Stationary ARCH models: dependence structure and Central Limit Theorem *

LIUDAS GIRAITIS † — London School of Economics

PIOTR KOKOSZKA — The University of Liverpool

REMIGIJUS LEIPUS ‡ — The University of Liverpool

First version January 1998
Revised September 1998

Abstract

This paper studies a broad class of non–negative ARCH(∞) models. Sufficient conditions for the existence of a stationary solution are established and an explicit representation of the solution as a Volterra type series is found. Under our assumptions, the covariance function can decay slowly like a power function, falling just short of the long memory structure. A moving average representations in martingale differences is established and the Central Limit Theorem is proved.

Running head: Stationary ARCH models

Corresponding Author: Piotr Kokoszka, Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 3BX, United Kingdom
E–mail: P.S.Kokoszka@liverpool.ac.uk, Telephone: +44 (151) 794 4759

*Research partially supported by an EPSRC grant GR/L/78222 at the University of Liverpool
†Supported by ESRC grant R000235892

‡On leave from Vilnius University and Institute of Mathematics and Informatics, Lithuania

1 Introduction

In the present paper we study general non-negative ARCH(∞) sequences defined as follows:

Definition 1.1 A random sequence $\{X_k,\ k \in \mathbf{Z}\}$ is said to satisfy ARCH(∞) equations if there exists a sequence of iid non–negative random variables $\{\xi_k,\ k \in \mathbf{Z}\}$ such that

$$X_k = \rho_k \xi_k, \tag{1.1}$$

$$\rho_k = a + \sum_{j=1}^{\infty} b_j X_{k-j}, \tag{1.2}$$

where $a \geq 0,\ b_j \geq 0,\ j = 1, 2, \ldots$.

The general framework leading to the model (1.1), (1.2) was introduced by Robinson (1991) in the context of testing for strong serial correlation and has been subsequently developed by several authors, see Section 4.

This class of models includes the classical ARCH(∞) model

$$r_k^2 = \sigma_k^2 \varepsilon_k^2, \quad \sigma_k^2 = a + \sum_{j=1}^{\infty} b_j r_{k-j}^2,$$

as well as a more general model

$$|r_k|^\delta = \sigma_k^\delta |\varepsilon_k|^\delta, \tag{1.3}$$

$$\sigma_k^\delta = a + \sum_{j=1}^{\infty} b_j |r_{k-j}|^\delta, \tag{1.4}$$

where $\{\varepsilon_k,\ k \in \mathbf{Z}\}$ is a sequence of zero mean iid random variables and $\delta > 0$. In fact, one can assume that $\xi_k = f(\varepsilon_k)$ for some non-negative function f which may be of interest. This flexibility of the model (1.1), (1.2) is important in light of recent empirical findings which show that while the returns r_k on most speculative assets are basically uncorrelated, the covariance structure of the powers $|r_k|^\delta$ is quite rich, see e.g. Dacorogna *et al.* (1993) and a review paper of Baillie (1996).

An important question is to find conditions on the coefficients $a \geq 0,\ b_j \geq 0,\ j = 1, 2, \ldots$ and the random variables $\{\xi_k\}$ which guarantee the existence of a stationary solution to equations (1.1), (1.2). Such conditions will, in particular, lead to corresponding conditions for special cases like e.g. equations (1.3), (1.4). We show in Section 2 that if $E(\sum_{j=1}^{\infty} b_j \xi_j) < 1$, then equations (1.1), (1.2) have a unique strictly stationary solution with finite first moment EX_k. Under a stronger condition, see Theorem 2.1, this solution has also finite variance and, consequently, is weakly stationary as well. An important point is that we basically require only the summability of the b_j to show the existence and to obtain an explicit solution to equations (1.1) and (1.2). Thus the class of models studied in this paper is larger than the class of GARCH models defined by equations

$$r_k^2 = \sigma_k^2 \varepsilon_k^2, \qquad \sigma_k^2 = \alpha_0 + \sum_{i=1}^{p} \beta_i \sigma_{k-i}^2 + \sum_{j=1}^{q} \alpha_j r_{k-j}^2. \tag{1.5}$$

Under additional constraints on the polynomials $\beta(z)$ and $\alpha(z)$ the conditional variance σ_k^2 can be rewritten as a moving average $\sigma_k^2 = a + \sum_{j=1}^{\infty} b_j r_{k-j}^2$ with positive *exponentially* decaying weights b_j (see Nelson and Cao (1992)). The existence of a stationary solution to GARCH equations has been studied in depth by Nelson (1990) and Bougerol and Picard (1992).

Section 3 deals with the asymptotic behaviour of the covariance function of the series $\{X_k\}$ defined by (1.1), (1.2). Bollerslev (1988) and He and Teräsvirta (1999), among others, studied the covariance function of GARCH models. What is relevant to our paper is that the exponential decay of the coefficients b_j implies the exponential decay of the covariance function of the sequence $\{r_k^2\}$. This result remains valid for the general model (1.1), (1.2), see Kokoszka and Leipus (2000). We show, in particular, that if $b_j \sim Cj^{-\gamma}$ for some $\gamma > 1$, then the covariance function of $\{X_k\}$ also decays at the rate $k^{-\gamma}$. Thus, as γ approaches 1, this covariance structure becomes very close to the long memory structure which is typically defined by requiring that the autocovariance function γ_k, say, is not absolutely summable, i.e. $\sum_k |\gamma_k| = \infty$, or, more specifically, by the condition

$$\gamma_k \sim k^{2q-1} L(k), \tag{1.6}$$

for some $0 < q < 1/2$ and a slowly varying function $L(k)$. ($L(s), s \geq 0$ is called a slowly varying function if $L(ts)/L(s) \to 1$ as $s \to \infty$ for any $t > 0$.)

The question of the existence of long memory ARCH models is taken up in Section 4. We discuss relevant recent research in light of the results of Sections 2 and 3. The main conclusion is that more research in this direction is called for. We also establish a martingale representation of the solution to (1.1), (1.2) and apply it to obtain an example of an ARCH process whose covariance function is arbitrarily close to the covariance function of a long memory sequence at any finite number of lags.

Finally, we show in Section 5 that the Central Limit Theorem holds for the ARCH sequence $\{X_k\}$ under the same assumption (2.4) which guarantees the existence of a weakly stationary solution to (1.1), (1.2). As for a Gaussian sequence, no additional mixing assumptions are needed.

The results of this paper are proved using a Volterra series type expansion of the ARCH process. In general, if X_k is some non–linear function $X_k = f(\xi_k, \xi_{k-1}, \ldots)$ of the shocks ξ_k, it is sometimes convenient to work with its Volterra series expansion:

$$f(\xi_k, \xi_{k-1}, \ldots) = \sum_{j_1=0}^{\infty} c_{j_1} \xi_{k-j_1} + \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} c_{j_1,j_2} \xi_{k-j_1} \xi_{k-j_1-j_2}$$

$$+ \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \sum_{j_3=0}^{\infty} c_{j_1,j_2,j_3} \xi_{k-j_1} \xi_{k-j_1-j_2} \xi_{k-j_1-j_2-j_3} + \ldots \tag{1.7}$$

(see Priestley (1988)). For the ARCH models studied in this paper the weights $c_{j_1,j_2,\ldots,j_m}$ have the form $c_{j_1,j_2,\ldots,j_m} = b_{j_1} \ldots b_{j_m}$.

2 Stationary ARCH(∞) models

Our goal in this section is to investigate the existence of a stationary solution to equations (1.1), (1.2). Since our main interest is in the memory structure of the model, we focus on the existence of solutions with finite variance.

Recall that a random sequence $\{Z_k\}$ is strictly stationary if for any integers m and N the distribution of the vector $(Z_m, Z_{m+1}, \ldots, Z_{m+N})$ does not depend on m. Note that no moment assumptions are needed to define strict stationarity. By weak stationarity we understand below covariance stationarity, which, obviously, requires finite variance.

In general, we shall assume that the first two moments $E\xi_0$ and $E\xi_0^2$ are finite. However, for the existence of a strictly stationary solution to (1.1), (1.2) we assume only the existence of the first moment $E\xi_0$ (see Theorem 2.1).

Formally, (1.1) and recursion relations (1.2) give

$$X_k = (a + \sum_{j_1=1}^{\infty} b_{j_1} X_{k-j_1})\xi_k = a\xi_k + \sum_{j_1=1}^{\infty} b_{j_1} \xi_k \xi_{k-j_1} \rho_{k-j_1}$$

$$= \ldots = a \sum_{l=0}^{\infty} \sum_{j_1,\ldots,j_l=1}^{\infty} b_{j_1} \ldots b_{j_l} \xi_k \xi_{k-j_1} \ldots \xi_{k-j_1-\ldots-j_l}, \tag{2.1}$$

which can be rewritten as

$$X_k = a \sum_{l=0}^{\infty} M_l(k), \tag{2.2}$$

where $M_0(k) = \xi_k$ and for $l \geq 1$

$$M_l(k) = \sum_{j_1,\ldots,j_l=1}^{\infty} b_{j_1} \ldots b_{j_l} \xi_k \xi_{k-j_1} \ldots \xi_{k-j_1-\ldots-j_l} \tag{2.3}$$

$$= \sum_{j_l < j_{l-1} < \ldots < j_1 < k}^{\infty} b_{k-j_1} b_{j_1-j_2} \ldots b_{j_{l-1}-j_l} \xi_k \xi_{j_1} \ldots \xi_{j_l}.$$

However, to show that the sequence $\{X_k\}$ given by (2.2) and (2.3) is a solution to (1.1), (1.2) additional restrictions on the weights b_j and the sequence $\{\xi_j\}$ must be imposed. Kokoszka and Leipus (2000) proved that $\{X_k\}$ is a strictly and weakly stationary solution if $(E\xi_0^2)^{1/2} \sum_{j=1}^{\infty} b_j < 1$. We refine this result in Theorem 2.1 below and prove the uniqueness of the stationary solution. We say that a solution to (1.1), (1.2) is *non-anticipative* if X_k is independent of the ξ_n, $n > k$ for every integer k.

THEOREM 2.1 *Suppose that* $E\xi_0 < \infty$ *and*

$$E\xi_0 \sum_{j=1}^{\infty} b_j < 1. \tag{2.4}$$

Then (2.2) *is a strictly stationary solution to* (1.1), (1.2) *with finite first moment* EX_k. *Moreover, such a solution with finite first moment is a unique non–anticipative solution. If the constant* a *in* (1.2) *is zero, this unique solution is* $X_k = 0$.

If, in addition, $E\xi_0^2 < \infty$ *and*

$$(E\xi_0^2)^{1/2} \sum_{j=1}^{\infty} b_j < 1, \tag{2.5}$$

then (2.2) *is also a unique weakly stationary solution to* (1.1), (1.2).

Remark 2.1 In the sequel, we use the notation

$$\lambda_1 = E\xi_0, \ \lambda_2 = E\xi_0^2$$

and

$$B = \sum_{j=1}^{\infty} b_j.$$

In this notation, conditions (2.4) and (2.5) become $\lambda_1 B < 1$ and $\lambda_2 B^2 < 1$, respectively. Condition (2.5) implies (2.4) because $\lambda_1^2 \leq \lambda_2$.

Remark 2.2 In general, the conditions of Theorem 2.1 cannot be weakened, if the existence of appropriate moments is required. We illustrate this point by considering ARCH(1) model defined by equations (1.1), (1.2) with $b_1 = b$ and $b_j = 0$ for $j \geq 2$. Suppose that $\xi_k = \varepsilon_k^2$, where the ε_k are standard normal. Then condition (2.4) becomes $b < 1$. If $b = 1$, then there is a unique stationary solution satisfying $\lim_{x\to\infty} xP(X_k > x) = c$, where c is a positive constant. It follows that in this case $EX_k = \infty$. Condition (2.5) becomes $\sqrt{3}b < 1$ (since $E\varepsilon_k^4 = 3$), and a similar argument shows that if $b \geq 1/\sqrt{3}$, then the unique stationary solution satisfies $EX_k^2 = \infty$. The details are presented in Section 8.4.2 of Embrechts *et al.* (1997).

Remark 2.3 Note that GARCH(1,1) model

$$r_k = \sigma_k \varepsilon_k,$$

$$\sigma_k^2 = \alpha_0 + \beta_1 \sigma_{k-1}^2 + \alpha_1 r_{k-1}^2,$$

can be written in the form (1.1), (1.2) with $\rho_k = \sigma_k^2$, $\xi_k = \varepsilon_k^2$, $X_k = r_k^2$ and the coefficients $a = \alpha_0(1+\beta_1+\beta_1^2+\ldots) = \alpha_0/(1-\beta_1)$, $b_j = \beta_1^{j-1}\alpha_1$. The conditions for strict stationarity of the sequence $\{r_k^2\}$ are studied in Nelson (1990) who uses the following representation of σ_k^2:

$$\sigma_k^2 = \alpha_0(1 + \sum_{l=1}^{\infty} \prod_{j=1}^{l} \eta_{k-j}), \tag{2.6}$$

where $\eta_j = \beta_1 + \alpha_1 \varepsilon_j^2$. Representation (2.6) and the strong law of large numbers applied to $S_n = \log \eta_{k-1} + \ldots + \log \eta_{k-l}$ yield conditions of the existence of strongly stationary solution $\{r_k^2\}$. This idea is not directly applicable to our model (1.1), (1.2) or (2.1).

He and Teräsvirta (1999) study general GARCH(p,q) model (1.5) and establish conditions for stationarity and the existence of higher order moments. They use the following representation of conditional variance:

$$\sigma_k^2 = \alpha_0 + \sum_{i=1}^{p} \eta_{i,k-i} \sigma_{k-i}^2, \tag{2.7}$$

where $\eta_{i,k-i} = \beta_i + \alpha_i \varepsilon_{k-i}^2$. Recursive application of equality (2.7) yields a representation analogous to (2.6).

Proof of Theorem 2.1: We first verify that condition (2.4) implies the existence of a strictly stationary solution with finite expectation. Since all factors of the product $\xi_k \xi_{k-j_1} \ldots \xi_{k-j_1-\ldots-j_l}$ in (2.3) are independent we obtain

$$EM_l(k) = \lambda_1^{l+1} \sum_{j_1,\ldots,j_l=1}^{\infty} \prod_{m=1}^{l} b_{j_m} = \lambda_1 (\lambda_1 B)^l$$

and

$$EX_k = a \sum_{l=0}^{\infty} EM_l(k) = a\lambda_1 \sum_{l=0}^{\infty} (\lambda_1 B)^l < \infty. \tag{2.8}$$

Therefore $M_l(k)$ and X_k are well defined: for any k and l, $M_l(k) < \infty$ a.s. and $X_k < \infty$ a.s. One can easily verify that $\{X_k\}$ (2.2) is a strictly stationary sequence satisfying (1.1), (1.2).

To prove uniqueness, assume that Y_k is any strictly stationary solution to (1.1), (1.2) with finite first moment $EY_0 < \infty$. Then, applying recursively relation (1.2), after m steps we obtain:

$$Y_k = a\xi_k \sum_{l=0}^{m} \sum_{j_1,\ldots,j_l=1}^{\infty} b_{j_1} \ldots b_{j_l} \xi_{k-j_1} \ldots \xi_{k-j_1-\ldots-j_l}$$

$$+\xi_k \sum_{j_1,\ldots,j_{m+1}=1}^{\infty} b_{j_1} \ldots b_{j_{m+1}} \xi_{k-j_1} \ldots \xi_{k-j_1-\ldots-j_m} Y_{k-j_1-\ldots-j_{m+1}}.$$

Therefore, we have that

$$Y_k - X_k = \xi_k \sum_{j_1,\ldots,j_{m+1}=1}^{\infty} b_{j_1} \ldots b_{j_{m+1}} \xi_{k-j_1} \ldots \xi_{k-j_1-\ldots-j_m} Y_{k-j_1-\ldots-j_{m+1}} \tag{2.9}$$

$$-\xi_k a \sum_{l=m+1}^{\infty} \sum_{j_1,\ldots,j_l=1}^{\infty} b_{j_1} \ldots b_{j_l} \xi_{k-j_1} \ldots \xi_{k-j_1-\ldots-j_l}.$$

For the first term on the right-hand side of (2.9), by Chebyshev's inequality and the assumption that $\{Y_k\}$ is a non–anticipative solution, we have

$$\epsilon P\{\xi_k \sum_{j_1,\ldots,j_{m+1}=1}^{\infty} \left[b_{j_1} \ldots b_{j_{m+1}} \xi_{k-j_1} \ldots \xi_{k-j_1-\ldots-j_m} Y_{k-j_1-\ldots-j_{m+1}} \right] > \epsilon\}$$

$$\leq EY_k \ (\lambda_1 B)^{m+1}.$$

By (2.4) and the Borel-Cantelli lemma, this implies almost sure convergence to zero as $m \to \infty$. Since $\sum_{l=0}^{\infty} M_l(k) < \infty$ a.s., choosing m large enough, the second term on the right-hand side of (2.9) can be made arbitrarily small with probability 1. Thus $Y_k = X_k$ a.s. (Note that if instead of assuming non–anticipativity, we assume $EY_k^2 < \infty$ and (2.5), then the bound on the right–hand side of in the Chebyshev's inequality above becomes

$(EY_k^2)^{1/2}(\lambda_2^{1/2}B)^{m+1}$. Consequently under (2.5), (2.2) is a unique solution in the class of all stationary solutions with finite second moment, i.e not only in the class of non–anticipative solutions.)

We prove now that under assumption (2.5) the solution X_k given by (2.2) is weakly stationary. Since (2.5) implies (2.4), by the first part of the theorem, X_k is a unique strictly stationary solution with finite first moment. It remains to prove that the variance $\operatorname{Var} X_k$ is finite. This follows from a number of more general facts we state and prove below. These lemmas also play an important role in the subsequent sections of the paper.

Our first lemma establishes a convenient representation for $\operatorname{Cov}(X_{n+k}, X_n)$ and shows that $\operatorname{Cov}(X_{n+k}, X_n) \geq 0$.

LEMMA 2.1 *Under assumption* (2.5)

$$\operatorname{Cov}(X_{n+k}, X_n) = a^2 \sum_{l,p=0}^{\infty} R_k(l,p), \tag{2.10}$$

where

$$R_k(l,p) := \operatorname{Cov}(M_l(k), M_p(0)) \tag{2.11}$$

$$= \sum_{\substack{j_l<\ldots<j_1<k \\ s_p<\ldots<s_1<0;\gamma(k,j,s)\geq 1}} \Big[(\lambda_2^{\gamma(k,j,s)} - \lambda_1^{2\gamma(k,j,s)})\lambda_1^{2+l+p-2\gamma(k,j,s)}$$

$$\times b_{k-j_1} b_{j_1-j_2} \ldots b_{j_{l-1}-j_l} b_{-s_1} b_{s_1-s_2} \ldots b_{s_{p-1}-s_p} \Big]$$

with $\gamma(k,j,s) := |\{k, j_1, \ldots, j_l\} \cap \{0, s_1, \ldots, s_p\}|$, *the number of common elements of the sets* $\{k, j_1, \ldots, j_l\}$ *and* $\{0, s_1, \ldots, s_p\}$.

Since $\lambda_2 \geq \lambda_1^2$, *the right-hand side of* (2.10) *is non-negative*

PROOF: Observe that

$$\operatorname{Cov}(X_k, X_0) = a^2 \sum_{l,p=0}^{\infty} \operatorname{Cov}(M_l(k), M_p(0)) = a^2 \sum_{l,p=0}^{\infty} R_k(l,p), \tag{2.12}$$

where

$$R_k(l,p) = \operatorname{Cov}(M_l(k), M_p(0))$$

$$= \sum_{\substack{j_l<\ldots<j_1<k \\ s_p<\ldots<s_1<0}} b_{k-j_1} b_{j_1-j_2} \ldots b_{j_{l-1}-j_l} b_{-s_1} b_{s_1-s_2} \ldots b_{s_{p-1}-s_p} \operatorname{Cov}(\xi_k \xi_{j_1} \ldots \xi_{j_l}, \xi_0 \xi_{s_1} \ldots \xi_{s_p}). \tag{2.13}$$

Since the ξ_k are iid random variables, we get that

$$\operatorname{Cov}(\xi_k \xi_{j_1} \ldots \xi_{j_l}, \xi_0 \xi_{s_1} \ldots \xi_{s_p}) = E(\xi_k \xi_{j_1} \ldots \xi_{j_l}\ \xi_0 \xi_{s_1} \ldots \xi_{s_p}) - E(\xi_k \xi_{j_1} \ldots \xi_{j_l}) E(\xi_0 \xi_{s_1} \ldots \xi_{s_p})$$

equals to 0, if the sets of indexes $\{k, j_1, \ldots, j_l\}$ and $\{0, s_1, \ldots, s_p\}$ do not intersect, and equals to $(\lambda_2^{\gamma(k,j,s)} - \lambda_1^{2\gamma(k,j,s)})\lambda_1^{2+l+p-2\gamma(k,j,s)}$ otherwise. This implies (2.10) and (2.11). ∎

We now show that the right-hand side of (2.10) is finite. This is established in Proposition 2.1, which follows from relations (2.14) of Lemma 2.2 and (2.10) of Lemma 2.1. Relation (2.15) of Lemma 2.2 will be used in Section 3.

Lemma 2.2 *Consider the quantities* $R_k(l,p)$ *defined by* (2.11). *Under assumption* (2.5),

$$0 \le R_k(l,p) \le \lambda_2 D^{l+p}, \tag{2.14}$$

and if $k \ge 1$,

$$R_k(l,p) \le \sum_{u=1}^{l} \sum_{v=0}^{p} Q_k(u,v) \lambda_1^{u+v} \lambda_2 D^{l+p-u-v}, \tag{2.15}$$

where

$$D = B\lambda_2^{1/2} \tag{2.16}$$

stands for the left-hand side of (2.5) *and where*

$$Q_k(u,v) = \sum_{\substack{j_u<\ldots<j_1<k \\ j_u<s_{v-1}<\ldots<s_1<0}} b_{k-j_1} b_{j_1-j_2} \ldots b_{j_{u-1}-j_u} b_{-s_1} b_{s_1-s_2} \ldots b_{s_{v-2}-s_{v-1}} b_{s_{v-1}-j_u}. \tag{2.17}$$

Proof: Since

$$0 \le \mathrm{Cov}(\xi_k \xi_{j_1} \ldots \xi_{j_l}, \xi_0 \xi_{s_1} \ldots \xi_{s_p}) \le E[\xi_k \xi_{j_1} \ldots \xi_{j_l}][\xi_0 \xi_{s_1} \ldots \xi_{s_p}]$$
$$\le \Big(E[\xi_k \xi_{j_1} \ldots \xi_{j_l}]^2 E[\xi_0 \xi_{s_1} \ldots \xi_{s_p}]^2\Big)^{1/2} = \lambda_2 \lambda_2^{(l+p)/2} \tag{2.18}$$

we get from (2.13)

$$0 \le R_k(l,p) \le (\sum_{j=1}^{\infty} b_j)^{l+p} \lambda_2 \lambda_2^{(l+p)/2} = \lambda_2 D^{l+p}.$$

Thus (2.14) holds.

Now we prove (2.15).

Let $k \ge 1$. Since we have to consider in (2.13) only overlaping indexes $\{k, j_1, \ldots, j_l\}$ and $\{0, s_1, \ldots, s_p\}$, the set I of such indexes can be factorized into non-intersecting subsets $I_{u,v}$, $u = 1, \ldots, l, v = 0, \ldots, p$ consisting of indexes $(j) := \{j_1, \ldots, j_l\}$ and $(s) = \{s_1, \ldots, s_p\}$ such that sets $\{k, j_1, \ldots, j_{u-1}\}$ and $\{0, s_1, \ldots, s_{v-1}\}$ do not overlap and $j_u = s_v$. Using (2.18) and (2.13) we can rewrite

$$R_k(l,p) = \sum_{u=1}^{l} \sum_{v=0}^{p} \sum_{(j),(s)\in I_{u,v}} \Big[b_{k-j_1} b_{j_1-j_2} \ldots b_{j_{l-1}-j_l}$$
$$\times b_{-s_1} b_{s_1-s_2} \ldots b_{s_{p-1}-s_p}\Big] \mathrm{Cov}(\xi_k \xi_{j_1} \ldots \xi_{j_l}, \xi_0 \xi_{s_1} \ldots \xi_{s_p})$$
$$\le \sum_{u=1}^{l} \sum_{v=0}^{p} Q_k(u,v) \lambda_1^{u+v} \lambda_2 D^{l+p-u-v}.$$

■

The following corollary to Lemma 2.2 will be used in Section 3.

Corollary 2.1 *Under assumption* (2.5)

$$\sum_{k=0}^{\infty} R_k(l,p) \le D^{l+p}(l+1)(p+1)\lambda_2, \tag{2.19}$$

where $R_k(l,p)$ *is defined in* (2.11).

PROOF: Since indexes $k-j_1, j_1-j_2, \ldots, j_{u-1}-j_u, -s_1, s_1-s_2, \ldots, s_{v-2}-s_{v-1}, s_{v-1}-j_u$ in (2.17) are linearly independent and number $u+v$ of these variables is the same as number of variables $k, j_1, \ldots, j_u, s_1, \ldots, s_{v-1}$, it follows that

$$\lambda_1^{u+v} \sum_{k=0}^{\infty} Q_k(u,v) \leq (\sum_{j=1}^{\infty} b_j)^{u+v} \lambda_1^{u+v} \leq D^{u+v}$$

in view of $\lambda_1 \leq \lambda_2^{1/2}$. This and (2.15) imply (2.19). ∎

PROPOSITION 2.1 *Under assumption* (2.5)

$$0 \leq \mathrm{Cov}(X_{n+k}, X_n) \leq \lambda_2 a^2 \sum_{l=0}^{\infty} D^l < \infty, \tag{2.20}$$

where $\lambda_2 = E\xi_0^2$ *and* $D < 1$ *is the left-hand side of* (2.5).

PROOF: By Lemma 2.2, for all $k \geq 0$

$$R_k(l,p) \leq \lambda_2 D^{l+p}, \quad (l \geq p \geq 0).$$

Therefore

$$\sum_{l,p=0}^{\infty} R_k(l,p) \leq \lambda_2 (\sum_{l=0}^{\infty} D^l)^2.$$

Inserting this bound into (2.10) yields (2.20). ∎

3 Asymptotic properties of the covariance function

Condition (2.5), $(E\xi_0^2)^{1/2} \sum_{j=1}^{\infty} b_j < 1$, in Theorem 2.1 ensures that the process $\{X_k\}$ is weakly stationary, i.e. it has finite variance and covariances $\mathrm{Cov}(X_{n+k}, X_n) = \mathrm{Cov}(X_k, X_0)$ which are invariant with respect to the shift in n. This result, however, says nothing about the memory structure of the sequence $\{X_k\}$, i.e. how fast the covariance function of the process $\{X_k\}$ tends to 0, as $k \to \infty$.

We show in Proposition 3.1 below that the assumption (2.5), which ensures the existence of a weakly stationary solution of ARCH(∞) equations, also implies that

$$\sum_{k=-\infty}^{\infty} \mathrm{Cov}(X_k, X_0) < \infty. \tag{3.1}$$

Since the covariances $\mathrm{Cov}(X_k, X_0)$ are nonnegative, this shows that ARCH(∞) models with finite second moments satisfying condition (2.5) are *short memory* processes, in the sense that the covariance function is absolutely summable.

PROPOSITION 3.1 *If* $D = (E\xi_0^2)^{1/2} \sum_{j=1}^{\infty} b_j < 1$, *then*

$$\sum_{k=-\infty}^{\infty} \mathrm{Cov}(X_k, X_0) < \infty. \tag{3.2}$$

(Recall that $\mathrm{Cov}(X_k, X_0) \geq 0$.*)*

PROOF: By (2.10) and Corollary 2.1,

$$\sum_{k=-\infty}^{\infty} \mathrm{Cov}(X_k, X_0) = \mathrm{Cov}(X_0, X_0) + 2\sum_{k=1}^{\infty} \mathrm{Cov}(X_k, X_0)$$

$$= a^2 \sum_{k=-\infty}^{\infty} \sum_{l,p=0}^{\infty} R_k(l,p) \le 2\lambda_2 a^2 \sum_{l,p=0}^{\infty} (l+1)(p+1)D^{l+p} = 2\lambda_2 a^2 (\sum_{l=0}^{\infty}(l+1)D^l)^2 < \infty.$$

■

We show now that the rate of decay of the coefficients b_j controls the asymptotic behaviour of $\mathrm{Cov}(X_k, X_0)$. Kokoszka and Leipus (2000) showed that if the b_j decay exponentially, then the covariance function is exponentially decaying as well, in particular, if $b_j = \beta\alpha^j$ with $\beta \ge 0$, $0 < \alpha < 1$, and if condition (2.5) is satisfied, then

$$\mathrm{Cov}(X_k, X_0) \asymp (\alpha(1+\beta))^k,$$

where $\alpha_k \asymp \beta_k$ means that there are positive constants C_1 and C_2 such that $C_1\beta_k < \alpha_k < C_2\beta_k$, $k > k_0$ for some $k_0 > 0$.

In Proposition 3.2 below we show that hyperbolic decay $b_j \le Cj^{-\gamma}$, $\gamma > 1$, implies the hyperbolic decay of the covariances $\mathrm{Cov}(X_k, X_0)$. Proposition 3.2 also implies that if $b_j \sim Cj^{-\gamma}$, $\gamma > 1$, ($\alpha_j \sim \beta_j$ means that $\alpha_j/\beta_j \to 1$ as $j \to \infty$), then $\mathrm{Cov}(X_k, X_0) \asymp k^{-\gamma}$. Thus, even though condition (2.5) implies absolute summability of the covariances, it allows for a very slow rate of decay of the autocorrelation function when $\gamma > 1$ (fixed) is chosen close to 1.

PROPOSITION 3.2 *Suppose that* (2.5) *holds and assume that for some $\gamma > 1$, $c > 0$*

$$b_j \le cj^{-\gamma}, \ j = 1, 2, \ldots \tag{3.3}$$

Then there exists $C > 0$ such that for $k \ge 1$

$$\mathrm{Cov}(X_k, X_0) \le Ck^{-\gamma}. \tag{3.4}$$

If $c_1 j^{-\gamma} \le b_j \le c_2 j^{-\gamma}$ for sufficiently large j, then

$$C_1 k^{-\gamma} \le \mathrm{Cov}(X_k, X_0) \le C_2 k^{-\gamma} \tag{3.5}$$

for sufficiently large k.

PROOF: In the proof C denotes a generic positive constant which may change from line to line.

We will show that under assumption (3.3) the terms $R_k(l,p)$ in the expansion (2.10) of the covariance $\mathrm{Cov}(X_k, X_0)$ can be estimated as

$$R_k(l,p) \le Ck^{-\gamma}(l+1)^{\gamma+2}(p+1)^{\gamma+2}D^{l+p}. \tag{3.6}$$

Once (3.6) has been verified, (3.4) and the upper bound in (3.5) follow immediately because

$$\mathrm{Cov}(X_k, X_0) = a^2 \sum_{l,p=0}^{\infty} R_k(l,p) \leq Ck^{-\gamma} \sum_{l,p=0}^{\infty} (l+1)^{\gamma+2}(p+1)^{\gamma+2} D^{l+p}$$

$$= Ck^{-\gamma}(\sum_{l=0}^{\infty}(l+1)^{\gamma+2}D^l)^2 \leq Ck^{-\gamma}.$$

The lower bound in (3.5) is easy to establish. Since the $R_k(l,p)$ in (2.10) and the b_j are non-negative

$$\mathrm{Cov}(X_k, X_0) \geq a^2 R_k(1,1) = a^2 \sum_{\substack{j_1<k \\ s_1<0}} b_{k-j_1} b_{-s_1} \mathrm{Cov}(\xi_k \xi_{j_1}, \xi_0 \xi_{s_1})$$

$$= a^2(\lambda_1^2 \lambda_2 - \lambda_1^4)(\sum_{j<0} b_{k-j} b_{-j} + \sum_{s<0} b_k b_{-s}).$$

Since $\lambda_1^2\lambda_2 - \lambda_1^4 > 0$ and $b_j \geq 0$, we get

$$R_k(1,1) \geq (\lambda_1^2\lambda_2 - \lambda_1^4) b_k \sum_{s<0} b_{-s} \geq Ck^{-\gamma} \tag{3.7}$$

for sufficiently large k.

It remains to verify (3.6), which will follow from inequality (2.15) once we have shown that

$$\lambda_1^{u+v} Q_k(u,v) \leq C(l+1)^{\gamma+1}(p+1)^{\gamma+1} k^{-\gamma} D^{u+v}. \tag{3.8}$$

To prove (3.8), we have to estimate

$$Q_k(u,v) = \sum_{\substack{j_u<\ldots<j_1<k \\ s_v<\ldots<s_1<0: j_u=s_v}} b_{k-j_1} b_{j_1-j_2} \ldots b_{j_{u-1}-j_u} b_{-s_1} b_{s_1-s_2} \ldots b_{s_{v-2}-s_{v-1}} b_{s_{v-1}-j_u}.$$

Set $n_1 = k - j_1$, $n_2 = j_1 - j_2, \ldots, n_u = j_{u-1} - j_u$, $n_{u+1} = -s_1$, $n_{u+2} = s_1 - s_2, \ldots, n_{u+v-1} = s_{v-2} - s_{v-1}$, $n_{u+v} = s_{v-1} - s_v$, where $j_u = s_v$. Since $n_1 + \ldots + n_u - (n_{u+1} + \ldots + n_{u+v}) = k$ there exists $1 \leq i_0 \leq u+v$ such that $n_{i_0} \geq k/(u+v)$. Thus

$$b_{n_{i_0}} \leq C n_{i_0}^{-\gamma} \leq C(u+v)^{\gamma} k^{-\gamma}.$$

The indexes $n_1, \ldots, n_{i_0-1}, n_{i_0+1}, \ldots, n_{u+v}$ are linearly independent. Thus

$$Q_k(u,v) \leq C \sum_{i_0=1}^{u+v} (u+v)^{\gamma} k^{-\gamma} \prod_{\substack{i=1 \\ i\neq i_0}}^{u+v} (\sum_{n_i=1}^{\infty} b_{n_i}) = C(u+v)^{\gamma+1} k^{-\gamma} B^{u+v-1}.$$

This implies (3.8), since $(u+v)^{\gamma+1} \leq (l+1)^{\gamma+1}(p+1)^{\gamma+1}$ and

$$\lambda_1 \sum_{j=1}^{\infty} b_j \leq (\lambda_2)^{1/2} \sum_{j=1}^{\infty} b_j = D < 1.$$

∎

4 Long memory and martingale representation

We have seen in Section 3 (see Proposition 3.1) that it is not possible to construct a stationary ARCH model (1.1), (1.2) satisfying (2.5) with autocovariance function which is not absolutely summable. Long memory ARCH models are however desirable in light of the observed covariance structure of many real-life time series, see e.g. Dacorogna *et al.* (1993), Baillie *et al.* (1996), Ding and Granger (1996), Breidt *et al.* (1998). Consequently, many authors, see e.g. Robinson (1991), Baillie *et al.* (1996), Ding and Granger (1996), Robinson and Zaffaroni (1997), Robinson and Zaffaroni (1998), Robinson and Henry (1998), discussed possible ways of constructing stationary long memory ARCH models. The approach generally taken is to argue that some fractional power of the difference operator applied to the series $|r_k|^\delta$ forms a sequence of martingale differences, and hence, by analogy with linear sequences, the series $|r_k|^\delta$ exhibits long memory as defined by (1.6). In this section we take a closer look at this approach in light of the results established in the previous sections and develop a martingale representation for the model (1.1), (1.2).

To focus attention, consider the Long Memory ARCH, LM(q)-ARCH, model of Ding and Granger (1996). The model is defined as follows: $\{r_k^2\}$ is the observable time series, which may represent e.g. squares of log-returns, and $\{\varepsilon_k\}$ is a sequence of iid mean zero random variables with unit variance. Suppose that the ε_k have finite fourth moment $E\varepsilon_k^4 = \lambda_2$. The LM($q$)-ARCH equations are:

$$r_k^2 = \sigma_k^2 \varepsilon_k^2, \quad \sigma_k^2 = \sigma^2(1-\mu) + \mu \sum_{j=1}^{\infty} a_j r_{k-j}^2 \tag{4.1}$$

where $0 \le \mu \le 1$, $0 < q < 1$ and

$$a_j = \frac{q\Gamma(j-q)}{\Gamma(1-q)\Gamma(j+1)} \sim \frac{q}{\Gamma(1-q)} j^{-q-1} \quad (j \to \infty) \tag{4.2}$$

are the coefficients in the expansion of the fractional differencing operator $(1-L)^q = -\sum_0^\infty a_j L^j$; L is the lag operator defined by $LX_k = X_{k-1}$. The coefficients a_j clearly satisfy $\sum_1^\infty a_j = 1$. Ding and Granger (1996) argue that under conditions which guarantee the existence of the fourth moment $Er_t^4 < \infty$,

$$\mathrm{Cov}(r_{t+k}^2, r_t^2) \sim Ck^{2q-1}, \tag{4.3}$$

whenever $0 < q < 1/2$. The results we obtained show that the long memory property (4.3) is not valid for μ satisfying $0 \le \mu < 1/\sqrt{\lambda_2} < 1$. (If the ε_k are standard normal, (4.3) does not hold for $\mu < 1/\sqrt{3} = .577$.) Indeed, upon setting $X_k = r_k^2$, $\sigma_k^2 = \rho_k$, $\xi_k = \varepsilon_k^2$, $a = \sigma^2(1-\mu)$, $b_j = \mu a_j$, the above model coincides with our model defined by (1.1), (1.2). If $\mu < 1$ condition (2.4) holds because $\sum_{j=1}^\infty a_j = 1$ implies $\sum_{j=1}^\infty b_j = \mu$, which combined with $E\varepsilon_k^2 = 1$ implies

$$E\xi_0 \sum_{j=1}^{\infty} b_j = \mu E\xi_0 = \mu < 1.$$

Therefore, by Theorem 2.1 the unique stationary solution exists and is given by (2.2), (2.3), whenever $\mu < 1$. Under the additional assumption $\mu < 1/\sqrt{\lambda_2}$ this solution is also covariance stationary because condition (2.5) is satisfied:

$$(E\xi_0^2)^{1/2} \sum_{j=1}^{\infty} b_j = (\lambda_2)^{1/2}\mu < 1.$$

Moreover, by (4.2), $b_j \sim Cj^{-q-1}$. Thus, using Proposition 3.2, we get the rate of decay of the covariance function

$$\mathrm{Cov}(X_k, X_0) \asymp k^{-q-1},$$

which is different from (4.3).

The case $\mu = 1$ discussed by Ding and Granger (1996) has drawn particular attention because it yields equations

$$(1-L)^q r_k^2 = \nu_k, \tag{4.4}$$

where the $\nu_k = r_k^2 - \sigma_k^2$ are martingale differences (provided a stationary solution $\{r_k^2\}$ *exists*). Because of the similarity of (4.4) and fractional ARIMA equations, it is argued that for $0 < q < 1/2$ the covariance function of $\{r_k^2\}$ decays at the rate k^{2q-1}. Note that in the case $\mu = 1$ condition (2.4) for the existence of a stationary solution is not satisfied, therefore our results do not apply to this case suggesting a need for further research in this direction.

In the remainder of this section, we discuss the existence of a martingale representation for the ARCH model studied in the present paper.

Consider a general ARCH(∞) model (1.1), (1.2) and assume that condition (2.5) holds. We also assume $a > 0$, because under our assumptions, $a = 0$ implies that the only stationary solution is $X_k \equiv 0$. We reparametrize the model introducing parameters ψ_j by setting

$$\psi_j = \lambda_1 b_j.$$

Condition (2.4) is then equivalent to

$$\sum_{j=1}^{\infty} \psi_j < 1. \tag{4.5}$$

Define the function

$$\psi(z) = 1 - \psi_1 z - \psi_2 z^2 - \ldots, \tag{4.6}$$

which, by (4.5), is analytic in the open unit disc $\{z\colon |z| < 1\}$ and has no zeroes in the closed disc $\{z\colon |z| \le 1\}$. Equations (1.1),(1.2) can now be rewritten in the following form

$$X_k = \rho_k \xi_k, \ \ \rho_k = a + \sum_{j=1}^{\infty} b_j X_{k-j} \equiv a + \lambda_1^{-1}(1 - \psi(L))X_k. \tag{4.7}$$

According to Theorem 2.1, under assumption (2.5) ($(\lambda_2)^{1/2} \sum_1^{\infty} \lambda_1^{-1}\psi_j < 1$), there is a unique weakly stationary solution to equations (1.1), (1.2), given by (2.2), which by (2.8) satisfies

$$EX_k = \frac{a\lambda_1}{1 - \lambda_1 B} = \frac{a\lambda_1}{\psi(1)}. \tag{4.8}$$

Hence (4.7) can be rewritten as

$$\psi(L)(X_k - EX_k) = \nu_k, \tag{4.9}$$

where

$$\nu_k = \rho_k(\xi_k - \lambda_1), \;\; k \in \mathbf{Z} \tag{4.10}$$

is a sequence of square integrable martingale differences, i.e.

$$E(\nu_k|\mathcal{F}_{k-1}) = 0 \quad \text{and} \quad E\nu_k^2 < \infty. \tag{4.11}$$

In (4.11) $\mathcal{F}_k = \sigma\{\xi_j, j \le k\}$ is the σ-algebra generated by the random variables $\xi_k, \xi_{k-1}, \ldots$ (It is easy to verify that $E\nu_k^2 = \lambda_2^{-1}(\lambda_2 - \lambda_1^2)EX_0^2$.)

Note that by condition (4.5) the function $\psi(z)$ has absolutely summable coefficients, and both functions $\psi(z)$ and $1/\psi(z)$ are analytic in the open unit disc. It is well known that if $\psi(z)^{-1} = \sum_{j=0}^{\infty} \tilde{\psi}_j z^j$ has square summable coefficients $\tilde{\psi}_j$, then X_k has a causal (nonanticipating) moving average representation

$$X_k = EX_k + \psi^{-1}(L)\nu_k \equiv EX_k + \sum_{j=0}^{\infty} \tilde{\psi}_j \nu_{k-j} \tag{4.12}$$

(see Brockwell and Davis (1991) Theorem 4.10.1).

We introduce now a bit stronger assumptions on function $\psi(z)$. Note that if $\psi(z)$ is analytic in the disc $|z| < 1 + \delta$ for some $\delta > 0$ then $\psi^{-1}(z)$ has absolutelly sumable coefficients $\tilde{\psi}_j$.

Proposition 4.1 *Suppose assumption*

$$(\lambda_2)^{1/2} \sum_{j=1}^{\infty} \lambda_1^{-1}\psi_j < 1$$

holds and the function $\psi(z)$ is analytic and has no zeroes in the disc $|z| < 1 + \delta$ for some $\delta > 0$. Then the solution $\{X_k\}$ of the ARCH equations (1.1) *and* (1.2) *admits representation* (4.12) *with the ν_k satisfying* (4.11).

Recall that Proposition 4.1 cannot be applied to the operator $\psi(L) = (1 - L)^d$, $0 < d < 1/2$ because $\psi(1) = 1 - \lambda_1 \sum_1^{\infty} b_j = 0$, what violates condition (2.4).

The following example discusses a simple but interesting application of Proposition 4.1.

Example 4.1 Consider the operator

$$\psi(L) = (1 - rL)^{\alpha},$$

where $0 < r < 1$ and $0 < \alpha < 1$. Then condition (2.4) is satisfied and assumption (2.5) reads as follows

$$(\lambda_2)^{1/2}\lambda_1^{-1}(1 - (1 - r)^{\alpha}) < 1. \tag{4.13}$$

In this case

$$\tilde{\psi}_j = \frac{\Gamma(j+\alpha)}{\Gamma(\alpha)\Gamma(j+1)}\, r^j \sim \Gamma^{-1}(\alpha) j^{\alpha-1} r^j \quad (j \to \infty),$$

thus $\sum_{j=0}^{\infty} \tilde{\psi}_j < \infty$ and by Proposition 4.1 X_k admits representation (4.12), which implies that

$$\mathrm{Cov}(X_k, X_0) = E\nu_0^2 \sum_{j=0}^{\infty} \tilde{\psi}_j \tilde{\psi}_{j+|k|}.$$

One can easily verify that

$$\mathrm{Cov}(X_k, X_0) \sim K k^{\alpha-1} r^k. \tag{4.14}$$

Indeed, for $k > 0$

$$\frac{\mathrm{Cov}(X_k, X_0)}{k^{\alpha-1} r^k} = E\nu_0^2 \sum_{j=0}^{\infty} \tilde{\psi}_j \, \frac{\tilde{\psi}_{j+k}}{k^{\alpha-1} r^k}.$$

Since

$$\frac{\tilde{\psi}_{j+k}}{k^{\alpha-1} r^k} \to \Gamma^{-1}(\alpha) r^j \text{ as } k \to \infty,$$

it follows from the Dominated Convergence Theorem that (4.14) holds with

$$K = E\nu_0^2 \, \Gamma^{-1}(\alpha) \sum_{j=0}^{\infty} \tilde{\psi}_j r^j = E\nu_0^2 \, \Gamma^{-1}(\alpha)(1-r^2)^{-\alpha}.$$

Analogously to the known term "semi heavy-tails" applied to distribution densities asymptotically behaving as $x^{-\alpha}\beta^{-x}$ ($\alpha > 0$, $\beta > 0$) when $x \to \infty$, the dependence structure in the example above could be characterized as "semi long memory".

The model described in Example 4.1 has empirical relevance, see e.g. Figure 6 in Dacorogna *et al.* (1993), which shows that the absolute value of the returns on USD-DEM exchange rate have autocorrelation function which exhibits behavior similar to the one given in (4.14).

5 Central Limit Theorem

We prove in this section the Central Limit Theorem (CLT) for the ARCH(∞) sequence $\{X_k\}$.

It is well-known that for moving averages $X_k = \sum_{j=-\infty}^{\infty} a_{k-j}\varepsilon_j$, where $\sum_k a_k^2 < \infty$ and ε_j are iid random variables with $E\varepsilon_j = 0$, $E\varepsilon_j^2 = 1$, or Gaussian processes X_k, the condition $\sum_k |\mathrm{Cov}(X_k, X_0)| < \infty$, implies the CLT: $N^{-1/2}\sum_1^N X_k \Rightarrow N(0, \sigma^2)$ with $\sigma^2 = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(X_k, X_0)$ (see Theorem 18.6.5 in Ibragimov and Linnik (1971)).

Theorem 5.1 shows that although the ARCH(∞) sequence (2.2) has much more complicated structure than moving averages, the short memory condition (3.2) still effectively controls its dependence structure and implies the CLT. For the CLT to hold neither the rate of decay of the b_j need be specified nor any other "short memory assumptions", like e.g. mixing conditions, need be imposed.

The fact that the CLT holds for the ARCH sequences studied in this paper can be interpreted as yet another manifestation of their "short memory nature"; partial sums of long memory sequences converge at a rate different than $\sqrt{N}$, c.f. (5.2) below.

In the present paper, we prove the convergence of finite dimensional distributions.

THEOREM 5.1 *Suppose that $\{X_k\}$ is an ARCH(∞) sequence satisfying* (1.1), (1.2) *and assumption* (2.5), *i.e.*

$$(E\xi_0^2)^{1/2}\sum_{j=1}^{\infty} b_j < 1. \tag{5.1}$$

Then, as $N \to \infty$

$$N^{-1/2}\sum_{j=1}^{[Nt]}(X_j - EX_j) \Rightarrow \sigma W(t) \quad (0 \le t \le 1), \tag{5.2}$$

where

$$\sigma^2 = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(X_k, X_0) < \infty \tag{5.3}$$

and $\{W(t), 0 \le t \le 1\}$ is the standard Wiener process with zero mean and covariances $EW(t)W(s) = \min(t,s)$. Here $\Rightarrow$ denotes the convergence of finite dimensional distributions.

PROOF: By Theorem 2.1 the unique stationary solution to (1.1), (1.2) exists and can be represented as the sum (2.2). Fix $L > 1$ large enough and split X_k (2.2) into two parts:

$$X_k \equiv a\sum_{l=0}^{\infty} M_l(k) = a\sum_{l=0}^{L} M_l(k) + a\sum_{l=L+1}^{\infty} M_l(k)$$

$$= \{a\sum_{l=0}^{L} M_l^-(k)\} + \{a\sum_{l=0}^{L} M_l^+(k) + a\sum_{l=L+1}^{\infty} M_l(k)\} =: X_k^- + X_k^+.$$

Here $M_l^-(k), M_l^+(k)$ are defined by splitting $M_l(k)$ as

$$M_l(k) = \sum_{j_l<\ldots<j_1<k} b_{k-j_1}b_{j_1-j_2}\ldots b_{j_{l-1}-j_l}\xi_k\xi_{j_1}\ldots\xi_{j_l}$$

$$= \sum_{j_l<\ldots<j_1<k}[\ldots]\mathbf{1}_{\{k-j_1\le L, j_1-j_2\le L,\ldots,j_{l-1}-j_l\le L\}}$$

$$+ \sum_{j_l<\ldots<j_1<k}[\ldots](1-\mathbf{1}_{\{k-j_1\le L, j_1-j_2\le L,\ldots,j_{l-1}-j_l\le L\}}) =: M_l^-(k) + M_l^+(k).$$

Set

$$S_{N,l}(t) = N^{-1/2}\sum_{j=1}^{[Nt]} a(M_l(k) - EM_l(k)),$$

$$S_{N,l}^-(t) = N^{-1/2}\sum_{j=1}^{[Nt]} a(M_l^-(k) - EM_l^-(k)), \; S_{N,l}^+(t) = N^{-1/2}\sum_{j=1}^{[Nt]} a(M_l^+(k) - EM_l^+(k))$$

and rewrite

$$N^{-1/2}\sum_{j=1}^{[Nt]}(X_j - EX_j) \equiv \sum_{l=0}^{\infty} S_{N,l}(t) = \{\sum_{l=0}^{L} S^-_{N,l}(t)\} + \{\sum_{l=0}^{L} S^+_{N,l}(t) + \sum_{l=L+1}^{\infty} S_{N,l}(t)\}$$

$$=: S^-_N(t) + S^+_N(t).$$

To prove (5.2) it is sufficient to show that for any fixed $L > 0$,

$$S^-_N(t) \Rightarrow \sigma^2_- W(t) \quad (N \to \infty) \tag{5.4}$$

with $\sigma^2_- := \sum_{k=-\infty}^{\infty} \mathrm{Cov}(X^-_k, X^-_0)$,

$$\sigma^2_- \to \sigma^2 = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(X_k, X_0) \quad (L \to \infty) \tag{5.5}$$

and that for any fixed $t > 0$, $\epsilon > 0$

$$P\{|S^+_N(t)| > \epsilon\} \to 0 \quad (L \to \infty) \tag{5.6}$$

uniformly in N.

To prove (5.4) note that

$$S^-_N(t) = N^{-1/2}\sum_{k=1}^{[Nt]}(X^-_k - EX^-_k),$$

where $X^-_k = a\sum_{l=0}^{L} M^-_l(k)$ is a stationary sequence of m-dependent random variables with $m = L^2$, i.e. X^-_k, X^-_s are independent if $|k-s| > L^2$. This follows from the definition of $M^-_l(k)$ and the fact that the ξ_k are independent. Since the random variables X^-_k have finite second moment, (5.4) follows from the CLT for an m-dependent sequence (see Billingsley (1968), Theorem 20.1).

We prove now (5.5). The covariance $\mathrm{Cov}(X^-_k, X^-_0)$ has a representation analogous to (2.10), (2.11) with the b_j replaced by $b_j\mathbf{1}\{j \le L\}$. Since for any k, $\mathrm{Cov}(X^-_k, X^-_0) \to \mathrm{Cov}(X_k, X_0)$ monotonically in L

$$\sigma^2_- := \sum_{k=-\infty}^{\infty} \mathrm{Cov}(X^-_k, X^-_0) \to \sum_{k=-\infty}^{\infty} \mathrm{Cov}(X_k, X_0) = \sigma^2 \ \ (L \to \infty).$$

It remains to verify (5.6). It is sufficient to prove that for any fixed $t > 0$

$$\mathrm{Var}\, S^+_N(t) \to 0 \ \ (L \to \infty) \tag{5.7}$$

uniformly in N. From the definitions of X^-_k and X^+_0 it follows that covariances $\mathrm{Cov}(X^-_k, X^-_s)$, $\mathrm{Cov}(X^+_k, X^-_s)$ $\mathrm{Cov}(X^+_k, X^+_s)$ are nonnegative. Thus

$$\mathrm{Cov}(X_k, X_s) = \mathrm{Cov}(X^+_k + X^-_k, X^+_s + X^-_s) \ge \mathrm{Cov}(X^+_k, X^+_s) + \mathrm{Cov}(X^-_k, X^-_s)$$

and hence

$$\operatorname{Cov}(X_k^+, X_s^+) \le \operatorname{Cov}(X_k, X_s) - \operatorname{Cov}(X_k^-, X_s^-).$$

This implies

$$\operatorname{Var} S_N^+(t) = N^{-1} \sum_{k,s=1}^{[Nt]} \operatorname{Cov}(X_k^+, X_s^+) \le N^{-1} \sum_{k,s=1}^{N} (\operatorname{Cov}(X_k, X_s) - \operatorname{Cov}(X_k^-, X_s^-))$$

$$\le \sum_{k=-\infty}^{\infty} \operatorname{Cov}(X_k, X_0) - \sum_{k=-\infty}^{\infty} \operatorname{Cov}(X_k^-, X_0^-) = \sigma^2 - \sigma_-^2 \to 0 \quad (L \to \infty)$$

by (5.5). This completes proof of (5.7). ∎

Acknowledgement. The authors would like to thank the referees for their insightful comments and many useful suggestions.

References

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, **73,** 5–59.

Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **74,** 3–30.

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

Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis*, **9,** 121–131.

Bougerol, P. and Picard, N. (1992). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics*, **52,** 115–127.

Breidt, F. J., Crato, N. and de Lima, P. (1998). On the detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, **83,** 325–348.

Brockwell, P. J. and Davis, R. A. (1991). *Time series: Theory and methods*, Second edn. Springer-Verlag, New York.

Dacorogna, M. M., Müller, U.A., Nagler, R. J., Olsen, R. B. and Pictet, O. V. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance*, **12,** 413–438.

Ding, Z. and Granger, C. W. J. (1996). Modeling volatility persistence of speculative returns: a new approach. *Journal of Econometrics*, **73,** 185–215.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance.* Springer–Verlag, Berlin.

He, C. and Teräsvirta, T. (1999). Fourth moment structure of the GARCH(p, q) process. *Econometric Theory (forthcoming)*, **15,** 824–846.

Ibragimov, I. A. and Linnik, Yu. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters–Noordhoff, Groningen.

Kokoszka, P. S. and Leipus, R. (2000). Change-point estimation in ARCH models. *Bernoulli*, **6,** 1–28.

Nelson, D. B. (1990). Stationarity and persistence in the GARCH$(1, 1)$ model. *Econometric Theory*, **6,** 318–334.

Nelson, D. B. and Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model. *Journal of Business and Economic Statistics*, **10,** 229–235.

Priestley, M. B. (1988). *Non-linear and non-stationary time series analysis.* Academic Press, New York.

Robinson, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics*, **47,** 67–84.

Robinson, P. M. and Henry, M. (1998). Long memory conditionally heteroscedastic errors in semiparametric estimation of long memory. Preprint.

Robinson, P. M. and Zaffaroni, P. (1997). Modelling nonlinearity and long memory in time series. *Fields Institute Communications*, **11,** 161–170.

Robinson, P. M. and Zaffaroni, P. (1998). Nonlinear time series with long memory: a model for stochastic volatility. *Journal of Statistical Planning and Inference*, **68,** 359–371.