birkhäuser.

[23] M. S. Taqqu, V. Teverovsky, and W. Willinger. Estimators for long-range dependence: an empirical study. *Fractals*, 3(4):785–798, 1995. Reprinted in *Fractal Geometry and Analysis*, C.J.G. Evertsz, H-O Peitgen and R.F. Voss, editors. World Scientific Publishing Co., Singapore, 1996.

[24] V. Teverovsky, M. S. Taqqu, and W. Willinger. A critical look at Lo's modified R/S statisitic. *Journal of Statistical Planning and Inference*, 80:211–227, 1999.

[25] W. Whitt. Some useful functions for functional limit theorems. *Mathematics of Operations Research*, 5:67–85, 1980.

[26] W. Willinger, M. S. Taqqu, and V. Teverovsky. Stock market prices and long-range dependence. *Finance and Stochastics*, 3:1–13, 1999.

Florin Avram, Department of Actuarial Mathematics and Statistics, Heriot Watt University, Edinburg, EH14 4AS, United Kingdom. *Email:* F.Avram@ma.hw.ac.uk

Murad S. Taqqu, Department of Mathematics and Statistics, Boston University, 111 Cummington Street, Boston, MA 02215, USA. *Email:* murad@math.bu.edu