Aggregation and memory of models of changing volatility

Paolo Zaffaroni *

Banca d' Italia

August 2001

Abstract

In this paper we study the effect of contemporaneous aggregation of an arbitrarily large number of processes featuring dynamic conditional heteroskedasticity with short memory when heterogeneity across units is allowed for. We look at the memory properties of the limit aggregate. General, necessary, conditions for long memory are derived. More specific results relative to certain stochastic volatility models are also developed, providing some examples of how long memory volatility can be obtained by aggregation.

JEL classification: C43

Keywords: stochastic volatility, contemporaneous aggregation, long memory.

*Address correspondence:
Servizio Studi, Banca d'Italia, Via Nazionale 91, 00184 Roma Italy, tel. + 39 06 4792 4178, fax. + 39 06 4792 3723, email `zaffaroni.paolo@insedia.interbusiness.it`

1 Introduction

Contemporaneous aggregation, in the sense of averaging across units, of stationary heterogeneous autoregressive moving average (ARMA) processes can lead to a limit stationary process displaying long memory, in the sense of featuring non summable autocovariance function, when the number of units grows to infinity (see Robinson (1978) and Granger (1980)).

Relatively recent research in empirical finance indicates that the long memory paradigm represents a valid description of the dependence of volatility of financial asset returns (see Ding, Granger, and Engle (1993), Granger and Ding (1996) and Andersen and Bollerslev (1997) among others). In most studies the time series of stock indexes has been used, such as the Standard & Poor's 500, to support this empirical evidence, naturally suggesting that the aggregation mechanism could be the ultimate source of long memory in the volatility of portfolio returns.

The strong analogies of the generalized autoregressive conditionally heteroskedasticity (GARCH) model of Bollerslev (1986) with ARMA naturally suggests that arithmetic averaging of an arbitrary large number of heterogeneous GARCH could lead to long memory ARCH, namely the ARCH(∞) with long memory parameterizations (see Robinson (1991) and Baillie, Bollerslev, and Mikkelsen (1996)). It turns out that under particular 'singularity' conditions on the GARCH coefficients, yielding perfect negative covariation between the latter, the squared limit aggregate is characterized by an hyperbolically decaying autocovariance function (acf) yet summable, a situation of 'quasi' long memory (see Ding and Granger (1996) and Leipus and Viano (1999)). A closer analysis shows that the result is quite disappointing though. Under no conditions long memory, in the sense of non-summable acf of the squared aggregate (hereafter long memory for brevity), could be obtained by aggregation of GARCH. Moreover the acf of the squared limit aggregate decays exponentially fast in general, except when the above mentioned 'singularity' condition holds. Third, the limit aggregate is not ARCH(∞). (See Zaffaroni (2000) for details.) This is the outcome of the severe parametric structure of GARCH processes by which finiteness of moments defines a region for the GARCH parameters which varies with the order of the moment. In particular, the larger is the order of the moment the smaller is this 'stationary' region. For example, for ARCH(1) with unit variance innovations the autoregressive parameter region for bounded second moment is the interval $(0, 1)$ and for bounded fourth moment is $(0, 1/\sqrt{3})$, strictly included in $(0, 1)$. We will refer to this outcome as the fourth-moment restriction.

In this paper we study the memory implications of the aggregation mechanism in a wider perspective. GARCH by no means represent the only successful way to describe dynamic conditional heteroskedasticity and other nonlinear time series models have been nowadays popular in fitting financial asset returns such as stochastic volatility (SV) models (see Ghysels, Harvey, and Renault (1996) for a complete survey). In contrast to ARCH-type models SV models are characterized by a latent state variable determining the dynamics of the volatility of the return process. This feature induces a great deal of issues related to estimation of SV models. However, as we are not concerned here with estimation and filtering, it would be convenient to study the aggregation mechanism within a large class of volatility models which nests both ARCH-type and SV-type models. With this respect, a convenient approach is given by the class of the square root stochastic autoregressive volatility (SR-SARV) models, introduced by Andersen (1994) and generalized by Meddahi and Renault (1996). A unified analysis of SV models is also developed in Robinson (1999) who provides an asymptotic expansion for the acf of a large class of SV models just requiring Gaussian innovations. This class excludes many ARCH-type models, including GARCH.

For the case of a finite number of units n, aggregation of GARCH has been analyzed by Nijman and Sentana (1996) and generalized by Meddahi and Renault (1996) to aggregation of SR-SARV. These results establish the conditions under which the aggregate maintains the same parametric structure of the micro units. In contrast, the main focus of this paper is to characterize the conditions under which the aggregate displays different features from the ones of the micro units, such as long memory, by letting $n \to \infty$.

First, we derive a set of necessary conditions for long memory with respect to the SR-SARV class. The first finding is that a necessary condition for long memory is that the micro units must be sufficiently cross-sectionally correlated. For instance, aggregation of independent and identically distributed (*i.i.d.*) cross-sectionally units yields under mild conditions a Gaussian noise limit, a case of no-memory. Second, we provide a necessary condition such that the fourth-moment restriction does not arise. Many volatility models used nowadays in empirical finance happen to violate such condition. These conditions are not sufficient, though, for long memory. In fact, we then focus on three particular models, all belonging to the SR-SARV class. The models are the exponential SV model of Taylor (1986), a linear SV and the nonlinear moving average model (nonlinear MA) of Robinson and Zaffaroni (1998). For all cases we assume that the innovations are common across units, yielding highly cross-correlated units. Moreover, for the three models the fourth-moment restrictions is not binding. Therefore, the necessary conditions for long memory are satisfied. However, these models deliver very different outcomes in terms of aggregation. In fact, a further, important feature of the aggregation mechanism emerges, meaning the shape of the nonlinearity specific to the model. This determines the minimal conditions required for existence and strict stationarity of the limit aggregate which can influence the possibility of long memory. These conditions are expressed in terms of the shape of the cross-sectional distribution of the parameters of the micro processes. It turns out that for both the exponential SV and linear SV the strict stationarity conditions rule out long memory. In particular, the long memory SV models of Harvey (1998), Comte and Renault (1998) and Breidt, Crato, and de Lima P. (1998) cannot be obtained by aggregation of short memory exponential SV. However, whereas for the exponential SV the acf will always decay exponentially fast, the limit aggregate of linear SV can exhibit hyperbolically decaying acf. In contrast aggregation of short memory nonlinear MA can yield a limit aggregate displaying long memory, with a non summable acf of the squares or, equivalently, an unbounded spectral density of the squares at frequency zero. In particular, the long memory nonlinear MA of Robinson and Zaffaroni (1998) is obtained through aggregation of short memory nonlinear MA.

This paper proceeds as follows. Section 2 develops a set of necessary conditions for long memory with respect the SR-SARV class. Section 3 focuses on the three models above described, exponential SV, linear SV and nonlinear MA. Concluding remarks are in section 4. The results are stated in propositions whose proofs are reported in the final appendix.

2 Some general results

Let us first recall the simplest definition of SR-SARV model. Summarizing Meddahi and Renault (1996, Definition 3.1) a stationary square integrable process $\{x_t\}$ is called a SR-SARV(1) process with respect to the increasing filtration J_t if x_t is J_t-adapted, $E(x_t \mid J_{t-1}) = 0$ and $\text{var}(x_t \mid J_{t-1}) := f_{t-1}$

satisfying

$$f_t = \omega + \gamma f_{t-1} + v_t, \tag{1}$$

with the sequence $\{v_t\}$ satisfying $E(v_t \mid J_{t-1}) = 0$. ω and γ are constant non-negative coefficients with $\gamma < 1$. This implies $Ex_t^2 < \infty$.

In order to study the impact of aggregation over an arbitrarily large number of units, we assume that at each point in time we observe n units $x_{i,t}\,(1 \leq i \leq n)$, each parameterized as a SR-SARV(1), with coefficients ω_i, γ_i being *i.i.d.* random draws from a joint distribution such that $\gamma_i < 1$ almost surely (*a.s*). The $x_{i,t}$ represent random coefficients SR-SARV.

We now establish two set of conditions which, independently, rule out the possibility of long memory. The violation of such conditions would then represent necessary conditions for inducing long memory. We first consider the impact of the assumed degree of cross-sectional dependence across units for the $x_{i,t}$.

Proposition 2.1 *Assume the $x_{i,t}$ are i.i.d. across units. When*

$$E(x_{i,t}^2) =: \sigma^2 < \infty, \tag{2}$$

then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_{i,t} \rightarrow_d X_t, \quad \text{as } n \rightarrow \infty,$$

where the X_t are $N(0, \sigma^2)$ distributed, mutually independent, and $\rightarrow_d$ denotes convergence in the sense of the finite dimensional distribution.

Remarks.

(a) Under *i.i.d.* $x_{i,t}$, the limit aggregate features no memory.

(b) Proposition 2.1 could be extended to the case of cross-sectionally dependent $x_{i,t}$ as long as the central limit theorem (CLT) holds, as e.g. for many cases of association (see Esary, Proschan, and Walkup (1967)).

(c) Asymptotic normality follows by *i.i.d.*-ness and the bounded second moment assumption. Then, mutual independence of the limit aggregate follows by the martingale difference assumption, which represents the key assumption, and normality.

(d) Condition (2)

$$\sigma^2 = E\frac{\omega_i}{1-\gamma_i} < \infty,$$

can fail when ω_i and γ_i are random variables, even if $\gamma_i < 1$ *a.s.* For instance, this happens (see Lippi and Zaffaroni (1998, Lemma 1)) when ω_i and γ_i are independent one another and γ_i is absolutely continuous with density

$$f_\gamma(\gamma_i) \sim c\,(1-\gamma_i)^\delta, \quad \text{as } \gamma_i \to 1^- \tag{3}$$

for real $-1 < \delta \leq 0$, where $\sim$ denotes asymptotic equivalence ($a(x) \sim b(x)$ as $x \to x_0$ when $a(x)/b(x) \to 1$ as $x \to x_0$).

When (2) fails, a result analog to Proposition 2.1 holds. For this, we need to assume that the SR-SARV admits a solution so that x_t can be written as a nonlinear moving average

$$x_t = u_t f(\epsilon_{t-1},\, \epsilon_{t-2}, ..., \epsilon_1; \omega,\, \gamma) + u_t g(\epsilon_0,\, \epsilon_{-1}, ...; \omega,\, \gamma), \tag{4}$$

for suitable functions $f(\cdot, .., \cdot; \omega,\, \gamma)$, $g(\cdot, .., \cdot; \omega,\, \gamma)$ such that $g(0, ..., 0; \omega,\, \gamma) = 0$, where $\{u_t, \epsilon_t\}$ denote a bivariate sequence satisfying

$$\begin{pmatrix} u_t \\ \epsilon_t \end{pmatrix} \sim i.i.d. \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{u\epsilon} \\ \sigma_{u\epsilon} & \sigma_\epsilon^2 \end{pmatrix} \right). \tag{5}$$

The u_t and ϵ_t can be perfectly correlated, including case $u_t = \epsilon_t$. Existence of the solution (4) for x_t is not implied by the definition of SR-SARV but requires stronger conditions. This is shown in four examples below.

Set the conditional model, allowing for heterogeneity, equal to

$$\tilde{x}_{i,t} := u_{i,t}\, f(\epsilon_{i,t-1},\, \epsilon_{i,t-2}, ..., \epsilon_{i,1}; \omega_i,\, \gamma_i),$$

for sequences $\{u_{i,t}, \epsilon_{i,t}\}$ satisfying (5) for each i and *i.i.d.* across units.

When $\sigma_t^2 := Ef^2(\epsilon_{i,t-1}, ..., \epsilon_{i,1}; \omega_i,\, \gamma_i) < \infty$ for any $t < \infty$ and *i.i.d.*-ness,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{x}_{i,t} \to_d \tilde{X}_t \quad \text{as } n \to \infty, \tag{6}$$

where the $\tilde{X}_t$ are $N(0, \sigma_t^2)$ distributed, mutually uncorrelated and σ_t^2 converges to σ^2 or diverges to infinity, as $t \to \infty$, depending on whether (2) holds or fails.

Example 1 (Zaffaroni 2000). The ARCH(1)

$$x_t = u_t \left(\mu + \alpha x_{t-1}^2 \right)^{\frac{1}{2}} \tag{7}$$

belongs to the SR-SARV(1) class setting $\omega = \mu$, $\gamma = \alpha$ and

$$v_t = \alpha\,(u_t^2 - 1) f_{t-1}.$$

When (2) holds then Proposition 2.1 applies. When (2) fails, setting $\epsilon_t = u_t$ and

$$\tilde{x}_{i,t} = u_{i,t}\,\omega_i^{\frac{1}{2}} \left(\sum_{k=0}^{t-1} \gamma_i^k \prod_{j=1}^{k} u_{i,t-j}^2 \right)^{\frac{1}{2}},$$

(6) holds with

$$\sigma_t^2 = E\left(\omega_i \frac{1-\gamma_i^t}{1-\gamma_i} \right).$$

Example 2. The simplest version of SR-SARV(1) is the linear SV, given by

$$x_t = u_t \sqrt{f_{t-1}} \tag{8}$$

and

$$f_t = \mu + \alpha f_{t-1} + \epsilon_t, \tag{9}$$

where

$$-\nu \leq \ \epsilon_t < \infty \ \ a.s. \tag{10}$$

for some constant $0 < \nu \leq \mu < \infty$ ensuring $f_t \geq 0$ $a.s.$ Finally u_t and ϵ_t are assumed mutually independent, implying $\sigma_{u\epsilon} = 0$. This rules out the possibility that linear SV nests GARCH$(1,1)$ although the former admits a weak GARCH$(1,1)$ representation (see Meddahi and Renault (1996, Proposition 2.7)) when further assuming $E(u_t^3) = 0$. Proposition 2.1 applies when (2) holds setting $\omega = \mu$, $\gamma = \alpha$ and $v_t = \epsilon_t$. In contrast, when the latter fails, set

$$\tilde{x}_{i,t} = u_{i,t} \left(\omega_i \frac{1-\gamma_i^{t-1}}{1-\gamma_i} + \sum_{k=0}^{t-2} \gamma_i^k \epsilon_{i,t-k-1} \right)^{\frac{1}{2}}.$$

Then (6) holds with

$$\sigma_t^2 = E\left(\omega_i \frac{1-\gamma_i^{t-1}}{1-\gamma_i} \right).$$

Example 3. The nonlinear moving average model, henceforth nonlinear MA, introduced by Robinson and Zaffaroni (1998), given in its simplest formulation by

$$x_t = u_t \, (\sum_{k=1}^{\infty} \alpha^k \epsilon_{t-k}), \tag{11}$$

with $|\alpha| < 1$, belongs to the SR-SARV(1). Define the stationary AR(1)

$$h_t = \alpha h_{t-1} + \epsilon_t. \tag{12}$$

Then consider (1) setting $f_t = \alpha^2 h_t^2$ with

$$v_t = \alpha^2(\epsilon_t^2 - \sigma_\epsilon^2) + 2\alpha^3 \epsilon_t h_{t-1}$$

and $\omega = \alpha^2 \sigma_\epsilon^2$, $\gamma = \alpha^2$. When (2) holds, then Proposition 2.1 applies whereas when the former fails set

$$\tilde{x}_{i,t} = u_{i,t} \left(\sum_{k=1}^{t-1} \alpha_i^k \epsilon_{i,t-k} \right)$$

and (6) holds with

$$\sigma_t^2 = E\left(\omega_i \frac{1 - \gamma_i^{t-1}}{1 - \gamma_i} \right).$$

Example 4. The exponential SV(1) of Taylor (1986)

$$x_t = u_t e^{\frac{1}{2} h_{t-1}}, \tag{13}$$

with h_t defined in (12) and

$$\text{Gaussian } \epsilon_t, \tag{14}$$

belongs to the SR-SARV(∞) class (see Meddahi and Renault (1996, section 6) for details). Gaussianity of the ϵ_t suggested the standard denomination of model (12)-(13) as log normal SV.

Again, when (2) holds then Proposition 2.1 applies whereas when the former fails, setting

$$\tilde{x}_{i,t} = u_{i,t} \exp\left(\frac{1}{2} \sum_{k=0}^{t-2} \alpha_i^k \epsilon_{i,t-k-1} \right),$$

(6) holds with

$$\sigma_t^2 = E \exp\left(\frac{\sigma_\epsilon^2}{2} \frac{1 - \alpha_i^{t-1}}{1 - \alpha_i} \right).$$

We shall now evaluate the fourth-moment restriction.

Proposition 2.2 *Let*
$$E(x_t^4 \mid J_{t-1}) = c\, f_{t-1}^2$$
for some $0 < c < \infty$ *and the* v_t *be strictly stationary with*
$$E(v_t^2 \mid J_{t-1}) = g_{t-1} + \kappa\, f_{t-1}^2, \tag{15}$$
for real $\kappa \geq 0$ *and a* J_t*-measurable function* g_t *satisfying* $Eg_t < \infty$ *for any* $0 < \gamma < 1$. *Then*
$$\gamma^2 + \kappa < 1$$
represents the necessary condition for $Ex_t^4 < \infty$.

Remarks.
(a) Assuming heterogeneity, Proposition 2.2 implies that when $\kappa_i \geq c_\kappa > 0$ $a.s.$ for some constant c_κ, the possibility of long memory by aggregation will always be ruled out. In fact, the key feature which characterizes the memory of the volatility of the limit aggregate is the behaviour of $E\gamma_i^k$ as $k \to \infty$. Covariance stationary levels requires $0 \leq \gamma_i < 1$ $a.s.$ and
$$E\gamma_i^k \sim c\, k^{-(b+1)} \quad \text{as } k \to \infty$$
under (3) by Lemma 2. Therefore, a sequence of hyperbolic decaying coefficients is obtained as a by-product of aggregation, representing the ultimate source of long memory. However, when requiring stationarity of the conditional variance, then $0 \leq \gamma_i < (1 - c_\kappa)^{\frac{1}{2}}$ $a.s.$ and the support of the γ_i is strictly included in $[0, 1)$ for $0 < c_\kappa < 1$. Under these conditions
$$E\gamma_i^k \sim c\left((1 - c_\kappa)^{\frac{1}{2}}\right)^k k^{-(b+1)} = O(a^k) \quad \text{as } k \to \infty$$
for some $0 < a < 1$ by Lemma 2. Therefore, condition $c_\kappa > 0$ imparts the exponential behaviour of the $E\gamma_i^k$ which in turn gives rise to short memory of the limit aggregate.
(b) In most cases
$$g_t = c + c'\, f_t,$$
for constants $0 \leq c, c' < \infty$. Note that $Eg_t < \infty$ for any $0 < \gamma < 1$ rules out the case of g_t being an affine function of f_t^2.
Example 1 (cont.). For ARCH(1)
$$E(v_t^2 \mid J_{t-1}) = \gamma^2\, E(u_t^2 - 1)^2 f_{t-1}^2,$$

and (15) holds with $\kappa = \gamma^2 E(u_t^2 - 1)^2$ and $g_t = 0$. Long memory by aggregation is ruled out for ARCH (see Zaffaroni (2000, Theorem 4) for details).
Example 2 (cont.). For linear SV

$$E(v_t^2 \mid J_{t-1}) = Ev_t^2 = \sigma_\epsilon^2,$$

and (15) holds with $\kappa = 0$ and constant g_t. In this case boundedness of the fourth moment does not imply any restrictions of the parameter space.
Example 3 (cont.). For nonlinear MA, assuming for simplicity's sake $E\epsilon_t^3 = 0$,

$$E(v_t \mid J_{t-1}) = \alpha^4 \text{var}(\epsilon_t^2) + 4\alpha^4\, \sigma_\epsilon^2\, f_{t-1},$$

and (15) holds with $\kappa = 0$ and g_t being an affine function of f_t. Again, boundedness of the fourth moment does not require any restrictions of the parameter space.
Example 4 (cont.). For exponential SV(1) it can be easily seen that no parameter space restriction arises when imposing bounded fourth moment, given Gaussianity of the variance innovation ϵ_t. By direct calculations

$$Ex_t^2 = \exp(\frac{\sigma_\epsilon^2}{2}\frac{1}{1-\alpha^2})$$

and

$$Ex_t^4 = Eu_t^4\, \exp(2\,\sigma_\epsilon^2 \frac{1}{1-\alpha^2}),$$

and for both cases boundedness requires $\mid \alpha \mid < 1$.

To formally establish within to the SR-SARV framework the distribution of the limit aggregate, and therefore to determine its memory properties, many additional assumptions are required which would greatly restrict the generality of the approach. For this reason, in the next section, we rather study the effect of aggregation for the three specific SV models defined in Examples 2,3 and 4, all belonging to the SR-SARV class, and explore the possibility of long memory in details.

3 Some particular results

In this section we study in more detail the outcome of aggregation, characterizing the limit of

$$X_{n,t} := \frac{1}{n}\sum^{n} x_{i,t}$$

as $n \to \infty$. We focus on three models for the $x_{i,t}$: exponential SV, linear SV and nonlinear MA. These models all belong to the SR-SARV class and do not satisfy the conditions of Proposition 2.2 which yield the bounded fourth moment restriction. We allow for a sufficiently strong degree of cross-sectional dependence of the heterogeneous units so that Propositions 2.1 does not apply. In fact, we show that for any of the three model, the limit aggregate will not be normally distributed but conveys the basic feature of a volatility model, uncorrelatedness in levels with dependence in squares.

As previously indicated, it will be assumed that the micro parameters governing the $x_{i,t}$ are *i.i.d.* draws from some distribution. This will suitably describe a framework made of an arbitrarily large number of heterogeneous units. All of the models we consider share the same parameter α_i which expresses the memory of the volatility of the $x_{i,t}$. It will be assumed that $\alpha_i < 1$ *a.s.* so that the $x_{i,t}$ are covariance stationary with probability one. A parametric specification of the cross-sectional distribution of the parameters is not required and we rely on milder assumptions which define only the local behaviour of the cross-sectional distribution of the α_i around unity. Hereafter, let c define a bounded constant, not necessarily the same, and $\sim$ asymptotic equivalence. Let γ be a finite positive constant.

Assumption $A(\gamma)$**.** *The α_i are i.i.d. draws with an absolutely continuous distribution with support $[0, \gamma)$ and density*

$$f(\alpha) \sim c\, L(\frac{1}{\gamma - \alpha})\, (\gamma - \alpha)^{\delta}\, e^{-\frac{\beta}{(\gamma - \alpha^2)}}, \quad \text{as } \alpha \to \gamma^-, \tag{16}$$

for real $\beta \geq 0$ and $\delta > -1$ and slowly varying function $L(\cdot)$.

Remarks.

(a) Assumption $A(\gamma)$ includes a large class of parametric specifications of $f(\alpha)$ as particular cases. Simple examples are the uniform distribution, for $L(\cdot) = 1$ and $\beta = \delta = 0$, and the Beta distribution, for $L(\cdot) = 1$ and $\beta = 0$.

(b) When $\beta = 0$ (16) becomes

$$f(\alpha) \sim c\, L(\frac{1}{\gamma - \alpha})\, (\gamma - \alpha)^{\delta} \quad \text{as } \alpha \to \gamma^-, \tag{17}$$

and $\delta > -1$ ensures integrability.

(c) When $\beta > 0$ $f(\alpha)$ has a zero of exponential order at γ and

$$f(\alpha) = O\left((\gamma - \alpha)^c\right) = o(1) \quad \text{as } \alpha \to \gamma^-$$

for any $c > 0$ and $\delta > -\infty$.
(d) The case of no-heterogeneity across parameters, such as $\alpha_i = \alpha$, represents a particular case of our setting and will not be discussed.

Finally the following is required.

Assumption B. *The* $\{u_t, \epsilon_t\}$ *satisfy (5) with*

$$E\, u_t^4 < \infty, \;\; E\, \epsilon_t^4 < \infty.$$

3.1 Exponential SV

Assume that the $x_{i,t}$ are described by (12) and (13). We now derive the limit of $X_{n,t}$ in mean square. Under $A(\gamma)$, with $\gamma \leq 1$, the $x_{i,t}$ are strictly stationary, ergodic and, by Gaussianity of the ϵ_t, covariance stationary. The following result is based on the Hermite expansion of $X_{n,t}$ in function of the Hermite polynomials $H_m(\cdot)\,(m = 0, 1, ..)$ defined by

$$H_m(s)\phi(s) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (it)^m \phi(t) e^{-ist} dt \quad s \in \mathbb{R},$$

where $\phi(\cdot)$ denotes the standard normal density function and i is the complex unit ($i^2 = -1$).

Proposition 3.1 *Assume* $A(\gamma)$, *B and (14).*
(i) When $\gamma < 1$ *or* $\gamma = 1$, $\beta > \sigma_\epsilon^2/4$

$$X_{n,t} \to_2 X_t \quad \text{as } n \to \infty$$

where $\to_r$ *denotes convergence in* r*-th mean and*

$$X_t := u_t \sum_{r=0}^{\infty} \left(\frac{\sigma_\epsilon}{2}\right)^r N_r(t-1), \tag{18}$$

with

$$N_r(t) \quad := \sum_{\substack{i_0,..,i_r=1 \\ i_0+...+i_r=r \\ 1\,i_1+...+r\,i_r=r}}^{\infty} \frac{1}{0!^{i_0} ... r!^{i_r}}$$

$$\times \sum_{{}^1j_1 \neq ... \neq {}^1j_{i_1} \neq ... \neq {}^rj_1 \neq ... \neq {}^rj_{i_r}}^{\infty\,*} \zeta_{({}^1j_1+..+{}^1j_{i_1})+2({}^2j_1+..+{}^2j_{i_2})+...+r({}^rj_1+..+{}^rj_{i_r})}$$

$$\times \prod_{h_1=1}^{i_1} H_1(\tilde{\epsilon}_{t-{}^1j_{h_1}})... \prod_{h_r=1}^{i_r} H_r(\tilde{\epsilon}_{t-{}^rj_{h_r}}),$$

where $\sum_{{}^aj_0}^{*} = 1\,(a = 1, ..., r)$, $\tilde{\epsilon}_t := \epsilon_t/\sigma_\epsilon$ *and, for real* $k \geq 0$,

$$\zeta_k := E\,\exp(\frac{\sigma_\epsilon^2}{8}\frac{1}{(1-\alpha_i^2)})\alpha_i^k,$$

Under the above conditions

$$|\, X_t \,| < \infty \;\; a.s.$$

and the X_t *are both strictly and weakly stationary and ergodic.*
(ii) When $\gamma < 1$ *or* $\gamma = 1$, $\beta > \sigma_\epsilon^2/2$ *the* X_t^2 *are covariance stationary. Under these conditions*

$$\text{cov}(X_t^2, X_{t+h}^2) = O(c^h) \quad as\ h \to \infty,$$

for some $0 < c < 1$.

Remarks.
(a) Exponential SV do not exhibit the fourth-moment restriction, formalized in Proposition 2.2, allowing $\gamma = 1$. However the X_t^2 display short memory, with an exponentially decaying acf, for any possible shape of the cross-sectional distribution of the α_i. This rules out the the possibility of obtaining long memory exponential SV by aggregation of heterogeneous short memory exponential SV. The reason for this is that the exponential function characterizing (13) requires a compensation effect in terms of the behaviour of $f(\alpha)$. The latter must in fact decay exponentially fast toward zero as $\alpha \to 1^-$, requiring $\beta > 0$. When $\beta = 0$ the $|\, X_{n,t}\,|$ diverge to infinity in probability.
(b) No distributional assumption is required for the u_t so that the $x_{i,t}$ need not be conditionally Gaussian.
(c) The Hermite expansion was permitted by the Gaussianity assumption for the ϵ_t. This could be relaxed and use the more general expansion in terms of Appell polynomials (see Giraitis and Surgailis (1986)). The analysis would be more involved, given that Appell polynomials are in general non-orthogonal

(unlike Hermite polynomials) but the same restrictions in terms of shape of the coefficients' cross-sectional distribution $f(\alpha)$ are likely to arise.
(d) This result could be easily extended to the case of aggregation of exponential GARCH(1) (see Nelson (1991)), the one-shock analog to exponential SV. In this case $\epsilon_t = g(u_t)$ for some smooth function $g(\cdot)$ known as the news-impact-curve, a simple case of which is $g(u_t) = u_t$. More in general, the degree of asymmetry of the model plays no role in terms of the memory of the limit aggregate.
(e) The asymptotic behaviour of $X_{n,t}$ is prominently different from the one of the geometric mean aggregate

$$G_{n,t} := u_t \left(\prod_{i=1}^{n} e^{\frac{1}{2} h_{i,t-1}} \right)^{\frac{1}{n}}. \tag{19}$$

It turns out that $\mid G_{n,t} \mid$ represents a very mild lower bound for $\mid X_{n,t} \mid$. In fact, when $\beta = 0$ and (17) holds, the limit of $X_{n,t}$ is not well-defined, being unbounded in probability, whereas $G_{n,t}$ might still have a well-defined, strictly stationary, limit. For this purpose, simply note that

$$G_{n,t} = u_t \exp\left(\frac{1}{2\,n} \sum_{i=1}^{n} \frac{1}{1 - \alpha_i L} \epsilon_{t-1}\right),$$

(L denotes the lag operator: $L\,\epsilon_t = \epsilon_{t-1}$), and the aggregation results developed for linear ARMA models applies to the exponent in $(\cdot)$ brackets. This was noted in Andersen and Bollerslev (1997). When $\gamma = 1$, $\beta = 0$ and $\delta > -1/2$

$$G_{n,t} \to_p u_t e^{\frac{1}{2} \sum_{j=0}^{\infty} \nu_j \, \epsilon_{t-j-1}} =: G_t \quad \text{as } n \to \infty,$$

$\to_p$ denoting convergence in probability, as $1/n \sum_{i=1}^{n} h_{i,t}$ converges in mean-square to the linear stationary process $\sum_{j=0}^{\infty} \nu_j \, \epsilon_{t-j}$ (see Lippi and Zaffaroni (1998, Theorem 9)), with

$$\nu_k := E\alpha_i^k \sim c\, k^{-(\delta+1)} \quad \text{as } k \to \infty \tag{20}$$

by Lemma 2. Long memory is obtained when $\delta < 0$. G_t is a semiparametric generalization of the long memory SV model of Harvey (1998).

3.2 Linear SV

Assume that the $x_{i,t}$ satisfy (8) and (9). Assume that the μ_i are *i.i.d.* draws, mutually independent from the α_i, with $0 < c_\mu \leq \mu_i < \infty$ *a.s.*, for some positive constant $c_\mu > 0$, and $E(\mu_i) < \infty$. The following re-parameterization is useful. Setting

$$\pi_i := \mu_i - \nu,$$

with $0 \leq \pi_i < \infty$, and

$$\eta_t := \ln(\nu + \epsilon_t),$$

with support $(-\infty, \infty)$ under (10), yields

$$f_{i,t} = \frac{\pi_i}{1-\alpha_i} + \sum_{k=0}^{\infty} \alpha_i^k e^{\eta_{t-k}}. \tag{21}$$

By this re-parameterization both terms on the right hand side of (21) are non-negative. Existence of the second moment (and in general of the rth moment) of the ϵ_t implies a suitable restriction on the shape of distribution of the η_t but we will not make this explicit.

Proposition 3.2 *Assume $A(\gamma)$, B and (10).*
There exist processes $\{\underline{X}_{n,t}, \overline{X}_{n,t}, t \in \mathbb{Z}\}$ such that

$$\min[\underline{X}_{n,t}, \overline{X}_{n,t}] \leq X_{n,t} \leq \max[\underline{X}_{n,t}, \overline{X}_{n,t}] \ \ a.s., \tag{22}$$

satisfying the following.
(i) When $\gamma < 1$ or $\gamma = 1$, $\beta > 0$ or $\gamma = 1$, $\beta = 0$, $\delta > -1/2$ for $\underline{X}_{n,t}$ and $\gamma = 1$, $\beta = 0$, $\delta > -1/2$ for $\overline{X}_{n,t}$:

$$\underline{X}_{n,t} \to_1 \underline{X}_t, \ \ \overline{X}_{n,t} \to_1 \overline{X}_t \quad as\ n \to \infty, \tag{23}$$

setting

$$\underline{X}_t := u_t \left(E^2 \left(\frac{\pi_i}{1-\alpha_i} \right)^{\frac{1}{2}} + \sum_{k=0}^{\infty} \nu_{\frac{k}{2}}^2 e^{\eta_{t-k}} \right)^{\frac{1}{2}}, \tag{24}$$

$$\overline{X}_t := u_t \left(E \left(\frac{\pi_i}{1-\alpha_i} \right)^{\frac{1}{2}} + \sum_{k=0}^{\infty} \nu_{\frac{k}{2}} e^{\frac{1}{2}\eta_{t-k}} \right).$$

Under the above conditions the $\{\underline{X}_t, \overline{X}_t\}$ *satisfy*

$$|\underline{X}_t| < \infty, \quad |\overline{X}_t| < \infty \quad a.s.$$

and are both weakly and strictly stationary and ergodic.
(ii) When $\gamma < 1$ *or* $\gamma = 1$, $\beta > 0$ *the* $\underline{X}_t^2$ *and* $\overline{X}_t^2$ *are covariance stationary with*

$$\mathrm{cov}(\underline{X}_t^2, \underline{X}_{t+h}^2) = O(c^h), \quad \mathrm{cov}(\overline{X}_t^2, \overline{X}_{t+h}^2) = O(c^h) \quad as\ h \to \infty$$

for some $0 < c < 1$.
When $\gamma = 1$, $\beta = 0$ *the* $\underline{X}_t^2$, *for* $\delta > -1/2$, *and the* $\overline{X}_t^2$, *for* $\delta > 0$, *are covariance stationary with*

$$\mathrm{cov}(\underline{X}_t^2, \underline{X}_{t+h}^2) \sim c h^{-2(\delta+1)}, \quad \mathrm{cov}(\overline{X}_t^2, \overline{X}_{t+h}^2) \sim c\, h^{-(\delta+1)} \quad as\ h \to \infty.$$

Remarks.
(a) Linear SV represent another case where, despite the fourth moment restriction is not binding and thus $\gamma = 1$ is feasible, long memory is ruled out. The non-negativity constraint represents the key factor ruling the degree of memory of the limit aggregate. Note that the acf of the squared limit aggregate decays hyperbolically when $\gamma = 1$ although fast enough to achieve summability.
(b) No distributional assumption on the the volatility innovations ϵ_t was imposed. In fact, thanks to the simple structure of (21), we have characterized the limit of the 'envelope' processes $\{\underline{X}_{n,t}, \overline{X}_{n,t}\}$, rather than looking directly at the limit of $X_{n,t}$. The latter would be highly involved, as Proposition 3.1 indicates for the case of exponential SV, besides requiring distributional assumptions. This route was not permitted for the exponential SV problem of section 3.1. In fact, finding envelope processes well approximating the statistical properties of the limit aggregate seems unfeasible for exponential SV, due to the severe nonlinearity of the exponential function (cf. remark (e) to Proposition 3.1).
(c) Meddahi and Renault (1996, Theorem 4.1) show that aggregation of linear SV, for finite n, maintains the parametric structure yielding a higher-order linear SV. In contrast, it turns out that the limit aggregate, as $n \to \infty$, will not belong to the SR-SARV class. In fact,

$$\nu_k \sim c\,\gamma^k\, k^{-\delta+1} \quad \text{as } k \to \infty$$

when $\beta = 0$ by Lemma 2. Therefore, due to the hyperbolic factor, the ν_k cannot be obtained as the coefficients in the expansion of the ratio of finite order rational polynomials in the lag operator.
(d) It easily follows that when $\delta > 0$

$$\frac{1}{n}\sum_{i=1}^{n} f_{i,t} \rightarrow_1 E\frac{\pi_i}{1-\alpha_i} + \sum_{k=0}^{\infty} \nu_k e^{\eta_{t-k}} \quad \text{as } n \rightarrow \infty.$$

This implies summability of the ν_k ruling out long memory of the limit aggregate variance. The complication of Proposition 3.2 arises from the fact that it characterizes the limit aggregate of the observables $x_{i,t}$ rather than the limit aggregate of the $f_{i,t}$.

3.3 Nonlinear MA

Assume that the $x_{i,t}$ are described by (11).

Proposition 3.3 *Assume $A(\gamma)$, B.*
(i) When $\gamma < 1$ or $\gamma = 1$, $\beta > 0$ or $\gamma = 1$, $\beta = 0$, $\delta > -1/2$

$$X_{n,t} \rightarrow_2 X_t \quad \textit{as } n \rightarrow \infty$$

with

$$X_t := u_t \left(\sum_{j=1}^{\infty} \nu_j \epsilon_{t-j-1} \right). \tag{25}$$

Under the above conditions

$$\mid X_t \mid < \infty \ \textit{a.s.}$$

and X_t is both strictly and weakly stationary and ergodic.
(ii) Under the above conditions the X_t^2 are covariance stationary.
When $\gamma < 1$ or $\gamma = 1$, $\beta > 0$

$$\text{cov}(X_t^2, X_{t+h}^2) = O(c^h) \quad \textit{as } u \rightarrow \infty,$$

for some $0 < c < 1$.
When $\gamma = 1$, $\beta = 0$

$$\text{cov}(X_t^2, X_{t+h}^2) \sim c\, h^{-4\delta-2} \quad \textit{as } u \rightarrow \infty.$$

Remarks.
(a) When $\gamma = 1$, $\beta = 0$ the acf of the squared limit aggregate decays at an hyperbolic rate and long memory is achieved when $-1/2 < \delta < -1/4$. When $\delta > -1/4$ the acf of the squared limit aggregate will be summable although it still decays hyperbolically.
(b) The limit aggregate (25) is precisely the long memory nonlinear MA introduced by Robinson and Zaffaroni (1998) and Proposition 3.3 represents a sound rationalization for this model.
(c) The assumption of independence between the u_t and the ϵ_t is irrelevant for the result. One can consider the case $\epsilon_t = u_t$ and (25) expresses the (one-shock) long memory nonlinear MA of Robinson and Zaffaroni (1997).

4 Concluding comments

We analyze the outcome of contemporaneous aggregation of heterogeneous SV models, when the cross-sectional dimension gets arbitrarily large. Some general results are obtained, in the sense of necessary conditions for long memory. These conditions are not sufficient though. In fact, we focus on three, well known, SV models which satisfy these necessary conditions. Long memory is ruled out when aggregating exponential SV and permitted when aggregating nonlinear MA. Linear SV represent an intermediate case which allows 'quasi' long memory, in the sense of hyperbolically decaying yet summable acf of the limit squares. The key feature driving the results is the shape of the cross-sectional distribution of the micro coefficients, in turn defined by the specific form of the model nonlinearity.

The $X_{n,t}$ could be interpreted as the return portfolio of n heterogeneous assets with return $x_{i,t}$, each modeled as a SV model. Therefore, this paper obtains the statistical properties of the return portfolio, based on an arbitrary large number of assets. Alternatively, $X_{n,t}$ could represent the return of a single asset, such as a single stock return or a foreign exchange rate return. In this case $X_{n,t}$ can be viewed as the arithmetic average of n heterogeneous components $x_{i,t}$, each characterized by a different degree of persistence of their conditional variance, as suggested in Ding and Granger (1996). Alternatively, many equilibrium models of speculative trading assume that observed asset prices are average of traders' reservation prices with the single asset return implicitly defined as an aggregate (see Tauchen and Pitts (1983)).

Our results still applies when a number m ($m < n$) of units exhibits different properties from the ones assumed here as long as these units are bounded *a.s.* and represent a degenerate fraction of units ($1/m + m/n \to 0$ *a.s.* for $n \to \infty$). Under these conditions, the aggregate properties will be entirely determined by the non-degenerate fraction of units described by the our assumptions.

Several generalizations are possible. Among many, aggregation of continuous-time SV, higher-order SV and models with a time-varying conditional mean. These generalizations could be obtained by a suitable extension of our framework. Several of the conditions for long memory and, more in general, for existence of the limit aggregate of the models here considered, provides a rich set of testable implications on which developing empirical applications.

Appendix

We recall that c denotes an arbitrary positive constant not necessarily the same, the symbol $\sim$ denotes asymptotic equivalence and $P(A)$, 1_A, respectively, the probability and the indicator function of any event A. Finally $\mathrm{E}_n(\cdot)$, $\mathrm{var}_n(\cdot)$ define the expectation and the variance operator conditional on the random coefficients μ_i, α_i ($1 \leq i \leq n$).

Lemma 1 *Let z be a random variable (r.v.) with support $[0,1]$ and density*

$$g(z) \sim c \ \exp(-\frac{\beta}{(1-z)}) \quad \text{as } z \to 1^-,$$

for real $0 < \beta < \infty$. Then

$$E(z^k) \sim c\, k^{-\frac{1}{2}} (1+\beta)^{-k} e^{-k(1+\beta/2)} \quad \text{as } k \to \infty.$$

Proof. All the equivalence below hold for $k \to \infty$. Then

$$E(z^k) \sim c \int_0^1 x^k \exp(-\frac{\beta}{(1-x)})dx = c \int_1^\infty t^{-(k+2)}(t-1)^k e^{-\beta t} dt$$
$$= c\,\Gamma(k+1) e^{-\frac{\beta}{2}} W_{-(k+1),0.5}(\beta(k+1)),$$

by the change of variable $t = 1/(1-x)$ and using Gradshteyn and Ryzhik (1994, # 3.383-4), where $\Gamma(\cdot)$ denote the Gamma function and $W_{\lambda,\mu}(\cdot)$ the

Whittaker function (see Gradshteyn and Ryzhik (1994, section 8.31 p.942 and 9.22-9.23 p.1086). Finally, from Gradshteyn and Ryzhik (1994, # 9.222-2),

$$\begin{aligned}
W_{-(k+1),0.5}(\beta(k+1)) &= \frac{e^{-\frac{\beta(k+1)}{2}}}{\Gamma(k+2)} \int_0^\infty \frac{t^{k+1}}{(\beta(k+1)+t)^{k+1}} e^{-t} dt \\
&= \frac{e^{-\frac{\beta(k+1)}{2}}}{\Gamma(k+2)\,(\beta(k+1))^{k+1}} \int_0^\infty \frac{t^{k+1}}{(1+\frac{t}{\beta(k+1)})^{k+1}} e^{-t} dt \\
&\sim \frac{e^{-\frac{\beta(k+1)}{2}}}{\Gamma(k+2)\,(\beta(k+1))^{k+1}} \int_0^\infty t^{k+1} e^{-t(1+1/\beta)} dt \\
&= \frac{e^{-\frac{\beta(k+1)}{2}}}{\Gamma(k+2)\,(\beta(k+1))^{k+1}} \frac{1}{(1+1/\beta)^{k+2}} \int_0^\infty t^{k+1} e^{-t} dt \\
&= \frac{e^{-\frac{\beta(k+1)}{2}}}{(1+1/\beta)^{k+2}\,(\beta(k+1))^{k+1}},
\end{aligned}$$

using $(1+a/k)^k \sim e^a$. The latter equivalence holds uniformly for any $a \in [0, M]$ for any constant $M < \infty$ but split the integral $\int_0^\infty t^{k+1} e^{-t} dt$ as $\int_0^M t^{k+1} e^{-t} dt + \int_M^\infty t^{k+1} e^{-t} dt$ and note that $\int_M^\infty t^{k+1} e^{-t} dt = O(e^{-cM})$ for some $0 < c < 1$ as $M \to \infty$. Using Stirling's formula (see Brockwell and Davis (1987, p.522)) and combining terms concludes. □

Lemma 2 *Let z be a r.v. with support $[0, \gamma)$ and density*

$$g(z) \sim c\ (\gamma - z)^\delta \quad \text{as } z \to \gamma^-,$$

for real $-1 < \delta < \infty$. Then

$$E\, z^k \sim c\, \gamma^k\, k^{-(\delta+1)} \quad \text{as } k \to \infty.$$

Proof. The result follows by Stirling's formula (see Brockwell and Davis (1987, p.522)) and by the change of variable $t = x/\gamma$, yielding

$$\begin{aligned}
Ez^k &= \int_0^\gamma x^k g(x) \underset{\cdot}{\text{x}} \sim c \int_0^\gamma x^k (\gamma - x)^\delta dx = c\, \gamma^{\delta+k+1} \frac{\Gamma(k+1)\Gamma(\delta+1)}{\Gamma(\delta+k+2)} \\
&\sim c\, \gamma^k\, k^{-(\delta+1)} \quad \text{as } k \to \infty,
\end{aligned}$$

where the constant c is not always the same and $\Gamma(\cdot)$ indicates the Gamma function. □

Lemma 3 *Let z_i be i.i.d. draws from a distribution with support $[0,\gamma)$ and density*

$$g(z) \sim c\ \ L(\frac{1}{\gamma - z})(\gamma - z)^{\delta}\ \exp(-\frac{\beta}{(\gamma - z^2)}) \quad \textit{as } z \to \gamma^-,$$

for real $0 \leq \beta < \infty$, $-1 < \delta < \infty$ and slowly varying $L(\cdot)$. For real $0 < \theta < \infty$ set

$$w_i := \exp(\frac{\theta}{1 - z_i^2}), \ \ d := \frac{\beta}{\theta}.$$

When $d > 0$, for a non-degenerate r.v. $\Gamma > 0$ a.s., as $n \to \infty$

$$\begin{array}{lll} n^{-\frac{1}{d}} \sum_{i=1}^{n} w_i \to_d & \Gamma & \textit{for } 0 < d < 1, \\ n^{-1} \sum_{i=1}^{n} w_i \to_{a.s.} & Ew_1 & \textit{for } d > 1. \end{array}$$

When $d = 0$

$$P(\sum_{i=1}^{n} w_i/n^c < c') \to 0 \quad \textit{as } n \to \infty,$$

for any $0 < c, c' < \infty$.

Proof. We use results of classical extreme value theory. By simple calculations the w_i have density

$$f_w(w_i) = \frac{\theta}{2} \frac{g(\sqrt{1 - \theta/\ln w_i})}{w_i\,(\ln\ w_i)^2 \sqrt{1 - \theta/\ln w_i}},$$

and as $t \to \infty$

$$f_w(t) \sim \left\{ \begin{array}{ll} L_w(t)\, t^{-1}, & \beta = 0, \\ L_w(t) t^{-1-d}, & \beta > 0, \end{array} \right.$$

for slowly varying $L_w(\cdot)$ satisfying $\int_1^\infty L_w(x)\, x^{-1}\, dx < \infty$.

Then, setting $m_n := max(w_1, .., w_n)$, from Embrechts, Klüppelberg, and Mikosch (1997, section 6.2.6 and Theorem A36)

$$\frac{m_n}{(w_1 + ... + w_n)} \to_p 1 \quad \text{as } n \to \infty$$

when $\beta = 0$ and $f_w(\cdot)$ will belong to no maximum domain of attraction, viz. $P(m_n/n^c < c') \to 0$ for any $c, c' > 0$ as $n \to \infty$. When $0 < d < 1$ the distribution of the w_i belongs to the domain of attraction of Γ and Feller (1966, Theorem IX.8.1) applies. Finally, when $d > 1$ the w_i have bounded first moment so by $i.i.d.$-ness the ergodic theorem applies. □

Proof of Proposition 2.1. Given the i.i.d.ness and bounded variance of the $x_{i,t}$ the Lindeberg-Lévy CLT applies, as $n \to \infty$. Moreover, for any n by the martingale difference property

$$\mathrm{cov}_n(\frac{1}{n^{\frac{1}{2}}}\sum_{i=1}^{n} x_{i,t}, \frac{1}{n^{\frac{1}{2}}}\sum_{i=1}^{n} x_{i,t+u}) = 0 \quad \text{for any } u \neq 0. \quad \square$$

Proof of Proposition 2.2. $Ex_t^4 < \infty$ when $Ef_{t-1}^2 < \infty$. Evaluating the expectation of

$$f_t^2 = \omega^2 + \gamma^2 f_{t-1}^2 + v_t^2 + 2\omega\gamma f_{t-1} + 2\omega v_t + 2\gamma f_{t-1} v_t,$$

given J_{t-1}, yields

$$E(f_t^2 \mid J_{t-1}) = \omega^2 + (\gamma^2 + \kappa) f_{t-1}^2 + 2\omega\gamma f_{t-1} + g_{t-1}.$$

By Stout (1974, Theorem 3.5.8) the f_t are strictly stationary yielding $Ef_t^2 = Ef_{t-1}^2$ when they are finite. Then collect terms and use the law of iterated expectations. □

Proof of Proposition 3.1. (i) Any instantaneous transformation of a normally distributed r.v. $g(Z)$, where Z is Gaussian, could be expanded in terms of Hermite polynomials when $E\,g(Z)^2 < \infty$ (see Hannan (1970)). Hence, given

$$\prod_{k=0}^{\infty} E_n\, e^{\alpha_i^k \epsilon_{t-k}} = e^{\frac{\sigma_\epsilon^2}{2(1-\alpha_i^2)}} < \infty \;\; a.s.,$$

expanding the $\exp(\alpha_i^k \epsilon_{t-k}/2)$ yields

$$X_{n,t} = u_t \sum_{\substack{m_j=0 \\ j=0,1,..}}^{\infty} \left(\frac{\sigma_\epsilon}{2}\right)^{\sum_{j=0}^{\infty} m_j} \frac{1}{\prod_{j=0}^{\infty} m_j!}\, \hat{\zeta}_{\sum_{j=0}^{\infty} j m_j} \prod_{j=0}^{\infty} H_{m_j}(\tilde{\epsilon}_{t-j-1}), \qquad (26)$$

setting

$$\hat{\zeta}_k := \frac{1}{n}\sum_{i=1}^{n} \exp(\frac{\sigma_\epsilon^2}{8}\frac{1}{(1-\alpha_i^2)})\alpha_i^k, \;\; k = 0, 1, ...$$

When $\gamma = 1$, $\beta > \sigma_\epsilon^2/4$ the ζ_k are finite and the law of iterated logarithm (LIL) for *i.i.d.* variates (see Stout (1974, Corollary 5.2.1)) applies yielding

$$| \hat{\zeta}_k - \zeta_k |= O\left(\frac{\theta_k^{\frac{1}{2}}}{n^{\frac{1}{2}}}(\ln\ln(n\,\theta_k))^{\frac{1}{2}}\right) \quad a.s. \text{ for } k \to \infty, \tag{27}$$

setting

$$\theta_k := E\ \exp(\frac{\sigma_\epsilon^2}{4}\frac{1}{(1-\alpha_i^2)})\alpha_i^{2k}$$

bounded by assumption for any $k \geq 0$.

By the independence of the ϵ_t and the orthogonality of the Hermite polynomials

$$\text{cov}\left(\prod_{j=0}^{\infty} H_{m_j}(\tilde{\epsilon}_{t-j-1}), \prod_{h=0}^{\infty} H_{n_h}(\tilde{\epsilon}_{t-h-1})\right) = \prod_{j=0}^{\infty}\delta(n_j, m_j)m_j! - \prod_{j=0}^{\infty}\delta(n_j, 0)\delta(m_j, 0),$$

where $\delta(\cdot, \cdot)$ is the Kronecker delta. Thus,

$$\begin{aligned}
E_n(X_{n,t} - X_t)^2 \quad &= \sum_{\substack{m_j, n_j=0 \\ j=0,1,\ldots}}^{\infty} (\frac{\sigma_\epsilon}{2})^{\sum_{j=0}^{\infty}(m_j+n_j)} \frac{1}{\prod_{j=0}^{\infty} m_j! n_j!} \times \\
&(\hat{\zeta}_{\sum_{j=0}^{\infty} jm_j} - \zeta_{\sum_{j=0}^{\infty} jm_j})(\hat{\zeta}_{\sum_{j=0}^{\infty} jn_j} - \zeta_{\sum_{j=0}^{\infty} jn_j}) \times \\
&\text{cov}\left(\prod_{j=0}^{\infty} H_{m_j}(\tilde{\epsilon}_{t-j-1}), \prod_{j=0}^{\infty} H_{n_j}(\tilde{\epsilon}_{t-j-1})\right) \\
&= O\left(\frac{\ln\ \ln\ n}{n} \sum_{\substack{m_j=0 \\ j=0,1,\ldots}}^{\infty} (\frac{\sigma_\epsilon}{2})^{2\sum_{j=0}^{\infty} m_j} \frac{1}{\prod_{j=0}^{\infty} m_j!} c^{\sum_{j=0}^{\infty} jm_j}\right) \\
&= O\left(\frac{\ln\ \ln\ n}{n} \exp(\frac{\sigma_\epsilon^2}{4(1-c)})\right),
\end{aligned}$$

for some $0 < c = c(b, \sigma_\epsilon^2) < 1$ using Lemma 1.

The nonlinear moving average representation (18) of X_t follows replacing $\hat{\zeta}_k$ with ζ_k in (26) and re-arranging terms, given the equivalence

$$\sum_{\substack{m_1,\ldots,m_q=0 \\ m_1+\ldots+m_q=q}}^{\infty} \frac{1}{m_1!\ldots m_q!} = \sum_{\substack{n_0,\ldots,n_q=0 \\ n_0+\ldots+n_q=q \\ 1n_1+\ldots qn_q=q}}^{\infty} \frac{1}{0!^{n_0}\ldots q!^{n_q}} \frac{q!}{n_0!\ldots n_q!}.$$

We show covariance stationarity of the X_t. By Schwarz inequality

$$E_n\, X_{n,t}^2 = \frac{1}{n^2}\sum_{i=1}^{n} E_n\, e^{h_{i,t-1}} + \frac{1}{n^4}\sum_{i\neq j=1}^{n} E_n\, e^{0.5(h_{i,t-1}+h_{j,t-1})} \tag{28}$$

$$\leq \frac{1}{n^2}\sum_{i=1}^{n} E_n\, e^{h_{i,t-1}} + \left(\frac{1}{n}\sum_{i=1}^{n}(E_n\, e^{h_{i,t-1}})^{\frac{1}{2}}\right)^2. \tag{29}$$

By Lemma 3 $E_n(X_{n,t}^2)$ diverges to infinity in probability at rate $n^{2(\frac{\sigma_\epsilon^2}{4\beta}-2)}$ when $\beta < \sigma_\epsilon^2/4$. In fact, the first term on the right hand side of (28) and the second term on the right hand side of (29) have the same asymptotic behaviour, diverging at rate $n^{2(\frac{\sigma_\epsilon^2}{4\beta}-2)}$. This follows considering that given any r.v. Z with distribution tail regularly varying with index $-c\,(c \geq 0)$ (see Embrechts, Klüppelberg, and Mikosch (1997, Appendix A3.1)) then $Z^{\frac{1}{2}}$ has distribution tail regularly varying with index $-2c$. Similarly, $E_n\, X_{n,t}^2$ converges $a.s.$ to $EX_{n,t}^2 < \infty$ when $\beta > \sigma_\epsilon^2/4$. Finally, by Lemma 3, $E_n\, X_{n,t}^2 \to_d \Gamma$ as $n \to \infty$, when $\beta = \sigma_\epsilon^2/4$ for a non-degenerate r.v. Γ. Strict stationarity and ergodicity follows by using Stout (1974, Theorem 3.5.8) and Royden (1980, Proposition 5 and Theorem 3) to (18).

(ii) Let us focus for simplicity's sake on case $\gamma = 1$. Case $\gamma < 1$ easily follows. By Schwarz inequality

$$E_n\, X_{n,t}^4 = \frac{Eu_t^4}{n^4}\sum_{i=1}^{n} E_n\, e^{2\,h_{i,t-1}} + \frac{Eu_t^4}{n^4}\sum_{\substack{i\neq j\neq a\neq b\\=1}}^{n} E_n\, e^{0.5(h_{i,t-1}+h_{j,t-1}+h_{a,t-1}+h_{b,t-1})} \tag{30}$$

$$\leq \frac{Eu_t^4}{n^4}\sum_{i=1}^{n} E_n\, e^{2\,h_{i,t-1}} + Eu_t^4\left(\frac{1}{n}\sum_{i=1}^{n}(E_n\, e^{2\,h_{i,t-1}})^{\frac{1}{4}}\right)^4. \tag{31}$$

By the same arguments used in part (i) one gets that $E_nX_{n,t}^4$ converges to a bounded constant when $\beta > \sigma_\epsilon^2/2$, diverges to infinity in probability at rate $n^{2(\sigma_\epsilon^2/\beta-2)}$ when $\beta < \sigma_\epsilon^2/2$ and converges to a non-degenerate r.v. when $\beta = \sigma_\epsilon^2/2$.

Let us deal with the acf. By the cumulants' theorem (see Leonov and Shiryaev (1959))

$$\mathrm{cov}_n(X_{n,t}^2, X_{n,t+h}^2) = \frac{1}{n^4}\sum_{a,b,c,d=1}^{n} e^{\frac{\sigma_\epsilon^2}{8}\sum_{k=0}^{h-1}(\alpha_c^k+\alpha_d^k)^2}\sum_{k=0}^{\infty} e^{\frac{\sigma_\epsilon^2}{8}\sum_{j=0}^{k-1}(\alpha_a^j+\alpha_b^j+\alpha_c^{j+h}+\alpha_d^{j+h})^2}$$

$$\times\left[e^{\frac{\sigma_\epsilon^2}{8}(\alpha_a^k+\alpha_b^k+\alpha_c^{k+h}+\alpha_d^{k+h})^2} - e^{\frac{\sigma_\epsilon^2}{8}(\alpha_a^k+\alpha_b^k)^2}e^{\frac{\sigma_\epsilon^2}{8}(\alpha_c^{k+h}+\alpha_d^{k+h})^2}\right]$$
$$\times e^{\frac{\sigma_\epsilon^2}{8}\sum_{j=k+1}^\infty(\alpha_a^j+\alpha_b^j)^2}e^{\frac{\sigma_\epsilon^2}{8}\sum_{j=k+1}^\infty(\alpha_c^{j+h}+\alpha_d^{j+h})^2},$$

for any integer $h > 0$, using

$$\text{cov}(\prod_{i=1}^m C_i, \prod_{i=1}^m D_i) = \sum_{k=1}^m \prod_{j=1}^{k-1} E(C_jD_j)\text{cov}(C_k, D_k)E(\prod_{j=k+1}^m C_j)E(\prod_{j=k+1}^m D_j),$$

which holds for any sequence of independent bivariate r.vs $\{C_i, D_i\}$ with $EC_i\,D_i < \infty$. Re-arranging terms yields

$$\text{cov}_n(X_{n,t}^2, X_{n,t+h}^2) = \frac{1}{n^4}\sum_{a,b,c,d=1}^n e^{\frac{\sigma_\epsilon^2}{8}\sum_{j=0}^\infty[(\alpha_a^j+\alpha_b^j)^2+(\alpha_c^j+\alpha_d^j)^2]}$$
$$\times\sum_{k=0}^\infty e^{\frac{\sigma_\epsilon^2}{4}\sum_{j=0}^{k-1}(\alpha_a^j+\alpha_b^j)(\alpha_c^{j+h}+\alpha_d^{j+h})}\left[e^{\frac{\sigma_\epsilon^2}{4}(\alpha_a^k+\alpha_b^k)(\alpha_c^{k+h}+\alpha_d^{k+h})} - 1\right]$$
$$\le \frac{1}{n^4}\sum_{a,b,c,d=1}^n e^{\frac{\sigma_\epsilon^2}{8}\sum_{j=0}^\infty(\alpha_a^j+\alpha_b^j+\alpha_c^j+\alpha_d^j)^2}\sum_{k=0}^\infty\left[e^{\frac{\sigma_\epsilon^2}{4}(\alpha_a^k+\alpha_b^k)(\alpha_c^{k+h}+\alpha_d^{k+h})} - 1\right].$$

Note that $\text{cov}_n(X_{n,t}^2, X_{n,t+h}^2)$ is non-negative. Expanding the exponential term in $[\,\cdot\,]$ brackets

$$\text{cov}_n(X_{n,t}^2, X_{n,t+h}^2) = \sum_{k=0}^\infty\sum_{r=1}^\infty(\frac{\sigma_\epsilon^2}{4})^r\frac{1}{r!}$$
$$\times\left(\frac{1}{n^4}\sum_{a,b,c,d=1}^n e^{\frac{\sigma_\epsilon^2}{8}\sum_{j=0}^\infty(\alpha_a^j+\alpha_b^j+\alpha_c^j+\alpha_d^j)^2}\,(\alpha_a^k+\alpha_b^k)^r\,(\alpha_c^{k+h}+\alpha_d^{k+h})^r\right).$$

Using

$$(a+b+c+d)^2 < 4(a^2+b^2+c^2+d^2),$$

which holds for any real a, b, c, d except for case $a{=}b{=}c{=}d$, when $\beta > \sigma_\epsilon^2/2$, we can find a $0 < c < \infty$ such that

$$E\,\text{cov}_n(X_{n,t}^2, X_{n,t+h}^2)$$
$$= O\left(\int_0^1\int_0^1\int_0^1\int_0^1(\alpha_a^k+\alpha_b^k)^r(\alpha_c^{k+h}+\alpha_d^{k+h})^r e^{-c(\frac{1}{1-\alpha_a}+\frac{1}{1-\alpha_b}+\frac{1}{1-\alpha_c}+\frac{1}{1-\alpha_d})}d\alpha_a d\alpha_b d\alpha_c d\alpha_d\right).$$

Expanding the two binomial terms and using Lemma 1 repeatedly yields, for some $0 < \bar{c} < 1$,

$$E\,\mathrm{cov}_n(X_{n,t}^2, X_{n,t+h}^2) = O(\sum_{r=1}^{\infty} \frac{\sigma_\epsilon^{2r}}{r!} \frac{\bar{c}^{hr}}{1-\bar{c}^{2r}}) = O(e^{\sigma_\epsilon^2 \bar{c}^h} - 1) = O(\bar{c}^h) \quad \text{as } h \to \infty. \ \square$$

Proof of Proposition 3.2. From generalizations of Minkowski's inequality (see Hardy, Littlewood, and Polya (1964, Theorems 24 and 25)), for any sequence $a_{i,j}$ ($i = 1, \ldots,$ $j = 1, \ldots, n$) one obtains:

$$\left(\sum_{i=0}^{\infty} (\frac{1}{n} \sum_{j=1}^{n} (a_{i,j})^{\frac{1}{2}})^2 \right)^{\frac{1}{2}} \le \frac{1}{n} \sum_{j=1}^{n} \left(\sum_{i=0}^{\infty} a_{i,j} \right)^{\frac{1}{2}} \le \left(\sum_{i=0}^{\infty} \frac{1}{n} \sum_{j=1}^{n} (a_{i,j})^{\frac{1}{2}} \right), \qquad (32)$$

yielding (22) with

$$\underline{X}_{n,t} := u_t \left(\left(\frac{1}{n} \sum_{i=1}^{n} \left(\frac{\pi_i}{1-\alpha_i} \right)^{\frac{1}{2}} \right)^2 + \sum_{k=0}^{\infty} \hat{\nu}_{\frac{k}{2}}^2 e^{\eta_{t-k}} \right)^{\frac{1}{2}}$$

$$\overline{X}_{n,t} := u_t \left(\frac{1}{n} \sum_{i=1}^{n} \left(\frac{\pi_i}{1-\alpha_i} \right)^{\frac{1}{2}} + \sum_{k=0}^{\infty} \hat{\nu}_{\frac{k}{2}} e^{\frac{1}{2}\eta_{t-k}} \right),$$

setting

$$\hat{\nu}_k := \frac{1}{n} \sum_{i=1}^{n} \alpha_i^k.$$

(i) Let us initially focus on $\overline{X}_{n,t}$. Let us focus on case $\gamma = 1$, $\beta = 0$ for simplicity's sake. The other cases easily follows. We show that $E_n \mid \overline{X}_{n,t} - \overline{X}_t \mid = o(1)$ as $n \to \infty$. Consider

$$|\overline{X}_{n,t} - \overline{X}_t| \le |u_t| \left(\left| \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\pi_i}{1-\alpha_i} \right)^{\frac{1}{2}} - E \left(\frac{\pi_i}{1-\alpha_i} \right)^{\frac{1}{2}} \right| + \sum_{k=0}^{\infty} |\hat{\nu}_k - \nu_k| \, e^{\frac{1}{2}\eta_{t-k}} \right). \qquad (33)$$

When $\delta > -1/2$ then $E\pi_i^{\frac{1}{2}}/(1-\alpha_i)^{\frac{1}{2}} < \infty$ by Lippi and Zaffaroni (1998, Lemma 1) and using the LIL for *i.i.d.* variates one gets for the first term on the right hand side of (33)

$$\left| \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\pi}{1-\alpha_i} \right)^{\frac{1}{2}} - E \left(\frac{\pi}{1-\alpha_i} \right)^{\frac{1}{2}} \right| = O\left((\frac{\ln \ln n}{n})^{\frac{1}{2}} \right) \quad a.s.$$

For the second term on the right hand side of (33), setting $\rho_k := E\, e^{k\eta_t}$ for real k, consider

$$\begin{aligned}
&\sum_{k=0}^{\infty} |\, \hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}} \,|\; e^{\frac{1}{2}\eta_t} \\
&= \rho_{\frac{1}{2}} \sum_{k=0}^{\infty} |\, \hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}} \,| + \sum_{k=0}^{\infty} |\, \hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}} \,|\; (e^{\frac{1}{2}\eta_t} - \rho_{\frac{1}{2}}). \qquad (34)
\end{aligned}$$

Let us focus on the first term on the right hand side of (34). By the LIL for *i.i.d.* variates, for real k

$$\hat{\nu}_k - \nu_k \;\sim\; 2^{\frac{1}{2}}\, \mathrm{var}^{\frac{1}{2}}(\alpha_i^k) \frac{\ln\, \ln(n\, \mathrm{var}(\alpha_i^k))}{n}^{\frac{1}{2}} \quad a.s. \text{ for } n \to \infty. \qquad (35)$$

When $\beta > 0$ by Lemma 1 $\mathrm{var}^{\frac{1}{2}}(\alpha_i^{\frac{k}{2}}) \leq E^{\frac{1}{2}}(\alpha_i^k) = O(c^k)$ as $k \to \infty$ for some $0 < c < 1$ and thus the result easily follows. When $\beta = 0$ then by Lemma 2 $\mathrm{var}(\alpha_i^k) \sim E(\alpha_i^{2k}) \sim c\, k^{-(\delta+1)}$ as $k \to \infty$ and summability of $\mathrm{var}^{\frac{1}{2}}(\alpha_i^{\frac{k}{2}})$ is not guaranteed anymore. We consider the following truncating argument. For arbitrary $s < \infty$

$$\rho_{\frac{1}{2}} \sum_{k=0}^{\infty} |\, \hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}} \,| = \rho_{\frac{1}{2}} \sum_{k=0}^{s} |\, \hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}} \,| + \rho_{\frac{1}{2}} \sum_{k=s+1}^{\infty} |\, \hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}} \,| \,.$$

For the first term above

$$\sum_{k=0}^{s} |\, \hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}} \,| = O\left((\frac{\ln\, \ln\, n}{n})^{\frac{1}{2}} \right) \quad a.s.\;\; n \to \infty$$

by (35) and for the second term

$$\begin{aligned}
&\sum_{k=s+1}^{\infty} |\, \hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}} \,| \leq \sum_{k=s+1}^{\infty} \hat{\nu}_{\frac{k}{2}} + \sum_{k=s+1}^{\infty} \nu_{\frac{k}{2}} \qquad (36) \\
&= \frac{1}{n} \sum_{i=1}^{n} \frac{\alpha_i^{\frac{(s+1)}{2}}}{1 - \alpha_i^{\frac{1}{2}}} + \sum_{k=s+1}^{\infty} \nu_{\frac{k}{2}} \\
&= O\left(E \frac{\alpha_i^{\frac{(s+1)}{2}}}{1 - \alpha_i} + \sum_{k=s+1}^{\infty} \nu_{\frac{k}{2}} \right) \quad a.s. \quad \text{for } n \to \infty. \qquad (37)
\end{aligned}$$

By Lippi and Zaffaroni (1998, Lemma 1) and Lemma 2 it easily follows that both terms on the right hand side of (37) are well defined when $\delta > 0$. Moreover both are $O(s^{-\delta})$ which can be made arbitrarily small by choosing s large enough.

Let us deal with the second term on the right hand side of (34). Note that $\{e^{\frac{1}{2}\eta_t} - \rho_{\frac{1}{2}}\}$ is an *i.i.d.* sequence with zero mean and finite variance. By the same truncating argument just used, for arbitrary s,

$$\begin{aligned}
&\sum_{k=0}^{\infty} |\hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}}| \, (e^{\frac{\eta_{t-k}}{2}} - \rho_{\frac{1}{2}}) \\
&= \sum_{k=0}^{s} |\hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}}| \, (e^{\frac{\eta_{t-k}}{2}} - \rho_{\frac{1}{2}}) + \sum_{k=s+1}^{\infty} |\hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}}| \, (e^{\frac{1}{2}\eta_{t-k}} - \rho_{\frac{1}{2}}).
\end{aligned}$$

For the first term above

$$\begin{aligned}
&\mathrm{var}_n(\sum_{k=0}^{s} |\hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}}| \, (e^{\frac{1}{2}\eta_{t-k}} - \rho_{\frac{1}{2}})) \\
&= \mathrm{var}(e^{\frac{1}{2}\eta_t}) \sum_{k=0}^{s} \left(\hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}}\right)^2 = O\left(\frac{\ln\ln n}{n}\right) \quad a.s. \quad \text{for } n \to \infty.
\end{aligned}$$

For the second term, using $\mathrm{var}(A-B) \le (\sqrt{\mathrm{var}(A)} + \sqrt{\mathrm{var}(B)})^2$ for any r.vs A, B with finite variance ,

$$\begin{aligned}
&\mathrm{var}_n(\sum_{k=s+1}^{\infty} |\hat{\nu}_{\frac{k}{2}} - \nu_{\frac{k}{2}}| \, (e^{\frac{1}{2}\eta_{t-k}} - \rho_{\frac{1}{2}})) \\
&= O\left(\frac{1}{n^2}\sum_{i,j=1}^{n} \frac{\alpha_i^{\frac{(s+1)}{2}} \alpha_j^{\frac{(s+1)}{2}}}{1-(\alpha_i\,\alpha_j)^{\frac{1}{2}}} + \sum_{k=s+1}^{\infty} \nu_{\frac{k}{2}}^2\right) \\
&= O\left(E\,\frac{\alpha_i^{\frac{(s+1)}{2}} \alpha_j^{\frac{(s+1)}{2}}}{1-(\alpha_i\,\alpha_j)^{\frac{1}{2}}} + \sum_{k=s+1}^{\infty} \nu_{\frac{k}{2}}^2\right) \quad a.s. \quad \text{for } n \to \infty. \qquad (38)
\end{aligned}$$

Both terms on the left hand side of (38) are well defined when $\delta > -1/2$. In fact

$$\frac{1}{n^2}\sum_{i,j=1}^{n} \frac{\alpha_i^k\,\alpha_j^k}{1-(\alpha_i\,\alpha_j)^{\frac{1}{2}}} \le \frac{2}{n^2}\sum_{i,j=1}^{n} \frac{\alpha_i^k\,\alpha_j^k}{1-\alpha_i\,\alpha_j} \le \left(\frac{1}{n}\sum_{i=1}^{n} \frac{\alpha_i^k}{(1-\alpha_i)^{\frac{1}{2}}}\right)^2,$$

since $(1-\alpha_i\alpha_j)^2 \geq (1-\alpha_i^2)(1-\alpha_j^2)$. Boundedness follows by Lippi and Zaffaroni (1998, Lemma 1). Secondly, $\sum_{k=s+1}^{\infty} \nu_{\frac{k}{2}}^2$ is summable for $\delta > -1/2$ by Lemma 2. Both terms are $O(s^{-2\delta})$ and can be made arbitrarily for s large enough.

Combining terms, convergence in first mean of $\overline{X}_{n,t}$ to $\overline{X}_t$ occurs when $\delta > 0$. Boundedness of the $\overline{X}_t$ follows re-writing

$$\overline{X}_t = u_t \left(E\left(\frac{\pi_i}{1-\alpha_i}\right)^{\frac{1}{2}} + \rho_{\frac{1}{2}} \sum_{k=0}^{\infty} \nu_{\frac{k}{2}} + \sum_{k=0}^{\infty} \nu_{\frac{k}{2}} (e^{\frac{1}{2}\eta_{t-k}} - \rho_{\frac{1}{2}}) \right). \tag{39}$$

Then $\sum_{k=0}^{\infty} \nu_{\frac{k}{2}} < \infty$ when $\delta > 0$ and $\mid \sum_{k=0}^{\infty} \nu_{\frac{k}{2}} (e^{\frac{1}{2}\eta_{t-k}} - \rho_{\frac{1}{2}}) \mid < \infty$ $a.s.$ when $\delta > -1/2$ given independence and finite variance of the $\{e^{\frac{1}{2}\eta_t} - \rho_{\frac{1}{2}}\}$ and $\sum_{k=0}^{\infty} \nu_{\frac{k}{2}}^2 < \infty$ by Billingsley (1986, Theorem 22.6). Strict stationarity and ergodicity follows by Stout (1974, Theorem 3.5.8). Finally it easily follows that covariance stationarity for both levels and squares requires $\sum_{k=0}^{\infty} \nu_{\frac{k}{2}} < \infty$ and thus $\delta > 0$.

The same results apply for $\underline{X}_{n,t}$ when $\sum_{k=1}^{\infty} \nu_k^2 < \infty$ which requires $\delta > -1/2$.

(ii) When $\gamma < 1$ or $\gamma = 1$, $\beta > 0$, then $\nu_k = O(c^k)$ as $k \to \infty$ for some $0 < c < 1$ and the result easily follows. When $\gamma = 1$, $\beta = 0$ easy calculations show that $E\overline{X}_t^4 < \infty$ requires $\sum_{k=0}^{\infty} \nu_k < \infty$ and thus $\delta > 0$. Moreover

$$\begin{aligned} &\operatorname{cov}(\overline{X}_t^2, \overline{X}_{t+h}^2) \\ &\sim c \sum_{k=0} \nu_{\frac{k}{2}} \nu_{\frac{(k+h)}{2}} + c' (\sum_{k=0} \nu_{\frac{k}{2}} \nu_{\frac{(k+h)}{2}})^2 + c'' \sum_{k=0} \nu_{\frac{k}{2}}^2 \nu_{\frac{(k+h)}{2}}^2 \text{ as } h \to \infty, \end{aligned}$$

and by

$$\sum_{k=0}^{\infty} \nu_{\frac{k}{2}}^2 \nu_{\frac{(k+h)}{2}}^2 = O(\max_{k > h/2} \nu_{\frac{k}{2}}^2),$$

the first terms on the right hand side above dominates yielding

$$\operatorname{cov}(\overline{X}_t^2, \overline{X}_{t+h}^2) \sim c \sum_{k=0}^{\infty} \nu_{\frac{k}{2}} \nu_{\frac{(k+h)}{2}} \sim c' \nu_{\frac{h}{2}} \sim c''\, h^{-(\delta+1)} \quad \text{as } h \to \infty$$

by (20) and summability of the ν_k. Along the same lines, it follows that covariance stationarity of the $\underline{X}_t^2$ requires $\sum_{k=1}^{\infty} \nu_{\frac{k}{2}}^2 < \infty$ and thus $\delta > -1/2$, yielding

$$\text{cov}(\overline{X}_t^2, \overline{X}_{t+h}^2) \sim c \sum_{k=1}^{\infty} \nu_{\frac{k}{2}}^2 \nu_{\frac{(k+h)}{2}}^2 \sim c' \nu_{\frac{h}{2}}^2 \sim c'' \, h^{-2(\delta+1)} \quad \text{as } h \to \infty. \square$$

Proof of Proposition 3.3. Adapting the arguments developed in part (i) of the proof of Proposition 3.2 above to

$$E_n (X_{n,t} - X_t)^2 = \sigma_\epsilon^2 \sum_{k=1}^{\infty} (\hat{\nu}_k - \nu_k)^2 ,$$

part (i) follows. The limit aggregate coincides with the long memory nonlinear MA of Robinson and Zaffaroni (1998) who establish the memory properties of part (ii), given the asymptotic behaviour of the ν_j characterized in Lemma 1 and 2. □

References

Andersen, T. (1994): "Stochastic autoregressive volatility: a framework for volatility modelling," *Mathematical Finance*, 4/2, 75–102.

Andersen, T., and T. Bollerslev (1997): "Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns," *Journal of Finance*, 52, 975–1005.

Baillie, R. T., T. Bollerslev, and H. O. A. Mikkelsen (1996): "Fractionally integrated generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 74/1, 3–30.

Billingsley, P. (1986): *Probability and Measure*. New York: Wiley, second edn.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31, 302–327.

Breidt, F., N. Crato, and de Lima P. (1998): "Modelling long-memory stochastic volatility," *Journal of Econometrics*, 73, 325–334.

Brockwell, P., and R. Davis (1987): *Time series: theory and methods*. New York: Springer-Verlag.

Comte, F., and E. Renault (1998): "Long memory in continuous-time stochastic volatility models," *Mathematical Finance*, 8/4, 291–323.

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

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

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997): *Modelling extremal events in insurance and finance*. Berlin: Springer.

Esary, J., F. Proschan, and D. Walkup (1967): "Association of random variables with applications," *Annals of Mathematical Statistics*, 38, 1466–1474.

Feller, W. (1966): *An introduction to probability theory and its applications*, vol. II. New York: Wiley.

Ghysels, E., A. Harvey, and E. Renault (1996): "Stochastic Volatility," in *Statistical methods of finance. Handbook of Statistics*, ed. by G. Maddala, and C. Rao, vol. 14, pp. 119–191. Amsterdam; New York and Oxford: Elsevier, North Holland.

Giraitis, L., and D. Surgailis (1986): "Multivariate Appell polynomials and the central limit theorem," in *Dependence in probability and Statistics*, ed. by M. Taqqu, and A. Eberlein. Birkhauser.

Gradshteyn, I., and I. Ryzhik (1994): *Table of integrals, series and products*. San Diego: Academic Press, 5 edn.

Granger, C. (1980): "Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics*, 14, 227–238.

Granger, C. W. J., and Z. Ding (1996): "Varieties of long memory models," *Journal of Econometrics*, 73, 61–77.

Hannan, E. (1970): *Multiple time series*. New York: Wiley.

Hardy, G., J. Littlewood, and G. Polya (1964): *Inequalities*. Cambridge (UK): Cambridge University Press.

Harvey, A. (1998): "Long memory in stochastic volatility," in *Forecasting volatility in financial markets*, ed. by J. Knight,and S. Satchell. London: Butterworth-Heineman, forthcoming.

Leipus, R., and M. Viano (1999): "Aggregation in ARCH models," *Preprint*.

Leonov, V., and A. Shiryaev (1959): "On a method of calculation of semi-invariants," *Theory of Probability Applications*, 4, 319–329.

Lippi, M., and P. Zaffaroni (1998): "Aggregation of simple linear dynamics: exact asymptotic results," Econometrics Discussion Paper 350, STICERD-LSE.

Meddahi, N., and E. Renault (1996): "Aggregation and marginalization of GARCH and Stochastic Volatility models," *Preprint*.

Nelson, D. (1991): "Conditional heteroscedasticity in asset pricing: a new approach," *Econometrica*, 59, 347–370.

Nijman, T., and E. Sentana (1996): "Marginalization and contemporaneous aggregation in multivariate GARCH processes," *Journal of Econometrics*, 71, 71–87.

Robinson, P. M. (1978): "Statistical inference for a random coefficient autoregressive model," *Scandinavian Journal of Statistics*, 5, 163–168.

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. (1999): "The memory of stochastic volatility models," *Preprint*.

Robinson, P. M., and P. Zaffaroni (1997): "Modelling nonlinearity and long memory in time series," in *Nonlinear Dynamics and Time Series*, ed. by C. Cutler,and D. Kaplan. Providence (NJ): American Mathematical Society.

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

Royden, H. (1980): *Real analysis.* London: Macmillan.

Stout, W. (1974): *Almost sure convergence.* Academic Press.

Tauchen, G., and M. Pitts (1983): "The price variability volume relationship on speculative markets," *Econometrica*, 51, 485–505.

Taylor, S. (1986): *Modelling Financial Time Series.* Chichester (UK): Wiley.

Zaffaroni, P. (2000): "Contemporaneous aggregation of GARCH processes," Econometrics Discussion Paper 383, STICERD-LSE.