Evaluating the Extremal Index in GARCH Processes through Double Random Walk

Fabrizio Laurini
Department of Statistical Science, University of Padova,
Via Cesare Battisti 241/243, 35121, Padova,
Italy.
Email: flaurini@stat.unipd.it

15th February 2003

Abstract

The extremal index of a time series describes its tendency to cluster by observations above high thresholds. This phenomena, very common in financial time series, is referred to as volatility clustering. Recent research has shown the theoretical feasibility of computing the extremal index for a GARCH(1,1) process, though it has so far not been done in practice. Any GARCH(1,1) process can be expressed in the context of stochastic difference equations, but, unlike the case of ARCH(1), this does not help in computing the cluster-size distribution of those observations exceeding high thresholds. I approach computing this distribution by simulating the joint behaviour of the sequence of returns and volatility-returns, coercing the first observation of both processes to be extreme in some sense. This method is related to the notion of rejection sampling, and it avoids the intensive computation required in generating GARCH observations. Through a double random walk embedded in the stochastic difference equations, I first generate independent chains and then evaluate the empirical cluster-size distribution of observations above the thresholds.

Keywords: *GARCH processes; Regular variation; Bivariate extremes; Rejection sampling.*

1 Introduction

For discrete-time continuous-valued stationary time series the parameter which describes the strength of short-range dependence is the extremal index θ, with $0 \leq \theta \leq 1$, where the level of extremal dependence increases with decreasing θ. Formally, if $X_1, \ldots, X_n$ is a stationary sequence of random variables satisfying the long-range asymptotic independence condition $D(u_n)$ of Leadbetter *et al.* (1983) then

$$\Pr\{\max(X_1, \ldots, X_n) \leq u_n\} \to \{G(x)\}^\theta \text{ as } n \to \infty.$$

The normalizing sequence $u_n = a_n x + b_n$, with $a_n > 0$ and $b_n \in \mathbb{R}$, is selected to ensure that the limit distribution is non-degenerate, and $G(x)$ is the limit distribution of $\Pr\{\max(X_1^*, \ldots, X_n^*) \leq u_n\}$. Here, $X_1^*, \ldots, X_n^*$ represents a sequence of random variables independent, but with the

same univariate marginal distribution as $X_1, \ldots, X_n$. This result shows that θ is the key measure of short-range dependence in extreme values, with the value $\theta = 1$ being of particular importance as the asymptotic behaviour of $\max(X_1, \ldots, X_n)$ and $\max(X_1^*, \ldots, X_n^*)$ is identical even in presence of dependence in the sequence $X_1, \ldots, X_n$.

Subject-matter specialists for financial time series decompose the series into a volatility term and a innovation series. There are three standard types of volatility dependence model: autoregressive conditional heteroskedastic (ARCH), stochastic volatility (SV) and generalised ARCH (GARCH). The extremal index is known for the first two of these, see de Haan *et al.* (1989) and Breidt and Davis (1998) respectively, whereas Mikosch and Stărică (2000) shows that the extremal index in GARCH$(1,1)$ can be computed. They refrain from numerical evaluation of θ in GARCH$(1,1)$ processes, but they show several nice properties of regular variation. Here we develop an algorithm for the evaluation of the extremal index for the GARCH process.

2 Technical conditions for the square of GARCH$(1,1)$

Given a multiplicative process $X_t = \sigma_t \times z_t$, we can define the GARCH$(1,1)$ processes, introduced by Bollerslev (1986), as the volatility σ_t takes the form

$$\sigma_t = (\beta + \lambda X_{t-1}^2 + \gamma \sigma_{t-1}^2)^{1/2}$$

and z_t is referred to as a White Noise term $z_t \sim WN(0, \sigma_z^2)$. For stationarity, of both X_t and σ_t, we need some specific constraint for λ and γ. If no otherwise stated, we assume the stronger condition of Gaussianity for z_t, thus $z_t \sim N(0, \sigma_z^2)$.

In a recent work Mikosch and Stărică (2000) show that computing the extremal index for a GARCH$(1,1)$ process is feasible, even though they refrain from calculating numerically the values of θ given the values (λ, γ). To prove that the extremal index can be computed in GARCH$(1,1)$ processes it is necessary to go through the work of Davis and Hsing (1995) extended by Davis and Mikosch (1998).

The strategy adopted by de Haan *et al.* (1989), for the evaluation of θ in ARCH(1) processes, is based on the idea suggested by Kesten (1973). He shows how to compute the needed cluster probabilities π_i, $(i = 1, 2, \ldots)$ for one dimensional random difference equations. Notice that as we can compute cluster probabilities, we immediately derive the extremal index: Given the sequence $\{\pi_i\}$ we have $\theta^{-1} = \lim_{n\to\infty} \sum_{i=1}^{n} i\pi_i$. The idea of de Haan *et al.* is adopted by Turkman and Turkman (1997) for bilinear models. Unfortunately, the same approach does not seem to have the same efficiency in the GARCH$(1,1)$ setting.

We notice that, in the wider context of stochastic difference equations, any GARCH$(1,1)$ can be written as

$$\mathbf{X}_t = \mathbf{A}_t \mathbf{X}_{t-1} + \mathbf{B}_t, \qquad t \in \mathbb{Z} \tag{1}$$

where $\mathbf{A}_t$ are *iid* random $d \times d$ matrices and $\mathbf{B}_t$ *iid* $d-$dimensional random vectors. In addition, $(\mathbf{A}_t, \mathbf{B}_t)$ is an *iid* sequence independent of $\mathbf{X}_{t-1}$ for any t. The stochastic difference equation setting for the GARCH$(1,1)$ model is met if we choose

$$\mathbf{X}_t = \begin{pmatrix} X_t^2 \\ \sigma_t^2 \end{pmatrix}, \qquad \mathbf{A}_t = \begin{pmatrix} \lambda z_t^2 & \gamma z_t^2 \\ \lambda & \gamma \end{pmatrix}, \qquad \mathbf{B}_t = \begin{pmatrix} \gamma z_t^2 \\ \gamma \end{pmatrix} \tag{2}$$

and $d = 2$. The literature on this topic is abundant; Goldie (1991) gives a wide range of model that can be set in random recurrences context.

With notation of equations (2), consider $A_t = \lambda z_{t-1}^2 + \gamma$, $t \in \mathbb{Z}$; let us indicate with A any general term. Loosely, assume $\mathbb{E}(\ln A) < 0$, $\mathbb{P}(A > 1) > 0$ and there exists $h_0 \leq \infty$ such that $\mathbb{E}A^h$ is finite for all $h < h_0$ and $\mathbb{E}A^{h_0} = \infty$. Then the following holds (Mikosch and Stărică, 2000):

1. The function $\mathbb{E}A^h = \mathbb{E}(\lambda z^2 + \gamma)^h = \psi(h)$ is continuous and convex in h and the equation

$$\mathbb{E}A^\nu = 1 \tag{3}$$

has a unique positive solution. With some algebra, it might be showed that the solution to equation (3) is given by

$$E(A^\nu) = \frac{1}{\sqrt{\pi}} \int_0^\infty (2\lambda s + \gamma)^\nu s^{-1/2} e^{-s} ds = 1. \tag{4}$$

2. There exists a positive constant δ_0 such that

$$\mathbb{P}(\sigma_t > x) \sim \delta_0 x^{-2\nu} \quad \text{and } \mathbb{P}(|X_t| > x) \sim \mathbb{E}|z_t|^{2\nu} \mathbb{P}(\sigma_t > x) \quad \text{for } x \to \infty.$$

3. Moreover, the joint tail behaviour of the lagged vector $\mathbf{Y}_h^{(2)} = (|X_0|^2, \sigma_0^2, \ldots, |X_h|^2, \sigma_h^2)$ is jointly regularly varying with index 2ν, i.e.

$$\frac{\Pr\big(|\mathbf{Y}_h^{(2)}| > xq, \mathbf{Y}_h^{(2)}/|\mathbf{Y}_h^{(2)}| \in \cdot\big)}{\Pr(|\mathbf{Y}_h^{(2)}| > q)} \to x^{-\nu} \Pr\{\Omega \in \cdot\} \quad \text{for } q \to \infty, \tag{5}$$

with this result holding for any norm on $\mathbf{Y}_h^{(2)}$, and with Ω being a random variable on the space defined by this norm. The distribution of Ω, termed the spectral measure, is given by

$$\Pr\{\Omega \in \cdot\} = \frac{E\left\{|\mathbf{Z}_h^{(2)}|^\nu I_{\{\mathbf{Z}_h^{(2)}/|\mathbf{Z}_h^{(2)}| \in \cdot\}}\right\}}{E\left\{|\mathbf{Z}_h^{(2)}|^\nu\right\}}, \tag{6}$$

where $\mathbf{Z}_h^{(2)} = \Big((z_0^2, 1), \ldots, \prod_{i=1}^h (\lambda z_{i-1}^2 + \gamma)(z_h^2, 1)\Big)$.

3 About the function $\mathbb{E}(\lambda z^2 + \gamma)^h = \psi(h)$

Given GARCH(1,1) parameters λ and γ, the first step is to evaluate ν to satisfy expression (4). The shape of $\psi(h)$ is sketched in Figure 1 for $\lambda = 0.11$ and $\gamma = 0.88$.

There is no closed analytical form for expression (4) but numerical integration is trivial and solution to find $\nu > 0$ is immediate as the function is convex with one root being $h = 0$. For a range of parameter values (λ, γ), ν is given in Table 1.

We will see that for small values of the roots in Table 1 we observe small values of the extremal index. Loosely, it seems that the parameter λ is more preeminent: Suppose a value for λ not necessarily small (like $\lambda \approx 0.5$). We have $\nu \approx 2$ regardless the value of γ. On the contrary, for values of $\gamma \approx 0.5$ we observe a root $\nu \approx 2$ only for $\lambda \geq 0.3$.

The roots given in Table 1 represent the starting point for obtaining the cluster probability distribution π_i and the extremal index θ of a GARCH$(1,1)$ process. Like for ARCH(1), we

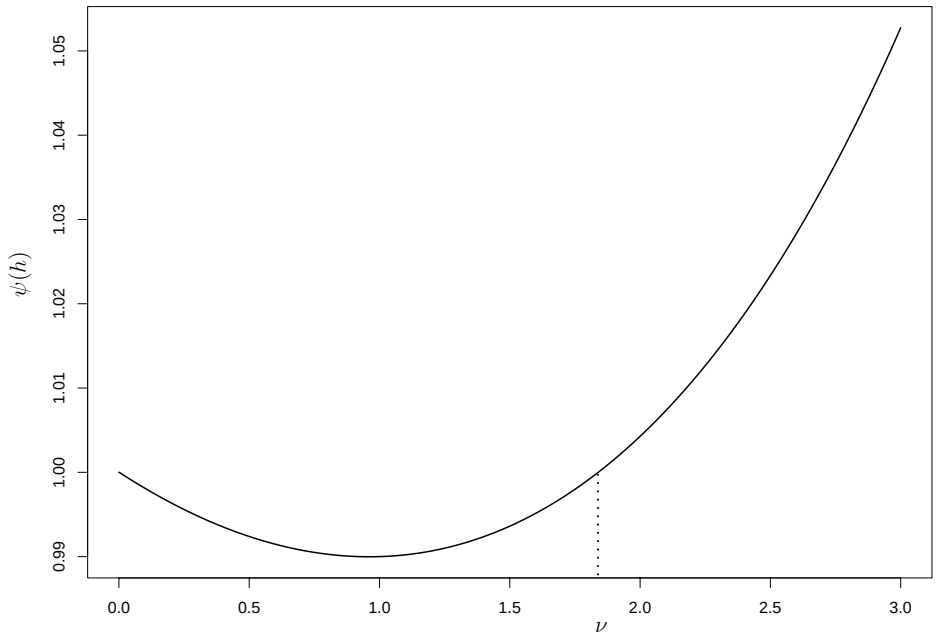

Figure 1: Plot of $\mathbb{E}(\lambda z^2 + \gamma)^h = \psi(h)$ for $\lambda = 0.11$ and $\gamma = 0.88$; the vertical dotted line represents the value $\nu = \nu(\lambda, \gamma)$ for which $\psi(h) = 1$.

cannot get the extremal index theoretically, but we can only have an approximation through simulations using random walks.

As stated above, to find the extremal index for GARCH$(1, 1)$ we need to develop a technique for solving the system of random equations (2). Apparently, the problem of having a closed form solution for $d > 1$ does not seem solved yet. To understand which is the cluster size distribution for excesses in a GARCH$(1, 1)$ processes is not straightforward. Potentially, this could be done in two ways.

a) The first method would consist in computing the extremal index from the product of σ_t and z_t of the GARCH$(1, 1)$ process. We know that z_t is *iid* and that σ_t satisfies the recurrence (1) for $d = 1$, $\mathbf{X}_t = \sigma_t$, $\mathbf{A}_t = \lambda z_{t-1}^2 + \gamma$ and $\mathbf{B}_t = \gamma$. In theory, we could easily get the cluster probabilities for the sequence z_t and, modifying slightly the approach of de Haan *et al.* (1989), we can get the cluster probabilities for the volatility process σ_t. Some research need to be developed in this direction, since the overall effect obtained by combining the extremal properties of two processes is unknown.

b) A second approach, which we follow, consists in simulating the joint distribution of the sequence (X_t^2, σ_t^2) keeping only the points for which X_t^2 is big enough. This is related to the idea of *rejection sampling* (e.g. Ripley, 1987) and avoid the intensive computation which an extremely long-range simulation GARCH would require. For financial time series, we are not aware of any work done in this framework.

GARCH Parameters		γ 0.1	0.3	0.5	0.7	0.9
	0.05	25.781	23.507	20.908	17.651	10.544
	0.10	12.697	11.505	10.097	8.098	–
	0.15	8.334	7.497	6.458	4.743	–
	0.20	6.151	5.486	4.604	2.956	–
	0.25	4.841	4.272	3.463	1.813	–
	0.40	2.869	2.415	1.654	–	–
λ	0.45	2.502	2.061	1.296	–	–
	0.60	1.764	1.329	–	–	–
	0.65	1.588	1.153	–	–	–
	0.80	1.194	–	–	–	–
	0.85	1.092	–	–	–	–

Table 1: Roots $\nu = \nu(\lambda, \gamma)$ for the expression $\mathbb{E}(\lambda z^2 + \gamma)^h = \psi(h) = 1$.

4 Simulating the joint distribution of X_t^2 and σ_t^2 in GARCH$(1,1)$

Very often it may be desirable simulating directly a cluster of exceedances instead of simulating the whole time series and extracting clusters of values above a threshold.

The technique we are going to set up comes from the study of bivariate extremes, which are a special case of multivariate extremes of Barnett (1976). Throughout the whole section, for a given vector $\mathbf{x}$, we use the $\mathcal{L}^1$ norm, thus $|\mathbf{x}| = \sum |x_i|$.

We have seen that expression (5) is the product of independent radial and angular pseudo-polar components, where the radial is $x^{-\nu}$ and the angular is $\mathbb{P}\{\Omega \in \cdot\}$.

Our aim is to simulate contemporaneously a couple (X_1^2, σ_1^2) but, in practice we can concentrate on the vector $\mathbf{Z}_h^{(2)}$. To understand this argument observe that the original space of interest (X_t^2, σ_t^2) can be considered as $(X_t^2, \sigma_t^2) = \sigma_t^2(z_t^2, 1)$; this motivates why we need to focus only on the couple $(z_t^2, 1)$. Due to the connections with bivariate and multivariate extremes of Coles and Tawn (1991) and Coles and Tawn (1994), we also need to define the pseudo-polar coordinates as

$$R = 1 + z \quad \text{and} \quad \Omega = \frac{1}{1+z}.$$

We justify the introduction of pseudo-polar coordinates recalling a very important result for bivariate extremes.

Theorem 1 *Let $(X_i, Y_i)_{i \in \mathbb{N}}$ be a sequence of independent bivariate observations from a distribution with standard Fréchet margins that satisfies the convergence for componentwise maxima*

$$\mathbb{P}\left\{ \frac{\max\limits_{i=1,\dots,n}(X_i)}{n} \leq x, \frac{\max\limits_{i=1,\dots,n}(Y_i)}{n} \leq y \right\} \to G(x,y)$$

where G is the bivariate extreme value distribution. Let (N_n) be a sequence of point processes defined by

$$N_n = \left\{ \left(\frac{X_1}{n}, \frac{Y_1}{n} \right), \ldots, \left(\frac{X_n}{n}, \frac{Y_n}{n} \right) \right\}.$$

Then, on regions bounded from the origin $(0,0)$

$$N_n \xrightarrow{d} N$$

where N *is a non-homogeneous Poisson process on* $(0,\infty) \times (0,\infty)$. *Moreover, transforming* X_i *and* Y_i *in pseudo-polar components* (R, Ω) *as above, the intensity function of* N *can be written as*

$$\lambda(r, \omega) = 2 \frac{dH(\omega)}{r^2} dr$$

for a distribution function H *on* $[0,1]$ *(related to* G*) which satisfies some mean constraint.*

What is relevant to observe is that intensity of the limit process factorizes across radial and angular components; therefore the angular distribution of points is independent of radial distance. This property explains the need to transform in pseudo-polar coordinates.

After simulating from the radial and angular component we trace back to the original space (X_t^2, σ_t^2) through the transformations $X_t^2 = R(1-\Omega)$ and $\sigma_t^2 = R\Omega$.

Simulating from the joint distribution (5) can be done in two steps: get $x^{-\nu}$ and, independently, simulating from $\mathbb{P}\{\Omega \in (0,\omega)\}$. The radial component can be simulated directly from inversion. We generate from a uniform b_i, in $[0,1]$, $i = 1, \ldots, M$. Inverting with $x = b^{-1/\nu}$ we get the radial component R of the pseudo-polar coordinates.

On the contrary, a Monte Carlo technique is required from simulating the angular component in (6). For our purposes it suffices to chose $h = 0$.

Replacing theoretical expectation of expression (6) with empirical long-range average, we have the empirical distribution of $\mathbb{P}\big\{\Omega \in (0,\omega)\big\}$. Therefore, we can build a lookup table of the function $f : (0,\omega) \to \mathbb{P}_\Omega(0,\omega)$. Afterward, we generate a vector $u_i, i = 1, \ldots, M$ of independent observations from a uniform in $(0,1)$ that represent our values on the $y-$axis of our lookup table. For each u_i we take the correspondent $x-$value which is $\omega_i = \inf\{f \geq u_i\}$. This is like generating ω_i values from the random variable Ω which represents the angular component of our pseudo-polar coordinates. The shape of the distribution function of $\Pr\{\Omega \in (0,\omega)\}$, for some values of ν, is sketched in Figure 2.

In the (X_t^2, σ_t^2) space we would keep only those observations for which X_t^2 is fairly big. Indeed, we originally simulate R from a conditional distribution. Expression (5) tells us that R and Ω are independent and that, in the limit, the regular variation property holds since

$$\frac{\mathbb{P}(R > xq)}{\mathbb{P}(R > q)} = x^{-\nu}.$$

This can be thought as the conditional probability $\mathbb{P}(R > xq | R > q)$, for $x > q$; for this reason we can argue that sampling from R correspond to sample from the conditional distribution

$$F_{\frac{R}{q}>1}(x;\nu) = \mathbb{P}\left\{ \frac{R}{q} > x \mid \frac{R}{q} > 1 \right\} = x^{-\nu}.$$

This argument suggests to adopt $q = 1$ as threshold for X_t^2; values in the (X_t^2, σ_t^2) space for which $X_t^2 \leq 1$ will be then discarded. Some improvement can be achieved if we could directly

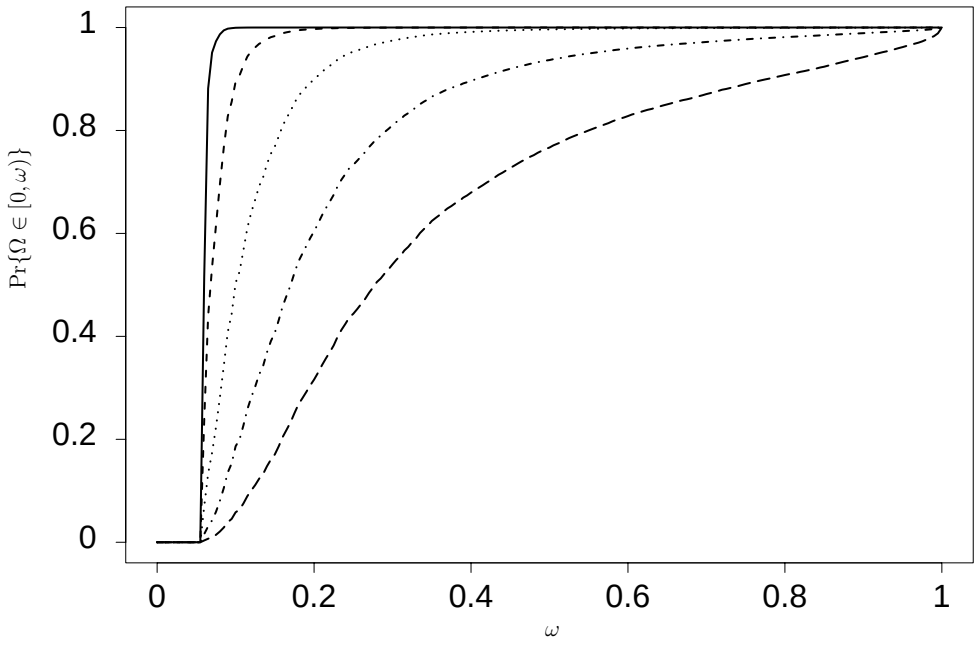

Figure 2: Distribution function of the spectral measure $\Pr\{\Omega \in (0,\omega)\}$. From top to bottom respectively, lines represent all values of λ in Table 1 which give decreasing roots ν when $\gamma = 0.7$.

simulate from the distribution of R and Ω such that $X_t^2 = (1-\Omega)R > 1$.

Suppose to have drawn a value $r^\star$ from R; it means that the values of Ω for which $X_t^2 > 1$ must lie in $(0,\omega^\star)$ where $\omega^\star = 1 - (1/r^\star)$. It follows that drawing $u^\star$ from a uniform in $U(0, f^{-1}(\omega^\star))$ and mapping through the lookup table we have a value from the Ω distribution such that $(1-\Omega)R > 1$.

5 Extremal index for GARCH$(1,1)$

Once generated the seed (X_1, σ_1) for the chain (2) we can build a cluster of exceedances. After the generation of the innovation term from a χ_1^2 we then build the matrix A_t in expression (2). Obtaining the process (X_t^2, σ_t^2), $t = 1, \ldots, T$ conditioned on having a big X_1^2 is straightforward; in practice we generate a double random walk chains of length $T = 100$. For each random walk chain we can count the frequency of observing points $X_1, \ldots, X_{100}$ bigger than $q = 1$. Therefore $\sum I_{X_j > 1}$ is the number of exceedances in each cluster. Averaging for all the clusters generated we obtain the extremal index for the square of a GARCH$(1,1)$ with related cluster probabilities.

It can be shown that, given the extremal index θ_{X^2} of the square of a GARCH$(1,1)$, $\{X_t^2\}$ (with cluster probabilities defined as $\tilde{\pi}_i$), the extremal index of the GARCH$(1,1)$ process can be computed with

$$\theta_X = 2\theta_{X^2}\big(1 - \tilde{\Pi}(0.5)\big) \tag{7}$$

where $\tilde{\Pi}$ is the probability generating function corresponding to the limiting compound Poisson process of the point processes of exceedances of thresholds u_n^2 by the square process $\{X_t^2\}$.

Mikosch and Stărică (2000) show that θ_{X^2} can be evaluated by simulation. Through the proof of Theorem 4.1 in their paper, it can be seen that

$$\theta_{X^2} = \frac{\lim\limits_{k\to\infty} \mathbb{E}\left\{ |z_1|^{2\nu} - \max\limits_{j=2,\ldots,k+1} \left| z_j^2 \prod_{i=2}^{j} (\lambda z_{i-1}^2 + \gamma) \right|^{\nu} \right\}_+}{\mathbb{E}|z_1|^{2\nu}}$$

where $\{0, x\}_+ = \max(0, x)$. A very convenient way to write cluster probabilities $\tilde{\pi}_i$ is

$$\tilde{\pi}_i = \frac{\theta_i^{(2)} - \theta_{i+1}^{(2)}}{\theta_{X^2}}, \qquad i \in \mathbb{N}.$$

Each difference $\theta_i^{(2)} - \theta_{i+1}^{(2)}$ tells us the number of clusters over the number of exceedances with *exactly* i values above the threshold u_n^2. Dividing $\theta_i^{(2)} - \theta_{i+1}^{(2)}$ by the total number of clusters, we get the required quantities.

Once computed probabilities $\tilde{\pi}_i$ we can evaluate the extremal index for any GARCH$(1,1)$ process through expression (7). Moreover, we can move from $\tilde{\pi}_i$, of a square GARCH$(1,1)$, to the cluster size distribution π_i, of the corresponding GARCH$(1,1)$ process. Considering excess over u_n of a GARCH process having half probability to occur than excess of the square GARCH over u_n^2, with some combinatorial arguments, it can be shown that

$$\pi_i = \left(1 - \tilde{\Pi}(0.5)\right)^{-1} \sum_{m=i}^{\infty} \binom{m}{m-i} \tilde{\pi}_m 2^{-m}, \qquad i = 1, 2, \ldots.$$

Evaluating the extremal index through simulations can arise problems of accuracy; we achieve any desired level of accuracy with the following arguments. We know that, in general, θ can be thought as the ratio between the total number of clusters $\bar{Z} = \sum_j^M Z_j$ over the total number of exceedances $N = \sum_i^T \sum_j^M I_{X_{i,j}>q}$; here $i = 1, \ldots, T$ is the length of a cluster (typically $T = 100$), while $j = 1, \ldots, M$ is the number of clusters. In our procedure we can control the number of clusters M. Once given the number of clusters the number of exceedances is random and related to the parameters space of the GARCH process.

Our target is to be accurate at second decimal position, at least for the extremal index of the square of GARCH processes, for which $\theta_{X^2} = \theta_1^{(2)}$. Since $\theta_1^{(2)}$ is computed as the ratio between clusters number and exceedances number we want

$$\hat{\theta}_1^{(2)} \pm \underbrace{2\sqrt{Var(\hat{\theta}_1^{(2)})}}_{0.001}$$

where $Var(\hat{\theta}_1^{(2)}) = \hat{\theta}_1^{(2)}(1-\hat{\theta}_1^{(2)})/\bar{Z}$. Notice that $\bar{Z}$ is the number of clusters, and not the number of exceedances, but, as stated above, is the only variable we can control in our simulation procedure. In the worst case of $\theta_1^{(2)} = 0.5$ we have to generate $\bar{Z} = 1,000,000$ clusters and from each cluster start generating the double random walk chain for the square of a GARCH$(1,1)$. Notice that for assessing the accuracy of all $\hat{\theta}_i^{(2)}$ we would proceed as above, and evaluating the minimal number of clusters which allows the specified two-decimal position accuracy. Nevertheless, we believe that this has little effect for GARCH processes, being $\theta_1^{(2)}$ the leading parameter for all clusters probabilities $\tilde{\pi}_i$.

We report some results from our simulation scheme in Table 2 indicating directly the value of the extremal index and cluster probabilities for GARCH$(1,1)$ process. Notice that $\hat{\theta}_X$ is computed with relation (7) and not through the use of cluster probabilities π_i with the obvious

λ	γ	ν	$\hat{\theta}_X$	π_1	π_2	π_3	π_4	π_5
0.1	0.3	11.505	0.997	0.9974	0.0026	0.0000	0.0000	0.0000
0.1	0.5	10.097	0.995	0.9950	0.0049	0.0000	0.0000	0.0000
0.09	0.7	9.184	0.989	0.9898	0.0098	0.0004	0.0000	0.0000
0.15	0.1	8.334	0.991	0.9909	0.0089	0.0002	0.0000	0.0000
0.15	0.7	4.743	0.901	0.9135	0.0694	0.0125	0.0031	0.0009
0.25	0.5	3.463	0.851	0.8682	0.1012	0.0218	0.0059	0.0018
0.2	0.7	2.956	0.674	0.7427	0.1505	0.0547	0.0244	0.0121
0.4	0.3	2.415	0.775	0.8007	0.1405	0.0387	0.0127	0.0045
0.45	0.3	2.061	0.707	0.7451	0.1636	0.0540	0.0209	0.0088
0.11	0.88	1.838	0.171	0.3244	0.1404	0.0907	0.0662	0.0512
0.25	0.7	1.813	0.354	0.4926	0.1821	0.0991	0.0618	0.0413
0.65	0.3	1.153	0.412	0.5004	0.2011	0.1073	0.0637	0.0401

Table 2: The extremal index $\hat{\theta}_X$ and the cluster probabilities π_i, $i = 1, \ldots, 5$ for the GARCH$(1, 1)$ process for a range of values of λ and γ.

relation $\hat{\theta}_X^{-1} = \sum j\pi_j$; however, results are broadly consistent and evaluating $\hat{\theta}_X$ on both ways does not make any difference.

6 Discussion

Table 2 gives the extremal index θ_X for the GARCH$(1, 1)$ process, illustrating that θ_X depends, in a rather complicated way, on parameters λ and γ. For fixed γ, θ_X decreases monotonically with increasing λ. Similarly, for fixed λ, θ_X decreases monotonically with increasing γ. Though θ_X changes in a largely consistent way with ν, with smaller θ_X occurring for smaller ν, this property breaks down when $\lambda + \gamma \approx 1$, i.e. when the process is close to non-stationary.

Some generalization can be given for more complicated GARCH processes, since in a recent work Basrak *et al.* (2002) set any GARCH(p, q) process in the context of stochastic difference equations. Moreover, they also proof that high-order GARCH processes share the same properties of regularity of variations in the tails as for the GARCH$(1, 1)$ case.

Final Remark

This paper represents part of a wider work which is in progress and co-authored by Jonathan A. Tawn, Department of Mathematics and Statistics, Lancaster University, UK.

References

Barnett, V. (1976) The ordering of multivariate data (with discussion). *Journal of the Royal Statistical Society, Series A*, **139**, 318–355.

Basrak, B., Davis, R. A. and Mikosch, T. (2002) Regular variation of GARCH processes. *Stochastic Processes and their Applications*, **99**, 95–115.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.

Breidt, J. F. and Davis, R. A. (1998) Extremes of stochastic volatility models. *The Annals of Applied Probability*, **8**, 664–675.

Coles, S. G. and Tawn, J. A. (1991) Modelling extreme multivariate events. *Journal of the Royal Statistical Society, Series B*, **53**, 377–392.

Coles, S. G. and Tawn, J. A. (1994) Statistical methods for multivariate extremes: An application to structural design (with discussion). *Applied Statistics*, **43**, 1–48.

Davis, R. A. and Hsing, T. (1995) Point process and partial sum convergence for weakly dependent random variables with infinite variance. *The Annals of Probability*, **23**, 879–917.

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

de Haan, L., Resnick, S. I., Rootzén, H. and de Vries, C. G. (1989) Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Processes and their Applications*, **32**, 213–224.

Goldie, C. M. (1991) Implicit renewal theory and tails of solutions of random equations. *Annals of Applied Probability*, **1**, 126–166.

Kesten, H. (1973) Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, **131**, 207–248.

Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes*. New York: Springer-Verlag.

Mikosch, T. and Stărică, C. (2000) Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *The Annals of Statistics*, **28**, 1427–1451.

Ripley, B. D. (1987) *Stochastic Simulation*. Wiley.

Turkman, K. F. and Turkman, M. A. A. (1997) Extremes of bilinear time series models. *Journal of Time Series Analysis*, **18**, 305–319.