"The Annals of Statistics", Vol. 24 (1996) 1880-1913.

Parameter Estimation for Infinite Variance Fractional ARIMA [*†‡]

Piotr S. Kokoszka
University of Utah

Murad S. Taqqu
Boston University

Revised November 1995

Abstract

Consider the fractional ARIMA time series with innovations that have infinite variance. This is a finite parameter model which exhibits both long-range dependence (long memory) and high variability. We prove the consistency of an estimator of the unknown parameters which is based on the periodogram and derive its asymptotic distribution. This shows that the results of Mikosch, Gadrich, Klüppelberg and Adler (1995) for ARMA time series remain valid for fractional ARIMA with long-range dependence. We also extend the limit theorem for sample autocovariances of infinite variance moving averages developed in Davis and Resnick (1985) to moving averages whose coefficients are not absolutely summable.

1 Introduction and main results

This paper is concerned with the estimation of the parameters of the fractional ARIMA time series $\{X_n\}$ defined by the equations

$$\Phi(B)X_n = \Theta(B)\Delta^{-d}Z_n, \tag{1.1}$$

where the innovations Z_n have *infinite variance* and where d is a positive fractional number. B and Δ denote the backward and differencing operator respectively. Because of the presence of the fractional d, the times series (1.1) has not only infinite variance but also exhibits long-range dependence (long memory). For more details, see Samorodnitsky and Taqqu (1994), Kokoszka and Taqqu (1995*a*) and (1995*b*).

Our goal is to estimate both d and the coefficients of the polynomials Φ and Θ, by using a variant of Whittle's method. For a stationary Gaussian time series with spectral density $g(\lambda, \beta)$, $-\pi < \lambda < \pi$, Whittle's method, which provides an estimate of β, requires replacing the inverse covariance matrix that appears in the Gaussian likelihood by a Toeplitz (covariance) matrix with spectral density $1/g$ and then maximizing the quadratic form. Hannan (1973) applied Whittle's method to finite variance ARMA time series, that is to (1.1) with $d = 0$. He proved that the estimator is consistent and asymptotically normal. An ARMA time series, however, has short range dependence because the correlations decrease exponentially fast. Fox and Taqqu (1986) extended this result to Gaussian time series with long-range dependence such as fractional Gaussian noise or fractional ARIMA by appealing

[*] The second author was partially supported by the NSF Grants DMS-9404093 and NCR-9404931 at Boston University.
[†] *AMS Subject classification.* 60E07, 62F12. *Short title:* Estimation for Infinite Variance FARIMA.
[‡] *Keywords and phrases:* estimation, fractional ARIMA, long memory, stable distributions, heavy tails.

to a central limit theorem for weighted quadratic forms whose weights are chosen in such a way as to compensate for the long-range dependence. Fox and Taqqu's result, which was later generalized to the full maximum likelihood by Dahlhaus (1989), is the basis of one of the most commonly used techniques for estimating the intensity of long-range dependence in Gaussian time series (see Beran (1994)). Giraitis and Surgailis (1990) extended Fox and Taqqu's result to finite variance innovations without Gaussian assumptions and Heyde and Gay (1993) to random fields.

When the innovations are in the domain of attraction of an infinite variance stable random variable, covariances stop making sense. One can, however, still use the same estimator as in the Gaussian case. Doing so has the advantage of not having to determine beforehand the often unknown distributions of the innovations. It is necessary, however, to verify that these estimators have good properties in the infinite variance case as well. Mikosch, Gadrich, Klüppelberg and Adler (1995) showed that this is the case for ARMA time series. In this paper we extend the result of Mikosch *et al.* to fractional ARIMA time series which have long-range dependence. We prove that the estimator is consistent and determine its asymptotic distribution. Because of the slow decay of the coefficients in the fractional ARIMA time series, very few of the technical arguments used by Mikosch et al. (1995) carry over to our setting and hence significantly different proofs of the basic lemmas had to be developed.

Assume then that the innovations Z_n in (1.1) are i.i.d. with *mean zero* and are in the domain of attraction of an α-stable law with $1 < \alpha < 2$, i.e.

$$P(|Z_n| > x) = x^{-\alpha} L(x), \quad \text{as} \quad x \to \infty, \tag{1.2}$$

where L is a slowly varying function, and

$$P(Z_n > x)/P(|Z_n| > x) \to a, \; P(Z_n < -x)/P(|Z_n| > x) \to b, \tag{1.3}$$

where a and b are nonnegative numbers satisfying $a + b = 1$. It has been shown in Kokoszka (1995) (and Kokoszka and Taqqu (1995a) in the case of stable innovations) that for the Z_n as above, there is a unique moving average

$$X_n = \sum_{j=0}^{\infty} c_j Z_{n-j} \tag{1.4}$$

satisfying Equations (1.1), provided that the polynomials Φ and Θ have no zeros in the closed unit disk $D = \{z : |z| \le 1\}$ and no zeros in common and that $d < 1 - \frac{1}{\alpha}$. (This is why we suppose α greater than 1.) The coefficients c_j in (1.4) are defined by

$$\sum_{j=0}^{\infty} c_j z^j = \frac{\Theta(z)}{\Phi(z)(1-z)^d}\,, \quad |z| < 1, \tag{1.5}$$

and are asymptotically proportional to j^{d-1} as $j \to \infty$. Therefore they do not satisfy the fundamental assumption

$$\sum_{j=0}^{\infty} j|c_j|^{1 \wedge \mu} < \infty \quad \text{for some } 0 < \mu < \alpha \tag{1.6}$$

of Mikosch *et al.* (1995). The fact that the c_j's are not absolutely summable turns out to be a major source of difficulties.

We want to estimate the $(p+q+1)$-dimensional vector

$$\beta_0 = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, d)$$

where $\phi_1, \ldots, \phi_p$ and $\theta_1, \ldots, \theta_q$ are the coefficients of the autoregressive polynomial $\Phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p$ and the moving average polynomial $\Theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q$ respectively, and d is the differencing parameter in (1.1). We assume that the true value of d is positive and hence lies in the open interval $(0, 1 - \frac{1}{\alpha})$. In the case of Gaussian innovations, positive d corresponds to a spectral density that diverges at zero (1/f noise). The preceding discussion motivates the choice of our parameter space E, namely a compact set contained in

$$\Big\{(\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, d) : \ \phi_p \neq 0, \ \theta_q \neq 0, \ \Phi(z) \text{ and } \Theta(z) \text{ have no common zeros,}$$
$$\Phi(z)\Theta(z) \neq 0 \text{ for } |z| \leq 1, \ d \in \Big(0, 1 - \frac{1}{\alpha}\Big)\Big\}.$$

We introduce now some additional notation which will be used throughout the paper. The time series (1.4), called fractional ARIMA, will be referred to as FARIMA(p,d,q). The elements of E are denoted β, possibly with some sub- and/or superscripts. The last coordinate of the vector β is the difference parameter d. The true value of the parameter vector is denoted β_0. The sample autocovariance and autocorrelation functions are defined respectively by

$$\gamma_n(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} X_t X_{t+|h|} \tag{1.7}$$

and

$$\rho_n(h) = \Big(\sum_{t=1}^{n} X_t^2\Big)^{-1} \Big(\sum_{t=1}^{n-|h|} X_t X_{t+|h|}\Big) = (\gamma_n(0))^{-1} \gamma_n(h). \tag{1.8}$$

We will frequently use the corresponding deterministic quantities

$$\gamma(h) = \sum_{j=0}^{\infty} c_j c_{j+|h|} \ , \quad \rho(h) = (\gamma(0))^{-1} \gamma(h). \tag{1.9}$$

The *normalized periodogram* is defined as follows:

$$\tilde{I}_n(\lambda) = \Big(\sum_{t=1}^{n} X_t^2\Big)^{-1} \Big|\sum_{t=1}^{n} X_t e^{-i\lambda t}\Big|^2 = \sum_{|h|<n} \rho_n(h) e^{-i\lambda h}, \ -\pi \leq \lambda \leq \pi. \tag{1.10}$$

For $\beta \in E$, the *power transfer function* is

$$g(\lambda, \beta) = \left| \frac{\Theta(e^{-i\lambda}, \beta)}{\Phi(e^{-i\lambda}, \beta)(1 - e^{-i\lambda})^{d(\beta)}} \right|^2 = \Big|\sum_{j=0}^{\infty} c_j(\beta) e^{-i\lambda j}\Big|^2. \tag{1.11}$$

Following Fox and Taqqu (1986) who consider *Gaussian* fractional ARIMA and Mikosch *et al.* (1995) who study *infinite variance* ARMA processes, we define the estimator β_n based on the sample $X_1, \ldots, X_n$ as the value of β minimizing

$$\sigma_n^2(\beta) = \int_{-\pi}^{\pi} \frac{\tilde{I}_n(\lambda)}{g(\lambda, \beta)} d\lambda, \quad \beta \in E. \tag{1.12}$$

Notice that under our assumptions, the function $1/g(\cdot,\cdot)$ is continuous on $[-\pi, \pi] \times E$, and hence, in particular, the integral in (1.12) is well-defined. The following consistency result holds:

Theorem 1.1 *If β_0 is the true parameter and β_n is the value of β minimizing $\sigma_n^2(\beta)$, then*

$$\beta_n \xrightarrow{P} \beta_0 \tag{1.13}$$

and

$$\sigma_n^2(\beta_n) \xrightarrow{P} \frac{2\pi}{\gamma(0)}. \tag{1.14}$$

This theorem is proved in Section 2.2. As part of the proof, we extend the limit theorem for sample autocovariances of infinite variance moving averages developed in Davis and Resnick (1985) to moving averages whose coefficients are not absolutely summable.

We now turn to the asymptotic distribution of the estimator β_n. Theorem 1.2 below, which generalizes Theorem 2.2 of Mikosch *et al.* (1995) and which is an infinite variance analog of Theorem 2 of Fox and Taqqu (1986), is valid under a more restrictive assumption on the innovations Z_n. We now assume that the Z_n are *symmetric* and are in the *domain of normal attraction* of a $S\alpha S$ law i.e.

$$n^{-1/\alpha} \sum_{j=1}^{n} Z_j \xrightarrow{\mathcal{D}} Y, \tag{1.15}$$

where $E\exp(i\theta Y) = \exp\{-\sigma^\alpha|\theta|^\alpha\}$. Relation (1.15) is equivalent to (see e.g. Gnedenko and Kolmogorov (1954))

$$\lim_{\lambda\to\infty} \lambda^\alpha P(Z < -\lambda) = \frac{C_\alpha \sigma^\alpha}{2} \quad \text{and} \quad \lim_{\lambda\to\infty} \lambda^\alpha P(Z > \lambda) = \frac{C_\alpha \sigma^\alpha}{2}, \tag{1.16}$$

where

$$C_\alpha = \left(\alpha \int_0^\infty (1-\cos x)\frac{dx}{x^{\alpha+1}}\right)^{-1} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} & \text{if } \alpha \neq 1, \\ 2/\pi & \text{if } \alpha = 1. \end{cases} \tag{1.17}$$

We make these additional assumptions on the Z_n because our proofs depend heavily on the results of Rosinski and Woyczynski (1987) which require that the Z_n (the X_i in their paper) be symmetric and satisfy $\limsup_{\lambda\to\infty} \lambda^\alpha P(|Z_n| > \lambda) \le C < \infty$.

In order to state our result we introduce the $(p+q+1)\times(p+q+1)$ matrix $W(\beta_0)$ with entries

$$w_{ij} = \int_{-\pi}^{\pi} g(\lambda,\beta_0)\frac{\partial^2}{\partial\beta_i\partial\beta_j} g^{-1}(\lambda,\beta_0)d\lambda, \quad i,j = 1,\ldots,p+q+1, \tag{1.18}$$

and the $(p+q+1)$-dimensional vectors $b_k,\ k \in \mathbb{Z}$, whose jth coordinate is

$$(b_k)_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ik\lambda} g(\lambda,\beta_0)\frac{\partial}{\partial\beta_j} g^{-1}(\lambda,\beta_0)d\lambda, \quad j = 1,\ldots,p+q+1. \tag{1.19}$$

The following theorem shows that, as in the Gaussian case, the asymptotic result for $d=0$ extends to $d>0$.

Theorem 1.2 *If the innovations Z_n are symmetric and satisfy*

$$\lim_{\lambda\to\infty} \lambda^\alpha P(|Z| > \lambda) = C_\alpha \sigma^\alpha, \tag{1.20}$$

then

$$\left(\frac{n}{\log n}\right)^{1/\alpha}(\beta_n-\beta_0)\xrightarrow{\mathcal{D}}4\pi W^{-1}(\beta_0)\sum_{k=1}^{\infty}\frac{Y_k}{Y_0}b_k, \tag{1.21}$$

where Y_0 is positive $\frac{\alpha}{2}$-stable with scale parameter $C_{\alpha/2}^{-2/\alpha}$ and the $Y_k,\ k\geq 1$ are i.i.d. $S\alpha S$ with scale parameter $C_\alpha^{-1/\alpha}$. The random variables $Y_0, Y_1, Y_2, \ldots$ are independent and C_α is given in (1.17).

Setting $c_k = 4\pi W^{-1}(\beta_0)b_k$, observe that the j^{th} coordinate of the limiting random vector $\sum_{k=1}^{\infty} c_k Y_k/Y_0$ is distributed as $(\sum_{k=1}^{\infty}|(c_k)_j|^{\alpha})^{1/\alpha}Y_1/Y_0$, that is, as the ratio of two independent stable random variables. Observe also that the result is similar to the one in the ARMA case (see Theorem 2.2 in Mikosch *et al.* (1995). In that theorem, the scale parameter of $Y_k,\ k\geq 1$, should be $C_\alpha^{-1/\alpha}$ and not $C_\alpha^{1/\alpha}$).

Theorem 1.2 is a first step in the development of statistical procedures for time series that exhibit both infinite variance and long-range dependence. Its proof is presented in Section 3. In Section 4 we describe the results of a small simulation study.

2 Consistency of the estimator

The proof of the Consistency Theorem 1.1, which is presented in Section 2.1 below, follows in its main outline that of Theorem 2.1 of Mikosch *et al.* (1995). In our case, however, the power transfer function $g(\lambda, \beta_0)$ diverges to infinity at $\lambda = 0$, so the arguments developed for continuous g do not carry over. By working with the compact parameter space E, we are able to avoid some technical complications.

We first establish the following extention of Theorem 4.2 of Davis and Resnick (1985).

Theorem 2.1 *Suppose the innovations Z_n have mean zero and satisfy* (1.2) *and* (1.3) *and*

$$\sum_{j=0}^{\infty}|c_j|^{\alpha-\epsilon}<\infty \tag{2.1}$$

for some $\epsilon > 0$. Then, for the moving average

$$X_n=\sum_{j=0}^{\infty}c_jZ_{n-j},$$

we have

$$\left(a_n^{-2}\sum_{t=1}^{n-|h|}X_tX_{t+|h|},\ |h|<m\right)\xrightarrow{\mathcal{D}}\left((\sum_{j=0}^{\infty}c_jc_{j+|h|})Y_0,\ |h|<m\right), \tag{2.2}$$

where Y_0 is as in Theorem 1.2 and the a_n are determined by the condition

$$\forall x>0\quad \lim_{n\to\infty}nP(|Z_1|>a_nx)=x^{-\alpha}. \tag{2.3}$$

Observe that in Theorem 2.1 we do not assume the absolute summability of the c_j if $\alpha > 1$, which was a global assumption in the paper of Davis and Resnick (1985). A careful study of their proofs shows that the result depends on the relation:

$$\lim_{t\to\infty}\frac{P(|\sum_{j=0}^{\infty}c_jZ_j|>t)}{P(|Z_1|>t)}=\sum_{j=0}^{\infty}|c_j|^{\alpha}, \tag{2.4}$$

and on Condition (2.1), which guarantees that the process X_n is well-defined (see e.g. Avram and Taqqu (1986)). Relation (2.4) was proved by Cline (1983) under the assumption that $\sum_{j=0}^{\infty}|c_j|<\infty$, if $\alpha>1$. Because the coefficients c_j in the moving average representation of fractional ARIMA processes behave like j^{d-1} as $j\to\infty$, they are not absolutely summable if $d>0$, and hence we cannot use Cline's result here. We will show, however, that (2.4) continues to hold under Condition (2.1) if the Z_j have mean zero. As this fact is of central importance to the present paper and is also of independent interest we formulate it as a separate theorem.

Theorem 2.2 *Suppose* $1<\alpha<2$ *and the* Z_n *have mean zero and satisfy* (1.2) *and* (1.3). *Then Condition* (2.1) *implies Relation* (2.4).

The preceding discussion and Theorem 2.2 imply Theorem 2.1.

The proof of Theorem 2.2 which utilizes the ideas of the proof of Lemma 4.2 of Resnick (1987) is given below. (The theorem holds, with the same proof, for two-sided moving averages.)

2.1 Proof of Theorem 2.2

We start by describing the basic idea. While our argument essentially follows the one presented in Resnick (1987) pp. 228-230, the crucial difference is that in order to find an effective upper bound for

$$P(|\sum_j c_j Z_j 1_{[|Z_j|\le x|c_j|^{-1}]}|>x),$$

we use the Chebyshev Inequality rather than the Markov Inequality. This makes it unnecessary to use Jensen's inequality to reduce the case $\alpha\ge 1$ to the case $\alpha<1$, a procedure which required the assumption $\sum_j|c_j|<\infty$ ($\sum_j$ stands for $\sum_{j=0}^{\infty}$ or $\sum_{j=-\infty}^{\infty}$). Observe first that

$$\begin{aligned}(2.5)\quad P(|\sum_j c_jZ_j|>x) &= P(|\sum_j c_jZ_j|>x,\ \sup_j|c_jZ_j|>x)+P(|\sum_j c_jZ_j|>x,\ \sup_j|c_jZ_j|\le x)\\ &\le P(\bigcup_j\{|c_jZ_j|>x\})+P(|\sum_j c_jZ_j1_{[|c_jZ_j|\le x]}|>x)\\ &\le \sum_j P(|Z_j|>x|c_j|^{-1})+x^{-2}E|\sum_j c_jZ_j1_{[|Z_j|\le x|c_j|^{-1}]}|^2.\end{aligned}$$

We first verify that the series $\sum_j Y_j$, $Y_j=c_jZ_j1_{[|c_jZ_j|\le x]}$ converges in L^2. Observe that the Y_j need not be orthogonal. It suffices to show that $\sum_j|EY_j|<\infty$ and $\sum_j E|Y_j-EY_j|^2<\infty$. In the arguments below we often use Potter's theorem (see Theorem 1.5.6(c) of Bingham et al. (1987)). Since $EZ_j=0$, we have for sufficiently large j.

$$\begin{aligned}(2.6)\qquad |EY_j| &= |E\{c_jZ_j1_{[|c_jZ_j|>x]}\}|\\ &\le \int_0^x P(|c_jZ_j|>x)dt+\int_x^{\infty}P(|c_jZ_j|>t)dt\\ &\le x(1+\epsilon)x^{-\alpha+\epsilon}|c_j|^{\alpha-\epsilon}+(1+\epsilon)|c_j|^{\alpha-\epsilon}\int_x^{\infty}t^{-\alpha+\epsilon}dt\\ &= (1+\epsilon)\frac{\alpha-\epsilon}{\alpha-1-\epsilon}x^{1-\alpha+\epsilon}|c_j|^{\alpha-\epsilon},\end{aligned}$$

which shows that the series $\sum_j EY_j$ converges absolutely. Since $\sum_j E|Y_j - EY_j|^2 \le \sum_j EY_j^2$, it remains to verify that $\sum_j EY_j^2 < \infty$. It follows from the well-known relation

$$\frac{E[Z_1^2 1_{[|Z_1|\le x]}]}{x^2 P(|Z_1|>x)} \to \frac{\alpha}{2-\alpha} \tag{2.7}$$

(see Section 8.1 of Bingham et al. (1987)) that the function $U(x) := E[|Z_1|^2 1_{[|Z_1|\le x]}]$ is regularly varying with index $2-\alpha$, and, consequently, for sufficiently large x and some constant K,

$$\frac{EY_j^2}{x^2 P(|Z_1|>x)} = |c_j|^2 \frac{E[Z_1^2 1_{[|Z_1|\le x|c_j|^{-1}]}]}{x^2 P(|Z_1|>x)} \tag{2.8}$$

$$= |c_j|^2 \frac{U(x|c_j|^{-1})}{U(x)} \frac{U(x)}{x^2 P(|Z_1|>x)} \le K|c_j|^2 (|c_j|^{-1})^{2-\alpha+\epsilon} = K|c_j|^{\alpha-\epsilon}.$$

Since $E|\sum_j Y_j|^2 = \mathrm{Var}(\sum Y_j) + (E\sum Y_j)^2 \le \sum_j EY_j^2 + (\sum_j |EY_j|)^2$, we obtain from (2.5), (2.6) and (2.8),

$$\frac{P(|\sum_j c_j Z_j|>x)}{P(|Z_1|>x)} \le S_1(x) + S_2(x) + S_3(x), \tag{2.9}$$

where

$$S_1(x) = \sum_j \frac{P(|Z_1|>x|c_j|^{-1})}{P(|Z_1|>x)}, \quad S_2(x) = \sum_j \frac{EY_j^2}{x^2 P(|Z_1|>x)}, \quad S_3(x) = \frac{(\sum_j |EY_j|)^2}{x^2 P(|Z_1|>x)}.$$

Since $P(|Z_1|>x)$ is regularly varying at infinity with index $-\alpha$, for sufficiently large j the summands in the sum defining $S_1(x)$ do not exceed $(1+\epsilon)|c_j|^{\alpha-\epsilon}$, and so $\lim_{x\to\infty} S_1(x) = \sum_j |c_j|^\alpha$. By (2.7), (2.8) and the Dominated Convergence Theorem, $\lim_{x\to\infty} S_2(x) = \frac{\alpha}{2-\alpha}\sum_j |c_j|^\alpha$. Finally, by (2.6), $S_3(x) \le c(\alpha,\epsilon)(\sum_j |c_j|^{\alpha-\epsilon})^2 x^{-\alpha+\epsilon}$. The inequality (2.9) and these relations yield

$$\limsup_{x\to\infty} \frac{P(|\sum_j c_j Z_j|>x)}{P(|Z_1|>x)} \le \frac{2}{2-\alpha}\sum_j |c_j|^\alpha. \tag{2.10}$$

Now, for any $0<r<1$ and any positive integer m,

$$P(|\sum_j c_j Z_j|>x) \ge P(|\sum_{|j|\le m} c_j Z_j| > (1+r)x) - P(|\sum_{|j|>m} c_j Z_j| > rx).$$

Since the hypothesis of the theorem holds for finite sums (see Proposition on p. 278 of Feller (1971)), (2.10) yields

$$\liminf_{x\to\infty} \frac{P(|\sum_j c_j Z_j|>x)}{P(|Z_1|>x)} \ge (1+r)^{-\alpha} \sum_{|j|\le m} |c_j|^\alpha - \frac{2}{(2-\alpha)r^\alpha} \sum_{|j|>m} |c_j|^\alpha. \tag{2.11}$$

Similarly,

$$\limsup_{x\to\infty} \frac{P(|\sum_j c_j Z_j|>x)}{P(|Z_1|>x)} \le (1-r)^{-\alpha} \sum_{|j|\le m} |c_j|^\alpha + \frac{2}{(2-\alpha)r^\alpha} \sum_{|j|>m} |c_j|^\alpha. \tag{2.12}$$

Letting first $m\to\infty$ and then $r\to 0$ in (2.11) and (2.12), we get (2.4). ∎

2.2 Proof of the Consistency Theorem 1.1

The proof of Theorem 1.1 uses two lemmas. The first extends Proposition 10.8.1 of Brockwell and Davis (1991); the second extends Lemma 1 of Fox and Taqqu (1986) and Lemma 6.1 of Mikosch *et al.* (1995).

Lemma 2.1 *Suppose $\beta_1, \beta_2 \in E$. If $\beta_1 \neq \beta_2$, then*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{g(\lambda,\beta_1)}{g(\lambda,\beta_2)}d\lambda > 1. \tag{2.13}$$

PROOF: For $|z| < 1$ and $\beta \in E$ define

$$C(z,\beta) = \frac{\Theta(z,\beta)}{\Phi(z,\beta)(1-z)^{d(\beta)}} = \sum_{j=0}^{\infty} c_j(\beta) z^j,$$

$$H(z,\beta) = 1/C(z,\beta) = \sum_{j=0}^{\infty} h_j(\beta) z^j.$$

Let $\{\epsilon_n\}$ be a sequence of i.i.d. $N(0,1)$ random variables and consider the Gaussian fractional ARIMA process

$$X_n(\beta_1) = \sum_{j=0}^{\infty} c_j(\beta_1)\epsilon_{n-j}.$$

It is well known (see e.g. Brockwell and Davis (1991) §13.2) that $\mathrm{Var}(X_{n+1}(\beta_1) - \sum_{j=0}^{\infty} u_j X_{n-j}(\beta_1))$ is minimized if and only if $u_j = -h_{j+1}(\beta_1)$ and the minimum value of the variance is 1. Since $\beta_2 \neq \beta_1$, we have $H(\cdot,\beta_2) \neq H(\cdot,\beta_1)$, and so $\mathrm{Var}(X_{n+1}(\beta_1) + \sum_{j=0}^{\infty} h_{j+1}(\beta_2)X_{n-j}(\beta_1)) > 1$. This concludes the proof because that variance equals

$$\sum_{j=0}^{\infty}\left|\sum_{k=0}^{j} h_k(\beta_2)c_{j-k}(\beta_1)\right|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H(e^{-i\lambda},\beta_2)C(e^{-i\lambda},\beta_1)|^2 d\lambda$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{C(e^{-i\lambda},\beta_1)}{C(e^{-i\lambda},\beta_2)}\right|^2 d\lambda = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{g(\lambda,\beta_1)}{g(\lambda,\beta_2)}d\lambda. \quad \blacksquare$$

Remark. The lemma holds whenever $|d| < \frac{1}{2}$, thus not only for positive values of d. Indeed, since for $d < \frac{1}{2}$, $C(e^{-i\cdot},\beta) \in L^2[-\pi,\pi]$ and for $d > -\frac{1}{2}, H(e^{-i\cdot},\beta) \in L^2[\pi,\pi]$, Parseval's relation applies whenever $|d| < 1/2$.

Lemma 2.2 *Let $f(\lambda,\beta)$ be any continuous (and hence uniformly continuous) function on $[-\pi,\pi]\times E$. Then, as $n \to \infty$,*

$$\sup_{\beta\in E}\left|\int_{-\pi}^{\pi} f(\lambda,\beta)\tilde{I}_n(\lambda)d\lambda - \frac{1}{\gamma(0)}\int_{-\pi}^{\pi} f(\lambda,\beta)g(\lambda,\beta_0)d\lambda\right| \xrightarrow{P} 0. \tag{2.14}$$

PROOF: The proof is similar to the proof of Lemma 6.1 of Mikosch *et al.* (1995). Note, however, that for fractional ARIMA processes the function $g(\lambda, \beta_0)$ diverges at $\lambda = 0$.

Let $K_m(e^{i\lambda}) = \sum_{|h|<m}(1 - \frac{|h|}{m})e^{i\lambda h}$ denote the Fejér kernel. Fix $\epsilon > 0$ and choose m so large that for all λ and β,

$$|K_m * f(\lambda, \beta) - f(\lambda, \beta)| < \frac{\epsilon}{4\pi}. \tag{2.15}$$

(To verify (2.15), repeat the proof of Fejér's theorem for continuous functions and use the uniform continuity of f in both variables.) Hence, for any β,

$$\Big| \int_{-\pi}^{\pi} f(\lambda, \beta) \tilde{I}_n(\lambda) d\lambda - \int_{-\pi}^{\pi} K_m * f(\lambda, \beta) \tilde{I}_n(\lambda) d\lambda \Big| < \frac{\epsilon}{4\pi} \int_{-\pi}^{\pi} \tilde{I}_n(\lambda) d\lambda = \frac{\epsilon}{2}.$$

Consequently,

$$P\left(\sup_\beta \left| \int_{-\pi}^{\pi} f(\lambda, \beta) \tilde{I}_n(\lambda) d\lambda - \gamma(0)^{-1} \int_{-\pi}^{\pi} f(\lambda, \beta) g(\lambda, \beta_0) d\lambda \right| \geq \epsilon \right) \tag{2.16}$$

$$\leq P\left(\sup_\beta \left| \int_{-\pi}^{\pi} K_m * f(\lambda, \beta) \tilde{I}_n(\lambda) d\lambda - \gamma(0)^{-1} \int_{-\pi}^{\pi} f(\lambda, \beta) g(\lambda, \beta_0) d\lambda \right| \geq \frac{\epsilon}{2} \right)$$

$$= P\left(\sup_\beta \left| 2\pi \sum_{|h|<m} \left(1 - \frac{|h|}{m}\right) \hat{f}(h, \beta) \rho_n(h) - \gamma(0)^{-1} \int_{-\pi}^{\pi} f(\lambda, \beta) g(\lambda, \beta_0) d\lambda \right| \geq \frac{\epsilon}{2} \right),$$

where $\rho_n(h)$ is given in (1.8) and $\hat{f}(h, \beta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ih\lambda} f(\lambda, \beta) d\lambda$. The last quantity in (2.16) is bounded above by

$$P\left(\sup_\beta \left| 2\pi \sum_{|h|<m} (\rho_n(h) - \rho(h)) \left(1 - \frac{|h|}{m}\right) \hat{f}(h, \beta) \right| \geq \frac{\epsilon}{4} \right) \tag{2.17}$$

$$+P\left(\sup_\beta \left| 2\pi \sum_{|h|<m} \rho(h) \left(1 - \frac{|h|}{m}\right) \hat{f}(h, \beta) - \gamma(0)^{-1} \int_{-\pi}^{\pi} f(\lambda, \beta) g(\lambda, \beta_0) d\lambda \right| \geq \frac{\epsilon}{4} \right).$$

The first term in (2.17) tends to zero because, by Theorem 2.1, $\rho_n(h) \xrightarrow{\mathcal{D}} \rho(h)$ and $|\hat{f}(h, \beta)| \leq \sup_{-\pi \leq \lambda \leq \pi} \sup_{\beta \in E} |f(\lambda, \beta)| < \infty$. To evaluate the second term in (2.17) observe that

$$2\pi \sum_{|h|<m} \rho(h)(1 - \frac{|h|}{m})\hat{f}(h, \beta) = \gamma(0)^{-1} \int_{-\pi}^{\pi} \left[\sum_{|h|<m} \gamma(h)(1 - \frac{|h|}{m})e^{i\lambda h} \right] f(\lambda, \beta) d\lambda$$

and

$$\sum_{|h|<m} \gamma(h)(1 - \frac{|h|}{m})e^{i\lambda h} = K_m * g(\lambda, \beta_0).$$

Therefore the second term in (2.17) equals

$$P\left(\sup_\beta |\gamma(0)^{-1} \int_{-\pi}^{\pi} [K_m * g(\lambda, \beta_0) - g(\lambda, \beta_0)] f(\lambda, \beta) d\lambda| \geq \frac{\epsilon}{4} \right),$$

which is zero for sufficiently large m because $f(\lambda,\beta)$ is uniformly bounded on $[-\pi,\pi]\times E$ and, by Fejér's theorem (see e.g. Helson (1983) p. 14),

$$\int_{-\pi}^{\pi} |K_m * g(\lambda,\beta_0) - g(\lambda,\beta_0)| d\lambda \to 0, \text{ as } m\to\infty,$$

since $g(\cdot,\beta_0)\in L^1[-\pi,\pi]$. This concludes the proof. ■

PROOF OF THE CONSISTENCY THEOREM 1.1: Since $d(\beta)>0$, the function $1/g(\lambda,\beta)$ is continuous on $[-\pi,\pi]\times E$, and hence Lemma 2.2 applies. We get

$$\sigma_n^2 = \int_{-\pi}^{\pi} \frac{\tilde{I}_n(\lambda)}{g(\lambda,\cdot)} d\lambda \xrightarrow{P} \sigma^2 := \frac{1}{\gamma(0)} \int_{-\pi}^{\pi} \frac{g(\lambda,\beta_0)}{g(\lambda,\cdot)} d\lambda, \tag{2.18}$$

where σ_n^2 and σ^2 are random elements of the function space $C(E)$ of continuous functions on E equipped with the sup-norm. The remainder of the proof, included for completeness, is a variation on the proof of Theorem 2.1 of Mikosch *et al.* (1995).

Since β_0 is a constant, to prove (1.13) it suffices to show that $\beta_n \xrightarrow{\mathcal{D}} \beta_0$. As E is compact, the sequence $\{\beta_n\}$ is tight, and hence $\beta_n \xrightarrow{\mathcal{D}} \beta_0$ if and only if every weakly convergent subsequence of $\{\beta_n\}$ converges weakly to β_0.

Let then $\{\beta_m\}$ be a subsequence of $\{\beta_n\}$ such that $\beta_m \xrightarrow{\mathcal{D}} \beta'$ (β' is a random variable). We want to show that $\beta'=\beta_0$ a.s.

By (2.18) and Theorem 4.4 of Billingsley (1968),

$$(\sigma_m^2,\beta_m) \xrightarrow{\mathcal{D}} (\sigma^2,\beta') \quad (\text{in } C(E)\times E), \tag{2.19}$$

since σ^2 is a non-random element in $C(E)$. By the Continuous Mapping Theorem, (2.19) implies

$$\sigma_m^2(\beta_m) \xrightarrow{\mathcal{D}} \sigma^2(\beta'). \tag{2.20}$$

and hence, for any t,

$$\limsup_m P(\sigma_m^2(\beta_m)\le t) \le P(\sigma^2(\beta')\le t). \tag{2.21}$$

Now, the definition of β_n and (2.18) yield

$$\sigma_m^2(\beta_m) \le \sigma_m^2(\beta_0) \xrightarrow{P} t_0 := \frac{2\pi}{\gamma(0)}.$$

Consequently,

$$\liminf_m P(\sigma_m^2(\beta_m) < t) \ge \liminf_m P(\sigma_m^2(\beta_0) < t) \ge 1, \quad \forall t > t_0, \tag{2.22}$$

and so $\limsup_m P(\sigma_m^2(\beta_m)\le t) = 1$ for all $t>t_0$. Using (2.21) we obtain

$$P(\sigma^2(\beta')\le t) = 1, \qquad \forall t>t_0. \tag{2.23}$$

Moreover, by Lemma 2.1, $\sigma^2(\beta')>t_0$ whenever $\beta'\ne\beta_0$, and so

$$\lim_{t\searrow t_0} P\{\sigma^2(\beta')\le t,\ \beta'\ne\beta_0\} = P\{\sigma^2(\beta')\le t_0,\ \beta'\ne\beta_0\} = 0. \tag{2.24}$$

The equality (2.23) implies that for any $t > t_0$,

$$\begin{aligned} 1 &= P(\sigma^2(\beta') \le t,\ \beta' = \beta_0) + P(\sigma^2(\beta') \le t,\ \beta' \ne \beta_0) \\ &\le P(\beta' = \beta_0) + P(\sigma^2(\beta') \le t,\ \beta' \ne \beta_0), \end{aligned}$$

which together with (2.24) yields $P(\beta' = \beta_0) = 1$. This is what we wanted to establish.

Finally, to prove (1.14), write

$$P(|\sigma_n^2(\beta_n) - \sigma^2(\beta_0)| \ge \epsilon) \le P(|\sigma_n^2(\beta_n) - \sigma^2(\beta_n)| \ge \epsilon/2) + P(|\sigma^2(\beta_n) - \sigma^2(\beta_0)| \ge \epsilon/2). \tag{2.25}$$

The first term in the right-hand side of (2.25) tends to zero by (2.18), the second because $\beta_n \xrightarrow{P} \beta_0$ and σ^2 is continuous on E. ∎

3 Asymptotic distribution of the estimator

The crucial element in the proof of Theorem 1.2 is Proposition 3.1 below, which is established in Section 3.2. We show first that this proposition yields Theorem 1.2.

Proposition 3.1 *For real numbers $u_1, u_2, \ldots, u_{p+q+1}$ (which are fixed but arbitrary) set*

$$\eta(\lambda) = \sum_{j=1}^{p+q+1} u_j \frac{\partial}{\partial \beta_j} g^{-1}(\lambda, \beta_0). \tag{3.1}$$

Then, as $n \to \infty$,

$$\left(\frac{n}{\log n}\right)^{1/\alpha} \int_{-\pi}^{\pi} \tilde{I}_n(\lambda)\eta(\lambda)d\lambda \xrightarrow{\mathcal{D}} \frac{4\pi}{\gamma(0)} \sum_{k=1}^{\infty} \frac{Y_k}{Y_0} \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ik\lambda} g(\lambda, \beta_0)\eta(\lambda)d\lambda\right) \tag{3.2}$$

where the random variables Y_0, Y_1, Y_2 are as in Theorem 1.2.

Proof of Theorem 1.2: Let $\frac{\partial}{\partial \beta}$ denote the column vector with entries $\frac{\partial}{\partial \beta_j}$, $j = 1, \ldots, p+q+1$ and $\frac{\partial^2}{\partial \beta^2}$ the matrix with entries $\frac{\partial^2}{\partial \beta_i \partial \beta_j}$, $i, j = 1, \ldots, p+q+1$. Relation (3.2) implies

$$\left(\frac{n}{\log n}\right)^{1/\alpha} \frac{\partial}{\partial \beta} \sigma_n^2(\beta_0) \xrightarrow{\mathcal{D}} \frac{4\pi}{\gamma(0)} \sum_{k=1}^{\infty} \frac{Y_k}{Y_0} b_k \tag{3.3}$$

(cf. (1.12)). Also,

$$\frac{\partial^2}{\partial \beta^2} \sigma_n^2(\beta_0) = \int_{-\pi}^{\pi} \tilde{I}_n(\lambda) \frac{\partial^2}{\partial \beta^2} g^{-1}(\lambda, \beta_0) d\lambda. \tag{3.4}$$

Differentiation under the integral in (3.3) and (3.4) is justified because the function $g(\lambda, \beta)$ is proportional to the spectral density of Gaussian fractional ARIMA and consequently satisfies Condition (A3) on p. 521 of Fox and Taqqu (1986), namely, for any $\delta > 0$ there are constants $K_1(\delta)$ and $K_2(\delta)$ such that

$$\left|\frac{\partial}{\partial \beta_j} g^{-1}(\lambda, \beta)\right| \le K_1(\delta)|\lambda|^{2d(\beta)-\delta},\ j = 1, \ldots, p+q+1 \tag{3.5}$$

and

(3.6) $$\left|\frac{\partial^2}{\partial\beta_i\partial\beta_j}g^{-1}(\lambda,\beta)\right| \le K_2(\delta)|\lambda|^{2d(\beta)-\delta},\ i,j=1,\ldots,p+q+1.$$

Since E is compact, we can clearly assume that there is a $\delta > 0$ such that

(3.7) $$\inf_{\beta\in E}(2d(\beta)-\delta) > 0.$$

Condition (3.7) together with (3.6) show that $\partial^2/\partial\beta^2(g^{-1}(\lambda,\beta_0))$ has continuous components, and so, by Lemma 2.2 and (3.4),

(3.8) $$\sup_{\beta\in E}\left\|\frac{\partial^2}{\partial\beta^2}\sigma_n^2(\beta)-\frac{1}{\gamma(0)}W(\beta_0)\right\|\xrightarrow{P} 0.$$

Since β_n minimizes $\sigma_n^2(\beta)$, there is some β_n^* satisfying $\|\beta_n^*-\beta_0\| < \|\beta_n-\beta_0\|$ such that

(3.9) $$\frac{\partial}{\partial\beta}\sigma_n^2(\beta_0) = -\frac{\partial^2}{\partial\beta^2}\sigma_n^2(\beta_n^*)(\beta_n-\beta_0).$$

Multiplying both sides of (3.9) by $(n/\log n)^{1/\alpha}$ and using (3.3) and (3.8) together with Theorem 4.4 of Billingsley (1968), yields (1.21).

3.1 Tools

We state here several results on which we rely extensively.

The following proposition follows from Theorem 3.1 of Rosinski and Woyczynski (1987).

Proposition 3.2 *Suppose $Z_1, Z_2, \ldots$ is a sequence of i.i.d. symmetric random variables satisfying*

(3.10) $$\limsup_{\lambda\to\infty}\lambda^\alpha P(|Z_1|>\lambda)\le C<\infty,$$

where $0<\alpha<2$. Consider the sequence of bilinear forms

(3.11) $$Q_n = \sum_{\substack{j,k=1\\ j\neq k}}^{n} q(j,k)Z_jZ_k,$$

and set

(3.12) $$N_\alpha^{(n)} = \sum_{\substack{j,k=1\\ j\neq k}}^{n}|q(j,k)|^\alpha(1+\log_+|q(j,k)|^{-1}).$$

Then there is a constant D_α such that

(3.13) $$P(|Q_n|>\lambda)\le D_\alpha\lambda^{-\alpha}(1+\log_+\lambda)N_\alpha^{(n)}.$$

Moreover, if $N_\alpha = \lim_{n\to\infty}N_\alpha^{(n)} < \infty$, then the sequence $\{Q_n\}$ converges in L^p, for any $0<p<\alpha$, to the limit

(3.14) $$Q = \sum_{\substack{j,k=1\\ j\neq k}}^{\infty} q(j,k)Z_jZ_k$$

which satisfies

(3.15) $$P(|Q|>\lambda)\le D_\alpha\lambda^{-\alpha}(1+\log_+\lambda)N_\alpha.$$

By $\log_+ x$ we mean $\log x$ if $x \geq 1$ and 0 otherwise. Combining this proposition with the inequality

$$x^\alpha(1 + \log_+ \frac{1}{x}) < (\alpha - \mu)^{-1} x^\mu, \tag{3.16}$$

valid for $1 < \mu < \alpha$ and $0 < x < 1$, we get

Corollary 3.1 *Suppose p is a real number. If for some $1 < \mu < \alpha$,*

$$\lim_{n\to\infty} \sum_{\substack{j,k \\ j\neq k}} |n^p q_n(j,k)|^\mu = 0,$$

then

$$n^p \sum_{\substack{j,k \\ j\neq k}} q_n(j,k) Z_j Z_k \xrightarrow{P} 0 \quad \text{as } n \to \infty.$$

The next proposition is a direct consequence of Theorem 3.3 of Davis and Resnick (1986).

Proposition 3.3 *Suppose the Z_n and the $Y_k,\ k \geq 0$ are as in Theorem 1.2. Then, for any $m > 0$, as $n \to \infty$,*

$$(n^{-2/\alpha} \sum_{t=1}^{n} Z_t^2,\ (n\log n)^{-1/\alpha} \sum_{t=1}^{n-1} Z_t Z_{t+1}, \ldots, (n\log n)^{-1/\alpha} \sum_{t=1}^{n-m} Z_t Z_{t+m}) \tag{3.17}$$

$$\xrightarrow{\mathcal{D}} (C_\alpha^{2/\alpha}\sigma^2)\ (Y_0, Y_1, \ldots, Y_m).$$

We shall often use the following lemma.

Lemma 3.1 *Suppose the c_j are defined by* (1.5)*. Then for $0 < |\lambda| < \pi$ and any integers $n_1 < n_2$*

$$\left|\sum_{j=n_1}^{n_2} c_j e^{i\lambda j}\right| \leq K n_1^{d-1} |\lambda|^{-1} \tag{3.18}$$

and

$$\left|\sum_{j=n_1}^{n_2} c_j e^{i\lambda j}\right| \leq K |\lambda|^{-d}, \tag{3.19}$$

where K does not depend on n_1, n_2 and λ.

Proof: Inequality (3.18) follows immediately from the fact that $\lim_{j\to\infty} c_j/j^{d-1}$ exists (cf. Bingham, Goldie and Teugels (1987) p. 208). Inequality (3.19) can be proved by modifying slightly the proof of Theorem 2.6 and using inequalities 2.26 and 2.27 on p. 191 of Zygmund (1979). ∎

3.2 Proof of Proposition 3.1

Proposition 3.1 follows from a number of lemmas which are proved below, some of which are of independent interest.

Lemma 3.2 *Consider the function η defined in Proposition 3.1 and set*

$$\chi(\lambda) = \eta(\lambda) g(\lambda, \beta_0)$$

and

$$f_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i\lambda k} \chi(\lambda) d\lambda.$$

Then $f_k = O(|k|^{-1})$ as $k \to \infty$.

PROOF: By (1.11),

$$\begin{aligned}
\chi(\lambda) &= \left(\sum_{j=1}^{p+q+1} u_j \frac{\partial g^{-1}(\lambda,\beta_0)}{\partial \beta_j}\right) g(\lambda,\beta_0) \\
&= -\sum_{j=1}^{p+q+1} u_j \frac{\partial}{\partial \beta_j} \log g(\lambda,\beta_0) \\
&= -\sum_{j=1}^{p+q+1} u_j \frac{\partial}{\partial \beta_j} \left\{ \log\left|\frac{\Theta(e^{-i\lambda},\beta_0)}{\Phi(e^{-i\lambda},\beta_0)}\right|^2 - 2d(\beta_0)\log|1-e^{-i\lambda}| \right\}.
\end{aligned}$$

Thus, we can write

$$\chi(\lambda) = \chi_1(\lambda) + A(\beta_0)\log|2\sin\frac{\lambda}{2}|, \tag{3.20}$$

where

$$\chi_1(\lambda) = -\sum_{j=1}^{p+q+1} u_j \frac{\partial}{\partial \beta_j} \log\left|\frac{\Theta(e^{-i\lambda},\beta_0)}{\Phi(e^{-i\lambda},\beta_0)}\right|^2 \tag{3.21}$$

and

$$A(\beta_0) = 2\sum_{j=0}^{p+q+1} u_j \frac{\partial}{\partial \beta_j} d(\beta_0). \tag{3.22}$$

By (3.20), $f_k = f_{1k} + A(\beta_0) f_{2k}$, where f_{1k} is the Fourier coefficient of χ_1 and

$$f_{2k} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\lambda k}\log\left|2\sin\frac{\lambda}{2}\right| d\lambda. \tag{3.23}$$

As the function χ_1 has continuous derivative on $[-\pi,\pi]$, we have $f_{1k} = O(|k|^{-1})$ and so it remains to show that $f_{2k} = O(|k|^{-1})$. Integrating the right hand side of (3.23) by parts and setting $\lambda = 2\mu$, we get

$$f_{2k} = -\frac{1}{\pi k}\int_0^{\pi/2} \frac{\sin(2k\mu)\cos\mu}{\sin\mu} d\mu. \tag{3.24}$$

Consequently, we must show that the integrals

$$I_j = \int_0^{\pi/2} \frac{\sin(j\mu)\cos\mu}{\sin\mu} d\mu$$

are uniformly bounded in j. To verify this, decompose I_j as $I_j = I_{1j} + I_{2j}$, where

$$I_{1j} = \int_0^{\pi/2} \frac{\sin(j\mu)\cos\mu}{\mu} d\mu$$

and

$$I_{2j} = \int_0^{\pi/2} \sin(j\mu)\cos\mu\left(\frac{1}{\sin\mu} - \frac{1}{\mu}\right) d\mu.$$

Since $|1/\sin\mu - 1/\mu| = O(\mu)$, as $\mu \to 0$, the sequence $\{I_{2j}\}$ is bounded. To see that the sequence $\{I_{1j}\}$ is bounded, observe that

$$I_{1j} = \frac{1}{2}\left\{\int_0^{\pi/2} \frac{\sin(j+1)\mu}{\mu} d\mu + \int_0^{\pi/2} \frac{\sin(j-1)\mu}{\mu} d\mu\right\}$$

and

$$\int_0^{\pi/2} \frac{\sin(j\mu)}{\mu} d\mu = \int_0^{\pi j/2} \frac{\sin x}{x} dx \to \int_0^{\infty} \frac{\sin x}{x} dx,$$

as $j \to \infty$, the last integral converging conditionally. ∎

Lemma 3.2 is used in the proof of Proposition 3.4 below, in which $\tilde{I}_{n,Z}(\lambda)$ denotes the self-normalized periodogram of the process $\{Z_t\}$, i.e.

$$\tilde{I}_{n,Z}(\lambda) = \left(\sum_{t=1}^{n} Z_t^2\right)^{-1} \left|\sum_{t=1}^{n} Z_t e^{-i\lambda t}\right|^2. \tag{3.25}$$

Proposition 3.4 *Suppose the Z_n and the Y_k, $k \geq 0$, are as in Theorem 1.2 and the function χ and the sequence $\{f_k\}$ are as in Lemma 3.2. Then*

$$\left(\frac{n}{\log n}\right)^{1/\alpha} \int_{-\pi}^{\pi} \tilde{I}_{n,Z}(\lambda)\chi(\lambda)d\lambda \xrightarrow{\mathcal{D}} 4\pi \sum_{k=1}^{\infty} \frac{Y_k}{Y_0} f_k, \tag{3.26}$$

as $n \to \infty$.

Proof: This proof follows closely that of Lemma 6.3 of Mikosch *et al.* (1995); we use here $\alpha > 1$ and Lemma 3.2 rather than the assumption $\sum |f_k|^\mu < \infty$ for some $\mu \in (0, 1 \wedge \alpha)$, which holds in the ARMA case.

By Theorem 4.2 of Billingsley (1968), Relation (3.26) will follow, once we have verified that, as $m \to \infty$,

$$2\pi \sum_{|k|\leq m} \frac{Y_k}{Y_0} f_k \xrightarrow{P} 4\pi \sum_{k=1}^{\infty} \frac{Y_k}{Y_0} f_k, \tag{3.27}$$

and, as $n \to \infty$,

$$\left(\frac{n}{\log n}\right)^{1/\alpha} \int_{-\pi}^{\pi} \tilde{I}_{n,Z}(\lambda)\chi_m(\lambda)d\lambda \xrightarrow{\mathcal{D}} 2\pi \sum_{|k|\leq m} \frac{Y_k}{Y_0} f_k, \tag{3.28}$$

for each fixed m, where $\chi_m(\lambda) = \sum_{|k|\leq m} f_k e^{i\lambda k}$, and

$$\lim_{m\to\infty} \limsup_{n\to\infty} P\left\{\left(\frac{n}{\log n}\right)^{1/\alpha} \left|\int_{-\pi}^{\pi} \tilde{I}_{n,Z}(\lambda)(\chi(\lambda) - \chi_m(\lambda))d\lambda\right| > \epsilon\right\} = 0, \tag{3.29}$$

for any $\epsilon > 0$.

In order to verify (3.27) observe first that

$$\int_{-\pi}^{\pi} \log g(\lambda, \beta) d\lambda = 0, \quad \forall \beta \in E. \tag{3.30}$$

Relation (3.30) follows from the remark on p. 520 of Fox and Taqqu (1986) and the fact that $g(\lambda,\beta) = 2\pi f(\lambda,\beta)$, where $f(\lambda,\beta)$ is the spectral density of the Gaussian fractional ARIMA. Using Condition (A.1) of Fox and Taqqu (1986) with g in place of f, we get, by (3.30),

$$f_0 = \sum_{j=1}^{p+q+1} u_j \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\partial}{\partial\beta_j} \log g^{-1}(\lambda,\beta_0)d\lambda = \sum_{j=1}^{p+q+1} u_j \frac{\partial}{\partial\beta_j}\left(-\frac{1}{2\pi}\int_{-\pi}^{\pi} \log g(\lambda,\beta_0)d\lambda\right) = 0.$$

In view of the above relation and the fact that $f_k = f_{-k}$, it remains to show that

$$\sum_{|k|>m} \frac{Y_k}{Y_0} f_k \xrightarrow{P} 0, \quad \text{as } m\to\infty. \tag{3.31}$$

Since Y_0 is independent of the remaining Y_k's, it suffices to verify that $\sum_{|k|>m} f_k Y_k \xrightarrow{P} 0$. The latter relation follows from the Three Series Theorem and Lemma 3.2.

Direct verification, moreover, shows that

$$\left(\frac{n}{\log n}\right)^{1/\alpha} \int_{-\pi}^{\pi} \tilde{I}_{n,Z}(\lambda)\chi_m(\lambda)d\lambda = 2\pi\left(n^{-2/\alpha}\sum_{t=1}^{n} Z_t^2\right)^{-1} \sum_{|k|\le m} f_k \left[(n\log n)^{-1/\alpha}\sum_{t=1}^{n-|k|} Z_t Z_{t+|k|}\right],$$

so Relation (3.28) follows from Proposition 3.3.

Finally, to verify (3.29), notice that for $m<n$,

$$\begin{aligned}\int_{-\pi}^{\pi} \tilde{I}_{n,Z}(\lambda)(\chi(\lambda)-\chi_m(\lambda))d\lambda &= \int_{-\pi}^{\pi}\left(\sum_{|h|<n}\rho_{n,Z}(h)e^{-i\lambda h}\right)\left(\sum_{|k|>m} f_k e^{i\lambda k}\right)d\lambda \\ &= \sum_{|h|<n}\rho_{n,Z}(h)\int_{-\pi}^{\pi}\left(\sum_{|k|>m} f_k e^{i\lambda(k-h)}\right)d\lambda \\ &= 2\pi\sum_{m<|h|<n}\rho_{n,Z}(h)f_h,\end{aligned} \tag{3.32}$$

where

$$\rho_{n,Z}(h) = \left(\sum_{t=1}^{n} Z_t^2\right)^{-1}\left(\sum_{t=1}^{n-|h|} Z_t Z_{t+h}\right). \tag{3.33}$$

(The last equality in (3.32) is justified by the fact that $\sum_{|k|>m}|f_k|^2<\infty$.) Equalities (3.32) and (3.33) and a change of indices yield

$$\left(\frac{n}{\log n}\right)^{1/\alpha}\int_{-\pi}^{\pi}\tilde{I}_{n,Z}(\lambda)\left(\chi(\lambda)-\chi_m(\lambda)\right)d\lambda \tag{3.34}$$

$$= 4\pi\left(n^{-2/\alpha}\sum_{t=1}^{n} Z_t^2\right)^{-1}(n\log n)^{-1/\alpha}\sum_{t=1}^{n-m-1} Z_t \sum_{h=m+t+1}^{n} f_{h-t}Z_h.$$

By Proposition 3.3, $n^{-2/\alpha}\sum_{t=0}^{n} Z_t^2 \xrightarrow{\mathcal{D}} Y_0$, so (3.28) will follow once we have verified that

$$\lim_{m\to\infty}\limsup_{n\to\infty} P\left\{(n\log n)^{-1/\alpha}\left|\sum_{t=1}^{n-m-1} Z_t \sum_{h=m+t+1}^{n} f_{h-t}Z_h\right| > \epsilon\right\} = 0. \tag{3.35}$$

By (3.13) and the inequality (3.16), the probability in (3.35) is bounded above by

$$K_{\alpha,\mu}\epsilon^{-\alpha}n^{-1}\sum_{t=1}^{n-m-1}\sum_{h=m+t+1}^{h} |f_{h-t}|^{\mu}. \tag{3.36}$$

Since, by Lemma 3.2,

$$n^{-1}\sum_{t=1}^{n-m-1}\sum_{h=m+t+1}^{n} |f_{h-t}|^{\mu} \le n^{-1}\sum_{t=1}^{n}\sum_{j=m+1}^{\infty} |f_j|^{\mu} \le Km^{1-\mu},$$

we see that (3.35) holds. ∎

Our next goal is to establish a relationship between the right-hand side of (3.2) and the right-hand side of (3.26). The first step in this direction is to relate the sample variances of the processes $\{X_t\}$ and $\{Z_t\}$.

Lemma 3.3 *Suppose the Z_n, c_j, X_n and a_n are as in Theorem 2.1. Then*

$$a_n^{-2}\sum_{k=1}^{n} X_k^2 = \left(\sum_{j=0}^{\infty} c_j^2\right) a_n^{-2}\sum_{k=1}^{n} Z_k^2(1+o_P(1)). \tag{3.37}$$

PROOF: We only sketch the proof since it is similar to the proofs in Davis and Resnick (1985) and (1986) which rely on point processes techniques. Let

$$\lambda(dx) = \alpha p x^{-\alpha-1}\mathbf{1}_{(0,\infty)}(x)dx + \alpha q(-x)^{-\alpha-1}\mathbf{1}_{(-\infty,0)}(x)dx$$

be the Lévy measure of a stable random variable and set $\mu(dt,dx) = dt\times\lambda(dx)$, $t>0$, where dt stands for the Lebesgue measure. If $\sum_{k=1}^{\infty}\epsilon_{(t_k,j_k)}$ is the Poisson random measure with mean measure μ, then Theorem 2.2 of Davis and Resnick (1985) asserts that

$$\sum_{k=1}^{\infty}\epsilon_{(k/n,\ a_n^{-1}\mathbf{Z}^{(k)})} \Rightarrow \sum_{k=1}^{\infty}\sum_{i=1}^{m}\epsilon_{(t_k,\ j_k\mathbf{e}_i)}, \tag{3.38}$$

where "$\Rightarrow$" denotes the weak convergence of random measures on $(0,\infty)\times\mathbb{R}^m\backslash\{0,\cdots,0\}$, $\mathbf{Z}^{(k)} = (Z_k, Z_{k-1},\ldots,Z_{k-m+1})$, and where $\mathbf{e}_i\in\mathbb{R}^m$ is the basis element with ith component equal to one and the rest zero. Instead of applying, as in Davis and Resnick (1985), the continuous map $(z_k, z_{k-1},\ldots,z_{k-m+1})\mapsto\sum_{i=0}^{m-1}c_i z_{k-i}$ to both sides of (3.38), we shall apply the continuous map $(z_k, z_{k-1},\ldots,z_{k-m+1})\mapsto(\sum_{i=0}^{m-1}c_i z_{k-i}, z_k)$. Thus using Theorems 4.2 and 5.1 of Billingsley (1968), and the arguments of the proof of Theorem 2.4 of Davis and Resnick, we get

$$\sum_{k=1}^{n}\epsilon_{(a_n^{-1}(X_k,\ Z_k))} \Rightarrow \sum_{k=1}^{\infty}\left(\epsilon_{(c_o j_k,\ j_k)} + \sum_{i=1}^{\infty}\epsilon_{(c_i j_k,\ 0)}\right). \tag{3.39}$$

where now the space is $\mathbb{R}^2 \backslash \{0,0\}$. Proceeding as in the proof of Theorem 4.2 of Davis and Resnick (1985), it can be verified that (3.39) yields

$$a_n^{-2} \sum_{k=1}^{n} (X_k^2, Z_k^2) \xrightarrow{\mathcal{D}} \left(\sum_{j=0}^{\infty} c_j^2, 1 \right) S, \tag{3.40}$$

where the $\alpha/2$-stable random variable S is as in Theorem 2.1. Finally, applying the continuous map $h(u,v) = (u - (\sum_{j=0}^{\infty} c_j^2)v)/v$ to both sides of (3.40), we obtain (3.37). ∎

Remark. Lemma 3.3 extends Lemma 5.2 of Mikosch *et al.*, which was proved under the assumption (1.6). We have shown that Condition (2.1) is sufficient for (3.37) to hold.

Lemma 3.3 yields the following useful corollary:

Corollary 3.2 *For any fractional ARIMA process whose innovations satisfy* (1.2) *and* (1.3),

$$\begin{aligned} \int_{-\pi}^{\pi} \tilde{I}_n(\lambda)\eta(\lambda)d\lambda &= \gamma(0)^{-1}(1+o_P(1)) \int_{-\pi}^{\pi} \tilde{I}_{n,Z}(\lambda) g(\lambda,\beta_0)\eta(\lambda)d\lambda \\ &\quad + \gamma(0)^{-1}(1+o_P(1)) \left(a_n^{-2} \sum_{k=1}^{n} Z_k^2 \right)^{-1} \int_{-\pi}^{\pi} R_n(\lambda)\eta(\lambda)d\lambda, \end{aligned} \tag{3.41}$$

where $\gamma(0) = \sum_{j=0}^{\infty} c_j^2$ *and*

$$\begin{aligned} R_n(\lambda) &= |Y_n(\lambda)|^2 \\ &+ Y_n(\lambda) \left(\sum_{j=0}^{\infty} c_j e^{i\lambda j} \right) \left(a_n^{-1} \sum_{k=1}^{n} Z_k e^{i\lambda k} \right) + Y_n(-\lambda) \left(\sum_{j=0}^{\infty} c_j e^{-i\lambda j} \right) \left(a_n^{-1} \sum_{k=1}^{n} Z_k e^{-i\lambda k} \right), \end{aligned} \tag{3.42}$$

and where

$$Y_n(\lambda) = a_n^{-1} \sum_{j=0}^{\infty} c_j e^{-i\lambda j} \left(\sum_{k=1-j}^{n-j} Z_k e^{-i\lambda k} - \sum_{k=1}^{n} Z_k e^{-i\lambda k} \right). \tag{3.43}$$

PROOF: Defining

$$I_{n,X}(\lambda) = \frac{1}{a_n^2} \left| \sum_{t=1}^{n} X_t e^{-i\lambda t} \right|^2 \quad \text{and} \quad I_{n,Z}(\lambda) = \frac{1}{a_n^2} \left| \sum_{t=1}^{n} Z_t e^{-i\lambda t} \right|^2,$$

it can be verified identically as in the case of finite variance ARMA processes (cf. e.g. the proof of Theorem 10.3.1 of Brockwell and Davis (1991)) that

$$I_{n,X}(\lambda) = \left| \sum_{j=0}^{\infty} c_j e^{-i\lambda j} \right|^2 I_{n,Z}(\lambda) + R_n(\lambda). \tag{3.44}$$

Now, using Lemma 3.3, we have

$$\int_{-\pi}^{\pi} \tilde{I}_n(\lambda)\eta(\lambda)d\lambda = \int_{-\pi}^{\pi} \left(\sum_{k=1}^{n} X_k^2 \right)^{-1} \left| \sum_{k=1}^{n} X_k e^{-i\lambda k} \right|^2 \eta(\lambda)d\lambda$$

$$
\begin{aligned}
&= \left(a_n^{-2}\sum_{k=1}^n X_k^2\right)^{-1}\int_{-\pi}^{\pi} I_{n,X}(\lambda)\eta(\lambda)d\lambda \\
&= \left[\gamma(0)a_n^{-2}\sum_{k=1}^n Z_k^2(1+o_P(1))\right]^{-1} \times \\
&\quad \int_{-\pi}^{\pi}\left[g(\lambda,\beta_0)I_{n,Z}(\lambda)+R_n(\lambda)\right]\eta(\lambda)d\lambda \\
&= \gamma(0)^{-1}(1+o_P(1))\int_{-\pi}^{\pi}\tilde{I}_{n,Z}(\lambda)g(\lambda,\beta_0)\eta(\lambda)d\lambda \\
&\quad + \gamma(0)^{-1}(1+o_P(1))\left(a_n^{-2}\sum_{k=1}^n Z_k^2\right)^{-1}\int_{-\pi}^{\pi} R_n(\lambda)\eta(\lambda)d\lambda. \quad \blacksquare
\end{aligned}
$$

Since $g(\lambda,\beta_0)\eta(\lambda)=\chi(\lambda)$ and $a_n^{-2}\sum_{k=1}^n Z_k^2 \stackrel{\mathcal{D}}{\to} S$, by (3.40), we obtain

Corollary 3.3 *The relation*

$$
\left(\frac{n}{\log n}\right)^{1/\alpha}\int_{-\pi}^{\pi} R_n(\lambda)\eta(\lambda)d\lambda \stackrel{P}{\to} 0, \quad \text{as} \quad n\to\infty \tag{3.45}
$$

together with Proposition 3.4 will imply Relation (3.2) *and hence will complete the proof of Proposition 3.1.*

Therefore, it remains to establish (3.45). In view of (3.42) and the fact that under the assumptions of Theorem 1.2, a_n is proportional to $n^{1/\alpha}$, Relation (3.45) will follow once we have proved the following two Lemmas:

Lemma 3.4 *Setting* $Y_n(\lambda)=a_n^{-1}\Delta_n(\lambda)$, *we have,*

$$
(n\log n)^{-1/\alpha}\int_{-\pi}^{\pi}|\Delta_n(\lambda)|^2\eta(\lambda)d\lambda \stackrel{P}{\to} 0, \quad \text{as} \quad n\to\infty. \tag{3.46}
$$

Lemma 3.5 *Setting* $C(\lambda):=\sum_{j=0}^{\infty} c_j e^{i\lambda j}$, *we have*

$$
(n\log n)^{-1/\alpha}\int_{-\pi}^{\pi}\Delta_n(\lambda)C(\lambda)\left(\sum_{k=1}^n Z_k e^{i\lambda k}\right)\eta(\lambda)d\lambda \stackrel{P}{\to} 0, \quad \text{as} \quad n\to\infty. \tag{3.47}
$$

The proof of these lemmas involves delicate bounds where the assumption $0<d<1-1/\alpha$ plays a crucial role.

3.3 Proof of the two lemmas

Proof of Lemma 3.4: We shall show that

$$
n^{-1/\alpha}\int_{-\pi}^{\pi}|\Delta_n(\lambda)|^2\eta(\lambda)d\lambda \stackrel{P}{\to} 0. \tag{3.48}
$$

Observe that $\Delta_n(\lambda)$ can be conveniently split into four sums as follows:

$$\Delta_n(\lambda) = \sum_{j=0}^{\infty} c_j e^{-i\lambda j} \left[\sum_{k=1-j}^{n-j} Z_k e^{-i\lambda k} - \sum_{k=1}^{n} Z_k e^{-i\lambda k} \right] \tag{3.49}$$

$$= \Gamma_{1n}(\lambda) + \Gamma_{2n}(\lambda) + \Gamma_{3n}(\lambda) + \Gamma_{4n}(\lambda),$$

where

$$\Gamma_{1n}(\lambda) = \sum_{k=0}^{n-2} \left(\sum_{j=k+1}^{n+k} c_j e^{i(k-j)\lambda} \right) Z_{-k}, \tag{3.50}$$

$$\Gamma_{2n}(\lambda) = \sum_{k=n-1}^{\infty} \left(\sum_{j=k+1}^{n+k} c_j e^{i(k-j)\lambda} \right) Z_{-k}, \tag{3.51}$$

$$\Gamma_{3n}(\lambda) = -e^{-in\lambda} \sum_{k=0}^{n-2} \left(\sum_{j=k+1}^{n-1} c_j e^{i(k-j)\lambda} \right) Z_{n-k}, \tag{3.52}$$

$$\Gamma_{4n}(\lambda) = - \left(\sum_{j=n}^{\infty} c_j e^{-i\lambda j} \right) \sum_{t=1}^{n} Z_t e^{-i\lambda t}. \tag{3.53}$$

Clearly, (3.48) will follow once we have verified that

$$n^{-1/\alpha} \int_{-\pi}^{\pi} |\Gamma_{un}(\lambda)|^2 |\eta(\lambda)| d\lambda \xrightarrow{P} 0, \tag{3.54}$$

for $u = 1, 2, 3, 4$.

1) We first verify (3.54) with $u = 1$. Write

$$\int_{-\pi}^{\pi} |\Gamma_{1n}(\lambda)|^2 |\eta(\lambda)| d\lambda = \sum_{t=0}^{n-2} Z_{-t}^2 \left(\int_{-\pi}^{\pi} \left| \sum_{j=t+1}^{n+t} c_j e^{i\lambda j} \right|^2 |\eta(\lambda)| d\lambda \right) \tag{3.55}$$

$$+ \sum_{\substack{t,k=0 \\ t \neq k}}^{n-2} Z_{-t} Z_{-k} \int_{-\pi}^{\pi} \left(\sum_{j=k+1}^{n+k} c_j e^{i(k-j)\lambda} \right) \left(\sum_{j=t+1}^{n+t} c_j e^{i(j-t)\lambda} \right) |\eta(\lambda)| d\lambda$$

$$=: \sum_{t=0}^{n-2} \nu_n(t) Z_{-t}^2 + \sum_{\substack{t,k=0 \\ t \neq k}}^{n-2} \kappa_n(k,t) Z_{-t} Z_{-k}.$$

In order to establish upper bounds on the coefficients $\nu_n(t)$ and $\kappa_n(k,t)$, observe that by Condition (A.3) on p. 521 of Fox and Taqqu (1986), for any $\delta > 0$

$$|\eta(\lambda)| = O(|\lambda|^{2d-\delta}), \quad \text{as} \quad \lambda \to 0. \tag{3.56}$$

Relation (3.56) and the two inequalities of Lemma 3.1 yield

$$|\nu_n(t)| \leq K_1 t^{d-1} \int_{-\pi}^{\pi} |\lambda|^{-d-1+2d-\delta} d\lambda \leq K_2 t^{d-1}, \tag{3.57}$$

if δ is small enough. The same argument shows that

$$|\kappa_n(k,t)| \le K_2(t \vee k)^{d-1}. \tag{3.58}$$

Using (3.57), we obtain, for $1 < \mu < \alpha$,

$$\begin{aligned} P\left\{ n^{-1/\alpha} \sum_{t=0}^{n-2} \nu_n(t) Z_{-t}^2 > \epsilon \right\} &\le \epsilon^{-\mu/2} n^{-\mu/2\alpha} E \left| \sum_{t=0}^{n-2} \nu_n(t) Z_{-t}^2 \right|^{\mu/2} \\ &\le \epsilon^{-\mu/2} n^{-\mu/2\alpha} E \sum_{t=0}^{n-2} |\nu_n(t)|^{\mu/2} |Z_{-t}|^{\mu} \\ &\le \epsilon^{-\mu/2} E|Z_1|^{\mu} K_2^{\mu/2} n^{-\mu/2\alpha} \sum_{t=0}^{n-2} t^{(d-1)\mu/2} \\ &= O\left(n^{-\frac{\mu}{2\alpha} + \frac{(d-1)\mu}{2} + 1} \right). \end{aligned} \tag{3.59}$$

Observe that the condition $d < 1 - 1/\alpha$ implies that the exponent $-\frac{\mu}{2\alpha} + \frac{(d-1)\mu}{2} + 1$ is negative for μ sufficiently close to α. Therefore

$$n^{-1/\alpha} \sum_{t=0}^{n-2} \nu_n(t) Z_{-t}^2 \xrightarrow{P} 0. \tag{3.60}$$

By Corollary 3.1, to show

$$n^{-1/\alpha} \sum_{\substack{t,k=1 \\ t \ne k}}^{n-2} \kappa_n(k,t) Z_{-t} Z_{-k} \xrightarrow{P} 0 \tag{3.61}$$

it suffices to verify that for some $1 < \mu < \alpha$,

$$\sum_{1 \le t < k \le n-2} (n^{-1/\alpha} k^{(d-1)})^{\mu} \to 0, \quad \text{as} \quad n \to \infty. \tag{3.62}$$

Note that the left-hand side of (3.62) is bounded above by a multiple of

$$n^{-\mu/\alpha} \int_1^n \left(\int_t^n k^{(d-1)\mu} dk \right) dt = O(n^{(d-1)\mu + 2 - \mu/\alpha}),$$

which tends to 0 as $n \to \infty$ for μ sufficiently close to α. Indeed if $\mu = \alpha$, the exponent of n becomes $-1 + (d-1)\alpha + 2 = (d-1)\alpha + 1$ which is negative since $d < 1 - 1/\alpha$.

Relations (3.60), (3.61) and (3.55) prove (3.54) with $u = 1$.

2) To verify (3.54) with $u = 2$, we must show that

$$n^{-1/\alpha} \sum_{t=n-1}^{\infty} \nu_n(t) Z_{-t}^2 \xrightarrow{P} 0 \tag{3.63}$$

and

$$n^{-1/\alpha} \sum_{\substack{t,k=n-1 \\ t \ne k}}^{\infty} \kappa_n(t,k) Z_{-t} Z_{-k} \xrightarrow{P} 0 \tag{3.64}$$

with the $\nu_n(t)$ and the $\kappa_n(k,t)$ as in the case $u=1$. Here, however, we need more delicate bounds than (3.57) and (3.58). Write

$$|\kappa_n(k,t)| \le 2(I_1 + I_2),$$

where

$$I_1 = \int_0^{\pi/n} \Big| \sum_{j=t+1}^{n+t} c_j e^{i\lambda j} \Big| \Big| \sum_{j=k+1}^{n+k} c_j e^{i\lambda j} \Big| |\eta(\lambda)| d\lambda; \tag{3.65}$$

$$I_2 = \int_{\pi/n}^{\pi} \Big| \sum_{j=t+1}^{n+t} c_j e^{i\lambda j} \Big| \Big| \sum_{j=k+1}^{n+k} c_j e^{i\lambda j} \Big| |\eta(\lambda)| d\lambda. \tag{3.66}$$

Note that by (3.56), for sufficiently small $\delta > 0$,

$$I_1 = O\left(\int_0^{\pi/n} n^2 t^{d-1} k^{d-1} \lambda^{2d-\delta} d\lambda \right) = O(t^{d-1} k^{d-1} n^{1-2d+\delta}). \tag{3.67}$$

To establish an upper bound on I_2, set $x = n\lambda$ in (3.66). Then,

$$I_2 = O\left(t^{d-1} k^{d-1} n^{1-2d+\delta} \int_{\pi}^{n\pi} x^{2d-2-\delta} dx \right) = O(t^{d-1} k^{d-1} n^{1-2d+\delta}). \tag{3.68}$$

Combining (3.67) and (3.68), we get

$$\kappa_n(k,t) = O(t^{d-1} k^{d-1} n^{1-2d+\delta}) \tag{3.69}$$

and, in particular,

$$\nu_n(t) = O(t^{2(d-1)} n^{1-2d+\delta}). \tag{3.70}$$

Now we verify (3.63). By (3.70), for $1 < \mu < \alpha$, we have,

$$\begin{aligned} P\left\{ n^{-1/\alpha} \Big| \sum_{t=n-1}^{\infty} \nu_n(t) Z_{-t}^2 \Big| > \epsilon \right\} &\le (\epsilon n^{1/\alpha})^{-\mu/2} E \Big| \sum_{t=n-1}^{\infty} \nu_n(t) Z_{-t}^2 \Big|^{\mu/2} \\ &\le K_n^{-\mu/2\alpha} \sum_{t=n-1}^{\infty} (t^{2(d-1)} n^{1-2d+\delta})^{\mu/2} \\ &= O(n^{-\mu/(2\alpha)+(d-1)\mu+1+(1-2d+\delta)\mu/2}). \end{aligned} \tag{3.71}$$

Notice that if $\delta = 0$, and $\mu = \alpha$, the exponent is equal to $-\frac{1}{2} + (d-1)\alpha + 1 + (1-2d)\alpha/2 = (1-\alpha)/2$ and is negative iff $\alpha > 1$. This completes the proof of (3.63).

To prove (3.64), we use Corollary 3.1. We have

$$\begin{aligned} \sum_{\substack{t,k=n-1 \\ t\neq k}}^{\infty} |n^{-1/\alpha} \kappa_n(k,t)|^{\mu} &= O\left(\int_n^{\infty} \left(\int_n^k |n^{-1/\alpha} t^{d-1} k^{d-1} n^{1-2d+\delta}|^{\mu} dt \right) dk \right) \\ &= O\left(n^{(1+\delta-1/\alpha-2d)\mu} \int_n^{\infty} k^{(d-1)\mu} \left(\int_n^{\infty} t^{(d-1)\mu} dt \right) dk \right) \\ &= O(n^{(1+\delta-1/\alpha-2d)\mu+2(d-1)\mu+2}), \end{aligned} \tag{3.72}$$

which again tends to 0 as $n \to \infty$ for μ sufficiently close to α. This completes the proof of (3.64) and (3.54) with $u = 2$.

3) The proof of (3.54) with $u = 3$ is the same as in the case $u = 1$.

4) The case $u = 4$ is dealt with similarly as the previous three cases. Write

$$\int_{-\pi}^{\pi} |\Gamma_{4n}(\lambda)|^2 |\eta(\lambda)| d\lambda = \nu_n \sum_{t=1}^{n} Z_t^2 + \sum_{1 \le t \ne k \le n} \kappa_n(t,k) Z_t Z_k,$$

where now

$$\nu_n = \int_{-\pi}^{\pi} \left| \sum_{j=n}^{\infty} c_j e^{-i\lambda j} \right|^2 |\eta(\lambda)| d\lambda$$

and

$$\kappa_n(k,t) = \int_{-\pi}^{\pi} \left| \sum_{j=n}^{\infty} c_j e^{-i\lambda j} \right|^2 |\eta(\lambda)| e^{i\lambda(k-t)} d\lambda.$$

By Lemma 3.1 and (3.56) we obtain

$$\nu_n = O(n^{d-1}) \quad \text{and} \quad \kappa_n(t,k) = O(n^{d-1}).$$

Consequently, Proposition 3.3 and the condition $d < 1 - 1/\alpha$ imply

$$n^{-1/\alpha} \sum_{t=1}^{n} \nu_n Z_t^2 \le K n^{(1/\alpha)+d-1} \left(n^{-2/\alpha} \sum_{t=1}^{n} Z_t^2 \right) \xrightarrow{P} 0. \tag{3.73}$$

The relation

$$n^{-1/\alpha} \sum_{1 \le t < k \le n} \kappa_n(k,t) Z_t Z_k \xrightarrow{P} 0 \tag{3.74}$$

follows, by Corollary 3.1, from the relation

$$\sum_{1 \le t < k \le n} |n^{-1/\alpha} n^{d-1}|^{\mu} \to 0,$$

which is easily seen to hold for μ sufficiently close to α.

This proves (3.54) with $u = 4$, completing the proof of Lemma 3.4. ∎

Proof of Lemma 3.5: Here it is convenient to split $\Delta_n(\lambda)$ as follows:

$$\Delta_n(\lambda) = \Delta_{1n}(\lambda) + \Delta_{2n}(\lambda) + \Delta_{3n}(\lambda),$$

where

$$\begin{aligned}
\Delta_{1n}(\lambda) &= \Gamma_{1n}(\lambda) + \Gamma_{2n}(\lambda) = \sum_{t=1}^{n-1} e^{-i\lambda t} \left(\sum_{j=t}^{\infty} c_j Z_{t-j} \right), \\
\Delta_{2n}(\lambda) &= \Gamma_{3n}(\lambda) = -e^{-i\lambda n} \sum_{t=1}^{n-1} e^{-i\lambda t} \left(\sum_{j=t}^{n-1} c_j Z_{n+t-j} \right), \\
\Delta_{3n}(\lambda) &= \Gamma_{4n}(\lambda) = -\left(\sum_{j=n}^{\infty} c_j e^{-i\lambda j} \right) \sum_{t=1}^{n} Z_t e^{-i\lambda t}.
\end{aligned}$$

We verify below that

$$n^{-1/\alpha}\int_{-\pi}^{\pi}\Delta_{un}(\lambda)\left(\sum_{k=1}^{n}Z_k e^{i\lambda k}\right)C(\lambda)\eta(\lambda)d\lambda \xrightarrow{P} 0, \tag{3.75}$$

for $u = 1, 2, 3$.

To prove (3.75) for $u = 1$ and $u = 2$, we need the following lemma.

Lemma 3.6 *Suppose $C(\lambda) = \sum_{j=-\infty}^{\infty} c_j e^{i\lambda j}$ and $E(\lambda) = \sum_{k=-\infty}^{\infty} e_k e^{i\lambda k}$ are functions on $[-\pi, \pi]$ whose Fourier coefficients satisfy*

$$c_j = O(|j|^{d-1}), \qquad e_k = O(|k|^{-\epsilon-1}), \tag{3.76}$$

for some $0 < d < 1/2$ and $\epsilon > 0$. Then the Fourier coefficients of the product CE satisfy

$$h_l := \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\lambda l}C(\lambda)E(\lambda)d\lambda = O(l^{d-1}). \tag{3.77}$$

PROOF: Conditions (3.76) imply that both C and E are in $L^2[-\pi,\pi]$, so $h_l = \sum_{j=-\infty}^{\infty} c_j e_{l-j}$. We have, for $l > 0$,

$$|h_l| \le \sum_{|j|\le l/2}|c_j||e_{l-j}| + \sum_{|j|>l/2}|c_j||e_{l-j}|.$$

By (3.76),

$$\begin{aligned}
\sum_{|j|\le l/2}|c_j||e_{l-j}| &= O\left(\sum_{|j|\le l/2}|j|^{d-1}|l-j|^{-\epsilon-1}\right)\\
&= O\left(\int_0^{l/2} j^{d-1}(l-j)^{-\epsilon-1}d_j\right)\\
&= O\left(l^{d-1-\epsilon}\int_0^{1/2} x^{d-1}(1-x)^{-\epsilon-1}dx\right)\\
&= O(l^{d-1-\epsilon}) = O(l^{d-1}),
\end{aligned}$$

and

$$\sum_{|j|>l/2}|c_j||e_{l-j}| = O\left(\sum_{|j|>l/2}|j|^{d-1}|e_{l-j}|\right) = O\left(l^{d-1}\sum_{k=-\infty}^{\infty}|e_k|\right) = O(l^{d-1}). \quad \blacksquare$$

Now introduce the function

$$H(\lambda) = C(\lambda)\eta(\lambda) = \sum_{j=1}^{p+q+1} u_j C(\lambda)\frac{\partial}{\partial\beta_j}g^{-1}(\lambda,\beta_0).$$

By Lemma 5 on p. 526 of Fox and Taqqu (1986), the Fourier coefficients of the functions $\frac{\partial}{\partial\beta_j}g^{-1}(\lambda,\beta_0)$ are $O(|k|^{-2d-1+\delta})$, so applying Lemma 3.6 (with $E = \eta$) we see that $H(\lambda) = \sum_{l=-\infty}^{\infty} h_l e^{i\lambda l}$, where

$$h_l = O(|l|^{d-1}). \tag{3.78}$$

Using (3.78) and the elementary identity

$$\left(\sum_{t=1}^{n-1} a_t e^{-i\lambda t}\right)\left(\sum_{k=1}^{n} b_k e^{i\lambda k}\right) = \sum_{s=1}^{n-1} A_s e^{i\lambda s} + \sum_{s=0}^{n-2} A_{-s} e^{-i\lambda s}, \tag{3.79}$$

where

$$A_s = \sum_{t=1}^{n-s} a_t b_{t+s}, \quad s > 0, \qquad \text{and} \qquad A_{-s} = \sum_{t=1}^{n-1-s} b_t a_{t+s}, \quad s \geq 0,$$

we shall now show that for $u = 1$ and $u = 2$,

$$n^{-1} \int_{-\pi}^{\pi} H(\lambda)\Delta_{un}(\lambda) \left(\sum_{k=1}^{n} Z_k e^{i\lambda k}\right) d\lambda \xrightarrow{P} O. \tag{3.80}$$

1) Consider first the case $u = 1$. By (3.79), $\Delta_{1n}(\lambda)(\sum_{k=1}^{n} Z_k e^{i\lambda k}) = \sum_{s=1}^{n-1} A_s e^{i\lambda s} + \sum_{s=0}^{n-2} A_{-s} e^{-i\lambda s}$, where

$$A_s = \sum_{t=1}^{n-s} \left(\sum_{j=t}^{\infty} c_j Z_{t-j}\right) Z_{t+s} \qquad \text{and} \qquad A_{-s} = \sum_{t=1}^{n-1-s} Z_t \left(\sum_{j=t+s}^{\infty} c_j Z_{t+s-j}\right).$$

Consequently,

$$\begin{aligned} \frac{1}{2\pi}\int_{-\pi}^{\pi} H(\lambda)\Delta_{1n}(\lambda)\left(\sum_{k=1}^{n} Z_k e^{i\lambda k}\right) d\lambda &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(\sum_{l=-\infty}^{\infty} h_l e^{i\lambda l}\right)\left(\sum_{s=1}^{n-1} A_s e^{i\lambda s} + \sum_{s=0}^{n-2} A_{-s} e^{-i\lambda s}\right) d\lambda \\ &= \sum_{s=1}^{n-1} h_{-s} A_s + \sum_{s=0}^{n-2} h_s A_{-s}. \end{aligned}$$

We will show that

$$n^{-1/\alpha} \sum_{s=1}^{n-1} h_{-s} A_s = n^{-1/\alpha} \sum_{s=1}^{n-1} h_{-s} \sum_{t=1}^{n-s} \left(\sum_{j=t}^{\infty} c_j Z_{t-j}\right) Z_{t+s} \xrightarrow{P} 0, \tag{3.81}$$

the verification of

$$n^{-1/\alpha} \sum_{s=0}^{n-2} h_s A_{-s} = n^{-1/\alpha} \sum_{s=0}^{n-2} h_s \sum_{t=1}^{n-1-s} Z_t \left(\sum_{j=t+s}^{\infty} c_j Z_{t+s-j}\right) \xrightarrow{P} 0 \tag{3.82}$$

being similar.

By Corollary 3.1, since the bilinear form in (3.81) has no diagonal elements, it suffices to verify that for some $\mu < \alpha$

$$n^{-\mu/\alpha} \sum_{s=1}^{n-1} |h_{-s}|^{\mu} \sum_{t=1}^{n-s} \sum_{j=t}^{\infty} |c_j|^{\mu} \to 0. \tag{3.83}$$

By (3.78), this reduces to showing

$$n^{-\mu/\alpha} \int_1^n s^{(d-1)\mu} ds \int_1^{n-s} dt \int_t^{\infty} j^{(d-1)\mu} dj \to 0. \tag{3.84}$$

An elementary computation shows that the LHS of (3.84) is $O(n^{-\mu/\alpha+(d-1)\mu+2})$. The condition $(d-1)\alpha+1<0$ guarantees that the exponent is negative for μ sufficiently close to α.

2) Now consider the case $u=2$. It is convenient to express $\Delta_{2n}(\lambda)$ as $\Delta_{2n}(\lambda)=-e^{-i\lambda n}\Delta'_{2n}(\lambda)$. By (3.79), $\Delta'_{2n}(\lambda)(\sum_{k=1}^{n} Z_k e^{i\lambda k})=\sum_{s=1}^{n-1} A_s e^{i\lambda s}+\sum_{s=0}^{n-2} A_{-s}e^{-i\lambda s}$, where

$$A_s=\sum_{t=1}^{n-s}\left(\sum_{j=t}^{n-1} c_j Z_{n+t-j}\right) Z_{t+s} \quad \text{and} \quad A_{-s}=\sum_{t=1}^{n-1-s} Z_t\left(\sum_{j=t+s}^{n-1} c_j Z_{n+t+s-j}\right).$$

Consequently,

$$\begin{aligned}
-\frac{1}{2\pi}\int_{-\pi}^{\pi} H(\lambda)\Delta_{2n}(\lambda)\left(\sum_{k=1}^{n} Z_k e^{i\lambda k}\right) d\lambda &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\lambda n} H(\lambda)\Delta'_{2n}(\lambda)\left(\sum_{k=1}^{n} Z_k e^{i\lambda k}\right) d\lambda \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{j=-\infty}^{\infty} h_{j+n}e^{i\lambda j}\right)\left(\sum_{s=1}^{n-1} A_s e^{i\lambda s}+\sum_{s=0}^{n-2} A_{-s}e^{-i\lambda s}\right) d\lambda \\
&= \sum_{s=1}^{n-1} h_{n-s}A_s+\sum_{s=0}^{n-2} h_{n+s}A_{-s}.
\end{aligned}$$

We first show that

$$n^{-1/\alpha}\sum_{s=1}^{n-1} h_{n-s}A_s = n^{-1/\alpha}\sum_{s=1}^{n-1} h_{n-s}\sum_{t=1}^{n-s}\left(\sum_{j=t}^{n-1} c_j Z_{n+t-j}\right) Z_{t+s} \xrightarrow{P} 0. \tag{3.85}$$

The bilinear form in (3.85) contains diagonal elements, so we prove separately that

$$n^{-1/\alpha}\sum_{s=1}^{n-1} h_{n-s}\sum_{t=1}^{n-s} c_{n-s}Z_{t+s}^2 \xrightarrow{P} 0 \tag{3.86}$$

and

$$n^{-1/\alpha}\sum_{s=1}^{n-1} h_{n-s}\sum_{t=1}^{n-1}\left(\sum_{\substack{j=t\\ j\neq n-s}}^{n-1} c_j Z_{n+t-j}\right) Z_{t+s} \xrightarrow{P} 0. \tag{3.87}$$

To prove (3.86), observe that

$$\begin{aligned}
\sum_{s=1}^{n-1} h_{n-s}\sum_{t=1}^{n-s} c_{n-s}Z_{t+s}^2 &= \sum_{j=1}^{n-1} h_j c_j \sum_{t=1}^{j} Z_{t+n-j}^2 \\
&= \sum_{m=2}^{n} Z_m^2\left(\sum_{j=n-m+1}^{n-1} h_j c_j\right) \\
&= \sum_{m=2}^{n} Z_m^2\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} H(\lambda)\left(\sum_{j=n-m+1}^{n-1} c_j e^{-i\lambda j}\right)\right\} \\
&=: \sum_{m=2}^{n} \nu_n(m) Z_m^2.
\end{aligned} \tag{3.88}$$

Since $H(\lambda) = C(\lambda)\eta(\lambda) = O(|\lambda|^{-d}|\lambda|^{2d-\delta}) = O(|\lambda|^{d-\delta})$ by (3.56), Inequality (3.18) yields $\nu_n(m) = O((n-m+1)^{d-1})$. Consequently, for any $\mu < \alpha$,

$$\begin{aligned}
E\left|n^{-1/\alpha}\sum_{m=2}^{n}\nu_n(m)Z_m^2\right|^{\mu/2} &\le n^{-\mu/2\alpha}E|Z_1|^{\mu}\sum_{m=2}^{n}|\nu_n(m)|^{\mu/2} \\
&\le Kn^{-\mu/2\alpha}\sum_{m=2}^{n}(n-m+1)^{(d-1)\mu/2} \\
&= Kn^{-\mu/2\alpha}\sum_{l=1}^{n-1}l^{(d-1)\mu/2} \\
&\le K_1 n^{-\mu/2\alpha+(d-1)(\mu/2)+1}. \qquad (3.89)
\end{aligned}$$

The condition $(d-1)\alpha+1<0$, again guarantees that the exponent in the last expression in (3.89) is negative, which proves (3.86).

To verify (3.87), it suffices to show

$$n^{-\mu/\alpha}\sum_{s=1}^{n-1}|h_{n-s}|^{\mu}\sum_{t=1}^{n-s}\sum_{j=t}^{n-1}|c_j|^{\mu}\to 0. \qquad (3.90)$$

It is easy to check that the LHS of (3.90) is $O(n^{-\mu/\alpha+2(d-1)\mu+3})$ and that the exponent is negative for $\mu<\alpha$ sufficiently close to α.

We must also check that

$$n^{-1/\alpha}\sum_{s=0}^{n-2}h_{n+s}A_{-s} = n^{-1/\alpha}\sum_{s=0}^{n-2}h_{n+s}\sum_{t=1}^{n-1-s}Z_t\left(\sum_{j=t+s}^{n-1}c_jZ_{n+t+s-j}\right)\xrightarrow{P}0. \qquad (3.91)$$

Since the bilinear form in (3.91) has no diagonal elements, one checks as above that

$$n^{-\mu/\alpha}\sum_{s=0}^{n-2}|h_{n+s}|^{\mu}\sum_{t=1}^{n-1-s}\sum_{j=t+s}^{n-1}|c_j|^{\mu}\to 0, \qquad (3.92)$$

if $\mu<\alpha$ is sufficiently close to α.

3) It remains to verify (3.75) with $u=3$. Denote the left-hand side of (3.75) by $-n^{-1/\alpha}I_{4n}$. Then $I_{4n} = I_{4n1} + 2I_{4n2}$, where $I_{4n1} = \nu_n\sum_{t=1}^{n}Z_t^2$ with $\nu_n = \int_{-\pi}^{\pi}C(\lambda)\left(\sum_{j=n}^{\infty}c_je^{-i\lambda j}\right)\eta(\lambda)d\lambda$ and

$$I_{4n2} = \sum_{k=1}^{n-1}\sum_{t=1}^{n-k}\kappa_n(k)Z_tZ_{t+k}$$

with

$$\kappa_n(k) = \int_{-\pi}^{\pi}C(\lambda)\left(\sum_{j=n}^{\infty}c_je^{-i\lambda j}\right)\cos(\lambda k)\eta(\lambda)d\lambda.$$

One can verify that $n^{-1/\alpha}I_{4ni}\xrightarrow{P}0$, $i=1,2$, in the same way as relations (3.73) and (3.74). This completes the proof of Lemma 3.5. ∎

4 Simulation

The estimator β_n of the unknown parameter vector β minimizes the function $\sigma_n^2(\beta)$ in (1.12). To find β_n, one can use without modification programs for Gaussian time series, for example, the one given in Section 12.1.3 of Beran (1994). These programs follow the minimization procedure described in Fox and Taqqu (1986). This procedure differs from the one discussed in Section 1 in two respects, neither of which affects the results. The division by $\sum_{i=1}^n X_i^2$ in (1.10) can be ignored because this quantity does not depend on the unknown parameter vector β. There is also no need for subtracting $\int_{-\pi}^{\pi} \log g(\lambda, \beta) d\lambda$ as in Fox and Taqqu (1986), because, in the case of FARIMA, this integral equals a constant independent of β.

Mikosch *et al.* (1995) ran a simulation using ARMA sequences. Focusing on long-range dependence, we generate here FARIMA $(0, d, 0)$ sequences with $S\alpha S$ innovations. In the Gaussian case $\alpha = 2$, one can apply the Durbin-Levinson algorithm (see Brockwell and Davis (1991)) to generate an exact FARIMA, using for example, the *arima.fracdiff.sim* function in S-Plus. Because there is no known technique to generate an exact FARIMA in the stable case, we will approximate the infinite moving average (1.4) by the finite one

$$X_t = \sum_{j=0}^{J} c_j Z_{t-j}, \quad t = 1, \ldots, n. \tag{4.1}$$

Here $c_j = \Gamma(j + d)/(\Gamma(d)\Gamma(j + 1))$ (see relation (7.13.1) in Samorodnitsky and Taqqu (1994)), and therefore, these coefficients can be easily obtained by using the recursion relation:

$$c_0 = 1, \qquad c_{j+1} = \frac{j+d}{j+1} c_j.$$

The $S\alpha S$ innovations are obtained through the S-Plus function *rstab* which uses a version of the Chambers, Mallow and Stuck (1976) algorithm described in Section 1.7 of Samorodnitsky and Taqqu (1994).

We set $J = 1000$ in (4.1) and simulate series with parameters

$$(\alpha, d) = (1.2, 0.1), \quad (1.5, 0.2), \quad (2, 0.1), \quad (2, 0.2)$$

and sample sizes $n = 100,\ 1000$ and 10,000. Observe that $0 < d < 1 - 1/\alpha$. We included the Gaussian $\alpha = 2$, so that one can compare the results with this known case. Gaussian series are generated with the S-Plus function *arima-fraidiff.sim* referred to earlier.

For each kind of time series, we generated 50 independent samples and reported the average values of the estimates of d, the corresponding sample standard deviations and the square root of the sample MSEs. The following notation is used. If d_0 is the nominal value of d and d_i is the estimate for sample i then,

$$\bar{d} = \frac{1}{50} \sum_{i=1}^{50} d_i, \quad \hat{\sigma}^2 = \frac{1}{49} \sum_{i=1}^{50} (d_i - \bar{d})^2, \quad \text{MSE} = \frac{1}{50} \sum_{i=1}^{50} (d_i - d_0)^2.$$

The results are summarized in Table 1. Figure 1 displays the corresponding boxplots and shows the relative scatter of the 50 estimates. The vertical axis in the figure indicates the deviations from the nominal values of d. For each time series we have (1) a thick line for the median; (2) a box representing the middle 50% of the data; (3) "Whiskers" encompassing approximately 95% of the data, and designated by dashed lines; (4) Outliers that fall beyond the whiskers.

n		$\alpha = 1.2$,$d = 0.1$	$\alpha = 1.5$,$d = 0.2$	$\alpha = 2$,$d = 0.1$	$\alpha = 2$,$d = 0.2$
100	Average	.066	.161	.087	.161
	$\hat{\sigma}$	.079	.059	.089	.095
	$\sqrt{\text{MSE}}$	.085	.071	.089	.102
1000	Average	.096	.195	.098	.196
	$\hat{\sigma}$	.021	.030	.027	.026
	$\sqrt{\text{MSE}}$	.021	.030	.027	.026
10000	Average	.099	.200	.101	.200
	$\hat{\sigma}$	.005	.006	.008	.007
	$\sqrt{\text{MSE}}$	.005	.006	.008	.007

Table 1: Estimation results for d using 50 replications.

The parameter d, as is well known, is hard to estimate when the time series is short. As in the Gaussian case, $\hat{\sigma}$ and $\sqrt{MSE}$ are relatively large for $n = 100$. The estimates improve dramatically for large sample sizes. They are very good when $n = 1000$ and excellent when $n = 10,000$.

Acknowledgment. The authors would like to thank Thomas Mikosch for pointing out an error in the original proof of Lemma 3.5, Vadim Teverovsky for performing the simulations, and also the two anonymous referees.

References

Avram, F. & Taqqu, M. S. (1986), 'Weak convergence of moving averages with infinite variance', *Dependence in Probability and Statistics* pp. 399–415. E. Eberlein and M. S. Taqqu, eds., Boston:Birkhauser.

Beran, J. (1994), *Statistics for Long-Memory Processes*, Chapman & Hall, New York.

Billingsley, P. (1968), *Convergence of Probability Measures*, Wiley, New York.

Bingham, N. H., Goldie, C. M. & Teugels, J. L. (1987), *Regular Variation*, Cambridge University Press.

Brockwell, P. J. & Davis, R. A. (1991), *Time Series: Theory and Methods*, 2nd edn, Springer-Verlag, New York.

Chambers, J. M., Mallows, C. & Stuck, B. W. (1976), 'A method for simulating stable random variables', *Journal of the American Statistical Association* **71**(354), 340–344. Theory and Methods Section.

Cline, D. B. H. (1983), Estimation and linear prediction for regression, autoregression and ARMA with infinite variance data, PhD thesis, Colorado State University.

Dahlhaus, R. (1989), 'Efficient parameter estimation for self similar processes', *The Annals of Statistics* **17**(4), 1749–1766.

Davis, R. A. & Resnick, S. I. (1985), 'Limit theory for moving averages of random variables with regularly varying tail probabilities', *The Annals of Probability* **13**(1), 179–195.

Davis, R. A. & Resnick, S. I. (1986), 'Limit theory for the sample covariance and correlation functions of moving averages', *The Annals of Statistics* **14**(2), 533–558.

Feller, W. (1971), *An Introduction to Probability Theory and its Applications*, Vol. 2, 2nd edn, Wiley, New York.

Fox, R. & Taqqu, M. S. (1986), 'Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series', *The Annals of Statistics* **14**, 517–532.

Giraitis, L. & Surgailis, D. (1990), 'A central limit theorem for quadratic forms in strongly dependent linear variables and application to asymptotical normality of Whittle's estimate', *Probability Theory and Related Fields*.

Gnedenko, B. V. & Kolmogorov, A. N. (1954), *Limit distributions for sums of independent random variables*, Addison-Wesley, Reading, MA.

Hannan, E. J. (1973), 'The asymptotic theory of linear time series models', *Journal of Applied Probability* **10**, 130–145.

Helson, H. (1983), *Harmonic Analysis*, Addison-Wesley, Reading, MA.

Heyde, C. C. & Gay, R. (1993), 'Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence', *Stochastic Processes and their Applications* **45**, 169–182.

Kokoszka, P. S. (1995), 'Prediction of infinite variance fractional ARIMA', *Probability and Mathematical Statistics.* To appear.

Kokoszka, P. S. & Taqqu, M. S. (1995*a*), 'Fractional ARIMA with stable innovations', *Stochastic Processes and their Applications.* To appear.

Kokoszka, P. S. & Taqqu, M. S. (1995*b*), 'Infinite variance stable moving averages with long memory', *Annals of the Journal of Econometrics.* To appear.

Mikosch, T., Gadrich, T., Klüppelberg, C. & Adler, R. J. (1995), 'Parameter estimation for ARMA models with infinite variance innovations', *The Annals of Statistics* **23**, 305–326.

Resnick, S. I. (1987), *Extreme Values, Regular Variation and Point Processes*, Springer-Verlag, New York.

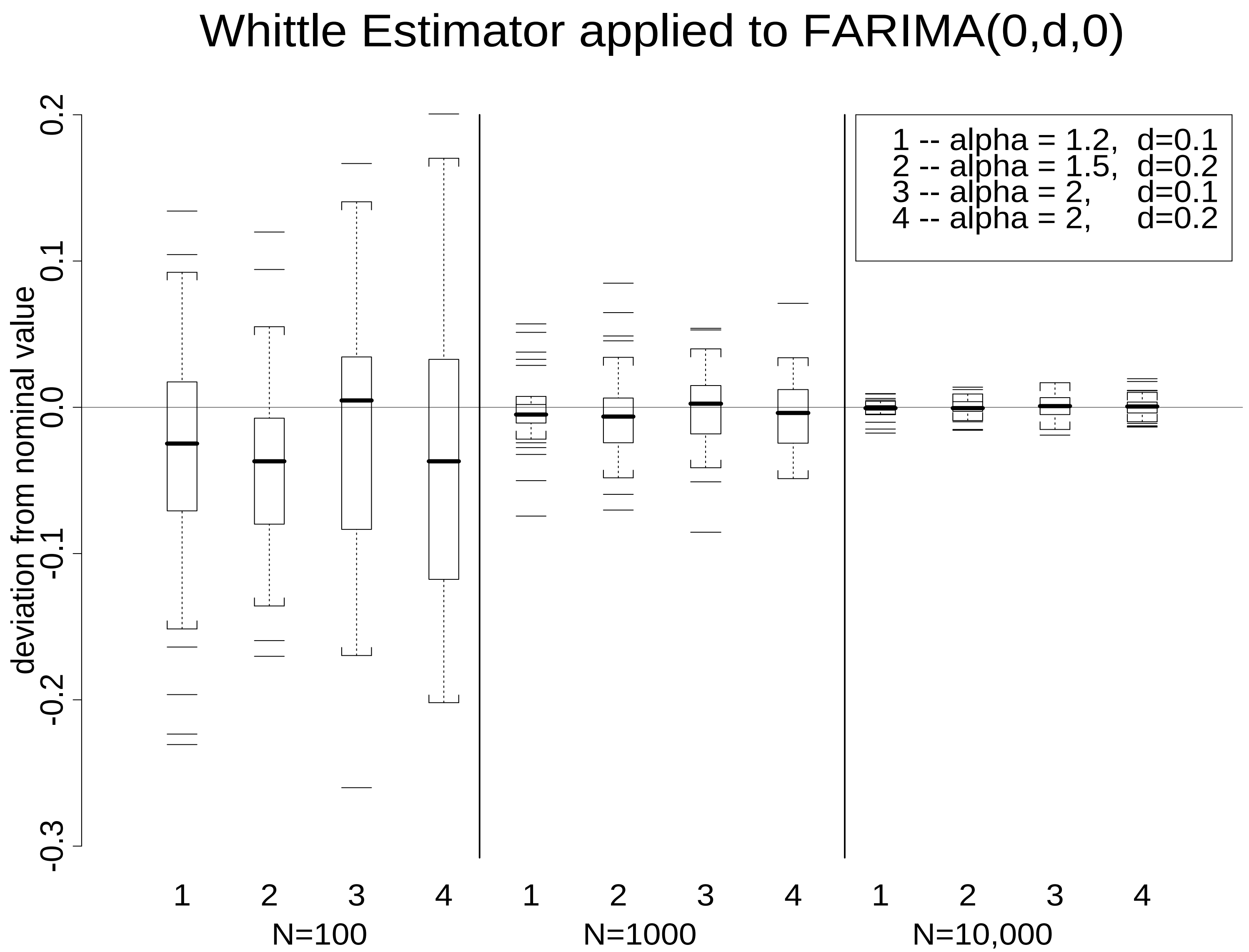
Rosiński, J. & Woyczynski, W. A. (1987), 'Multilinear forms in Pareto-like random variables and product random measures', *Colloquium Mathematicum* **51**, 303–313.

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

Zygmund, A. (1979), *Trigonometric Series*, Cambridge University Press, Cambridge. Volumes I,II.

Piotr Kokoszka, University of Utah, Department of Mathematics, 233 Widtsoe Building, Salt Lake City, UT 84112, USA.
email: kokoszka@math.utah.edu

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


Whittle Estimator applied to FARIMA(0,d,0)
deviation from nominal value
0.2
0.1
0.0
-0.1
-0.2
-0.3
1 -- alpha = 1.2, d=0.1
2 -- alpha = 1.5, d=0.2
3 -- alpha = 2, d=0.1
4 -- alpha = 2, d=0.2
1 2 3 4 1 2 3 4 1 2 3 4
N=100
N=1000
N=10,000