# Extremes of high-order IGARCH processes

## *Valori estremi per processi IGARCH di ordine elevato*

Fabrizio Laurini

**Abstract** The extremal properties of GARCH processes are of wide interest for market risk management. Only for simple GARCH(1,1) extremes have been widely discussed. Much remains to be found about the dependence structure of extreme values for higher order GARCH. Although recent research has identified the multivariate regular variation property of stationary GARCH($p,q$) processes, currently there are no methods for numerically evaluating extreme components, like the average length of an extreme period. Only very simple special cases are well understood, but these are of little practical relevance, as bounded distribution of the error term is assumed. We present a unified toolkit that tackles the above critics and it is usable for Integrated GARCH($p,q$) processes, assuming innovations with unbounded support or asymmetry. With our method we are able to generate the forward tail chain of the process to derive all extremal features. The convergence of our numerical algorithm is very fast due to an efficient implementation of a particle filtering simulation technique.

**Abstract** *Le proprietà estreme dei processi GARCH sono di interesse centrale per la gestione del rischio di mercato. Soltanto per semplici GARCH(1,1) i valori estremi sono stati caratterizzati. Molto resta da studiare riguardo la dipendenza dei valori estremi per processi GARCH di ordine più elevato. Sebbene recentemente si siano usate connessioni con le proprietà di variazione regolare multivariata dei processi GARCH(p,q), al momento non ci sono metodi in grado di quantificare tali caratteristiche estreme. Pertanto, soltanto casi speciali sono stati compresi a fondo, ma questi sono spesso irrilevanti da un punto di vista pratico, visto che viene assunta una forma sul termine d'errore con distribuzione limitata. Si presenta un insieme di tecniche unificato volto a superare tutti questi inconvenienti ed è usabile anche per processi GARCH(p,q) Integrati con innovazioni a supporto illimitato e asimmetriche. Il metodo si basa sulla cosiddetta forward tail chain per derivare tutti gli*

F. Laurini
Department of Economics and Management, University of Parma, Via J. F. Kennedy, 6, Parma, e-mail: fabrizio.laurini@unipr.it

*aspetti rilevanti nel processo dei valori estremi. La convergenza dell'algoritmo è veloce grazie all'utilizzo di un'implementazione efficiente di particle filtering.*

**Key words:** Extremes, IGARCH, Particle filtering, Regular variation

## 1 Introduction

The risk management in the stock markets, commonly called *market risk management*, suggests the use of statistical tools and models which aim at reducing the potential size of losses, occurring by sudden drops or growth in the stock market. Such losses are even amplified when the volatility of the stock market is substantial. Hence, modeling and forecasting the temporal evolution of the market volatility is of great concerns for financial institutions.

Consider the daily log-returns $X_t = \log P_t - \log P_{t-1}$, ($X_t \in \mathbb{R}$) where $P_t$, $t = 1,2,\ldots,$ is the price of a generic asset. Then a broad class of models, mostly adopted to describe the market volatility, is the generalized autoregressive conditionally heteroscedastic (GARCH) introduced by [3]. For market risk management one of the most important issue is the presence of extreme values of daily log-returns. Therefore, understanding the extreme properties for such processes is fundamental, and this can be achieved by considering the marginal and the clustering properties of GARCH processes.

GARCH$(p,q)$ models, for integers $p$ and $q$, have the form

$$X_t = \sigma_t Z_t \quad \text{with} \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \quad t = 1,2,\ldots, \quad \text{and} \quad \alpha_0 > 0. \tag{1}$$

For fixed $t$ $Z_t$ and $\sigma_t$ are independent. The independent and identically distributed (IID) sequence $\{Z_t\}_{t\geq 1}$ is assumed to be symmetric with $E(Z_t^2) = 1$. Conditions on the parameters $\alpha_i, i = 1,\ldots,q$ and $\beta_j, j = 1,\ldots,p$ are discussed in Section 2.

GARCH$(p,q)$ processes $\{X_t\}_{t\geq 1}$ satisfies mixing conditions, so that the key parameter for quantifying the impact of extreme values is the extremal index $\theta_X \in (0,1]$. The extremal index $\theta_X$ measures the level of clustering of extreme values, with the clustering of extreme increasing for $\theta_X$ decreasing.

An important interpretation of the extremal index is provided using the cluster size distribution $\pi_X(i), i = 1,2,\ldots$, since $\sum_{i=1}^{\infty} i\pi_X(i) = (\theta_X)^{-1}$, so $\theta_X$ is the reciprocal of the limiting mean cluster size of extreme values. The special case $\theta_X = 1$ means no clustering of extremes. Extremes of GARCH$(p,q)$ models have been studied by [1], but formulae for $\theta_X$ do not exist.

[2] were the first to propose computational algorithms for the evaluation of $\theta_X$. Their algorithms make a very strict assumption that the innovation $Z_t$ has bounded support, ruling out many important distributions used by practitioners, e.g., $Z_t$ being Gaussian or $t$-distributed. We propose an entirely new algorithm that does not re-

quire these assumptions, and critically allows unbounded support for $Z_t$, as we take $Z_t \sim \mathrm{ST}(0,1,\lambda,\nu)$ scaled to have unit variance.

The relevant feature of the proposed algorithm is to simulate directly the tail chain of a GARCH$(p,q)$ model. The tail chain of a process, introduced by [2], will be discussed later. The approach used here extends the algorithm of [4] for GARCH$(1,1)$ processes.

We derive the theory for obtaining the extremal index of a GARCH$(p,q)$ process, and we provide a Monte Carlo algorithm for the numerical evaluation $\theta_X$ and associated cluster size distribution. Precisely, with our algorithm we first obtain the cluster size distribution $\pi_{X^2}(\cdot)$ and the extremal index $\theta_{X^2}$ for the square of the process, and then derive $\pi_X(\cdot)$ and $\theta_X$. All results do not require the symmetry of $Z$.

## 2 Technical background, notation and assumptions

### *2.1 SRE representation for GARCH*

Let us start by defining stationarity for GARCH$(p,q)$ processes. We focus on the squared GARCH process, $X_t^2{}_{t\geq 1}$, and rewrite the process as a stochastic recurrence equation (SRE) as this enables the exploitation of a range of established results for such processes, e.g., the existence of results for the marginal distribution.

Let the $(p+q)$ vector $\mathbf{Y}_t$, the $(p+q)\times(p+q)$ matrix $\mathbf{A}_t$ and the $(p+q)$ vector $\mathbf{B}_t$ be

$$\mathbf{Y}_t = \begin{pmatrix} X_t^2 \\ \vdots \\ X_{t-q+1}^2 \\ \sigma_t^2 \\ \vdots \\ \sigma_{t-p+1}^2 \end{pmatrix} \quad \mathbf{A}_t = \begin{pmatrix} \alpha^{(q-1)}Z_t^2 & \alpha_q Z_t^2 & \beta^{(p-1)}Z_t^2 & \beta_p Z_t^2 \\ I_{q-1} & 0_{q-1} & I_{q-1} & 0_{q-1} \\ \alpha^{(q-1)} & \alpha_q & \beta^{(p-1)} & \beta_p \\ 0_{q-1} & 0_{p-1} & I_{p-1} & 0_{p-1} \end{pmatrix} \quad \mathbf{B}_t = \begin{pmatrix} \alpha_0 Z_t^2 \\ 0_{q-1} \\ \alpha_0 \\ 0_{q-1} \end{pmatrix} \tag{2}$$

where $\alpha^{(s)} = (\alpha_1,\ldots,\alpha_s) \in \mathbb{R}^{(s)}$, $\beta^s = (\beta_1,\ldots,\beta_s) \in \mathbb{R}^s$, $I_s$ is the identity matrix of size $s$, $0_s$ is a square matrix of zeros of size $s$ and $0_s$ is a column vector of zeros having length $s$. In each case here if $s<0$ then these terms are to be interpreted as being dimensionless. Then it follows that the squared GARCH$(p,q)$ processes satisfies the SRE

$$\mathbf{Y}_t = \mathbf{A}_t\mathbf{Y}_{t-1} + \mathbf{B}_t, \qquad t \geq 1, \tag{3}$$

where $\{\mathbf{A}_t\}_{t\geq 1}$ and $\{\mathbf{B}_t\}_{t\geq 1}$ are each sequences of IID stochastic matrices and vectors. The formulation of the SRE via (2) is less parsimonious than that of [1], but has the benefit of covering all GARCH$(p,q)$ processes, even when $p=q=1$.

It is necessary and sufficient that there is a negative top Lyapunov exponent of $\mathbf{A}_t$ for the existence of a unique, strictly stationary solution of SRE (3). Under the con-

dition $E\ln^+\|\mathbf{A}_t\| < \infty$ (here $\ln^+ x = \ln x$, if $x \geq 1$ and 0 otherwise), the top Lyapunov exponent is

$$\gamma = \lim_{t\to\infty} \frac{1}{t} \ln\|\mathbf{A}_t \mathbf{A}_{t-1} \cdots \mathbf{A}_1\| \tag{4}$$

almost surely, so that, via expression (4) a relatively simple simulation can be performed to obtain $\gamma$.

If $\sum_{j=1}^{p} \beta_j < 1$ then $\gamma < 0$. This stationary condition covers various forms of GARCH process including the IGARCH$(p,q)$ process which has the property that

$$\sum_{i=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j = 1. \tag{5}$$

For second-order stationarity a stronger condition is required, namely the left hand side of equation (5) is required to be strictly less than 1. This condition implies that $\gamma < 0$, and the second moment of $X_t^2$ are finite, and so is the fourth moment of $X_t$. So an IGARCH$(p,q)$ is strictly stationary but has infinite variance and so is not second-order stationary.

## *2.2 Tail chain process and regular variation for squared GARCH*

Taking a heavy tailed process $\{\mathbf{Y}_t\}_{t\geq 1}$ as strictly stationary, the tail chain is defined in the following way. When $u \to \infty$, if for any $t \geq 1$

$$(\mathbf{Y}_0/u, \mathbf{Y}_1/X_0, \ldots, \mathbf{Y}_t/\mathbf{Y}_0) \mid \|\mathbf{Y}_0\| > u,$$

converges weakly to $(\hat{\mathbf{Y}}_0, \hat{\mathbf{Y}}_1, \ldots \hat{\mathbf{Y}}_t)$. The tail process $\{\hat{\mathbf{Y}}_t\}_{t\geq 1}$ exists if and only if $\{\mathbf{Y}_t\}_{t\geq 1}$ is jointly regularly varying.

[1] show that there exists a unique stationary solution to the SRE (3) and this solution exhibits a multivariate regular variation property, i.e., for any $t \geq 1$, any norm $\|\cdot\|$ and all $r > 0$,

$$\frac{\Pr(\|\mathbf{Y}_t\| > rx, \mathbf{Y}_t/\|\mathbf{Y}_t\| \in \cdot)}{\Pr(\|\mathbf{Y}_t\| > x)} \xrightarrow{v} r^{-\kappa} \Pr(\mathbf{D}_t \in \cdot), \qquad \text{as } x \to \infty, \tag{6}$$

where $\xrightarrow{v}$ denotes vague convergence, $\kappa \geq 0$, and $\mathbf{D}$ is $p+q$ dimensional random vector in the unit sphere (with respect to a norm $\|\cdot\|$) defined by $\mathbb{S}^{p+q} \subset \mathbb{R}^{p+q}$. If condition (6) holds $\mathbf{Y}_t$ exhibits multivariate regularly variation with index $\kappa$ and $\mathbf{D}$ is termed the spectral tail process of the vector $\mathbf{Y}_t$.

The subsequent results link $\gamma$ to $\kappa$. There is structure imposed on both $\kappa$ and $\mathbf{D}$ by the GARCH$(p,q)$ process. In particular, $\kappa$ is the unique positive solution of the equation

$$\lim_{t\to\infty} \frac{1}{t} \ln E\left(\|\mathbf{A}_t \mathbf{A}_{t-1} \cdots \mathbf{A}_1\|^{\kappa}\right) = 0. \tag{7}$$

For all the numerical evaluations, we will use the norm $\|\mathbf{A}\| = \sum|a_{ij}|$. In general $\kappa$ can be found only by numerical solution of equation (7) via Monte Carlo. However for any IGARCH$(1,1)$ process $\kappa = 1$; see [1].

A consequence of the multivariate regular variation property (6) is that all the marginal variables of $\mathbf{Y}_t$ have regularly varying tails with index $\kappa$, so in particular for $r \geq 1$

$$\Pr(X_t^2 > rx \mid X_t^2 > x) \to r^{-\kappa}, \qquad \text{as } x \to \infty. \tag{8}$$

## 3 Tail chain for IGARCH process with asymmetric $Z$

### *3.1 Particle filter algorithm*

In this Section an algorithm to sample from $\mathbf{D}$ is presented. The property $E(\|\mathbf{AD}_0\|^\kappa) = 1$, proved by [2], is extensively used to this aim. After obtaining $\kappa$ via Monte Carlo (7), that require as inputs the coefficients and a sample from $Z^2$, the steps are the following

- Initialize with the estimate of the empirical distribution of $\widetilde{\mathbf{D}}_0$. The empirical distribution is estimated using $m$ extreme values from a simulated squared GARCH$(p,q)$, where the extremes are the $m$ largest values of the squared sequence.
- From the empirical distribution $\widetilde{\mathbf{D}}_0$ we initialize the procedure by taking $J$ particles with equal weight $w_0^{(j)} = 1/J$ and sample from $\widetilde{\mathbf{D}}_0$ with probabilities given by $w_0^{(j)}$. At this step of the sampling is made with replacement, i.e. it is possible to set $J > m$.
- The empirical distribution after the first run of the algorithm is computed by first using the transition
$$\mathbf{D}_1^\star = \mathbf{A}^{(j)}\widetilde{\mathbf{D}}_0, \qquad j = 1,\dots,J, \tag{9}$$
where the $\mathbf{A}^{(j)}$ are independent copies of $\mathbf{A}$. Since we neglect the random vector $\mathbf{B}$ when computing the transition (9), then $\mathbf{D}_1^\star$ has to be normalized. A proper (yet empirical) distribution can be obtained by simply scaling the $\mathbf{D}_1^\star$, i.e.
$$\widetilde{\mathbf{D}}_1 = \frac{\mathbf{D}_1^\star}{\|\mathbf{D}_1^\star\|}.$$
- The new particles weights are subsequently updated exploiting $E(\|\mathbf{AD}_0\|^\kappa) = 1$. We take advantage of that property by first storing
$$w_j^\star = \|\mathbf{D}_1^\star\|^\kappa, \quad \text{normalized with} \quad w_j^{(1)} = \frac{w_j^\star}{\sum_{j=1}^J w_j^\star}, \qquad j = 1,\dots,J. \tag{10}$$

- In general particle weights $w_j^{(1)}$ from iteration 1 are no longer identical, and that will be the case in all iterations.
- Iterate $S$ times using recursions similar to (9) and (10), where at each fixed replicate $s$, ($s = 1,\ldots,S$) the $J$ particles of $\widetilde{\mathbf{D}}_s$ are sampled, with replacement, with updated weights $w_j^{(s)}$, $j = 1,\ldots,J$.
- In our algorithm we check that these conditions hold at convergence. We noticed that with small $S$ the weights stabilize, and we have a good sample from $\mathbf{D}_t$, i.e. $\widetilde{\mathbf{D}}_S \to \mathbf{D}_t$.

### *3.2 Random thinning for the IGARCH*

Mapping from the squared process to the "original" process require to rule out extremes in the $X_t^2$ which do not belong to the $X_t$ process. First note that if $X_t^2$ is regularly varying with index $\kappa > 0$ then if

$$\Pr(Z_t > x \mid |Z_t| > x) \to \delta \qquad \text{as } x \to \infty, \tag{11}$$

where $0 < \delta < 1$ then it follows that $X_t$ is a regularly varying random variable, with index $\kappa/2$, in both its upper and lower tails.

To translate results about the tail chain of the squared GARCH process we take the first component of $\hat{\mathbf{Y}}_t$ to be denoted as $\{\hat{Y}_t\}$. To study properties for the tail chains of the GARCH $\{\hat{X}_t\}_{t\geq 1}$ process we adopt a similar strategy to [4]. It is key to recognise that there are two tails chains for $\hat{X}_t$, a lower and an upper tail chain $\hat{X}_t^l$ and $\hat{X}_t^u$ respectively, with $\hat{X}_t^u = B_t(\hat{Y}_t)^{1/2}$ and $\hat{X}_t^l = -B_t(\hat{Y}_t)^{1/2}$ where $B_t$ is a sequence of IID Bernoulli($\delta$) variables, with $\delta$ given by limit (11), where $B_t = \{-1, 1\}$ with respective probabilities $\{1-\delta, \delta\}$.

An extreme event for the tail chain $\{\hat{Y}_t\}_{t\geq 1}$ of the squared GARCH does not occur in the upper tail $\{\hat{X}_t^u\}_{t\geq 1}$ and lower tail $\{\hat{X}_t^l\}_{t\geq 1}$ with respective probabilities $P^u(\delta)$ and $P^l(1-\delta)$ where

$$P^u(\delta) = \sum_{i=1}^{\infty} \pi_{X^2}(i)\delta^i$$

and $\pi_{X^2}(i)$ is the probability that a cluster of length $i$ is in the $\{X_t^2\}$ series.

For the upper and lower tail behaviour of $\{X_t\}$ it follows that the respective extremal indices are

$$\theta_X^u(\delta) = \delta^{-1}\theta_{X^2}\{1 - P^u(\delta)\} \tag{12}$$

and $\theta_X^l(1-\delta)$ respectively, where $\theta_{X^2}$ is the extremal index of the squared GARCH process. Similarly, the upper and lower tail limiting cluster size distributions of $\{X_t\}$ are given by

$$\pi_X^u(i,\delta) = \Big\{1 - P(\delta)\Big\}^{-1} \sum_{j=i}^{\infty} \pi_{X^2}(j)\binom{j}{i}\delta^i(1-\delta)^{j-i} \text{ for } i = 1,2,\ldots. \tag{13}$$

and $\pi_X^l(i, 1-\delta)$. Thus once we have derived $\pi_{X^2}(\cdot)$, obtaining $\pi_X(\cdot)$ is immediate for both tails of the GARCH process.

## 4 Some results for the GARCH(2,2)

In this example a GARCH$(2,2)$ process is considered with standard Gaussian innovation $Z_t$ and parameters $\alpha_1 = \alpha_2 = 0.2$, $\beta_1 = 0.3$ and $\beta_2 = 0.25$.

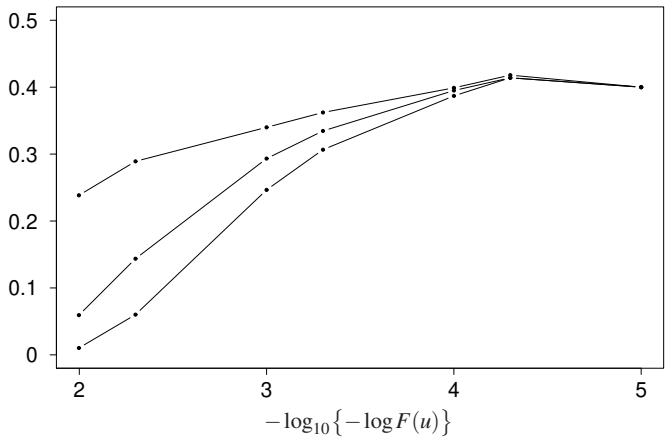


**Fig. 1** Runs estimation of the extremal index for a stationary GARCH$(2,2)$ process.

To check our results, we simulate $10^7$ values from such a configuration. An estimate of the extremal index with the runs method, which require run length (here $m = 100, 500, 1000$) and a sequence of high thresholds $u_n$, is considered. The length of the sample size reduces considerably the bias of runs method, even if only one simulation is considered. In Figure 1 we plot $u_n$ on the $-\log_{10}\{-\log F(u_n)\}$ scale. Such choice standardises the upper tail of the distribution. From this plot we can conjecture that the extremal index might lie in the interval 0.25–0.45, but we cannot say more.

From the same GARCH$(2,2)$ structure we now turn to our algorithm, with Monte Carlo value of $\kappa = 1.83$. Figure 2 shows the numerical evaluation of $\theta$ and the (little) sensitivity of our algorithm to the number of initial seeds and threshold selection. For each fixed level of threshold, darker points in Figure 2 correspond to lower number of initial seeds. More precisely, white empty circles show results $\theta$ with a number of initial seeds higher than other circles and thus more reliable (in this example white filled circles correspond to nearly 250000 initial cluster of extremes, while black filled circles are about 10000 initial seeds).

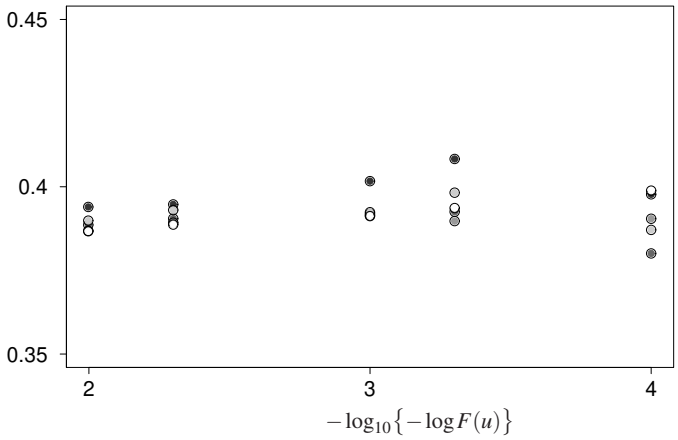


**Fig. 2** Sensitivity to $\theta_X$ for the GARCH$(2,2)$ process for a range of thresholds and a variety of initial seeds. For a fixed threshold darker circles correspond to a smaller number of initial seeds.

## 5 Summary and final remarks

We simulate a number of independent seeds conditioning on the event of being at extreme levels. For each initial seed we exploit the autoregressive property of GARCH processes and simulate only clusters of extreme values. The accuracy of our method relies on the number of such cluster of extremes. The usefulness of simulating from within a cluster is two-fold, as we avoid computation inefficiency derived by simulating long sequences of GARCH processes and reduce the influence on the subjective choice of a suitable high threshold.

**Acknowledgements** The work is part of a broader project involving Paul Fearnhead and Jonathan Tawn, Lancaster University - UK, which are kindly acknowledged. Simone Padoan read carefully an early draft and made a number of useful comments.

## References

1. Basrak, B., Davis, R. A., Mikosch, T.: Regular variation of GARCH processes. Stoch Proc Applicat, **99**, 95–115 (2002)
2. Basrak, B. and Segers, J.: Regularly varying multivariate time series. Stoch Proc Applicat, **119**, 1055–1080 (2009).
3. Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, **31**, 307–327.
4. Laurini, F, Tawn, J.A.: The extremal index for GARCH(1, 1) processes. Extremes, **15**, 511–529 (2012)