The Annals of Applied Probability
2003, Vol. 13, No. 2, 641–668

ASYMPTOTIC RESULTS FOR LONG MEMORY LARCH SEQUENCES

BY ISTVÁN BERKES[1] AND LAJOS HORVÁTH

Hungarian Academy of Sciences and University of Utah

For a LARCH ("linear ARCH") sequence (y_n, σ_n) exhibiting long range dependence, we determine the limiting distribution of sums $\sum f(y_n)$, $\sum f(\sigma_n)$ for smooth functions f satisfying $E(y_0 f'(y_0)) \neq 0$, $E(\sigma_0 f'(\sigma_0)) \neq 0$. We also give an approximation formula for the above sums, providing the first term of the asymptotic expansions of $\sum f(y_n)$, $\sum f(\sigma_n)$.

1. Introduction. Since their introduction by Engle (1982) and Bollerslev (1986), ARCH and GARCH sequences have been used extensively to model financial time series, such as asset returns and exchange rates. A common property of ARCH(p) and GARCH(p, q) sequences is that they are defined by finite recursions and their autocorrelations decrease very rapidly, implying short memory behavior of these sequences. Short memory behavior holds even for ARCH(∞) models defined by the infinite recursion

$$y_k = \sigma_k \varepsilon_k, \tag{1.1}$$

$$\sigma_k^2 = a + \sum_{i=1}^{\infty} b_i y_{k-i}^2, \qquad k \in \mathbf{Z}, \tag{1.2}$$

where (b_k) is a sequence of nonnegative numbers and (ε_k) is an independent, identically distributed sequence of random variables having suitable moments. In fact, under reasonable conditions implying the existence of a covariance stationary solution of (1.1)–(1.2), we automatically have

$$\sum_{k=1}^{\infty} |\mathrm{Cov}(y_0, y_k)| < +\infty$$

implying short range dependence of the sequence [compare with Giraitis, Kokoszka and Leipus (2000)]. On the other hand, empirical evidence suggests in many typical financial situations a much greater degree of persistence in the process, indicating a long memory behavior of (y_k). A model describing such long

Received April 2001; revised May 2002.
[1]Supported by the Hungarian National Foundation for Scientific Research Grants T 29621 and T 37886.
AMS 2000 subject classifications. Primary 60F17; secondary 60K99.
Key words and phrases. LARCH sequences, long range dependence, asymptotic distribution, fractional Brownian motion.

memory behavior was suggested by Giraitis, Robinson and Surgailis (2000). They introduced a model called LARCH ("Linear ARCH") defined by

$$y_k = \sigma_k \varepsilon_k, \tag{1.3}$$

$$\sigma_k = a + \sum_{i=1}^{\infty} b_i y_{k-i}, \qquad k \in \mathbf{Z}. \tag{1.4}$$

As they showed, if $a \neq 0$, the ε_i are independent, identically distributed with $E\varepsilon_0 = 0$, $E\varepsilon_0^2 = 1$ and

$$\sum_{i=1}^{\infty} b_i^2 < 1, \tag{1.5}$$

then (1.3)–(1.4) has a unique stationary solution admitting a Volterra expansion

$$\sigma_n = a + a \sum_{k=1}^{\infty} \sum_{j_1,\dots,j_k=1}^{\infty} b_{j_1} \cdots b_{j_k} \varepsilon_{n-j_1} \cdots \varepsilon_{n-j_1-\cdots-j_k} \tag{1.6}$$

and if

$$b_j \sim c j^{-\beta}, \qquad 1/2 < \beta < 1 \tag{1.7}$$

then (σ_n) has long memory behavior. (Here $\sim$ means that the ratio of the left- and right-hand sides tends to 1.) In particular,

$$N^{-(3/2-\beta)} \sum_{j=1}^{[Nt]} (\sigma_j - a), \qquad 0 \le t \le 1$$

converges, as $N \to \infty$, not to the standard Brownian motion as it should be the case in a short memory situation, but to the fractional Brownian motion $W_{3/2-\beta}$. Let us recall that $\{W_\gamma(t), t \ge 0\}$, the fractional Brownian motion with parameter γ $(0 < \gamma < 1)$, is a Gaussian process with mean 0 and covariance

$$E W_\gamma(s) W_\gamma(t) = \tfrac{1}{2}(|s|^{2\gamma} + |t|^{2\gamma} - |s-t|^{2\gamma}).$$

See, for example, Samorodnitsky and Taqqu [(1994), Chapter 7].

The first profound analysis of a long memory situation in the probabilistic literature is due to Taqqu (1975, 1979) and Dobrushin and Major (1979) in the case of Gaussian processes. Specifically, they obtained the limit distribution of sums $\sum f(\xi_k)$ for a centered stationary Gaussian sequence (ξ_k) with covariance function $r_k \sim k^{-\alpha}$, $\alpha > 0$. The case $\alpha > 1$ is classical: in this case the sequence (ξ_k) has short memory and $N^{-1/2} \sum_{k=1}^{[Nt]} f(\xi_k)$ converges weakly to a multiple of the Wiener process for any f with $Ef(\xi_0) = 0$, $Ef^2(\xi_0) < +\infty$. In the difficult case $0 < \alpha < 1$ the limiting behavior of $\sum f(\xi_k)$ depends essentially on f; specifically, if $Ef(\xi_0) = 0$, $Ef^2(\xi_0) < +\infty$ and c_m is the first nonzero term in the Hermite expansion $f(x) = \sum c_k H_k(x)$, then with suitable norming a_N the sequence

$a_N^{-1}\sum_{k=1}^{[Nt]} f(\xi_k)$ converges weakly to a process $Z_m(t)$ (Hermite process) defined in terms of multiple Wiener integrals. For $m=1$, $Z_1(t)$ is fractional Brownian motion, but for $m\ge 2$, $Z_m(t)$ is non-Gaussian. In a subsequent paper Dehling and Taqqu (1989) determined the asymptotic behavior of the empirical process of (ξ_k). In the short memory case $\alpha>1$ the limiting process is a Gaussian process with mean 0 and covariance function $R(x,y)=\sum_{n\in\mathbf{Z}}\operatorname{Cov}(I\{\xi_0\le x\}, I\{\xi_n\le y\})$, which appears in many typical weakly dependent situations, but in the case $0<\alpha<1$ we get a totally different type of limiting process, whose trajectories are semi-deterministic, that is, are random multiples of a fixed deterministic function.

Surgailis (1982) [see also Giraitis and Surgailis (1986, 1989, 1999)] extended the Dobrushin–Major–Taqqu theory to linear (moving average) processes defined by

$$\xi_n=\sum_{j\in\mathbf{Z}} b_j\varepsilon_{n-j}, \qquad n\in\mathbf{Z},$$

where $\{\varepsilon_i, i\in Z\}$ are independent, identically distributed random variables with $E\varepsilon_0=0$, $E\varepsilon_0^2=1$ and b_j are real numbers with $\sum b_j^2<+\infty$. Long memory behavior holds here if

$$b_j\sim cj^{-\beta}, \qquad 1/2<\beta<1,$$

and Surgailis (1982) showed that the class of the limiting processes of $a_N^{-1}\times\sum_{k=1}^{[Nt]} f(\xi_k)$ is the same as in the Gaussian case, just the role of the Hermite polynomials is played by another polynomial sequence, the so-called Appell polynomials. Specifically, if f is smooth and $m\ge 1$ is the smallest integer with $E(f^{(m)}(\xi_0))\ne 0$, then the limiting process of $a_N^{-1}\sum_{k=1}^{[Nt]} f(\xi_k)$ is $Z_m(t)$. As a consequence, we get the same asymptotic behavior of the empirical process of the (ξ_k) as in the Gaussian case.

The purpose of our paper is to investigate the limiting behavior of sums $\sum f(\sigma_n)$, $\sum f(y_n)$, where (y_n,σ_n) is a long memory LARCH sequence defined by (1.3)–(1.4), where $a\ne 0$, ε_i are independent, identically distributed random variables with $E\varepsilon_0=0$, $E\varepsilon_0^2=1$ and b_j are positive numbers satisfying (1.7). The strong similarity between linear sequences and the basic recursion formula for LARCH sequences [see relation (1.21)] was exploited by Giraitis, Robinson and Surgailis (2000) to determine the limit distribution of $\sum(\sigma_n-a)$. In fact, they proved that the asymptotic behavior of $\sum(\sigma_n-a)$ is the same as that of $\sum\sigma_n^*$, where σ_n^* is the linear process defined by

$$\sigma_n^*=\sum_{j=1}^{\infty} b_j\delta_{n-j}, \qquad n\in\mathbf{Z}, \tag{1.8}$$

where δ_j are i.i.d. random variables with $\delta_0\overset{\mathcal{D}}{=}\varepsilon_0\sigma_0$. We will see that this phenomenon does not extend to $\sum f(\sigma_n)$ with general f: the variances of

$\sum_{n=1}^{N} f(\sigma_n)$ and $\sum_{n=1}^{N} f(\sigma_n^*)$ can grow with different speed and the sums have, in general, different limit distributions. However, we will show that $\sum f(\sigma_n)$, $\sum f(y_n)$ exhibit a long memory behavior, similar to that of Gaussian and linear processes, if $E(\sigma_n f'(\sigma_n)) \neq 0$ and $E(y_n f'(y_n)) \neq 0$. In this case the variances of the sums grow as $CN^{3-2\beta}$ and the sums, properly normalized, converge to the fractional Brownian motion. More precisely, we have:

THEOREM 1.1. *Assume that* $E\varepsilon_0 = 0$, $E\varepsilon_0^2 = 1$, $E|\varepsilon_0|^p < +\infty$ *for some* $p > 4$ *and that* (1.7) *holds with*

$$b^2 = \sum_{n=1}^{\infty} b_n^2 < \frac{p-1}{3(6p)^3 \|\varepsilon_0\|_p^2}, \tag{1.9}$$

where $\|\cdot\|$ *denotes the* L_p *norm. Let* $f : \mathbf{R} \to \mathbf{R}$ *be a twice continuously differentiable function with*

$$|f''(x)| \le C(|x|^\alpha + 1), \qquad x \in \mathbf{R}, \tag{1.10}$$

where

$$0 < \alpha < (p-4)^2/(2p). \tag{1.11}$$

Then

$$N^{-(3/2-\beta)} \sum_{n=1}^{[Nt]} \big(f(\sigma_n) - Ef(\sigma_n)\big) \xrightarrow{\mathcal{D}[0,1]} \gamma d W_{3/2-\beta}(t) \tag{1.12}$$

and

$$N^{-(3/2-\beta)} \sum_{n=1}^{[Nt]} \big(f(y_n) - Ef(y_n)\big) \xrightarrow{\mathcal{D}[0,1]} \gamma_1 d W_{3/2-\beta}(t), \tag{1.13}$$

where

$$\gamma = \frac{1}{a} E\big(\sigma_0 f'(\sigma_0)\big), \qquad \gamma_1 = \frac{1}{a} E\big(y_0 f'(y_0)\big) \tag{1.14}$$

and

$$d = \left(\frac{B(1-\beta, 2\beta-1)}{(1-\beta)(3-2\beta)}\right)^{1/2} \frac{ac}{(1-b^2)^{1/2}}. \tag{1.15}$$

Here a, b, c *are from* (1.4), (1.7), (1.9) *and* $B(\cdot,\cdot)$ *is the beta function.*

In particular, we have

$$N^{-(3/2-\beta)} \sum_{n=1}^{N} \big(f(\sigma_n) - Ef(\sigma_n)\big) \xrightarrow{\mathcal{D}} N(0, d^2\gamma^2), \tag{1.16}$$

$$N^{-(3/2-\beta)} \sum_{n=1}^{N} \big(f(y_n) - Ef(y_n)\big) \xrightarrow{\mathcal{D}} N(0, d^2\gamma_1^2). \tag{1.17}$$

As a comparison, we note that for a long-memory moving average process (y_n), relation (1.13) holds with $\gamma_1 d$ replaced by const $\cdot\, E(f'(y_0))$. As a consequence, there is a change in the behavior of the corresponding empirical processes, see our remarks below. The appearance of $E(\sigma_0 f'(\sigma_0))$ and $E(y_0 f'(y_0))$ in Theorem 1.1 indicates a new situation: the connection with the Appell expansions, underlying the structure of moving average processes, is lost. The fractional Brownian limits in (1.12), (1.13) correspond to the main term in the asymptotic formulas given by Theorem 1.2 below and become degenerate if $E(\sigma_0 f'(\sigma_0)) = 0$ or $E(y_0 f'(y_0)) = 0$. To determine the limit distribution in these degenerate cases would require finding the further terms in the asymptotic expansions of $\sum f(\sigma_n)$, $\sum f(y_n)$, a problem we will not deal with in the present paper, although the path of doing it is clearly indicated by the proof of our theorems.

Note that for large p we assume that $b^2 = \sum_{n=1}^{\infty} b_n^2$ is small. The actual bound in (1.9) is needed for the moment estimates in Lemmas 2.4, 2.5 and is similar to the bound on b required in Giraitis, Robinson and Surgailis (2000) for their asymptotic covariance estimates. Observe also that for large p, relations (1.10) and (1.11) permit polynomials f of degree $\sim p/2$ in our limit theorems. As a comparison, in the case of linear processes, the existence of finite p moments of the generating i.i.d. sequence permits to apply the corresponding limit theorem for polynomials of order $p/2$ [see Avram and Taqqu (1987)].

We note finally that $E(y_0) = E(\sigma_0\varepsilon_0) = 0$ and thus for $f(x) = x$ the γ_1 in (1.14) becomes 0 and thus the limit in (1.13) becomes degenerate. The explanation is that the y_n are orthogonal and thus $\sum_{n=1}^{N} y_n = O_P(\sqrt{N})$. Clearly, we get a nondegenerate limit if $f(x) = x^{2k}$, $k \in \mathbf{N}$.

While in the present paper we do not investigate the empirical processes of $(\sigma_n), (y_n)$, their asymptotic behavior is easy to obtain heuristically from Theorem 1.1. Let, for example, μ_N be the empirical process of $\{y_1, \ldots, y_N\}$ defined by

$$\mu_N\big((-\infty, x]\big) = N^{-(3/2-\beta)} \sum_{n=1}^{N} \big(I(y_n \le x) - P(y_n \le x)\big).$$

Relation (1.17) can be written equivalently as

$$\int_{-\infty}^{+\infty} f(x)\mu_N(dx) \xrightarrow{\mathcal{D}} N(0, d^2\gamma_1^2).$$

In particular, with $f(x) = e^{itx}$ we get

$$\int_{-\infty}^{+\infty} e^{itx}\mu_N(dx) \xrightarrow{\mathcal{D}} it\frac{d}{a}E(y_0 e^{ity_0})Z \tag{1.18}$$

for any $t \in \mathbf{R}$, where Z is a standard normal r.v. Assuming that y_0 has a density $\varphi(x)$ and introducing the process

$$\xi(x) = \frac{d}{a} x\varphi(x)Z, \qquad x \in \mathbf{R},$$

it is easy to verify that the right-hand side of (1.18) is $-\int_{-\infty}^{+\infty} e^{itx}\xi(dx)$ and thus, observing that $\xi(x)$ and $-\xi(x)$ have the same distribution, (1.18) suggests that

$$\mu_N\big((-\infty, x]\big) \overset{\mathcal{D}[-\infty,+\infty]}{\longrightarrow} \xi(x). \tag{1.19}$$

A similar heuristic suggests, in view of (1.13), the two-parameter convergence

$$\mu_{[Nt]}\big((-\infty, x]\big) \longrightarrow \frac{d}{a} x\varphi(x) W_{3/2-\beta}(t) \tag{1.20}$$

in $\mathcal{D}([0,1] \times [-\infty, +\infty])$. However, the precise verification of (1.19) and (1.20) is quite laborious, requiring the chaining technique employed in Dehling and Taqqu (1989) and Ho and Hsing (1996) and will be postponed to a subsequent paper.

In contrast to the semideterministic limit process $\xi(x)$ appearing in (1.19), the empirical processes of ARCH and GARCH models in the short memory case converge to nondegenerate Gaussian processes and the limit of $\mu_{[Nt]}((-\infty, x])$ in (1.20) in the short term memory case is a Kiefer process exhibiting random behavior in both parameters. See Berkes and Horváth (2001).

We formulate now a stronger version of Theorem 1.1, which also reveals the reason for the relations (1.12)–(1.13). For this purpose, we introduce some new LARCH type sequences associated with (σ_n). Clearly, the sum of those terms in the doubly infinite sum in (1.6) which contain $\varepsilon_{n-\ell}$, but no ε_ν with $\nu > n - \ell$ is

$$\begin{aligned} b_\ell \varepsilon_{n-\ell} & \left(1 + \sum_{k=1}^{\infty} \sum_{j_1,\dots,j_k=1}^{\infty} b_{j_1} \cdots b_{j_k} \varepsilon_{n-\ell-j_1} \cdots \varepsilon_{n-\ell-j_1-\cdots-j_k}\right) \\ &= b_\ell \varepsilon_{n-\ell}\left(1 + \frac{\sigma_{n-\ell} - a}{a}\right) = \frac{1}{a} b_\ell \varepsilon_{n-\ell}\sigma_{n-\ell}. \end{aligned}$$

Hence

$$\sigma_n - a = b_1\varepsilon_{n-1}\sigma_{n-1} + b_2\varepsilon_{n-2}\sigma_{n-2} + \cdots. \tag{1.21}$$

Let f be a function satisfying (1.10), (1.11) and define the sequences $\{\sigma_n^{(f)}, n \in \mathbf{Z}\}$, $\{\bar{\sigma}_n^{(f)}, n \in \mathbf{Z}\}$ by

$$\sigma_n^{(f)} = B_1\varepsilon_{n-1}\sigma_{n-1} + B_2\varepsilon_{n-2}\sigma_{n-2} + \cdots, \tag{1.22}$$

$$\bar{\sigma}_n^{(f)} = \bar{B}_1\varepsilon_{n-1}\sigma_{n-1} + \bar{B}_2\varepsilon_{n-2}\sigma_{n-2} + \cdots, \tag{1.23}$$

where

$$B_\ell = E\big(f'(\sigma_0)\zeta_\ell\big), \qquad \bar{B}_\ell = E\big(f'(y_0)\varepsilon_0\zeta_\ell\big) \tag{1.24}$$

with

$$\zeta_\ell = \sum_{\substack{r\ge 1\\ j_1,\dots,j_r\ge 1\\ j_1+\cdots+j_r=\ell}} b_{j_1}\cdots b_{j_r}\varepsilon_{-j_1}\cdots\varepsilon_{-j_1-\cdots-j_{r-1}}. \tag{1.25}$$

That is, $\sigma_n^{(f)}$ and $\bar{\sigma}_n^{(f)}$ are obtained from $\sigma_n - a$ by replacing the coefficients b_j in (1.21) by B_j and $\bar{B}_j$, respectively. Now we have:

THEOREM 1.2. *Under the assumptions of Theorem* 1.1 *we have*

$$\sum_{n=1}^{N}\big(f(\sigma_n)-Ef(\sigma_n)\big)=\sum_{n=1}^{N}\sigma_n^{(f)}+CN^{3/2-\beta-\varepsilon}\xi_N \tag{1.26}$$

and

$$\sum_{n=1}^{N}\big(f(y_n)-Ef(y_n)\big)=\sum_{n=1}^{N}\bar{\sigma}_n^{(f)}+CN^{3/2-\beta-\varepsilon}\eta_N \tag{1.27}$$

for some $C>0$, $\varepsilon>0$ *where* $E\xi_N^2\le 1$, $E\eta_N^2\le 1$. *Moreover, the* B_ℓ, $\bar{B}_\ell$ *in definitions* (1.22) *and* (1.23) *satisfy*

$$B_\ell \sim \gamma b_\ell, \qquad \bar{B}_\ell \sim \gamma_1 b_\ell \qquad \text{as } \ell\to\infty \tag{1.28}$$

with the γ *and* γ_1 *in* (1.14).

Relations (1.26) and (1.27) are invariance principles for $\sum_{n=1}^N f(\sigma_n)$ and $\sum_{n=1}^N f(y_n)$ and reduce their study to those of $\sum_{n=1}^N \sigma_n^{(f)}$, $\sum_{n=1}^N \bar{\sigma}_n^{(f)}$. For the original (σ_n), Giraitis, Robinson and Surgailis (2000) proved that

$$N^{-(3/2-\beta)}\sum_{n=1}^{[Nt]}(\sigma_n-a)\xrightarrow{\mathcal{D}[0,1]} dW_{3/2-\beta}(t) \tag{1.29}$$

with the d in (1.15). [Actually, they showed only the convergence of finite dimensional distributions in (1.29), but the tightness follows from

$$E\left(\sum_{n=1}^{N}(\sigma_n-a)\right)^2\sim C_1N^{3-2\beta}$$

which, in turn, is a consequence of their Corollary 2.1, and Theorem 15.6 of Billingsley (1968). See the analogous argument for $N^{-(3/2-\beta)}R_{[Nt]}$ at the end of our paper.] Using (1.28), the same proof shows that

$$N^{-(3/2-\beta)}\sum_{n=1}^{[Nt]}\sigma_n^{(f)}\xrightarrow{\mathcal{D}[0,1]} d\gamma W_{3/2-\beta}(t) \tag{1.30}$$

and

$$N^{-(3/2-\beta)} \sum_{n=1}^{[Nt]} \bar{\sigma}_n^{(f)} \xrightarrow{\mathcal{D}[0,1]} d\gamma_1 W_{3/2-\beta}(t) \tag{1.31}$$

and thus Theorem 1.2 implies Theorem 1.1 (see Section 2).

2. Proof of the theorems. As we have already noted, the asymptotic behavior of $\sum(\sigma_n - a)$ is the same as that of $\sum \sigma_n^*$, where σ_n^* is the linear process defined by

$$\sigma_n^* = b_1\delta_{n-1} + b_2\delta_{n-2} + \cdots, \qquad n \in \mathbf{Z}, \tag{2.1}$$

where δ_j are i.i.d. random variables with $\delta_0 \overset{\mathcal{D}}{=} \varepsilon_0\sigma_0$. While this similarity does not extend to $\sum f(\sigma_n)$, we will make an essential use of the theory of linear processes in our arguments. In particular, we will utilize the martingale decomposition technique used by Ho and Hsing (1996) to give an Edgeworth expansion of the empirical process of long memory moving average processes.

Let us first note that

$$\{\varepsilon_{\nu_1}\cdots\varepsilon_{\nu_r},\ 1 \le \nu_1 < \cdots < \nu_r,\ r = 1, 2, \ldots\}$$

is an orthonormal system and also that $\sum b_j^2 < 1$ implies that the sum of squares of the coefficients in the sum in (1.6) is finite. Thus the series in (1.6) converges in L_2 norm under any ordering of its terms. Since the above orthonormal system is also complete, its L_2 sum is independent of the order of its terms. The same remark will apply to all infinite sums of r.v.'s appearing in the sequel.

Let $\mathcal{F}_\ell = \sigma\{\varepsilon_\nu, \nu \le \ell\}$ and

$$X_{n,\ell} = E\big(f(\sigma_n)|\mathcal{F}_{n-\ell}\big) - E\big(f(\sigma_n)|\mathcal{F}_{n-\ell-1}\big). \tag{2.2}$$

Then

$$\begin{aligned}\sum_{\ell=1}^{L} X_{n,\ell} &= E\big(f(\sigma_n)|\mathcal{F}_{n-1}\big) - E\big(f(\sigma_n)|\mathcal{F}_{n-L-1}\big)\\ &= f(\sigma_n) - E\big(f(\sigma_n)|\mathcal{F}_{n-L-1}\big)\end{aligned} \tag{2.3}$$

since $f(\sigma_n)$ is $\mathcal{F}_{n-1}$ measurable by (1.6). For fixed n and $L \to \infty$, the last conditional expectation in (2.3) converges to $Ef(\sigma_n)$ by the martingale convergence theorem and thus

$$f(\sigma_n) - Ef(\sigma_n) = \sum_{\ell=1}^{\infty} X_{n,\ell}. \tag{2.4}$$

Our first lemma gives an approximation formula for $X_{n,\ell}$.

LEMMA 2.1. *Under the conditions of Theorem* 1.1 *we have*

$$X_{n,\ell} = E\big(f'(\sigma_n)\zeta_{n,\ell}\big)\sigma_{n-\ell}\varepsilon_{n-\ell} + R_{n,\ell}, \tag{2.5}$$

where

$$\zeta_{n,\ell} = \sum_{\substack{r\geq 1\\ j_1,\dots,j_r\geq 1\\ j_1+\cdots+j_r=\ell}} b_{j_1}\cdots b_{j_r}\varepsilon_{n-j_1}\cdots\varepsilon_{n-j_1-\cdots-j_{r-1}} \stackrel{\mathcal{D}}{=} \zeta_\ell \tag{2.6}$$

and

$$\begin{aligned} R_{n,\ell} &= \big\{E\big(f'(\sigma_n)\zeta_{n,\ell}|\mathcal{F}_{n-\ell}\big) - E\big(f'(\sigma_n)\zeta_{n,\ell}\big)\big\}\sigma_{n-\ell}\varepsilon_{n-\ell} \\ &\quad + c_p\theta\sigma_{n-\ell}^2\big\{(\varepsilon_{n-\ell}^2+1)E(|\sigma_n|^\alpha\zeta_{n,\ell}^2|\mathcal{F}_{n-\ell}) \\ &\quad + (|\varepsilon_{n-\ell}|^{p/2}+\varepsilon_{n-\ell}^2+2)E\zeta_{n,\ell}^2 + E(|\sigma_n|^{\alpha p/(p-4)}\zeta_{n,\ell}^2|\mathcal{F}_{n-\ell-1})\big\}, \end{aligned} \tag{2.7}$$

where $c_p = Cp^{-1}2^{p/2}E|\varepsilon_0|^{p/2}$ *with the* C *in* (1.10), *and* $|\theta|\leq 1$.

For $r=1$ we get the constant term b_ℓ in (2.6). Actually, in the case when $\sigma_n = \sum_{j=1}^\infty b_j\varepsilon_{n-j}$ is a linear process, the analogue of Lemma 2.1 holds with $\zeta_{n,\ell} = b_\ell$ and thus the effect of the nonlinear terms in (1.6) is given by the nonconstant terms of $\zeta_{n,\ell}$ in (2.6).

Adding (2.5) for $\ell = 1,2,\dots$ and $n = 1,\dots,N$ and observing that the coefficient $E(f'(\sigma_n)\zeta_{n,\ell})$ in (2.5) equals B_ℓ in (1.24) by stationarity, we get, in view of (1.22) and (2.4),

$$\sum_{n=1}^N\big(f(\sigma_n) - Ef(\sigma_n)\big) = \sum_{n=1}^N \sigma_n^{(f)} + \sum_{n=1}^N\sum_{\ell=1}^\infty R_{n,\ell}. \tag{2.8}$$

Hence the proof of the theorems will be reduced to an asymptotic evaluation of $\sum_{n=1}^N \sigma_n^{(f)}$ and $\sum_{n=1}^N\sum_{\ell=1}^\infty R_{n,\ell}$ which will be done in a series of lemmas.

PROOF OF LEMMA 2.1. We have seen above that the sum

$$\sum_{k=1}^\infty \sum_{j_1,\dots,j_k=1}^\infty b_{j_1}\cdots b_{j_k}\varepsilon_{n-j_1}\cdots\varepsilon_{n-j_1-\cdots-j_k} \tag{2.9}$$

converges in L_2 with any ordering of its terms. Actually, this remains valid if $\ell\geq 1$ and we replace $\varepsilon_{n-1},\dots,\varepsilon_{n-\ell}$ by arbitrary real numbers $u_1,\dots,u_\ell$. For example, if ε_{n-1} and ε_{n-2} are replaced by u_1 and u_2, then the resulting series in (2.9) can be broken into 4 series, according as their terms contain both u_1 and u_2, only u_1, only u_2 and none of u_1, u_2, respectively. Factoring out u_1u_2, u_1, u_2, respectively, in the first 3 series, their convergence can be seen directly, in analogy with (2.9), proving our claim. In other words, letting (formally)

$$\psi(x_1, x_2, \dots) = a + a \sum_{k=1}^{\infty} \sum_{j_1,\dots,j_k=1}^{\infty} b_{j_1} \cdots b_{j_k} x_{j_1} \cdots x_{j_1+\cdots+j_k}, \tag{2.10}$$

the expressions $\psi(u_1, \dots, u_\ell, \varepsilon_{n-\ell-1}, \varepsilon_{n-\ell-2}, \dots)$ are well defined for any $\ell \geq 1$ and any real $u_1, \dots, u_\ell$. Clearly

$$\sigma_n = \psi(\varepsilon_{n-1}, \varepsilon_{n-2}, \dots).$$

Keeping in mind that we will use the functions $\psi(x_1, x_2, \dots)$ only when there exists an n such that $x_j = \varepsilon_{n-j}$ with finitely many exceptions, it is clear that the sum of terms in the infinite sum in (2.10) containing x_ℓ but no x_j with $j < \ell$ is

$$\begin{aligned} b_\ell x_\ell & \left(1 + \sum_{k=1}^{\infty} \sum_{j_1,\dots,j_k=1}^{\infty} b_{j_1} \cdots b_{j_k} x_{\ell+j_1} \cdots x_{\ell+j_1+\cdots+j_k}\right) \\ &= b_\ell x_\ell \left(1 + \frac{\psi(x_{\ell+1}, x_{\ell+2}, \dots) - a}{a}\right) \\ &= \frac{1}{a} b_\ell x_\ell \psi(x_{\ell+1}, x_{\ell+2}, \dots). \end{aligned}$$

On the other hand, the sum of terms in the sum in (2.10) containing x_ℓ is

$$x_\ell \sum_{\substack{r \geq 1 \\ j_1, \dots, j_r \geq 1 \\ j_1 + \cdots + j_r = \ell}} b_{j_1} \cdots b_{j_r} x_{j_1} \cdots x_{j_1+\cdots+j_{r-1}} \left(1 + \frac{\psi(x_{\ell+1}, x_{\ell+2}, \dots) - a}{a}\right).$$

Clearly $E(f(\sigma_n)|\mathcal{F}_{n-\ell})$ is obtained by integrating $f(\sigma_n) = f(\psi(\varepsilon_{n-1}, \varepsilon_{n-2}, \dots))$ with respect to $\varepsilon_{n-1}, \varepsilon_{n-2}, \dots, \varepsilon_{n-\ell+1}$, more precisely,

$$\begin{aligned} & E\big(f(\sigma_n)|\mathcal{F}_{n-\ell}\big) \\ &\quad = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f\big(\psi(u_1, \dots, u_{\ell-1}, \varepsilon_{n-\ell}, \dots)\big)\, dG(u_1) \cdots dG(u_{\ell-1}), \end{aligned}$$

where G denotes the distribution function of ε_0. Similarly,

$$\begin{aligned} & E\big(f(\sigma_n)|\mathcal{F}_{n-\ell-1}\big) \\ &\quad = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f\big(\psi(u_1, \dots, u_{\ell-1}, v, \varepsilon_{n-\ell-1}, \dots)\big) \\ &\qquad \times dG(u_1) \cdots dG(u_{\ell-1})\, dG(v). \end{aligned}$$

Thus

$$\begin{aligned} X_{n,\ell} = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} & \big[f\big(\psi(u_1, \dots, u_{\ell-1}, \varepsilon_{n-\ell}, \dots)\big) \\ & - f\big(\psi(u_1, \dots, u_{\ell-1}, v, \varepsilon_{n-\ell-1}, \dots)\big)\big]\, dG(u_1) \cdots dG(u_{\ell-1})\, dG(v). \end{aligned} \tag{2.11}$$

Using a two-term Taylor expansion, the integrand in (2.11) becomes

$$\begin{aligned} &f'\big(\psi(u_1,\dots,u_{\ell-1},\varepsilon_{n-\ell},\dots)\big)\\ &\quad\times[\psi(u_1,\dots,u_{\ell-1},\varepsilon_{n-\ell},\dots)-\psi(u_1,\dots,u_{\ell-1},v,\varepsilon_{n-\ell-1},\dots)]\\ &\quad+\tfrac{1}{2}f''(\tau^*)[\psi(u_1,\dots,u_{\ell-1},\varepsilon_{n-\ell},\dots)\\ &\qquad-\psi(u_1,\dots,u_{\ell-1},v,\varepsilon_{n-\ell-1},\dots)]^2, \end{aligned}\tag{2.12}$$

where τ^* lies between $\psi(u_1,\dots,u_{\ell-1},\varepsilon_{n-\ell},\dots)$ and $\psi(u_1,\dots,u_{\ell-1},v,\varepsilon_{n-\ell-1},\dots)$. By the above remarks on the structure of ψ we see that

$$\begin{aligned} &\psi(u_1,\dots,u_{\ell-1},\varepsilon_{n-\ell},\dots)-\psi(u_1,\dots,u_{\ell-1},v,\varepsilon_{n-\ell-1},\dots)\\ &\quad=a\big(1+\big(\psi(\varepsilon_{n-\ell-1},\dots)-a\big)/a\big)(\varepsilon_{n-\ell}-v)S\\ &\quad=\sigma_{n-\ell}(\varepsilon_{n-\ell}-v)S, \end{aligned}\tag{2.13}$$

where

$$S=S(u_1,\dots,u_{\ell-1})=\sum_{\substack{r\ge 1\\ j_1,\dots,j_r\ge 1\\ j_1+\cdots+j_r=\ell}} b_{j_1}\cdots b_{j_r}u_{j_1}u_{j_1+j_2}\cdots u_{j_1+\cdots+j_{r-1}}.$$

Thus using $\int_{-\infty}^{+\infty}dG(v)=1$ and $\int_{-\infty}^{+\infty}v\,dG(v)=0$, we see that the contribution of the first term of the Taylor expansion (2.12) in the integral (2.11) is

$$\begin{aligned} &\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f'\big(\psi(u_1,\dots,u_{\ell-1},\varepsilon_{n-\ell},\dots)\big)\\ &\qquad\times\sigma_{n-\ell}(\varepsilon_{n-\ell}-v)S(u_1,\dots,u_{\ell-1})\,dG(u_1)\cdots dG(u_{\ell-1})\,dG(v)\\ &\quad=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f'\big(\psi(u_1,\dots,u_{\ell-1},\varepsilon_{n-\ell},\dots)\big)\\ &\qquad\times\sigma_{n-\ell}\varepsilon_{n-\ell}S(u_1,\dots,u_{\ell-1})\,dG(u_1)\cdots dG(u_{\ell-1})\\ &\quad=E\left\{f'\big(\psi(\varepsilon_{n-1},\varepsilon_{n-2},\dots)\big)\sum_{\substack{r\ge 1\\ j_1,\dots,j_r\ge 1\\ j_1+\cdots+j_r=\ell}} b_{j_1}\cdots b_{j_r}\varepsilon_{n-j_1}\cdots\varepsilon_{n-j_1-\cdots-j_{r-1}}\bigg|\mathcal{F}_{n-\ell}\right\}\\ &\qquad\times\sigma_{n-\ell}\varepsilon_{n-\ell}\\ &\quad=E\big(f'(\sigma_n)\zeta_{n,\ell}|\mathcal{F}_{n-\ell}\big)\sigma_{n-\ell}\varepsilon_{n-\ell}. \end{aligned}$$

Relations (2.13) and (1.10) show that the second term in (2.12) is at most

$$\frac{C}{2}\{|\psi(u_1,\dots,u_{\ell-1},\varepsilon_{n-\ell},\dots)|^\alpha+|\psi(u_1,\dots,u_{\ell-1},v,\varepsilon_{n-\ell-1},\dots)|^\alpha+1\}\\ \times\sigma_{n-\ell}^2(\varepsilon_{n-\ell}-v)^2S^2$$

and thus the contribution of the second Taylor term in the integral (2.11) is at most $I_1+I_2+I_3$, where

$$I_1=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\frac{C}{2}|\psi(u_1,\dots,u_{\ell-1},\varepsilon_{n-\ell},\dots)|^\alpha\sigma_{n-\ell}^2(\varepsilon_{n-\ell}-v)^2S^2\\ \times dG(u_1)\cdots dG(u_{\ell-1})\,dG(v),$$

$$I_2=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\frac{C}{2}|\psi(u_1,\dots,u_{\ell-1},v,\varepsilon_{n-\ell-1},\dots)|^\alpha\sigma_{n-\ell}^2(\varepsilon_{n-\ell}-v)^2S^2\\ \times dG(u_1)\cdots dG(u_{\ell-1})\,dG(v),$$

$$I_3=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\frac{C}{2}\sigma_{n-\ell}^2(\varepsilon_{n-\ell}-v)^2S^2\,dG(u_1)\cdots dG(u_{\ell-1})\,dG(v).$$

Using $\int_{-\infty}^{+\infty}v\,dG(v)=0$, $\int_{-\infty}^{+\infty}v^2\,dG(v)=1$ again, we get

$$\begin{aligned}I_1&=\tfrac12\sigma_{n-\ell}^2\int_{-\infty}^{+\infty}(\varepsilon_{n-\ell}-v)^2\,dG(v)\\ &\qquad\times\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}|\psi(u_1,\dots,u_{\ell-1},\varepsilon_{n-\ell},\dots)|^\alpha S^2\,dG(u_1)\cdots dG(u_{\ell-1})\\ &=\tfrac12\sigma_{n-\ell}^2(\varepsilon_{n-\ell}^2+1)E\big(|\psi(\varepsilon_{n-1},\dots,\varepsilon_{n-\ell+1},\varepsilon_{n-\ell},\dots)|^\alpha\zeta_{n,\ell}^2|\mathcal{F}_{n-\ell}\big)\\ &=\tfrac12\sigma_{n-\ell}^2(\varepsilon_{n-\ell}^2+1)E(|\sigma_n|^\alpha\zeta_{n,\ell}^2|\mathcal{F}_{n-\ell}).\end{aligned}$$

On the other hand, applying the inequality

$$|xy|\le|x|^s/s+|y|^t/t,\qquad s>1,\ s^{-1}+t^{-1}=1$$

[see, e.g., Loève (1977), page 157], we see that the integrand in I_2 is bounded by

$$\frac12\sigma_{n-\ell}^2S^2\bigg(\frac4p|\varepsilon_{n-\ell}-v|^{p/2}+\frac{p-4}{p}|\psi(u_1,\dots,u_{\ell-1},v,\varepsilon_{n-\ell-1},\dots)|^{\alpha p/(p-4)}\bigg)\\ =: J_1+J_2.$$

The contribution of J_1 in the integral I_2 is

$$\begin{aligned}
&\frac{2}{p}\sigma_{n-\ell}^2 \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} S^2 |\varepsilon_{n-\ell} - v|^{p/2}\, dG(u_1) \cdots dG(u_{\ell-1})\, dG(v) \\
&\quad = \frac{2}{p}\sigma_{n-\ell}^2 \int_{-\infty}^{+\infty} |\varepsilon_{n-\ell} - v|^{p/2}\, dG(v) \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} S^2\, dG(u_1) \cdots dG(u_{\ell-1}) \\
&\quad = \frac{2}{p}\sigma_{n-\ell}^2 \int_{-\infty}^{+\infty} |\varepsilon_{n-\ell} - v|^{p/2}\, dG(v) \cdot E\zeta_{n,\ell}^2 \\
&\quad \le \frac{2}{p} 2^{p/2-1} \sigma_{n-\ell}^2 (E\zeta_{n,\ell}^2) \int_{-\infty}^{+\infty} (|\varepsilon_{n-\ell}|^{p/2} + |v|^{p/2})\, dG(v) \\
&\quad \le c_p \sigma_{n-\ell}^2 (|\varepsilon_{n-\ell}|^{p/2} + 1) E\zeta_{n,\ell}^2,
\end{aligned}$$

where

$$c_p = \frac{1}{p} 2^{p/2} E|\varepsilon_0|^{p/2}.$$

Here we used the inequality $|x+y|^\gamma \le 2^{\gamma-1}(|x|^\gamma + |y|^\gamma)$ $(\gamma \ge 1)$, following from the convexity of $|x|^\gamma$. On the other hand, the contribution of J_2 in the integral I_2 is

$$\begin{aligned}
&\frac{p-4}{2p}\sigma_{n-\ell}^2 \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} |\psi(u_1, \ldots, u_{\ell-1}, v, \varepsilon_{n-\ell-1}, \ldots)|^{\alpha p/(p-4)} \\
&\qquad\qquad \times S^2\, dG(u_1) \cdots dG(u_{\ell-1})\, dG(v) \\
&\quad = \frac{p-4}{2p}\sigma_{n-\ell}^2 E\big(|\psi(\varepsilon_{n-1}, \ldots)|^{\alpha p/(p-4)} \zeta_{n,\ell}^2 | \mathcal{F}_{n-\ell-1}\big) \\
&\quad = \frac{p-4}{2p}\sigma_{n-\ell}^2 E(|\sigma_n|^{\alpha p/(p-4)} \zeta_{n,\ell}^2 | \mathcal{F}_{n-\ell-1}).
\end{aligned}$$

Finally,

$$\begin{aligned}
I_3 &= \frac{C}{2}\sigma_{n-\ell}^2 (\varepsilon_{n-\ell}^2 + 1) \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} S^2\, dG(u_1) \cdots dG(u_{\ell-1}) \\
&= \frac{C}{2}\sigma_{n-\ell}^2 (\varepsilon_{n-\ell}^2 + 1) E\zeta_{n,\ell}^2.
\end{aligned}$$

Collecting the terms, we get Lemma 2.1. □

We next give an asymptotic formula for B_ℓ in (1.24). To this end we prove the following elementary lemma.

LEMMA 2.2. *If* (1.7) *holds and* $\sum_{n=1}^{\infty} b_n^2 < 1$, *then*

$$\sum_{\substack{r\ge 1\\ i_1,\dots,i_r\ge 1\\ i_1+\cdots+i_r=n}} b_{i_1}^2\cdots b_{i_r}^2 \le C b_n^2 \tag{2.14}$$

with some constant $C > 0$.

PROOF. Let $b^2 = \sum_{n=1}^{\infty} b_n^2 < 1$ and let S_n denote the sum in (2.14). Clearly, for any fixed $1 \le \ell \le n-1$ the contribution of the terms in S_n with $i_1 = \ell$ is $b_\ell^2 S_{n-\ell}$ and thus

$$S_n = \sum_{i=1}^{n} b_i^2 S_{n-i}, \tag{2.15}$$

where we put $S_0 = 1$. Let further $\Sigma_N = S_1 + \cdots + S_N$ for $N \ge 1$. Then by (2.15) we have

$$\Sigma_N \le \left(\sum_{i=1}^{\infty} b_i^2\right)(1+\Sigma_N) \qquad \text{for all } N \ge 1$$

and thus

$$\Sigma_\infty \le \frac{b^2}{1-b^2} < +\infty.$$

Choose $\delta > 0$ so small that $b^2(1+\delta) < 1$. Let $g_t = \sup_{i \ge t} b_i^2$ for $t > 0$, then g_t is nonincreasing and by (1.7) we have $g_n/b_n^2 \to 1$ as $n \to \infty$. Thus we can choose a small $0 < \varepsilon < 1$ so that

$$g_{n-n\varepsilon} \le (1+\delta) b_n^2, \qquad g_{n\varepsilon} < \frac{2}{\varepsilon^2}\, b_n^2 \tag{2.16}$$

for $n \ge n_0$. Let $C > 0$ be so large that

$$C \ge \frac{2}{\varepsilon^2}\,(1+\Sigma_\infty)\,\frac{1}{1-b^2(1+\delta)} \tag{2.17}$$

and that $S_n \le C b_n^2$ holds for $1 \le n \le n_0$. We show by induction that $S_n \le C b_n^2$ for all $n \ge 1$. Indeed, if $n > n_0$ and $S_k \le C b_k^2$ holds for $1 \le k \le n-1$, then we get, by (2.15)–(2.17) and the induction hypothesis,

$$\begin{aligned} S_n &= \sum_{i \le n\varepsilon} b_i^2 S_{n-i} + \sum_{i > n\varepsilon} b_i^2 S_{n-i} \\ &\le \left(\sum_{i=1}^{\infty} b_i^2\right) \max_{n-n\varepsilon \le j \le n-1} S_j + \left(\sup_{i > n\varepsilon} b_i^2\right)\left(\sum_{j=0}^{\infty} S_j\right) \end{aligned}$$

$$
\begin{aligned}
&\le b^2 C\Big(\max_{n-n\varepsilon\le j\le n-1} b_j^2\Big) + g_{n\varepsilon}(1+\Sigma_\infty)\\
&\le b^2 C g_{n-n\varepsilon} + g_{n\varepsilon}(1+\Sigma_\infty)\\
&\le b^2 C(1+\delta)b_n^2 + \frac{2}{\varepsilon^2}(1+\Sigma_\infty)\,b_n^2\\
&\le C b_n^2.
\end{aligned}
$$

This completes the induction step and the proof of Lemma 2.2. □

REMARK. The previous argument shows that for any $\eta > 0$ we have

$$\sum_{i=1}^{n-1} b_i^2 b_{n-i}^2 < 2(1+\eta)b^2 b_n^2 \qquad \text{for } n \ge n_0(\eta).$$

Indeed, let ε, δ denote the quantities introduced above and split the sum I_n on the left-hand side of the last relation into 3 sums $I_{n,1}$, $I_{n,2}$, $I_{n,3}$ containing the terms $i \le n\varepsilon$, $n\varepsilon < i < n - n\varepsilon$, $i \ge n - n\varepsilon$. Then we get, using the estimates above,

$$I_{n,1} \le b^2 g_{n-n\varepsilon}, \qquad I_{n,2} \le n g_{n\varepsilon}^2, \qquad I_{n,3} \le b^2 g_{n-n\varepsilon},$$

so that

$$I_n \le 2b^2(1+\delta)b_n^2 + \frac{4}{\varepsilon^4}\, n b_n^4 \le 2b^2(1+2\delta)b_n^2$$

for sufficiently large n since $nb_n^2 \to 0$ by (1.7). Since δ can be chosen arbitrary small, our claim is proved.

We can now prove the following.

LEMMA 2.3. *We have*

$$B_\ell \sim \gamma b_\ell \qquad \text{as } \ell \to \infty, \tag{2.18}$$

where γ is defined by (1.14).

PROOF. The constant term of the sum ζ_ℓ in (1.25) (obtained for $r = 1$) is b_ℓ and thus we can write

$$
\begin{aligned}
\zeta_\ell &= b_\ell + \sum_{\substack{s\ge 1\\ j_1,\dots,j_s\ge 1\\ j_1+\cdots+j_s<\ell}} b_{j_1}\cdots b_{j_s} b_{\ell-j_1-\dots-j_s}\varepsilon_{-j_1}\cdots\varepsilon_{-j_1-\dots-j_s}\\
&=: b_\ell + \zeta_\ell^*.
\end{aligned}
$$

Next we observe that (1.10) implies by integration

$$|f'(x)| \le C_1 + C_1|x|^{\alpha+1}, \qquad x \in \mathbf{R},$$

for some constant $C_1 > 0$ and thus

$$Ef'(\sigma_0)^2 \le E(C_1 + C_1|\sigma_0|^{\alpha+1})^2 < +\infty$$

since $2\alpha + 2 \le p$ by (1.11) and $E|\sigma_0|^p < +\infty$ (see Lemma 2.5). Since the sequence

$$\{\varepsilon_{\nu_1} \cdots \varepsilon_{\nu_s},\ s \ge 1, 1 \le \nu_1 < \cdots < \nu_s\}$$

is orthonormal, it follows that

$$\sum_{\substack{s \ge 1 \\ j_1, \dots, j_s \ge 1}} E^2\big(f'(\sigma_0)\varepsilon_{-j_1} \cdots \varepsilon_{-j_1 - \cdots - j_s}\big) < +\infty$$

and thus for any $\delta > 0$ there exists a $K(\delta) > 0$ such that $\lim_{\delta \to 0} K(\delta) = +\infty$ and

$$\sum_{\substack{s \ge 1 \\ j_1, \dots, j_s \ge 1 \\ j_1 + \cdots + j_s \ge K(\delta)}} E^2\big(f'(\sigma_0)\varepsilon_{-j_1} \cdots \varepsilon_{-j_1 - \cdots - j_s}\big) < \delta. \tag{2.19}$$

Let $\ell > K(\delta)$ and write

$$\zeta_\ell^* = \zeta_\ell^{(1)} + \zeta_\ell^{(2)},$$

where

$$\zeta_\ell^{(1)} = \sum_{\substack{s \ge 1 \\ j_1, \dots, j_s \ge 1 \\ j_1 + \cdots + j_s < K(\delta)}} b_{j_1} \cdots b_{j_s} b_{\ell - j_1 - \cdots - j_s} \varepsilon_{-j_1} \cdots \varepsilon_{-j_1 - \cdots - j_s},$$

$$\zeta_\ell^{(2)} = \sum_{\substack{s \ge 1 \\ j_1, \dots, j_s \ge 1 \\ K(\delta) \le j_1 + \cdots + j_s < \ell}} b_{j_1} \cdots b_{j_s} b_{\ell - j_1 - \cdots - j_s} \varepsilon_{-j_1} \cdots \varepsilon_{-j_1 - \cdots - j_s}.$$

By the Cauchy–Schwarz inequality, (2.19) and Lemma 2.2, we get

$$\begin{aligned}
&\big|E\big(f'(\sigma_0)\zeta_\ell^{(2)}\big)\big| \\
&\quad\le \sum_{\substack{s\ge 1\\ j_1,\dots,j_s\ge 1\\ K(\delta)\le j_1+\dots+j_s<\ell}} |b_{j_1}\cdots b_{j_s} b_{\ell-j_1-\dots-j_s}|\big|E\big(f'(\sigma_0)\varepsilon_{-j_1}\cdots\varepsilon_{-j_1-\dots-j_s}\big)\big| \\
&\quad\le \left(\sum_{\substack{s\ge 1\\ j_1,\dots,j_s\ge 1\\ j_1+\dots+j_s<\ell}} b_{j_1}^2\cdots b_{j_s}^2 b_{\ell-j_1-\dots-j_s}^2\right)^{1/2} \\
&\qquad\times\left(\sum_{\substack{s\ge 1\\ j_1,\dots,j_s\ge 1\\ K(\delta)\le j_1+\dots+j_s}} E^2\big(f'(\sigma_0)\varepsilon_{-j_1}\cdots\varepsilon_{-j_1-\dots-j_s}\big)\right)^{1/2} \\
&\quad\le \sqrt{C}\,b_\ell\sqrt{\delta},
\end{aligned}$$

where C is the constant in (2.14). On the other hand, for every fixed $j_1,\dots,j_s$ with $j_1+\dots+j_s<K(\delta)$ we have

$$b_{\ell-j_1-\dots-j_s}\sim b_\ell \qquad \text{as } \ell\to\infty$$

and thus

$$\begin{aligned}
&E\big(f'(\sigma_0)\zeta_\ell^{(1)}\big) \\
&\quad\sim b_\ell E\left(\sum_{\substack{s\ge 1\\ j_1,\dots,j_s\ge 1\\ j_1+\dots+j_s<K(\delta)}} b_{j_1}\cdots b_{j_s} f'(\sigma_0)\varepsilon_{-j_1}\cdots\varepsilon_{-j_1-\dots-j_s}\right) \qquad \text{as } \ell\to\infty.
\end{aligned}$$

As $\delta\to 0$, the last expected value tends to

$$E\left(\sum_{\substack{s\ge 1\\ j_1,\dots,j_s\ge 1}} b_{j_1}\cdots b_{j_s} f'(\sigma_0)\varepsilon_{-j_1}\cdots\varepsilon_{-j_1-\dots-j_s}\right)=E\left(\frac{\sigma_0-a}{a}f'(\sigma_0)\right)$$

and thus

$$E\big(f'(\sigma_0)\zeta_\ell^*\big) \sim E\bigg(\frac{\sigma_0 - a}{a} f'(\sigma_0)\bigg) b_\ell \qquad \text{as } \ell \to \infty.$$

Since $\zeta_\ell = \zeta_\ell^* + b_\ell^*$, Lemma 2.3 follows. $\square$

Lemma 2.3 describes the asymptotic behavior of the first summand in (2.5) and it remains to estimate the remainder term $R_{n,\ell}$, which will be broken into several steps. We first give some moment estimates for $\zeta_{n,\ell}$ and the tail sums of σ_n in (1.6). Asymptotic estimates for the moments and product moments of the sequence (σ_n) were given in Giraitis, Robinson and Surgailis (2000) by using a diagram formalism. In our estimates we will not use this technique. Instead, we will use an induction argument combined with martingale inequalities, which will yield the desired results quite simply, without combinational difficulties. Whether our method is capable to give optimal constants [as the diagram technique in Giraitis, Robinson and Surgailis (2000) gives asymptotically precise estimates] is unclear.

LEMMA 2.4. *For any* $n, \ell \geq 1$ *we have*

$$E|\zeta_{n,\ell}|^p \leq C b_\ell^p \tag{2.20}$$

with some constant $C > 0$, *independent of* n, ℓ.

PROOF. We will use the fact that if $p > 1$ and $\{\xi_i, 1 \leq i \leq N\}$ is a martingale difference sequence with $E|\xi_i|^p \leq K$ $(1 \leq i \leq N)$, then for any real numbers $c_1, \dots, c_N$ we have

$$E\Bigg(\Bigg|\sum_{i=1}^N c_i \xi_i\Bigg|^p\Bigg) \leq A_p K \Bigg(\sum_{i=1}^N c_i^2\Bigg)^{p/2}, \tag{2.21}$$

where $A_p = (18p)^p (p/(p-1))^{p/2}$. Indeed, by Burkholder's square function inequality [see, e.g., Hall and Heyde (1980), page 23] the left-hand side of (2.21) is bounded by

$$A_p E\Bigg(\sum_{i=1}^N c_i^2 \xi_i^2\Bigg)^{p/2},$$

which, by Minkowski's inequality, cannot exceed

$$\begin{aligned} A_p \Bigg(\sum_{i=1}^N \|c_i^2 \xi_i^2\|_{p/2}\Bigg)^{p/2} &= A_p \Bigg(\sum_{i=1}^N c_i^2 \|\xi_i\|_p^2\Bigg)^{p/2} \\ &\leq A_p \Big(\max_{1 \leq i \leq N} \|\xi_i\|_p^p\Big) \Bigg(\sum_{i=1}^N c_i^2\Bigg)^{p/2} \leq A_p K \Bigg(\sum_{i=1}^N c_i^2\Bigg)^{p/2}. \end{aligned}$$

Let $1 \le s \le \ell$. Clearly the sum of terms in (2.6), where $j_1 = s$ is b_ℓ if $s = \ell$ and is

$$b_s \varepsilon_{n-s} \sum_{\substack{r \ge 2 \\ j_2, \dots, j_r \ge 1 \\ j_2 + \cdots + j_r = \ell - s}} b_{j_2} \cdots b_{j_r} \varepsilon_{n-s-j_2} \cdots \varepsilon_{n-s-j_2-\cdots-j_{r-1}} = b_s \varepsilon_{n-s} \zeta_{n-s,\ell-s}$$

if $1 \le s \le \ell - 1$. Thus

$$\zeta_{n,\ell} = b_\ell + \sum_{s=1}^{\ell-1} b_s \varepsilon_{n-s} \zeta_{n-s,\ell-s} = b_\ell + \sum_{s=1}^{\ell-1} b_s b_{\ell-s} \varepsilon_{n-s} \zeta^*_{n-s,\ell-s}, \tag{2.22}$$

where $\zeta^*_{n,\ell} = b_\ell^{-1} \zeta_{n,\ell}$. Noting that $\zeta_{n-s,\ell-s}$ contains only ε_ν's with $\nu < n - s$, it follows that

$$\{\varepsilon_{n-s} \zeta^*_{n-s,\ell-s}, s = \ell - 1, \ell - 2, \dots, 1\}$$

is a martingale difference sequence. Next we note that by (1.9) and the Remark after the proof of Lemma 2.2 we have

$$\sum_{i=1}^{\ell-1} b_i^2 b_{\ell-i}^2 \le (1 - \delta) \frac{p - 1}{324 p^3 \|\varepsilon_0\|_p^2} b_\ell^2 \tag{2.23}$$

for some $0 < \delta < 1$ and $\ell \ge \ell_0$. Observing that the distribution of $\zeta_{n,\ell}$ does not depend on n, one can find a constant $C \ge (1 - (1 - \delta)^{1/2})^{-p}$ such that $E|\zeta_{n,\ell}|^p \le C b_\ell^p$ holds for $1 \le \ell \le \ell_0$ and all n. We show by induction that (2.20) holds for all n, ℓ. Indeed, if $\ell > \ell_0$ and $E|\zeta_{n,j}|^p \le C b_j^p$ holds for $1 \le j \le \ell - 1$ and all n, then for $1 \le s \le \ell - 1$

$$E(|\varepsilon_{n-s} \zeta^*_{n-s,\ell-s}|^p) = E|\varepsilon_{n-s}|^p E|\zeta^*_{n-s,\ell-s}|^p \le C E|\varepsilon_0|^p \tag{2.24}$$

and thus using (2.21)–(2.24) and the Minkowski inequality we get

$$\begin{aligned} \|\zeta_{n,\ell}\|_p &\le b_\ell + \left\{ E\left(\left| \sum_{s=1}^{\ell-1} b_s b_{\ell-s} \varepsilon_{n-s} \zeta^*_{n-s,\ell-s} \right|^p \right) \right\}^{1/p} \\ &\le b_\ell + \left\{ A_p C (E|\varepsilon_0|^p) \left(\sum_{s=1}^{\ell-1} b_s^2 b_{\ell-s}^2 \right)^{p/2} \right\}^{1/p} \\ &\le b_\ell + (1 - \delta)^{1/2} C^{1/p} b_\ell \le C^{1/p} b_\ell \end{aligned}$$

by $C^{1/p} \ge (1 - (1 - \delta)^{1/2})^{-1}$, showing that $E|\zeta_{n,\ell}|^p \le C b_\ell^p$, completing the induction step and the proof of Lemma 2.4. □

We now introduce partial sums and tail sums of σ_n defined by

$$\tilde{\sigma}_n^{(\ell)} = a + a \sum_{k=1}^{\infty} \sum_{\substack{j_1, \dots, j_k \ge 1 \\ j_1 + \cdots + j_k < \ell}} b_{j_1} \cdots b_{j_k} \varepsilon_{n-j_1} \cdots \varepsilon_{n-j_1-\cdots-j_k}$$

and

$$\hat{\sigma}_n^{(\ell)} = a \sum_{k=1}^{\infty} \sum_{\substack{j_1, \dots, j_k \geq 1 \\ j_1 + \cdots + j_k \geq \ell}} b_{j_1} \cdots b_{j_k} \varepsilon_{n-j_1} \cdots \varepsilon_{n-j_1-\cdots-j_k}.$$

LEMMA 2.5. *We have*

$$E|\hat{\sigma}_n^{(\ell)}|^p \leq C_1 \left(\sum_{j=\ell}^{\infty} b_j^2 \right)^{p/2} \tag{2.25}$$

for some constant $C_1 > 0$, *independent of* n, ℓ. *In particular*, $E|\sigma_0|^p < +\infty$.

PROOF. Observe that the sum of terms in the infinite series in (1.6) containing $\varepsilon_{n-\ell}$, but no ε_ν with $\nu < n - \ell$ is

$$\sum_{\substack{k \geq 1 \\ j_1, \dots, j_k \geq 1 \\ j_1 + \cdots + j_k = \ell}} b_{j_1} \cdots b_{j_k} \varepsilon_{n-j_1} \cdots \varepsilon_{n-j_1-\cdots-j_{k-1}} \varepsilon_{n-\ell} = \varepsilon_{n-\ell} \zeta_{n,\ell}.$$

Thus

$$\hat{\sigma}_n^{(\ell)} = a \sum_{j=\ell}^{\infty} \varepsilon_{n-j} \zeta_{n,j}. \tag{2.26}$$

Recalling that $\zeta_{n,j}^* = b_j^{-1} \zeta_{n,j}$, the sequence

$$\{\varepsilon_{n-j} \zeta_{n,j}^*, \, j = \ell, \ell+1, \dots\}$$

is clearly a martingale difference sequence and, by Lemma 2.4, we have

$$E(|\varepsilon_{n-j} \zeta_{n,j}^*|^p) = E|\varepsilon_{n-j}|^p E|\zeta_{n,j}^*|^p \leq C E|\varepsilon_0|^p,$$

where C is the constant in Lemma 2.4. Thus by (2.21) we have for any $L > \ell$

$$E\left(\left| \sum_{j=\ell}^{L} \varepsilon_{n-j} \zeta_{n,j} \right|^p \right) = E\left(\left| \sum_{j=\ell}^{L} b_j \varepsilon_{n-j} \zeta_{n,j}^* \right|^p \right) \leq C^* \left(\sum_{j=\ell}^{L} b_j^2 \right)^{p/2},$$

where $C^* = A_p C E|\varepsilon_0|^p$. Letting $L \to \infty$ and using Fatou's lemma we get

$$E\left(\left| \sum_{j=\ell}^{\infty} \varepsilon_{n-j} \zeta_{n,j} \right|^p \right) \leq C^* \left(\sum_{j=\ell}^{\infty} b_j^2 \right)^{p/2}$$

which is identical with (2.25) in view of (2.26). □

LEMMA 2.6. *Let $r = 2p/(p-2)$. Then*

$$\big\| E\big(f'(\sigma_n)\zeta_{n,\ell}|\mathcal{F}_{n-\ell}\big) - E\big(f'(\sigma_n)\zeta_{n,\ell}\big)\big\|_r \le C_2 b_\ell \left(\sum_{j=\ell}^{\infty} b_j^2\right)^{1/2} \tag{2.27}$$

with some constant $C_2 > 0$.

PROOF. Clearly, the left-hand side of (2.27) is not greater than

$$\begin{aligned}&\big\| E\big(f'(\sigma_n)\zeta_{n,\ell}|\mathcal{F}_{n-\ell}\big) - E\big(f'(\tilde{\sigma}_n^{(\ell)})\zeta_{n,\ell}|\mathcal{F}_{n-\ell}\big)\big\|_r \\ &\qquad + \big\| E\big(f'(\tilde{\sigma}_n^{(\ell)})\zeta_{n,\ell}|\mathcal{F}_{n-\ell}\big) - E\big(f'(\sigma_n)\zeta_{n,\ell}\big)\big\|_r =: I_1 + I_2.\end{aligned}$$

Letting $\Delta = f'(\sigma_n) - f'(\tilde{\sigma}_n^{(\ell)})$, we get by the conditional Cauchy–Schwarz inequality

$$|E(\Delta\zeta_{n,\ell}|\mathcal{F}_{n-\ell})| \le E^{1/2}(\Delta^2|\mathcal{F}_{n-\ell})E^{1/2}(\zeta_{n,\ell}^2|\mathcal{F}_{n-\ell}) \le C_3 b_\ell E^{1/2}(\Delta^2|\mathcal{F}_{n-\ell}).$$

Here we used the fact that $\zeta_{n,\ell}$ is independent of $\mathcal{F}_{n-\ell}$ and thus by Lemma 2.4 and the monotonicity of the L_p norm in p we have

$$E(\zeta_{n,\ell}^2|\mathcal{F}_{n-\ell}) = E\zeta_{n,\ell}^2 \le (E|\zeta_{n,\ell}|^p)^{2/p} \le C_4 b_\ell^2.$$

Thus

$$\begin{aligned}I_1 &= \|E(\Delta\zeta_{n,\ell}|\mathcal{F}_{n-\ell})\|_r \le C_3 b_\ell \|E^{1/2}(\Delta^2|\mathcal{F}_{n-\ell})\|_r \\ &= C_3 b_\ell \|E(\Delta^2|\mathcal{F}_{n-\ell})\|_{r/2}^{1/2} \le C_3 b_\ell \|\Delta^2\|_{r/2}^{1/2} = C_3 b_\ell \|\Delta\|_r.\end{aligned} \tag{2.28}$$

Now by (1.10) and the mean value theorem we get

$$\begin{aligned}|\Delta| = |f''(\rho_n)||\sigma_n - \tilde{\sigma}_n^{(\ell)}| &\le C_5(|\rho_n|^\alpha + 1)|\sigma_n - \tilde{\sigma}_n^{(\ell)}| \\ &\le C_5(|\sigma_n|^\alpha + |\tilde{\sigma}_n^{(\ell)}|^\alpha + 1)|\sigma_n - \tilde{\sigma}_n^{(\ell)}|,\end{aligned}$$

where ρ_n lies between σ_n and $\tilde{\sigma}_n^{(\ell)}$. Lemma 2.5 implies that there is a constant C_6 such that $\|\sigma_n\|_p \le C_6$, $\|\sigma_n - \tilde{\sigma}_n^{(\ell)}\|_p \le C_6$ for all $n \ge 1$, $\ell \ge 1$ and thus using Hölder's inequality and Lemma 2.5 again we get

$$\begin{aligned}\|\Delta\|_r &\le C_5 \big\| |\sigma_n|^\alpha + |\tilde{\sigma}_n^{(\ell)}|^\alpha + 1\big\|_{rp/(p-r)} \|\sigma_n - \tilde{\sigma}_n^{(\ell)}\|_p \\ &\le C_7 \left(\sum_{i=\ell}^{\infty} b_i^2\right)^{1/2} \big(\big\||\sigma_n|^\alpha\big\|_{rp/(p-r)} + \big\||\tilde{\sigma}_n^{(\ell)}|^\alpha\big\|_{rp/(p-r)} + 1\big) \\ &\le C_7 \left(\sum_{i=\ell}^{\infty} b_i^2\right)^{1/2} \big(\|\sigma_n\|_{rp\alpha/(p-r)}^\alpha + \|\tilde{\sigma}_n^{(\ell)}\|_{rp\alpha/(p-r)}^\alpha + 1\big) \\ &\le C_8 \left(\sum_{i=\ell}^{\infty} b_i^2\right)^{1/2}\end{aligned} \tag{2.29}$$

since $rp\alpha/(p-r) \le p$ by (1.11). Relations (2.28) and (2.29) together yield

$$I_1 \le C_9 b_\ell \left(\sum_{i=\ell}^{\infty} b_i^2\right)^{1/2}. \tag{2.30}$$

On the other hand, $\tilde{\sigma}_n^{(\ell)}$ and $\zeta_{n,\ell}$ are independent of $\mathcal{F}_{n-\ell}$ and thus

$$I_2 = \left|E\left(f'(\tilde{\sigma}_n^{(\ell)})\zeta_{n,\ell}\right) - E\left(f'(\sigma_n)\zeta_{n,\ell}\right)\right| = |E(\Delta\zeta_{n,\ell})| \le \|\Delta\zeta_{n,\ell}\|_r.$$

Similarly to (2.28) we get $\|\Delta\zeta_{n,\ell}\|_r \le C_3 b_\ell \|\Delta\|_r$ and thus (2.29) yields

$$I_2 \le C_9 b_\ell \left(\sum_{i=\ell}^{\infty} b_i^2\right)^{1/2},$$

completing the proof of Lemma 2.6. □

We are now in a position to estimate the remainder term $R_{n,\ell}$ in (2.7). We prove the following.

LEMMA 2.7. *We have*

$$\|R_{n,\ell}\|_2 \le C_{10} \ell^{-(2\beta-1/2)}.$$

PROOF. (2.7) gives the decomposition

$$R_{n,\ell} = J_1 + J_2 + J_3 + J_4$$

where

$$\begin{aligned}
J_1 &= \left\{E\left(f'(\sigma_n)\zeta_{n,\ell}|\mathcal{F}_{n-\ell}\right) - E\left(f'(\sigma_n)\zeta_{n,\ell}\right)\right\}\sigma_{n-\ell}\varepsilon_{n-\ell},\\
J_2 &= c_p\theta\sigma_{n-\ell}^2(\varepsilon_{n-\ell}^2+1)E(|\sigma_n|^\alpha \zeta_{n,\ell}^2|\mathcal{F}_{n-\ell}),\\
J_3 &= c_p\theta\sigma_{n-\ell}^2(|\varepsilon_{n-\ell}|^{p/2}+\varepsilon_{n-\ell}^2+2)E\zeta_{n,\ell}^2,\\
J_4 &= c_p\theta\sigma_{n-\ell}^2 E(|\sigma_n|^{\alpha p/(p-4)}\zeta_{n,\ell}^2|\mathcal{F}_{n-\ell-1}).
\end{aligned}$$

We estimate J_1, J_2, J_3, J_4 separately. Since $\varepsilon_{n-\ell}$ and $\sigma_{n-\ell}$ are independent and $E|\varepsilon_0|^p < +\infty$, $E|\sigma_0|^p < +\infty$, we have

$$\|\sigma_{n-\ell}\varepsilon_{n-\ell}\|_p = \|\sigma_{n-\ell}\|_p \|\varepsilon_{n-\ell}\|_p \le C_{11}$$

and thus letting $r = 2p/(p-2)$, Lemma 2.6, Hölder's inequality and (1.7) give

$$\begin{aligned}
\|J_1\|_2 &\le \left\|E\left(f'(\sigma_n)\zeta_{n,\ell}|\mathcal{F}_{n-\ell}\right) - E\left(f'(\sigma_n)\zeta_{n,\ell}\right)\right\|_r \|\sigma_{n-\ell}\varepsilon_{n-\ell}\|_p\\
&\le C_{12} b_\ell \left(\sum_{j=\ell}^{\infty} b_j^2\right)^{1/2} \le C_{13}\ell^{-(2\beta-1/2)}.
\end{aligned}$$

On the other hand, $E\zeta_{n,\ell}^2 \le C_{14} b_\ell^2$ by Lemma 2.4 and the monotonicity of the L_p norm, and thus by the independence of $\varepsilon_{n-\ell}$ and $\sigma_{n-\ell}$ and (1.7) we have

$$\begin{aligned}\|J_3\|_2 &\le C_{15}\|\sigma_{n-\ell}^2\|_2\bigl(\bigl\||\varepsilon_{n-\ell}|^{p/2}\bigr\|_2 + \|\varepsilon_{n-\ell}^2\|_2 + 2\bigr)b_\ell^2 \\ &= C_{15}\|\sigma_{n-\ell}\|_4^2\bigl(\|\varepsilon_{n-\ell}\|_p^{p/2} + \|\varepsilon_{n-\ell}\|_4^2 + 2\bigr)b_\ell^2 \le C_{16} b_\ell^2 \le C_{17}\ell^{-2\beta}.\end{aligned}$$

To estimate J_2 we first use the conditional Hölder inequality and Lemma 2.4 to get

$$\begin{aligned}&\bigl|E(|\sigma_n|^\alpha \zeta_{n,\ell}^2|\mathcal{F}_{n-\ell})\bigr| \\ &\quad\le \{E(|\sigma_n|^{\alpha p/(p-2)}|\mathcal{F}_{n-\ell})\}^{(p-2)/p}\{E(|\zeta_{n,\ell}|^p|\mathcal{F}_{n-\ell})\}^{2/p} \\ &\quad\le C_{18} b_\ell^2\{E(|\sigma_n|^{\alpha p/(p-2)}|\mathcal{F}_{n-\ell})\}^{(p-2)/p}\end{aligned}$$

since $\zeta_{n,\ell}$ is independent of $\mathcal{F}_{n-\ell}$. Thus by the Hölder inequality and the independence of $\sigma_{n-\ell}$ and $\varepsilon_{n-\ell}$ we have

$$\begin{aligned}\|J_2\|_2 &\le C_{19}\bigl\|\sigma_{n-\ell}^2(\varepsilon_{n-\ell}^2+1)\bigr\|_{p/2}\bigl\|E(|\sigma_n|^\alpha\zeta_{n,\ell}^2|\mathcal{F}_{n-\ell})\bigr\|_{2p/(p-4)} \\ &\le C_{20}\|\sigma_{n-\ell}^2\|_{p/2}\|\varepsilon_{n-\ell}^2+1\|_{p/2} b_\ell^2\bigl\|E^{(p-2)/p}(|\sigma_n|^{\alpha p/(p-2)}|\mathcal{F}_{n-\ell})\bigr\|_{2p/(p-4)} \\ &= C_{20}\|\sigma_{n-\ell}\|_p^2\|\varepsilon_{n-\ell}^2+1\|_{p/2} b_\ell^2\bigl\|E(|\sigma_n|^{\alpha p/(p-2)}|\mathcal{F}_{n-\ell})\bigr\|_{(2p-4)/(p-4)}^{(p-2)/p} \\ &\le C_{20}\|\sigma_{n-\ell}\|_p^2(\|\varepsilon_{n-\ell}\|_p^2+1) b_\ell^2\bigl\||\sigma_n|^{\alpha p/(p-2)}\bigr\|_{(2p-4)/(p-4)}^{(p-2)/p} \\ &\le C_{21} b_\ell^2\|\sigma_n\|_{2\alpha p/(p-4)}^\alpha \le C_{22} b_\ell^2 \le C_{23}\ell^{-2\beta}\end{aligned}$$

since $2\alpha p/(p-4) \le p$ by (1.11). Finally, the estimate of J_4 is the same as that of J_2, just α should be replaced by $\alpha p/(p-4)$ in all steps and we get

$$\|J_4\|_2 \le C_{24} b_\ell^2\|\sigma_n\|_{2\alpha p^2/(p-4)^2}^{\alpha p/(p-4)} \le C_{25}\ell^{-2\beta}$$

since $2\alpha p^2/(p-4)^2 \le p$ by (1.11). Collecting the estimates for J_1, J_2, J_3, J_4 we get Lemma 2.7. □

The following lemma is a variant of Lemma 6.4 in Ho and Hsing (1996).

LEMMA 2.8. *We have*

$$E(R_{n,\ell}R_{n',\ell'}) = 0 \qquad \textit{if } n-\ell \ne n'-\ell'.$$

PROOF. Since $R_{n,\ell} = X_{n,\ell} - B_\ell\sigma_{n-\ell}\varepsilon_{n-\ell}$ by (2.5), (1.24) and stationarity, it suffices to show that for $n-\ell \ne n'-\ell'$ we have

$$E(X_{n,\ell}X_{n',\ell'}) = 0, \tag{2.31}$$

$$E(\sigma_{n-\ell}\varepsilon_{n-\ell}X_{n',\ell'}) = 0, \tag{2.32}$$

$$E(\sigma_{n-\ell}\varepsilon_{n-\ell}\sigma_{n'-\ell'}\varepsilon_{n'-\ell'}) = 0. \tag{2.33}$$

Assume $n' - \ell' < n - \ell$. Then $X_{n',\ell'}$ is $\mathcal{F}_{n-\ell-1}$ measurable and thus the conditional expectation of $X_{n,\ell}X_{n',\ell'}$ with respect to $\mathcal{F}_{n-\ell-1}$ is

$$X_{n',\ell'}E(X_{n,\ell}|\mathcal{F}_{n-\ell-1}) \\ = X_{n',\ell}\big[E\big(f(\sigma_n)|\mathcal{F}_{n-\ell-1}\big) - E\big(f(\sigma_n)|\mathcal{F}_{n-\ell-1}\big)\big] = 0.$$

On the other hand, $\sigma_{n-\ell}$ is $\mathcal{F}_{n-\ell-1}$ measurable by (1.6), and thus the conditional expectation of $\sigma_{n-\ell}\varepsilon_{n-\ell}X_{n',\ell'}$ with respect to $\mathcal{F}_{n-\ell-1}$ is

$$\sigma_{n-\ell}X_{n',\ell'}E(\varepsilon_{n-\ell}|\mathcal{F}_{n-\ell-1}) = 0.$$

Finally, the conditional expectation of $\sigma_{n-\ell}\varepsilon_{n-\ell}\sigma_{n'-\ell'}\varepsilon_{n'-\ell'}$ with respect to $\mathcal{F}_{n-\ell-1}$ is

$$\sigma_{n-\ell}\sigma_{n'-\ell'}\varepsilon_{n'-\ell'}E(\varepsilon_{n-\ell}|\mathcal{F}_{n-\ell-1}) = 0.$$

Thus (2.31)–(2.33) are valid. □

LEMMA 2.9. *We have*

$$E\left(\sum_{n=1}^{N}\sum_{\ell=1}^{\infty} R_{n,\ell}\right)^2 = O(N^{3-2\beta-\varepsilon}) \tag{2.34}$$

for some $\varepsilon > 0$.

PROOF. In view of Lemma 2.8, relation (2.34) is equivalent to

$$\sum_{\substack{1 \le n, n' \le N \\ \ell, \ell' \ge 1 \\ n-\ell = n'-\ell'}} E(R_{n,\ell}R_{n',\ell'}) = O(N^{3-2\beta-\varepsilon}). \tag{2.35}$$

By Lemma 2.7 and the Cauchy–Schwarz inequality we have

$$|E(R_{n,\ell}R_{n',\ell'})| \le \|R_{n,\ell}\|_2\|R_{n',\ell'}\|_2 \le C_{26}(\ell\ell')^{-(2\beta-1/2)}.$$

Thus to prove (2.35) it suffices to show that

$$\sum_{\substack{1 \le n, n' \le N \\ \ell, \ell' \ge 1 \\ n-\ell = n'-\ell'}} (\ell\ell')^{-(2\beta-1/2)} = O(N^{3-2\beta-\varepsilon}). \tag{2.36}$$

We note that, as proved in Lemma 6.5 of Ho and Hsing (1996), we have, for any integer $m \ge 1$,

$$\sum_{j=1}^{\infty} \frac{1}{(j(m+j))^{\alpha}} \le \begin{cases} Cm^{-2\alpha+1}, & \text{if } \frac{1}{2} < \alpha < 1, \\ C\dfrac{\log m}{m}, & \text{if } \alpha = 1, \\ Cm^{-\alpha}, & \text{if } \alpha > 1. \end{cases} \tag{2.37}$$

Fix $m \in \mathbf{Z}$ and add those terms in the sum in (2.36) where $n' - n = m$. Then automatically $\ell' - \ell = m$, that is, $\ell\ell' = \ell(m+\ell)$. Clearly n' can take at most N values and once it is fixed, n is uniquely determined. Thus the sum of the considered terms in (2.36) is not greater than

$$N \sum_{\ell=1}^{\infty} \frac{1}{(\ell(m+\ell))^{2\beta-1/2}}$$

and the total sum in (2.36) cannot exceed

$$N \sum_{|m|\le N} \sum_{\ell=1}^{\infty} \frac{1}{(\ell(m+\ell))^{2\beta-1/2}}. \tag{2.38}$$

Here the contribution of the terms with $m = 0$ is not greater than $N \times \sum_{\ell=1}^{\infty} \ell^{-(4\beta-1)} = O(N)$ by $4\beta - 1 > 1$, which is smaller than the remainder term in (2.36) if ε is small enough. Hence it suffices to consider the terms with $m \neq 0$ and for reasons of symmetry we may assume $m > 0$. Note that the exponent $2\beta - 1/2$ in (2.38) lies in $(1/2, 3/2)$. By (2.37) the inner sum in (2.38) is $O(m^{-(4\beta-2)})$, $O(\log m/m)$ and $O(m^{-(2\beta-1/2)})$ according as $\beta < 3/4$, $\beta = 3/4$ or $\beta > 3/4$, respectively. Thus the expression in (2.38) is at most

$$\begin{aligned} &N \sum_{m=1}^{N} m^{-(4\beta-2)} = O(N^{4-4\beta}) && \text{if } \frac{1}{2} < \beta < \frac{3}{4},\\ &N \sum_{m=1}^{N} \frac{\log m}{m} = O(N \log^2 N) && \text{if } \beta = \frac{3}{4},\\ &N \sum_{m=1}^{N} m^{-(2\beta-1/2)} = O(N) && \text{if } \frac{3}{4} < \beta < 1. \end{aligned} \tag{2.39}$$

A simple calculation shows that all remainder terms in (2.39) are $O(N^{3-2\beta-\varepsilon})$ if ε is small enough and thus (2.36) is proved. □

PROOF OF THEOREMS 1.1 AND 1.2. Relation (1.26) of Theorem 1.2 is immediate from (2.8) and Lemma 2.9. Letting

$$R_N = \sum_{n=1}^{N} \big(f(\sigma_n) - Ef(\sigma_n)\big) - \sum_{n=1}^{N} \sigma_n^{(f)},$$

relation (1.26) and stationarity imply for any $N \ge 1$ and any $0 \le t_1 \le t \le t_2 \le 1$,

$$\begin{aligned} &E(|R_{[Nt]} - R_{[Nt_1]}||R_{[Nt_2]} - R_{[Nt]}|)\\ &\quad \le \|R_{[Nt]} - R_{[Nt_1]}\|_2 \|R_{[Nt_2]} - R_{[Nt]}\|_2\\ &\quad = \|R_{[Nt]-[Nt_1]}\|_2 \|R_{[Nt_2]-[Nt]}\|_2\\ &\quad \le C^2([Nt] - [Nt_1])^{3/2-\beta}([Nt_2] - [Nt])^{3/2-\beta}. \end{aligned} \tag{2.40}$$

If $t_2 - t_1 < 1/N$, then either $[Nt] = [Nt_1]$ or $[Nt] = [Nt_2]$ and thus the last expression in (2.40) is 0; if $t_2 - t_1 \geq 1/N$, then the last expression in (2.40) is at most

$$C^2([Nt_2] - [Nt_1])^{3-2\beta} \leq C^2\big(N(t_2 - t_1) + 1\big)^{3-2\beta} \leq 4C^2\big(N(t_2 - t_1)\big)^{3-2\beta}.$$

Thus we showed that the process

$$X_N(t) = N^{-(3/2-\beta)} R_{[Nt]}, \qquad 0 \leq t \leq 1,\ N = 1, 2, \dots$$

satisfies

$$E\big(|X_N(t) - X_N(t_1)||X_N(t_2) - X_N(t)|\big) \leq 4C^2(t_2 - t_1)^{3-2\beta},$$

and consequently we have for any $\lambda > 0$

$$P\big(|X_N(t) - X_N(t_1)| \geq \lambda, |X_N(t_2) - X_N(t)| \geq \lambda\big) \leq 4C^2 \frac{1}{\lambda^2}(t_2 - t_1)^{3-2\beta}.$$

The last relation implies by $3 - 2\beta > 1$, and Theorem 15.6 of Billingsley (1968) and its proof, that the sequence $\{X_N(t),\ N = 1, 2, \dots\}$ is tight in $\mathcal{D}[0, 1]$. By (1.30), the sequence

$$N^{-(3/2-\beta)} \sum_{n=1}^{[Nt]} \sigma_n^{(f)}, \qquad 0 \leq t \leq 1,\ N = 1, 2, \dots$$

is also tight and thus we can conclude the tightness of the processes in (1.12). Finally, the convergence of the finite dimensional distributions in (1.12) follows from (1.26), (1.30).

To prove relation (1.27) of Theorem 1.2 we use the decomposition, similar to (2.4),

$$f(y_n) - Ef(y_n) = \sum_{\ell=0}^{\infty} Y_{n,\ell}, \tag{2.41}$$

where

$$Y_{n,\ell} = E\big(f(y_n)|\mathcal{F}_{n-\ell}\big) - E\big(f(y_n)|\mathcal{F}_{n-\ell-1}\big).$$

Note that the summation in (2.41) starts with $\ell = 0$, but the contribution $\sum_{n=1}^{N} Y_{n,0}$ of the $Y_{n,0}$'s in the sum $\sum_{n=1}^{N}(f(y_n) - Ef(y_n))$ is

$$\sum_{n=1}^{N}\big(f(y_n) - E\big(f(y_n)|\mathcal{F}_{n-1}\big)\big) = O_P(N^{1/2}) = O_P(N^{3/2-\beta-\varepsilon})$$

if ε is small enough, since $\{f(y_n) - E(f(y_n)|\mathcal{F}_{n-1}), n \geq 1\}$ is a square integrable martingale difference sequence [the finiteness of $Ef^2(y_n)$ follows from the fact that $|f(x)| \leq C'|x|^{\alpha+2}$ for sufficiently large x by (1.10) and $2\alpha + 4 \leq p$ by (1.11)] and hence it is orthogonal. For $\ell \geq 1$, $Y_{n,\ell}$ satisfies the following approximation formula, analogous to Lemma 2.1. □

LEMMA 2.10. *Under the conditions of Theorem* 1.1 *we have, for* $\ell \geq 1$,

$$Y_{n,\ell} = E\big(f'(y_n)\varepsilon_n \zeta_{n,\ell}\big)\sigma_{n-\ell}\varepsilon_{n-\ell} + \bar{R}_{n,\ell}, \tag{2.42}$$

where

$$\begin{aligned}\bar{R}_{n,\ell} = {} & \{E\big(f'(y_n)\varepsilon_n\zeta_{n,\ell}|\mathcal{F}_{n-\ell}\big) - E\big(f'(y_n)\varepsilon_n\zeta_{n,\ell}\big)\}\sigma_{n-\ell}\varepsilon_{n-\ell} \\ & + c^*\theta\sigma_{n-\ell}^2\{(\varepsilon_{n-\ell}^2 + 1)E(|\sigma_n|^\alpha \varepsilon_n^2\zeta_{n,\ell}^2|\mathcal{F}_{n-\ell}) \\ & \qquad + (|\varepsilon_{n-\ell}|^{p/2} + \varepsilon_{n-\ell}^2 + 2)E(\varepsilon_n^2\zeta_{n,\ell}^2) \\ & \qquad + E(|\sigma_n|^{\alpha p/(p-4)}\varepsilon_n^2\zeta_{n,\ell}^2|\mathcal{F}_{n-\ell-1})\},\end{aligned} \tag{2.43}$$

where $|\theta| \leq 1$ *and* c^* *is a positive constant depending on* p *and the sequence* (ε_n).

Note that the terms in (2.42) and (2.43) are the same as in (2.5) and (2.7), just $f'(\sigma_n)$ is replaced by $f'(y_n)$ and $\zeta_{n,\ell}$ is replaced by $\varepsilon_n\zeta_{n,\ell}$. The proof of (2.42)–(2.43) follows from that of Lemma 2.1. Since $y_n = \varepsilon_n\sigma_n = \varepsilon_n\psi(\varepsilon_{n-1}, \dots)$, formula (2.11) gets replaced by

$$\begin{aligned}Y_{n,\ell} = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} & \big[f\big(u_0\psi(u_1, \dots, u_{\ell-1}, \varepsilon_{n-\ell}, \dots)\big) \\ & - f\big(u_0\psi(u_1, \dots, u_{\ell-1}, v, \varepsilon_{n-\ell-1}, \dots)\big)\big] \\ & \times dG(u_0)\,dG(u_1)\cdots dG(u_{\ell-1})\,dG(v)\end{aligned} \tag{2.44}$$

and thus (2.12) becomes

$$\begin{aligned} & f'\big(u_0\psi(u_1, \dots, u_{\ell-1}, \varepsilon_{n-\ell}, \dots)\big) \\ & \quad \times[\psi(u_1, \dots, u_{\ell-1}, \varepsilon_{n-\ell}, \dots) - \psi(u_1, \dots, u_{\ell-1}, v, \varepsilon_{n-\ell-1}, \dots)]u_0 \\ & \quad + \tfrac{1}{2} f''(\tau^*)[\psi(u_1, \dots, u_{\ell-1}, \varepsilon_{n-\ell}, \dots) \\ & \qquad\qquad - \psi(u_1, \dots, u_{\ell-1}, v, \varepsilon_{n-\ell-1}, \dots)]^2 u_0^2.\end{aligned} \tag{2.45}$$

Using (2.44) and (2.45) instead of (2.11) and (2.12), the proof of Lemma 2.1 yields (2.42) and (2.43) with obvious changes.

Similarly to (2.18), we have also

$$\bar{B}_\ell \sim \gamma_1 b_\ell \qquad \text{as } \ell \to \infty$$

(the proof is the same) and Lemma 2.6 remains valid with σ_n replaced by y_n and $\zeta_{n,\ell}$ replaced by $\varepsilon_n\zeta_{n,\ell}$; the proof is again similar, with $\tilde{\sigma}_n^{(\ell)}$ replaced by $\tilde{y}_n^{(\ell)} = \varepsilon_n\tilde{\sigma}_n^{(\ell)}$. The remaining changes in the argument leading to (1.26) are obvious and we get (1.27). The implication (1.27) $\Rightarrow$ (1.13) can be proved in the same way as (1.26) $\Rightarrow$ (1.12). □

Acknowledgment. The authors are indebted to the referee for his valuable remarks which led to an improvement of the results of the paper.

REFERENCES

AVRAM, F. and TAQQU, M. (1987). Noncentral limit theorems and Appell polynomials. *Ann. Probab.* **15** 767–775.

BERKES, I. and HORVÁTH, L. (2001). Strong approximation of the empirical process of GARCH sequences. *Ann. Appl. Probab.* **11** 789–809.

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

BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31** 307–327.

DEHLING, H. and TAQQU, M. (1989). The empirical process of some long-range dependent sequences with an application to U-statistics. *Ann. Statist.* **17** 1767–1783.

DOBRUSHIN, R. L. and MAJOR, P. (1979). Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z. Wahrsch. Verw. Gebiete* **50** 27–52.

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

GIRAITIS, L., KOKOSZKA, P. and LEIPUS, R. (2000). Stationary ARCH models: Dependence structure and central limit theorem. *Econom. Theory* **16** 3–22.

GIRAITIS, L., ROBINSON, P. and SURGAILIS, D. (2000). A model for long memory conditional heteroskedasticity. *Ann. Appl. Probab.* **10** 1002–1024.

GIRAITIS, L. and SURGAILIS, D. (1986). Multivariate Appell polynomials and the central limit theorem. In *Dependence in Probability and Statistics* (E. Eberlein and M. S. Taqqu, eds.) 21–71. Birkhäuser, Boston.

GIRAITIS, L. and SURGAILIS, D. (1989). A limit theorem for polynomials of a linear process with long-range dependence. *Lithuanian Math. J.* **29** 128–145.

GIRAITIS, L. and SURGAILIS, D. (1999). Central limit theorem for the empirical process of a linear sequence with long memory. *J. Statist. Plann. Inference* **80** 81–93.

HALL, P. and HEYDE, C. (1980). *Martingale Limit Theory and Its Application.* Academic Press, New York.

HO, H. C. and HSING, T. (1996). On the asymptotic expansion of the empirical process of long-memory moving averages. *Ann. Statist.* **24** 992–1024.

LOÈVE, M. (1977). *Probability Theory* **1**, 4th ed. Springer, New York.

SAMORODNITSKY, G. and TAQQU, M. (1994). *Stable Non-Gaussian Random Processes.* Chapman and Hall, New York.

SURGAILIS, D. (1982). Zones of attraction of self-similar multiple integrals. *Lithuanian Math. J.* **22** 327–340.

TAQQU, M. (1975). Weak convergence to fractional Brownian motion and the Rosenblatt process. *Z. Wahrsch. Verw. Gebiete* **31** 287–302.

TAQQU, M. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch. Verw. Gebiete* **50** 53–83.

A. RÉNYI INSTITUTE OF MATHEMATICS
HUNGARIAN ACADEMY OF SCIENCES
P.O. BOX 127
H-1364 BUDAPEST
HUNGARY
E-MAIL: berkes@renyi.hu

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
155 SOUTH 1440 EAST
SALT LAKE CITY, UTAH 84112-0090
E-MAIL: horvath@math.utah.edu