A MODEL FOR LONG MEMORY CONDITIONAL HETEROSCEDASTICITY

by

Liudas Giraitis* and Peter Robinson**
London School of Economics and Political Science

and

Donatas Surgailis
Institute of Mathematics and Informatics, Vilnius, Lithuania

<u>Contents:</u>

Discussion Paper
No. EM/00/382
March 2000

The Suntory Centre
Suntory and Toyota International Centres
for Economics and Related Disciplines
London School of Economics and Political
Science
Houghton Street
London WC2A 2AE
Tel.: 020-7405 7686

---

* Research supported by ESRC Grant R000235892.
** Research supported by ESRC Grant R000235892 and a Leverhulme Trust Personal Research Professorship.

NB: The authors regret that the missing figures on page 13 could not be included in this PDF; they are only available in the printed version.

Abstract

In Giraitis, Robinson, and Samarov (1997), we have shown that the optimal rate for memory parameter estimators in semiparametric long memory models with degree of 'local smoothness' β is $n^{-r(\beta)}$, $r(\beta) = \beta/(2\beta+1)$, and that a log-periodogram regression estimator (a modified Geweke and Porter-Hudak (1983) estimator) with maximum frequency $m = m(\beta) \asymp n^{2r(\beta)}$ is rate optimal. The question which we address in this paper is what is the best obtainable rate when β is unknown, so that estimators cannot depend on β. We obtain a lower bound for the asymptotic quadratic risk of any such adaptive estimator, which turns out to be larger than the optimal nonadaptive rate $n^{-r(\beta)}$ by a logarithmic factor. We then consider a modified log-periodogram regression estimator based on tapered data and with a data-dependent maximum frequency $m = m(\hat{\beta})$, which depends on an adaptively chosen estimator $\hat{\beta}$ of β, and show, using methods proposed by Lepskii (1990) in another context, that this estimator attains the lower bound up to a logarithmic factor. On one hand, this means that this estimator has nearly optimal rate among all adaptive (free from β) estimators, and, on the other hand, it shows near optimality of our data-dependent choice of the rate of the maximum frequency for the modified log-periodogram regression estimator. The proofs contain results which are also of independent interest: one result shows that data tapering gives a significant improvement in asymptotic properties of covariances of discrete Fourier transforms of long memory time series, while another gives an exponential inequality for the modified log-periodogram regression estimator.

Keywords: Long range dependence; semiparametric model; rates of convergence; adaptive bandwidth selection.
JEL classification nos: C13.

1. Introduction.

A principle stylized fact emerging from the analysis of many financial time series (such as asset returns and exchange rates) is the approximate uncorrelatedness of the 'return' series r_t (often a first difference of logarithms of the basic observed series) alongside pronounced autocorrelation in certain instantaneous nonlinear functions of r_t, such as r_t^2. Such behaviour is consistent with the property that the conditional mean is zero (almost surely),

$$E(r_t|\mathcal{G}_{t-1}) = 0, \tag{1.1}$$

where $\mathcal{G}_t$ is the σ- field of events generated by $r_s, s \leq t$, whereas the conditional variance

$$\sigma_t^2 = \mathrm{Var}(r_t|\mathcal{G}_{t-1}) \tag{1.2}$$

is stochastic.

The earliest models of this form assumed that

$$\sigma_t^2 = a + \sum_{j=1}^{\infty} b_j r_{t-j}^2, \quad t \in \mathbf{Z}, \tag{1.3}$$

for constants $a > 0$ and $b_j \geq 0$ (to ensure that $\sigma_t^2 > 0$), where the b_j also satisfy some summability condition, easily achieved in both the ARCH(p) model of Engle (1982) (wherein $b_j = 0, j > p$) and its GARCH extension of Bollerslev (1986). However, these latter models imply exponential decay in the autocorrelations of the r_t^2, whereas empirical evidence has frequently suggested a much greater degree of persistence, possibly consistent with long memory in r_t^2, where autocorrelations are not summable (see e.g. Whistler, 1990, Ding, Granger and Engle, 1993). Such behaviour could arise from heavy - tailedness or structural breaks (see e.g. Davis and Mikosch (1999), Lobato and Savin, 1998), but it might also be explained by (1.3), since considerable flexibility is possible in the choice of the b_j. Robinson (1991) referred to the possibility of b_j in (1.3) that correspond to long memory in r_t^2 and developed tests for no - ARCH with optimal efficiency against parametric alternatives in the class (1.3), while Granger and Ding (1995), Ding and Granger (1996) have discussed such models further. On the other hand, the sufficient conditions established by Giraitis, Kokoszka and Leipus (1998) for existence of a covariance stationary solution in versions of (1.1), (1.2) given by

$$r_t = \varepsilon_t \sigma_t, \tag{1.4}$$

where ε_t is an independent and identically distributed (iid) sequence having suitable moments, and σ_t is the positive square root of σ_t^2 in (1.3), rule out long memory autocorrelation in r_t^2, so that a full account of the long memory potential of (1.3) is lacking.

Fortunately it is easy to find alternative models for which conditions for stationary long memory of squares and other instantaneous functions are available. In particular in models of form (1.4) with

$$\sigma_t = f(\eta_t), \tag{1.5}$$

where η_t is a possibly vector-valued, possibly Gaussian, unobservable long memory process, the memory properties of instantaneous functions such as r_t^l for integer $l \geq 2$, or $|r_t|^\alpha$ for real $\alpha > 0$, depend on the character of the function f. Models of this type with long memory properties have already been discussed by, for example, Andersen and Bollerslev (1997), Breidt, Crato and De Lima (1998), Harvey (1998), Robinson and Zaffaroni (1997, 1998).

Here we consider the long memory potential of an alternative class of models of form (1.4) that is more similar to the ARCH form (1.3). We consider the model, which one might call LARCH ("Linear ARCH"),

$$\sigma_t = a + \sum_{j=1}^{\infty} b_j r_{t-j}, \quad t \in \mathbf{Z}. \tag{1.6}$$

Thus with (1.4), we have a special case of the model consisting of (1.1) and

$$\sigma_t^2 = (a + \sum_{j=1}^{\infty} b_j r_{t-j})^2, \quad t \in \mathbf{Z}, \tag{1.7}$$

for the first and second conditional moments that was considered by Robinson (1991) (equation (16)). Indeed (1.6) with (1.4) is also a special case of the general class of bilinear models referred to by Granger and Andersen (1978) (equation (4.1)), though these authors and the subsequent literature on bilinear time series models, focussed on forms that specifically exclude the combination of (1.6) with (1.4). Robinson (1991) contrasted the implications for third moment behaviour of r_t under (1.3) and (1.6). Notice also that (1.6), unlike (1.3), is not constrained to be non-negative, so that σ_t is not a standard deviation and lacks something of the usual volatility interpretation. However constraints on a and b_j, of the type needed for (1.3), are not thereby necessary, leading to some convenience of theoretical analysis. Whereas Robinson (1991) considered weights b_j of long memory type in (1.7), this was in connection with testing for no-ARCH against general parametric alternatives of form (1.2), including short memory ones. Short memory versions of (1.6) (such as when $b_j = 0, j > p$) may deserve further study, but our results, except for Theorem 2.1, focus on long memory type b_j. Here, we examine the structure of σ_t (Theorem 2.1) and its possible long memory behaviour (Corollary 2.1), and give conditions under which powers r_t^l, for integer $l \geq 2$, have long memory autocorrelation (Theorem 2.2) and their normalised partial sums converge to fractional Brownian motion (Theorem 2.3). These results and the relevant conditions are presented in the following section, which also gives the proofs of Theorem 2.1, Corollary 2.1 and Theorem 2.3, but only the main steps of the proofs of Theorem 2.2, the remaining details appearing in the following three sections of the paper.

2. Main results

We introduce first

Assumption 1 (i) (1.4) and (1.6) hold.

(ii) $\{\varepsilon_t\}$ is a sequence of iid random variables with zero mean and unit variance.

(iii) $a \neq 0$.

Assumption 2.

$$b = \{\sum_{j=1}^{\infty} b_j^2\}^{1/2} < 1.$$

Let $\mathcal{F}_t$ be the σ- field of events generated by ε_s, $s \leq t$.

Theorem 2.1. *Let Assumption 1 hold. Then a covariance stationary $\mathcal{F}_{t-1}$-measurable solution $\sigma_t, t \in \mathbf{Z}$, of (1.4), (1.6) exists if and only if Assumption 2 holds, in which case, for $t \in \mathbf{Z}$, we have the Volterra expansion*

$$\sigma_t = a \sum_{k=0}^{\infty} \sum_{j_1,\ldots,j_k=1}^{\infty} b_{j_1} \ldots b_{j_k} \varepsilon_{t-j_1} \ldots \varepsilon_{t-j_1-\ldots-j_k}, \tag{2.1}$$

and

$$E(\sigma_t) = a, \tag{2.2}$$

$$\mathrm{Cov}(\sigma_0, \sigma_t) = \frac{a^2}{1-b^2} \sum_{j=1}^{\infty} b_j b_{j+t}. \tag{2.3}$$

Proof. If σ_t is a covariance stationary $\mathcal{F}_{t-1}$-measurable solution of (1.4), (1.6) then r_t is also covariance stationary with $E(r_0) = 0$, $\mathrm{Cov}(r_0, r_t) = 0$, $a \neq 0$. Thus $E(\sigma_0) = a$ and

$$E(\sigma_0^2) = a^2 + b^2 E(\sigma_0^2), \tag{2.4}$$

to give the first statement of the theorem. We thus have, under Assumption 2,

$$E(r_0^2) = E(\sigma_0^2) = \frac{a^2}{1-b^2}. \tag{2.5}$$

We also deduce from (1.6) and stationarity that

$$\mathrm{Cov}(\sigma_0, \sigma_t) = E(r_0^2) \sum_{j=1}^{\infty} b_j b_{j+t} \tag{2.6}$$

to give (2.3). Finally (2.1) is obtained by iteration of (1.4), (1.6) as in Nelson (1990), Giraitis, Kokoszka and Leipus (1998), and is clearly also strictly stationary. ■

The iid requirement can be relaxed to a martingale difference one, on the ε_t and $\varepsilon_t^2 - 1$. There is no loss of generality in fixing $\mathrm{Var}(\varepsilon_0) = 1$. If Assumption 1(iii) does not hold, so $a = 0$, we deduce from (2.4) that $b = 1$, so Assumption 2 cannot hold. Then, for example, in case $b_1 = 1, b_j = 0,\ j > 1$, we have instead, subject to convergence, $\sigma_t = \prod_{j=1}^{\infty} \varepsilon_{t-j}$, which is a sequence of uncorrelated variables with zero mean and unit variance, as is r_t; trivially $r_t = \sigma_t \equiv 0$ is also a solution. Hence we discuss only the case $a \neq 0$. The Volterra expansion (2.1) plays a basic role in the proofs of Theorems 2.2 and 2.3 below.

From (2.3), (2.5) we can also write

$$Corr(\sigma_0, \sigma_t) = \frac{\sum_{j=1}^{\infty} b_j b_{j+t}}{b^2},$$

which we recognize as the usual formula for the autocorrelation function in terms of Wold decomposition weights. We can thus control the memory of σ_t by choice of b_j. We introduce

Assumption 3. For

$$0 < c < \infty, \quad 0 < \theta < 1, \tag{2.7}$$

we have

$$b_t \sim c t^{-(1+\theta)/2}, \quad as \quad t \to \infty, \tag{2.8}$$

where $' \sim '$ indicates that the ratio of left and right sides tends to 1.

Corollary 2.1. Let Assumptions 1-3 hold. Then

$$\mathrm{Cov}(\sigma_0, \sigma_t) \sim c_1^2 t^{-\theta}, \quad as \quad t \to \infty \tag{2.9}$$

where

$$c_1 = ac\{\frac{B(\frac{1-\theta}{2}, \theta)}{1 - b^2}\}^{1/2}. \tag{2.10}$$

Proof. Standard from the long memory literature, using (2.8) and (2.3). ■

An example of b_t satisfying Assumption 3 is

$$b_t = c\frac{\Gamma(t + \frac{1-\theta}{2})}{\Gamma(t+1)}, \tag{2.11}$$

which is proportional to the moving average weights in a standard fractional ARIMA$(0, \frac{1}{2}(1-\theta), 0)$ model (see e.g. Adenstedt, 1974, Samorodnitsky and Taqqu, 1994, p.381), so that (1.6) becomes $\sigma_t = a + ((1-L)^{(\theta-1)/2} - 1)r_t$, L being the lag operator. More general b_t include the fractional ARIMA$(p, \frac{1}{2}(1-\theta), q)$ weights. Notice that many of the latter models have b_j that are not all non-negative, and so could not be used in connection with (1.3).

It will be found that, for integer $l \geq 2$, r_t^l has autocorrelations decaying at the same rate as those of σ_t when σ_t has long memory, and that the normalized partial sums of r_t^l (like those of σ_t) converge to fractional Brownian motion. Notice that typically σ_t is unobservable, whereas r_t is observable, so that its autocovariances can likely be consistently estimated under suitable conditions. There is thus the possibility of drawing inferences on the presence and extent of long memory in r_t^l. The choice of l likely to be of most interest to empirical workers is $l = 2$, especially as finiteness of low order moments of financial time series has frequently been questioned. However, subject to finiteness of moments, the extent to which our approximation to $\mathrm{Corr}(r_0^l, r_t^l)$ depends of l may be helpful in validating the model from real data.

To establish the properties of r_t^l, we impose also:

Assumption 4(l). ε_t has finite $2l^{th}$ moments such that

$$(4^l - 2l - 1)\mu_{2l}^{1/l} b^2 < 1, \tag{2.12}$$

where $\mu_j = E(\varepsilon_0^j)$.

For given l, (2.12) is a tighter restriction on b than Assumption 2, that is, a tighter restriction on c in case (2.10), while (2.12) becomes more stringent as l increases, since $(4^l - 2l - 1)$ and $\mu_{2l}^{1/l}$ are increasing functions, so that (2.12) holds also for $j < l$. When ε_t is Gaussian $\mu_{2l} = (2l-1)(2l-3)\ldots 3 \cdot 1$, though in this case is likely that the factor $(4^l - 2l - 1)$ can be reduced, (2.12) being only a sufficient condition for the following results.

Theorem 2.2. *Let Assumptions 1, 2, 3 and 4(l) hold. Then, for $j = 2, \ldots, l$*

$$\mathrm{Cov}(r_0^j, r_t^j) \sim c_j^2 t^{-\theta}, \text{ as } t \to \infty, \tag{2.13}$$

where

$$c_j = \frac{c_1}{a} j E(r_0^j).$$

Proof. It suffices to take $j = l$. Write $\nu_{tl} = (\varepsilon_t^l - \mu_l)\sigma_t^l$, so

$$r_t^l = \varepsilon_t^l \sigma_t^l = \mu_l \sigma_t^l + \nu_{tl}. \tag{2.14}$$

Since we may write

$$c_l = \mu_l \frac{c_1 l}{a} E(\sigma_0^l)$$

and Assumptions 1 and 4(l) imply that $\mathrm{Cov}(\nu_{0l}, \nu_{tl}) = \mathrm{Cov}(\sigma_0^l, \nu_{tl}) = 0$ for $t > 0$, it suffices to show that

$$\mathrm{Cov}(\sigma_0^l, \sigma_t^l) \sim c_1^2 d_l^2 t^{-\theta} \tag{2.15}$$

where

$$d_l = \frac{lE(\sigma_0^l)}{a}$$

and that

$$\mathrm{Cov}(\nu_{0l}, \sigma_t^l) = o(t^{-\theta}). \tag{2.16}$$

To consider (2.15), introduce the 'remainder' term

$$y_{tl} = \sigma_t^l - d_l \sigma_t. \tag{2.17}$$

Then (2.15) will be a consequence of Corollary 2.1 and

$$\mathrm{Cov}(y_{0l}, y_{tl}) = o(t^{-\theta}), \quad \mathrm{Cov}(y_{0l}, \sigma_t) = o(t^{-\theta}), \quad \mathrm{Cov}(\sigma_0, y_{tl}) = o(t^{-\theta}).$$

These are easy consequences of

$$\mathrm{Cov}(\sigma_0^{l'}, \sigma_t^{l''}) \sim d_{l'} d_{l''} \mathrm{Cov}(\sigma_0, \sigma_t), \quad 1 \leq l', l'' \leq l, \tag{2.18}$$

and Corollary 2.1, noting that $d_1 = 1$. To show (2.18) we introduce an 'intermediate' term

$$\zeta_{t,j} := j \sum_{k=1}^{\infty} \sum_{s_k < \ldots < s_1 < t} a_{t-s_1,j} b_{s_1 - s_2} \ldots b_{s_{k-1} - s_k} \varepsilon_{s_1} \ldots \varepsilon_{s_k}, \tag{2.19}$$

where

$$a_{t,j} := aE[\sigma_0^{j-1}] b_t + \sum_{0<s<t} G_{t-s,j} b_s, \tag{2.20}$$

$$G_{t-s,j} := H_{t-s,j} - H_{t-s-1,j}, \tag{2.21}$$

$$H_{t-s,j} = E(\sigma_t^{j-1} E(\sigma_t | \mathcal{F}_s^+)), \tag{2.22}$$

$\mathcal{F}_t^+$ being the σ-algebra of events generated by ε_s, $s \geq t$. Then we prove (2.18) by showing that

$$\mathrm{Cov}(\sigma_0^{l'}, \sigma_t^{l''}) \sim \mathrm{Cov}(\zeta_{0,l'}, \zeta_{t,l''}) \tag{2.23}$$

and

$$\mathrm{Cov}(\zeta_{0,l'}, \zeta_{t,l''}) \sim d_{l'} d_{l''} \mathrm{Cov}(\sigma_0, \sigma_t). \tag{2.24}$$

We prove (2.23) in Lemma 5.1 and (2.24) in Corollary 4.3. Finally, (2.16) is proved in Lemma 4.4.

The proof of Theorem 2.2 rests on the approximations

$$\sigma_t^l \asymp d_l \sigma_t, \quad r_t^l \asymp \mu_l d_l \sigma_t, \tag{2.25}$$

'$\asymp$' meaning that left and right sides have the same autocovariance function, at long lags j to order $o(j^{-\theta})$. The typical dominance of the linear term in approximating the autocovariance of stochastic volatility models also arose in Andersen and Bollerslev (1997), Robinson and Zaffaroni (1997, 1998). On the other hand, Ding and Granger (1996) found significant variation with α in sample autocorrelations of $|r_t|^\alpha$ computed from stock returns and exchange rates. To the extent that this phenomenon pertains to long lags, Theorem 2.2 can only explain it in respect of the asymptotic scale factor c_l^2 of $\mathrm{Cov}(r_0^l, r_t^l)$ which varies with l), not in respect of the decay rate $t^{-\theta}$ (which is constant with respect to l).

Nevertheless the approximations (2.25) are quite remarkable and also provide the leading term in the limit distribution of normalized partial sums of the r_t^l. Let $W_\theta(t), t \geq 0$ be fractional Brownian motion, that is a zero-mean Gaussian process with covariance

$$EW_\theta(s)W_\theta(t) = \frac{1}{2}(|s|^{2-\theta} + |t|^{2-\theta} - |t-s|^{2-\theta})$$

(see Samorodnitsky and Taqqu (1994), Chapter 7). Let $[\![\cdot]\!]$ denote integer part, and $\Rightarrow$ the convergence of finite dimensional distributions.

Theorem 2.3 *Under Assumptions 1,2,3 and 4(l), for $j = 2, \ldots, l$, as $N \to \infty$*

$$N^{\theta/2-1} \sum_{s=1}^{[\![Nt]\!]} (r_s^j - E[r_s^j]) \;\Rightarrow\; \chi_\theta c_j W_\theta(t), \quad t \geq 0, \tag{2.26}$$

where

$$\chi_\theta = \{\frac{2}{(1-\theta)(2-\theta)}\}^{1/2}.$$

Proof. Again we can take $j = l$. Considering again (2.14), from uncorrelatedness of ν_{tl},

$$\mathrm{Var}(\sum_{t=1}^{N} \nu_{tl}) \leq \mu_{2l} E(\sigma_0^{2l}) N = O(N) = o(N^{2-\theta}),$$

$E(\sigma_0^{2l})$ being finite from Lemma 3.1 (replacing l by $2l$ there and noting assumption $4(l)$), so we can replace r_s^l by $\mu_l \sigma_s^l$. Now employing again (2.17), Corollary 5.3 below implies that $\mathrm{Var}(\sum_{t=1}^N y_{tl}) = o(N^{2-\theta})$, so it remains to show that

$$N^{\theta/2-1} \sum_{s=1}^{[\![Nt]\!]} (\sigma_s - a) \Rightarrow \chi_\theta c_1 W_\theta(t). \tag{2.27}$$

For $K > 0$, (1.4) and (1.6) give

$$\begin{aligned} \sigma_t - a = \sum_{s<t} b_{t-s}\varepsilon_s\sigma_s &= \sum_{s<t} b_{t-s}\varepsilon_s E[\sigma_s|\mathcal{F}_{s-K}^+] + \sum_{s<t} b_{t-s}\varepsilon_s(\sigma_s - E[\sigma_s|\mathcal{F}_{s-K}^+]) \\ &=: z_t^- + z_t^+. \end{aligned}$$

Thus,

$$Z_N(t) := \sum_{s=1}^{[\![Nt]\!]}(\sigma_s - a) = \sum_{s=1}^{[\![Nt]\!]} z_s^- + \sum_{s=1}^{[\![Nt]\!]} z_s^+ =: Z_N^-(t) + Z_N^+(t).$$

We show first that the term $Z_N^+ := Z_N^+(1)$ is negligible. We have

$$N^{\theta-2}\mathrm{Var}(Z_N^+) = \{N^{\theta-2}\sum_{t',t''=1}^{N}\sum_{s<t'\wedge t''} b_{t'-s}b_{t''-s}\}E[\varepsilon_0^2(\sigma_0 - E[\sigma_0|\mathcal{F}_{-K}^+]))^2] \tag{2.28}$$

where the factor in braces is, from Corollary 2.1,

$$\frac{c_1^2}{E(r_0^2)}N^{\theta-2}\sum_{t',t''=1}^{N}|t'-t''|^{-\theta}(1+o(1)) \to \frac{2c_1^2}{E(r_0^2)}\int_0^1 (1-x)^{-\theta}xdx = \frac{c_1^2\chi_\theta^2}{E(r_0^2)}. \tag{2.29}$$

Thus (2.28) is $O(\delta_K)$, where

$$\delta_K := E[(\sigma_0 - E[\sigma_0|\mathcal{F}_{-K}^+])^2] \to 0 \quad (K\to\infty).$$

Hence (2.21) follows from

$$N^{\theta/2-1}Z_N^-(t) \Rightarrow d_K W_\theta(t), \tag{2.30}$$

if $d_K^2 := \lim_{N\to\infty} N^{\theta-2}\sum_{i,j=1}^N \mathrm{Cov}(z_i^-, z_j^-)$ satisfies

$$\lim_{K\to\infty} d_K = \chi_\theta c_1. \tag{2.31}$$

To prove (2.31), using the fact that $\eta_t := \varepsilon_t E[\sigma_t|\mathcal{F}_{t-K}^+], t\in \mathbf{Z}$, are uncorrelated, we obtain

$$\begin{aligned} d_K^2 &= E[\eta_0^2]\lim_{N\to\infty} N^{\theta-2}\{\sum_{t',t''=1}^{N}\sum_{s<t'\wedge t''} b_{t'-s}b_{t''-s}\} \\ &= E[\eta_0^2]\frac{c_1^2\chi_\theta^2}{E(r_0^2)} \end{aligned}$$

from (2.29), where

$$E[\eta_0^2] = E[\varepsilon_0^2]E[(E[\sigma_0|\mathcal{F}_{-K}^+])^2] \to E[\sigma_0^2] = E(r_0^2) = a^2/(1-b^2) \quad (K\to\infty).$$

To prove the convergence (2.29), note that z_t^- is a of form $z_t^- = \sum_{s<t} b_{t-s}\eta_s$ where η_s, $s\in\mathbf{Z}$ is a stationary sequence of uncorrelated K-dependent random variables. Hence the central limit theorem (2.29) follows using the same argument as in the case of an iid sequence $\{\eta_s, s\in\mathbf{Z}\}$ (see e.g. Davydov (1970), Giraitis and Surgailis (1991)). ■

Sections 4 and 5 provide the proofs of the outstanding results (2.16), (2.23) and (2.24) needed for the proof of Theorem 2.2. First, however, the following section establishes finiteness of the moments of powers σ_0.

3. Moments and diagrams.

In this section we discuss diagram formalism for the moments $E[\sigma_t^l]$, $l = 2, 3, \dots,$ of the Volterra series (2.1).

Let $\sum_S^{k,t}$ denote the sum over all subsets $S = \{s_k, s_{k-1}, \dots, s_1\} \subset \mathbf{Z}$, $s_k < s_{k-1} < \dots < s_1 < s_0 = t, k = 0, 1, \dots$. With any such S we associate the function

$$b^S := \prod_{i=1}^k b_{s_{i-1}-s_i} = b_{t-s_1} b_{s_1-s_2} \dots b_{s_{k-1}-s_k}$$

and the random variable

$$\varepsilon^S := \prod_{s \in S} \varepsilon_s = \varepsilon_{s_1} \dots \varepsilon_{s_k},$$

$b^\emptyset = \varepsilon^\emptyset := 1$. Then

$$\sigma_t = a \sum_{k=0}^{\infty} \sum\nolimits_S^{k,t} b^S \varepsilon^S \tag{3.1}$$

and

$$\begin{aligned} \sigma_t^l &= a^l \sum_{k_1,\dots,k_l} \sum\nolimits_{S_1}^{k_1,t} \dots \sum\nolimits_{S_l}^{k_l,t} b^{S_1} \dots b^{S_l} \varepsilon^{S_1} \dots \varepsilon^{S_l} \\ &=: a^l \sum_{(k)_l} \sum\nolimits_{(S)_l}^{(k)_l,t} b^{(S)_l} \varepsilon^{(S)_l}. \end{aligned} \tag{3.2}$$

In (3.2), the sum $\sum_{(k)_l}$ is taken over all collections $(k)_l = (k_1, \dots, k_l) \in \mathbf{Z}_+^l, \mathbf{Z}_+ := \{0, 1, \dots, \}$, $\sum_{(S)_l}^{(k)_l,t} := \sum_{S_1}^{k_1,t} \dots \sum_{S_l}^{k_l,t}$, $(S)_l := (S_1, \dots, S_l)$ and we put $b^{(S)_l} := b^{S_1} \dots b^{S_l}$, $\varepsilon^{(S)_l} := \varepsilon^{S_1} \dots \varepsilon^{S_l}$. Then

$$E[\sigma_t^l] = a^l \sum_{(k)_l} \sum\nolimits_{(S)_l}^{(k)_l,t} b^{(S)_l} \mu_{(S)_l}, \tag{3.3}$$

where

$$\mu_{(S)_l} := E[\varepsilon^{(S)_l}] = E[\varepsilon^{S_1} \dots \varepsilon^{S_l}].$$

In a similar way, for any integers $l', l'' \ge 1$

$$\mathrm{Cov}(\sigma_t^{l'}, \sigma_0^{l''}) = a^{l'+l''} \sum_{(k')_{l'}} \sum_{(k'')_{l''}} \sum\nolimits_{(S')_{l'}}^{(k')_{l'},t} \sum\nolimits_{(S'')_{l''}}^{(k'')_{l''},0} b^{(S')_{l'}} b^{(S'')_{l''}} \bar{\mu}_{(S')_{l'},(S'')_{l''}}, \tag{3.4}$$

where

$$\bar{\mu}_{(S')_{l'},(S'')_{l''}} := \mathrm{Cov}(\varepsilon^{(S')_{l'}}, \varepsilon^{(S'')_{l''}}),$$

for any collections $(S')_{l'} = (S'_1, \dots, S'_{l'}), (S'')_{l''} = (S''_1, \dots, S''_{l''})$ of subsets of $\mathbf{Z}$. To study the convergence and the asymptotics as $t \to \infty$ of the formal series (3.3-4), we introduce below a diagram formalism. Observe, by the independence of $\varepsilon_i, i \in \mathbf{Z}$,

$$\mu_{(S)_l} = 0 \qquad \text{unless} \quad \triangle(S)_l := \cup_{i=1}^l (S_i \setminus \cup_{j \ne i} S_j) = \emptyset \tag{3.5}$$

and

$$\bar{\mu}_{(S')_{l'},(S'')_{l''}} = 0 \qquad \text{unless} \quad \triangle(S')_{l'} \subset \cup_{i=1}^{l''} S''_i \quad \text{and} \quad \triangle(S'')_{l''} \subset \cup_{i=1}^{l'} S'_i. \tag{3.6}$$

Let $(k)_l = (k_1, \ldots, k_l) \in \mathbf{Z}_+^l$ be given. Let $I \equiv I((k)_l)$ be the table consisting of l rows $I_j \equiv I_j((k)_l) = \{(k_j, j), \ldots, (1, j)\}$ of length $k_j \geq 0, j = 1, \ldots, l$. (Some of these rows may be empty, too.) A *diagram* is an ordered partition $\gamma = (V_1, \ldots, V_r)$ of the table I by nonempty subsets (*edges*) $V_q, q = 1, \ldots, r$, $r = 1, 2, \ldots$, containing at most one element of any row: $|V_q \cap I_j| \leq 1$, $q = 1 \ldots, r, j = 1, \ldots, l$.

Let $f(s_{i,j} : (i,j) \in I)$ be a function defined on (collections of) ordered integers:

$$s_{k_j,j} < s_{k_{j-1},j} < \ldots < s_{1,j} < s_{0,j}, \quad j = 1, \ldots, l, \tag{3.7}$$

where $(s_{0,1}, \ldots, s_{0,l}) := (s_0)_l \in \mathbf{Z}^l$ is fixed. With any such $f(s_{i,j} : (i,j) \in I)$ and any diagram $\gamma = (V_1, \ldots, V_r)$ we associate the sum

$$\sum\nolimits_\gamma^{(s_0)_l} f(s_{i,j} : (i,j) \in I) \ = \sum_{s_{i,j} = \tilde{s}_q, (i,j) \in V_q, q = 1, \ldots, r} f(s_{i,j} : (i,j) \in I) \tag{3.8}$$

over all integers $s_{i,j}, i = 1, \ldots, k_j, j = 1, \ldots, l$ satisfying the inequalities (3.7) and

$$s_{i,j} = s_{i',j'} := \tilde{s}_q, \quad (i,j), (i',j') \in V_q, \ q = 1, \ldots, r \tag{3.9}$$

and

$$\tilde{s}_1 < \ldots < \tilde{s}_r. \tag{3.10}$$

In general, the inequalities (3.7) and (3.10) may be incompatible in which case the sum (3.8) is zero by definition. It is convenient to picture edges of a diagram as 'vertical sets' connected by curve segments right, as well as to connect the vertices lying on the same row, thus making γ a graph. The 'vertical edges' $V_1, \ldots, V_r$ should be placed horizontally in increasing order. For example, the graph in Fig. 1 corresponds to $(k)_4 = (1,2,2,3)$ and $\gamma = (V_1, \ldots, V_5), V_1 = \{(3,4)\}, V_2 = \{(2,4)\}, V_3 = \{(2,2),(2,3),(1,4)\}, V_4 = \{(1,3)\}, V_5 = \{(1,1),(1,2)\}$. According to (3.8), a diagram determines the choice of 'coinciding diagonals' in the summation over integers (3.7). For example, for γ shown in Fig. 1 and $(s_0)_4 = (0,0,t,0)$, $\sum_\gamma^{(s_0)_4}$ denotes the sum over integers $s_{1,1} < 0, s_{2,2} < s_{1,2} < 0, s_{2,3} < s_{1,3} < t, s_{3,4} < s_{2,4} < s_{1,4} < 0$ satisfying $s_{3,4} =: \tilde{s}_1 < s_{2,4} =: \tilde{s}_2 < s_{2,2} = s_{2,3} = s_{1,4} =: \tilde{s}_3 < s_{1,3} =: \tilde{s}_4 < s_{1,1} = s_{1,2} =: \tilde{s}_5$.

Graph of Example

Fig. 1

Write Γ_I for the class of all diagrams $\gamma = (V_1, \ldots, V_r)$ over $I = I((k)_l)$ such that $|V_q| > 1 \, \forall q = 1, \ldots, r$. Then from (3.3), (3.5) one obtains

$$E[\sigma_t^l] = a^l \sum_{(k)_l} \sum_{\gamma \in \Gamma_{I((k)_l)}} \mu_\gamma \sum\nolimits_\gamma^{(t)_l} b^{(S)_l}, \tag{3.11}$$

where $(t)_l := \underbrace{(t, \ldots, t)}_{l}$ and where

$$\mu_\gamma := \mu_{(S)_l}$$

for $(S)_l = (S_1, \ldots, S_l), S_j = \{s_{k_j,j}, \ldots, s_{1,j}\}$, $j = 1, \ldots, l$ satisfying (3.9-10), depends on γ only. Similarly, for any $l', l'' \geq 1$,

$$\mathrm{Cov}(\sigma_t^{l'}, \sigma_{\mathbf{0}}^{l''}) = a^{l'+l''} \sum_{(k')_{l'}} \sum_{(k'')_{l''}} \sum_{\gamma \in \Gamma_{I((k',k'')_{l',l''})}} \bar{\mu}_\gamma \sum\nolimits_\gamma^{(t,0)_{l',l''}} b^{(S')_{l'}} b^{(S'')_{l''}}. \tag{3.12}$$

In (3.12), $I((k', k'')_{l',l''}) := I = I' \cup I''$ is the table having $l' + l''$ rows and consisting of two *blocks* $I' := I((k')_{l'}), I'' := I((k'')_{l''})$, $(t, 0)_{l',l''} := (\underbrace{t, \ldots, t}_{l'}, \underbrace{0, \ldots, 0}_{l''})$, and

$$\bar{\mu}_\gamma := \bar{\mu}_{(S')_{l'},(S'')_{l''}} = \mathrm{Cov}(\varepsilon^{(S')_{l'}}, \varepsilon^{(S'')_{l''}}) \tag{3.13}$$

depends on γ only. Property (3.6) of the last covariance translates to the diagram language as follows. Call a diagram $\gamma = (V_1, \ldots, V_r) \in \Gamma_{I((k',k'')_{l',l''})}$ *block-connected* if there is an edge V_q which has an nonempty intersection with both blocks I', I'' of the table I: $V_q \cap I' \neq \emptyset$, $V_q \cap I'' \neq \emptyset$. By (3.6), the last sum on the right hand side of (3.12) vanishes for each diagram which is not block-connected so that (3.12) involves summation over block-connected diagrams only.

Lemma 3.1. *Let Assumption 1 hold and*

$$(2^l - l - 1)^{1/2} |\mu|_l^{1/l} b < 1 \qquad (3.14)$$

where $|\mu|_j = E(|\varepsilon|_0^j)$. *Then the series (3.3) converges absolutely and defines a finite moment* $E[\sigma_t^l]$.

Proof. By Hölder's inequality,

$$|\mu_{(S)_l}| = |E[\varepsilon^{S_1} \ldots \varepsilon^{S_l}]| \le |\mu|_l^{(|S_1|+\ldots+|S_l|)/l} \qquad (3.15)$$

where $|\mu|_j = E(|\varepsilon|_0^j)$. Then the lemma follows from

$$\sum\nolimits_{\gamma}^{(t)_l} |b^{(S)_l}| \le b^{|S_1|+\ldots+|S_l|} \qquad (3.16)$$

and Lemma 3.2 below. To show (3.16), consider a diagram $\gamma = (V_1, \ldots, V_r) \in \Gamma_I$, $I = I((k)_l)$. Then by the Cauchy - Schwarz inequality, V_1 being the leftmost edge of γ,

$$\sum\nolimits_{\gamma}^{(t)_l} |b^{(S)_l}| \le b^{|V_1|} \sum\nolimits_{\gamma'}^{(t)_l} |b^{(S')_l}|, \qquad (3.17)$$

where $\gamma' := (V_2, \ldots, V_r), S'_j := S_j \backslash V_1, |S'_j| = k'_j, j = 1, \ldots, l$ and $\gamma' \in \Gamma_{I'}, I' := I((k')_l)$, $(k')_l = (k'_1, \ldots, k'_l)$. Indeed, let $V_1, |V_1| = m$ connect the first m rows $S_i, 1 \le i \le m, 2 \le m \le l$. Then the summation over $s_{k_1,1} = \ldots = s_{k_m,m} =: \tilde{s}$ in the sum $\sum_{\gamma}^{(t)_l}$ contributes to

$$\sum_{\tilde{s}} \prod_{i=1}^{m} |b_{s_{k_i-1,i}-\tilde{s}}| \le \Big(\sum_{\tilde{s}} b^2_{s_{k_1-1,1}-\tilde{s}}\Big)^{1/2} \Big(\sum_{\tilde{s}} \prod_{i=2}^{m} b^2_{s_{k_i-1,i}-\tilde{s}}\Big)^{1/2} \le b^m. \qquad (3.18)$$

Thus, (3.16) follows by repeated use of (3.17). ■

Lemma 3.2.

$$|\Gamma_{I((k)_l)}| \le (2^l - l - 1)^{(k_1+\ldots+k_l)/2}.$$

Proof. According to (3.10), edges of a diagram $\gamma = (V_1, \ldots, V_r)$ are ordered, and any edge $V_q, 2 \le |V_q| \le l$ may be chosen in $\sum_{i=2}^{l} \binom{l}{i} = 2^l - l - 1$ ways. The number r of edges does not exceed $(k_1 + \ldots + k_l)/2$. This proves the lemma. ■

Remark 3.1. If $l = 2$ and $\{\varepsilon_t\}$ is a Gaussian sequence, then condition (3.14) of Lemma 3.1 can be replaced by

$$\mu_2 b^2 < 1$$

and Assumption 4 (2) by

$$7\mu_4^{1/2} b^2 < 1.$$

This easily follows, noting that $E\varepsilon_0 = E\varepsilon_0^3 = 0$ implies that

$$|\gamma \in \Gamma_{I((k)_4)} : \mu_\gamma \neq 0| \leq (\binom{4}{2} + \binom{4}{4})^{(k_1+\ldots+k_4)/2} = 7^{(k_1+\ldots+k_4)/2}.$$

Write

$$x_{i,j} := s_{i-1,j} - s_{i,j}, \quad i = 1, \ldots, k_j, \quad j = 1, \ldots, l$$

for the differences of the arguments (3.7). Below, we need

Lemma 3.3. *Assume that*

$$\sup_{t \geq 1} t^{(1+\theta)/2} |b_t| < \infty, \tag{3.19}$$

where $0 < \theta < 1$. *Let* $\gamma = (V_1, \ldots, V_r) \in \Gamma_I$ *be a diagram, and* (i_1, j_1), (i_2, j_2) *be arbitrary elements of the table* $I = I((k)_l)$, *which do not belong to the same row or the same edge. Then for any* $L_1, L_2 > 0$

$$\sum\nolimits_{\gamma}^{(s_0)_l} |b^{(S)_l}| \mathbf{1}(|x_{i_1,j_1}| > L_1, |x_{i_2,j_2}| > L_2) \;\leq\; C b^{|(k)_l|} L_1^{-\theta} L_2^{-\theta}. \tag{3.20}$$

In (3.20), the constant C *does not depend on* $(k)_l, (s_0)_l, \gamma$, *and* $|(k)_l| := |I| = k_1 + \ldots + k_l$.

Proof. This follows that of (3.16), where we use the Cauchy-Schwarz inequality (3.17) for any edge which contains (i_1, j_1) or (i_2, j_2). Consider an edge $\{(i_1, j_1), (i, j)\}$, $(i, j) \neq (i_2, j_2), j \neq j_1$. Assume first $s := s_{i-1,j} \geq s_{i_1-1,j_1} =: s_1$. Then the summation over $s_{i_1,j_1} = s_{i,j} =: \tilde{s}$ contributes to

$$\begin{aligned} &\sum_{\tilde{s}} |b_{s_1-\tilde{s}} b_{s-\tilde{s}}| \mathbf{1}(|s_1 - \tilde{s}| > L_1) \\ \leq\;& C \sum_{u > L_1} u^{-(1+\theta)/2} (u - (s_1 - s))^{-(1+\theta)/2} \leq C \sum_{v > 0} (v + L_1)^{-(1+\theta)/2} v^{-(1+\theta)/2} \leq C L_1^{-\theta}. \end{aligned}$$

For $s \leq s_1$, a similar bound follows easily. By evaluating in a similar way the sum over s_{i_2,j_2}, one obtains (3.20). ■

4. The intermediate term $\zeta_{t,l}$.

From (2.19), it follows that $\zeta_{t,l}, t \in \mathbf{Z}$ is strictly stationary, with zero mean and (cross)autocovariance

$$\mathrm{Cov}(\zeta_{t,l'}, \zeta_{0,l''}) = \frac{l' l''}{1 - b^2} \sum_{j=1}^{\infty} a_{j,l'} a_{t+j,l''}, \tag{4.1}$$

determined by the last convolution. It turns out, that the weights $a_{t,l}$ have similar asymptotic behaviour to b_t under Assumption 3.

Lemma 4.1. *Assume conditions (3.14) and (3.19). Then*

$$|G_{t,l}| \le C t^{-1-\theta}. \tag{4.2}$$

Furthermore, under Assumption 3,

$$a_{t,l} \;=\; E[\sigma_0^l] b_t + o(b_t). \tag{4.3}$$

Proof. Let us first prove (4.3). By (2.21), (4.2),

$$\sum_{s=1}^{\infty} G_{s,l} = H_{\infty,l} - H_{0,l} = E[\sigma_0^l] - aE[\sigma_0^{l-1}].$$

Hence, (4.3) follows from (2.20),

$$\sum_{s>t} |G_{s,l}| = o(1) \tag{4.4}$$

and

$$\sum_{0<s<t} |G_{s,l}||b_t - b_{t-s}| = o(b_t). \tag{4.5}$$

Here, (4.4) is obvious from (4.2). To show (4.5), write

$$\begin{aligned}\sum_{0<s<t} |G_{s,l}||b_t - b_{t-s}| \;&\le\; \sum_{0<s<t/2} |G_{s,l}||b_t - b_{t-s}| + \sum_{t/2\le s<t} |G_{s,l}||b_{t-s}| + |b_t| \sum_{t/2\le s<t} |G_{s,l}| \\ &=: \quad J_1 + J_2 + J_3.\end{aligned}$$

From (4.2) and (3.19), the estimates $J_i = O(t^{-(1+3\theta)/2}) = o(b_t), i = 2, 3$ easily follow. Next, $J_1 = |b_t| \sum_{0<s<t/2} |G_{s,l}| h_t(s)$, where $h_t(s) := |1-(b_{t-s}/b_t)|$ vanishes as $t \to \infty$ for each fixed $s \ge 1$, and $h_t(s)$ is uniformly bounded for $0 < s < t/2$, implying $\sum_{0<s<t/2} |G_{s,l}| h_t(s) = o(1)$ by (4.2). This proves (4.5) and (4.3).

It remains to prove (4.2). Observe, that

$$E[\sigma_t|\mathcal{F}_s^+] = \sum_{k=0}^{t-s} \sum_{s\le s_k<\ldots<s_1<t} b_{t-s_1} b_{s_1-s_2} \ldots b_{s_{k-1}-s_k} \varepsilon_{s_1} \ldots \varepsilon_{s_k}$$

has similar structure to σ_t. Therefore the expectation $H_{t-s,l}$ can be written similarly to (3.11):

$$H_{t-s,l} = \sum_{(k)_l} \sum_{\gamma\in\Gamma_{I((k)_l)}} \mu_\gamma \sum\nolimits_{\gamma}^{(t)_l} b^{(S)_l} \mathbf{1}(S_1 \subset [s,t)).$$

Therefore

$$G_{t-s,l} = \sum_{(k)_l} \sum_{\gamma \in \Gamma_{I((k)_l)}} \mu_\gamma \sum\nolimits_\gamma^{(t)_l} b^{(S)_l} \mathbf{1}(\wedge(S_1) = s), \tag{4.6}$$

where $\wedge(S) = \min\{s : s \in S\}$. With (3.14), (3.15) in mind, the bound (4.2) and the lemma follow from (4.6) and Lemma 4.2 below. Lemma 4.1 is proved. ■

Lemma 4.2. *Let $b_i, i \geq 1$ satisfy the condition*

$$\sup_{i \geq 1} i^{(1+\theta)/2} |b_i| \leq D, \tag{4.7}$$

where $D \geq 1$. Then for any $(k)_l \in \mathbf{Z}_+^l$ and any diagram $\gamma \in \Gamma_{I((k)_l)}$,

$$\sum\nolimits_\gamma^{(t)_l} |b^{(S)_l}| \, \mathbf{1}(\wedge(S_1) = s) \; \leq \; D^3 |k|^3 b^{|k|} |t-s|^{-1-\theta}, \tag{4.8}$$

where $|k| = k_1 + \ldots + k_l$.

Proof. Write $N_{\gamma,t-s}$ for the left hand side of (4.8). By homogeneity of both sides of (4.8) with respect to b, it suffices to show the lemma for $b = 1$, in which case according to (3.16)

$$\sum_{s<t} N_{\gamma,t-s} \leq 1. \tag{4.9}$$

We prove (4.8) by induction in the number r of edges of $\gamma = (V_1, \ldots, V_r)$. For $r = 1$, it follows easily; indeed, in this case, $N_{\gamma,t-s} = |b_{t-s}|^{|V_1|} \leq b_{t-s}^2 \leq D^2 |t-s|^{-1-\theta}$.

To show the induction step $r-1 \to r$, let $V_{q^*}, 1 \leq q^* \leq r$ be the edge which contains the element $(k_1, 1)$ (= the far left element of the first row of the table $I = I((k)_l)$. There are two possibilities: (1) $q^* > 1$ and (2) $q^* = 1$. In the case (1), use the Cauchy - Schwarz inequality as in (3.17), to obtain

$$N_{\gamma,t-s} = \sum\nolimits_\gamma^{(t)_l} |b^{(S)_l}| \, \mathbf{1}(\wedge(S_1) = s) \; \leq \; \sum\nolimits_{\gamma'}^{(t)_l} |b^{(S')_l}| \, \mathbf{1}(\wedge(S'_1) = s) = N_{t-s,\gamma'}.$$

The diagram γ' has $r' = r - 1 < r$ edges and therefore satisfies the inductive assumption, thereby proving the induction step.

Let now $V_{q^*} = V_1$ be the far left edge of γ. Without loss of generality, assume $|V_1| = m \geq 2$ connects the first m rows $I_i, 1 \leq i \leq m$ of the table $I = I((k)_l)$. Then using the notation of (3.17-18), one can rewrite the product $|b_l^{(S)}|$ in (4.8) as

$$|b^{(S)_l}| = |b^{(S')_l}| \prod_{i=1}^m |b_{s_{k_i-1,i}-s}| = |b^{(S')_l}| \prod_{i=1}^m |b_{s_{k'_i,i}-s}|. \tag{4.10}$$

Using the inequalities $|b_{s_{k'_1,1}-s} b_{s_{k'_2,2}-s}| \le (1/2)(b^2_{s_{k'_1,1}-s} + b^2_{s_{k'_2,2}-s})$, $|b_i| \le 1\, \forall i$, from (4.10) one obtains

$$\begin{aligned} 2N_{\gamma,t-s} &\le \sum_{s<\tilde{s}<t} b^2_{\tilde{s}-s} \sum\nolimits_{\gamma'}^{(t)_l} |b^{(S')_l}| \left(\mathbf{1}(\wedge(S'_1) = \tilde{s}) + \mathbf{1}(\wedge(S'_2) = \tilde{s})\right) \\ &\le 2 \sum_{s<\tilde{s}<t} b^2_{\tilde{s}-s} N_{\gamma',t-\tilde{s}}. \end{aligned}$$

Put $|k'| := k'_1 + \ldots + k'_l$, then $|k'| \le |k| + 2$ and

$$N_{\gamma,t} \le \sum_{t/|k'|<u<t} b^2_u N_{\gamma',t-u} + \sum_{t-t/|k'|\le t-u<t} b^2_u N_{\gamma',t-u}. \tag{4.11}$$

Here, $b^2_u \mathbf{1}(u > t/|k'|) \le D^2 |u|^{-1-\theta} \mathbf{1}(u > t/|k'|) \le D^2 |k'|^{1+\theta} t^{-1-\theta}$. Similarly, by the inductive assumption,

$$\begin{aligned} N_{\gamma',t-u}\mathbf{1}(t-u \ge t - t/|k'|) &\le D^3|k'|^3|t-u|^{-1-\theta}\mathbf{1}(t-u \ge t-t/|k'|) \\ &\le D^3|k'|^3(|k'|/(|k'|-1))^{1+\theta} t^{-1-\theta}. \end{aligned}$$

Substituting these inequalities into the right hand side of (4.11) and using (4.9) and $b = 1$, we obtain

$$t^{1+\theta} N_{\gamma,t} \le D^2|k'|^{1+\theta} + D^3|k'|^3(|k'|/(|k'|-1))^{1+\theta} \le D^3(|k'|^2 + |k'|^3(|k'|/(|k'|-1)^2).$$

In view of the inequality $n^2 + n^3(n/(n-1))^2 \le (n+2)^3$, which is true for any integer $n \ge 2$, this proves the induction step $r-1 \to r$ and Lemma 4.2 also. ∎

From (4.1) and Lemma 4.1 we deduce:

Corollary 4.3. *Under Assumptions 3 and 4(l), for any $1 \le l', l'' \le l$*

$$\mathrm{Cov}(\zeta_{t,l'}, \zeta_{0,l''}) = l'l'' E[\sigma_0^{l'}] E[\sigma_0^{l''}] \mathrm{Cov}(\sigma_t, \sigma_0)(1 + o(1)). \tag{4.12}$$

Lemma 4.4. *Under the assumptions of Theorem 2.2, relation (2.16) holds.*

Proof. Note from (2.1) that for $t \ge 1$ σ_t can be written as

$$\begin{aligned} \sigma_t &= a \sum_{k=0}^{\infty} \sum_{s_k<\ldots<s_1<t} b_{t-s_1} \ldots b_{s_{k-1}-s_k} \varepsilon_{s_1} \ldots \varepsilon_{s_k} \\ &= a^{-1} \sum_{s=0}^{t-1} (E[\sigma_t|\mathcal{F}_s^+] - E[\sigma_t|\mathcal{F}_{s+1}^+]) E[\sigma_s|\mathcal{F}_{-1}] + E[\sigma_t|\mathcal{F}_{-1}], \end{aligned}$$

where

$$E[\sigma_t|\mathcal{F}_s^+] - E[\sigma_t|\mathcal{F}_{s+1}^+] = a\sum_{k=1}^{t} \sum_{s=s_k<\ldots<s_1<t} b_{t-s_1}\ldots b_{s_{k-1}-s_k}\varepsilon_{s_1}\ldots\varepsilon_{s_k} \tag{4.13}$$

and

$$E[\sigma_s|\mathcal{F}_{-1}] = a\sum_{k=0}^{\infty} \sum_{s_k<\ldots<s_1<0} b_{s-s_1}\ldots b_{s_{k-1}-s_k}\varepsilon_{s_1}\ldots\varepsilon_{s_k}.$$

Set $g_s^+ := a^{-1}(E[\sigma_s|\mathcal{F}_s^+] - E[\sigma_s|\mathcal{F}_{s+1}^+]), 0 \le s \le t-1, g_t^+ := 1, g_s^- := E[\sigma_s|\mathcal{F}_{-1}], 0 \le s \le t$. Then

$$\sigma_t = \sum_{s=0}^{t} g_s^+ g_s^-.$$

Hence

$$\begin{aligned}\mathrm{Cov}(\nu_{0l},\sigma_t^l) &= E[(\varepsilon_0^l - E\varepsilon_0^l)\sigma_0^l(\sigma_t^l - E\varepsilon_t^l)] \\ &= \sum_{s_1,\ldots,s_l=0:\min s_i=0}^{t} E[(\varepsilon_0^l - E\varepsilon_0^l)\prod_{i=1}^{l} g_{s_i}^+]E[\sigma_0^l\prod_{i=1}^{l} g_{s_i}^-],\end{aligned} \tag{4.14}$$

where we use the fact that $\sigma_0^l, g_s^-, 0 \le s \le t$ are $\mathcal{F}_{-1}-$measurable, and $g_s^+, 1 \le s \le t$ are $\mathcal{F}_1^+-$measurable. Note that $E(g_s^-)^{2l} \le E\sigma_s^{2l} = E\sigma_0^{2l} < \infty$ by Assumption $4(l)$ and Lemma 3.1, and therefore

$$\left|E[\sigma_0^l\prod_{i=1}^{l} g_{s_i}^-]\right| \le (E\sigma_0^{2l})^{1/2}\prod_{i=1}^{l}(E(g_{s_i}^-)^{2l})^{1/2l} \le E\sigma_0^{2l} \tag{4.15}$$

is bounded uniformly in $s_i, 1 \le i \le l$. On the other hand, taking into account the definition of g_s^+ and (4.13), the first expectation on the right hand side of (4.14) can be written similarly to (3.12) and (3.13) with the help of diagrams, yielding

$$\begin{aligned}\sum_{s_1,\ldots,s_l=0:\min s_i=0}^{t}\left|E[(\varepsilon_0^l - E\varepsilon_0^l)\prod_{i=1}^{l} g_{s_i}^+]\right| &\le C\sum_{(k)_l}\mu_{2l}^{|k|/l}\sum_{\gamma\in\Gamma_{I((k)_l)}}\sum_{\gamma}^{(t)_l}|b^{(S)_l}|\mathbf{1}(\wedge(S_1)=0) \\ &\le Ct^{-1-\theta}\sum_{(k)_l}\mu_{2l}^{|k|/l}(2^l - l - 1)^{|k|/2},\end{aligned}$$

according to Lemmas 3.1 and 4.2. The last sum being finite under Assumption $4(l)$, this completes by (4.14), (4.15) the proof of Lemma 4.4. ■

5. The remainder term y_{tl}.

In this section we study the asymptotic behaviour of the autocovariances of the difference $y_{tl} = \sigma_t^l - lE[\sigma_0^l]\sigma_t$, to which end, we first study the (cross)autocovariances $\mathrm{Cov}(\sigma_t^{l'}, \sigma_0^{l''})$ for $l', l'' = 1,\ldots,l$.

Lemma 5.1. *Under Assumptions 3 and* 4(l), *for any* $1 \le l', l'' \le l$

$$\mathrm{Cov}(\sigma_t^{l'}, \sigma_0^{l''}) \;=\; \mathrm{Cov}(\zeta_{t,l'}, \zeta_{0,l''}) + O(t^{-\theta-\lambda}), \tag{5.1}$$

where $\lambda = \theta(1-\theta)/(1+\theta) > 0$.

Proof. To prove the lemma, we write the covariances in terms of diagrams and perform cancellation in the corresponding expressions, leaving out terms which are of order $O(t^{-\theta-\lambda})$.

We start by recalling the diagram formula (3.12) for the covariance $\mathrm{Cov}(\sigma_t^{l'}, \sigma_0^{l''})$, where the summation is taken over block-connected diagrams. We compare this formula with $\mathrm{Cov}(\zeta_{t,l'}, \zeta_{0,l''})$ which we rewrite in a similar way, using a special type of diagram which we call *regular*. Roughly speaking, a regular diagram connects the two $l-$blocks I', I'' of the table $I = I((k')_l, (k'')_l)$ only by edges having two elements and all belonging to the same pair of rows.

To give a formal definition, let $\mathcal{V}_{i',i''}\,(1 \le i' \le l', 1 \le i'' \le l'')$ denote the class of edges $V \subset I = I' \cup I''$ such that $|V| = 2$ and $V \cap I'_{i'} \ne \emptyset, V \cap I''_{i''} \ne \emptyset$, where $I'_{i'}, I''_{i''}, 1 \le i' \le l', 1 \le i'' \le l''$ denote rows of $I' = I((k')_{l'}), I'' = I((k'')_{l''})$, respectively.

Definition 5.2. *A diagram* $\gamma = (V_1, \ldots, V_r) \in \Gamma_I$ *will be said regular if it is block-connected and there exist* $1 \le i' \le l', 1 \le i'' \le l''$ *such that, for any* $q = 1, \ldots, r$, *either* $V_q \in \mathcal{V}_{i',i''}$, *or* $V_q \subset I'$, *or* $V_q \subset I''$ *hold, and, moreover, if* $V_q \in \mathcal{V}_{i',i''}$ *for some* $1 \le q \le r$ *then* $V_{q'} \in \mathcal{V}_{i',i''}$ *for any* $1 \le q' < q$ *such that* $V_{q'} \cap (I'_{i'} \cup I''_{i''}) \ne \emptyset$.

The last property says that the edges $V_q \in \mathcal{V}_{i',i''}$ connecting the blocks I', I'', connect pairwise consecutive elements of the corresponding rows $I'_{i'}, I''_{i''}$, starting from the left. A block-connected diagram $\gamma \in \Gamma_I$ which is not regular will be called *irregular*. Write $\Gamma_I^{\mathrm{reg}}, \Gamma_I^{\mathrm{irreg}}$ for the corresponding classes of diagrams. By definition,

$$\Gamma_I^{\mathrm{reg}} = \cup_{i'=1}^{l'} \cup_{i''=1}^{l''} \Gamma_I^{\mathrm{reg}}(i', i'') \tag{5.4}$$

is the union of disjoint classes $\Gamma_I^{\mathrm{reg}}(i', i'')$ corresponding to given i', i'' in Definition 5.2. In general, given a table I, the class Γ_I^{reg} may be empty as well.

Let us introduce one more class of diagrams. Namely, the class $\tilde{\Gamma}_I^{\mathrm{irreg}} \subset \Gamma_I^{\mathrm{irreg}}$ consists of irregular diagrams $\gamma = (V_1, \ldots, V_r)$ which are obtained from a regular diagram $\tilde{\gamma} = (\tilde{V}_1, \ldots, \tilde{V}_{\tilde{r}}) \in \Gamma_I^{\mathrm{reg}}(i', i'')\,(1 \le i' \le l', 1 \le i'' \le l''), \tilde{r} > r$ as follows: any edge of γ either coincides with some edge of $\tilde{\gamma}$, or is a union of an edge of $\tilde{\gamma}$ which intersects both blocks I', I'', and one or two other edges of $\tilde{\gamma}$ lying entirely in one or two different blocks, respectively.

It is not hard to verify, using (2.19-2.22) and (4.6), that the covariance (4.1) can be written as

$$\mathrm{Cov}(\zeta_{t,l'}, \zeta_{0,l''}) \;=\; \sum_{(k')_{l'}} \sum_{(k'')_{l''}} \sum_{\gamma \in \Gamma_I^{\mathrm{reg}} \cup \tilde{\Gamma}_I^{\mathrm{irreg}}} \tilde{\mu}_\gamma \; {\sum_\gamma}^{(t,0)_{l',l''}} b^{(S')_{l'}} b^{(S'')_{l''}}, \tag{5.5}$$

where $\tilde{\mu}_\gamma := \bar{\mu}_\gamma$ for $\gamma \in \Gamma_I^{\rm reg}$, and, in the case when $\gamma \in \tilde{\Gamma}_I^{\rm irreg}$ is obtained from a diagram $\tilde{\gamma}$ as described above, $\tilde{\mu}_\gamma := \bar{\mu}_{\tilde{\gamma}}$. It follows from (3.15) that $\tilde{\mu}_\gamma$ and $\bar{\mu}_\gamma$ satisfy a similar inequality:

$$\max(|\bar{\mu}_\gamma|, |\tilde{\mu}_\gamma|) \le 2|\mu|_l^{|I|/l} = 2|\mu|_l^{(|(k')_{l'}|+|(k'')_{l''}|)/l}, \tag{5.6}$$

$|(k')_{l'}| = k'_1 + \ldots + k'_{l'}, |(k'')_{l''}| = k''_1 + \ldots + k''_{l''}$ being the number of elements of the blocks I', I'', respectively, and $l = l' + l''$. Then, by comparing (3.12) and (5.5), it is easily seen that the relation (5.2) follows from

$$\sum_{(k')_{l'}} \sum_{(k'')_{l''}} \sum_{\gamma \in \Gamma_I^{\rm irreg}} |\mu|_l^{|I|/l} \sum\nolimits_\gamma^{(t,0)_{l',l''}} |b^{(S')_{l'}} b^{(S'')_{l''}}| \; = \; O(t^{-\theta-\lambda}). \tag{5.7}$$

With Lemma 3.2 in mind, (5.7) follows from

Lemma 5.2. *For any $A > b$, there is a constant $C < \infty$ such that, for any $(k')_{l'} \in \mathbf{Z}_+^{l'}, (k'')_{l''} \in \mathbf{Z}_+^{l''}$ and any $\gamma \in \Gamma_I^{\rm irreg}, I = I((k', k'')_{l',l''})$,*

$$\sum\nolimits_\gamma^{(t,0)_{l',l''}} |b^{(S')_{l'}} b^{(S'')_{l''}}| \; \le \; CA^{|I|} t^{-\theta-\lambda}. \tag{5.8}$$

Proof. Let $\gamma \in (V_1, \ldots, V_r) \in \Gamma_I^{\rm irreg}, I = I' \cup I'', I' = I((k')_{l'}), I'' = I((k'')_{l''})$ be given. Put $q_* = \max\{q = 1, \ldots, r : V_q \cap I' \ne \emptyset, V_q \cap I'' \ne \emptyset\}$. In other words, V_{q_*} is the first edge from the right which connects I', I''. There are two possibilities:
(c.1) $|V_{q_*}| = 2$;
(c.2) $|V_{q_*}| \ge 3$.

Consider the case (c.1). Let

$$V_{q_*} = \{(i'_*, j'_*), (i''_*, j''_*)\}, \quad s_* := s_{i'_*, j'_*} = s_{i''_*, j''_*},$$

$(i'_*, j'_*) \in I', (i''_*, j''_*) \in I''$. Choose $L := t^{(1-\theta)/(1+\theta)} = o(t)$. Then

$$\begin{aligned} w_t \;\; : & = \sum\nolimits_\gamma^{(t,0)_{l',l''}} \left| b^{(S')_{l'}} b^{(S'')_{l''}} \right| = \sum\nolimits_\gamma^{(t,0)_{l',l''}} \left| b^{(S')_{l'}} b^{(S'')_{l''}} \right| \mathbf{1}\,(s_* < -L) \\ & + \sum\nolimits_\gamma^{(t,0)_{l',l''}} \left| b^{(S')_{l'}} b^{(S'')_{l''}} \right| \mathbf{1}\,(s_* \ge -L) \\ = & \; : w_{t,L}^- + w_{t,L}^+. \end{aligned} \tag{5.9}$$

Consider $w_{t,L}^+$. Let

$$I_* \; := \; \cup_{q=q_*}^r V_q \; = \; \{V_{q_*}\} \cup I'_* \cup I''_*,$$

where, by the definition of V_{q_*},

$$I'_* := \cup_{q=q_*+1,\ldots,r: V_q \subset I'} V_q, \quad I''_* := \cup_{q=q_*+1,\ldots,r: V_q \subset I''} V_q.$$

Then $I'_* = I((k'_*)_{l'}), I''_* = I((k''_*)_{l''})$, where $(k'_*)_{l'} = (k'_{1,*}, \dots, k'_{l',*}), (k''_*)_{l''} = (k''_{1,*}, \dots, k''_{l'',*})$ are the vectors of lengths of rows of the tables $I'_* \subset I', I''_* \subset I''$, respectively. Then, by applying the Cauchy - Schwarz inequality as in (3.17), one obtains

$$
\begin{aligned}
w^+_{t,L} &\le B^{|I \setminus I_*|} \sum_{s_* > -L} \sum\nolimits^{(t)_{l'}}_{\gamma'_*} \sum\nolimits^{(0)_{l''}}_{\gamma''_*} |b_{s_{i'_*-1,j'_*}-s_*}|\,|b_{s_{i''_*-1,j''_*}-s_*}|\,|b^{(S'_*)_{l'}}|\,|b^{(S''_*)_{l''}}| \\
&= B^{|I \setminus I_*|} \sum_{s_* > -L} \sum_{s_* < s'_* < t} \sum_{s_* < s''_* < 0} |b_{s'_*-s_*}||b_{s''_*-s_*}| \qquad (5.10) \\
&\qquad \times \sum\nolimits^{(t)_{l'}}_{\gamma'_*} |b^{(S'_*)_{l'}}| \mathbf{1}(\wedge(S'_{j',*}) = s'_*) \sum\nolimits^{(0)_{l''}}_{\gamma''_*} |b^{(S''_*)_{l''}}| \mathbf{1}(\wedge(S''_{j'',*}) = s''_*),
\end{aligned}
$$

where $(S'_*)_{l'} := (S'_{1,*}, \dots, S'_{l',*}), (S''_*)_{l''} := (S''_{1,*}, \dots, S''_{l'',*}),\ S'_{i,*} := \{s_{i,j} : (i,j) \in I'_*\},\ S''_{i,*} := \{s_{i,j} : (i,j) \in I''_*\}$ are the corresponding sub-collections of integers (3.9) determined by the diagrams $\gamma'_* = (V_q : V_q \subset I'_*), \gamma''_* = (V_q : V_q \subset I''_*)$. Applying Lemma 4.2 to (5.10), one obtains

$$
\begin{aligned}
w^+_{t,L} &\le CB^{|I|}|I'|^3|I''|^3 \sum_{-L < s_* < 0} \sum_{s_* < s'_* < t} \sum_{s_* < s''_* < 0} |s'_* - s_*|^{-(1+\theta)/2} |s''_* - s_*|^{-(1+\theta)/2} \\
&\qquad \times |t - s'_*|^{-1-\theta} |s''_*|^{-1-\theta}. \qquad (5.11)
\end{aligned}
$$

Hence, by applying the inequality

$$
\sum_{s'} |s' - s|_+^{-(1+\theta)/2} |t - s'|_+^{-1-\theta} \le C|t - s|_+^{-(1+\theta)/2}, \qquad (5.12)
$$

(where $|t|_+^{-\beta} := |t|^{-\beta} \wedge 1, \beta > 0, t \in \mathbf{Z}$) and using $|I'|^3|I''|^3 \le |I|^3$, we get

$$
\begin{aligned}
w^+_{t,L} &\le Cb^{|I|}|I|^3 \sum_{-L < s_* < 0} |s_*|^{-(1+\theta)/2} |t - s_*|^{-(1+\theta)/2} \\
&\le Cb^{|I|}|I|^3 t^{-(1+\theta)/2} \sum_{s=1}^{L} s^{-(1+\theta)/2} \qquad (5.13) \\
&\le Cb^{|I|}|I|^3 L^{(1-\theta)/2} t^{-(1+\theta)/2} = Cb^{|I|}|I|^3 t^{-\theta-\lambda},
\end{aligned}
$$

where $\lambda = \theta(1-\theta)/(1+\theta)$.

Consider $w^-_{t,L}$. As γ is irregular, by definition it contains an edge $V_{q_{**}}$, $q_{**} < q_*$ which either:

(c.1.1) connects the blocks I', I'', but does not belong to the same class $\mathcal{V}_{j'_*, j''_*}$ as V_{q_*} does; or

(c.1.2) $V_{q_{**}}$ belongs to the block I' and and contains an element from the line $I'_{j'_*}$, or $V_{q_{**}}$ belongs to the block I'' and and contains an element from the line $I''_{j''_*}$.

Consider case (c.1.1). Assume for simplicity $|V_{q_{**}}| = 2, V_{q_{**}} = \{(i'_{**}, j'_{**}), (i''_{**}, j''_{**})\}$, where $j'_{**} = j'_*, j''_{**} \ne j''_*$ (the remaining cases can be treated similarly). Put $s_{**} :=$

$s_{i'_{**},j'_{**}} = s_{i''_{**},j''_{**}}$. As the common arguments (3.9) are ordered according to the ordering of edges, we have in the sum $w^-_{t,L}$ the inequalities

$$s_{**} < s_* < -L. \tag{5.14}$$

As $s_{0,j'_*} = t$, among the intervals $x_{i,j'_*} = s_{i-1,j'_*} - s_{i,j'_*}, i = 1, \ldots, i'_*$ in the sum $w^-_{t,L}$, there exists at least one "large" interval $x_{i'_\bullet,j'_*}$ of length

$$x_{i'_\bullet,j'_*} > (t+L)/|I'_{j'_*}| \ \ge\ t/|I'| \qquad (\exists i'_\bullet = 1, \ldots, i'_*). \tag{5.15}$$

In a similar way it follows from (5.14) that among the intervals $x_{i,j''_{**}} = s_{i-1,j''_{**}} - s_{i,j''_{**}}, i = 1, \ldots, i''_{**}$ there is at least one "large" interval $x_{i''_\bullet,j''_{**}}$ of length

$$x_{i''_\bullet,j''_{**}} > L/|I''_{j''_{**}}| \ \ge\ L/|I''| \qquad (\exists i''_\bullet = 1, \ldots, i''_{**}). \tag{5.16}$$

Moreover, the two vertices $(i'_\bullet, j'_*), (i''_\bullet, j''_{**})$ do not belong to the same edge V_q of our diagram. (Indeed, as they belong to the different blocks, so such an edge, if it exists, must be necessarily be either V_{q_*}, or $V_{q_{**}}$, which is clearly impossible. Hence, by (5.15-16) and Lemma 3.3,

$$w^-_{t,L} \le CB^{|I|}(t/|I'|)^{-\theta}(L/|I''|)^{-\theta} \le CB^{|I|}|I|^2(Lt)^{-\theta} = CB^{|I|}|I|^2 t^{-\theta-\lambda}. \tag{5.17}$$

Relations (5.12), (5.17) prove the lemma in case (c.1.1).
Case (c.1.2). Assuming again for simplicity that $|V_{q_{**}}| = 2$,

$$V_{q_{**}} = \{(z''_{**}, u''_{**}), (i''_{**}, j''_{**})\} \subset I'',$$

where $u''_{**} = j''_*$ (the remaining cases can be treated similarly). Put $s_{**} := s_{z'_{**},u'_{**}} = s_{i''_{**},j''_{**}}$. Again (5.14) holds, and (5.15-16) are valid. Therefore, we get (5.17) using the same argument as above.

It remains case (c.2). Assume for simplicity $V_{q_*} = \{(i'_*, j'_*), (i''_*, j''_*), (i''_{**}, j''_{**})\}$, where the last two vertices belong to I''. Then similarly to (5.11) and three times using (5.12) one obtains

$$\begin{aligned} w_t &\le Cb^{|I|}|I|^3 \sum_{s_*<0}\ \sum_{s_*<s'_*<t}\ \sum_{s_*<s''_*<0} |s'_* - s_*|^{-(1+\theta)/2}|s''_* - s_*|^{-1-\theta}|t - s'_*|^{-1-\theta}|s''_*|^{-1-\theta} \\ &\le Cb^{|I|}|I|^3 t^{-(1+\theta)/2}, \end{aligned}$$

where $(1+\theta)/2 > \theta + \lambda\, (0 < \theta < 1)$. Lemma 5.2 is proved. ■

Corollary 5.3. *Under the conditions of Lemma 5.1,*

$$\mathrm{Cov}(\sigma^l_t, \sigma^l_0) = d_l^2 \mathrm{Cov}(\sigma_t, \sigma_0)(1 + o(1))$$

and

$$\mathrm{Cov}(y_{tl}, y_{0l}) = o(t^{-\theta}).$$

Proof. The first relation follows from Lemma 5.1 and Corollary 4.3. To show the second one, write

$$\mathrm{Cov}(y_{tl}, y_{0l}) = \mathrm{Cov}(\sigma_t^l, \sigma_0^l) - d_l(\mathrm{Cov}(\sigma_t^l, \sigma_0) + \mathrm{Cov}(\sigma_t, \sigma_0^l)) + d_l^2 \mathrm{Cov}(\sigma_t, \sigma_0),$$

and again apply Lemma 5.1 and Corollary 4.3. ■

References

Adenstedt, R. (1974). On large - sample estimation for the mean of a stationary random sequence. *Annals of Statistics*, **2,** 1095-1107.

Andersen, T.G. and Bollerslev, T. (1997). Heterogeneous information arrivals of return volatility dynamics: increasing the long run in high frequency returns. *Journal of Finance*, **52,** 975 - 1006.

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

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31,** 307–327.

Bollerslev, T. (1987). A conditional heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, **69,** 542 – 547.

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.

Davis, R.A. and Mikosch, T. (1999) The sample ACF of heavy-tailed stationary processes with applications to ARCH. *The Annals of Statistics, forthcoming.*

Davydov, Y. A (1970). The invariance principle for stationary processes. *Theory of Probability and its Applications*, **15,** 145–180.

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

Ding, Z., Granger, C.W.J. and Engle, R.F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, **1**, 83-106.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50,** 987–1008.

Giraitis, L. and Surgailis, D. (1989). Limit theorem for polynomials of linear processes with long range dependence. *Lietuvos Matematikos Rinkinys*, **29**, 128-145.

Giraitis, L., Kokoszka, P. and Leipus, R. (1998) Stationary ARCH models: dependence structure and central limit theorem. *Econometric Theory*, forthcoming.

Granger, C.W.J. and Andersen, A. P. (1978). *An Introduction to Bilinear Time Series models*. Vandenhoek & Ruprecht, Gottingen.

Granger, C.W.J. and Ding, Z. (1995). Some properties of absolute returns, an alternative measure of risk. *Annales d'Economie et de Statistique*, **40** , 67–91.

Harvey, A. (1998) Long memory in stochastic volatility. In: *Forecasting Volatility in the Financial Markets* (J. Knight and S. Satchell, eds.), Butterworth & Heinemann, London.

Lobato, I.N. and Savin, N. E. (1998). Real and spurious long-memory properties of stock market data (with comments). *Journal of Business and Economic Statistics* , **16,** 261–283.

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

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 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.

Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable Non-Gaussian Random Processes.* Chapman & Hall, New York.

Whistler, D.E.N. (1990). *Semiparametric Models of Daily and Intra-daily Exchange Rate Volatility.* Ph.D. thesis, University of London.

Liudas Giraitis, London School of Economics, Department of Economics, Houghton Street, London WC2A 2AE. *email:* L.Giraitis@lse.ac.uk.

Peter M. Robinson, London School of Economics, Department of Economics, Houghton Street, London WC2A 2AE. *email:* P.M.Robinson@lse.ac.uk

Donatas Surgailis, Institute of Mathematics and Informatics, Akademijos 4, 2600 Vilnius, Lithuania. *email:* sdonatas@ktl.mii.lt