Extreme distribution for a generalized stochastic volatility model *

Aliou DIOP[a,b] † Dominique GUEGAN[c]

[a] *UFR de Mathématiques Appliquées et d'Informatique, B.P. 234*
Université Gaston Berger, Saint-Louis, Sénégal,
e-mail : adiop@ugb.sn
and
[b] *Université de Reims, Laboratoire de Mathématiques, UPRESA 6056*
B.P. 1039, 51687 Reims Cedex2 France,
aliou.diop@univ-reims.fr

[c] *E.N.S. Cachan, GRID UMR CNRS C8534,*
61 Avenue du President Wilson, 94235, Cachan Cedex, France,
e-mail : dguegan@wanadoo.fr

Abstract

We give the asymptotic behavior of the extreme values of Stochastic Volatility model $(Y_t)_t$ when the noise follows a generalized error distribution (GED). This class of distribution whose a presentation has been studied in Box and Tiao (1973) for instance includes in particular the Gaussian law. In this paper, we show that in the general context, the normalized extremes of a log-transformation of $(Y_t)_t$ converges in distribution to the double exponential distribution. We investigate the importance of the different assumptions using Monte Carlo simulations. We also deal with the finite distance behavior of the normalized maxima. The influence of the parameters of the models is discussed.

keywords : Extreme value theory, tail behavior, stochastic volatility model, GED distribution.

---

*running title : Extremes of stochastic volatility

†Corresponding author.

The first author gratefully acknowledges the financial support of ACMAD

1 Introduction

The class of stochastic volatility models has its roots and applications in finance and finance econometrics. Indeed, volatility plays a central role in the analysis of a lot of phenomenon in these domains. There exists a lot of versions of stochastic volatility models in the literature. Here we are interested by a discrete time version ($t \in \mathbb{Z}$) of volatility model introduced first by Taylor (1986). This model appears as a particular model of the SARV model introduced by Andersen in 1994. The univariate volatility model $(Y_t)_t$ that we will consider here can be defined by the following equation :

$$Y_t = \sigma exp(\frac{\alpha_t}{2})\varepsilon_t \tag{1}$$

with

$$\alpha_t = \sum_{j=0}^{\infty} \theta_j Z_{t-j} \tag{2}$$

where σ is a positive constant, $(\varepsilon_t)_t$ a sequence of independent and identically distributed (iid) random variable (r.v.) and $(Z_t)_t$ a sequence of iid Gaussian r.v. with mean zero, variance σ_Z^2 and independent of the sequence $(\varepsilon_t)_t$. The parameters (θ_j) are such that $\sum_{j\geq 0} |\theta_j|^2 < \infty$. In the literature, generally it is assumed that ε_t and Z_t are not correlated with each other for all t. For simplicity here we will assume independence between the two sequences. The model can pick up the kind of asymmetric behavior which is often found in stock prices, and a negative correlation between ε_t and Z_t induces a leverage effect. This explains why practitioners often use this model. The behavior of the autocorrelation function and the power-moments for the process defined by (1)-(2) are well known when ε_t follows a Gaussian law or a Student law, see for instance Harvey (1993), Taylor (1986), Ghysels *et al.* (1996) or Shepard (1996).

Recently banks, insurance companies faced with questions concerning extremal rare events. In insurance, these extremal events may clearly correspond to individual claims which by far exceed the capacity of a single insurance company. In finance these extremal events can present themselves spectacularly whenever major stock market crashes like the one in 1987 occurs for instance. The first preoccupation for these institutions is to define a well-functioning risk management and control system to caution to these problems. Thus, they need stochastic methodology for the construction of various components of such tools.

In that context, it is important to know, for instance the extremal distribution of the different process which are used, until now, to characterize the behavior of certain asset prices. Our contribution consists in studying the extremal behavior of the distribution of a class of transformation of the process $(Y_t)_t$ defined by (1)-(2) when the noise $(\varepsilon_t)_t$ follows a generalized error distribution (GED). This class of distribution whose a presentation has been studied in Box and Tiao (1973) for instance, includes in particular the Gaussian law. This last case has been studied by Breidt and Davis (1998). Here we show, in a general context specified in the next section, that the normalized extremes of a log-transformation of $(Y_t)_t$ converge in distribution to the double exponential distribution. Since we are interested by a feasible application of these results on real data, we investigate the importance of the different assumptions using Monte Carlo simulations. Thus, we are able to show that some assumptions are not so important in the reality context. Nevertheless our results obtained in an asymptotic context can be very bad at finite distance. This last situation is classical in extreme value theory, indeed, some results obtained in an asymptotic context are not always valid with finite samples. Obviously, this implies some difficulties to use directly these results in view to construct a risk management theory.

Our paper is organized in the following way : Section two contains statistical properties of the process (1)-(2) when $(\varepsilon_t)_t$ follows a GED distribution. The extremal behavior of a log-transformation of $(Y_t)_t$ is presented in Section three. In Section four we focus on finite samples behavior for the normalized maximum of the log-transformation of $(Y_t)_t$. Section five is devoted to the conclusion. The proofs are postponed in an Appendix.

2 Stationarity of the process $(Y_t)_t$.

In this section we precise some properties of the autocorrelation function and the power-moments of the process $(Y_t)_t$ defined in (1)-(2) when $(\varepsilon_t)_t$ follows a GED distribution. The GED density is defined by

$$f_\varepsilon(x) = c_0 exp(-k|x|^\gamma), \tag{3}$$

with $c_0 > 0$, $\gamma > 0$ and $k > 0$. The expression (3) can also be written :

$$f_\varepsilon(x) = \frac{\gamma \exp(-\frac{1}{2}|\frac{x}{\tau}|^\gamma)}{\tau 2^{1+\frac{1}{\gamma}}\Gamma(\frac{1}{\gamma})}, \qquad \gamma > 0 \tag{4}$$

with

$$\tau = [\frac{2^{-\frac{2}{\gamma}}\Gamma(\frac{1}{\gamma})}{\Gamma(\frac{3}{\gamma})}]^{\frac{1}{2}}.$$

This class of density contains the normal density ($\gamma = 2$), the Laplace density ($\gamma = 1$) and has the uniform density as a limit ($\gamma \to +\infty$). It was firstly introduced by Subbotin (1923) as the exponential power distribution. The tail behavior of the innovations process $(\varepsilon_t)_t$ which is characterized by such a density depends on the tail-thickness parameter γ. For instance, if $\gamma = 2$, then $\varepsilon_t \sim \mathcal{N}(0,1)$, while for $\gamma < 2$ the distribution has thicker tails than the Gaussian distribution. Now we precise some results concerning the moments of the process (1)-(2) assuming that the distribution of $(\varepsilon_t)_t$ is given by (3).

Proposition 1 (Covariance and strict stationarity)
If the process $(Y_t)_t$ follows the model (1)-(2) driven by a GED noise $(\varepsilon_t)_t$ with index γ then,
i) the power-moments of the process $(Y_t)_t$ are given by

$$E(Y_t^r) = \begin{cases} \sigma^r \frac{\Gamma(\frac{1}{\gamma})^{\frac{r}{2}-1}\Gamma(\frac{r+1}{\gamma})}{\Gamma(\frac{3}{\gamma})^{\frac{r}{2}}} exp(\frac{r^2}{8}\sigma_\alpha^2) & \text{if } r \text{ is even} \\ 0 & \text{if } r \text{ is odd} \end{cases} \tag{5}$$

and

$$var(Y_t^r) = \begin{cases} \sigma^{2r}\frac{\Gamma(\frac{1}{\gamma})^{r-2}\Gamma(\frac{r+1}{\gamma})^2}{\Gamma(\frac{3}{\gamma})^r} exp(\frac{r^2}{4}\sigma_\alpha^2)[K_r exp(\frac{r^2}{4}\sigma_\alpha^2) - 1] & \text{if } r \text{ is even} \\ \\ \sigma^{2r}\frac{\Gamma(\frac{1}{\gamma})^{r-1}\Gamma(\frac{2r+1}{\gamma})}{\Gamma(\frac{3}{\gamma})^r} exp(\frac{r^2}{2}\sigma_\alpha^2) & \text{if } r \text{ is odd} \end{cases} \tag{6}$$

where

$$K_r = \frac{\Gamma(\frac{1}{\gamma})\Gamma(\frac{2r+1}{\gamma})}{\Gamma(\frac{r+1}{\gamma})^2}. \tag{7}$$

ii) The excess kurtosis of Y_t is given by

$$\kappa = 3(\frac{K_2}{3} exp(\sigma_\alpha^2) - 1), \tag{8}$$

where K_2 is given by (7) with $r = 2$.

iii) The r-th autocorrelation function is equal to:

$$\rho_h^{(r)} = \begin{cases} \frac{exp(\frac{r^2}{4}\gamma_\alpha(h)) - 1}{K_r exp(\frac{r^2}{4}\gamma_\alpha(0)) - 1} & \text{if } r \text{ is even} \\ \\ exp(\frac{r^2}{4}(\gamma_\alpha(h) - \gamma(0))) & \text{if } r \text{ is odd} \end{cases} \tag{9}$$

where $\gamma_\alpha(.)$ is the autocovariance function of the stationary process $(\alpha_t)_t$. Hence, the process $(Y_t^r)_y$ is both stationary and strict stationary.

Proof: See the Appendix.

This proposition ensures that the process $(Y_t)_t$ defined by (1)- (2) and the powers of this process are stationary. We are going to use this result in the next section.

3 Asymptotic behavior of the maxima of the process $(X_t)_t$.

In this section we study the asymptotic behavior of the distribution of a class of transformations of the process $(Y_t)_t$. First of all, we define $\forall t$,

$$X_t = \ln(\frac{Y_t}{\sigma})^2 = \alpha_t + \ln \varepsilon_t^2. \tag{10}$$

Now we put:

$$\zeta_t = \ln \varepsilon_t^2 \text{ for all } t \in \mathbb{Z}, \tag{11}$$

thus the process $(X_t)_t$ becomes a sum of two independent processes :

$$X_t = \alpha_t + \zeta_t. \tag{12}$$

The model (12) can be considered as a regression model with log-GED noise (ζ_t). The asymptotic behavior of the maxima for a regression model with particular noises has been studied by Diop and Guégan (2000). We refer also to Horowitz (1980), Ballerini and Mc Cormick (1989) and Niu (1997) for regression models involving non stationarities. When $(\varepsilon_t)_t$ is characterized by a GED distribution, we denote F the distribution function of $(X_t)_t$. First of all, we study the asymptotic behavior of the distribution F.

Proposition 2 *Let $(X_t)_t$ be the process defined by (12), assume that the distribution of $(\varepsilon_t)_t$ is (3), then the asymptotic behavior of F is :*

$$\begin{aligned}\bar{F}(x) &= P\{X_t > x\} \qquad (13)\\ &\sim \frac{\sigma_\alpha}{\sqrt{\pi}} exp\Big\{ - \frac{x^2}{2\sigma_\alpha^2} + \frac{axlng(x)}{\sigma_\alpha^2} - \frac{2x(1+R)}{\gamma\sigma_\alpha^2} + Alng(x) - \frac{bxlng(x)}{\sigma_\alpha^2 g(x)} \\ &\quad - \frac{2ln^2g(x)}{\sigma_\alpha^2\gamma^2} - \frac{cx}{\sigma_\alpha^2 g(x)} + \frac{1}{8\sigma_\alpha^2}(\frac{1}{2}\gamma - 1)^2) + \frac{1}{\sigma_\alpha^2}(\frac{1}{2} - \frac{1}{\gamma}) + C + \mathcal{O}(\frac{ln^2 g(x)}{g(x)})\Big\}\end{aligned}$$

where

$$a = \frac{2}{\gamma}, \qquad b = (\frac{2}{\gamma})^3, \qquad c = -(\frac{2}{\gamma})^3 R - \frac{2}{\gamma^2}\sigma_\alpha^2 + \frac{\sigma_\alpha^2}{\gamma},$$

$$R = lnk\sigma_\alpha^2, \qquad A = \frac{4}{\gamma^2\sigma_\alpha^2}(R+1) + \frac{1}{\gamma} - \frac{3}{2},$$

$$C = -\frac{2}{\gamma^2\sigma_\alpha^2}R^2 + (\frac{1}{2} - \frac{1}{\gamma} - \frac{4}{\gamma^2\sigma_\alpha^2})R - \frac{1}{\gamma} + \frac{1}{2} + ln\frac{4\Gamma(\frac{1}{2})c_0}{\sqrt{k}\gamma^2},$$

and

$$g(x) = \frac{2x+1}{\gamma} - 1/2.$$

Proof : See the Appendix.

Remark : When the noise $(\varepsilon_t)_t$ follows a Gaussian distribution, then $c_0 = \frac{1}{\sqrt{2\pi}}$, $\gamma = 2$, $k = \frac{1}{2}$ and we find the results of Breidt and Davis (1998, p.667).

We define now $M_n = \max(X_1, X_2, \ldots, X_n)$, $n \geq 2$, the maxima of the process $(X_t)_t$. Here, we investigate the asymptotic distribution of M_n.

Theorem 1 *Let $(X_t)_t$ be a process defined in (12). Assume that the density of $(\varepsilon_t)_t$ is given by (3). Assume also that $\rho_\alpha(h)\ln h \longrightarrow 0$, as $h \longrightarrow +\infty$, where $\rho_\alpha(h)$ denotes the autocorrelation function of the process $(\alpha_t)_t$ defined in (2) and $\gamma \geq \frac{1}{2}$. Then there exist normalizing constants $(a_n > 0)$ and (b_n) such that*

$$P[a_n(M_n - b_n) \leq x] \longrightarrow exp(-e^{-x}), \tag{14}$$

where

$$a_n = \sigma_\alpha(2\ln n)^{\frac{1}{2}}, \qquad d_n = (\ln n)^{\frac{1}{2}}$$

and

$$b_n = c_1 d_n + c_2 \ln d_n + c_3 + c_4\frac{\ln d_n}{d_n} + \frac{c_5}{d_n},$$

with

$$c_1 = (2\sigma_\alpha^2)^{\frac{1}{2}}, \qquad c_2 = a, \qquad c_3 = a[\frac{3}{2}\ln 2 + \ln\frac{2}{\gamma} - \frac{\ln\sigma_\alpha^2}{2} - 1],$$

$$c_4 = (\frac{1}{\gamma} - \frac{3}{2})\frac{\sigma_\alpha}{\sqrt{2}},$$

$$c_5 = \frac{-1}{2(2\sigma_\alpha^2)^{\frac{1}{2}}}\Big\{a^2 + (1 - R + \frac{1}{2}ln2\sigma_\alpha^2)\sigma_\alpha^2(1-a) + (\frac{\gamma^2}{4}(\frac{1}{2} - \frac{1}{\gamma}) + 2 + 2\sigma_\alpha^2)(\frac{1}{\gamma} - \frac{1}{2})$$
$$+(3-a)\sigma_\alpha^2 lna - 2\sigma_\alpha^2\ln(\frac{a^2c_0\sqrt{\pi}}{\sqrt{k}}) + \sigma_\alpha^2\ln(2\pi)\Big\}.$$

Proof : See the Appendix.

Remark : We note that the coefficient a_n in (3.12) of Breidt and Davis (1998) obtained with a Gaussian law for $(\varepsilon_t)_t$ seems to contain a computational error which is also highlighted by simulation results in the next section.

4 Behavior of the distribution of the maxima of $(X_t)_t$ using finite samples.

In this section we will study the behavior of the distribution of the maxima of $(X_t)_t$ given by (11)-(12) for different versions of the process $(\alpha_t)_t$. We precise now these versions. First we consider the iid stochastic volatility model with a GED noise $(\varepsilon_t)_t$: the process $(\alpha_t)_t$ is defined by :

$$\alpha_t = Z_t, \qquad \{Z_t\} \text{ is an iid } \mathcal{N}(0,\ \sigma_\alpha^2). \tag{15}$$

Secondly we study the first-order autoregressive stochastic volatility model (ARSV) with a GED noise $(\varepsilon_t)_t$: the process $(\alpha_t)_t$ is defined by :

$$\alpha_t = \phi\alpha_{t-1} + Z_t, \qquad \{Z_t\}\ iid\ \mathcal{N}(0,\ (1-\phi^2)\sigma_\alpha^2), \tag{16}$$

with $|\phi| < 1$. Finally, we consider the long memory stochastic volatility model (LMSV), (see Breidt *et al.* (1998)) with a GED noise $(\varepsilon_t)_t$ and the process $(\alpha_t)_t$ is defined by :

$$(1-B)^d\alpha_t = Z_t, \qquad \{Z_t\}\ iid\ \mathcal{N}(0,\ \frac{\sigma_\alpha^2\Gamma^2(1-d)}{\Gamma(1-2d)}), \tag{17}$$

with $|d| < 1/2$. In the following, we highlight the influence of the parameters on the limiting distribution for these different models. We study also the importance of the different assumptions made in the theorem 1.

4.1 Assumptions

The result stated in the theorem 1 requires the condition $\rho_\alpha(h)\ln h \to 0$ as $h \to \infty$. We illustrate this convergence for finite samples. The model (15) corresponds to an iid process, thus $\rho_\alpha(h) = 0,\ h \neq 0$. The model (16) is an AR(1) process, thus $\rho_\alpha(h)$ decreases with the rate of $\phi^{|h|}$. The model (17) is a FARIMA(*0* , *d*, *0*) process: it is known that $\rho_\alpha(h) \sim C(d)h^{2d-1}$, as $h \to \infty$, where $C(d)$ is some constant which depends on d, thus the speed of convergence of $\rho_\alpha(h)\ln h$ towards 0 is very slow.

Now we investigate empirically the behavior of $\rho_\alpha(h)\ln h$, as $h \to \infty$ for the model (16) and (17). In Figure 1, the graphes (a), (b), (c), (d) correspond to the model (16) for different values of the parameter ϕ ($\phi = 0.2,\ 0.5,\ 0.95,\ 0.99$). The graphes (e), (f), (g), (h) correspond to the model (17) for different values of the parameter d ($d = 0.1,\ 0.2,\ 0.3,\ 0.4$). For the model (16), when we are far from the non stationary case ($\phi = 0.2,\ 0.5$), the decreasing of $\rho_\alpha(h)\ln h$ to zero is very fast. Now when d is small ($d = 0.1,\ 0.2$) the decreasing of $\rho_\alpha(h)\ln h$ is not so fast than the one obtained with the AR(1) model. But this decreasing becomes quicker as soon as $d = 0.3$. For $d = 0.4$, which corresponds to a strong long memory behavior, $\rho_\alpha(h)\ln h$ does not decrease to zero for finite samples ($n = 1000$).

4.2 Influence of the parameters

In view to compare our results with those of Breidt and Davis (1998), we use in our simulations the same value as them for σ^2_α ($\sigma^2_\alpha = 2.3976$, 0.6933, 0.0953), for ϕ ($\phi = 0.95$) and for d ($d = 0.4$) .

In Figures 2-6, we give in solid lines the empirical distribution functions for 1000 normalized maxima $[a_n(M_n - b_n)]$ obtained from 1000 replications of samples of size 1000 and in dotted lines the limiting double exponential distribution. The quality of the convergence depends on σ_α and a_n (the coefficient a_n is given by $a_n = \sigma_\alpha(2\ln n)^{\frac{1}{2}}$). The influence of the tail-thickness parameter γ of the GED distribution is also studied in Figures 2-4. We present the limiting distribution when $\gamma = 2$ (Gaussian case), $\gamma = 1$ (Laplace distribution which is thicker tail than the Gaussian distribution) and $\gamma = 3$ (thinner tail than the Gaussian distribution). The asymptotic behavior of the normalized maxima is closely related to the value of σ_α as shown in Figures 2-4. We also study in Figures 5-6 the sensitivity of the convergence obtained in (14), with respect to the parameters ϕ and d respectively when $(\alpha_t)_t$ follows the models (16)-(17). We detail now these results.

4.2.1 Gaussian noise ($\gamma = 2$)

The behavior of the asymptotic distribution of $a_n(M_n - b_n)$ when we use a Gaussian driven noise $(\varepsilon_t)_t$ is given in Figure 2. In each row, σ^2_α is constant ($\sigma^2_\alpha = 2.3976$, 0.6933, 0.0953). First column corresponds to the iid stochastic volatility model with $(\alpha_t)_t$ defined in (15), second column to the ARSV model with $(\alpha_t)_t$ defined in (16), third column to the LMSV model with $(\alpha_t)_t$ defined in (17). When $\sigma^2_\alpha = 2.3976$, the approximation to the empirical distribution function of $a_n(M_n - b_n)$, $n = 1000$ by the double exponential distribution is very poor whatever the dependence structure taken for $(\alpha_t)_t$. For finite samples, the approximation gives better results when $\sigma^2_\alpha = 0.6933$ and $\sigma^2_\alpha = 0.0953$ as shown in the lower panels of Figure 2.

4.2.2 GED noise with $\gamma = 3$

We turn now to the case of the stochastic volatility model (12) driven with a noise $(\varepsilon_t)_t$ characterized by tail-thickness parameter $\gamma = 3$, see Figure 3: the solid lines represent the empirical distribution of the normalized maxima $a_n(M_n - b_n)$, with $n = 1000$ and the dotted lines represent the double exponential distribution. Looking at the nine graphes of the figure 3, we see that for finite samples the approximation of the empirical distribution with the double exponential distribution is better when $(\alpha_t)_t$ follows the model (17) with $d = 0.4$ and $\sigma^2_\alpha = 0.6933$. This reveals that the condition $\rho_\alpha(h)\ln h \to 0$ as $h \to \infty$ stated in theorem 1 is not so important in finite distance, indeed we have a good approximation between our distributions, although in that last case, the condition $\rho_\alpha(h) \ln h \to 0$ as $h \to \infty$, fails.

4.2.3 Laplace noise $\gamma = 1$

We now abstract from the thinner tailed case and consider a stochastic volatility model driven by Laplace noise $\gamma = 1$. The result is given in Figure 4. By reading across the rows of this picture, the panels again show the influence of the value of the parameter σ_α^2. More σ_α^2 is small, better is the rate of convergence of the normalized maxima of $(X_t)_t$ defined in (12) to the double exponential distribution. The dependence structure in $(\alpha_t)_t$ seems not to have a great influence on this convergence. A comparaison between Figure 2 and Figure 4 shows clearly that the convergence of the normalized maxima is better when the driven noise $(\varepsilon_t)_t$ is Gaussian. Despite the change of the scale in the x-axis of Figure 4, the comparaison between Figure 3 and Figure 4 corresponding respectively to $\gamma = 3$ and $\gamma = 1$ reveals a better approximation when $\gamma = 3$.

4.2.4 Influence of the autoregressive parameter ϕ

Here we assume for simplicity that the noise $(\varepsilon_t)_t$ follows a Gaussian law and we investigate the influence of the autoregressive parameter ϕ to the asymptotic behavior of $a_n(M_n - b_n)$ for the process $(X_t)_t$ defined in (12) when the process $(\alpha_t)_t$ follows the model (16). To determine the sensitivity of the results with respect to this parameter, different alternative values for ϕ have been used : $\phi = 0.2$, $\phi = 0.95$ and $\phi = 0.99$. The results are given in Figure 5. By looking across the rows of Figure 5, we note that more ϕ is small, better is the rate of convergence. When the autoregressive parameter ϕ is equal to 0.99, the approximation is dramatically poor whatever the value of σ_α as shown in the last column of Figure 5. A possible explanation is the fact that the underlying process tends to become nonstationary when the parameter ϕ is close to 1. It seems that this situation needs other investigations. This case has been already partially studied, see for instance, Horowitz (1980), Ballerini and Mc Cormick (1989) and Niu (1997).

4.2.5 Importance of the long memory parameter d.

Figure 6 shows the influence of the long memory parameter d to the convergence of the normalized maxima of the process $(X_t)_t$ defined in (12). For sake of concisness, we assume that the noise $(\varepsilon_t)_t$ follows a Gaussian distribution. We compare the empirical distribution function of $a_n(M_n - b_n)$ when the process $(X_t)_t$ follows the LMSV model and the double exponential distribution when $d = 0.1$, $d = 0.2$ and $d = 0.4$. It seems that the values of the parameter d have small influence on the convergence (14). With finite samples, when $d = 0.4$ the condition $\rho_\alpha(h)\log(h) \to 0$ fails as shown in Figure 1. However we have approximately the same behavior as in the other cases when σ_α is equal to 0.0953. We note that when σ_α is small (so does a_n), the asymptotic behavior of the normalized maximum is the same whatever the values of d.

4.2.6 Tail comparaison

We give here a brief comparison between the tail behavior of the two components of the process $(X_t)_t$ defined in (12). The density of the r. v. ζ_t defined in (11) when the r.v. ε_t follows a GED distribution with shape parameter γ is given by

$$h(x) = c_0 exp(\frac{x}{2} - ke^{\frac{\gamma}{2}x}).$$

Figure 7 shows the tails of the distribution of the Gaussian linear process $(\alpha_t)_t$ defined in (2) and the noise $(\zeta_t)_t$ defined in (11). The three panels correspond respectively to $\sigma_\alpha^2 = 2.3976$, $\sigma_\alpha^2 = 0.6933$ and $\sigma_\alpha^2 = 0.0953$. When $\sigma_\alpha^2 = 2.3976$, the Gaussian upper tail dominates the tail of ζ_t for $\gamma = 0.5,\ 1,\ 2,\ 3$. When $\sigma_\alpha^2 = 0.0953$, the Gaussian upper tail is dominated by the upper tail of ζ_t. Finally when $\sigma_\alpha^2 = 0.6933$, all the tails are tendency to have the same behavior as the Gaussian upper tail. It is evident by looking accross the panels of Figure 7 that the lower tail of ζ_t dominate the Gaussian tail whatever the value of σ_α^2.

5 Conclusion

In this paper, we get the asymptotic distribution of the maxima of a log-transformation of a process $(X_t)_t$ driven by a stochastic volatility model and we study the empirical behavior of its asymptotic distribution . We point out the influence of the tail-thickness parameter γ of the driven noise $(\varepsilon_t)_t$. Our findings are that the choice of γ, σ_α^2 and the dependence structure of $(\alpha_t)_t$ does affect the goodness-of-fit of the limiting distribution.

In practice, the results stated in theorem 1 allow us to calculate quantile risk measures using the maximum block method, see for instance Embrecht *et al.* (1997). However in finite samples care must be taken because of the poorness of the approximation. This is one of the limitations to evaluate for instance the Value at Risk for real data using this part of the extreme value theory. It would be interesting to investigated the asymptotic behavior for more complicated nonstationary models, this will be done in a companion paper.

6 Appendix

Proof of Proposition 1 : We can establish using Devroye (1986) p. 175 that the GED noise $(\varepsilon_t)_t$ defined in (3) satisfies

$$\varepsilon_t \overset{d}{=} k^{-\frac{1}{\gamma}} V G^{\frac{1}{\gamma}} \tag{18}$$

where $\overset{d}{=}$ denote the equality in distribution, V a uniform r.v. on $[-1,\ 1]$ and G a r.v. following a Gamma distribution $\Gamma(1+\frac{1}{\gamma},\ 1)$, independent of V. The identity (18) is used in Section 4 to draw samples from a GED distribution. From (18), it can be shown that

$$E(\varepsilon_t^r) = \begin{cases} \frac{k^{-\frac{r}{\gamma}}}{r+1} \frac{\Gamma(\frac{r+1}{\gamma}+1)}{\Gamma(\frac{1}{\gamma}+1)}, & \text{if } r \text{ is even} \\ \\ 0 & \text{if } r \text{ is odd.} \end{cases} \tag{19}$$

i) Since $(\varepsilon_t)_t$ has a symmetric distribution and $(\alpha_t)_t$ defined in (2) is a Gaussian linear process independent of $(\varepsilon_t)_t$, (5) and (6) are obtained directly from the properties of the lognormal distribution.
ii) The excess kurtosis of $(Y_t)_t$ defined by $\frac{E[Y_t^4]}{E[Y_t^2]^2} - 3$ follows directly from (5).
iii) The r-th autocorrelation function is defined by

$$\rho_h^{(r)} = \frac{E[Y_t^r Y_{t+h}^r] - E[Y_t^r]^2}{E[Y_t^{2r}] - E[Y_t^r]^2}. \tag{20}$$

From (1) and using again the properties of lognormal distribution, we have for all r :

$$E(Y_t^r Y_{t+h}^r) = \sigma^{2r} exp(\frac{r^2}{4}(\sigma_\alpha^2 + \gamma_\alpha(h)))E(\varepsilon_t^{2r}). \tag{21}$$

The remainder of the proof follows directly from (5) and (21).

Proof of Proposition 2: We avoid here to go through the details of the computations which need some great efforts. We focus only on the important steps. Breidt and Davis (1998) use a Tauberian argument in the Gaussian case to express the asymptotic approximation to the tail distribution of $(X_t)_t$, we use it again. In step 1, we give some expansions of the two derivatives $m(\lambda)$ and $S(\lambda)$ of the log-moment generating function of $(X_t)_t$ defined in (12). In the second step, we define an inverse function $m^{-1}(.)$ of $m(\lambda)$ and we justify the inverse notation. Some useful expansions of functions in terms of $m^{-1}(x)$ are also provided. In the third step, we establish the asymptotic normality for normalized transform Esscher of F

which permits us to prove (13).

Step 1 :
We give here the expansions of the two first derivatives of the log-moment generating of the process $(X_t)_t$ defined in (12).
The log of the moment generating function of $(\zeta_t)_t$ is given by :

$$\ln E\exp(\lambda\zeta_t) = \ln 2c_0 - \ln\gamma - \frac{2\lambda+1}{\gamma}\ln k + \ln\Gamma(\frac{2\lambda+1}{\gamma}).$$

Thus, the log of the moment-generating function of $(X_t)_t$ is equal to

$$\begin{aligned} \ln C(\lambda) &= \frac{\lambda^2\sigma_\alpha^2}{2} + \ln 2c_0 - \ln\gamma - \frac{2\lambda+1}{\gamma}\ln k + \ln\Gamma(\frac{2\lambda+1}{\gamma}) \quad (22)\\ &= \frac{\lambda^2\sigma_\alpha^2}{2} - g(\lambda)\ln k - g(\lambda) + g(\lambda)\ln g(\lambda) + \ln\frac{2\sqrt{2}\Gamma(\frac{1}{2})c_0}{\sqrt{k}\gamma} + \mathcal{O}(\frac{1}{\lambda}). \end{aligned}$$

where

$$g(\lambda) = \frac{2\lambda+1}{\gamma} - \frac{1}{2}.$$

The first two derivatives of $\ln C(\lambda)$ are

$$\begin{aligned} m(\lambda) &= \frac{d}{d\lambda}\ln C(\lambda) = \lambda\sigma_\alpha^2 - \frac{2}{\gamma}\ln k + \frac{2}{\gamma}\Psi(\frac{2\lambda+1}{\gamma})\\ &= \lambda\sigma_\alpha^2 - \frac{2}{\gamma}\ln k + \frac{2}{\gamma}\ln g(\lambda) + \mathcal{O}(\frac{1}{\lambda^2}) \quad (23) \end{aligned}$$

and

$$\begin{aligned} S^2(\lambda) = \frac{d^2}{d\lambda^2}\ln C(\lambda) &= \sigma_\alpha^2 + (\frac{2}{\gamma})^2\Psi'(\frac{2\lambda+1}{\gamma})\\ &= \sigma_\alpha^2 + \mathcal{O}(\frac{1}{\lambda}), \quad (24) \end{aligned}$$

where $\Psi(.)$ and $\Psi'(.)$ are the digamma and trigamma functions respectively.

Step 2 :
We set

$$m^{-1}(x) = \frac{\gamma g(x)}{2\sigma_\alpha^2} - \frac{a\ln g(x)}{\sigma_\alpha^2} + \frac{b\ln g(x)}{\sigma_\alpha^2 g(x)} + \frac{c}{\sigma_\alpha^2 g(x)} - \frac{d}{\sigma_\alpha^2}, \quad (25)$$

where

$$d = -\frac{2}{\gamma}\ln(k\sigma_\alpha^2).$$

It follows that

$$\ln g(m^{-1}(x)) = \ln(\frac{g(x)}{\sigma_\alpha^2}) - \frac{2a\ln g(x)}{g(x)} + \frac{D_1}{\gamma g(x)} + \mathcal{O}(\frac{\ln^2 g(x)}{g(x)^2}), \quad (26)$$

where

$$D_1 = -2d + \sigma_\alpha^2 - \frac{\gamma \sigma_\alpha^2}{2}.$$

After some computations, using (23), (25) and (26), we get

$$m(m^{-1}(x)) = x + \mathcal{O}(\frac{\ln^2 g(x)}{g(x)^2}) \tag{27}$$

which justifies the inverse notation.

Step 3 :
We establish the asymptotic normality for the normalized transform Esscher of the distribution F (see Feigin and Yashchin (1983)). We write

$$\bar{F}(m(\lambda)) \sim \frac{exp(-\lambda m(\lambda))C(\lambda)}{(2\pi)^{\frac{1}{2}}\lambda S(\lambda)}. \tag{28}$$

Using the inverse of m defined in (25), we have

$$\bar{F}(x) \sim \frac{exp(-x m^{-1}(x))C(m^{-1}(x))}{(2\pi)^{\frac{1}{2}} m^{-1}(x) S(m^{-1}(x))}. \tag{29}$$

The remainder of the proof needs the following expansions.

$$\begin{aligned}\frac{\sigma_\alpha^2}{2}(m^{-1}(x))^2 = \frac{\gamma^2 g(x)^2}{8\sigma_\alpha^2} &+ \frac{a^2 \ln^2 g(x)}{2\sigma_\alpha^2} - \frac{\gamma a g(x) \ln g(x)}{2\sigma_\alpha^2} + \frac{(\gamma b + 2ad)\ln g(x)}{2\sigma_\alpha^2} \\ &- \frac{\gamma d g(x)}{2\sigma_\alpha^2} + \frac{d^2 + \gamma c}{2\sigma_\alpha^2} + \mathcal{O}(\frac{\ln^2 g(x)}{g(x)^2}),\end{aligned}$$

and

$$\begin{aligned}g(m^{-1}(x)) \ln g(m^{-1}(x)) &= -\frac{\ln(\sigma_\alpha^2)}{\sigma_\alpha^2} g(x) + D_2 \ln g(x) + \frac{g(x)\ln(g(x))}{\sigma_\alpha^2} - \frac{2a \ln^2 g(x)}{\gamma \sigma_\alpha^2} \\ &+ \frac{D_1}{\gamma \sigma_\alpha^2}(1 - \ln \sigma_\alpha^2) + \mathcal{O}(\frac{\ln^2 g(x)}{g(x)}),\end{aligned}$$

where

$$D_2 = \frac{-ad + a^2 \ln \sigma_\alpha^2 - 2a}{\sigma_\alpha^2}.$$

We conclude the proof by using (24), (25), (27), (29) and these expansions.

Proof of theorem1: Using the proposition 2, we get

$$F^n(u_n) \longrightarrow \exp(-e^{-x}), \qquad x \in \mathbb{R} \tag{30}$$

where $u_n = a_n^{-1} x + b_n$. Now, we need to show that

$$|P(M_n \leq u_n) - F^n(u_n)| \longrightarrow 0 \qquad \text{as} \qquad n \longrightarrow +\infty. \tag{31}$$

The proof of (31) uses the Normal comparaison lemma (Leadbetter *et al.* (1983, page 81)) and follows the great lines of Breidt and Davis (1998). First, we establish

$$|P(M_n \leq u_n) - F^n(u_n)| \leq nK \sum_{i=1}^{n} |\rho_\alpha(i)| (Eexp\{-\frac{(u_n-\zeta_1)^2}{2\sigma_\alpha^2(1+|\rho_\alpha(i)|)}\})^2 \tag{32}$$

where K is a positive constant, $\rho_\alpha(.)$ the autocorrelation function of the process (α_t) and $\zeta_t = \ln \varepsilon_t^2$. Using the asymptotic relationship for the tail probability of the GED distribution for $\gamma \geq \frac{1}{2}$, given by

$$\bar{F}_\varepsilon(x) \approx \frac{c_0}{\gamma k} x^{1-\gamma} e^{-kx^\gamma} \qquad \text{as} \qquad x \longrightarrow +\infty,$$

then, it can be shown that

$$Eexp\{-\frac{(u_n-\zeta_1)^2}{2\sigma_\alpha^2(1+|\rho_\alpha(i)|)}\} \leq K'(\ln n)^{[\frac{1}{2(1+|\rho_\alpha(i)|)}]} E(\bar{F}_\varepsilon^{\frac{1}{1+|\rho_\alpha(i)|}}(\sigma_\alpha^{-1}(u_n-\zeta_1))) + o(n^{-1}), \tag{33}$$

where [.] stands for the integer part and K' a positive constant. By Jensens's inequality and the relation (30), we have

$$(Eexp\{-\frac{(u_n-\zeta_1)^2}{2\sigma_\alpha^2(1+|\rho_\alpha(i)|)}\})^2 \leq K''(\ln n)^{\frac{1}{1+|\rho_\alpha(i)|}} n^{\frac{-2}{1+|\rho_\alpha(i)|}} \tag{34}$$

where K'' is a positive constant. The remainder of the proof is similar to the one of Breidt and Davis (1998, page 671); see also Leadbetter *et al.* (1983, page 86).

References

Andersen, T. G., Stochastic autoregressive volatility : a framework for volatility modelling, *Mathematical finance*, **4** (2), (1994) 75-102.

Ballerini, R. and Mc Cormick, W.P., Extreme value theory for processes with periodic variances. *Comm. Statist. Stochastic Models*, **5**,(1989), 45-51.

Box, G.E.P. and Tiao, G.C., Bayesian Inference in Statistical Analysis. Reading, MA : Addison-Wesley. (1973).

Breidt, F. J., Crato, N. and De Lima, P., The detection and estimation of long memory in stochastic volatility. *J. Econometrics*, **83**, (1998) 325-348.

Breidt, F.J. and Davis, R.A. Extremes of Stochastic volatility model. *The Annals of Applied Probability* , **8** (3) (1998) 664-675.

Davis, R. A. and Resnick, S. I., Extremes of moving averages of random

variables with finite endpoint. *Ann. Probab.* **19**, (1991) 312-328.
Devroye, L., *Non Uniform Random Variate Generation,* Springer, New York, (1986).
Diop A., Guégan D. *Asymptotoc behavior for the extreme values of a regression model*, Preprint 00-11, University of Reims, France, (2000).
Embrecht P, Kluppelberg C., Mikosch T., *Modelling extremal events for insurance and finance*, Springer, Applications of Mathematics, 33, (1997).
Feigin, P. D. and Yashchin, E., On a strong Tauberian result. *Z. Wahrsch. Verw. Gebiete* **65**, (1983) 35-48.
Ghysels, E., Harvey, A. and Renault, E. *Stochastic volatility in Handbook of Statistics*, Vol. 14, " Statistical Methods in finance", J. Wiley, (1996).
Harvey, A. C. Long memory in stochastic volatility, Discussion paper, L.S.E., London, (1993).
Horowitz, J., Extreme values for a nonstationary stochastic process : an application to air quality analysis. *Technometrics*, **22**, (1980), 469-478.
Leadbetter, M. R. Lindgren, G. and Rootzen, H., *Extremes and Related Properties of Random sequences and Processes*, Springer, New York, (1983).
Niu, X-F. Extreme value theory for a class of nonstationary time series with applications. *J. of Appl. Proba.*, **7**, (1988), 508-522.
Shepard, N. Statistical aspect of ARCH and stochastic volatility, in Time Series Models in Econometrics, Finance and other fields, eds. Cox, Hinkley, Barndorff-Nielsen, Chapman and Hall, Monographes on Statistics and applied probability, **65**, (1993) 2-67.
Subbotin, M.T., On the law of frequency errors. *Matematicheskii Sbornik*, **31**, (1923) , 296-300.
Taylor, S., *Modelling Financial Time series*, Wiley, New York, (1986).

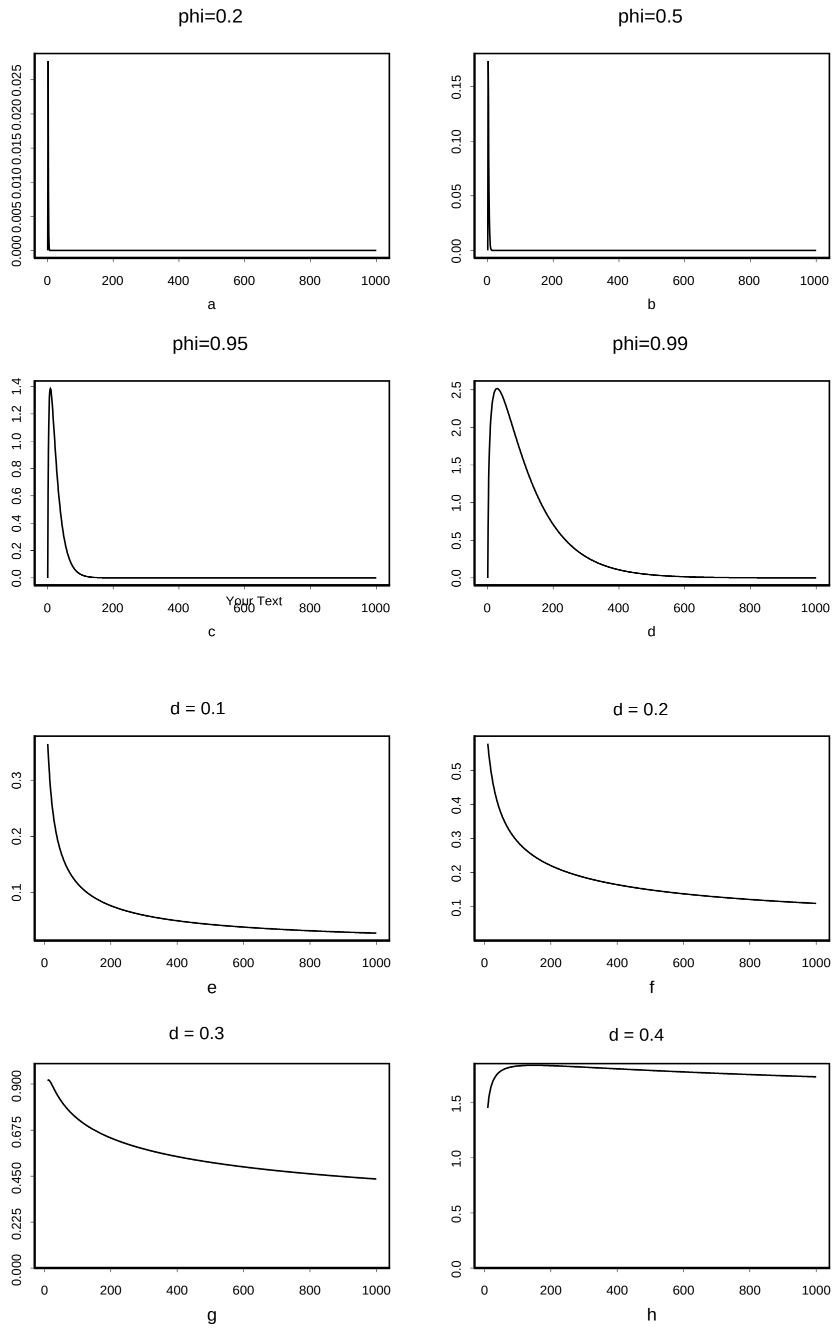

Figure 1: Rate of convergence of $\rho_\alpha(h)\ln h$ to 0 as $h \to \infty$ for different values of the autoregressive parameter ϕ ($\phi = 0.2,\ 0.5,\ 0.95,\ 0.99$) (upper plots : a, b, c, d) and different values for the long memory parameter d ($d = 0.1,\ 0.2,\ 0.3,\ 0.4$) (lower plots : e, f, g, h). The graph (h) shows the failure of the convergence for finite samples in the long memory case with $d = 0.4$.

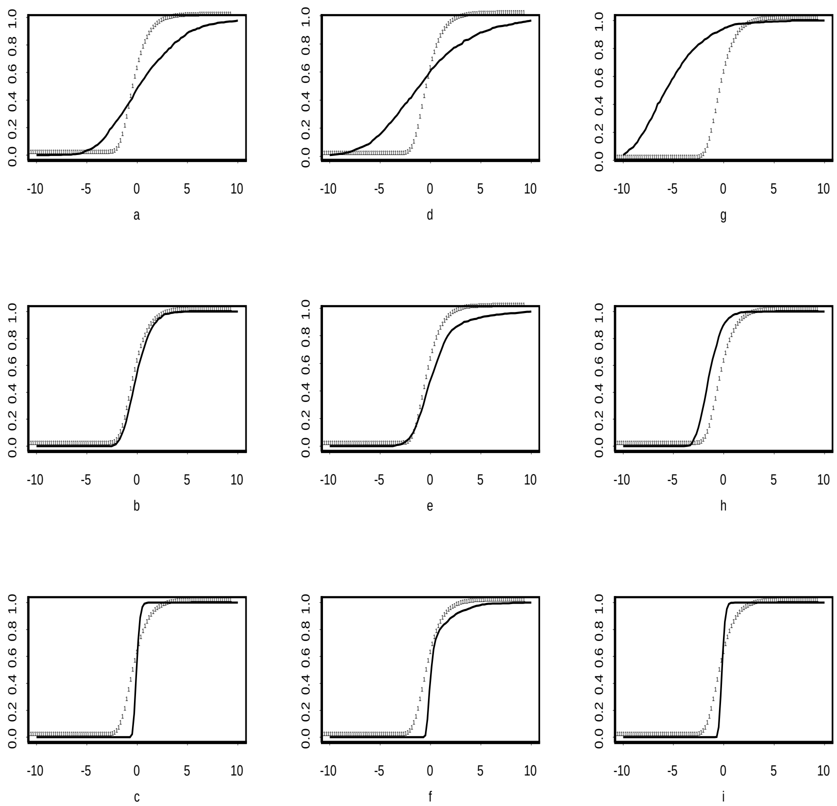

Figure 2: Comparaison between the empirical distribution function (solid lines) for 1000 normalized maxima and the double exponential distribution (dotted lines) for the process $(X_t)_t$ defined in (12) when the driven noise $(\varepsilon_t)_t$ is Gaussian. In first line for the three graphes $\sigma_\alpha^2 = 2.3976$, in the second line $\sigma_\alpha^2 = 0.6933$, in the third line $\sigma_\alpha^2 = 0.0953$. a, b, c : $\alpha_t \sim IID$; d, e, f : $\alpha_t \sim AR(1)$, $\phi = 0.95$; g, h, i : $\alpha_t \sim FARIMA(0, d, 0)$, $d = 0.4$.

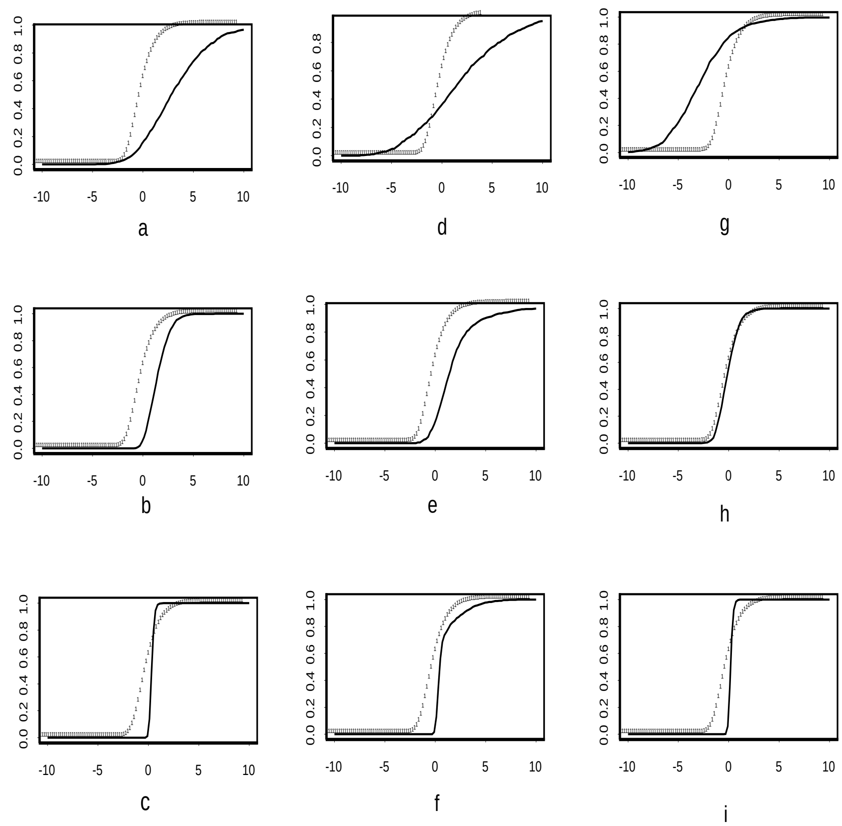

Figure 3: Comparison between the empirical distribution function and the double exponential distribution for the process $(X_t)_t$ defined in (12) with a GED driven noise $(\varepsilon_t)_t$ with $\gamma = 3$. In first line for the three graphes $\sigma_\alpha^2 = 2.3976$, in the second line $\sigma_\alpha^2 = 0.6933$, in the third line $\sigma_\alpha^2 = 0.0953$. a, b, c : $\alpha_t \sim IID$; d, e, f : $\alpha_t \sim AR(1)$, $\phi = 0.95$; g, h, i : $\alpha_t \sim FARIMA(0, d, 0)$, $d = 0.4$.

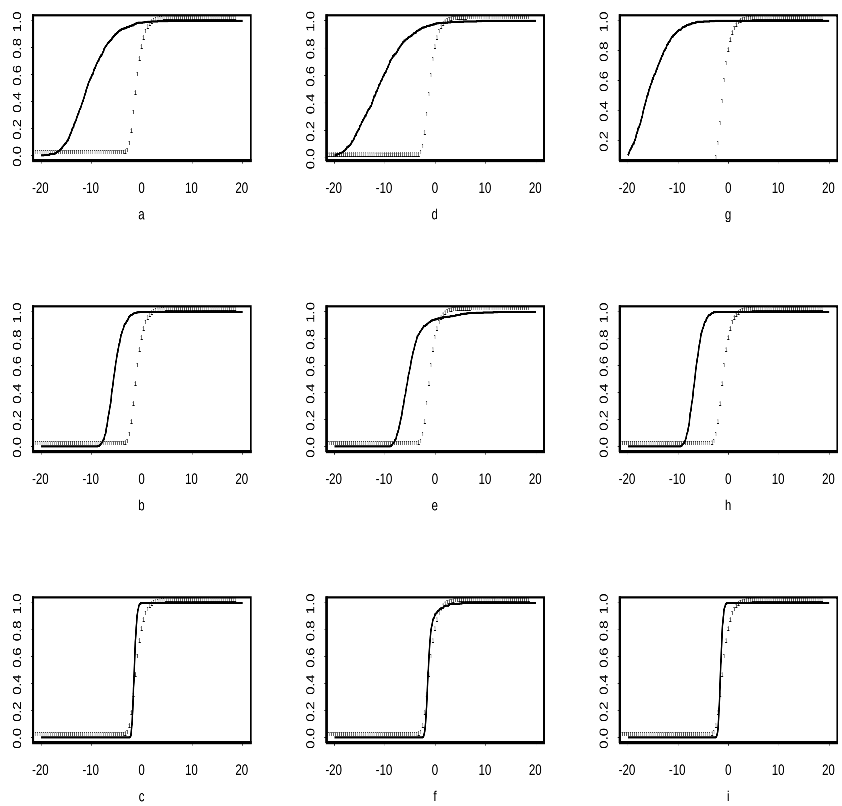

Figure 4: Comparison between the empirical distribution function (solid lines) for 1000 normalized maxima of $(X_t)_t$ defined in (12) and the double exponential distribution (dotted lines) with a GED driven noise $(\varepsilon_t)_t$ with $\gamma = 1$. In first line for the three graphes $\sigma_\alpha^2 = 2.3976$, in the second line $\sigma_\alpha^2 = 0.6933$, in the third line $\sigma_\alpha^2 = 0.0953$. a, b, c : $\alpha_t \sim IID$; d, e, f : $\alpha_t \sim AR(1)$, $\phi = 0.95$; g, h, i : $\alpha_t \sim FARIMA(0, d, 0)$, $d = 0.4$.

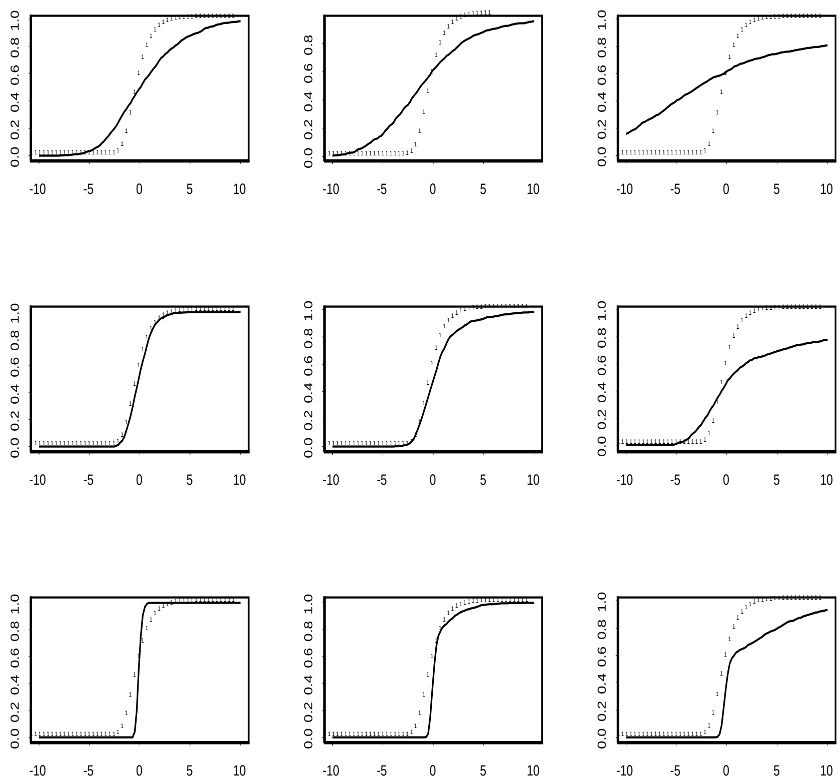

Figure 5: Comparaison between the empirical distribution function (solid lines) for 1000 normalized maxima of $(X_t)_t$ defined in (12) and the double exponential distribution (dotted lines) for ARSV when the driven noise $(\varepsilon_t)_t$ is Gaussian. For each line σ_α^2 is constant, the values of σ_α^2 are the same as in Figure 2. The columns correspond to $\phi = 0.2$, $\phi = 0.95$ and $\phi = 0.99$ respectively.

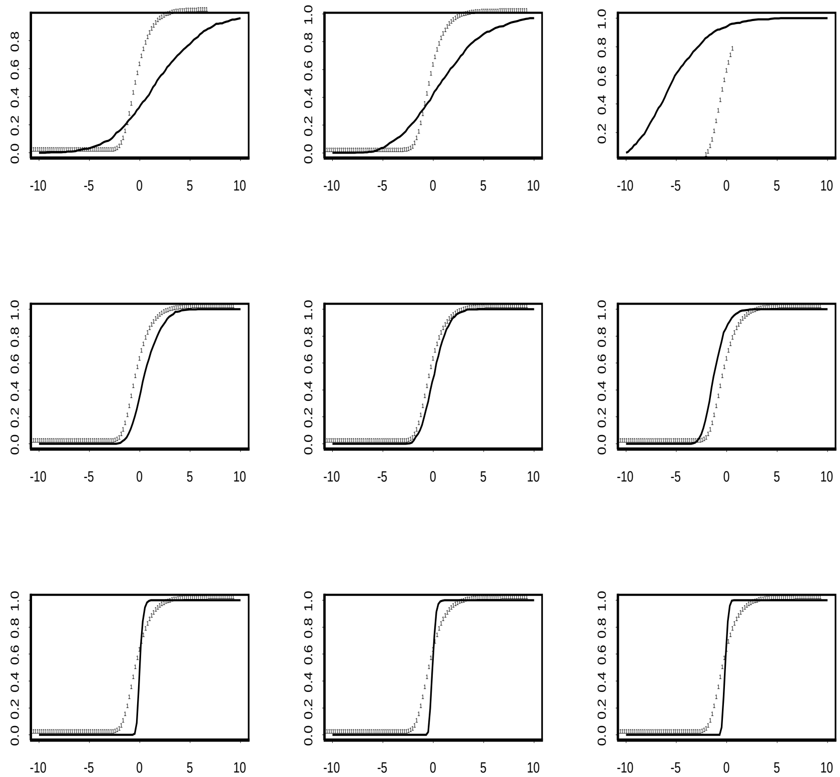

Figure 6: Comparaison between the empirical distribution function (solid lines) for 1000 normalized maxima of $(X_t)_t$ defined in (12) and the double exponential distribution (dotted lines) for LMSV when the driven noise $(\varepsilon_t)_t$ is Gaussian. For each line σ_α^2 is constant, the values of σ_α^2 are the same as in Figure 2. The columns correspond to $d = 0.1$, $d = 0.2$ and $d = 0.4$ respectively.

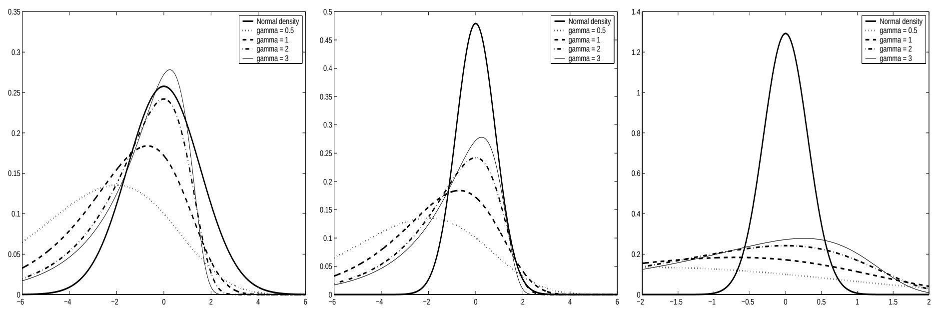

Figure 7: Comparaison between the densities of α_t defined in (2) (solid line) and ζ_t defined in (11) for different values of γ ($\gamma = 0.5,\ 1,\ 2,\ 3$). The three graphes correspond respectively to $\sigma_\alpha^2 = 2.3976$, $\sigma_\alpha^2 = 0.6933$ and $\sigma_\alpha^2 = 0.0953$.