Adv. Appl. Prob. **40,** 1198–1222 (2008)
Printed in Northern Ireland
© *Applied Probability Trust* 2008

A QUADRATIC ARCH(∞) MODEL WITH LONG MEMORY AND LÉVY STABLE BEHAVIOR OF SQUARES

DONATAS SURGAILIS,* *Institute of Mathematics and Informatics, Vilnius*

Abstract

We introduce a modification of the linear ARCH (LARCH) model (Giraitis, Robinson, and Surgailis (2000))—a special case of Sentana's (1995) quadratic ARCH (QARCH) model—for which the conditional variance is a sum of a positive constant and the square of an inhomogeneous linear combination of past observations. Necessary and sufficient conditions for the existence of a stationary solution with finite variance are obtained. We give conditions under which the stationary solution with infinite fourth moment can exhibit long memory, the leverage effect, and a Lévy-stable limit behavior of partial sums of squares.

Keywords: Linear ARCH model; quadratic ARCH model; long memory; partial sum; Lévy stable process; fractional Brownian motion

2000 Mathematics Subject Classification: Primary 62M10
Secondary 60G18; 60G52

1. Introduction

Sentana (1995) introduced a wide class of quadratic ARCH (QARCH) models $(X_t) = (X_t,\ t \in \mathbb{Z})$ with conditional variance $\sigma_t^2 = \mathrm{var}[X_t \mid X_s,\ s < t]$ given by a general quadratic form in the past p values of X_t:

$$\sigma_t^2 = \theta + \sum_{i=1}^{p} \psi_i X_{t-i} + \sum_{i,j=1}^{p} a_{ij} X_{t-i} X_{t-j}. \tag{1.1}$$

Sentana (1995) obtained conditions on the coefficients θ, ψ_i, and a_{ij} of (1.1) which ensure nonnegativity of the conditional variance and the existence of a covariance stationary solution to generalized QARCH(p, q) (GQARCH(p, q)) equations.

Although limited to short memory processes, Sentana's model nests many important ARCH models (see the discussion in Sentana (1995)), in particular, the ARCH(p) process corresponding to $\psi_i = 0$ and $a_{ij} = 0$ $(i \neq j)$. It also encompasses short memory linear ARCH (LARCH) models introduced in Robinson (1991) and studied in Giraitis *et al.* (2000), (2004), Giraitis and Surgailis (2002), Berkes and Horváth (2003), Doukhan *et al.* (2006), Beran (2006), and other papers.

The (general) LARCH(∞) model corresponds to

$$m_t = 0, \qquad \sigma_t^2 = \left(a + \sum_{j=1}^{\infty} a_j X_{t-j}\right)^2, \tag{1.2}$$

Received 27 September 2007; revision received 9 November 2008.
* Postal address: Institute of Mathematics and Informatics, Akademijos 4, 08663 Vilnius, Lithuania.
Email address: sdonatas@ktl.mii.lt

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

where $m_t = \mathrm{E}[X_t \mid X_s, s < t]$ is the conditional mean, and a and a_j ($j \geq 1$) are real coefficients with $\sum_{j=1}^{\infty} a_j^2 < 1$. The last condition allows for nonsummable (hyperbolically decaying) coefficients (a_j) and covariance long memory of squares (X_t^2), and other functionals of the LARCH model (see Giraitis *et al.* (2000), Berkes and Horváth (2003)). The LARCH model in (1.2) was further generalized in Giraitis and Surgailis (2002) to include a linear drift $m_t = b + \sum_{j=1}^{\infty} b_j X_{t-j}$ and the conditional variance σ_t^2 as in (1.2). It is defined as a stationary solution to the bilinear equation

$$X_t = \zeta_t \left(a + \sum_{j=1}^{\infty} a_j X_{t-j} \right) + b + \sum_{j=1}^{\infty} b_j X_{t-j}, \tag{1.3}$$

where (ζ_t) is a standard independent and identically distributed (i.i.d.) sequence.

The present paper generalizes the LARCH and the bilinear models to the case of strictly positive conditional variance (volatility):

$$m_t = b + \sum_{j=1}^{\infty} b_j X_{t-j}, \qquad \sigma_t^2 = \nu^2 + \left(a + \sum_{j=1}^{\infty} a_j X_{t-j} \right)^2, \tag{1.4}$$

where $\nu > 0$ and $b, b_j, a, a_j \in \mathbb{R}$ are real parameters. Following the idea in Sentana (1995, p. 658), we define the corresponding process (X_t) as a stationary causal solution to the bilinear equation

$$X_t = \kappa \eta_t + \zeta_t \sum_{j=1}^{\infty} a_j X_{t-j} + b + \sum_{j=1}^{\infty} b_j X_{t-j}, \tag{1.5}$$

where (η_s, ζ_s) are i.i.d. random vectors, with values in $\mathbb{R}^2$, with

$$\mathrm{E}[\eta_s] = \mathrm{E}[\zeta_s] = 0, \qquad \mathrm{E}[\eta_s^2] = \mathrm{E}[\zeta_s^2] = 1, \qquad \rho = \mathrm{E}[\eta_s \zeta_s], \tag{1.6}$$

and the parameters $\rho \in [-1, 1]$ and $\kappa > 0$ are related to the parameters $a \in \mathbb{R}$ and $\nu \geq 0$ in (1.4) by

$$\kappa\rho = a, \qquad \kappa^2 = a^2 + \nu^2. \tag{1.7}$$

Note that $\nu = \sqrt{\kappa^2(1-\rho^2)} > 0$ is equivalent to $|\rho| < 1$ and $\kappa \neq 0$, while the bilinear model in (1.3) corresponds to $|\rho| = 1$, or to completely correlated components (η_s, ζ_s). For obvious reasons, we refer to (η_t) as the 'homoscedastic noise' and to (ζ_t) as the 'heteroscedastic noise' of the bilinear equation, (1.5).

The bilinear model in (1.5) shares many properties of (1.3) and the LARCH model; however, it may exhibit some new important features (see below). In the present paper we focus on the zero-drift case, $m_t = 0$:

$$X_t = \kappa \eta_t + \zeta_t \sum_{j=1}^{\infty} a_j X_{t-j}, \tag{1.8}$$

to which the general case of (1.5) with $b = 0$ can be reduced (see Section 2). Assuming that conditions (1.6) are satisfied and that $\|a\|^2 = \sum_{j=1}^{\infty} a_j^2 < 1$, we show that there exists a unique stationary causal solution (X_t) of (1.8) given by convergent Volterra series which admits a stochastic volatility type representation $X_t = \sigma_t \varepsilon_t$ with martingale difference innovations $\varepsilon_t = X_t/\sigma_t$ (see Section 3). An important distinction of (1.4) from the LARCH model in (1.2) is the fact that σ_t^2 in (1.4) is almost surely separated from 0: $\sigma_t^2 \geq \nu^2 > 0$, similarly as in

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

the classical GARCH and ARCH(∞) models. To stress the importance of the last fact, we term (1.8) the *LARCH*$_+(\infty)$ *model.* The fact that in the original LARCH model the (squared) volatility process may vanish is a drawback of this model (see Giraitis *et al.* (2000), (2004) for a discussion) and also seems to contradict empirical data. Strict positivity of volatility clearly is very important for (quasi-)maximum likelihood estimation of parameters of the model (see the likelihood in (2.12), below).

Since the stationary solution (X_t) of (1.8) is uncorrelated, (covariance) long memory can be observed for the squared process (X_t^2), under some additional assumptions which include a regular decay of coefficients

$$a_t \sim c_0 t^{d-1} \quad \left(t \to \infty, \text{there exist } c_0 > 0 \text{ and } 0 < d < \tfrac{1}{2}\right) \tag{1.9}$$

and the existence of the finite fourth moment of noise variables, similarly as in the LARCH model. The present paper focuses on the case when $\mathrm{E}[X_t^4] = \infty$ and $\mathrm{E}[|X_t|^3] < \infty$. In the latter case, long memory of the LARCH$_+(\infty)$ model with hyperbolically decaying coefficients as in (1.9) is demonstrated as the corresponding decay of the 'leverage function' $H_t := \mathrm{E}[\sigma_t^2 X_0] = \mathrm{E}[X_t^2 X_0]$ as $t \to \infty$ (see Corollary 3.1(i)). The *leverage effect* (i.e. the fact that past returns and future volatilities are negatively correlated) is established for the LARCH$_+(\infty)$ model under similar assumptions on the sign of the coefficients as for the LARCH model (see Corollary 3.1(ii) and Giraitis *et al.* (2004, Theorem 4)).

The main results of the present paper refer to the limit behavior of partial sums of the squared process (X_t^2) in (1.8). Assume, in addition to (1.6) and (1.9), that

$$\mathrm{E}[\zeta^4 + \zeta^2\eta^2] < \infty, \qquad \mathrm{P}[\eta^2 > x] \sim c_\eta x^{-\alpha} \quad (x \to \infty, \text{there exist } 1 < \alpha < 2 \text{ and } c_\eta > 0). \tag{1.10}$$

The second condition in (1.10) implies that the i.i.d. random variables (η_t^2) belong to the domain of attraction of an α-stable law. It turns out that, under assumptions (1.6), (1.9), and (1.10), and some additional assumption on $\|a\|^2$, there exists a dichotomy in the limit distribution of partial sums: depending on whether $1/\alpha > d + \frac{1}{2}$ or $1/\alpha < d + \frac{1}{2}$ holds, partial sums of squares (X_t^2) of the conditionally heteroscedastic process in (1.8) converge to an α-stable Lévy motion, or to a fractional Brownian motion with $H = d + \frac{1}{2}$ (Corollary 4.1). For linear processes with long memory, a similar dichotomy was proved in Vaičiulis (2003).

Modeling financial data with the help of ARCH models is now well accepted in practice, and the literature on such models is vast. Probably the two most widely discussed empirical facts about financial data is the long memory of absolute returns and their squares, and the heavy tailedness of their marginal distribution. See Mikosch (2003) and the review paper Giraitis *et al.* (2008). The LARCH$_+(\infty)$ model can exhibit these two empirical facts and, therefore, might present an interest to financial modeling.

One of the motivations for our study was the problem of the limit distribution of the empirical covariances of long-memory and heavy-tailed ARCH processes and the squared processes in particular. The question of the limiting distribution of empirical covariances and correlations of heavy-tailed moving averages, ARCH, and other processes with short memory was studied in Davis and Resnick (1985a), (1985b), (1986), (1996), and Davis and Mikosch (1998), (2001). The results of the present paper can be easily extended to sample autocorrelations of the LARCH$_+(\infty)$ model (X_t) in Corollary 4.1. However, the more interesting (and more difficult) part of the above problem concerns the limit behavior of sample autocorrelations of the squared process (X_t^2). We plan to study this problem in a subsequent paper.

Several interesting questions pertaining to the model in (1.8) and the results of this paper remain open (some of them were raised by the referees). They concern the functional conver-

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

gence for partial sums (see also Remark 4.1), the tail behavior of X_t under condition (1.10), the possibility of extending (1.8) to infinite variance innovations, and other issues.

2. Stationary solution

In this section we obtain a (covariance) stationary solution of (1.5). The discussion follows Giraitis and Surgailis (2002), where the case $|\rho| = 1$, or (1.3) was considered. We also restrict ourselves to the zero-mean situation, $b = 0$, in (1.5) since the case in which $b \neq 0$ is different and most likely cannot lead to long memory. See Giraitis and Surgailis (2002).

We first formally derive a stationary solution by reducing (1.5) to (1.8), which in turn can be solved by iteration. A rigorous statement will follow immediately after this derivation. To this end, define

$$Y_t := \kappa\eta_t + \zeta_t A_t, \qquad A_t := \sum_{j=1}^{\infty} a_j X_{t-j}, \qquad m_t := \sum_{j=1}^{\infty} b_j X_{t-j}. \tag{2.1}$$

Then (1.5) reduces to the linear equation

$$X_t - \sum_{j=1}^{\infty} b_j X_{t-j} = Y_t,$$

which can be formally solved by inverting it, i.e.

$$X_t = \sum_{j=0}^{\infty} g_j Y_{t-j}, \tag{2.2}$$

where we use the notation

$$A(z) := \sum_{j=1}^{\infty} a_j z^j, \qquad B(z) := \sum_{j=1}^{\infty} b_j z^j, \qquad G(z) := (1 - B(z))^{-1} = \sum_{j=0}^{\infty} g_j z^j,$$

$$F(z) := \frac{A(z)}{1 - B(z)} = \sum_{j=1}^{\infty} (a \star g)_j z^j = \sum_{j=1}^{\infty} f_j z^j, \tag{2.3}$$

for the generating series, where $(a \star g)_j := \sum_{k=0}^{j} g_k a_{j-k}$ is the convolution. From (2.1) and (2.2), we obtain

$$A_t = \sum_{j=1}^{\infty} a_j X_{t-j} = \sum_{j=1}^{\infty} (a \star g)_j Y_{t-j},$$

$$Y_t = \kappa\eta_t + \zeta_t \sum_{j=1}^{\infty} (a \star g)_j Y_{t-j}, \tag{2.4}$$

$$= \kappa\left(\eta_t + \zeta_t \sum_{u<t} \sum_{k=0}^{\infty} \sum_{u<s_k<\cdots<s_1<t} (a \star g)_{t-s_1} \cdots (a \star g)_{s_{k-1}-s_k} \zeta_{s_1} \cdots \zeta_{s_k} \eta_u\right), \tag{2.5}$$

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

where we set $a_j := 0$, $b_j := 0$ ($j \le 0$), and $g_j := 0$ ($j < 0$). Note that (2.5) follows from (2.4) by iteration. From (2.5) and (2.2), we finally obtain a (formal) solution of the bilinear equation (1.8) in terms of multiple series of noise variables:

$$X_t = \kappa\Bigg(\sum_{j=0}^{\infty} g_j \eta_{t-j} + \sum_{j\ge 1}\sum_{u<t-j}\sum_{k=0}^{\infty}\sum_{u<s_k<\cdots<s_1<t-j} g_j (a\star g)_{t-j-s_1}\cdots \\ \times (a\star g)_{s_{k-1}-s_k}(a\star g)_{s_k-u}\zeta_{t-j}\zeta_{s_1}\cdots\zeta_{s_k}\eta_u\Bigg). \tag{2.6}$$

Next, we introduce some rigorous notions and definitions. Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space, and let $(\eta_s, \zeta_s) = (\eta_s, \zeta_s)_{s\in\mathbb{Z}}$ be a sequence of i.i.d. vectors defined on this space. Let $\mathcal{F}_t = \sigma\{\eta_s, \zeta_s,\ s \le t\}$, $t \in \mathbb{Z}$, be the increasing family of sub-σ-fields of $\mathcal{F}$. A random sequence (y_t) is called *adapted* if, for each $t \in \mathbb{Z}$, y_t is $\mathcal{F}_t$-measurable. Let $L^p(\Omega)$ $(1 \le p < \infty)$ denote the class of all complex-valued random variables ξ defined on $(\Omega, \mathcal{F}, \mathrm{P})$ such that $\mathrm{E}[|\xi|^p] < \infty$. Write l.i.m. for the limit in mean square. Set $\ell^p = \{\phi = (\phi_0, \phi_1, \ldots) \colon \|\phi\|_p < \infty\}$, $\|\phi\|_p = \{\sum_{j=0}^{\infty} |\phi_j|^p\}^{1/p}$, and $\|\phi\| := \|\phi\|_2$. We use C to designate generic quantities whose precise values are not important.

Assumption 2.1. *The generating functions $A(z)$ and $B(z)$ in (2.3) are analytic on $\{|z| < 1\}$ and $B(z) \ne 1$ ($|z| < 1$), $(g_j) \in \ell^2$, $(f_j) \in \ell^2$, and*

$$\|f\| = \left\{\sum_{j=1}^{\infty} f_j^2\right\}^{1/2} < 1.$$

Assumption 2.2. *We have*

$$\|(a^{(n)} - a)\star g\| \to 0, \qquad \|(b^{(n)} - b)\star g\| \to 0,$$

where $a_j^{(n)} := a_j \mathbf{1}(1 \le j \le n)$ and $b_j^{(n)} := b_j \mathbf{1}(1 \le j \le n)$.

Definition 2.1. By a solution of (1.5) we mean an adapted sequence (X_t) with finite second moment, $\mathrm{E}[X_t^2] < \infty$, such that, for every $t \in \mathbb{Z}$, the series A_t and B_t in (2.1) converge in mean square and (1.5) holds.

Note that the above definition implies the convergence $X_t = \text{l.i.m.}[\kappa\eta_t + \zeta_t \sum_{j=1}^{n} a_j X_{t-j} + b + \sum_{j=1}^{n} b_j X_{t-j}]$ in $L^2(\Omega)$.

Theorem 2.1. *Let Assumptions 2.1 and 2.2 be satisfied, and let $b = 0$ and $\kappa \ne 0$. Then there exists a solution (X_t) of (1.5) which is unique, strictly stationary, ergodic, and is given by the convergent orthogonal Volterra series in (2.6). Moreover,* $\mathrm{E}[X_t] = 0$ *and*

$$\mathrm{E}[X_0 X_t] = \frac{\kappa^2}{1 - \|f\|^2}\sum_{j=0}^{\infty} g_j g_{j+t}. \tag{2.7}$$

Proof. Let us check that the series in (2.6) converges and satisfies (2.7). Let $t \in \mathbb{N}$. Using the above notation, (2.6) can be rewritten as

$$X_t = \kappa\Bigg(\sum_{v\le t} g_{t-v}\eta_v + \sum_{u<v<0}\sum_{k=0}^{\infty}\sum_{u<s_k<\cdots<s_1<v<t} g_{t-v} f_{v-s_1}\cdots \\ \times f_{s_{k-1}-s_k} f_{s_k-u}\zeta_v\zeta_{s_1}\cdots\zeta_{s_k}\eta_u\Bigg). \tag{2.8}$$

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

By orthogonality of the last expansion,

$$\begin{aligned}\mathrm{E}[X_0X_t] &= \kappa^2\Bigg(\sum_{v\le 0} g_{t-v}g_{-v}\,\mathrm{E}[\zeta_v^2] + \sum_{u<v<0} g_{t-v}g_{-v}\sum_{k=0}^{\infty}\sum_{u<s_k<\cdots<s_1<v}^{\infty} f_{v-s_1}^2\cdots \\ &\qquad\qquad \times f_{s_{k-1}-s_k}^2 f_{s_k-u}^2\,\mathrm{E}[\zeta_v^2]\,\mathrm{E}[\zeta_{s_1}^2]\cdots\mathrm{E}[\zeta_{s_k}^2]\,\mathrm{E}[\eta_u^2]\Bigg) \\ &= \frac{\kappa^2}{1-\|f\|^2}\sum_{j=0}^{\infty} g_j g_{j+t},\end{aligned}$$

proving (2.7). It is clear that the process (X_t) in (2.8) is strictly stationary and adapted. The rest of the proof of the theorem is completely analogous to Giraitis and Surgailis (2002, Theorem 2.2).

Remark 2.1. In the case in which $b = 0$ and $\kappa \neq 0$, Assumptions 2.1 and 2.2 are also necessary for the existence of a solution in the sense of Definition 2.1. The necessity of Assumption 2.1 follows similarly as in Giraitis *et al.* (2004, proof of Theorem 1) and the necessity of Assumption 2.2 follows similarly as in Giraitis and Surgailis (2002, proof of Theorem 2.2).

Corollary 2.1. *Assume that the conditions of Theorem 2.1 are satisfied and that $\nu > 0$, and let (X_t) be the stationary solution of (2.6). Then, for any $t \in \mathbb{Z}$,*

$$X_t = m_t + \varepsilon_t\sigma_t, \tag{2.9}$$

where

$$\sigma_t := \sqrt{\nu^2 + (a + A_t)^2}, \qquad \varepsilon_t := \frac{\kappa\eta_t + \zeta_t A_t}{\sigma_t}, \tag{2.10}$$

and A_t and m_t are defined by the series in (2.1) convergent in $L^2(\Omega)$. Then $(\varepsilon_t, \mathcal{F}_t)$ forms a strictly stationary martingale difference sequence with

$$\mathrm{E}[\varepsilon_t \mid X_s,\ s < t] = 0, \qquad \mathrm{E}[\varepsilon_t^2 \mid X_s,\ s < t] = 1. \tag{2.11}$$

In particular, if (η, ζ) has a jointly Gaussian distribution with zero mean and covariance matrix

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

then (ε_t) forms a standard normal i.i.d. sequence and the conditional likelihood of $X_1 = x_1, \dots, X_n = x_n$ given $X_s,\ s \le 0$, equals

$$\prod_{t=1}^{n}(2\pi\sigma_t^2)^{-1/2}\exp\left\{-\frac{(x_t - m_t)^2}{2\sigma_t^2}\right\}. \tag{2.12}$$

Proof. The convergence in $L^2(\Omega)$ of the series defining A_t and B_t can be shown using Assumption 2.2, as in Giraitis and Surgailis (2002, proof of Theorem 2.2). Representation (2.9) is then immediate from the definitions in (2.10) and representation (1.5). The statements of the corollary about (ε_t) are easy consequences of the properties of (X_t) in Theorem 2.1 and assumptions (1.6) on the noise.

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

In the rest of the paper we will restrict ourselves to the zero-drift case, $b = b_j = 0$ $(j \geq 1)$, or the bilinear equation, (1.8). Note that in this case $g_t = \mathbf{1}(t = 0)$ is the δ-function and $f_j = a_j$ $(j \geq 1)$. The somewhat cumbersome formulae (2.6) and (2.8) can be rewritten in a more compact way using the notation

$$a_{u,t}^S := a_{t-s_1} a_{s_1-s_2} \cdots a_{s_k-u} \quad \text{and} \quad \zeta^S := \zeta_{s_1} \cdots \zeta_{s_k}$$

for

$$S = \{s_1, \dots, s_k\}, \qquad u < s_k < \cdots < s_1 < t, \qquad k = 1, \dots, \tag{2.13}$$

with $a_{u,t}^{\varnothing} := a_{t-u}$ and $\zeta^{\varnothing} := 1$. Then

$$X_t = \kappa \Bigg(\eta_t + \sum_{u<t} \sum_{S \subset (u,t)} a_{u,t}^S \zeta_t \zeta^S \eta_u \Bigg), \qquad A_t = \kappa \sum_{u<t} \sum_{S \subset (u,t)} a_{u,t}^S \zeta^S \eta_u, \tag{2.14}$$

where the sums over S are taken over all subsets, as in (2.13), including $k = 0$, or $S = \varnothing$.

Following the terminology in Giraitis *et al.* (2000), (2004) we consider two concrete examples of the LARCH$_+(\infty)$ models.

Example 2.1. (*GLARCH*$_+(1, 1)$ *model.*) Here $\sigma_t^2 = \nu^2 + (a + A_t)^2$ and $A_t = \alpha A_{t-1} + \beta X_{t-1}$, where ν, α, β, and a are real parameters. In the LARCH$_+(\infty)$ representation, (1.8), this model corresponds to $a_j = \beta\alpha^{j-1}$ $(j \geq 1)$, while Assumptions 2.1 and 2.2 are equivalent to $\alpha^2 + \beta^2 < 1$. It can be shown that the last condition is also necessary for the existence of a covariance stationary solution of the GLARCH$_+(1, 1)$ equations. Note that the GLARCH$_+(1, 1)$ model is different from Sentana's GQARCH(1, 1) model given by $\sigma_t^2 = \nu^2 + (a + \beta X_{t-1})^2 + b\sigma_{t-1}^2$ (see Sentana (1995)). Contrary to the GLARCH$_+(1, 1)$ model, the GQARCH(1, 1) model does not admit an explicit solution in Volterra series.

Example 2.2. (*GLARCH*$_+(0, d, 0)$ *model.*) Here $\sigma_t^2 = \nu^2 + (a + A_t)^2$, $A_t = c(1-L)^{-d} X_{t-1}$, or $A(z) = cz(1 - z)^{-d}$, where $c \in \mathbb{R}$ and $d \in (0, \frac{1}{2})$ are parameters. In this case, Assumptions 2.1 and 2.2 can be shown to be equivalent to $c^2 < \Gamma(1 - 2d)/\Gamma(1 - d)$.

3. Long memory and the leverage effect

Recall from the introduction the definition of the leverage function

$$H_t := \operatorname{cov}(\sigma_t^2, X_0), \qquad t = 1, 2, \dots, \tag{3.1}$$

which requires finiteness of the third moment, $\mathrm{E}[|X_t|^3] < \infty$, alone. Indeed, by (2.9) (with $B_t \equiv 0$) and (2.11),

$$H_t = \mathrm{E}[\varepsilon_t^2 \sigma_t^2 X_0] = \mathrm{E}[X_t^2 X_0],$$

implying that

$$|H_t| \leq \mathrm{E}^{2/3}[|X_t|^3]\, \mathrm{E}^{1/3}[|X_0|^3] = \mathrm{E}[|X_0|^3],$$

by stationarity and Hölder's inequality. The term 'leverage function' was introduced in Giraitis *et al.* (2004) to measure the so-called *leverage effect* (see Black (1976)), a tendency for volatility to move in the opposite direction to returns, after a delay, as happens when the conditional variance is negatively correlated with past returns. In the simplest case, the leverage effect can be measured by the absolute value of $\min(H_1, 0)$, although leverage of order $1 \leq k < \infty$ or $H_j < 0$ (for all $0 < j \leq k$) can be of interest also. See Giraitis *et al.* (2004, p. 178).

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

The subsequent discussion in this section resembles that in Giraitis *et al.* (2004, Section 1), and we refer to the above-mentioned paper for details. The discussion is technically simpler under the assumption that all mixed 'noise' moments of order three vanish:

$$\mathrm{E}[\eta^3] = \mathrm{E}[\eta^2\zeta] = \mathrm{E}[\eta\zeta^2] = \mathrm{E}[\zeta^3] = 0. \tag{3.2}$$

Conditions (3.2) together with $\mathrm{E}[|A_t|^3] < \infty$ imply that $\mathrm{E}[X_0^3] = 0$. In the last case, the leverage function in (3.1) satisfies a linear (Hilbert–Schmidt) equation, (3.3), below, whose derivation is completely similar to that in Giraitis *et al.* (2004, Equation (A.7)):

$$\begin{aligned} H_t &= \mathrm{E}\biggl[X_0\biggl(\nu^2 + \biggl(a + \sum_{j=1}^{\infty} a_j X_{t-j}\biggr)^2\biggr)\biggr] \\ &= \mathrm{E}\biggl[X_0\biggl(a + \sum_{s<t} a_{t-s} X_s\biggr)^2\biggr] \\ &= 2a\sigma^2 a_t + \sum_{0<s<t} a_{t-s}^2 H_s + 2a_t \sum_{s>0} a_{t+s} H_s. \end{aligned} \tag{3.3}$$

The (nontrivial) condition about finiteness of $\mathrm{E}[|A_t|^3]$ and $\mathrm{E}[|X_t|^3]$, which does not require finite fourth moment of the noise, is satisfied provided that (3.2) and

$$|\mu|_3^{1/3}\|a\|_3 + 3\theta\|a\| < 1 \tag{3.4}$$

hold, where $|\mu|_3 := \max(\mathrm{E}[|\eta|^3], \mathrm{E}[|\zeta|^3])$, $\|a\|_3 := \{\sum_{j=1}^{\infty} |a_j|^3\}^{1/3}$, and $\theta \approx 1.27$ is the solution of $3\theta^2 - 3\theta - 1 = 0$; see Giraitis *et al.* (2004, Proposition 1). Equation (3.3) coincides with Giraitis *et al.* (2004, Equation (A.7)), and the conclusions of the last paper about the sign of H_t (the 'leverage effect') and the asymptotics of H_t as $t \to \infty$ (the 'long memory') apply. These conclusions are formulated in Corollary 3.1, below, whose proof is omitted.

Corollary 3.1. *Assume that the conditions of Theorem 2.1 are satisfied, including (1.6), (1.7), and $\|a\| < 1$. Moreover, assume that conditions (3.2) and (3.4) are satisfied.*

(i) *Let (a_j) satisfy the regular decay condition in (1.9) with exponent $d \in (0, \frac{1}{2})$. Then (H_t) satisfies a similar condition, i.e.*

$$H_t \sim c_H t^{d-1}, \qquad c_H := \frac{2\sigma^4 c_0}{a}.$$

(ii) *Let $aa_1 < 0$ and $aa_j \le 0$, $j = 2, \dots, k$ for some $1 \le k < \infty$. Then $H_j < 0$, $j = 1, \dots, k$.*

4. Partial sums of the squared process

Now we come to our main problem—the limit distribution of the partial sums process $\sum_{t=1}^{[n\tau]}(X_t^2 - \mathrm{E}[X_t^2])$ ($\tau \in [0, 1]$), where (X_t) is a LARCH$_+(\infty)$ process in (1.8) with infinite fourth moment, $\mathrm{E}[X_t^4] = \infty$. The last fact is an easy consequence of assumptions (1.10) on the noise. Moreover, we also assume that conditions (1.6) are satisfied as in the previous sections, and set $\kappa = 1$ without loss of generality. Recall the Volterra representation of X_t in (2.14):

$$X_t = \eta_t + \sum_{u<t} \eta_u A_{u,t}\zeta_t, \qquad A_{u,t} := \sum_{S\subset(u,t)} a_{u,t}^S \zeta^S, \quad u < t.$$

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

Then

$$X_t^2 = \underbrace{\left((\eta_t^2 - \mathrm{E}[\eta_t^2]) + \sum_{u<t}(\eta_u^2 - \mathrm{E}[\eta_u^2])A_{u,t}^2\zeta_t^2\right)}_{(\text{'}\eta\text{-diagonal term'})}$$
$$+ \underbrace{2\left(\eta_t\zeta_t\sum_{u<t}\eta_u A_{u,t} + \sum_{u_2<u_1<t}\eta_{u_2}\eta_{u_1}A_{u_1,t}A_{u_2,t}\zeta_t^2\right) + 1 + \sum_{u<t}A_{u,t}^2\zeta_t^2}_{(\text{'}\eta\text{-off-diagonal term'})}$$
$$=: Y_t^{\mathrm{D}} + Y_t^{\mathrm{O}}. \tag{4.1}$$

The above decomposition separates the (centered) 'η-diagonal term', Y_t^{D}, with infinite variance and the remaining 'η-off-diagonal term', Y_t^{O}, with finite variance. We may expect that partial sums of these two terms have different limit distributions. The expectations are confirmed in the following theorems. Note that the second condition in (1.10) implies that

$$n^{-1/\alpha}\sum_{t=1}^{[n\tau]}(\eta_t^2 - \mathrm{E}[\eta_t^2]) \xrightarrow{\text{FDD}} L_\alpha(\tau), \tag{4.2}$$

where '$\xrightarrow{\text{FDD}}$' denotes weak convergence of finite-dimensional distributions, and $(L_\alpha(\tau))_{\tau\geq 0}$ is a homogeneous α-stable Lévy process with skewness parameter $\beta = 1$ and characteristic function

$$\mathrm{E}[\exp\{\mathrm{i}\theta L_\alpha(1)\}] = \exp\left\{-c|\theta|^\alpha\left(1 - \mathrm{i}\,\mathrm{sgn}(\theta)\tan\left(\frac{\alpha\pi}{2}\right)\right)\right\}, \qquad \theta \in \mathbb{R},$$

where the constant $c > 0$ depends on α and the asymptotic constant c_η in (1.10) (see, e.g. Ibragimov and Linnik (1971, Theorem 2.6.5)).

Theorem 4.1. ('Lévy α-stable limit'.) *Assume that conditions (1.10) are satisfied, that*

$$|a_j| \leq Cj^{d-1} \quad \left(\textit{there exist } C > 0 \textit{ and } 0 < d < \tfrac{1}{2}\right), \tag{4.3}$$

and that

$$(11)^{1/2}\mu_4^{1/4}\|a\| < 1. \tag{4.4}$$

Then

$$n^{-1/\alpha}\sum_{t=1}^{[n\tau]}Y_t^{\mathrm{D}} \xrightarrow{\text{FDD}} C_D L_\alpha(\tau), \tag{4.5}$$

where $(L_\alpha(\tau))$ is the α-stable Lévy process in (4.2),

$$C_D := \mathrm{E}^{1/\alpha}[Z_0^\alpha],$$

and

$$Z_t := 1 + \sum_{v>t}\left(\sum_{S\subset(t,v)} a_{t,v}^S\zeta^S\zeta_v\right)^2 \tag{4.6}$$

is a strictly stationary process with finite variance.

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

Theorem 4.2. ('Fractional Brownian motion limit'.) *Assume that*

$$\mathrm{E}[\zeta^4 + \zeta^2\eta^2] < \infty, \qquad \rho = \mathrm{E}[\zeta\eta] \neq 0,$$
$$a_j \sim c_0 j^{d-1} \quad \big(j \to \infty,\ \textit{there exist } c_0 \neq 0 \textit{ and } 0 < d < \tfrac{1}{2}\big),$$

and (4.4) are satisfied. Then

$$n^{-d-1/2} \sum_{t=1}^{[n\tau]} (Y_t^{\mathrm{O}} - \mathrm{E}[Y_t^{\mathrm{O}}]) \xrightarrow{\mathrm{FDD}} C_O B_{d+1/2}(\tau), \tag{4.7}$$

where $(B_{d+1/2}(\tau),\ \tau \geq 0)$ *is a standard fractional Brownian motion (FBM) with Hurst parameter* $d + \frac{1}{2}$ *and variance* $\mathrm{E}[B_{d+1/2}^2(\tau)] = \tau^{2d+1}$, *the asymptotic constant*

$$C_O := 2c_0 c(d)\rho(\mathrm{E}[X_0^2])^{3/2} = \frac{2c_0 c(d)\rho}{(1 - \|a\|^2)^{3/2}},$$

and $c(d) := (\int_{-\infty}^1 (\int_0^1 (t-s)_+^{d-1}\,\mathrm{d}t)^2\,\mathrm{d}s)^{1/2}$ *depends only on* d.

Corollary 4.1. *Assume that the conditions of Theorems 4.1 and 4.2 are satisfied. Then*

$$n^{-1/\alpha} \sum_{t=1}^{[n\tau]} (X_t^2 - \mathrm{E}[X_t^2]) \xrightarrow{\mathrm{FDD}} C_D L_\alpha(\tau) \quad \textit{if } \frac{1}{\alpha} > d + \frac{1}{2},$$
$$n^{-d-1/2} \sum_{t=1}^{[n\tau]} (X_t^2 - \mathrm{E}[X_t^2]) \xrightarrow{\mathrm{FDD}} C_O B_{d+1/2}(\tau) \quad \textit{if } \frac{1}{\alpha} < d + \frac{1}{2},$$

where the limiting quantities are the same as in Theorems 4.1 and 4.2.

Remark 4.1. The convergence (4.7) in Theorem 4.2 can be extended to functional convergence in the Skorokhod space $D[0, 1]$ by verifying the (Kolmogorov) tightness criterion for the variance of partial sums. However, the question of functional convergence in Theorem 4.1 and Corollary 4.1 remains open.

Let us present an example of a generic noise distribution satisfying conditions (1.10).

Example 4.1. Let $\zeta \sim N(0, 1)$ and $\eta = \zeta\sqrt{\xi}$, where $\xi > 0$ is a random variable, independent of ζ and such that $\mathrm{E}[\sqrt{\xi}] = \rho < 1$, $\mathrm{E}[\xi] = 1$, and $\mathrm{P}[\xi > x] \sim c_\xi x^{-\alpha}$ ($x \to \infty$, $c_\xi > 0$, $1 < \alpha < 2$). Then the pair (η, ζ) satisfies the conditions in (1.6) and (1.10).

5. Convergence to a Lévy stable process (proof of Theorem 4.1)

The proof of Theorem 4.1 is split into two steps.

Step 1. Reduction of the sum of the Y_t^{D}s to a corresponding sum of the stationary martingale transforms, $(\eta_t^2 - \mathrm{E}[\eta_t^2])Z_t$s, where the process (Z_t) is defined in Theorem 4.1.

Step 2. Proof of the central limit theorem (convergence to a stable law) for partial sums of the above martingale transform.

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

Step 2 is a particular case of a general central limit theorem (CLT) for stationary martingale transforms given in Appendix B, whose proof is based on the so-called 'principle of conditioning' due to Jakubowski (1986) (see also Kwapień and Woyczyński (1992, Chapter 5.8)). Step 1 means proving the relation

$$\sum_{t=1}^{n} Y_t^{\mathrm{D}} = \sum_{t=1}^{n} (\eta_t^2 - \mathrm{E}[\eta_t^2]) Z_t + o_p(n^{1/\alpha}), \tag{5.1}$$

and uses the diagram technique developed in Giraitis *et al.* (2000) for moments of multiple Volterra series, and some graph-theoretical argument. Technically, step 1 is the most involved part of the paper.

The proof of (5.1) (step 1) is as follows. Define $\tilde{\eta}_t := \eta_t^2 - \mathrm{E}[\eta_t^2]$. Then

$$Y_t^{\mathrm{D}} = \tilde{\eta}_t + \sum_{u<t} \tilde{\eta}_u A_{u,t}^2 \zeta_t^2, \qquad \tilde{\eta}_u Z_u = \tilde{\eta}_u + \tilde{\eta}_u \sum_{t>u} A_{u,t}^2 \zeta_t^2,$$

whence the difference $W_n := \sum_{t=1}^n Y_t^{\mathrm{D}} - \sum_{u=1}^n \tilde{\eta}_u Z_u$ of the sums in (5.1) can be rewritten as

$$\begin{aligned} W_n &= \sum_{u\le 0} \tilde{\eta}_u \sum_{t=1}^{n} A_{u,t}^2 \zeta_t^2 - \sum_{u=1}^{n} \tilde{\eta}_u \sum_{t>n} A_{u,t}^2 \zeta_t^2 \\ &=: W_{n1} - W_{n2}. \end{aligned}$$

Then (5.1) follows from

$$\mathrm{E}[|W_{ni}|^r] = o(n^{r_i/\alpha}), \qquad i = 1, 2, \tag{5.2}$$

for some $r > 0$. According to a well-known martingale inequality (due to von Bahr and Esséen (1965)), for any $1 \le r < \alpha$,

$$\begin{aligned} \mathrm{E}\,|W_{n1}|^r &\le C \sum_{u\le 0} \mathrm{E}\biggl[\biggl|\sum_{t=1}^{n} A_{u,t}^2 \zeta_t^2\biggr|^r\biggr] \\ &\le C \sum_{u\le 0} \mathrm{E}^{r/2}\biggl[\biggl(\sum_{t=1}^{n} A_{u,t}^2 \zeta_t^2\biggr)^2\biggr] \\ &\le C \mu_4^{r/2} \sum_{u\le 0} \mathrm{E}^{r/2}\biggl[\biggl(\sum_{t=1}^{n} A_{u,t}^2\biggr)^2\biggr], \end{aligned} \tag{5.3}$$

$$\mathrm{E}[|W_{n2}|^r] \le C \sum_{u=1}^{n} \mathrm{E}\biggl[\biggl|\sum_{t>n} A_{u,t}^2 \zeta_t^2\biggr|^r\biggr] \le C \mu_4^{r/2} \sum_{u=1}^{n} \mathrm{E}^{r/2}\biggl[\biggl(\sum_{t>n} A_{u,t}^2\biggr)^2\biggr], \tag{5.4}$$

where $\mu_4 := \mathrm{E}[\zeta^4]$.

We need to evaluate the right-hand sides of (5.3) and (5.4), which involve fourth mixed moments of Volterra series. To this end, we use the diagram approach developed in Giraitis *et al.* (2000), (2004), which is briefly described below.

Let a collection $(k)_4 = (k_1, \dots, k_4) \in \mathbb{Z}_+^4$ of integers be given; $\mathbb{Z}_+ := \{s \in \mathbb{Z}\colon s \ge 1\}$, $|k| := k_1 + \cdots + k_4$. Let $I = I(k)_4$ be a table having four rows of length k_j, $I_j = \{(k_j, j), \dots, (1, j)\}$, $j = 1, \dots, 4$. A *diagram* is an ordered partition $\gamma = (G_1, \dots, G_r)$ of

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

the table I by subsets (*edges*) G_q, $q = 1, \dots, r$, $r \geq 1$, having at least two elements and containing at most one element of any row (i.e. such that $2 \leq |G_q| \leq 4$ and $|G_q \cap I_j| \leq 1$, $q = 1, \dots, r$, $j = 1, \dots, 4$). The class of all such diagrams $\gamma = (G_1, \dots, G_r)$ over $I = I(k)_4$ will be denoted $\Gamma(k)_4$.

Let $S_j \subset \mathbb{Z}$, $|S_j| = k_j$, be a collection of ordered integers:

$$S_j = \{s_{k_j,j}, \dots, s_{1,j}\}, \qquad s_{k_j,j} < \cdots < s_{1,j}, \qquad j = 1, \dots, 4. \tag{5.5}$$

Set $(S)_4 := (S_1, \dots, S_4)$. Let

$$f = f((S)_4) = f(s_{i,j} : (i, j) \in I)$$

be a function defined on the set of all such collections $(S)_4$. With any such f and any diagram $\gamma = (G_1, \dots, G_r) \in \Gamma(k)_4$, we associate the sum

$$\sum_{(S)_4 \sim \gamma} f((S)_4) := \sum_{s_{i,j} = \hat{s}_q,\, (i,j) \in G_q,\, q = 1, \dots, r} f(s_{i,j} : (i, j) \in I)$$

over integers $s_{i,j}$, $i = 1, \dots, k_j$, $j = 1, \dots, 4$, satisfying the inequalities in (5.5) and such that

$$\hat{s}_1 < \cdots < \hat{s}_r.$$

Consider

$$\begin{aligned} J_1(u) &:= \mathrm{E}\biggl[\biggl(\sum_{t=1}^{n} A_{u,t}^2\biggr)^2\biggr] \\ &= \sum_{t',t''=1}^{n} \sum_{(S)_4} a_{u,t'}^{S_1} a_{u,t'}^{S_2} a_{u,t''}^{S_3} a_{u,t''}^{S_4} \,\mathrm{E}[\zeta^{S_1}\zeta^{S_2}\zeta^{S_3}\zeta^{S_4}] \\ &= \sum_{t',t''=1}^{n} \sum_{(k)_4} \sum_{\gamma \in \Gamma_{I(k)_4}} \mu(\gamma) \sum_{(S)_4 \sim \gamma} a_{u,(t)_4}^{(S)_4}, \end{aligned} \tag{5.6}$$

where we use the notation $(t)_4 := (t', t', t'', t'') \in \mathbb{Z}^4$ and

$$a_{u,(t)_4}^{(S)_4} := \begin{cases} a_{u,t'}^{S_1} a_{u,t'}^{S_2} a_{u,t''}^{S_3} a_{u,t''}^{S_4} & \text{if } S_i \in (u, t')\ (i = 1, 2),\ S_i \in (u, t'')\ (i = 3, 4), \\ 0 & \text{otherwise,} \end{cases}$$

and where

$$\mu(\gamma) := \mathrm{E}[\zeta^{S_1}\zeta^{S_2}\zeta^{S_3}\zeta^{S_4}]$$

in (5.6) depends only on $\gamma \in \Gamma(k)_4$ and vanishes unless the sets $S_1, \dots, S_4$ are 'coupled', i.e. $\Delta(S)_4 := \bigcup_{i=1}^4 (S_i \setminus \bigcup_{j \neq i} S_j) \neq \varnothing$; moreover, by Hölder's inequality,

$$|\mu(\gamma)| \leq \mu_4^{|k|/4} \quad (\mu_4 := \mathrm{E}[\zeta^4]) \tag{5.7}$$

(see Giraitis *et al.* (2000, Equations (3.5), (3.6), and (3.15))). Following Giraitis *et al.* (2000, p. 1013), we call $\gamma = (G_1, \dots, G_r) \in \Gamma(k)_4$ *block connected* if it contains an edge G_q $(1 \leq q \leq r)$ which has a nonempty intersection with blocks $I' := I_1 \cup I_2$ and $I'' := I_3 \cup I_4$. We also note the following simple bound on the number $|\Gamma(k)_4|$ of all diagrams:

$$|\Gamma(k)_4| \leq (11)^{|k|/2}; \tag{5.8}$$

see Giraitis *et al.* (2000, Lemma 3.2). To proceed, we will need an auxiliary lemma.

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

Lemma 5.1. *There exists a constant $C < \infty$ such that, for any integers $u, t', t'', u < t' \wedge t''$, any $(k)_4 \in \mathbb{Z}_+^4$, and any diagram $\gamma \in \Gamma((k)_4)$, the following inequality is true:*

$$\sum_{(S)_4 \sim \gamma} |a_{u,(t)_4}^{(S)_4}| \le C|k|^6 \|a\|^{|k|} |t' - u|^{2d-2} |t'' - u|^{2d-2}. \tag{5.9}$$

Proof. The proof of (5.9) is easy if γ is not block connected. Assuming that γ is not block connected, we necessarily have $S_1 = S_2 =: S'$ and $S_3 = S_4 =: S''$, and the left-hand side of (5.9) can be estimated by a product of two sums over rows 1 and 2 and over rows 3 and 4:

$$\sum_{(S)_4 \sim \gamma} |a_{u,(t)_4}^{(S)_4}| \le \Lambda_{k'+1}(t' - u) \Lambda_{k''+1}(t'' - u), \tag{5.10}$$

where

$$\Lambda_k(t) := \sum_{0 < s_{k-1} < \cdots < s_1 < t} a_{t-s_1}^2 \cdots a_{s_{k-1}}^2, \qquad t, k = 1, 2, \ldots, \tag{5.11}$$

$\Lambda_k(t) := 0$ $(t \le 0)$, is the kth convolution of the sequence $(a_t^2)_{t \ge 1}$. Condition (4.3) of Theorem 4.1 implies the following inequality: for any $t, k = 1, 2, \ldots$ (see Giraitis *et al.* (2000, Lemma 4.2)),

$$\Lambda_k(t) \le C k^3 \|a\|^k t^{2d-2}, \tag{5.12}$$

where the constant C is the same as in (4.3). From (5.10) and (5.12), we obtain

$$\begin{aligned} \sum_{(S)_4 \sim \gamma} |a_{u,(t)_4}^{(S)_4}| &\le C(k')^3 (k'')^3 \|a\|^{2(k'+k'')} |t' - u|^{2d-2} |t'' - u|^{2d-2} \\ &\le C|k|^6 \|a\|^{|k|} |t' - u|^{2d-2} |t'' - u|^{2d-2}, \end{aligned} \tag{5.13}$$

or the statement of Lemma 5.1, provided that γ is not block connected.

The proof of Lemma 5.1 for the case of a general (i.e. block-connected) diagram $\gamma \in \Gamma(k)_4$ can be reduced to the 'block-unconnected' case above by an argument which is explained below.

Let us first consider the situation when $\gamma = (G_1, \ldots, G_r)$ connects only pairs: $|G_1| = \cdots = |G_r| = 2$. The subclass of all such diagrams will be denoted $\Gamma_2(k)_4$. Later we will extend the proof to an arbitrary diagram $\gamma \in \Gamma(k)_4 \setminus \Gamma_2(k)_4$.

The idea of the proof is to split the product $|a_{u,(t)_4}^{(S)_4}| = |A' \cdot A''| \le \frac{1}{2}((A')^2 + (A'')^2)$, where $(A')^2 = a_{u,(t)_4}^{(S')_4}$ and $(A'')^2 = a_{u,(t)_4}^{(S'')_4}$ correspond to some new diagrams γ' and γ'' which are *not* block connected and for which the bound (5.9) follows similarly as in (5.13).

To obtain such a decomposition, we use a graph-theoretical fact in Lemma 5.2, below. The proof of Lemma 5.2 is carried over to Appendix A. Consider a (multi)graph $\mathcal{G} = (V, E)$ with a set $V = \{0, 1, \ldots, r, r+1, r+2\} \subset \mathbb{Z}$ of vertices and a set E of edges (a multiset of unordered pairs of distinct vertices). For a vertex $i \in V$, we define $\deg(i) = |\{(i, j) \in E\}|$, $\deg_+(i) = |\{(i, j) \in E,\ j > i\}|$, and $\deg_-(i) = |\{(i, j) \in E,\ j < i\}|$, so that $\deg(i) = \deg_+(i) + \deg_-(i)$ $(i \in V)$. An *Eulerian cycle* in $\mathcal{G}$ is a cycle which uses each edge $(i, j) \in E$ exactly once. A *Hamiltonian path* in $\mathcal{G}$ is a path which uses each vertex $i \in V$ exactly once.

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

Lemma 5.2. *Assume that a multigraph $\mathcal{G} = (V, E)$ with $V = \{0, 1, \dots, r+2\}$ satisfies the following conditions:*

$$\deg_+(i) = 4, \qquad \deg_-(i) = 0, \qquad i = 0, \tag{5.14}$$

$$\deg_+(i) = \deg_-(i) \in \{2, 3, 4\}, \qquad i = 1, \dots, r, \tag{5.15}$$

$$\deg_-(i) = 2, \qquad \deg_+(i) = 0, \qquad i = r+1, r+2, \tag{5.16}$$

$$\deg(i) = 6 \implies (i, i \pm 1) \in E, \qquad i = 1, \dots, r, \tag{5.17}$$

$$\deg(i) = 8 \implies (i', i'') \notin E, \qquad i' < i < i'', \ i = 1, \dots, r. \tag{5.18}$$

Then $\mathcal{G}$ contains two distinct paths, $\check{H}$ and $\hat{H}$, which start at $r+2$, go to 0, and return to $r+1$, and whose union $\check{H} \cup \hat{H}$ forms an Eulerian cycle. More precisely,

$$\check{H} : r+2 \to \check{i}_1 \to \cdots \to \check{i}_{\check{k}-1} \to \check{i}_{\check{k}} = 0 \to \check{i}_{\check{k}+1} \to \cdots \to \check{i}_{\check{p}} \to r+1, \tag{5.19}$$

$$\hat{H} : r+2 \to \hat{i}_1 \to \cdots \to \hat{i}_{\hat{k}-1} \to \hat{i}_{\hat{k}} = 0 \to \hat{i}_{\hat{k}+1} \to \cdots \to \hat{i}_{\hat{p}} \to r+1, \tag{5.20}$$

with the following properties:

$$r \ge \check{i}_1 > \cdots > \check{i}_{\check{k}-1} > \check{i}_{\check{k}} = 0 < \check{i}_{\check{k}+1} < \cdots < \check{i}_{\check{p}} \le r, \tag{5.21}$$

$$r \ge \hat{i}_1 > \cdots > \hat{i}_{\hat{k}-1} > \hat{i}_{\hat{k}} = 0 < \hat{i}_{\hat{k}+1} < \cdots < \hat{i}_{\hat{p}} \le r, \tag{5.22}$$

$$\{\check{i}_1, \dots, \check{i}_{\check{k}-1}\} \cup \{\check{i}_{\check{k}+1}, \dots, \check{i}_{\check{p}}\} = \{1, 2, \dots, r\}, \tag{5.23}$$

$$\{\hat{i}_1, \dots, \hat{i}_{\hat{k}-1}\} \cup \{\hat{i}_{\hat{k}+1}, \dots, \hat{i}_{\hat{p}}\} = \{1, 2, \dots, r\}. \tag{5.24}$$

Moreover, the respective lengths $\check{p}+1$ and $\hat{p}+1$ of $\check{H}$ and $\hat{H}$ satisfy

$$\check{p} + 1 = \hat{p} + 1 = \frac{|E|}{2} \quad \text{if } |E| \text{ is even}, \tag{5.25}$$

$$\check{p} + 1 + \hat{p} + 1 = |E| \quad \text{and} \quad |\check{p} - \hat{p}| = 1 \quad \text{if } |E| \text{ is odd}. \tag{5.26}$$

Remark 5.1. Conditions (5.14)–(5.18) arise from the relationship between diagrams and multigraphs, as defined in the proof of Lemma 5.1 for arbitrary diagrams, below. Figures 1 and 2 help us to understand this relationship and the meaning of Lemma 5.2.

Proof of Lemma 5.1 for arbitrary diagrams. With any diagram $\gamma = (G_1, \dots, G_r) \in \Gamma(k)_4$ we can associate a graph $\mathcal{G} = (V, E)$ with $r+3$ vertices as follows. The set V of vertices of this graph,

$$V := \{G_0, G_1, \dots, G_r, G', G''\},$$

can obviously be identified with the set $\tilde{V} := \{0, 1, 2, \dots, r, r+1, r+2\} \subset \mathbb{N}_+$. The first $r+1$ elements of $\tilde{V}$ correspond to the ordered indices $u := \tilde{s}_0 < \tilde{s}_1 < \cdots < \tilde{s}_r$ and the last two elements correspond to $\tilde{s}_{r+1} := t' < \tilde{s}_{r+2} := t''$ in (5.9). An edge $e \in E$ in $\mathcal{G}$ between vertices $i, j \in \tilde{V}$ $(i < j)$ is equivalent to the existence of a corresponding factor $a_{\tilde{s}_i - \tilde{s}_j} =: a_e$ in the product on the left-hand side of (5.9), which can be rewritten as $a^{(S)_4}_{u,(t)_4} = \prod_{e \in E} a_e$. It is easy to see that the constructed graph $\mathcal{G} = (V, E)$ satisfies the conditions of Lemma 5.2. Let $E = \check{E} \cup \hat{E}$ be the decomposition of E into disjoint subsets corresponding to the paths $\check{H}$ and $\hat{H}$ in Lemma 5.2. Clearly, $|\prod_{e \in E} a_e| \le \frac{1}{2}(\prod_{e \in \check{E}} a_e^2 + \prod_{e \in \hat{E}} a_e^2)$. We thus obtain

$$\begin{aligned} \sum_{(S)_4 \sim \gamma} |a^{(S)_4}_{u,(t)_4}| &\le \frac{1}{2}\biggl(\sum_{(S)_4 \sim \gamma} \prod_{e \in \check{E}} a_e^2 + \sum_{(S)_4 \sim \gamma} \prod_{e \in \hat{E}} a_e^2 \biggr) \\ &\le \tfrac{1}{2}\{\Lambda_{\check{k}}(t'' - u)\Lambda_{\check{k}+1}(t' - u) + \Lambda_{\hat{k}}(t'' - u)\Lambda_{\hat{k}+1}(t' - u)\}, \end{aligned} \tag{5.27}$$

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

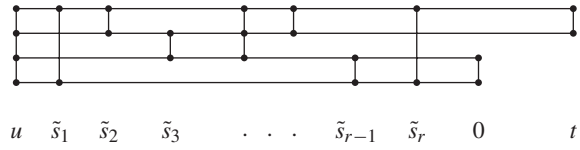

FIGURE 1: A diagram $\gamma = (G_1, \dots, G_r)$ with $r = 7$ edges.

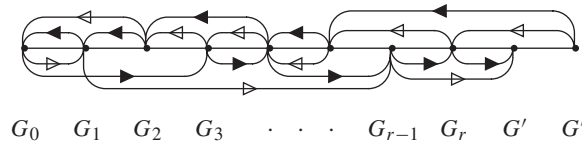

FIGURE 2: The graph $\mathcal{G}$ with $r+3 = 10$ vertices corresponding to the diagram γ in Figure 1. The Eulerian cycle $\check{H} \cup \hat{H}$ of Lemma 5.2 is shown (path $\check{H}$ is indicated by closed-headed arrows and path $\hat{H}$ is indicated by open-headed arrows).

where the function Λ_k is defined in (5.11). From (5.25), (5.26), and (5.27), the statement of the lemma follows exactly as in (5.13). Lemma 5.1 is proved.

Let us return to the proof of (5.2) and the estimates in (5.3) and (5.4). According to Lemma 5.1, (5.6), (5.7), and (5.8),

$$J_1(u) \le C \sum_{(k)_4} \mu_4^{|k|/4} (11)^{|k|/2} |k|^6 \|a\|^{|k|} \sum_{t',t''=1}^{n} |t'-u|^{2d-2} |t''-u|^{2d-2} \tag{5.28}$$

and, therefore, $\mathrm{E}[|W_{n1}|^r] (\le C \sum_{u \le 0} J_1^{r/2}(u))$, see (5.3), does not exceed

$$\begin{aligned} \mathrm{E}[|W_{n1}|^r] &\le C \left(\sum_{k=1}^{\infty} \mu_4^{k/4} (11)^{k/2} k^6 \|a\|^k \right)^{2r} \sum_{u=0}^{\infty} \left(\sum_{t=1}^{n} |t+u|^{2d-2} \right)^r \\ &\le C + C \int_1^{\infty} \left(\int_0^n (t+u)^{2d-2} \, \mathrm{d}t \right)^r \mathrm{d}u \\ &\le C(1 + n^{1+(2d-1)r}), \end{aligned} \tag{5.29}$$

since the last integral does not exceed

$$\int_1^n \left(\int_0^u u^{2d-2} \, \mathrm{d}t + \int_u^{\infty} t^{2d-2} \, \mathrm{d}t \right)^r \mathrm{d}u + \int_n^{\infty} \left(u^{2d-2} \int_0^n \mathrm{d}t \right)^2 \mathrm{d}u \le C n^{1+(2d-1)r}$$

for any $d < \frac{1}{2}$, $(2d-2)r > 1$.

Let $1 < \alpha \le 1/(1-2d)$. Then $1 + (2d-1)r > 0$ for $r < \alpha$ and (5.29) implies that

$$\mathrm{E}[|W_{n1}|^r] = O(n^{1+(2d-1)r}) = o(n^{r/\alpha})$$

provided that $1 + (2d-1)r < r/\alpha$ holds. The last inequality is clearly satisfied for $r = \alpha$ since $1 + (2d-1)\alpha < 1$ and, therefore, also for all $r < \alpha$ sufficiently close to α. Next, let $\alpha > 1/(1-2d)$. Then $1 + (2d-1)r < 0$ for all $r < \alpha$ sufficiently close to α and, hence, (5.29) implies the relation

$$\mathrm{E}[|W_{n1}|^r] = O(1) = o(n^{r/\alpha})$$

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

trivially. This proves the bound in (5.2) for $i = 1$ and a suitably chosen r. In a similar way,

$$J_2(u) \le C \sum_{(k)_4} \mu_4^{|k|/4} (11)^{|k|/2} |k|^6 \|a\|^{|k|} \sum_{t',t''>n} |t'-u|^{2d-2} |t''-u|^{2d-2},$$

similarly as in (5.6) and (5.28). Hence, $\mathrm{E}[|W_{n2}|^r] \le C \sum_{u=1}^n J_2(u)^{r/2}$ (see (5.4)) can be estimated as

$$\begin{aligned}\mathrm{E}[|W_{n2}|^r] &\le C\biggl(\sum_{k=1}^{\infty} \mu_4^{k/4} (11)^{k/2} k^6 \|a\|^k\biggr)^{2r} \sum_{u=1}^{n} \biggl(\sum_{t>n} |t-u|^{2d-2}\biggr)^r \\ &\le C \int_1^n \biggl(\int_n^{\infty} (t-u)^{2d-2}\,\mathrm{d}t\biggr)^r \mathrm{d}u \\ &\le C(1+n^{1+(2d-1)r}),\end{aligned}$$

as in (5.29). This proves (5.2) for $i = 2$ and a suitably chosen r. It also completes the proof of (5.1) (step 1).

Proof of Theorem 4.1. With (5.1) and stationarity of all processes involved in mind, relation (4.5) follows from

$$n^{-1/\alpha} \sum_{t=-1}^{[-n\tau]} (\eta_t^2 - \mathrm{E}[\eta_t^2]) Z_t \xrightarrow{\text{FDD}} C_D L_\alpha(\tau). \tag{5.30}$$

To show (5.30), apply Theorem B.1, with $\xi_t := \eta_{-t}^2 - \mathrm{E}[\eta_{-t}^2]$, $V_t := Z_{-t}$, and $\mathcal{H}_t := \sigma\{\eta_s, \zeta_s, s > -t\}$. Then the left-hand side of (5.30) can be rewritten as in (B.5), with U_i defined as in (B.1). Assumptions (U1) and (U2) in Appendix B clearly apply in our case: representation (B.2) is a consequence of the tail condition in (1.10) and the representation of the characteristic function in the domain of attraction of a stable law (see, e.g. Ibragimov and Linnik (1971, Theorem 2.6.5)). The ergodicity of the time-reversed process (Z_{-t}) follows from Stout (1974, Theorem 3.5.8). The moment condition (B.4) holds with $r = 2 > \alpha$, since the series in (4.6) converges in mean square. This concludes the proof of Theorem 4.1.

6. Convergence to FBM (proof of Theorem 4.2)

The proof of Theorem 4.2 resembles that in Giraitis *et al.* (2000, Theorem 2.3), but is simpler since in the present work we deal with squares only. Hence, we present a sketch of the proof, the details being similar as in the above-mentioned paper.

Introduce an 'intermediate' stationary process (ϖ_t) (the analog of Giraitis *et al.* (2000, Equation (2.19))), defined by

$$\begin{aligned}\varpi_t &:= 2\rho \sum_{k=0}^{\infty} \sum_{u<s_k<\cdots<s_1<v\le t} G_{t-v} a_{v-s_1} \cdots a_{s_{k-1}-s_k} a_{s_k-u} \zeta_{s_1} \cdots \zeta_{s_k} \eta_u \\ &= 2\rho \sum_{v\le t} \tilde{a}_{t-s} X_s,\end{aligned} \tag{6.1}$$

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

where

$$\tilde{a}_t := \sum_{0\le v<t} G_v a_{t-v},$$
$$G_{t-v} := \sum_{k=1}^{t-v-1} \sum_{v<s_k<\cdots<s_1<t} a_{t-s_1}^2 \cdots a_{s_k-v}^2, \qquad v < t'$$
$$G_0 := 1.$$

Then (see Giraitis *et al.* (2000, Lemma 4.1, Equation (4.3)))

$$\sum_{v\le t} G_{t-v} = 1 + \sum_{k=0}^{\infty} \sum_{t>v} \sum_{v<s_k<\cdots<s_1<t} a_{t-s_1}^2 \cdots a_{s_k-sv}^2 = \frac{1}{1-\|a\|^2} = \sigma^2,$$
$$\tilde{a}_t \sim \left(\sum_{i=0}^{\infty} G_i\right) a_t(1+o(1)) = \sigma^2 a_t(1+o(1)) \quad \text{as } t\to\infty. \tag{6.2}$$

By definition (6.1), (ϖ_t) is a moving average in a stationary martingale difference process (X_t) with variance $\sigma^2 = \mathrm{E}[X_0^2] = 1/(1-\|a\|^2)$. Then, using asymptotics (6.2), we can show the convergence

$$n^{-d-1/2} \sum_{t=1}^{[n\tau]} \varpi_t \xrightarrow{\text{FDD}} 2c_0 c(d)\rho\sigma^3 B_{d+1/2}(\tau). \tag{6.3}$$

See also the proof of Equation (2.27) in Giraitis *et al.* (2000). With (6.3) in mind, the statement of the theorem follows from the approximation

$$\operatorname{var}\left(\sum_{t=1}^{n} (Y_t^{\mathrm{O}} - \varpi_t)\right) = o(n^{2d+1}). \tag{6.4}$$

Relation (6.4) follows from

$$\operatorname{cov}(\varpi_t, \varpi_0) \sim \tilde{C} t^{2d-1} \quad \left(t\to\infty,\ \tilde{C} := 4\rho^2 c_0^2 \sigma^6 \tilde{c}(d),\ \tilde{c}(d) := \int_0^\infty (u(1+u))^{d-1}\,\mathrm{d}u\right)$$

(which is an easy consequence of (4.1), (6.1), and (6.2)) and the following relations:

$$\operatorname{cov}\left(\sum_{u<t} A_{u,t}^2 \zeta_t^2, \sum_{u<0} A_{u,0}^2 \zeta_0^2\right) = o(t^{2d-1}), \tag{6.5}$$
$$\operatorname{cov}(\tilde{Y}_t^{\mathrm{O}}, \tilde{Y}_0^{\mathrm{O}}) \sim \operatorname{cov}(\tilde{Y}_t^{\mathrm{O}}, \varpi_0) \sim \operatorname{cov}(\varpi_t, \tilde{Y}_0^{\mathrm{O}}) \sim \operatorname{cov}(\varpi_t, \varpi_0) \quad \text{as } t\to\infty, \tag{6.6}$$

where

$$\tilde{Y}_t^{\mathrm{O}} := Y_t^{\mathrm{O}} - \sum_{u<t} A_{u,t}^2 \zeta_t^2 = 2\left(\eta_t \zeta_t \sum_{u<t} \eta_u A_{u,t} + \sum_{u_2<u_1<t} \eta_{u_2}\eta_{u_1} A_{u_1,t} A_{u_2,t} \zeta_t^2\right).$$

The proofs of (6.5) and (6.6) use the diagram approach and the argument in Giraitis *et al.* (2000). In particular, (6.5) reduces to (a special case of) *irregular* diagrams in Giraitis *et al.* (2000, Lemma 5.2). The proof of (6.6) is very similar to Giraitis *et al.* (2000, Lemma 5.1). This completes the proof of Theorem 4.2.

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

Appendix A. Proof of Lemma 5.2

The statement and the proof of the lemma simplifies in the case when all 'inner' vertices $i = 1, 2, \dots, r$ have degree 4, i.e.

$$\deg_+(i) = \deg_-(i) = 2, \qquad 1 \le i \le r. \tag{A.1}$$

In such a case, $\check{H}$ and $\hat{H}$ are (distinct) Hamiltonian paths. Indeed, assume that $\check{H}$ visits a vertex $i \in \{1, 2, \dots, r\}$ twice. Therefore, $\check{H}$ traverses all four edges incident to i. Since $\check{H}$ and $\hat{H}$ have no edges in common, this means that $\hat{H}$ does not visit i and, therefore, violates condition (5.24). Note that in our case, it suffices to show the existence of a (Hamiltonian) path $\check{H}$ only, since the construction of $\hat{H}$ uses the edges of $\mathcal{G}$ which were not traversed by $\check{H}$ and, hence, it is trivial.

Assuming that (A.1) holds, the subsequent proof uses induction on r. For $r = 0, 1, 2$, the statement of the lemma is obvious by visual inspection.

Let us prove the induction step $r - 1 \to r$. To this end, we assume that the statement of the lemma holds for any multigraph $\mathcal{G}' = (V', E')$ having $r + 2$ vertices.

Define $\mathcal{G}' = (V', E')$ by $V' := \{1, 2, \dots, r + 2\} = V \setminus \{0\}$ and

$$E' := \{(i, j) \in E : i, j \in V'\} \cup \{(1, j) : j = 2, \dots, r + 2,\ (0, j) \in E\}.$$

Note that $\mathcal{G}'$ satisfies the conditions of the lemma, in particular, $\deg_+(1) = 4$ and $\deg_+(2) = \deg_-(2) = 2$. Let

$$\check{H}' : r + 2 \to i'_1 \to \cdots \to i'_{k'-1} \to i'_{k'} = 1 \to i_{k'+1} \to \cdots \to i'_{p'} \to r + 1$$

be a corresponding path in $\mathcal{G}'$ satisfying property (5.19) for this graph. Without loss of generality, we can assume that $i'_{k'-1} = 2$, which implies that $i'_{k'+1} > 2$ and $(1, i'_{k'+1}) \in E'$, and then $(0, i'_{k'+1}) \in E$ by the construction of $\mathcal{G}'$. Define $\check{H}$ as follows: $\check{k} := k' + 1$,

$$\begin{aligned}\check{H} : r + 2 &\to \check{i}_1 := i'_1 \to \cdots \to \check{i}_{\check{k}-1} := i'_{k'} = 1 \to \check{i}_{\check{k}} := 0 \\ &\to \check{i}_{\check{k}+1} := i'_{k'+1} \to \cdots \to \check{i}_{\check{p}} := i'_{p'} \to r + 1.\end{aligned}$$

Clearly, the constructed path $\check{H}$ visits each point of $\mathcal{G}$ exactly once. Because of the last fact, property (5.25) is obvious since $|E| = 2(r + 2)$ is even and $\hat{p} = \check{p} = r + 2$. This proves the lemma for the case in (A.1).

In the general case, the proof of the lemma can be obtained by induction on the number of vertices which do not satisfy condition (A.1). Let

$$\begin{aligned}q_3(\mathcal{G}) &:= |\{i = 1, \dots, n : \deg_+(i) = \deg_-(i) = 3\}|, \\ q_4(\mathcal{G}) &:= |\{i = 1, \dots, n : \deg_+(i) = \deg_-(i) = 4\}|.\end{aligned}$$

Assume that the statement of the lemma holds for any multigraph with $q_3(\mathcal{G}) \le q - 1$ ($q = 1, 2, \dots$) and $q_4(\mathcal{G}) = 0$. Let us show that it also holds for $q_3(\mathcal{G}) = q$ and $q_4(\mathcal{G}) = 0$.

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

Assume first that $|E|$ is odd. Note that the last assumption is equivalent to $q_3(\mathcal{G}) = q$ being odd. Indeed, since

$$\begin{aligned}|E| &= \frac{1}{2}\sum_{i\in V}\deg(i)\\ &= \frac{1}{2}\Bigg(\deg(0) + 4\sum_{i=1:\ \deg(i)=4}^{r} 1 + 6q + \sum_{i=r+1}^{r+2}\deg(i)\Bigg)\\ &= 2\Bigg(1 + \sum_{i=1:\ \deg(i)=4}^{r} 1 + 2\Bigg) + 3q,\end{aligned}$$

so $|E| - 3q$ is an even number and parities of $|E|$ and q must coincide.

Consider any vertex $i^* = 2, \dots, r$ with $\deg_+(i^*) = \deg_-(i^*) = 3$. Let $(i^*, j^*_\pm) \in E$, $j^*_- < i^* < j^*_+$. Define a new graph $\mathcal{G}^* = (V^*, E^*)$ by $V^* := V$ and

$$E^* := \{(i, j) \in E,\ (i, j) \neq (i^*, j^*_\pm)\} \cup \{(j^*_-, j^*_+)\}.$$

Then $\mathcal{G}^*$ satisfies the assumptions of the lemma and $q_3(\mathcal{G}^*) = q - 1$, implying that $|E^*|$ is even. Therefore, by the inductive assumption, it satisfies the statement of the lemma, i.e. it contains paths $\check{H}^*$ and $\hat{H}^*$ with properties as in (5.19)–(5.20) and (5.21)–(5.24), such that $\check{p}^* = \hat{p}^*$. Since $\check{H}^* \cup \hat{H}^*$ is an Eulerean cycle in $\mathcal{G}^*$, either $\check{H}^*$ or $\hat{H}^*$ contains the edge $(j^*_-, j^*_+) \in E^*$. Assume that $(j^*_-, j^*_+) \in \check{H}^*$ and that $\check{H}^*$ traverses this edge when going from $r + 2$ to 0, so that $j^*_+ = \check{i}^*_v$ and $j^*_- = \check{i}^*_{v+1}$ for some $1 \le v < \check{k}^*$ in the corresponding path

$$\begin{aligned}\check{H}^*\colon r + 2 &\to \cdots \to \check{i}^*_v = j^*_+ \to \check{i}^*_{v+1} = j^*_- \to \cdots \to \check{i}^*_{\check{k}^*} = 0 \to \check{i}^*_{\check{k}^*+1} \to \cdots\\ &\to \check{i}^*_{\check{p}^*} \to r + 1.\end{aligned}$$

Define $\check{H}$ by $\check{i}_u := \check{i}^*_u$ $(1 \le u \le v)$, $\check{i}_{v+1} := i^*$, and $\check{i}_u := \check{i}^*_{u-1}$ $(v + 2 \le u \le \check{p} := \check{p}_* + 1)$, i.e.

$$\begin{aligned}\check{H}\colon r + 2 &\to \cdots \to \check{i}_v = j^*_+ \to \check{i}_{v+1} = i^* \to \check{i}_{v+2} = j^*_- \to \cdots \to \check{i}_{\check{k}^*+1} = \check{i}^*_{\check{k}^*} = 0\\ &\to \check{i}_{\check{k}*+2} = \check{i}^*_{\check{k}^*+1} \to \cdots \to \check{i}_{\check{p}^*+1} = \check{i}^*_{\check{p}^*} \to r + 1,\end{aligned}$$

and $\hat{H} := \hat{H}^*$. Then the constructed paths $\check{H}$ and $\hat{H}$ can easily be shown to satisfy the statement of the lemma, including (5.26). This proves the induction step $q_3(\mathcal{G}) = q - 1 \to q$ in the case when $|E|$ is odd.

Next, assume that $|E|$ is even (and $q_3(\mathcal{G}) \ge 2$). Consider two vertices $i^*, i^{**} \in \{1, \dots, r\}$, $i^* < i^{**}$, with

$$\deg_+(i^*) = \deg_-(i^*) = \deg_+(i^{**}) = \deg_-(i^{**}) = 3.$$

Let $(i^*, j^*_\pm), (i^{**}, j^{**}_\pm) \in E$, $j^*_- < i^* < j^*_+$, $j^{**}_- < i^{**} < j^{**}_+$. Define a new graph $\mathcal{G}^* = (V^*, E^*)$ by $V^* := V$ and

$$E^* := \{(i, j) \in E,\ (i, j) \neq (i^*, j^*_\pm),\ (i, j) \neq (i^{**}, j^{**}_\pm)\} \cup \{(j^*_-, j^*_+), (j^{**}_-, j^{**}_+)\}.$$

Then $\mathcal{G}^*$ satisfies the assumptions of the lemma and $q_3(\mathcal{G}^*) = q - 2$ is even. Therefore, by the inductive assumption, it satisfies the statement of the lemma, i.e. it contains paths $\check{H}^*$ and $\hat{H}^*$ with properties as in (5.19)–(5.20) and (5.21)–(5.24), such that $\check{p}^* = \hat{p}^*$. Proceeding

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

with the construction as in the case when $|E|$ is odd results in disjoint paths $\check{H}$ and $\hat{H}$ which satisfy (5.19)–(5.20) and (5.21)–(5.24) of the lemma with the (possible) exception of (5.25): the construction implies a weaker fact, namely, that $\check{p} - \hat{p} \in \{-2, 0, 2\}$.

Let us show that $\check{H}$ and $\hat{H}$ can be further modified to the paths $\check{H}^\dagger$ and $\hat{H}^\dagger$ satisfying all the statements of the lemma, including (5.25), or $\check{p}^\dagger = \hat{p}^\dagger$.

Assume without loss of generality that $(0 <) i^{**} < i^*(\le r)$ and that $\check{p} = \hat{p} + 2$. The assumption implies that the path $\check{H}$ visits each point i^* and i^{**} *twice* (the first time going from $r+2$ to 0 and the second time going backwards from 0 to $r+1$), while $\hat{H}$ visits these points exactly *once*. (The fact is a consequence of the observation that conditions (5.21) and (5.23) imply that $\hat{p} = r + \hat{q} + 1$ and $\check{p} = r + \check{q} + 1$, where $\hat{q}$ and $\check{q}$ are respective numbers of vertices of E visited twice by $\hat{H}$ and $\check{H}$.)

The paths $\check{H}^\dagger$ and $\hat{H}^\dagger$ will be defined by 'switching' from $\check{H}$ to $\hat{H}$ and back to $\check{H}$, and vice versa, and 'pasting together' the corresponding parts of $\check{H}$ and $\hat{H}$.

Observe first that the edge $(i^* - 1, i^*) \in E$ (which exists because of (5.17)) must be traversed by the (longer) path $\check{H}$. Indeed, $\check{H}$ visits $i^* - 1$ at least once and i^* twice; there are two possibilities: either

(C1) $(i^* - 1, i^*)$ is traversed by $\check{H}$ on the way from $r+2$ to 0, or

(C2) $(i^* - 1, i^*)$ is traversed by $\check{H}$ on the way back from 0 to $r+1$.

In case (C2), define

$$\check{H}^\dagger := \check{H}_1 \cup \hat{H}_2 \cup \check{H}_3,$$

where $\check{H}_1$ is the part of $\check{H}$ from $r+2$ to i^*, $\hat{H}_2$ is the (reversed) part of $\hat{H}$ from i^* to 0 and back to $i^* - 1$, and $\check{H}_3$ is the part of $\check{H}$ from $i^* - 1$ to $r+1$. It is clear that the constructed path $\check{H}^\dagger$ visits point i^* twice and point i^{**} once. Case (C1) reduces to (C2) by exchanging vertices $r+1$ and $r+2$ and the directions of the paths $\hat{H}$ and $\check{H}$. Having constructed $\check{H}^\dagger$ with $\check{p}^\dagger = \check{p} - 1$, the construction of $\hat{H}^\dagger$ with $\hat{p}^\dagger = \hat{p} + 1$ (implying that $\check{p}^\dagger = \hat{p}^\dagger$) is obvious, as in the case $q_3(\mathcal{G}) = 0$. This proves the induction step $q_3(\mathcal{G}) = q - 1 \to q$ and the statement of the lemma in the case $q_4(\mathcal{G}) = 0$.

Finally, consider the case $q_4(\mathcal{G}) = 1$ (the argument below easily extends to the general case, $q_4(\mathcal{G}) \ge 1$). Let $1 \le i' \le r$ be a corresponding vertex with $\deg_+(i') = \deg_-(i') = 4$. In such a case, because of condition (5.18), the graph $\mathcal{G}$ is a union $\mathcal{G} = \mathcal{G}' \cup \mathcal{G}''$ of two components, $\mathcal{G}' = (V', E')$ and $\mathcal{G}'' = (V'', E'')$, where

$$V' := \{0, 1, \dots, i'\}, \qquad V'' := \{i'', i''+1, \dots, r+2\},$$
$$E' := \{(i, j) \in E : i, j \in V'\}, \quad \text{and} \quad E'' := \{(i, j) \in E : i, j \in V''\}.$$

Since $q_4(\mathcal{G}'') = 0$, the corresponding paths $\hat{H}''$ and $\check{H}''$ for the graph $\mathcal{G}''$ exist by the argument above. In a similar way, there exist paths $\hat{H}'$ and $\check{H}'$, both of which go from i' to 0 and back to i', satisfying all properties of the lemma except that r in (5.21)–(5.24) must be replaced by $i' - 1$ and the two 'outer' vertices $r+1$ and $r+2$ must be replaced by a single vertex i'. Constructing the paths $\hat{H}$ and $\check{H}$ from the 'components' $\hat{H}'$, $\check{H}'$, $\hat{H}''$, and $\check{H}''$ is obvious. Lemma 5.2 is proved.

Appendix B. A CLT for martingale transforms

Let $(\mathcal{H}_i)_{i\in\mathbb{Z}}$ be a filtration, and let $(U_i)_{i\in\mathbb{Z}}$ be a stationary sequence of the form

$$U_i = V_i \xi_i, \tag{B.1}$$

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

where the following conditions hold.

(U1) (ξ_i) is an i.i.d. sequence, adapted to $(\mathcal{H}_i)$ and such that ξ_s, $s > t$ are independent of $\mathcal{H}_t$ for any t. Moreover, $\mathrm{E}[\xi_0] = 0$ and the distribution of ξ_0 belongs to the domain of attraction of an α-stable law with index $1 < \alpha < 2$, in the sense that the characteristic function of ξ_0 in a neighborhood $|\theta| < \theta_0$ of $\theta = 0$ can be represented as (see also Ibragimov and Linnik (1971, Theorem 2.6.5))

$$\mathrm{E}[\exp(\mathrm{i}\theta\xi_0)] = \exp\left\{-c|\theta|^\alpha q(\theta)\left(1 - \mathrm{i}\beta\,\mathrm{sgn}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right)\right\} \tag{B.2}$$

for all $\theta \in \mathbb{R}$, $|\theta| < \theta_0$, where $\alpha \in (1, 2)$, $c > 0$, $\beta \in [-1, 1]$, $\theta_0 > 0$, and $q(\cdot)$ is continuous at 0:

$$\lim_{\theta\to 0} q(\theta) = q(0) = 1. \tag{B.3}$$

(U2) (V_i) is stationary, ergodic, and predictable (i.e. V_i is $\mathcal{H}_{i-1}$-measurable for any i), and

$$\mathrm{E}[|V_0|^r] < \infty \quad (\text{there exists } r > \alpha). \tag{B.4}$$

Theorem B.1. *Assume that conditions (U1) and (U2) are satisfied. Then*

$$n^{-1/\alpha}\sum_{i=1}^{[n\tau]} U_i \xrightarrow{\text{FDD}} \Lambda(\tau), \tag{B.5}$$

where $(\Lambda(\tau))_{\tau\in[0,1]}$ *is a homogeneous* α*-stable Lévy process with the characteristic function*

$$\mathrm{E}[\mathrm{e}^{\mathrm{i}\theta\Lambda(1)}] = \exp\left\{-c'|\theta|^\alpha\left(1 - \mathrm{i}\beta'\,\mathrm{sgn}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right)\right\}, \tag{B.6}$$

$$c' := c\,\mathrm{E}[|V_0|^\alpha], \qquad \beta' := \frac{\beta\,\mathrm{E}[|V_0|^\alpha\,\mathrm{sgn}(V_0)]}{\mathrm{E}[|V_0|^\alpha]}.$$

The proof of Theorem B.1, below, will be restricted to the 'one-dimensional' convergence at $\tau = 1$ in (B.5), since the convergence of general finite-dimensional distributions can be proved by a standard argument. As noted in Section 5, the proof uses the so-called 'principle of conditioning' due to Jakubowski (1986). The terminology and the formulation of this principle are taken from Kwapień and Woyczyński (1992). We will need the following definitions.

Definition B.1. (*Kwapień and Woyczyński (1992, Definition 4.3.1, p. 103).*) We say that two $(\mathcal{H}_i)$-adapted sequences (U_i) and (W_i) are *tangent* if, for each i, almost surely,

$$\mathrm{P}[U_i < c \mid \mathcal{H}_{i-1}] = \mathrm{P}[W_i < c \mid \mathcal{H}_{i-1}] \quad \text{for all } c \in \mathbb{R}.$$

Definition B.2. (*Kwapień and Woyczyński (1992, Definition 4.3.2, p. 104).*) An $(\mathcal{H}_i)$-adapted sequence (W_i) is said to satisfy the *conditional independence (CI) condition* if there exists a σ-field $\mathcal{G}$ such that, for each i, almost surely,

$$\mathrm{P}[W_i < c \mid \mathcal{H}_{i-1}] = \mathrm{P}[W_i < c \mid \mathcal{G}] \quad \text{for all } c \in \mathbb{R},$$

and such that (W_i) is a sequence of $\mathcal{G}$-conditionally independent random variables.

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

Theorem B.2. (The 'principle of conditioning' (Kwapień and Woyczyński (1992, Theorem 5.8.3, p. 144.)).) *Let $\{U_{nk} : n, k \in \mathbb{N}\}$ be an array of real random variables adapted to $\{\mathcal{H}_{nk}\}$, and, for each $n \in \mathbb{N}$, let $W_{n1}, W_{n2}, \dots$ be an $(\mathcal{H}_{ni})$-tangent sequence to $U_{n1}, U_{n2}, \dots$, which satisfies the CI condition with respect to a σ-field $\mathcal{G}_n$. Furthermore, define*

$$S_n = \sum_{k=1}^{\infty} U_{nk}, \qquad T_n = \sum_{k=1}^{\infty} W_{nk}.$$

Then

(i) *if the sequence* $(\mathrm{law}(T_n))$ *is compact then the sequence* $(\mathrm{law}(S_n))$ *is compact,*

(ii) *if* $\mathrm{law}(T_n \mid \mathcal{G}_n)$ *converges in probability to a nonrandom probability distribution μ on $\mathbb{R}$ with a nonvanishing characteristic function, then* $\mathrm{law}(S_n)$ *also converges to μ.*

Proof of Theorem B.1. Let (ξ_i^*) be an independent copy of (ξ_i), independent also of the sequence (V_i). This implies in particular that the ξ_i^*s are i.i.d. random variables in the domain of attraction of an α-stable law and having the same characteristic function in (B.2). Without loss of generality, we may assume that the sequence (ξ_i^*) is adapted to $(\mathcal{H}_i)$. Next, define, for $i, n \in \mathbb{N}$,

$$\begin{aligned}
W_i &:= V_i \xi_i^*, \\
U_{ni} &:= n^{-1/\alpha} U_i = n^{-1/\alpha} V_i \xi_i, \\
W_{ni} &:= n^{-1/\alpha} W_i = n^{-1/\alpha} V_i \xi_i^*, \\
S_n &:= \sum_{i=1}^{n} U_{ni} = n^{-1/\alpha} \sum_{i=1}^{n} U_i, \\
T_n &:= \sum_{i=1}^{n} W_{ni} = n^{-1/\alpha} \sum_{i=1}^{n} W_i, \\
\mathcal{G}_n &:= \sigma\{V_i,\ i \in \mathbb{Z}\}, \\
\mathcal{H}_{ni} &:= \mathcal{F}_i.
\end{aligned}$$

Let us check that (U_{ni}), (W_{ni}), $\mathcal{G}_n$, and $(\mathcal{H}_{ni})$, introduced above, satisfy the conditions of Theorem B.2. Definition B.1 (tangency of (U_{ni}) and (W_{ni})) is obvious by

$$\mathrm{P}[U_{ni} < c \mid \mathcal{H}_{n,i-1}] = \mathrm{P}[W_{ni} < c \mid \mathcal{H}_{n,i-1}] = F(cV_i n^{1/\alpha}),$$

where $F(c)$ is the (common) distribution function of the ξ_is and the ξ_i^*s. Definition B.2 (the CI condition) follows similarly since

$$\mathrm{P}[W_{ni} < c \mid \mathcal{H}_{n,i-1}] = F(cV_i n^{1/\alpha}) = \mathrm{P}[\xi_i^* < cV_i n^{1/\alpha} \mid \mathcal{G}_n],$$

by the independence of (ξ_i^*) and $\mathcal{G}_n$ defined above.

Let μ be the (stable) law of $\Lambda(1)$ in Theorem B.1. Note that the characteristic function $\hat{\mu}(\theta) = \mathrm{E}[\mathrm{e}^{\mathrm{i}\theta\Lambda(1)}]$ in (B.6) does not vanish: $|\hat{\mu}(\theta)| = \exp\{-c'|\theta|^\alpha\} \neq 0$ (for all $\theta \in \mathbb{R}$). Therefore, Theorem B.1 follows from Theorem B.2(ii) provided that we can verify the following

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

convergence: for any $\theta \in \mathbb{R}$,

$$\begin{aligned}\Phi_n &:= \mathrm{E}[\exp\{\mathrm{i}\theta T_n\} \mid \mathcal{G}_n] \\ &= \prod_{i=1}^{n} \mathrm{E}\left[\exp\left\{\frac{\mathrm{i}\theta \xi_i^* V_i}{n^{1/\alpha}}\right\} \,\middle|\, \mathcal{G}_n\right] \\ &\to \hat{\mu}(\theta) \quad \text{in probability.}\end{aligned} \tag{B.7}$$

In order to use the representation in (B.2) of the characteristic function of the ξ_i^*s (which is valid in a small neighborhood $\theta = 0$ only), we need to correspondingly restrict the magnitude of $|V_i|$, $i = 1, \dots, n$. Given $\theta_0 > 0$ in (B.2) and a $\theta \in \mathbb{R}$ in (B.7), let $K > 0$ be a (small) number such that

$$|\theta| K < \theta_0 \tag{B.8}$$

holds. To this end, consider two complementary events:

$$\begin{aligned}C'_{n,K} &:= \{\omega \in \Omega\colon |V_i| > K n^{1/\alpha} (\text{there exists } i = 1, \dots, n)\}, \\ C_{n,K} &:= \{\omega \in \Omega\colon |V_i| \le K n^{1/\alpha} (\text{for all } i = 1, \dots, n)\}.\end{aligned}$$

The probability of the first event tends to 0 for any fixed $K > 0$ according to condition (B.4) of the theorem. Indeed, since $r > \alpha$,

$$\begin{aligned}\mathrm{P}[C'_{n,K}] &\le \sum_{i=1}^{n} \mathrm{P}[|V_i| > K n^{1/\alpha}] \\ &\le \sum_{i=1}^{n} \frac{\mathrm{E}[|V_i|^r]}{K^r n^{r/\alpha}} \\ &= \frac{\mathrm{E}[|V_0|^r]}{n^{(r-\alpha)/\alpha} K^\alpha} \\ &\to 0 \quad \text{as } n \to \infty.\end{aligned} \tag{B.9}$$

Therefore, it suffices to verify (B.7) for Φ_n replaced by $\Phi_n \mathbf{1}(C_{n,K})$. Due to our choice of K in (B.8), we can use representation (B.2) in the product in (B.7) and rewrite $\Phi_n \mathbf{1}(C_{n,K})$ as

$$\begin{aligned}\Phi_n \mathbf{1}(C_{n,K}) &= \exp\left\{-\frac{c|\theta|^\alpha}{n} \sum_{i=1}^{n} |V_i|^\alpha q\left(\frac{\theta V_i}{n^{1/\alpha}}\right)\left(1 - \mathrm{i}\beta \operatorname{sgn}(\theta V_i) \tan\left(\frac{\pi\alpha}{2}\right)\right)\right\} \mathbf{1}(C_{n,K}) \\ &= \exp\{-c|\theta|^\alpha \Omega_{1n}\} \exp\{-c|\theta|^\alpha \Omega_{2n}\} \mathbf{1}(C_{n,K}),\end{aligned} \tag{B.10}$$

where

$$\begin{aligned}\Omega_{1n} &:= \frac{1}{n} \sum_{i=1}^{n} |V_i|^\alpha - \mathrm{i}\beta \tan\left(\frac{\pi\alpha}{2}\right) \operatorname{sgn}(\theta) \frac{1}{n} \sum_{i=1}^{n} |V_i|^\alpha \operatorname{sgn}(V_i), \\ \Omega_{2n} &:= \frac{1}{n} \sum_{i=1}^{n} |V_i|^\alpha \left(q\left(\frac{\theta V_i}{n^{1/\alpha}}\right) - q(0)\right)\left(1 - \mathrm{i}\beta \operatorname{sgn}(\theta V_i) \tan\left(\frac{\pi\alpha}{2}\right)\right).\end{aligned}$$

By the law of large numbers (ergodicity of the sequences $(|V_i|^\alpha)$ and $(|V_i|^\alpha \operatorname{sgn}(V_i))$, we have

$$\frac{1}{n} \sum_{i=1}^{n} |V_i|^\alpha \to \mathrm{E}[|V_0|^\alpha] \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^{n} |V_i|^\alpha \operatorname{sgn}(V_i) \to \mathrm{E}[|V_0|^\alpha \operatorname{sgn}(V_0)] \quad \text{in probability.}$$

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

Therefore, for any $\theta \in \mathbb{R}$,

$$\exp\{-c|\theta|^\alpha \Omega_{1n}\} \to \hat{\mu}(\theta) \quad \text{in probability.} \tag{B.11}$$

Also, note that

$$|\Omega_{2n}|\mathbf{1}(C_{n,K}) \le \sqrt{1+\tan^2\left(\frac{\pi\alpha}{2}\right)}\Big(\sup_{|z|<K}|q(z)-q(0)|\Big)\frac{1}{n}\sum_{i=1}^{n}|V_i|^\alpha$$

can be made arbitrarily small in probability uniformly in $n > n_0$, by using condition (B.3) and choosing K small enough. The last conclusion obviously applies also to the difference $\exp\{-c|\theta|^\alpha \Lambda_{2n}\}\mathbf{1}(C_{n,K}) - 1$. With (B.9), (B.10), and (B.11) in mind, this completes the proof of (B.7) and the proof of Theorem B.1 too.

Acknowledgements

The author is grateful to Adam Jakubowski and Olav Kallenberg for comments concerning Theorem B.1, and the two anonymous referees for their comments. The research was partially supported by the bilateral France–Lithuania scientific project Gilibert and the Lithuanian State Science and Studies Foundation grant T-25/08.

References

Beran, J. (2006). On location estimation for LARCH processes. *J. Multivariate Anal.* **97,** 1766–1782.

Berkes, I. and Horváth, L. (2003). Asymptotic results for long memory LARCH sequences. *Ann. Appl. Prob.* **13,** 641–668.

Black, F. (1976). Studies in stock price volatility changes. In *Proc. 1976 Meeting Amer. Statist. Assoc., Business Econom. Session,* American Statistical Association, Alexandria, VA, pp. 177–181.

Davis, R. A. and Mikosch, T. (1998). The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Ann. Statist.* **26,** 2049–2080.

Davis, R. A. and Mikosch, T. (2000). The sample autocorrelations of financial time series models. In *Nonlinear and Nonstationary Signal Processing*, eds W. J. Fitzgerald *et al.*, Cambridge University Press, pp. 247–274.

Davis, R. and Resnick, S. (1985a). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Prob.* **13,** 179–195.

Davis, R. and Resnick, S. (1985b). More limit theory for the sample correlation function of moving averages. *Stoch. Process. Appl.* **20,** 257–279.

Davis, R. and Resnick, S. (1986). Limit theory for the sample covariance and correlation functions moving averages. *Ann. Statist.* **14,** 533–558.

Davis, R. A. and Resnick, S. I. (1996). Limit theory for bilinear processes with heavy-tailed noise. *Ann. Appl. Prob.* **6,** 1191–1210.

Doukhan, P., Teyssière, G. and Winant, P. (2006). A LARCH(∞) vector valued process. In *Dependence in Probability and Statistics* (Lecture Notes Statist. **187**), eds. P. Bertail *et al.*, Springer, Berlin, pp. 245–258.

Giraitis, L. and Surgailis, D. (2002). ARCH-type bilinear models with double long memory. *Stoch. Process. Appl.* **100,** 275–300.

Giraitis, L., Leipus, R. and Surgailis, D. (2008). ARCH(∞) models and long memory properties. To appear in *Handbook of Financial Time Series*.

Giraitis, L., Robinson, P. M. and Surgailis, D. (2000). A model for long memory conditional heteroscedasticity. *Ann. Appl. Prob.* **10,** 1002–1024.

Giraitis, L., Leipus, R., Robinson, P. M. and Surgailis, D. (2004). LARCH, leverage and long memory. *J. Financial Econometrics* **2,** 177–210.

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

Jakubowski, A. (1986). Principle of conditioning in limit theorems for sums of random variables. *Ann. Prob.* **14,** 902–913.

Kwapień, S. and Woyczyński, W. A. (1992). *Random Series and Stochastic Integrals: Single and Multiple*. Birkhäuser, Boston, MA.

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press

Mikosch, T. (2003). Modelling dependence and tails of financial time series. In *Extreme Values in Finance, Telecommunications and the Environment*, eds B. Finkelstädt and H. Rootzén, Chapman & Hall/CRC, Boca Raton, FL, pp. 185–286.

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

Sentana, E. (1995). Quadratic ARCH models. *Rev. Econom. Stud.* **3,** 77–102.

Stout, W. F. (1974). *Almost Sure Convergence.* Academic Press, New York.

Vaičiulis, M. (2003). Convergence of sums of Appell polynomials with infinite variance. *Lithuanian Math. J.* **43,** 67–82.

Von Bahr, B. and Esséen, C.-G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.* **36,** 299–303.

https://doi.org/10.1239/aap/1231340170 Published online by Cambridge University Press