Probabilistic Properties of Stochastic Volatility Models

Richard A. Davis[1] and Thomas Mikosch[2]

[1] Department of Statistics, Colorado State University, Fort Collins, CO 80523, U.S.A., www.stat.colostate.edu/∼rdavis, `rdavis@stat.colostate.edu`
[2] Laboratory of Actuarial Mathematics, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark, www.math.ku.dk/∼mikosch, `mikosch@math.ku.dk`

We collect some of the probabilistic properties of a strictly stationary stochastic volatility process. These include properties about mixing, covariances and correlations, moments, and tail behavior. We also study properties of the autocovariance and autocorrelation functions of stochastic volatility processes and its powers as well as the asymptotic theory of the corresponding sample versions of these functions. In comparison with the GARCH model (see Lindner [26]) the stochastic volatility model has a much simpler probabilistic structure which contributes to its popularity.

1 The model

We consider a *stochastic volatility process* $(X_t)_{t\in\mathbb{Z}}$ given by the equations

$$X_t = \sigma_t \, Z_t \,, \quad t \in \mathbb{Z} \,, \tag{1}$$

where $(\sigma_t)_{t\in\mathbb{Z}}$ is a strictly stationary sequence of positive random variables which is independent of the iid *noise* sequence $(Z_t)_{t\in\mathbb{Z}}$.[3] We refer to $(\sigma_t)_{t\in\mathbb{Z}}$ as the *volatility* sequence. Following the tradition in time series analysis, we index the stationary sequences (X_t), (Z_t), (σ_t) by the set $\mathbb{Z}$ of the integers. For practical purposes, one would consider e.g., the sequence $(X_t)_{t\in\mathbb{N}}$ corresponding to observations at the times $t = 1, 2, \ldots$.

[3] It is common to assume the additional standardization conditions $EZ_t = 0$ and $\mathrm{var}(Z_t) = 1$. These conditions are important for example in order to avoid identification problems for the parameters of the model. In most parts of this article, these additional conditions are not needed. Moreover, in Sections 4 and 5 we will also consider results when $\mathrm{var}(Z) = \infty$ or $E|Z| = \infty$.

2 Stationarity, ergodicity and strong mixing

2.1 Strict stationarity

The independence of the noise (Z_t) and the volatility sequence (σ_t) allow for a much simpler probabilistic structure than that of a GARCH process which includes explicit feedback of the current volatility with previous volatilities and observations. This is one of the attractive features of stochastic volatility models. For example, the mutual independence of the sequences (Z_t) and (σ_t) and their strict stationarity immediately imply that (X_t) is strictly stationary.[4] Conditions for the existence of a stationary GARCH process are much more difficult to establish, see Nelson [33] and Bougerol and Picard [5], cf. Lindner [26].

A convenient way of constructing a positive stationary volatility sequence (σ_t) is to assume that $Y_t = \log \sigma_t$, $t \in \mathbb{Z}$, is a stationary sequence (Y_t) with certain nice properties. An obvious candidate is the class of causal linear processes given by

$$Y_t = \sum_{i=0}^{\infty} \psi_i \, \eta_{t-i} \,, \quad t \in \mathbb{Z} \,, \tag{2}$$

where (ψ_t) is a sequence of deterministic coefficients with $\psi_0 = 1$ and (η_t) is an iid sequence of random variables. It immediately follows by an application of Kolmogorov's 3-series theorem (cf. Billingsley [3], Petrov [34]) that, if $E\eta = 0$ and $\mathrm{var}(\eta) < \infty$, the infinite series (2) converges a.s. if and only if $\sum_j \psi_j^2 < \infty$. The class of processes (2) includes short memory ARMA as well as long memory fractional ARIMA processes. We refer to Brockwell and Davis [8] for an extensive treatment of such processes, see also the discussion in Section 3 below. Moreover, if we further specify η to be centered normal and if we assume that $\sum_j \psi_j^2 < \infty$ then the sequence (σ_t) is stationary with log-normal marginals.

In what follows, we always assume that (σ_t), hence (X_t), is a (strictly) *stationary sequence.*

2.2 Ergodicity and strong mixing

If (σ_t) is stationary ergodic, then the sequence (X_t) inherits ergodicity as well. This applies e.g., if $\sigma_t = f((\eta_{t+h})_{h\in\mathbb{Z}})$ for a measurable function f and an iid sequence (η_t), and in particular for the model (2). These facts follow from standard ergodic theory, cf. Krengel [24].

The process (X_t) also inherits more subtle properties of the volatility sequence such as certain mixing properties. We illustrate this in the case of

[4] We refer to stationarity as strict stationarity and we write A for a generic element of any stationary sequence (A_t).

strong mixing. Recall that a stationary sequence $(A_t)_{t\in\mathbb{Z}}$ is *strongly mixing* if it satisfies the property

$$\sup_{B\in\mathcal{F}_{-\infty}^0\,,\,C\in\mathcal{F}_t^\infty} |P(B\cap C)-P(B)\,P(C)| =: \alpha_t \to 0\,,\quad t\to\infty\,,$$

where $\mathcal{F}_a^b$ is the σ-field generated by A_t, $a\le t\le b$, with the obvious modifications for $a=-\infty$ and $b=\infty$. The function (α_t) is called the *mixing rate function* of (A_t). Its decay rate to zero as $t\to\infty$ is a measure of the range of dependence or of the memory in the sequence (A_t). If (α_t) decays to zero at an exponential rate, then (A_t) is said to be strongly mixing with *geometric rate*. In this case, the memory between past and future dies out exponentially fast. A recent survey on strong mixing and its interrelationship with other mixing conditions can be found in Bradley [6], see also the collection of surveys on dependence [17, 16] and the overviews on mixing properties of time series models in Fan and Yao [20].

The rate function (α_t) is closely related to the decay of the *autocovariance* and *autocorrelation functions* (ACVF and ACF) of the stationary process (A_t) given by

$$\gamma_A(h)=\operatorname{cov}(A_0,A_h)\quad\text{and}\quad \rho_A(h)=\operatorname{corr}(A_0,A_h)\,,\quad h\ge 0\,,\tag{3}$$

where we assume that $\gamma_A(0)=\operatorname{var}(A)<\infty$. For example, if $E(|A|^{2+\delta})<\infty$ for some $\delta>0$, then

$$|\rho_A(h)|\le c\,\alpha_h^{\delta/(2+\delta)}\quad h\ge 0\,,\tag{4}$$

for some constant $c>0$, see Ibragimov and Linnik [23], Theorem 17.2.2. In some special cases one can also conclude from the decay rate to zero of the ACVF about the convergence rate to zero of (α_t). For example, if (A_t) is a Gaussian causal ARMA process it is well known (cf. Brockwell and Davis [8]) that $\rho_A(h)\to 0$ exponentially fast which in turn implies that $\alpha_t\to 0$ exponentially fast; see Pham and Tran [35].[5]

Since strong mixing is defined via σ-fields, strong mixing of the log-volatility sequence (Y_t) immediately transfers to sequences of measurable functions of the Y_t's. For example, if (Y_t) is strongly mixing with rate function (α_t), so are (σ_t) and (σ_t^2) with the same rate function. Moreover, the stochastic volatility process (X_t) inherits strong mixing from (σ_t) essentially with the same rate. This can be established using the following simple calculation. Since (σ_t) and (Z_t) are independent, we have for any Borel sets $B\in\mathcal{F}_{-\infty}^0$ and $C\in\mathcal{F}_t^\infty$ that

$$\begin{aligned}&|P(B\cap C)-P(B)\,P(C)| \qquad (5)\\ &=|E[f(\ldots,\sigma_{-1},\sigma_0)g(\sigma_k,\sigma_{k+1},\ldots)]-E[f(\ldots,\sigma_{-1},\sigma_0)]\,E[g(\sigma_k,\sigma_{k+1},\ldots)]|\,,\end{aligned}$$

[5] Pham and Tran [35] proved the result for β-mixing which implies strong mixing. They also prove the result for general classes of linear processes with exponentially decaying coefficients ψ_i and iid noise (η_i) more general than Gaussian noise.

where

$$f(\ldots,\sigma_{-1},\sigma_0) = P\left((\ldots,X_{-1},X_0)\in A \mid \sigma_s,\, s\le 0\right)\,,$$
$$g(\sigma_t,\sigma_{t+1},\ldots) = P\left((X_t,X_{t+1},\ldots)\in B \mid \sigma_s,\, s\ge t\right)\,,$$

and standard results about strong mixing (cf. Doukhan [15]) show that the right-hand side in (5) is bounded by $4\alpha_t$. Finally, we conclude that all sequences (X_t), (σ_t), (σ_t^2), (Y_t) essentially have the same strong mixing properties.

Moreover, a sequence generated by measurable transformations of the form $f(\sigma_t,\ldots,\sigma_{t+h})$ or $g(X_t,\ldots,X_{t+h})$ for any $h\ge 0$ and measurable functions f, g is stationary and inherits the strong mixing property with the same rate as (Y_t). This immediately follows from the definitions of stationarity and strong mixing. In particular, the sequences of powers (σ_t^p) and $(|X_t|^p)$ for any positive p have mixing rates comparable to those of (σ_t) and (X_t), respectively.

3 The covariance structure

A first check of the dependence structure in a stationary sequence (A_t) usually focuses on the ACVF γ_A or the ACF ρ_A, see (3). Since a stochastic volatility process (X_t) is a highly non-Gaussian process its covariance structure is not very informative. In fact, as shown below, (X_t) is uncorrelated yet dependent. Insight about the nature of the dependence in stochastic volatility processes can be obtained by studying the correlation function of powers of the process and volatility process given by $(|X_t|^p)$ and (σ_t^p) for some $p>0$, respectively.

In what follows, we focus on volatility sequences (σ_t) of the form (2) with iid random variables η_i since it has the attractive property that we get explicit representations for γ_X. Assuming $EZ=0$, $\operatorname{var}(Z)=1$, and $E\exp\{2|Y|\}<\infty$, direct calculation yields

$$\gamma_X(h)=\rho_X(h)=0\,,\quad h>0\,,\quad \operatorname{var}(X)=\operatorname{var}(\sigma)\,, \tag{6}$$

$$\operatorname{var}(\sigma)=\prod_{j=0}^{\infty} m_\eta(2\,\psi_j)-\prod_{j=0}^{\infty} m_\eta^2(\psi_j)\,, \tag{7}$$

where $m_\eta(z)=E\mathrm{e}^{\,z\eta}$ denotes the moment generating function of η. If η is centered normal with variance $\tau^2>0$, then we have $m_\eta(z)=\exp\{0.5\,\tau^2\,z^2\}$. Hence

$$\operatorname{var}(\sigma)=\exp\{\operatorname{var}(Y)\}(\exp\{\operatorname{var}(Y)\}-1)\,. \tag{8}$$

We observe that (X_t) is a white noise sequence. This fact is very much in agreement with real-life return data. However, this observation is not very informative. Therefore it has become common in financial time series analysis to study the ACVFs/ACFs of the absolute values, squares and other powers of absolute return data as well. The present state of research on GARCH processes does not allow one to get explicit formulae for the ACVF of the absolute returns. In a stochastic volatility model one can exploit the independence between (σ_t) and (Z_t) in order to get explicit formulae for $\gamma_{|X|}$ at least in the model (2) with iid noise (η_i).

In what follows, we focus on this model with iid centered normal noise with variance $\tau^2 > 0$ and calculate the corresponding ACVFs and ACFs. Recall that the ACVF of (Y_t) is then given by

$$\gamma_Y(h) = \tau^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+h}\,, \quad h \geq 0\,. \tag{9}$$

Calculations similar to those leading to equations (6) and (8) yield for any $p > 0$ and $h > 0$,

$$\begin{aligned}
\gamma_{|X|^p}(h) &= (E(|Z|^p))^2\, \gamma_{\sigma^p}(h)\,, \\
\gamma_{\sigma^p}(h) &= E\mathrm{e}^{\,p\,(Y_0+Y_h)} - (E\mathrm{e}^{\,p\,Y})^2 \\
&= \exp\Big\{(p\tau)^2 \sum_{i=0}^{\infty} \psi_i^2\Big\}\Big[\exp\Big\{(p\tau)^2 \sum_{i=0}^{\infty} \psi_i\,\psi_{i+h}\Big\} - 1\Big] \\
&= \mathrm{e}^{\,p^2\,\gamma_Y(0)}\Big[\mathrm{e}^{\,p^2\,\gamma_Y(h)} - 1\Big]\,,
\end{aligned}$$

where we have assumed that $E(|Z|^p) < \infty$. Since $\gamma_Y(h) \to 0$ as $h \to \infty$ a Taylor series argument yields[6]

$$\gamma_{\sigma^p}(h) \sim p^2\,\mathrm{e}^{\,p^2\,\gamma_Y(0)}\,\gamma_Y(h)\,, \quad h \to \infty\,. \tag{10}$$

Similarly,

$$\begin{aligned}
\mathrm{var}(\sigma^p) &= \mathrm{e}^{\,p^2\,\gamma_Y(0)}\big(\mathrm{e}^{\,p^2\,\gamma_Y(0)} - 1\big)\,, \\
\mathrm{var}(|X|^p) &= E(|Z|^{2p})\,\mathrm{e}^{\,p^2\,\gamma_Y(0)}\big(\mathrm{e}^{\,p^2\,\gamma_Y(0)} - 1\big) + \mathrm{var}(|Z|^p)\,\mathrm{e}^{\,p^2\,\gamma_Y(0)}\,,
\end{aligned}$$

where we have assumed that $E(|Z|^{2p}) < \infty$. Finally, we can calculate the ACFs of $(|X_t|^p)$ and (σ_t^p) for $h > 0$:

$$\rho_{\sigma^p}(h) = \frac{\mathrm{e}^{\,p^2\,\gamma_Y(h)} - 1}{\mathrm{e}^{\,p^2\,\gamma_Y(0)} - 1} \sim p^2\,(\mathrm{e}^{\,p^2\,\gamma_Y(0)} - 1)^{-1}\,\gamma_Y(h)\,, \tag{11}$$

$$\rho_{|X|^p}(h) = \frac{E(|Z|^p)^2}{E(|Z|^{2p}) + \mathrm{var}(|Z|^p)\,(\mathrm{e}^{\,p^2\,\gamma_Y(0)} - 1)^{-1}}\,\gamma_{\sigma^p}(h)\,. \tag{12}$$

---

[6] We write $f(h) \sim g(h)$ as $h \to \infty$ for two functions f and g whenever $f(h)/g(h) \to 1$ as $h \to \infty$.

The ACVF (9) of the linear causal Gaussian process (Y_t) in (2) may decay to zero arbitrarily fast. In particular, if (Y_t) is a causal ARMA process, the ACVF decays to zero at an exponential rate. On the other hand, if (Y_t) is a FARIMA(p,d,q) process with $d \in (0,1)$, $\gamma_Y(h) \sim \text{const}\, h^{d-1}$. In particular, the sequence (Y_t) exhibits *long-range dependence* or *long memory* in the sense that the ACVF is not absolutely summable. Otherwise, as in the ARMA case, the sequence (Y_t) is referred to as a process with *short-range dependence* or *short memory.* We refer to Brockwell and Davis [8] and Samorodnitsky and Taqqu [37] for extensive discussions on long memory processes, in particular on FARIMA processes and their properties. See also the more recent treatment of long memory phenomena in Doukhan et al. [16].

We conclude from the discussion above and from formulae (10)–(12) that γ_{σ^p} inherits the asymptotic behavior of the ACVF γ_Y and, in turn, $\gamma_{|X|^p}$ inherits the asymptotic behavior of γ_{σ^p}. Since γ_Y may decay to zero at any rate we conclude that the ACVF of the processes $(|X_t|^p)$ may decay to zero at any rate as well. This allows one to model the ACVF behavior of an absolute return series in a flexible way, in contrast to the GARCH case. Indeed, under general conditions on the noise, a GARCH process (A_t) is ϕ-mixing with a geometric rate, see Mokkadem [32], Boussama [9], cf. Doukhan [15]. In particular, it is strongly mixing with a rate function (α_t) which decays to zero exponentially fast. Then an appeal to (4) yields for any measurable function f on $\mathbb{R}$ that $\gamma_{f(A)}(h) \to 0$ at an exponential rate as $h \to \infty$, given that $\gamma_{f(A)}$ is well defined.

4 Moments and tails

In this section we consider some of the marginal distributional characteristics of a stochastic volatility process. It is straightforward that for any $p > 0$,

$$E(|X|^p) = E(|Z|^p)\, E\sigma^p\,,$$

and this pth moment of X is finite if and only if the pth moments of Z and σ are finite. This naturally leads to some restrictions on the moments of the noise (η_i) in model (2): the tails of η must not be too heavy, otherwise $E\sigma^p = \infty$ for all $p > 0$. This excludes in particular subexponential distribution for η for which $m_\eta(p) = \infty$ for all $p > 0$. The subexponential class includes distributions with a power law tail (such as the student and Pareto distributions) as well as moderately heavy-tailed distributions such as the Weibull distribution $P(\eta > x) = \exp\{-c\, x^\tau\}$, $x > 0$, for some $\tau \in (0,1)$, $c > 0$, and the log-normal distributions. We refer to Embrechts at el. [19] for an extensive treatment of subexponential distributions.

In various cases the analysis of the moments of a stochastic volatility model can be refined by a study of the asymptotic tail behavior of the distribution of X. The close relation between the moments and the tails can be seen e.g., from the fact that for any non-negative random variable A,

$$EA = \int_0^\infty P(A > x)\, dx\,. \tag{13}$$

Our particular interest focuses on non-negative random variables with power law tails of the form

$$P(A > x) = x^{-\alpha}\, L(x)\,, \quad x > 0\,, \tag{14}$$

where $\alpha > 0$ and L is a *slowly varying function* which is defined by the asymptotic relation $L(cx)/L(x) \to 1$ as $x \to \infty$, for all $c > 0$. The class of slowly varying functions includes constants, logarithms, iterated logarithms, powers of logarithms. Since for every $\delta > 0$ there exist positive constants x_0 and c_1, c_2 such that

$$c_1\, x^{-\delta} \le L(x) \le c_2\, x^{\delta}\,, \quad x \ge x_0\,, \tag{15}$$

the contribution of L to the tail in (14) is negligible compared to the power law $x^{-\alpha}$. The function on the right-hand side of (14) is said to be *regularly varying with index* $-\alpha$, and we will also say that the *distribution* of A is *regularly varying with index* α. It is an easy exercise to combine relations (15) and (13) to conclude that

$$E(A^{\alpha+\delta}) \begin{cases} < \infty\,, & \delta > 0\,, \\ = \infty\,, & \delta < 0\,, \end{cases}$$

whereas $E(A^{\alpha})$ may be finite or infinite, depending on the slowly varying function L. For an extensive treatment of slowly varying and regularly varying functions and distributions and their properties, we refer the reader to the classical encyclopedia by Bingham et al. [4] and to Resnick [36].

From the definition it is evident that regularly varying distributions have heavy tails especially for small α. Therefore they are capable of capturing the probabilities of rare erratic events such as crashes, eruptions, bursts, and other phenomena which cannot be adequately described by commonly used distributions such as the normal, exponential and gamma. Examples of regularly varying distributions include the Pareto and Burr distributions which are standard models for large claims in (re)insurance applications (see Embrechts et al. [19], Chapter 2), the ON-OFF distributions of Internet teletraffic models (see Leland et al. [25], Willinger et al. [38], Mikosch et al. [29]), the one-dimensional marginals of GARCH and infinite variance stable processes (see Goldie [22], Basrak et al. [19] for the tails of GARCH processes, Feller [21] and Samorodnitsky and Taqqu [37] for the tails of stable processes). There exists empirical evidence that the distribution of log-returns is well approximated in its left and right tails by regularly varying distributions (possibly with different tail indices on the left and on the right) such as the generalized Pareto distribution with positive shape parameter (see Embrechts et al. [19], Chapter 6, Mikosch [28]).

The observation that log-return data have power law tails goes back at least to the 1960s. For example, Mandelbrot [27] suggested that infinite variance stable distributions might be appropriate models. The latter class of distributions is regularly varying with index $\alpha < 2$, see Feller [21] and Samorodnitsky and Taqqu [37]. Since Mandelbrot's contributions were not based on rigorous statistical analyses there has been an ongoing discussion about the value α and whether power law tails make sense for financial data at all.[7] A detailed statistical analysis of the tail index α of return data depends on conditions such as strict stationarity which is unlikely to be satisfied for large samples, whereas the estimation of α requires large samples (sizes of 1000 observations and more are desirable, see Embrechts et al. [19], Chapters 6 and 7). Since changes in the distribution of returns are likely in large samples it is a rather difficult task to decide a value of α that is appropriate for the entire segment of a long series. Nevertheless, there is a strong belief by many researchers that return data have power law tails. Using extreme value statistics for estimating the tail index α, one often finds α to be well below 5 or 6 yet greater than 2.

Since returns may have heavy-tailed distributions it is natural to ask for conditions on the stochastic volatility process which ensure the existence of power law tails in the marginal distribution. Recall that $X_t \stackrel{d}{=} X = \sigma Z$, where $\sigma_t \stackrel{d}{=} \sigma$ and $Z_t \stackrel{d}{=} Z$ are independent random variables. In this context it is useful to recall an elementary result of Breiman [7]. Let A, B be non-negative random variables such that A is regularly varying with index α and $E(B^{\alpha+\delta}) < \infty$ for some $\delta > 0$. Then

$$P(A\,B > x) \sim E(B^{\alpha})\,P(A > x)\,, \quad x \to \infty\,.$$

In particular, AB is regularly varying with index α. Thus the product inherits the heavier tail of the two factors.[8]

An immediate consequence is that

$$P(X > x) = P(Z_+\,\sigma > x) \sim E(\sigma^{\alpha})\,P(Z_+ > x)\,, \quad x \to \infty\,, \tag{16}$$

provided $Z_+ = \max(0, Z)$ has a regularly varying distribution with index $\alpha > 0$ and $E(\sigma^{\alpha}) < \infty$. The latter condition is satisfied in model (2) with iid normal noise. Then σ is lognormal, hence all moments $E(\sigma^p)$, $p > 0$, are finite. Analogously,

$$P(X \le -x) = P(Z_-\,\sigma \ge x) \sim E(\sigma^{\alpha})\,P(Z_- > x)\,, \quad x \to \infty\,,$$

provided $Z_- = \max(0, -Z)$ has a regularly varying distribution with index α and $E(\sigma^{\alpha}) < \infty$. The case of σ and Z with heavy tails of the same order

[7] The use of stable distributions for financial modeling is discussed in Mittnik [31].

[8] Of course, $E(B^{\alpha+\delta}) < \infty$ for some $\delta > 0$ and regular variation of A with index α imply that $P(B > x) = o(P(A > x))$ as $x \to \infty$.

of magnitude is rather involved. As a matter of fact, if both σ and Z_+ are regularly varying with index $\alpha > 0$, then X_+ is regularly varying with the same index but the form of the slowly varying function L in the tail is in general not known, see Embrechts and Goldie [18].

Breiman's result (16) tells us that a power law tail for X may result from a heavy tail of the volatility σ or of the noise Z. Since we observe neither σ nor Z we can only judge about their distributional tail on the basis of a model such as the stochastic volatility model or the GARCH process. In the GARCH case power law tails of $P(X_t > x)$ are more the rule than the exception:[9] even for light-tailed Z (such as Gaussian noise) the volatility σ will typically have power law tails.

The tail behavior of the marginal distribution of a stationary sequence (X_t) is essential for its extremal behavior. In particular, power law behavior for the tail $P(X > x)$ often results in the fact that scaled maxima $\max_{t=1,\ldots,n} X_t$ converge in distribution to a Fréchet distributed random variable. We study the convergence of the extremes of a stochastic volatility process in Davis and Mikosch [14]. There we consider the case of regularly varying X, but also some light-tailed X and the corresponding extreme value theory.

5 Asymptotic theory for the sample ACVF and ACF

In this section we briefly study the asymptotic behavior of the sample mean, and the sample ACVFs of the stochastic volatility process (X_t), its absolute values $(|X_t|)$ and its squares. Recall that the *sample ACVF* and the *sample ACF* of a stationary sequence (A_t) are given by

$$\widehat{\gamma}_A(h) = \frac{1}{n}\sum_{t=1}^{n-h}(A_t - \overline{A}_n)\,(A_{t+h} - \overline{A}_n)\,, \quad \widehat{\rho}_A(h) = \frac{\widehat{\gamma}_A(h)}{\gamma_A(0)}\,, \quad 0 \le h < n\,,$$

respectively, where $\overline{A}_n = n^{-1}\sum_{t=1}^{n} A_t$ denotes the mean of the sample $A_1, \ldots, A_n$. If (σ_t), hence (X_t), is stationary ergodic, then the ergodic theorem (cf. Krengel [24]) implies that the sample ACVFs at a fixed lag h, $\widehat{\gamma}_\sigma(h)$, $\widehat{\gamma}_X(h)$, $\widehat{\gamma}_{|X|^i}(h)$, $i = 1, 2$, converge a.s. to their corresponding deterministic counterparts $\gamma_\sigma(h)$, $\gamma_X(h)$, $\gamma_{|X|^i}(h)$, $i = 1, 2$, provided that the limiting covariances exist and are finite. The corresponding sample ACFs at a fixed lag h will then converge a.s. to their deterministic counterparts as well.

Central limit theory for functionals of a stochastic volatility process (X_t) is often easily established. In what follows, we give some examples which are not exhaustive but illustrative of the techniques that are involved. Assume that the log-volatility process (Y_t) is given by the representation (2) for an iid sequence (η_i) and that $\mathrm{var}(\sigma) < \infty$, $EZ = 0$ and $\mathrm{var}(Z) = 1$. We introduce

[9] See the article about the extremes of a GARCH process in Davis and Mikosch [13].

the filtration $\mathcal{G}_t = \sigma(Z_s\,, \eta_s\,, s \le t)$. Then (X_t) is adapted to $(\mathcal{G}_t)$, $\mathrm{var}(X) < \infty$ and (recall that $\psi_0 = 1$)

$$E(X_t \mid \mathcal{G}_{t-1}) = \mathrm{e}^{\sum_{i=1}^\infty \psi_i\, \eta_{t-i}}\, E(Z_t\, \mathrm{e}^{\eta_t}) = 0 \quad \text{a.s.}$$

Hence (X_t) constitutes a finite variance martingale difference sequence and therefore the central limit theorem for stationary martingale sequences applies (see Billingsley [2]):

$$\sqrt{n}\, \overline{X}_n \xrightarrow{d} N(0, E(\sigma^2))\,.$$

Similarly, for $h > 0$, $(X_t X_{t+h})$ is adapted to the filtration $(\mathcal{G}_{t+h})$, and if in addition $E(\sigma^4) < \infty$ we have

$$\mathrm{var}(X_t X_{t+h}) = E(\sigma_0^2 \sigma_h^2) < \infty$$

and

$$E(X_t\, X_{t+h} \mid \mathcal{G}_{t+h-1}) = X_t\, E(X_{t+h} \mid \mathcal{G}_{t+h-1}) = 0 \quad \text{a.s.}$$

Therefore $(X_t X_{t+h})$ is a mean zero finite variance stationary martingale difference sequence. The ergodic theorem and the central limit theorem for stationary martingale differences yield

$$\begin{aligned}
\sqrt{n}\, \widehat{\gamma}_X(h) &= \sqrt{n} \sum_{t=1}^n X_t\, X_{t+h} - \sqrt{n}(\overline{X}_n)^2 + o_P(1) \\
&= \sqrt{n} \sum_{t=1}^n X_t\, X_{t+h} + o_P(1) \\
&\xrightarrow{d} N(0, E(\sigma_0^2 \sigma_h^2))\,.
\end{aligned}$$

Central limit theory can be derived for the sample means of the absolute values and squares of (X_t) as well as for $\widehat{\gamma}_{|X|^i}$, $i = 1, 2$, under additional strong mixing conditions. In Section 2.2 we have learned that (X_t) inherits strong mixing with a certain rate function (α_t) from the log-volatility sequence (σ_t). Given that the rate condition

$$\sum_{t=1}^\infty \alpha_t^{\delta/(2+\delta)} < \infty \tag{17}$$

holds for some $\delta > 0$, one can apply a classical central limit theorem, see Ibragimov and Linnik [23]. Condition (17) is satisfied if $\alpha_t \to 0$ at an exponential rate. It is satisfied e.g., for Gaussian ARMA log-volatility processes (Y_t) in (2). The central limit theorem applies to any strongly mixing sequence (A_t) with rate function (α_t) satisfying the conditions (17) and $E(|A|^{2+\epsilon}) < \infty$

for some $\epsilon > 0$. In particular, it is applicable to $A_t = \sigma_t$ and $A_t = |X_t|^p$ for any $p > 0$, but also for $A_t = |X_t X_{t+h}|^p$ for any $p > 0$. We omit further details.

It is also possible to derive limit theory with non-Gaussian limits for the ACVFs/ACFs of stochastic volatility processes (X_t), their absolute values and squares when standard moment conditions such as $\mathrm{var}(X) < \infty$ fail. Davis and Mikosch [11] (see also [12]) prove for regularly varying Z with index $\alpha \in (0, 2)$ and a Gaussian log-volatility process (Y_t) in (2) that the scaled sample ACVF $\widehat{\gamma}_X(h)$ at the fixed lag $h \geq 0$ converges in distribution to an infinite variance α-stable limit (see Mittnik [31] for a discussion of stable distributions) at a rate which depends on the tail of Z. Notice that in this case, X is regularly varying with index α and therefore $\mathrm{var}(X) = \infty$, see Section 4. In particular, the notions of ACVF/ACF are not defined. However, the sample ACF at a fixed lag h, $\widehat{\rho}_X(h)$, converges to zero even when the ACF is not defined. The rate at which this convergence happens is of the order $n^{1/\alpha}$, hence it is much faster than the common $\sqrt{n}$-rates for Gaussian central limit theory. Analogous results apply to the sample ACFs of the absolute values $|X_t|$ and the squares X_t^2. In the case of the squares one has to alter the condition of regular variation: Z must be regularly varying with index $\alpha \in (0, 4)$. Since, on the one hand, it has become common to study the sample ACFs of squared returns and, on the other hand, return series may have infinite moments of low order, the limit theory for the sample ACVFs/ACFs can be quite important in situations when one lacks sufficiently high moments.

We mention in passing that the limit theory for the sample ACVFs/ACFs of a stochastic volatility process, its absolute values and squares very much parallels the corresponding theory for an iid sequence, also in the infinite variance situation. Moreover, the limit theory for a heavy-tailed GARCH process (X_t) is of a completely different nature, see Davis and Mikosch [10], Mikosch and Stărică [30] and Basrak et al. [1]; cf. Davis and Mikosch [12] and Mikosch [28] for overviews. In particular, if the marginal distribution of a GARCH process is regularly varying with index $\alpha \in (2, 4)$ the sample ACF $\widehat{\rho}_X$ converges to 0 at a rate much slower than $\sqrt{n}$ and if $\alpha \in (0, 2)$ the sample ACF has a non-degenerate limit distribution without any normalization. The latter property would lead to completely different graphs for the sample ACFs for disjoint time intervals. This is another property which highlights a crucial difference between the stochastic volatility and GARCH models for returns.

References

1. Basrak, B., Davis, R.A. and Mikosch, T. (2002) Regular variation of GARCH processes. *Stoch. Proc. Appl.* **99**, 95–116.
2. Billingsley, P. (1968) *Convergence of Probability Measures.* Wiley, New York.
3. Billingsley, P. (1995) *Probability and Measure.* 3rd edition. Wiley, New York.
4. Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987) *Regular Variation.* Cambridge University Press, Cambridge.

5. Bougerol, P. and Picard, N. (1992) Stationarity of GARCH processes and of some nonnegative time series. *J. Econometrics* **52**, 115–127.
6. Bradley, R. Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys.* **2**, 107–144.
7. Breiman, L. (1965) On some limit theorems similar to the arc-sin law. *Theory Probab. Appl.* **10**, 323–331.
8. Brockwell, P.J. and Davis, R.A. (1991) *Time Series: Theory and Methods,* 2nd edition. Springer, Berlin, Heidelberg, New York.
9. Boussama, F. (1998) *Ergodicité, mélange et estimation dans le modelès GARCH.* PhD Thesis, Université 7 Paris.
10. Davis, R.A. and Mikosch, T. (1998) Limit theory for the sample ACF of stationary process with heavy tails with applications to ARCH. *Ann. Statist.* **26** (1998), 2049–2080.
11. Davis, R.A. and Mikosch, T. (2001) Point process convergence of stochastic volatility processes with application to sample autocorrelations. *J. Appl. Probab. Special Volume: A Festschrift for David Vere-Jones* **38A**, 93–104.
12. Davis, R.A. and Mikosch, T. (2001) The sample autocorrelations of financial time series models. In: W.J. Fitzgerald, R.L. Smith, A.T. Walden and P.C. Young (Eds.) *Nonlinear and Nonstationary Signal Processing.* Cambridge University Press, Cambridge (2001), 247–274.
13. Davis, R.A. and Mikosch, T. (2007) Extreme value theory for GARCH models. This volume.
14. Davis, R.A. and Mikosch, T. (2007) Extreme value theory for stochastic volatility models. This volume.
15. Doukhan, P. (1994) *Mixing. Properties and Examples.* Lecture Notes in Statistics **85**. Springer Verlag, New York.
16. Doukhan, P., Oppenheim, G. and Taqqu, M.S. (Eds.) (2004) *Long Range Dependence.* Birkhauser, Boston.
17. Eberlein, E. and Taqqu, M.S. (Eds.) (1986) *Dependence in Probability and Statistics.* Birkhäuser, Boston.
18. Embrechts, P. and Goldie, C.M. (1980) On closure and factorization theorems for subexponential and related distributions. *J. Austral. Math. Soc. Ser. A* **29**, 243–256.
19. Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.
20. Fan, J. and Yao, Q. (2003) *Nonlinear Time Series.* Springer, Berlin, Heidelberg, New York.
21. Feller, W. (1971) *An Introduction to Probability Theory and Its Applications II.* Wiley, New York.
22. Goldie, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1**, 126–166.
23. Ibragimov, I.A. and Linnik, Yu.V. (1971) *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen.
24. Krengel, U. (1985) *Ergodic Theorems.* De Gruyter, Berlin.
25. Leland, W.E., Taqqu, M.S., Willinger, W. and Wilson, D.V. (1993) On the self-similar nature of Ethernet traffic. *ACM/SIGCOMM Computer Communications Review* 183–193.
26. Lindner, A. (2007) Fundamental properties of GARCH processes. This volume, pp.???.

27. MANDELBROT, B. (1963) The variation of certain speculative prices. *J. Busin. Univ. Chicago* **36**, 394–419.
28. MIKOSCH, T. (2003) Modelling dependence and tails of financial time series. In: B. FINKENSTÄDT AND H. ROOTZÉN (EDS.) *Extreme Values in Finance, Telecommunications and the Environment.* Chapman and Hall, pp. 185–286.
29. MIKOSCH, T., RESNICK, S., ROOTZÉN, H. AND STEGEMAN, A. (2002) Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.* **12**, 23–68.
30. MIKOSCH, T. AND STĂRICĂ, C. Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *Ann. Stat.* **28** (2000), 1427–1451.
31. MITTNIK, S. (2007) Infinite variance processes. This volume.
32. MOKKADEM, A. (1990) Propriétés de mélange des processus autoregréssifs polynomiaux. *Ann. Inst. H. Poincaré Probab. Statist.* **26**, 219–260.
33. NELSON, D.B. (1990) Stationarity and persistence in the GARCH$(1,1)$ model. *Econom. Theory* **6**, 318–334.
34. PETROV, V.V. (1995) *Limit Theorems of Probability Theory.* Oxford University Press, Oxford.
35. PHAM, T.D. AND TRAN, L.T. (1985) Some mixing properties of time series models. *Stoch. Proc. Appl.* **19**, 279–303.
36. RESNICK, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.
37. SAMORODNITSKY, G. AND TAQQU, M.S. (1994) *Stable Non–Gaussian Random Processes. Stochastic Models with Infinite Variance.* Chapman and Hall, London.
38. WILLINGER, W., TAQQU, M.S., SHERMAN, R. AND WILSON, D. (1995) Self-similarity through high variability: statistical analysis of ethernet lan traffic at the source level. Proceedings of the ACM/SIGCOMM'95, Cambridge, MA. *Computer Communications Review* **25**, 100–113.