LARCH, leverage, and long memory

Liudas Giraitis, Remigijus Leipus, Peter M. Robinson and Donatas Surgailis

This is an electronic version of an Article published in the Journal of financial econometrics, 2 (2). pp. 177-210 © 2004 Oxford University Press.
http://www.jfec.oupjournals.org

LARCH, Leverage and Long Memory

Liudas Giraitis[1]*, Remigijus Leipus[2,3], †Peter M. Robinson[1‡]
and Donatas Surgailis[3]
[1]London School of Economics
[2]Vilnius University
[3]Vilnius Institute of Mathematics and Informatics

August 18, 2003

Abstract

We consider the long memory and leverage properties of a model for the conditional variance V_t^2 of an observable stationary sequence X_t, where V_t^2 is the square of an inhomogeneous linear combination of X_s, $s < t$, with square summable weights b_j. This model, which we call linear ARCH (LARCH), specializes, when V_t^2 depends only on X_{t-1}, to the asymmetric ARCH model of Engle (1990), and, when V_t^2 depends only on finitely many X_s, to a version of the quadratic ARCH model of Sentana (1995), these authors having discussed leverage potential in such models. The model which we consider was suggested by Robinson (1991), for use as a possibly long memory conditionally heteroscedastic alternative to i.i.d. behaviour, and further studied by Giraitis, Robinson and Surgailis (2000), who showed that integer powers X_t^ℓ, $\ell \geq 2$, can have long memory autocorrelations. We establish conditions under which the cross-autocovariance function between volatility and levels, $h_t = \mathrm{Cov}(V_t^2, X_0)$, decays in the manner of moving average weights of long memory processes on suitable choice of the b_j. We also establish the leverage property that $h_t < 0$ for $0 < t \leq k$, where the value of k (which may be infinite) again depends on the b_j. Conditions for finiteness of third and higher moments of X_t are also established.

JEL Classification: C22

Corresponding Author: Donatas Surgailis

We thank two referees and the Associate Editor for helpful comments.

*Research supported by ESRC Grant R000238212.

†Research supported by Lithuanian State Science and Studies Foundation Grant K-014.

‡Research supported by a Leverhulme Trust Personal Research Professorship and ESRC Grant R000238212.

1 Introduction

Considerable activity has centred on modelling the dependence structure of asset returns. Empirical evidence suggests that these may have little or no autocorrelation, but are far from independent. One empirical observation, due to Black (1976), is the leverage effect, 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. As a related finding, nonlinear functions such as squares or absolute values can be notably autocorrelated. So far as squares are concerned, this arises if the series has conditional heteroscedasticity, so that not only can substantial autocorrelation at short lags be detected, but also such slow decay as lag length increases that there is said to be long memory conditional heteroscedasticity. In empirical studies this latter possibility was recognized as early as Whistler (1990), who applied to exchange rate series tests for independence that are directed against the alternative of long memory autocorrelation in squares.

Denote by X_t, $t = 0, \pm 1, ...$, the observable series (of asset returns, for example), assumed strictly stationary, such that $E|X_0|^3 < \infty$, and define the conditional variance

$$V_t^2 = \operatorname{Var}\left(X_t \,|\mathcal{G}_{t-1}\right), \qquad t = 0, \pm 1, ..., \tag{1.1}$$

where $\mathcal{G}_t$ denotes the σ-field of events generated by X_s, $s \leq t$. To measure leverage, define the function

$$h_t = \operatorname{Cov}\left(V_t^2, X_0\right), \qquad t \geq 1. \tag{1.2}$$

Alternative measures may be used, with V_t^2 replaced by other increasing functions of $|V_t|$, but (1.2) proves mathematically the most tractable. We shall say that X_t has leverage of order k ($X_t \in \ell(k)$), $1 \leq k < \infty$, if and only if

$$h_j < 0, \qquad 0 < j \leq k. \tag{1.3}$$

We shall also consider the long memory property

$$h_t \sim C t^{d-1}, \qquad C \neq 0, \quad 0 < d < \frac{1}{2}, \tag{1.4}$$

as $t \to \infty$, where "$\sim$" indicates that the ratio of left and right sides tends to 1 as $t \to \infty$. From other experience with time series analysis, it is easy to understand that both the leverage and long memory properties (1.3) and (1.4) can arise, because by nested conditional expectations

$$h_t = \operatorname{Cov}\left(X_t^2, X_0\right), \qquad t \geq 1, \tag{1.5}$$

if

$$E\left(X_t \,|\mathcal{G}_{t-1}\right) = 0, \text{ a.s.} \tag{1.6}$$

Thus h_t is simply the cross-autocovariance function between the levels X_t and squares X_t^2. Long memory in scalar and vector time series is familiar, as are negative autocovariances and cross-autocovariances.

To provide some evidence of the possibility of leverage and long memory in financial data, Figure 1 displays the sample cross-autocorrelation between levels and future squares (solid line) and the sample autocorrelation of squares (dashed line), for 900 S&P500 daily returns beginning in 1928. No interval estimates (such as ones based on a null hypothesis of independent and identically distributed X_t) are presented, but Figure 1 seems suggestive of a leverage effect at low lags, with some tendency for negative values of the estimated h_t to persist (with oscillation), as (1.4) also predicts. Of course this behaviour could have other sources, but a negative h_t, at finitely or infinitely many t, the slow decay of (1.4), and oscillation, are features which can be described by the 'LARCH' model class which will form the focus of the current paper. Giraitis, Robinson and Surgailis (2000) demonstrated the ability of LARCH to explain long memory decay (at rate t^{2d-1} for the same d as in (1.4)) in autocorrelations of squares, and again Figure 1 provides some evidence of this, though there are clearly other effects to be explained also.

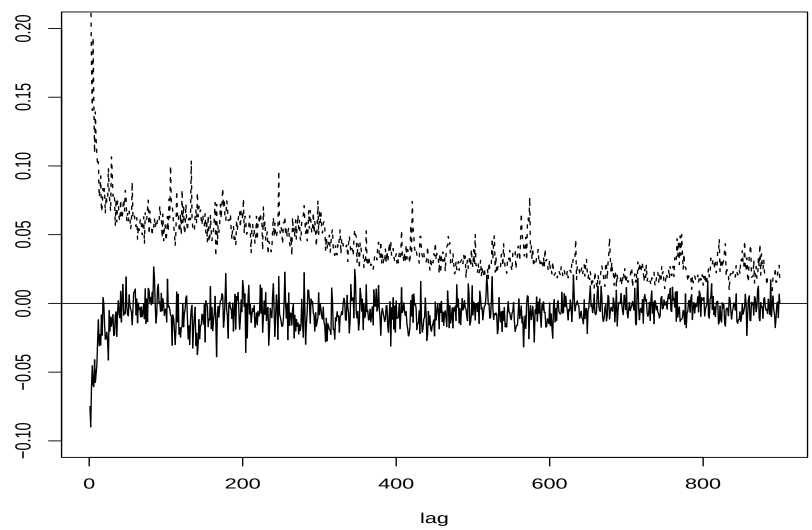

FIGURE 1

In fact, some other models have previously demonstrated to have both long memory conditional heteroscedasticity and leverage properties. Consider the model

$$X_t = \zeta_t\Big(a + \sum_{j=1}^{\infty} b_j \zeta_{t-j}\Big), \tag{1.7}$$

where

$$\zeta_t \text{ are i.i.d., } E\zeta_0 = 0, \text{ Var}\zeta_0 = 1, \tag{1.8}$$

and a and the b_t, $t \geq 1$, are non-stochastic with

$$\sum_{t=1}^{\infty} b_t^2 < \infty. \tag{1.9}$$

The non-linear moving average (MA) model (1.7) has the immediate property

$$\text{Cov}(X_0, X_t) = 0, \quad t \geq 1, \tag{1.10}$$

often believed true of asset returns, and was considered by Robinson and Zaffaroni (1997), who showed that if

$$b_t \sim ct^{d-1}, \qquad 0 < d < \frac{1}{2}, \quad 0 < c < \infty, \tag{1.11}$$

and also $E\zeta_0^4 < \infty$, then

$$\text{Cov}(X_0^2, X_t^2) \sim c' t^{2d-1}, \qquad 0 < c' < \infty, \quad a \neq 0, \tag{1.12}$$

$$\text{Cov}(X_0^2, X_t^2) \sim c'' t^{4d-2}, \qquad 0 < c'' < \infty, \quad a = 0, \tag{1.13}$$

as $t \to \infty$. The property (1.11) is characteristic of MA weights in long memory models, while the decay in (1.12) is consistent with the asymptotic autocovariance behaviour in such models; in case (1.13), there is long memory only when $\frac{1}{4} < d < \frac{1}{2}$. We can achieve (1.11) by, for example, taking

$$b_t = c_0 r_t(d), \quad r_t(d) = \frac{\Gamma(t+d)}{\Gamma(d)\Gamma(t+1)}, \quad c_0 > 0, \tag{1.14}$$

so the $r_t(d)$ are coefficients in the formal expansion

$$(1-z)^{-d} = 1 + \sum_{t=1}^{\infty} r_t(d) z^t \tag{1.15}$$

and the b_t are proportional to weights in the fractional autoregressive integrated moving average FARIMA(0,d,0) model of Adenstedt (1974) (see also Palma and Zevallos (2002) for generalizations). On the other hand, considering for simplicity only the case $E\zeta_0^3 = 0$, we have (cf. Zaffaroni (1998))

$$h_t = 2ab_t\Big(1 + \sum_{s=1}^{\infty} b_{t+s} b_s\Big), \qquad t \geq 1.$$

Thus under the sufficient conditions $\sum_{t=1}^{\infty} b_t^2 < 1$, or that all b_t have the same sign, $X_t \in \ell(k)$ if and only if

$$ab_t < 0, \qquad 1 \leq t \leq k, \tag{1.16}$$

that is if all b_t, $1 \leq t \leq k$, have the same sign and this differs from a's. It is possible to achieve this effect by choosing the b_t as weights in certain autoregressive integrated moving average (ARMA) models. Moreover, for b_t satisfying (1.11), for example (1.14), the long memory property (1.4) of h_t immediately follows. To estimate (1.7), with b_j depending on a finite dimensional vector of parameters, Robinson and Zaffaroni (1997) proposed a form of Whittle estimation based on the X_t^2, and Zaffaroni (1998) provides a central limit theorem for such estimates. One may then infer long memory if a test of $d = 0$ is rejected against positive alternatives, or infer leverage if a test of $b_j = 0$, $1 \leq j \leq k$, some k, is rejected against negative alternatives. The model (1.7) might be extended by, for example, replacing the second factor by some nonlinear function, cf. the stochastic volatility model of Taylor (1986).

Nonlinear MA models face the criticism, however, of being difficult to use in forecasting, being possibly non-invertible over a large portion of the parameter space, and having a likelihood that is relatively intractable. An alternative popular class that meets the above objections (albeit suffering other disadvantages) commences from functional forms for the first two conditional moments, $E(X_t | \mathcal{G}_{t-1})$, such as (1.6) (which implies (1.10)), and V_t^2 (1.1). Some popular choices of V_t^2 are special cases of

$$V_t^2 = a + \sum_{j=1}^{\infty} b_j X_{t-j}^2, \tag{1.17}$$

where $a \geq 0$, $b_j \geq 0$, and covariance stationarity of X_t implies the identity

$$a = EX_0^2\Big(1 - \sum_{t=1}^{\infty} b_t\Big). \tag{1.18}$$

The ARCH(p) model of Engle (1982) takes $b_t = 0$, $t > p$, in (1.17), the GARCH(p, q) model of Bollerslev (1986) entails exponentially decaying b_t, while the general "ARCH(∞)" form

(1.17) was considered by Robinson (1991) in connection with hypothesis testing. Robinson (1991), Ding and Granger (1996), also considered the possibility of long memory in squares resulting from (1.6), (1.17), as from taking $b_t = -r_t(-d)$ (see (1.14)) so that from (1.15)

$$V_t^2 = a + \left(1 - (1-L)^d\right) X_t^2, \tag{1.19}$$

where L denotes the lag operator. In fact (1.18) then implies $a = 0$, whence (1.19) corresponds to the FARIMA(0,d,0) model for X_t^2,

$$(1-L)^d X_t^2 = X_t^2 - V_t^2, \tag{1.20}$$

where the $X_t^2 - V_t^2$ are (conditionally heteroscedastic, given $EX_0^4 < \infty$) martingale differences. However (1.19) does not satisfy the sufficient conditions developed by Giraitis, Kokoszka and Leipus (2000) for a covariance stationary solution X_t of the equations

$$X_t = \zeta_t V_t, \quad t = 0, \pm 1, \ldots, \tag{1.21}$$

with V_t given by the positive square root of (1.17) and ζ_t satisfying (1.8); (1.21) satisfies (1.6) and thus (1.10). On the other hand, Baillie, Bollerslev and Mikkelsen (1996) consider a "FIGARCH" modification of (1.19) (allowing also for an ARMA factor) but with $a > 0$ (so (1.18) is not satisfied) whence a is added to the right side of (1.20) and X_t does not have finite variance for any $d > 0$.

Unlike (1.7), none of these ARCH-in-squares models is capable of explaining the leverage effect. Although some existing ARCH-type models, e.g. Glosten *et al.* (1993), Zakoian (1994), Müller *et al.* (1997), Schwert (1990), allow modelling of the leverage property, there is limited experience in extending them to model long memory properties such as (1.4) or (1.12). Another important class used for modelling asymmetry in financial data consists of exponential ARCH (EGARCH) models (see Nelson (1991), Karanasos and Kim (2001), He *et al.* (2002) for the properties of EGARCH process). Long memory stochastic volatility models were explored by Harvey (1998), Breidt *et al.* (1998), Comte and Renault (1998). Demos (2002) studied a model which nests both EGARCH and the stochastic volatility specification. Some of the above models, together with the FIEGARCH model of Bollerslev and Mikkelsen (1996) and FIAPARCH model of Tse (1998), have potential to explain both leverage and long memory but their theoretical properties are not established. Rigorous mathematical study of exponential models covering both the long memory and leverage effects can be found in Surgailis and Viano (2003). Finally note, that the slowly decaying component in the leverage function was advocated by Bouchaud *et al.* (2001), who investigated quantitatively the leverage effect for individual stocks and stock indices, and introduced the so-called "retarded volatility" model. Pagan (1996, p.30–31) also stressed the persistence of the leverage effect for some stock data.

One ARCH-type model for which long memory capability has already been established is the linear ARCH (LARCH) model suggested by Robinson (1991), which replaces (1.17) by

$$V_t^2 = \Big(a + \sum_{j=1}^{\infty} b_j X_{t-j}\Big)^2. \tag{1.22}$$

Note that (1.22) is satisfied by both

$$V_t = \Big|a + \sum_{j=1}^{\infty} b_t X_{t-j}\Big| \tag{1.23}$$

and

$$V_t = a + \sum_{j=1}^{\infty} b_t X_{t-j}. \tag{1.24}$$

In (1.23), but not (1.24), V_t has the familiar interpretation of a standard deviation, but (1.24) is mathematically the more tractable form, and unlike (1.17), requires no restrictions on the b_j to ensure nonnegativity of V_t^2. Short memory versions of (1.22) are given by

$$\text{“LARCH}(p)\text{”:} \qquad V_t = a + \sum_{j=1}^{p} b_j X_{t-j}, \tag{1.25}$$

to correspond to the ARCH(p) structure of Engle (1982), and

$$\text{“GLARCH}(p,q)\text{”:} \qquad V_t - \sum_{j=1}^{q} \beta_j V_{t-j} = a' + \sum_{j=1}^{p} \alpha_j X_{t-j}, \tag{1.26}$$

to compare with the GARCH(p,q) structure of Bollerslev (1986), where $a = a'/(1-\sum_{j=1}^{q}\beta_j)$ in (1.24). Indeed with $p = 1$, (1.25) first arose in Engle (1990), who considered a model for V_t^2 containing an additive constant, a term in X_{t-1}, and a term in $|X_{t-1}|^\theta$, for unknown θ, and then estimated θ from real data, noting the consequence of $\theta = 2$; see also Engle and Ng (1993). On supplementing (1.24) by (1.21) and imposing (1.8), (1.9), we have a kind of nonlinear AR dual of the nonlinear MA (1.7):

$$X_t = \zeta_t(a + \sum_{j=1}^{\infty} b_j X_{t-j}). \tag{1.27}$$

Of course, (1.27) satisfies the uncorrelatedness-in-levels assumption (1.10). So far as leverage properties are concerned, Robinson (1991) noted that for (1.27) h_t can be non-zero even if $E\zeta_0^3 = 0$, unlike in (1.17), while Engle (1990), Campbell and Hentschel (1992), Engle and Ng (1993) and Sentana (1995) explicitly discussed leverage and other asymmetry in models that overlap with (1.27); see also Barndorff-Nielsen and Shephards' (2001) discussion of leverage of a related continuous-time model. Sentana (1995) considered a number of issues relating to a model in which V_t^2 is a more general quadratic function of $X_{t-1}, \ldots, X_1$ than in our LARCH(p), along with a form of GARCH extension which, however, differs from our (1.26). The first of these latter types of model is given by

$$V_t^2 = \theta + \sum_{j=1}^{p} \psi_j X_{t-j} + \sum_{j=1}^{p}\sum_{k=1}^{p} \phi_{jk} X_{t-j} X_{t-k}, \tag{1.28}$$

where the parameters θ, ψ_j, ϕ_{jk} can vary freely and need not necessarily satisfy the constraints implicit in (1.25), so that (1.28) nests both the ARCH(p) model of Engle (1982) and (1.25). Whereas non-negativity of V_t^2 given by (1.25) is automatic, Sentana derived conditions for non-negativity in (1.28). Relative to (1.28), (1.27) has an advantage of parsimony, but at the same time it suffers from less flexibility, which may be an important drawback in modelling; for example in the LARCH(1), non-rejection of a test for $b_1 = 0$ suggests lack of both conditional heteroscedasticity and leverage, so we cannot examine both phenomena individually or conveniently interpret parameters as contributing primarily to one or the other. Sentana's interest in (1.28) was strongly motivated by a desire to explain leverage and other asymmetry, and he described conditions similar to some of ours (see Theorem 2.4 below) in case $E\zeta_0^3 = 0$, but taking for granted stationarity and existence of moments of X_t, aspects which we justify under primitive conditions on the ζ_t and b_t. Though Sentana referred to interest in long memory, he did not discuss its achievement in his models. Sentana also discussed the estimation and testing of his models, conditions for stationarity of his GARCH extension of (1.28), and multivariate extensions of his models, also exploring their ability to empirically explain leverage and other features of data.

Despite the appeal of Sentana's (1995) extension (1.28) of LARCH(p) and partial extension of LARCH(∞), a serious practical study of such models conveniently commences by focussing on LARCH(∞) (1.27), which has an aesthetically simple form with its automatic non-negativity of V_t^2, and significantly different mathematical structure from ARCH(∞) (1.17). Indeed, Giraitis, Robinson and Surgailis (2000) already gave conditions under which, for b_t given by (1.11), integer powers X_t^l, $l \geq 2$, of LARCH(∞) X_t have long memory autocorrelation, for example (1.12) holds. In the present paper we establish conditions for the alternative long memory property (1.4) in LARCH(∞), which (unlike (1.12)) is possible without finiteness of $X_t's$ fourth moment, and conditions for leverage (1.3). We assume in (1.27) that (1.8), (1.9) and also

$$a \neq 0 \tag{1.29}$$

hold; if $a = 0$ the model has a trivial character, as seen in Theorem 2.1 below. Theorem 2.1 provides conditions for existence and uniqueness of a stationary solution of (1.27). For the leverage and long memory properties, additional conditions are required on moments of ζ_t and on the b_t; we also give primitive sufficient conditions for finiteness of third and higher moments of X_t. In case of the simple LARCH(1) and GLARCH(1,1) models, we compare the leverage conditions we have obtained under a general LARCH(∞) model with ones that directly exploit the special structure of these models, also obtaining explicit formulae for h_t in these cases. Note that Sentana (1995) observes that "linear ARCH" is also used for the model (1.17), so our "LARCH" terminology is not ideal; there is, however, a plethora of existing ARCH acronyms and we have been unable to propose a simple alternative.

The following section presents conditions and results, while the proofs are developed in Appendices. Section 3 includes some final comments.

2 Main Results

Implicit in our results for h_t is the requirement that X_t have at least finite third moment. We begin by extending our conditions (1.8), (1.9) on ζ_t and the b_j in the LARCH model of the previous section for finite third or fourth moments (as is relevant to the results of Giraitis, Robinson and Surgailis (2000)), and then for finite even moments of any order; in general finiteness of $E\left|X_0\right|^r$ entails finiteness of $E\left|\zeta_0\right|^r$, so the latter case is relevant to Gaussian ζ_t, a possibility earlier stressed in estimation of (1.7) and (1.17).

Write $\|f\|_p = \left(\sum_{t=1}^{\infty}|f_t|^p\right)^{1/p}$ for $p > 1$ and a sequence $\{f_t; t \geq 1\}$; for brevity write $\|f\| = \|f\|_2$; as indicated by (1.9), $\|b\| < \infty$. Put $\mu_k = E\zeta_0^k$, $|\mu|_k = E|\zeta_0|^k$, $k \geq 1$; as indicated by (1.8), $\mu_1 = 0$, $\mu_2 = 1$. Define $\mathbf{Z} = \{0, \pm 1, \ldots\}$.

A preliminary result that is important to all that follows concerns the existence and nature of a unique stationary solution of (1.27). Let $\mathcal{F}_t$ be the $\sigma-$field of events generated by $\zeta_s, s \leq t$.

Definition 2.1: An $\mathcal{F}_t-$measurable sequence $X_t, t \in \mathbf{Z}$ is called a solution of (1.27) if

$$\sup_t EX_t^2 < \infty \tag{2.1}$$

and (1.27) holds for each $t \in \mathbf{Z}$. Note that (2.1) and $\|b\| < \infty$ imply that (1.24) converges in L^2.

Theorem 2.1: *(i) Let $a \neq 0$. Then a solution of* (1.27) *exists if and only if $\|b\| < 1$, in which case there is a unique covariance and strictly stationary solution for V_t given by the Volterra series*

$$V_t = a\Big(1 + \sum_{k=1}^{\infty} \sum_{s_k < \ldots < s_1 < t} b_{t-s_1} b_{s_1-s_2} \ldots b_{s_{k-1}-s_k} \zeta_{s_1} \ldots \zeta_{s_k}\Big). \tag{2.2}$$

(*ii*) *Let* $a = 0$ *and* $\|b\| < \infty$. *Then* (1.27) *admits a unique solution* $X_t = 0$ *a.s.*

Theorem 2.1 extends Theorem 2.1 of Giraitis, Robinson and Surgailis (2000) by demonstrating uniqueness in case $a \neq 0$, and by addressing also the case $a = 0$. All the results of the paper are readily extendable to allow for a nonstochastic, non-zero conditional mean of X_t, so that an additive constant is included in (1.27), leading to a constant, non-zero right hand side in (1.6).

Our results for h_t require higher moment conditions on ζ_t, and stronger conditions on b_t.

Assumption M$_3$: The third absolute moment $|\mu|_3 < \infty$ and

$$|\mu|_3^{1/3} \|b\|_3 + 3\theta \|b\| < 1, \tag{2.3}$$

where $\theta \approx 1.27$ is the solution of $3\theta^2 - 3\theta - 1 = 0$.

Proposition 2.1: *Under* (1.8), (1.9), (1.27), (1.29) *and assumption* M_3,

$$E|V_0|^3 < \infty, \quad E|X_0|^3 < \infty.$$

In view of frequently-expressed scepticism concerning the finiteness of fourth moments of much financial data, there is interest in Proposition 2.1, a finite third moment being a minimal condition for analysis of h_t (though statistical inference on h_t is liable to entail finiteness of at least the sixth moment of X_t). There is a trade-off between moment conditions and restrictions on the b_t, so we consider also:

Assumption M$_4$: The fourth moment $\mu_4 < \infty$ and

$$\mu_4 \|b\|_4^4 + 4|\mu_3| \|b\|_3^3 + 6\|b\|^2 < 1. \tag{2.4}$$

Proposition 2.2: *Under* (1.8), (1.9), (1.27), (1.29) *and assumption* M_4,

$$EV_0^4 < \infty, \quad EX_0^4 < \infty. \tag{2.5}$$

Assumption M$_{2k}$: For $k \geq 3$ the $2k$th moment $\mu_{2k} < \infty$ and

$$\sum_{p=2}^{2k} \binom{2k}{p} \|b\|_p^p |\mu_p| < 1. \tag{2.6}$$

Proposition 2.3: *Under* (1.8), (1.9), (1.27), (1.29) *and assumption* M_{2k}, $k \geq 3$,

$$EV_0^{2k} < \infty, \quad EX_0^{2k} < \infty. \tag{2.7}$$

Propositions 2.1 and 2.2 are proved in Appendix B; the proof of Proposition 2.3 is a development of these and is omitted as this proposition is not important to the rest of the paper. Assumption M$_4$ is weaker than the condition $11|\mu|_4^{1/2}\|b\|^2 < 1$ for (2.5) obtained in Giraitis, Robinson and Surgailis (2000) although it is not necessary. It should be noted that the question of finiteness of the third and other odd absolute moments of the LARCH model (and more general Volterra series) is more delicate than that of even moments; see Appendix B. Assumption M_{2k} is weaker than the condition $\left(4^k - 2k - 1\right) |\mu|_{2k}^{1/k} \|b\|^2 < 1$ for (2.7) obtained in Giraitis, Robinson and Surgailis (2000), indeed, as $\|b\|_p \le \|b\|$ and $\|b\|^2 \mu_{2k} \le 1$, so $\|b\|_p^p |\mu_p| \le (\|b\|^2 \mu_{2k}^{1/k})^{p/2} \le \|b\|^2 \mu_{2k}^{1/k}$ and therefore

$$\sum_{p=2}^{2k} \binom{2k}{p} \|b\|_p^p |\mu_p| \le \|b\|^2 \mu_{2k}^{1/k} \sum_{p=2}^{2k} \binom{2k}{p} = \left(4^k - 2k - 1\right) \mu_{2k}^{1/k} \|b\|^2 .$$

When the distribution of ζ_0 is unknown, the bounds (2.3), (2.4) and (2.6) cannot be used in practice. If, on the other hand, ζ_t is known to be Gaussian they may be evaluated using $|\mu|_3 = (8/\pi)^{1/2}$, $\mu_{2k} = (2k-1)(2k-3)\cdots 1$, $k \ge 1$, $\mu_k = 0$, k odd. Indeed, using also $\|b\|_p \le \|b\|$, $p \ge 2$ we can get the simplified sufficient conditions

$$\|b\| \le \left((8/\pi)^{1/6} + 3\theta\right)^{-1} \approx .2008 \tag{2.8}$$

for M_3 and

$$\|b\| < \sqrt{0.1547} \approx .3933 \tag{2.9}$$

for M_4. For example, the corresponding assumption to M_4 of Giraitis, Robinson and Surgailis (2000) gives $\|b\| \le .229$. Note that in case of the LARCH(1), $\|b\|_p = \|b\|$, for all $p \ge 2$, so (2.8) and (2.9) are precise versions of (2.3) and (2.4), respectively.

We now go on to study h_t directly. We assume either M_3 or M_4 holds in all that follows, implying in particular that $\|b\| < 1$. Then for $|z| \le 1$ we may define

$$\Phi(z) := \sum_{t=0}^{\infty} \phi_t z^t = \left(1 - \sum_{t=1}^{\infty} b_t^2 z^t\right)^{-1}, \tag{2.10}$$

so that

$$\phi_t := b_t^2 + \sum_{k=1}^{t-1} \sum_{0<s_k<\ldots<s_1<t} b_{s_k}^2 b_{s_{k-1}-s_k}^2 \cdots b_{s_1-s_2}^2 b_{t-s_1}^2, \quad (t \ge 1), \quad \phi_0 = 1. \tag{2.11}$$

Now, denoting

$$\sigma^2 = \mathrm{Var} X_t = \frac{a^2}{1 - \|b\|^2}, \qquad [b]_3 = \sum_{j=1}^{\infty} b_j^3, \tag{2.12}$$

introduce

$$g_t := 2a\sigma^2 \sum_{s=1}^{t} b_s \phi_{t-s} \quad (t \ge 1), \quad g_0 := a\mu_3 \left(a^2 + 3\sigma^2 \|b\|^2\right), \tag{2.13}$$

$$r_{tu} := 2 \sum_{s=1}^{t} b_s b_{s+u} \phi_{t-s} \quad (t, u \ge 1), \quad r_{t0} := \phi_t \quad (t \ge 1), \tag{2.14}$$

$$r_{0u} := 3\mu_3 \sum_{s=1}^{\infty} b_s^2 b_{s+u} \quad (u \ge 1), \quad r_{00} := \mu_3 [b]_3. \tag{2.15}$$

Then introduce h_t', $t \ge 1$, to be the unique square-summable solution of

$$h_t' = \phi_t + \sum_{u=1}^{\infty} r_{tu} h_u', \qquad t \ge 1; \tag{2.16}$$

such $\{h'_t\}$ exists because

$$\Big(\sum_{t,u=1}^{\infty} r_{tu}^2\Big)^{1/2} \le \frac{2\,\|b\|^2}{1-\|b\|^2} < 1, \tag{2.17}$$

as will be shown in the proof of the following theorem.

Theorem 2.2: *Let* (1.8), (1.9), (1.27), (1.29) *and either assumption* M_3 *or* M_4 *hold. Then*

$$r_{00} + \sum_{u=1}^{\infty} r_{0u} h'_u \ne 1 \tag{2.18}$$

is a necessary and sufficient condition for uniqueness of a solution $\{h_t, t \ge 0\}$ *satisfying*

$$\sum_{u=0}^{\infty} h_u^2 < \infty \tag{2.19}$$

of the linear equations

$$h_t = g_t + \sum_{u=0}^{\infty} r_{tu} h_u, \quad t \ge 0. \tag{2.20}$$

Condition (2.18) is automatically satisfied if $\mu_3 = 0$, because then (2.15) implies $r_{0u} = 0$, $u \ge 0$. A more general condition, obtained in Lemma A.1 below, is

$$|\mu_3|\,\|b\|^3 + 3\,\|b\|^2 < 1. \tag{2.21}$$

We now go on to establish the long memory property (1.4) for h_t, discussed in Section 1.

Theorem 2.3: *Under the assumptions of Theorem 2.1 and also* (1.11)*, it follows that* (1.4) *holds with* $C = 2\sigma^4 c/a$.

Sufficient conditions for the presence or absence of leverage are provided by:

Theorem 2.4: *Let the assumptions of Theorem 2.2 hold and also*

$$|\mu_3| \le \frac{2\Big(1-5\,\|b\|^2\Big)}{\|b\|\,\Big(1+3\,\|b\|^2\Big)}. \tag{2.22}$$

Then for any fixed k *such that* $1 \le k \le \infty$*:*
(i) if $ab_1 < 0$, $ab_j \le 0$, $j = 2, ..., k$, *then* $X_t \in \ell(k)$*;*
(ii) if $ab_1 > 0$, $ab_j \ge 0$, $j = 2, ..., k$, *then* $h_j > 0$, $j = 1, ..., k$.

Condition (2.22) implies that $\|b\||\mu_3|$ is bounded (by 2) and also that $\|b\| \le 1/5$. Note the similarity between the leverage conditions of Theorem 2.3 (i) for the LARCH(∞) model and condition (1.16) for the nonlinear MA model (1.7) (where $\mu_3 = 0$ was assumed). Note also that there is no loss of generality in taking $a > 0$, given that $a \ne 0$ has been assumed. Such a restriction leads to some simplification of our results, and indeed would be necessary to identify a and the b_j. Of course choosing $a > 0$ rather than $a < 0$ determines the sign of μ_3, when this is non-zero.

Below we discuss in more detail conditions for leverage in LARCH(1) and GLARCH(1, 1), where owing to the simple structure the function h_t can be explicitly found, and (2.22) can be relaxed.

LARCH(1): $V_t = a + \beta X_{t-1}$.

In this case, $\|b\| = |\beta|$ and the necessary condition $\|b\| < 1$ for the existence of the stationary solution (see Theorem 2.1) becomes $|\beta| < 1$. To obtain h_t, note from (2.10) $\phi_t = \beta^{2t}$; from (2.12) $\sigma^2 = a^2/(1-\beta^2)$, $[b]_3 = \beta^3$; from (2.13) $g_0 = a^3\mu_3(1+2\beta^2)/(1-\beta^2)$, $g_t = 2a^3\beta^{2t-1}/(1-\beta^2)$, $t \geq 1$; from (2.14), (2.15) $r_{tu} = 0$, $t \geq 0$, $u \geq 1$, and $r_{00} = \mu_3\beta^3$, $r_{t0} = \beta^{2t}$, $t \geq 1$. Equation (2.20) in this case becomes

$$h_0 = g_0 + \mu_3\beta^3 h_0, \quad h_t = g_t + \beta^{2t} h_0,$$

and has a unique solution $h_0 = g_0/(1-\mu_3\beta^3), h_t = g_t + \beta^{2t} g_0/(1-\mu_3\beta^3)$ provided $|\beta| < 1$ and $\mu_3\beta^3 \neq 1$ hold. More explicitly,

$$h_0 = \frac{a^3(2\beta^2+1)\mu_3}{(1-\beta^2)(1-\beta^3\mu_3)}, \quad h_1 = \frac{a^3\beta(2+\beta\mu_3)}{(1-\beta^2)(1-\beta^3\mu_3)}, \quad h_t = \beta^{2(t-1)}h_1 \quad (t \geq 2). \tag{2.23}$$

By Theorem 2.1, under assumptions M_3 and M_4 (2.23) provides (1.2) for the LARCH(1) model. (Note that in this case, either M_3 or M_4 imply (2.18) as well as the inequality $|\beta^3\mu_3| < 1$.) The above-mentioned assumptions become

M_3 : $\quad |\mu|_3 < \left(|\beta|^{-1} - 3\theta\right)^3$,

M_4 : $\quad \beta^4\mu_4 + 4|\beta|^3|\mu_3| \leq 1 - 6\beta^2$,

respectively. From Theorem 2.2 and (2.23) we derive:

Proposition 2.4: *Let X_t be LARCH(1), and assumptions M_3 or M_4 hold. Then $X_t \in \ell(1)$ if and only if either*

$$a\beta < 0 \quad \textit{and} \quad \beta\mu_3 > -2, \tag{2.24}$$

or

$$a\beta > 0 \quad \textit{and} \quad \beta\mu_3 < -2 \tag{2.25}$$

hold. Moreover, $X_t \in \ell(\infty)$ if and only if $X_t \in \ell(1)$ and $\beta > 0$.

Note that from Theorem 2.4 we have $X_t \in \ell(1)$ under stronger assumptions, namely under $a\beta < 0$ by imposing M_3 or M_4 together with

(2.22) : $\quad |\mu_3| \leq \dfrac{2(1-5\beta^2)}{|\beta|(1+3\beta^2)}$.

In his extension of LARCH(1), such that $p = 1$ in (1.28) and ψ_1 varies freely with θ and ϕ_{11}, Sentana (1995) obtained more heuristically, in case $\mu_3 = 0$, the condition $\psi_1 < 0$ for leverage, which corresponds to our condition (2.24). Sentana (1995) also examined the compatibility of empirical data with this condition.

GLARCH(1,1): $\quad V_t - \beta_1 V_{t-1} = a' + \alpha_1 X_{t-1}$.

As noted in Section 1, the above equation can be rewritten in the LARCH(∞) form (1.27) with $b_t = \alpha\beta^{t-1}$, $t \geq 1$, $\alpha = \alpha_1, \beta = \beta_1$, $a = a'(1-\beta_1)^{-1}$, and $\|b\| < 1$ is equivalent to $\gamma = \alpha^2 + \beta^2 < 1$. To find h_t, note that from (2.10) $\phi_t = \alpha^2\gamma^{t-1}$, $t \geq 1$. From (2.12), $\sigma^2 = a^2(1-\beta^2)/(1-\gamma)$ and $[b]_3 = \alpha^3/(1-\beta^3)$. From (2.13), (2.21),

$$g_0 = a^3\mu_3\left(1 + 3\frac{\alpha^2}{1-\gamma}\right), \tag{2.26}$$

$$g_t = \frac{2a^3\alpha(1-\beta^2)}{(\gamma-\beta)(1-\gamma)}\Big\{(\beta-1)\beta^t + \alpha^2\gamma^{t-1}\Big\}, \quad t \geq 1.$$

From (2.14), (2.15),

$$r_{tu} = 2\alpha^2\beta^u\gamma^{t-1}, \quad t, u \geq 1; \quad r_{t0} = \alpha^2\gamma^{t-1}, \quad t \geq 1;$$

$$r_{0u} = \frac{3\mu_3\alpha^3\beta^u}{1-\beta^3}, \quad u \ge 1; \quad r_{00} = \frac{\mu_3\alpha^3}{1-\beta^3}. \tag{2.27}$$

Thus from (2.20)

$$h_t = g_t + \alpha^2(2\bar{h} + h_0)\gamma^{t-1}, \quad t \ge 1, \tag{2.28}$$

$$h_0 = g_0 + \frac{\alpha^3}{1-\beta^3}\mu_3(3\bar{h} + h_0), \tag{2.29}$$

where $\bar{h} = \sum_{t=1}^{\infty} \beta^t h_t$. Defining $\bar{g} = \sum_{t=1}^{\infty} \beta^t g_t$, we have

$$\bar{g} = \frac{2a^3\alpha\beta(1-\beta^3)}{(1-\gamma)(1-\gamma\beta)}.$$

From (2.28) we deduce

$$\bar{h} = \bar{g} + \frac{\alpha^2\beta}{1-\beta\gamma}(2\bar{h} + h_0). \tag{2.30}$$

Assumptions M_3 and M_4 become

$$M_3: \qquad |\mu|_3 < \frac{(1-\beta^3)^{1/3}}{|\alpha|}\Big(1 - \frac{3\theta|\alpha|}{(1-\beta^2)^{1/2}}\Big),$$

$$M_4: \qquad \frac{\alpha^4}{1-\beta^4}\ \mu_4 + 4\frac{|\alpha|^3}{1-\beta^3}\ |\mu_3| \le 1 - \frac{6\alpha^2}{1-\beta^2},$$

respectively. Set $A := \dfrac{1-\beta^3-3\alpha^2\beta-\alpha^3\mu_3}{1-\beta\gamma}$. The proof of the following Proposition appears in Appendix A.

Proposition 2.5: *Let X_t be GLARCH(1,1), and assumptions M_3 or M_4 hold. Then*

$$h_t = g_t + \alpha^2 A^{-1}\Big(g_0 + \bar{g}(2 + [b]_3\mu_3)\Big)\gamma^{t-1}, \quad t \ge 1, \tag{2.31}$$

where $A > 0, 2 + [b]_3\mu_3 > 1$. In particular, $X_t \in \ell(1)$ if and only if

$$g_1 + \alpha^2 A^{-1}(g_0 + \bar{g}(2 + [b]_3\mu_3)) < 0. \tag{2.32}$$

Moreover, $X_t \in \ell(\infty)$ if $X_t \in \ell(1)$ and $a\alpha < 0, \beta > 0$ hold. In particular, $X_t \in \ell(\infty)$ if

$$a\alpha < 0, \quad a\mu_3 \le 0, \quad \beta > 0. \tag{2.33}$$

Theorem 2.4, on the other hand, implies $X_t \in \ell(\infty)$ for GLARCH(1,1) under the conditions

$$a\alpha < 0, \quad \beta > 0,$$

assumptions M_3 or M_4, and

$$(2.22): \qquad |\mu_3| \le \frac{2(1-\beta^2)^{1/2}(1-5\alpha^2-\beta^2)}{|\alpha|(1+3\alpha^2-\beta^2)},$$

which are stronger than the conditions of Proposition 2.5. In particular, Proposition 2.5 shows that leverage in the GLARCH(1,1) model may take place even if $|\mu_3|$ is arbitrarily large, as it may happen, for example when if $a > 0, \mu_3 \le 0$. On the other hand, for $a\mu_3 > 0$ (e.g. if $a > 0, \mu_3 > 0$), (2.31) is more difficult to analyze directly. In such a case, Theorem 2.4 can be applied, providing $|\mu_3|$ satisfies (2.22).

Also of interest is LARCH(∞) with b_j given by the FARIMA(0,d,0) weights (1.14) with $0 < d < 1/2$; we might call this a GLARCH(0,d,0) model, and write

$$V_t = a + (1 - L)^{-d} X_t \equiv a + \sum_{t=1}^{\infty} b_j X_{t-j}.$$

From Theorem 2.4, $X_t \in \ell(\infty)$ if $a < 0$ and (2.22) and either M_3 or M_4 hold; evaluation of these conditions in the present case is complicated and is thus omitted. Of course we deduce the long memory property (1.4) from Theorem 2.3.

3 Final remarks

The paper has derived, under primitive conditions, a number of properties of the LARCH(∞) model (1.27). We have developed a result of Giraitis, Robinson and Surgailis (2000) on conditions for existence and uniqueness of a solution of (1.27). We have provided conditions for finiteness of integer moments that again improve upon those of Giraitis, Robinson and Surgailis(2000). The paper is principally motivated by long memory and leverage properties. We have shown that if the weights b_j are chosen to decay like MA weights in linear long memory sequences, then the cross-autocovariance between the squares X_t^2 and past levels $X_{t-j}, j > 0$, decays in the same slow fashion in our model as it does in such linear models. Such a property may be available even if the fourth moment of X_t is infinite, in which sense it has an advantage over the long memory (of autocovariances of squares) property derived by Giraitis, Robinson and Surgailis (2000). We have given conditions for leverage properties, of various extents, and for lack of leverage. These latter conditions obtain for all members of our LARCH(∞) model, but we also directly analyzed two simple special cases of our model, thereby achieving some improvement in the conditions.

The LARCH(∞) model and its special cases are far from fully ready for practical use. We have not discussed estimation of (1.27), either in case of a parametric model such as (1.25), (1.26) or (1.14), or a nonparametric approach analogous to autoregressive spectral estimation. (Quasi)-maximum likelihood estimation based on a working Gaussian ζ_t assumption, as used by Sentana (1995), seems computationally relatively tractable. By analogy with results for ARCH and GARCH special cases of (1.17) (see Lee and Hansen (1994), Lumsdaine (1996)), it would be expected to be asymptotically normal and (if the Gaussianity holds) efficient, without stringent assumptions on unconditional moments of X_t, though the asymptotic theory would likely be difficult. A less elegant asymptotic theory should be available for Whittle estimates based on either X_t^2 or (making use of formulae in the current paper) the bivariate series (X_t, X_t^2) as considered for (1.7) by Zaffaroni (1998) or for short memory versions of (1.17) by Giraitis and Robinson (2001). More *ad hoc* methods include generalized methods-of-moments estimation, for example comparing h_t or its Fourier transform (the cross spectrum of X_t and X_t^2) with sample estimates.

However, a more basic question concerns the direct practical usefulness of the LARCH class. Though it provides equal scope for parsimony as the usual ARCH class (1.17), while at the same time offering more potential for leverage, it is restrictive relative to Sentana's (1995) class, which can nest both ARCH and LARCH models. The inability of LARCH(∞) to satisfactorily separate out parameters primarily describing conditional heteroscedasticity on the one hand, and leverage on the other, was not so much a problem in the original context of Robinson (1991), where it was used to provide Lagrange multiplier tests of i.i.d. behaviour. However, when conditional heteroscedasticity and leverage are to be quantified, a more flexible class like Sentana's (1995) may seem preferable to practitioners. On the other hand, the parsimony of the LARCH model makes it still of interest as a null hypothesis in such a context, and we believe our detailed theoretical investigation of the LARCH model is a necessary precursor to study of more general models.

A Appendix. Proofs of Theorems and Proposition 2.5

Proof of Theorem 2.1 (i) We first show the necessity of $\|b\| < 1$. Let X_t be a solution of (1.27), and let $t_0 < t$. Then

$$V_t = a + \sum_{u<t_0} b_{t-u} X_u + \sum_{t_0 \le s < t} b_{t-s} \zeta_s V_s. \tag{A.1}$$

By iterating (A.1), one obtains first

$$\begin{aligned} V_t &= a + \sum_{u<t_0} b_{t-u} X_u + a \sum_{t_0 \le s<t} b_{t-s}\zeta_s \\ &+ \sum_{u<t_0} \sum_{t_0 \le s_1 < t} b_{t-s_1} b_{s_1-u} \zeta_{s_1} X_u + \sum_{t_0 \le s_2 < s_1 < t} b_{t-s_1} b_{s_1-s_2} \zeta_{s_1}\zeta_{s_2} V_{s_2} \end{aligned}$$

and eventually

$$\begin{aligned} V_t &= a + \sum_{u<t_0} b_{t-u} X_u \\ &+ \sum_{k=1}^{t-t_0} \sum_{t_0 \le s_k < \ldots < s_1 < t} b_{t-s_1} \ldots b_{s_{k-1}-s_k} \zeta_{s_1} \ldots \zeta_{s_k} \Big(a + \sum_{u<t_0} b_{s_k-u} X_u \Big). \end{aligned} \tag{A.2}$$

Noting that $E[V_t|\mathcal{F}_{t_0-1}] = a + \sum_{u<t_0} b_{t-u} X_u$ and using the independence of $\zeta_s, s \ge t_0$ and $\mathcal{F}_{t_0-1}$, one has

$$\mathrm{Var}(V_t|\mathcal{F}_{t_0-1}) = \sum_{k=1}^{t-t_0} \sum_{t_0 \le s_k < \ldots < s_1 < t} b^2_{t-s_1} \ldots b^2_{s_{k-1}-s_k} \Big(a + \sum_{u<t_0} b_{s_k-u} X_u \Big)^2.$$

Therefore

$$\begin{aligned} E\mathrm{Var}(V_t|\mathcal{F}_{t_0-1}) &= \sum_{k=1}^{t-t_0} \sum_{t_0 \le s_k < \ldots < s_1 < t} b^2_{t-s_1} \ldots b^2_{s_{k-1}-s_k} \Big(a^2 + \sum_{u<t_0} b^2_{s_k-u} E X_u^2 \Big) \\ &\ge a^2 \sum_{k=1}^{t-t_0} \sum_{t_0 \le s_k < \ldots < s_1 < t} b^2_{t-s_1} \ldots b^2_{s_{k-1}-s_k}. \end{aligned}$$

For any $k \ge 1$, the last sum increases monotonically to $\|b\|^{2k}$ as $t_0 \to -\infty$. Therefore

$$\liminf_{t_0 \to -\infty} E\mathrm{Var}(V_t|\mathcal{F}_{t_0-1}) \ge a^2 \sum_{k=1}^{\infty} \|b\|^{2k}.$$

As $E\mathrm{Var}(V_t|\mathcal{F}_{t_0-1}) \le EV_t^2 < \infty$, this proves the necessity of the condition $\|b\| < 1$ in the case $a \ne 0$.

The sufficiency of this condition for the existence of the solution given by (2.2) was shown in Giraitis, Robinson and Surgailis (2000). To show uniqueness, let X_t', X_t'' be solutions of (1.21), (1.24), then $\tilde{X}_t = X_t' - X_t''$ is a solution of the homogeneous equation for V_t in (1.24) with $a = 0$, and therefore $X_t' = X_t''$ a.s. by part (ii).

(ii). Let $a = 0$. Noting that (A.2) still holds, we obtain

$$V_t = \sum_{u<t_0} b_{t-u} X_u + \sum_{u<t_0} \sum_{k=1}^{t-t_0} \sum_{t_0 \le s_k < \ldots < s_1 < t} b_{t-s_1} \ldots b_{s_{k-1}-s_k} b_{s_k-u} \zeta_{s_1} \ldots \zeta_{s_k} X_u. \tag{A.3}$$

Let $\mathcal{F}_{[s,t]} = \sigma\{\zeta_u : s \le u \le t\}$. The $\sigma-$fields $\mathcal{F}_{[t_0,t-1]}$ increase monotonically to $\mathcal{F}_{t-1}$ as $t_0 \to -\infty$ and therefore by a well-known property of conditional expectations,

$$E[V_t|\mathcal{F}_{[t_0,t-1]}] \to E[V_t|\mathcal{F}_{t-1}] = V_t \quad \text{a.s.} \tag{A.4}$$

as $t_0 \to -\infty$. On the other hand, the independence of X_u and $\mathcal{F}_{[t_0,t]}$ for $u < t_0$ implies $E[X_u|\mathcal{F}_{[t_0,t-1]}] = EX_u = 0$ and we obtain from (A.3)

$$E[V_t|\mathcal{F}_{[t_0,t-1]}] = 0 \quad \text{a.s.},$$

for each $t_0 < t$. Thus by (A.4) we obtain $V_t = 0$ and $X_t = 0$ a.s., thereby proving part (ii). $\square$

Proof of Theorem 2.2. For r_{tu} given by (2.14) let

$$(R^0 f)_t := \sum_{u=1}^{\infty} r_{tu} f_u, \quad t \ge 1 \tag{A.5}$$

be the linear operator in the Hilbert space $L^2(\mathbf{Z}^0_+)$, $\mathbf{Z}^0_+ := \{1, 2, \ldots\}$. It is easily seen from the proof of Lemma A.1 below that under Assumptions M_3 or M_4, the operator R^0 is well-defined on $L^2(\mathbf{Z}^0_+)$. We first show that its Hilbert-Schmidt norm $\|R^0\| = \big\{\sum_{t,u=1}^{\infty} r^2_{tu}\big\}^{1/2}$ satisfies (2.17). Put $b_t = 0$ $(t < 0)$. By the Minkowski inequality,

$$\begin{aligned} \|R^0\| &= 2\Big\{\sum_{t,u=1}^{\infty}\Big(\sum_{v=0}^{\infty}\phi_v b_{t-v} b_{t-v+u}\Big)^2\Big\}^{1/2} \\ &\le \sum_{v=0}^{\infty}\phi_v\Big\{\sum_{t,u=1}^{\infty} b^2_{t-v} b^2_{t-v+u}\Big\}^{1/2} = 2\Phi(1)\|b\|^2 = 2\|b\|^2/(1-\|b\|^2), \end{aligned} \tag{A.6}$$

which is less than 1 because both M_3 and M_4 imply $||b||^2 < 1/3$. We first derive (2.20) in the LARCH(N) case for $N < \infty$. From (1.21), (1.25), for $t > 0$,

$$h_t = E\Big[\Big(a + \sum_{s<t} b_{t-s} X_s\Big)^2 X_0\Big] = a^2 EX_0 + 2a\sum_{s<t} b_{t-s} EX_s X_0 + \sum_{s_1,s_2<t} b_{t-s_1} b_{t-s_2} EX_{s_1} X_{s_2} X_0.$$

Because $EX_0 = 0$ and $EX_s X_t = \sigma^2$, $s = t; = 0, s \ne t$; $EX_{s_1}X_{s_2}X_0 = 0$ if either $s_1 \ne s_2, \max(s_1, s_2) > 0$, or $s_1, s_2 < 0$, we deduce that

$$h_t = 2a\sigma^2 b_t + \sum_{0<s<t} b^2_{t-s} h_s + 2b_t \sum_{u>0} b_{t+u} h_u + b^2_t h_0 \tag{A.7}$$

and thence by iteration

$$\begin{aligned} h_t &= 2a\sigma^2 b_t + 2a\sigma^2 \sum_{0<s_1<t} b^2_{t-s_1} b_{s_1} + \sum_{0<s_2<s_1<t} b^2_{t-s_1} b^2_{s_1-s_2} h_{s_2} \\ &+ 2\sum_{u>0} h_u b_t b_{t+u} + 2\sum_{u>0} h_u \sum_{0<s_1<t} b^2_{t-s_1} b_{s_1} b_{s_1+u} \\ &+ \Big(b^2_t + \sum_{0<s_1<t} b^2_{t-s_1} b^2_{s_1}\Big) h_0, \end{aligned}$$

which yields

$$\begin{aligned} h_t &= 2a\sigma^2 \sum_{0<s\le t}\phi_{t-s} b_s + 2\sum_{u>0} h_u \sum_{0<s\le t}\phi_{t-s} b_{s+u} b_u + h_0\phi_t \\ &= g_t + \sum_{u=0}^{\infty} h_u r_{tu}. \end{aligned}$$

For $t = 0$,

$$\begin{aligned} h_0 &= EX_0^3 = \mu_3 E\Big(a + \sum_{s<0} b_{-s} X_s\Big)^3 \\ &= \mu_3\Big(a^3 + 3a\sigma^2\|b\|^2 + h_0[b]_3 + 3\sum_{s_2<s_1<0} b_{-s_2}^2 b_{-s_1} h_{s_2-s_2}\Big) \\ &= g_0 + \sum_{u=0}^{\infty} r_{0u} h_u. \end{aligned}$$

This proves the validity (2.20) in the LARCH(N), $N < \infty$ case.

Next we show that (2.20) has a unique solution $h \in L^2(\mathbf{Z}_+)$, $\mathbf{Z}_+ = \{0, 1, \ldots\}$. First fix an arbitrary value $h_0 \equiv \xi$ and solve the equation

$$h^\xi = g^\xi + R^0 h^\xi, \tag{A.8}$$

for $h^\xi = (h_t^\xi, t \geq 1)$, where $g^\xi = \xi\phi^0 + g^0$ and $\phi^0 = (\phi_t, t \geq 1) \in L^2(\mathbf{Z}_+^0)$, $g^0 = (g_t, t \geq 1) \in L^2(\mathbf{Z}_+^0)$. By (2.17), (A.8) admits a unique solution

$$h^\xi = (1 - R^0)^{-1} g^\xi =: \xi h' + h'',$$

where $h' := (1 - R^0)^{-1}\phi^0$, $h'' := (1 - R^0)^{-1} g^0$ belong to the space $L^2(\mathbf{Z}_+^0)$ and do not depend on ξ.

Next we solve the equation

$$\begin{aligned} \xi &= g_0 + r_{00}\xi + \sum_{u=1}^{\infty} r_{0u} h^\xi \\ &= g_0 + r_{00}\xi + \xi\sum_{u=1}^{\infty} r_{0u} h'_u + \sum_{u=1}^{\infty} r_{0u} h''_u \end{aligned}$$

for ξ, yielding

$$\xi = \frac{g_0 + \sum_{u=1}^{\infty} r_{0u} h''_u}{1 - r_{00} - \sum_{u=1}^{\infty} r_{0u} h'_u}, \tag{A.9}$$

by (2.18). Define $h = (h_t, t \geq 0)$ by

$$h_0 = \xi, \quad h_t = \xi h'_t + h''_t \quad (t \geq 1),$$

where $h' = (1 - R^0)^{-1}\phi^0, h'' = (1 - R^0)^{-1} g^0$, and ξ is given by (A.9). Then $h \in L^2(\mathbf{Z}_+)$ and satisfies (2.20).

Consider now the LARCH(∞) case. Put

$$b_{j,N} := \begin{cases} b_j, & \text{if } 1 \leq j \leq N, \\ 0, & \text{otherwise.} \end{cases}$$

Write $h_{t,N} := EX_{t,N}^2 X_{0,N}$, $t \geq 0$, where $\{X_{t,N}\}$ is the solution to the LARCH equations corresponding to $\{b_{j,N}\}$. According to Lemma B.3 below, for each $t \geq 0$,

$$h_t = \lim_{N\to\infty} h_{t,N}. \tag{A.10}$$

Let us show that h_t of (A.10) belongs to $L^2(\mathbf{Z}_+)$ and satisfies (2.20). Let $\phi_{t,N}$, $g_{t,N}$, $r_{tu,N}$ be defined as in (2.11), (2.13)–(2.15), with b_j replaced by $b_{j,N}$, and let R_N^0 be the corresponding operator in $L^2(\mathbf{Z}_+^0)$ given by (A.5). By (A.6), we obtain $\sup_{N\geq 1}\|R_N^0\| < 1$, $\|R^0 - R_N^0\| \to 0$, and the convergences $\phi_N^0 \to \phi$, $g_N^0 \to g^0$, $h'_N := (1 - R_N^0)^{-1}\phi_N^0 \to$

$(1-R^0)^{-1}\phi^0 = h'$, $h''_N := (1-R^0_N)^{-1}g^0_N \to (1-R^0)^{-1}g^0 = h''$ in $L^2(\mathbf{Z}^0_+)$ easily follow. Moreover, as $r_{00,N} + \sum_{u=1}^\infty r_{0u,N}h'_{u,N} \to r_{00} + \sum_{u=1}^\infty r_{0u}h'_u$ $(N\to\infty)$, so condition (2.18) implies $r_{00,N} + \sum_{u=1}^\infty r_{0u,N}h'_{u,N} \neq 1$ for all sufficiently large N, and therefore

$$\xi_N := \frac{g_{0,N} + \sum_{u=1}^\infty r_{0u,N}h'_{u,N}}{1 - r_{00,N} - \sum_{u=1}^\infty r_{0u,N}h'_{u,N}} \to \xi,$$

where ξ is defined by (A.9). The above relations imply the convergence $h_N \to h$ in $L^2(\mathbf{Z}_+)$ and the validity of (2.20). Thus sufficiency of (2.18) is established. To prove necessity, note that if (2.18) is not satisfied then $\lambda = 1$ is an eigenvalue of the operator $(Rf)_t = \sum_{u=0}^\infty r_{tu}f_u$, corresponding to the eigenfunction $\psi = (\psi_t, t\ge 0), \psi_0 = 1, \psi_t = h'_t (t \ge 1)$, and the solution h_t of (2.20) is not unique. This completes the proof of Theorem 2.2. □

Proof of Theorem 2.3. In view of (2.20) and (1.11), it suffices to show

$$g_t = 2a\sigma^2 \sum_{s=1}^t b_s\phi_{t-s} \sim \Big(2a\sigma^2 \sum_{s=0}^\infty \phi_s\Big) b_t \sim \frac{2\sigma^4 c}{a}\, t^{d-1} \tag{A.11}$$

and

$$\sum_{u=0}^\infty r_{tu}h_u = o(t^{d-1}) \quad \text{as } t\to\infty. \tag{A.12}$$

Here, (A.11) follows from (1.11) and the fact that $\phi_t^2 = O(b_t^2) = O(t^{2d-2})$, as in Giraitis, Robinson and Surgailis (2000, Lemma 4.1). It remains to show (A.12). Consider

$$J_t := \sum_{u=0}^\infty r_{tu}h_u = \sum_{u=1}^\infty r_{tu}h_u + \phi_t h_0 = 2\sum_{s=1}^t b_s\phi_{t-s}\sum_{u=1}^\infty b_{s+u}h_u + \phi_t h_0.$$

Since $\|h\| < \infty$,

$$|\sum_{u=1}^\infty b_{s+u}h_u| \le \left\{\sum_{u=1}^\infty b^2_{s+u}\right\}^{1/2} \|h\| \le K\left\{\sum_{u=s}^\infty u^{2d-2}\right\}^{1/2} \le Ks^{-1/2+d},$$

where K denotes a generic positive constant. As $\phi_s \le Ks^{2d-2}$, and $2d<1$, we obtain

$$|J_t| \le K\sum_{s=1}^t |t-s|^{2d-2}s^{2d-(3/2)} + O(|t|^{2d-2}) = O(t^{2d-(3/2)}) = o(t^{d-1})$$

which proves (A.12) and the theorem. □

The proof of Theorem 2.4 is preceded by the following lemma.

Lemma A.1: *Let* (2.21) *hold. Then* (2.18) *holds, and moreover*

$$|h_0| \le \frac{|a|^3|\mu_3|}{1 - 3\|b\|^2 - |\mu_3|\|b\|^3}. \tag{A.13}$$

Proof. To show (2.18), it suffices to verify the bound

$$|r_{00}| + \sum_{u=1}^\infty |r_{0u}h'_u| \le \frac{|\mu_3|\|b\|^3}{1-3\|b\|^2}, \tag{A.14}$$

whose right side is less than 1 by (2.21). By (2.15), $|r_{00}| \le |\mu_3| \|b\|_3^3 \le |\mu_3| \|b\|_2^3$ and

$$\sum_{u=1}^{\infty} |r_{0u} h'_u| \le 3|\mu_3| \sum_{s=1}^{\infty} b_s^2 \sum_{u=1}^{\infty} |b_{s+u} h'_u| \le 3|\mu_3| \|b\|^3 \|h'\|, \tag{A.15}$$

where

$$\|h'\| = \|(1 - R^0)^{-1} \phi^0\| \le \frac{\|\phi^0\|}{1 - \|R^0\|} \le \frac{\|b\|^2}{1 - 3\|b\|^2},$$

due to (2.17) and $\|\phi^0\| = \left\{\sum_{u=1}^{\infty} \phi_u^2\right\}^{1/2} \le \left\{\sum_{t,u=1}^{\infty} \phi_t \phi_u\right\}^{1/2} \le \sum_{t=1}^{\infty} \phi_t = \Phi(1) - 1 = \|b\|^2/(1 - \|b\|^2)$. Therefore,

$$|r_{00}| + \sum_{u=1}^{\infty} |r_{0u} h'_u| \le |\mu_3| \|b\|^3 + \frac{3|\mu_3| \|b\|^5}{1 - 3\|b\|^2},$$

proving (A.14).

The inequality (A.13) follows from

$$h_0 = \frac{g_0 + \sum_{u=1}^{\infty} r_{0u} h''_u}{1 - r_{00} - \sum_{u=1}^{\infty} r_{0u} h'_u},$$

(A.14), and the bound

$$|g_0| + \sum_{u=1}^{\infty} |r_{0u} h''_u| \le \frac{|a|^3 |\mu_3|}{1 - 3\|b\|^2}, \tag{A.16}$$

which we verify below. By (2.13),

$$|g_0| \le |a| |\mu_3| (a^2 + 3\sigma^2 \|b\|^2) = \frac{|a|^3 |\mu_3| (1 + 2\|b\|^2)}{1 - \|b\|^2}, \tag{A.17}$$

and, similarly to (A.15),

$$\sum_{u=1}^{\infty} |r_{0u} h''_u| \le 3|\mu_3| \|b\|^3 \|h''\|.$$

Here,

$$\|h''\| = \|(1 - R^0)^{-1} g^0\| \le \frac{(1 - \|b\|^2) \|g^0\|}{1 - 3\|b\|^2},$$

where

$$\|g^0\| \le 2|a| \sigma^2 \|b\| \sum_{t=0}^{\infty} \phi_t = \frac{2|a|^3 \|b\|}{(1 - \|b\|^2)^2}.$$

Consequently,

$$\sum_{u=1}^{\infty} |r_{0u} h''_u| \le \frac{6|a|^3 |\mu_3| \|b\|^4}{(1 - \|b\|^2)(1 - 3\|b\|^2)}. \tag{A.18}$$

Clearly, (A.17) and (A.18) imply (A.16). □

Proof of Theorem 2.4. Note that (2.22) implies (2.21) and therefore the validity of (2.18) and Theorem 2.1.

Let us prove the statements (i), (ii) for $k = 1$. From (A.7) it follows that

$$h_1 = 2a\sigma^2 b_1 + 2b_1 \sum_{u=1}^{\infty} h_u b_{1+u} + b_1^2 h_0,$$

where the last two terms do not exceed

$$2|b_1|\sum_{u=1}^{\infty}|h_u b_{1+u}| + b_1^2|h_0| \le 2|b_1|\|b\|\|h^0\| + |b_1|\|b\||h_0|.$$

Therefore $\operatorname{sgn}(h_1) = \operatorname{sgn}(ab_1)$ provided the inequality

$$2|a|\sigma^2 > 2\|h^0\|\|b\| + \|b\||h_0| \tag{A.19}$$

holds. From (A.7) it follows that

$$\|h^0\| \le 2|a|\sigma^2\|b\| + 3\|h^0\|\|b\|^2 + \|b\|^2|h_0|,$$

or

$$\|h^0\| \le \frac{2|a|\sigma^2\|b\| + \|b\|^2|h_0|}{1 - 3\|b\|^2}. \tag{A.20}$$

But (A.13) and (A.20) imply (A.19) and hence $\operatorname{sgn}(h_1) = \operatorname{sgn}(ab_1)$, or the statements (i), (ii) for $k = 1$. The general case $k \ge 1$ follows similarly by induction in k. Indeed, according to (A.7),

$$h_k = 2a\sigma^2 b_k + \sum_{s=1}^{k-1} b_{t-s}^2 h_s + 2b_k\sum_{u=1}^{\infty} b_{k+u}h_u + b_k^2 h_0.$$

To show (i), let $h_1, \dots, h_{k-1} < 0$ by the inductive hypothesis. Then $\sum_{s=1}^{k-1} b_{t-s}^2 h_s < 0$ and the inequality $h_k < 0$ follows from

$$\left|2b_k\sum_{u=1}^{\infty} b_{k+u}h_u + b_k^2 h_0\right| \le 2|a\sigma^2 b_k|, \tag{A.21}$$

where the left hand side does not exceed $2|b_k|\|b\|\|h^0\| + b_k^2|h_0| \le |b_k|(2\|b\|\|h^0\| + \|b\||h_0|)$. Then (A.21) follows from (A.19). The proof of (ii) is analogous. □

Proof of Proposition 2.5. We first show (2.31). Set $V = 2\bar{h} + h_0$. Adding $2\bar{h}$ to both sides of (2.29),

$$h_0 + 2\bar{h} = g_0 + 2\bar{h} + [b]_3\mu_3(V + \bar{h}),$$

and hence

$$V = g_0 + \bar{h}\Big(2 + [b]_3\mu_3\Big) + [b]_3\mu_3 V.$$

Replacing $\bar{h}$ in the above equation by (2.30) yields

$$VA = g_0 + \bar{g}(2 + [b]_3\mu_3).$$

We show that, under M_3 or M_4,

$$A > 0. \tag{A.22}$$

Indeed, note that M_3 or M_4 imply $0 < \gamma < 1, |\beta| < 1$ and (A.22) follows from

$$1 - \beta^3 - 3\alpha^2\beta - \alpha^3\mu_3 > 0, \quad \text{or} \quad i := \frac{|\alpha^3\mu_3|}{1-\beta^3} + \frac{3\alpha^2\beta}{1-\beta^3} < 1.$$

Here, $|\alpha^3\mu_3|/(1-\beta^3) = |[b]_3\mu_3| \le \|b\|_3^3|\mu|_3 < \|b\|_3|\mu|_3^{1/3} < 1$, where the last two inequalities hold under M_3. Next, $|\beta|/(1-\beta^3) \le 1/(1-\beta^2)$ implies

$$\frac{\alpha^2|\beta|}{1-\beta^3} \le \frac{\alpha^2}{1-\beta^2} = \|b\|^2 \le \|b\|,$$

as $\|b\| < 1$. Therefore, under M_3, $i \le \|b\|_3 |\mu|_3^{1/3} + 3\|b\| < 1$, while under M_4, $i \le \|b\|_3^3 |\mu_3| + 3\|b\|^2 < 1$. This proves (A.22) and (2.31). The inequality $2 + [b]_3\mu_3 > 1$ follows from $|[b]_3\mu_3| \le \|b\|_3^3|\mu_3| < 1$ (see (2.3), (2.4)).

To prove the second part of the proposition, it suffices to show that (2.32) together with $a\alpha < 0, \beta > 0$ imply $g_t < 0, t \ge 1$. Indeed, we have $0 < \gamma, \beta < 1$ and so $g_t < 0, t \ge 1$ follows from

$$\frac{(\beta-1)\beta^t + \alpha^2\gamma^{t-1}}{\gamma - \beta} > 0. \tag{A.23}$$

Let us check (A.23). Let $\gamma > \beta$, then $\alpha^2\gamma^{t-1} - (1-\beta)\beta^t \ge \alpha^2\gamma^{t-1} - (1-\beta)\beta\gamma^{t-1} = (\gamma-\beta)\gamma^{t-1}$ and (A.23) follows. The verification of (A.23) in the case $\beta > \gamma$ is similar. The fact that (2.33) implies (2.32) is immediate from (2.26). □

We remark that in the proof of Proposition 2.5 we directly verified that, in the GLARCH(1,1) model, the leverage equation (2.20), or (2.28), (2.29), has a unique square-summable solution (2.31) under assumptions M_3 or M_4 alone. Thus assumptions M_3 or M_4 imply also (2.18), as can be directly verified by using (2.27) and (2.16), (2.17).

B Appendix. Proofs of finiteness of moments

Proof of Proposition 2.1. This is contained in the following three lemmas. We first introduce some auxiliary notation. Consider integers $t_i \in \mathbf{Z}, k_i \ge 0, i = 1,2,3$ and a collection $f_{i,j} \in L^2(\mathbf{Z}_+^0), j = 1,\dots,k_i, i = 1,2,3$. Let

$$U_i := \sum_{s_{k_i} < \dots < s_1 < t_i} f_{i,1}(t_i - s_1)\dots f_{i,k_i}(s_{k_i-1} - s_{k_i})\zeta_{s_1}\dots\zeta_{s_{k_i}}, \tag{B.1}$$

if $k_i \ge 1$; $U_i := 1$ if $k_i = 0$. Then

$$EU_i^2 = \sum_{s_{k_i} < \dots < s_1 < t_i} f_{i,1}^2(t_i - s_1)\dots f_{i,k_i}^2(s_{k_i-1} - s_{k_i}) = \prod_{j=1}^{k_i} \|f_{i,j}\|^2 < \infty, \tag{B.2}$$

so that the series (B.1) converges in mean square. Put $I := \{(i,j) : i = 1,2,3, j = 1,\dots,k_i\}$. For each $f \in L^2(\mathbf{Z}_+^0)$, put

$$D(f) := |\mu|_3^{1/3}\|f\|_3 + 3\theta\|f\|_2,$$

where θ is defined as in Assumption M_3.

Lemma B.1: *For any collection $\{f_{i,j}, (i,j) \in I\}$,*

$$E|U_1U_2U_3| \le \prod_{(i,j)\in I} D(f_{i,j}).$$

Proof. By Fatou inequality, it suffices to prove the lemma for $f_{i,j}(s) = 0\, \forall s > N (\exists N < \infty)$, in other words, for finite sums U_i (B.1). Write the set I as the table

$$I = \left\{\begin{matrix} (1,1) & (1,2)\dots(1,k_1) \\ (2,1) & (2,2)\dots(2,k_2) \\ (3,1) & (3,2)\dots(3,k_3)\end{matrix}\right\}$$

consisting of three rows $I_i, i = 1,2,3$ (some of which may be empty) and having $|I| = k_1 + k_2 + k_3$ elements; $k_i \ge 0, i = 1,2,3$. Then

$$\prod_{i=1}^{3} U_i = \sum_{S_I} F_{S_I}((t))\zeta^{S_I}, \tag{B.3}$$

where $(t) := (t_1, t_2, t_3)$,

$$\zeta^{S_I} := \prod_{(i,j)\in I} \zeta_{s_{i,j}}, \qquad F_{S_I}((t)) := \prod_{(i,j)\in I} f_{i,j}(s_{i,j-1} - s_{i,j}), \tag{B.4}$$

and where $s_{i,0} := t_i$ and the sum $\sum_{S_I}$ is taken over all integers $s_{i,j}, (i,j) \in I$. As $f_{i,j}(s) := 0$ $(s \le 0)$, so $s_{i,j}$ in (B.3), (B.4) satisfy

$$s_{i,k_i} < \ldots < s_{i,1} < t_i, \quad i = 1, 2, 3.$$

To proceed, we need some terminology. Any subset $G \subset I, G \neq \emptyset$ such that $|G \cap I_j| \le 1, j = 1, 2, 3$ will be called an *edge*. Let Γ_I be the class of *ordered* partitions $\gamma = (G_1, \ldots, G_r)$ of I by edges. (Two partitions $\gamma = (G_1, \ldots, G_r) \in \Gamma_I, \gamma' = (G'_1, \ldots, G'_{r'}) \in \Gamma_I$ are equal $(\gamma = \gamma')$ if and only if $r = r'$ and $G_1 = G'_1, \ldots, G_r = G'_r$.) Then the sum in (B.3) can be rewritten as

$$\begin{aligned} \sum_{S_I} F_{S_I}((t))\zeta^{S_I} &= \sum_{\gamma\in\Gamma_I} \sum_{\tilde{s}_r<\ldots<\tilde{s}_1} F_{S_I}((t))\zeta^{S_I} \mathbf{1}\{s_{i,j} = \tilde{s}_q, \, (i,j) \in G_q, q = 1, \ldots, r\} \\ &\equiv \sum_{\gamma\in\Gamma_I} \sum\nolimits_{\tilde{S}}^{\gamma} F_{S_I}((t))\zeta^{S_I}, \end{aligned}$$

where $\sum_{\tilde{S}}^{\gamma}$ stands for the sum over all *ordered* integers $\tilde{s}_r < \ldots < \tilde{s}_1$ such that $s_{i,j} = \tilde{s}_q$ for $(i,j) \in G_q, q = 1, \ldots, r$.

Next, we split the sum $\sum_{\tilde{S}}^{\gamma}$ into "diagonal" and "off-diagonal" parts. To that end, for any $\gamma \in \Gamma_I$, put

$$I^0 := \bigcup_{q:|G_q|=1} G_q, \quad I^1 := \bigcup_{q:|G_q|>1} G_q.$$

Then

$$\sum\nolimits_{\tilde{S}}^{\gamma} F_{S_I}((t))\zeta^{S_I} = \sum\nolimits_{\tilde{S}^0}^{\gamma} \sum\nolimits_{\tilde{S}^1}^{\gamma} F_{S_I}((t))\zeta^{S_I},$$

where $\sum_{\tilde{S}^0}^{\gamma}$ stands for the sum over all ordered integers $\tilde{s}_q, q = 1, \ldots, r$ with $|G_q| > 1$, while $\sum_{\tilde{S}^1}^{\gamma}$ stands for the sum over all ordered integers $\tilde{s}_q, q = 1, \ldots, r$ with $|G_q| = 1$.

Write $\zeta^{S_I} = \zeta^{S_{I^0}}\zeta^{S_{I^1}}$, where $\zeta^{S_{I^0}} := \prod_{(i,j)\in I^0} \zeta_{s_{i,j}}$, $\zeta^{S_{I^1}} := \prod_{(i,j)\in I^1} \zeta_{s_{i,j}}$. Note that for fixed $S_{I^0} = \{s_{i,j} : (i,j) \in I^0\}$, $\zeta^{S_{I^0}}$ is independent of $\zeta^{S_{I^1}}$. Consequently,

$$\begin{aligned} E\Big|\sum\nolimits_{\tilde{S}^1}^{\gamma} F_{S_I}((t))\zeta^{S_I}\Big| &= E|\zeta^{S_{I^0}}|E\Big|\sum\nolimits_{\tilde{S}^1}^{\gamma} F_{S_I}((t))\zeta^{S_{I^1}}\Big| \\ &\le E|\zeta^{S_{I^0}}|E^{1/2}\Big\{\sum\nolimits_{\tilde{S}^1}^{\gamma} F_{S_I}((t))\zeta^{S_{I^1}}\Big\}^2. \end{aligned}$$

Now, as the sum $\sum_{\tilde{S}_1}^{\gamma}$ is taken over *ordered* sets of disjoint integers,

$$E\Big\{\sum\nolimits_{\tilde{S}^1}^{\gamma} F_{S_I}((t))\zeta^{S_{I^1}}\Big\}^2 = \sum\nolimits_{\tilde{S}^1}^{\gamma} \big|F_{S_I}((t))\big|^2.$$

We finally obtain

$$\begin{aligned} E|\prod_{i=1}^{3} U_i| &= E\Big|\sum_{S_I} F_{S_I}((t))\zeta^{S_I}\Big| \\ &\le \sum_{\gamma\in\Gamma_I} \sum\nolimits_{\tilde{S}^0}^{\gamma} E|\zeta^{S^0}|\Big(\sum\nolimits_{\tilde{S}^1}^{\gamma} \big|F_{S_I}((t))\big|^2\Big)^{1/2} =: p_I((t)). \end{aligned} \tag{B.5}$$

Put $p_I := \sup_{(t)} p_I((t))$. Now Lemma B.1 follows from the following lemma. □

Lemma B.2.

$$p_I \le D^I, \tag{B.6}$$

where $D^I := \prod_{(i,j)\in I} D_{i,j}$ *and* $D_{i,j} := D(f_{i,j})$.

Proof. In the case when I has one or two rows, (B.6) immediately follows from (B.2); indeed, $E|U_1U_2| \le \prod_{i=1}^2 E^{1/2}U_i^2 \le \prod_{(i,j)\in I} \|f_{i,j}\| \le D^I$. Let $k_1,\dots,k_3 \ge 1$. We prove (B.6) by induction in $|I| = k_1+k_2+k_3$. Let $\gamma = (G_1,\dots,G_r) \in \Gamma_I$ be a partition of the table I. Let G_1 be the first edge from the right. It may contain 1, 2, or 3 elements. Let $I' = I\backslash G_1$, so that $\gamma' = (G_2,\dots,G_r) \in \Gamma_{I'}$ is a partition of the table I'. Let $\sum_{\tilde S_0'}^{\gamma'}$ (respectively, $\sum_{\tilde S_1'}^{\gamma'}$) denote the sum over all ordered integers $\tilde s_q, q = 2,\dots,r$ such that $|G_q| > 1$ (respectively, $|G_q| = 1$). Let $p_{I,u}((t))(u = 1,2,3)$ be defined as in (B.5), where $\sum_{\gamma\in\Gamma_I}$ is replaced by the sum over all $\gamma \in \Gamma_I$ with $|G_1| = u$. Let $p_{I,u} := \sup_{(t)} p_{I,u}((t))$, then $p_I \le \sum_{u=1}^3 p_{I,u}$.

Consider first the case $|G_1| = 2, G_1 = \{(1,1),(2,1)\}$. Let $(t') = (\tau,\tau,t_3)$. Then

$$\begin{aligned}
\sup_{(t)} \quad & \sum\nolimits_{\tilde S_0}^{\gamma} E|\zeta^{S_0}|\Big(\sum\nolimits_{\tilde S_1}^{\gamma} \big|F_{S_I}((t))\big|^2\Big)^{1/2} \\
\le \quad & \sup_{(t)} \sum_{\tilde s_1} \sum\nolimits_{\tilde S_0'}^{\gamma'} E|\zeta^{S_0'}| \sup_{\tau} \Big(\sum\nolimits_{\tilde S_1'}^{\gamma'} \big|F_{S_{I'}}((t'))\big|^2\Big)^{1/2} |f_{1,1}(t_1-\tilde s_1) f_{2,1}(t_2-\tilde s_1)| \\
\le \quad & \|f_{1,1}\|_2 \|f_{2,1}\|_2 \sup_{(t')} \sum\nolimits_{\tilde S_0'}^{\gamma'} E|\zeta^{S_0'}|\Big(\sum\nolimits_{\tilde S_1'}^{\gamma'} \big|F_{S_{I'}}((t'))\big|^2\Big)^{1/2}.
\end{aligned} \tag{B.7}$$

Put $\hat I := I\backslash\{(1,1),(2,1),(3,1)\}$. By using the inductive assumption, we obtain

$$p_{I,2} \le C_0 D^{\hat I}\big(\|f_{1,1}\|_2\|f_{2,1}\|_2 D_{3,1} + \|f_{1,1}\|_2\|f_{3,1}\|_2 D_{2,1} + \|f_{2,1}\|_2\|f_{3,1}\|_2 D_{1,1}\big). \tag{B.8}$$

Next, let $|G_1| = 3, G_1 = \{(1,1),(2,1),(3,1)\}, (t') = (\tau,\tau,\tau)$. Then

$$\begin{aligned}
\sup_{(t)} \quad & \sum\nolimits_{\tilde S_0}^{\gamma} E|\zeta^{S_0}|\Big(\sum\nolimits_{\tilde S_1}^{\gamma} \big|F_{S_I}((t))\big|^2\Big)^{1/2} \\
\le \quad & \sup_{(t)} \sum_{\tilde s_1} \sum\nolimits_{\tilde S_0'}^{\gamma'} E|\zeta^{S_0'}||\mu|_3 \sup_{\tau} \Big(\sum\nolimits_{\tilde S_1'}^{\gamma'} \big|F_{S_{I'}}((t'))\big|^2\Big)^{1/2} \\
& \qquad \times |f_{1,1}(t_1-\tilde s_1) f_{2,1}(t_2-\tilde s_1) f_{3,1}(t_3-\tilde s_1)| \\
\le \quad & |\mu|_3 \|f_{1,1}\|_3 \|f_{2,1}\|_3 \|f_{3,1}\|_3 \sup_{(t')} \sum\nolimits_{\tilde S_0'}^{\gamma'} E|\zeta^{S_0'}|\Big(\sum\nolimits_{\tilde S_1'}^{\gamma'} \big|F_{S_{I'}}((t'))\big|^2\Big)^{1/2}.
\end{aligned}$$

By using the inductive assumption, we obtain

$$p_{I,3} \le D^{\hat I} |\mu|_3 \|f_{1,1}\|_3 \|f_{2,1}\|_3 \|f_{3,1}\|_3. \tag{B.9}$$

Finally, let $|G_1| = 1, G_1 = \{(1,1)\}$. Assume also $k_1 \ge 2$. Then

$$\sum\nolimits_{\tilde S_0}^{\gamma} E|\zeta^{S_0}|\Big(\sum\nolimits_{\tilde S_1}^{\gamma} \big|F_{S_I}((t))\big|^2\Big)^{1/2} \le \sum\nolimits_{\tilde S_0'}^{\gamma'} E|\zeta^{S_0'}|\Big(\sum\nolimits_{\tilde S_1'}^{\gamma'} \big|F'_{S_{I'}}((t))\big|^2\Big)^{1/2}, \tag{B.10}$$

where $F'_{S_{I'}}((t)) = \prod_{(i,j)\in I'} f'_{i,j}(s_{i,j-1} - s_{i,j})$ and $f'_{i,j} \in L^2(\mathbf{Z}), (i,j) \in I'$ are defined by $f'_{i,j} := f_{i,j}$ if $(i,j) \in I', (i,j) \ne (1,2)$,

$$f'_{1,2}(u) := \Big(\sum_v f_{1,1}^2(v) f_{1,2}^2(u-v)\Big)^{1/2}.$$

Observe that

$$\|f'_{1,2}\|_2 = \|f_{1,1}\|_2\|f_{1,2}\|_2 \tag{B.11}$$

and, by Minkowski or Young inequalities,

$$\|f'_{1,2}\|_3 \le \|f_{1,1}\|_2\|f_{1,2}\|_3. \tag{B.12}$$

From (B.11), (B.12) we obtain

$$\begin{aligned} D(f'_{1,2}) &= |\mu|_3^{1/3}\|f'_{1,2}\|_3 + 3\theta\|f'_{1,2}\|_2 \\ &\le |\mu|_3^{1/3}\|f_{1,1}\|_2\|f_{1,2}\|_3 + 3\theta\|f_{1,1}\|_2\|f_{1,2}\|_2 \end{aligned}$$

$$\le \|f_{1,1}\|_2\big(|\mu|_3^{1/3}\|f_{1,2}\|_3 + 3\theta\|f_{1,2}\|_2\big) = \|f_{1,1}\|_2 D(f_{1,2}).$$

Therefore, in the case $|G_1| = 1, |I_i| > 1, i = 1, 2, 3$,

$$p_{I,1} \le D^{\hat{I}}\Big(\|f_{1,1}\|_2 D_{2,1}D_{3,1} + \|f_{2,1}\|_2 D_{1,1}D_{3,1} + \|f_{3,1}\|_2 D_{1,1}D_{2,1}\Big). \tag{B.13}$$

By (B.8), (B.9), (B.13), the induction step $|I|-1 \to |I|$ in the case $|I_i| > 1, i = 1, 2, 3$ follows from

$$\begin{aligned} &\|f_1\|_2\|f_2\|_2 D_3 + \|f_1\|_2\|f_3\|_2 D_2 + \|f_2\|_2\|f_3\|_2 D_1 + |\mu|_3\|f_1\|_3\|f_2\|_3\|f_3\|_3 \\ &\quad + \|f_1\|_2 D_2 D_3 + \|f_2\|_2 D_1 D_3 + \|f_3\|_2 D_1 D_2 \ \le\ D_1 D_2 D_3, \end{aligned} \tag{B.14}$$

where we put $f_i \equiv f_{i,1}, D_i \equiv D_{i,1} = D(f_{i,1})$. To prove (B.14), put $x_i := \|f_i\|_2$, $y_i := |\mu|_3^{1/3}\|f_i\|_3$. Then (B.14) can be rewritten as

$$F(x_1, x_2, x_3, y_1, y_2, y_3) \ge 0, \tag{B.15}$$

where

$$\begin{aligned} F(x_1, x_2, x_3, y_1, y_2, y_3) &:= \prod_{i=1}^{3}(y_i + 3\theta x_i) - y_1y_2y_3 \\ &- \frac{1}{2}\sum_{i\ne j\ne k} x_i x_j (y_k + 3\theta x_k) \\ &- \frac{1}{2}\sum_{i\ne j\ne k} x_i (y_j + 3\theta x_j)(y_k + 3\theta x_k) \end{aligned}$$

and the sum $\sum_{i\ne j\ne k}$ is taken over all $i, j, k = 1, 2, 3, i \ne j \ne k$.

To prove (B.15), note that $F(x_1, x_2, x_3, 0, 0, 0) = 9\theta x_1x_2x_3(3\theta^2 - 1 - 3\theta) = 0$ by the definition of θ. Next, with $X_i = 3\theta x_i$,

$$\begin{aligned} F(x_1, x_2, x_3, y_1, y_2, y_3) &= F(x_1, x_2, x_3, 0, 0, 0) + \frac{1}{2}\sum_{i\ne j\ne k} y_i y_j (X_k - x_k) \\ &+ \frac{1}{2}\sum_{i\ne j\ne k} y_i (X_jX_k - x_jX_k - x_kX_j - x_jx_k) \end{aligned}$$

and (B.15) follows from the easily verified relations $X_i - x_i \ge 0,\ X_jX_k - x_jX_k - x_kX_j - x_jx_k \ge 0$.

It remains to prove the induction step $|I| - 1 \to |I|$ in the case when one of the rows $I_i, i = 1, 2, 3$ has only one element. Let, for example, $|I_1| = 1, |I_2| > 1, |I_3| > 1$. Then (B.10) becomes

$$\sum\nolimits_{\tilde{S}_0}^{\gamma} E|\zeta^{S_0}|\Big(\sum\nolimits_{\tilde{S}_1}^{\gamma}\big|F_{S_I}((t))\big|^2\Big)^{1/2} \le \|f_{1,1}\|_2 \sum\nolimits_{\tilde{S}'_0}^{\gamma'} E|\zeta^{S'_0}|\Big(\sum\nolimits_{\tilde{S}'_1}^{\gamma'}\big|F_{S_{I'}}((t))\big|^2\Big)^{1/2},$$

yielding again the bound (B.13). The remaining cases can be considered similarly. This proves Lemma B.2. □

Now we apply Lemma B.2 to estimate third moments of infinite Volterra series. Let $f_{i,j} \in L^2(\mathbf{Z}_+^0), i = 1,2,3, j = 1,2,\dots$ be an infinite collection of functions satisfying

$$\bar{f} := \sup_{i,j} \|f_{i,j}\|_2 < 1.$$

Let

$$\Phi_i := \sum_{k=1}^{\infty} \sum_{s_k < \dots < s_1 < t_i} f_{i,1}(t_i - s_1) \dots f_{i,k}(s_{k-1} - s_k)\zeta_{s_1} \dots \zeta_{s_k} \equiv \sum_{k=1}^{\infty} \Phi_i^{(k)}, \tag{B.16}$$

$i = 1,2,3, t_i \in \mathbf{Z}$ be (infinite) Volterra series, which converge in mean square by orthogonality:

$$E\Phi_i^2 = \sum_{k=1}^{\infty} \prod_{j=1}^{k} \|f_{i,j}\|_2^2 \le \bar{f}^2/(1 - \bar{f}^2) < \infty.$$

For given $N < \infty$, let $\Phi_{i,N}$ be defined analogously to Φ_i (B.16), where the $f_{i,j}$ are replaced by truncated functions

$$f_{i,j}^{(N)}(s) = \begin{cases} f_{i,j}(s), & \text{if } 1 \le s \le N, \\ 0, & \text{if } s > N. \end{cases}$$

Lemma B.3. *Assume that*

$$\bar{D} := \sup_{i,j} D(f_{i,j}) < 1. \tag{B.17}$$

Then

$$E\big|\Phi_1\Phi_2\Phi_3\big| \le (\bar{D}/(1 - \bar{D}))^3 < \infty. \tag{B.18}$$

Furthermore,

$$E\Phi_1\Phi_2\Phi_3 \;=\; \lim_{N\to\infty} E\Phi_{1,N}\Phi_{2,N}\Phi_{3,N}. \tag{B.19}$$

Proof. We have

$$E\prod_{i=1}^{3} |\Phi_i| \le \sum_{k_1,k_2,k_3=1}^{\infty} E\prod_{i=1}^{3} |\Phi_i^{(k_i)}|.$$

According to Lemma B.1, the last expectation does not exceed $\bar{D}^{k_1+k_2+k_3}$, thereby proving (B.18).

Next, note that relation (B.19) follows from

$$\lim_{N\to\infty} E\Phi_{1,N}\Phi_2\Phi_3 = E\Phi_1\Phi_2\Phi_3, \tag{B.20}$$

where $f_{i,j}, i = 2,3, j \ge 1$ may depend on N and satisfy (B.17).

To prove (B.20), for an integer L put $\Phi_{i,N}^- := \sum_{k=1}^{L} \Phi_{i,N}^{(k)}, \Phi_{i,N}^+ := \sum_{k=L+1}^{\infty} \Phi_{i,N}^{(k)}$. Then $\prod_{i=1}^{3} \Phi_{i,N} = \prod_{i=1}^{3} \Phi_{i,N}^- + R_N$, where, by Lemma B.1,

$$E|R_N| \le \sum_{k_1 > L, k_2, k_3 \ge 1} E\prod_{i=1}^{3} |\Phi_{i,N}^{(k_i)}| \le \sum_{k_1 > L, k_2, k_3 \ge 1} \bar{D}^{k_1+k_2+k_3} \le \frac{\bar{D}^{L+3}}{(1 - \bar{D})^3}$$

vanishes as $L \to \infty$ uniformly in N. Consequently, $|E(\Phi_1 - \Phi_{1,N})\Phi_2\Phi_3| \le |E(\Phi_1^- - \Phi_{1,N}^-)\Phi_2\Phi_3| + o(1)$ uniformly in N, so that (B.20) follows from $\lim_{N\to\infty} E(\Phi_1^- - \Phi_{1,N}^-)\Phi_2\Phi_3 = 0$ for each $L < \infty$. In turn, the last relation follows from

$$\lim_{N\to\infty} E(\Phi_1^{(k)} - \Phi_{1,N}^{(k)})\Phi_2\Phi_3 = 0. \tag{B.21}$$

for each $1 \le k < \infty$. The difference in (B.21) can be written as $\Phi_1^{(k)} - \Phi_{1,N}^{(k)} = \sum' U_{i_1,\ldots,i_k}$, where the sum $\sum'$ is taken over all $i_1, \ldots, i_k$ taking values 0,1 and such that $i_1 + \ldots + i_k \ge 1$,

$$U_{i_1,\ldots,i_k} := \sum_{s_k < \ldots < s_1 < t_1} g_1^{(i_1)}(t_1 - s_1) \ldots g_k^{(i_k)}(s_{k-1} - s_k)\zeta_{s_1} \ldots \zeta_{s_k}$$

and where $g_j^{(1)} := f_{1,j}^{(N)}, g_j^{(0)} := f_{1,j} - f_{1,j}^{(N)}$. Then $|E(\Phi_1^{(k)} - \Phi_{1,N}^{(k)})\Phi_2\Phi_3| \le \sum' \sum_{k_2,k_3=1}^{\infty} E|U_{i_1,\ldots,i_k}\Phi_2^{(k_2)}\Phi_3^{(k_3)}|$, where, by Lemma B.1, the last expectation does not exceed $\prod_{j=1}^{k} D(g_j^{(i_j)})\bar{D}^{k_2+k_3}$. Now, relation (B.21) follows from $D(g_j^{(i)}) \le D(f_{1,j}) \le \bar{D}\,(i = 0, 1)$ and $D(g_j^{(1)}) \to 0$ $(N \to \infty)$. This proves Lemma B.3. □

Proof of Proposition 2.2. It suffices to show $EV_0^4 < \infty$. We shall assume $a = 1$ without loss of generality. Let

$$U_N = 1 + \sum_{k=1}^{N-1} U_{k,N},$$

where

$$U_{k,N} := \sum_{-N < s_k < \ldots < s_1 < 0} b_{-s_1} b_{s_1 - s_2} \ldots b_{s_{k-1} - s_k} \zeta_{s_1} \ldots \zeta_{s_k}.$$

Then $U_N \to V_0$ in $L^2(\Omega)$ as $N \to \infty$ and therefore by Fatou lemma,

$$EV_0^4 \le \sup_{N \ge 1} EU_N^4.$$

Hence the proposition follows if for all $N \ge 1$

$$EU_N^4 \le K. \tag{B.22}$$

The inequality (B.22) follows from the following statement: there exist constants $K_1 < \infty, 0 < D < 1$ independent of $N \ge 1$ and such that for any integers $k_1, \ldots, k_4 \ge 0$

$$\big|E \prod_{i=1}^{4} U_{k_i,N}\big| \;\le\; K_1 D^{|(k)_4|}, \tag{B.23}$$

where $(k)_4 = (k_1, \ldots, k_4), |(k)_4| = k_1 + \ldots + k_4$, and $U_{N,0} := 1$. Indeed, if (B.23) is true, then

$$EU_0^4 \le \sum_{(k)_4} \big|E \prod_{i=1}^{4} U_{k_i,N}\big| \le K_1 \sum_{(k)_4} D^{|(k)_4|} = \frac{K_1}{(1 - D)^4} \le C < \infty.$$

To prove (B.23), write

$$E \prod_{i=1}^{4} U_{k_i,N} = \sum\nolimits_{(S)_4}^{(k)_4} b^{S_1} \ldots b^{S_4} E[\zeta^{S_1} \ldots \zeta^{S_4}],$$

where the sum $\sum_{(S)_4}^{(k)_4}$ is taken over all collections $(S)_4 = (S_1, \ldots, S_4), S_i \subset \{-N+1, \ldots, -1\} =: T_N \subset \mathbf{Z}$ consisting of sets of ordered indexes $S_i = \{s_{k_1}, s_{k_{i-1}}, \ldots, s_1\}$, $s_{k_1} < \ldots < s_1$, $i = 1, \ldots, 4$, and where, for each such subset $S = \{s_k, \ldots, s_1\} \subset T_N$,

$$b^S := b_{-s_1} b_{s_1 - s_2} \ldots b_{s_{k-1} - s_k}, \quad \zeta^S := \zeta_{s_1} \ldots \zeta_{s_k},$$

$b^{\emptyset} = \zeta^{\emptyset} := 1$. Put

$$p_{(k)_4} := \sum\nolimits_{(S)_4}^{(k)_4} |b^{S_1} \ldots b^{S_4}|\,|E[\zeta^{S_1} \ldots \zeta^{S_4}]|.$$

Obviously, it suffices to show (B.23) with the left hand side of (B.23) replaced by $p_{(k)_4}$. We use induction in $|(k)_4| = k_1 + \ldots + k_4$. Let first $k_1, \ldots, k_4 \geq 1$. Put $\underline{s}_i = s_{i,k_i}, i = 1, \ldots, 4$ then $E[\zeta^{S_1} \ldots \zeta^{S_4}] = 0$ if $\underline{s}_i, i = 1, \ldots, 4$ are all different. Hence

$$p_{(k)_4} = \sum_{2 \leq |A| \leq 4} |\mu_{|A|}| \sum\nolimits_{(S')}^{(k')_4} |b^{S'_1} \ldots b^{S'_4}| \, |E[\zeta^{S'_1} \ldots \zeta^{S'_4}]| \sum_{\underline{s}} \prod_{i \in A} |b_{s_{i,k'_i} - \underline{s}}|,$$

where the sum $\sum_{2 \leq |A| \leq 4}$ is taken over all subsets $A \subset \{1, \ldots, 4\}, 2 \leq |A| \leq 4$, and, for any such subset A, $k'_i := k_i - 1$ if $i \in A$ and $k'_i := k_i$ otherwise. By Hölder's inequality,

$$\sum_{\underline{s}} \prod_{i \in A} |b_{s_{i,k'_i} - \underline{s}}| \leq \sum_{s} |b_s|^{|A|} \equiv \|b\|_{|A|}^{|A|}.$$

Then we obtain for $p_{(k)_4}$ the following recursive relation

$$p_{(k)_4} \leq \mu_2 \|b\|_2^2 \sum_{|A|=2} p_{(k')_4} + |\mu_3| \|b\|_3^3 \sum_{|A|=3} p_{(k')_4} + \mu_4 \|b\|_4^4 \sum_{|A|=4} p_{(k')_4},$$

where the last sum consists of a single term, of course.

By using the inductive assumption for $p_{(k')_4}, |(k')_4| \leq |(k)_4| - 1$, we obtain

$$p_{(k)_4} \leq K_1 D^{|(k)_4|} \Big(6\mu_2 \|b\|_2^2 D^{-2} + 4|\mu_3| \|b\|_3^3 D^{-3} + \mu_4 \|b\|_4^4 D^{-4}\Big). \tag{B.24}$$

The constant D in (B.24) can be chosen arbitrarily close to 1, in particular, in view of (2.4), we can choose $D < 1$ such that

$$6\mu_2 \|b\|_2^2 D^{-2} + 4|\mu_3| \|b\|_3^3 D^{-3} + \mu_4 \|b\|_4^4 D^{-4} \leq 1. \tag{B.25}$$

Hence (B.24) implies $p_{(k)_4} \leq K_1 D^{|(k)_4|}$, thereby proving the induction step $|(k)_4| - 1 \to |(k)_4|$.

Assume now that $k_4 = 0, k_1, k_2, k_3 \geq 1$. Then in a similar way we obtain instead of (B.24) the relation

$$p_{(k)_4} \leq K_1 D^{|(k)_4|} \big(3\mu_2 \|b\|_2^2 D^{-2} + |\mu_3| \|b\|_3^3 D^{-3}\big),$$

which again proves the induction step $|(k)_4| - 1 \to |(k)_4|$ by (B.25). The case $k_3 = k_4 = 0, k_1 \geq 1, k_2 \geq 1$ follows easily. The above argument also proves the bound (B.23) for $0 \leq k_1, \ldots, k_4 \leq 1$, by verifying that $p_{(k)_4} \leq K_1 < \infty$. □

References

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

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

Barndorff-Nielsen, O.E. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck - based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society*, Series B, **63**, 167–241.

Black, F. (1976). Studies in stock price volatility changes. *Proceedings of the Business and Economic Statistics Section*, 177–181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, **3**, 307–327.

Bollerslev, T. and Mikkelsen, H.O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, **73**, 151–184.

Bouchaud, J.-P., Matacz, A. and Potters, M. (2001). Leverage effect in financial markets: the retarded volatility model. *Physical Review Letters*, **87**, 228701-1–228701-4.

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

Campbell, J.Y., and Hentschel, L. (1992). No news is good news. An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, **31**, 281–318.

Comte, F. and Renault, E. (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, **8**, 291–323.

Demos, A. (2002). Moments and dynamic structure of a time-varying parameter stochastic volatility in mean model. *Econometrics Journal*, **5**, 345–357.

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

Engle, R.F. (1990). Stock volatility and the crash of '87. Discussion. *The Review of Financial Studies*, **3**, 103–106.

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

Engle, R.F. and Ng, V.K. (1993). Measuring and testing the impact of news on volatility. *The Journal of Finance*, **48**, 1749–1778.

Giraitis, L. and Robinson, P. M. (2001). Whittle estimation of ARCH models. *Econometric Theory*, **17**, 608–631.

Giraitis, L., Kokoszka, P. and Leipus, R. (2000). Stationary ARCH models: dependence structure and Central Limit Theorem. *Econometric Theory*, **16**, 3–22.

Giraitis, L., Robinson, P.M. and Surgailis, D. (2000). A model for long memory conditional heteroskedasticity. *Annals of Applied Probability*, **10**, 1002–1024.

Glosten, L.R., Jagannathan R. and Runkle D.E. (1993). On the relation between the expected value of the volatility and the volatility of the nominal excess return on stocks. *The Journal of Finance*, **48**, 1779–1801.

Harvey, A. (1998). Long memory in stochastic volatility. In *Forecasting Volatility in the Financial Markets* (J. Knight and S. Satchell, eds.), pp. 307–320. Butterworth & Heineman, Oxford.

He, C., Teräsvirta, T. and Malmsten, H. (2002). Moment structure of a family of first-order exponential GARCH models. *Econometric Theory*, **18**, 868–885.

Karanasos, M. and Kim, J. (2001). Moments of the ARMA-EGARCH model. *Econometrics Journal*, **6**, 146-166.

Lee, A.W. and Hansen, B.E. (1994). Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory*, **10**, 29–52.

Lumsdaine, R. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models. *Econometrica*, **16**, 575–596.

Müller, U.A., Dacorogna, M. M., Dave, R.D., Olsen, R.B., Pictet, O.V. and von Weizsäcker, J.E. (1997). Volatilities of different time resolutions - Analyzing the dynamics of market components. *Journal of Empirical Finance.* **4**, 213–240.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, **59**, 347–370.

Pagan, A. (1996). The econometrics of financial markets. *Journal of Empirical Finance*, **3**, 15–102.

Palma, W. and Zevallos, M. (2002). Analysis of the correlation structure of square time series. *Preprint.*

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

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

Schwert, G.W. (1990). Stock volatility and the crash of '87. *The Review of Financial Studies*, **3**, 77–102.

Sentana, E. (1995). Quadratic ARCH models. *Review of Economic Studies*, **62**, 639–661.

Surgailis, D. and Viano, M.-C. (2002). Long memory properties and covariance structure of the EGARCH model. *ESAIM: Probability and Statistics*, **6**, 311–329.

Taylor, S. (1986). *Modelling Financial Time Series.* Wiley, New York.

Tse, Y.K. (1998). The conditional heteroskedasticity of the Yen-Dollar exchange rate. *Journal of Applied Econometrics*, **13**, 49–55.

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

Zaffaroni, P. (1998). Gaussian inference in certain long-range dependent volatility models. *Preprint.*

Zakoian, J.M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control.* **18**, 931–955.

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

Remigijus Leipus, Vilnius University, Department of Mathematics and Informatics, Naugarduko 24, Vilnius 2600, Lithuania. *email:* Remigijus.Leipus@maf.vu.lt

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

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