Multi-period Conditional Distribution Functions for Heteroscedastic Models with Applications to VaR

Raymond Brummelhuis [1] **Dominique Guégan** [2] [3]

Abstract

Let $(r_t)_t$ be a general GARCH(1, 1)-type process. We give explicit integral formula's for the probability densities of r_{t+k} conditional to given values of r_t, σ_t, where the latter denotes the variance. As an application, we study the extreme value asymptotics of r_{t+k} and of $r_{t+1} + \cdots + r_{t+k}$. These are relevant for the estimation of Value at Risk and other measures of financial risk in the context of GARCH-models.

Keywords: generalized autoregressive heteroskedastic process, conditional probability density functions, extreme value asymptotics, asymptotics of Laplace integrals, Value at Risk.

1 Introduction

The Value at Risk- or VaR associated with a position taken in the financial market is defined as the maximum expected loss within a chosen confidence and over a chosen time-frame. It's specification therefore requires the following input:

- a time window $[t, t+k]$.
- a confidence level c, typically close to 1.

[1]University of Reims, UPRESA 6056 -. E-mail : raymond.brummelhuis@univ-reims.fr

[2]University of Reims, UPRESA 6056 -. E-mail : dominique.guegan@univ-reims.fr

[3]The authors gratefully acknowledge the support of the Fondation de la Banque de France for this research, through the project "Mesures de risques extrêmes en finance et choix du temps" of the "appel d'offres nr. 4"

- a model for the behavior of (the financial assets making up) one's position over the chosen time frame.

With these data, VaR is simply the $(1-c)$-th lower quantile of the probability distribution of the Profit & Loss (P & L)-function over the period $[t, t+k]$; cf. RiskMetrics [20] and also Jorion [16] or Dowd [9]. Time-frame and confidence level are simple parameters, which are the user's discretion : choices for c of 95% or 99% and of time frames with k equal to 1 to 10 days are current, the last mentioned choices corresponding to the Basle Committee's recommendations. We note here that time will be discrete in this paper, measured in days or in multiples of some other basic unit.

The choice of model is of course crucial. Here, "model" is to be understood in a wide sense: ranging from straightforward Historical Simulation to (more or less sophisticated) parametric models. The disadvantages of HS are well-known: unreliable small quantile estimation due to lack of sufficient data for extreme events, the assumption that the future is almost exactly as the past and too much weight given to distant (in time) events [4]. Parametric models of course come with their own dangers, the most obvious one of which is trying to make the data fit a into straightjacket unsuitable for them. For example, a lot of VaR methodologies are still based on (conditionally) normally distributed returns, although the deviation of stock return data from the normal distribution, in particular as concerns the outliers, is by now well known and well documented: cf. Embrechts, Mikosch and Klüppelberg [10], Frey and McNeil [13], Mikosch and Starič̆a [19] and also Dowd [9] and it's references. Still, parametric models present the advantage of providing easily evaluated and more reliable extreme quantile-estimates, if the model's fit is right.

It is clear that VaR-estimates can vary hugely from one model to another, and that model risk is an important issue here. It is therefore important to thoroughly understand the quantitative and general mathematical implications of working with a certain stochastic model. Related to this, one of the issues with which we will be concerned in this paper is another kind of model risk, different from the one mentioned just now, which consists of naively applying rules of thumb valid for one

[4] The HS-method makes the hidden assumption that the P& L-process is stationary (which is of course one way of formalising that the "future is like the past") and non-parametrically estimates the *unconditional* probability distribution of the P&L process. Even if this process were in general stationary over periods of a year or so, which we frankly doubt, conditional and unconditional probability distributions may differ hugely, and it is the conditional distributions which are important for day to day risk management, as has been stressed by Frey and McNeil [13]

model to another, completely different one, without any (a priori or a posteriori) justification. This would seem to be too obvious a point to make, but is exactly the kind of thing which can be observed in some of the VaR methodologies, and in particular in the RiskMetrics one; cf. [20]. The RiskMetrics model is a particular example of the GARCH(1, 1)-models which are the main subject of this paper. It takes the form:

$$\begin{aligned} r_{t+1} &= \sigma_{t+1}\varepsilon_{t+1} \\ \sigma_{t+1}^2 &= \lambda\sigma_t^2 + (1-\lambda)r_t^2, \end{aligned} \tag{1}$$

where $r_t = \log(P_t/p_{t-1})$ is the logarithmic one-period return, P_t being the (monetary) value of one's position at time t, λ is a parameter to be estimated and where the ε_n are iid $N(0,1)$-distributed. In this model, the VaR over the period $[t, t+1]$ is easily computed to be

$$VaR_c(1) = \sigma_{t+1} q_{1-c}^N P_t,$$

where q_α^N denotes the lower α-th quantile of the standard normal distribution[5]. Note that it is the *conditional* VaR we are talking about. One next is interested in the k-period VaR. Following RiskMetrics, one easily computes that for any $\nu \geq 1$ the conditional expectation of $r_{t+\nu}^2$ given r_t and σ_t is simply σ_{t+1}^2 again, and that therefore the (conditional) variance of the k-period return is

$$E\left(r_{t+1}^2 + \cdots + r_{t+k}^2 | r_t, \sigma_t\right) = k\sigma_{t+1}^2,$$

the expectation being the one conditional to the (known) values of r_t and σ_t at time t. RiskMetrics then proposes to simply compute the k-period VaR over $[t, t+k]$ as

$$VaR_c(k) = \sqrt{k}VaR_c(1). \tag{2}$$

However, this makes only sense if the k-period return are (close to) normally distributed[6] and, as one of the main results of our paper shows, this is far from being the case, even for a k as small as 2. Therefore, baring accidental numerical coincidences for certain c, one should expect the "real" VaR (as given by the model) to be very different from (2). A phenomenon of this kind has been observed numerically, through Monte-Carlo simulation, by Frey and McNeil [13] (but for a GARCH(1, 1)-model with non-normal innovations ε_t, though). A similar preoccupation with variances can be observed in much of the empirical financial and

[5] we are making the usual approximation $e^r - 1 \simeq r$.

[6] In fact, (2) is the same as the k-period VaR in a simple random walk model for the r_t's, which should be enough to make one suspicious!

econometrical litterature. However, primary attention should be given to the distribution functions themselves and one of the merit's of the VaR (whatever it's adequacy as a risk management tool, cf. Artzner, Delbaen, Eber, and Heath [1]) is that it concentrates the minds on these[7]. Comparing variances strictly speaking only makes sense when all relevant probability distributions belong to a one-parameter family $\sigma^{-1} f(x/\sigma)$ of distributions (one might allow f to vary slightly). The main point of our paper, for risk analysis purposes, is this is certainly not the case for multi-period estimates in the context of a GARCH(1, 1)-models. To give an example, our results imply that for the RiskMetrics model (1) the logarithm of the probability distribution function of r_{t+k}, given $r_t = 1$ and $\sigma_t = 1$ at time t, behaves for large negative x as

$$(2\pi k)^{-1/2}(1-\lambda)^{(k-1)/k} \exp\left(\frac{k-1}{2}(\frac{\lambda}{1-\lambda} + \log 2)\right) \cdot$$

$$\cdot \frac{1}{|x|^{1-1/k}} \exp\left(\frac{k}{2}(1-\lambda)^{-(1-1/k)}|x|^{2/k}\right).$$

We have similar asymptotic estimates for the k-period returns, with the *same* type of x-dependence but different constants. Even for $k = 2$, this is asymptotically very different from the standard normal distribution which one obtains for r_{t+1} with the above values of the parameters. That the distribution function of r_{t+k} for big k's will be very much different from the normal one may be inferred from known results on the stationary distribution of a GARCH(1, 1), going back to Kesten [17]; cf. also Embrechts, Mikosch and Klüppelberg [10] and Mikosch and Starica [19] and their references. What may be unexpected about our results is that this non-normality shows up that quickly. Also, note that the RiskMetrics model is not second order stationary.

Autoregressive Conditionally Heteroskedastic or ARCH processes were introduced by Engle [11] and subsequently generalized by Bollerslev [2] to GARCH or Generalized ARCH processes: these are processes of the form (1), with a σ_{t+1} now more generally being given by

$$\sigma_{t+1} = (a_0 + \sum_{j=1}^{p} a_j r_{t-j}^2 + \sum_{j=1}^{q} b_j \sigma_{t-j}^2)^{1/2}$$

and with the $(\varepsilon_t)_t$ iid but not necessarily normal; this is called a GARCH(p, q). An ARCH(p) corresponds to no b's or, equivalently, to $q = 0$. More general

[7]For a riskmeasure like the expected shortfall the pre-eminece of the distribution function is even larger, making it still more model dependend. This might incidentally be somewhat of a disadvantage.

processes allowing more general functional dependences of σ_t on past values r_{t-j} of the processe and on past variances were introduced by Nelson [18], Guégan and Diebold [6], [7] and many others: we refer to Bollerslev, Engle and Nelson [3] or Guégan [14] for an overview.

GARCH processes are popular in empirical finance because of their capability to model a phenomenum like volatility clustering. One very often restricts oneself, as we will do here, to GARCH(1, 1) models, which have a low number of parameters to be estimated. Theoretical work on GARCH-processes has mainly concentrated on stationarity issues and on the behavior of secondary statistics like various moments. In this paper we concentrate on the distribution functions of the various r_{t+k}, and of the cumulated returns $r_{t+1} \cdots + r_{t+k}$, conditional to given values of r_t and σ_t at time t. We will in particular determine their asymptotic behavior for large values of the argument x, which will for both be shown to be of the type

$$C_k |x|^{-(1-1/k)} e^{-c_k |x|^{2/k}}$$

with explicit expressions for the constants in the case of r_{t+k}. For the cumulated returns we actually only show have asymptotic upper and lower bounds of this form. Note that it is exactly this kind of asymptotics, for fixed k and large $|x|$, which will be relevant for risk management practice, if r_t models a return. Stationarity considerations, which are more closely related with the $k \to \infty$ limit, will play no rôle in this paper.

The paper is organized as follows: first, in section 2, we derive a general representation formula for the probability density function of r_{t+k}, valid for a general GARCH(1, 1) with rather arbitrary dependence of σ_{t+1} on r_t and σ_t. Traditional GARCH(1, 1)'s à la Bollerslev [2] and EGARCH(1, 1) processes as introduced by Nelson [18] will provide illustrations. In section 3 we do the same for the cumulative returns $r_{t+1} + \cdots + r_{t+k}$. The remainder of the paper will then be concerned with the asymptotics of these distribution functions in the case of a classical GARCH(1, 1) with normal innovations[8]. In section 4 we first derive a technical result on asymptotics of Laplace transforms which will be needed in the sequel. In section 5 we will obtain rather precise asymptotics for the distribution functions of the returns r_{t+k} and, in section 6, a somewhat more qualitative result on the cumulative returns. We end the paper with an application to multi-period VaR estimation.

The main results of this paper were announced in the note [4].

[8]Non-normal innovations will be considered in another paper

2 Conditional distributions with a time-lag for GARCH(1, 1) processes

We consider a general $GARCH(1,1)$ process:

$$\left\{ \begin{array}{l} r_{t+1} = \sigma_{t+1}\varepsilon_{t+1} \\ \sigma_{t+1} = \varphi(r_t, \varepsilon_t) \end{array} \right. \tag{3}$$

For us, r_t will typically model a security return taken over a period $[t-1, t]$: if $(P_t)_{t\in N}$ is a discrete time series of financial asset prices, r_t will be the logarithmic return $\log(P_t/P_{t-1})$.

The function $\varphi : R^2 \to R_{\geq 0}$ will be measurable, and will have to be such that (4 below holds. The random shocks $(\varepsilon_t)_t$ will be, by hypothesis, iid, with mean 0 and variance 1. This could be weakened to independence only, with t-dependend distributions, at the cost of complicating the notations. One can also let φ explicitly depend on t. The independence of the ε_t is, however, essential. Since such changes can easily be incorporated afterwards, we will first concentrate on the model (3).

We are interested in the conditional probabilities of r_{t+k} given r_t, σ_t, for any $k \geq 1$ (the case of $k = 1$ being of course trivial). We will suppose that both the ε_t and the random variables $\varphi(u\varepsilon_t, u)$ have a probability distribution function, or pdf, for any fixed $u > 0$. Again, these hypotheses could be weakened, by replacing functions by measures at the appropriate places below. Specifically, if we use the notation $X \sim f$ to mean that a random variable X has pdf f, we will assume that

$$\begin{array}{rcl} \varepsilon_t & \sim & f \\ \varphi(u\varepsilon, u) & \sim & h_u \text{ if } \ \varepsilon \sim f, \ u > 0. \end{array} \tag{4}$$

The function h_u can easily be computed in practice. We want to compute the conditional pdf's

$$p_{t,k}(x; \rho, s) = \text{Prob } \left(r_{t+k} = x | r_t = \rho, \sigma_t = s\right). \tag{5}$$

using a somewhat informal "physicist's" notation. Define two integral operators F and H on $L^1(R)$ and $L^1(0,\infty)$, respectively, by

$$F(u)(x) = \int_0^\infty \frac{1}{s} f\left(\frac{x}{s}\right) u(s)\ ds \tag{6}$$

$$H(u)(s) = \int_0^\infty h_{s'}(s) u(s') ds' \tag{7}$$

It can easily be seen that F and H are positivity-preserving and of norm 1: for example,

$$\begin{aligned} ||F(u)||_1 &\leq \int\int \frac{1}{s} f\left(\frac{x}{s}\right) |u(s)| ds dx \\ &= \int \left(\frac{1}{s} f\left(\frac{x}{s}\right) dx\right) |u(s)| ds \\ &= ||u||_1, \end{aligned}$$

since f is a pdf: $f \geq 0$, $\int f(x)dx = 1$. Similary for H, since

$$\int_0^\infty h_{s'}(s) ds = \text{Prob } \left(h_{s'}(s'\varepsilon, s') \in R_{\geq 0}\right) = 1.$$

Note that F and H can also be defined on the spaces of finite Radon measures on R and $R_{>0}$, respectively, and are also positivity-preserving operators of norm 1 on these. We will occasionally use this observation to let them act on delta measures, when (notationally) convenient.

We can now state the main theorem of this section:

Theorem 2.1 *Let* (r_t) *be defined by (2). Under the above hypotheses on* ε *and the function* φ*, we have that*

$$P\left(r_{t+k} = x | r_t = \rho, \sigma_t = s\right) = F \circ H^{k-1}\left(\delta_{\varphi(\rho,s)}\right) \tag{8}$$

More explicitely, if $k > 1$ *then, using the notation* (5),

$$p_{t,k}(x:\rho,s) = \int_{(R_{\geq 0})^{k-1}} \frac{1}{s_k} f\left(\frac{x}{s_k}\right) h_{s_{k-1}}(s_k) h_{s_{k-2}}(s_{k-1}) \cdots h_{\varphi(\rho,s)}(s_2) ds_2 \cdots ds_k. \tag{9}$$

Proof. If $k = 1$ and $r_t = \rho$ and $\sigma_t = s$ are given, then r_{k+1} clearly has pdf

$$\frac{1}{\varphi(\rho,s)} f\left(\frac{x}{\varphi(\rho,s)}\right)$$

and (8) follows. If $k > 1$ then

$$\begin{aligned} & P\left(r_{t+k} = x | r_t = \rho, \sigma_t = s\right) \\ = \; & P\left(\sigma_{t+k}\varepsilon_{t+k} = x | r_t = \rho, \sigma_t = s\right) \\ = \; & \int_0^\infty P\left(\sigma_{t+k}\varepsilon_{t+k} = x | \sigma_{t+k} = s_k, r_t = \rho, \sigma_t = s\right) \cdot P\left(\sigma_{t+k} = s_t | r_t = \rho, \sigma_t = s\right) ds_k \\ = \; & \int_0^\infty \frac{1}{s_k} f\left(\frac{x}{s_k}\right) \cdot P\left(\sigma_{t+k} = s_k | r_k = \rho, \sigma_t = s\right) ds_k, \end{aligned} \tag{10}$$

since ε_{t+k} is independent of $\sigma_{t+k} = \varphi(r_{t+k-1}, \sigma_{t+k-1})$ in our model. Next, $r_{t+k-1} = \sigma_{t+k-1}\varepsilon_{t+k-1}$, the two factors on the right hand side being independent again. Hence

$$\begin{aligned}
& P\left(\sigma_{t+k} = s_k | r_k = \rho, \sigma_t = s\right) \\
= & \int_0^\infty P\left(\varphi(s_{k-1}\varepsilon_{t+k-1}, s_{k-1}) = s_k | \sigma_{t+k-1} = s_{k-1}, r_t = \rho, \sigma_t = s\right) \\
\cdot & P\left(\sigma_{t+k-1} = s_{t-1} | r_t = \rho, \sigma_t = s\right) ds_{k-1} \\
= & \int_0^\infty P\left(\varphi(s_{k-1}\varepsilon_{t+k-1}, s_{k-1}) = s_k\right) \cdot P\left(\sigma_{t+k-1} = s_{k-1} | r_t = \rho, \sigma_t = s\right) ds_{k-1} \\
= & \int_0^\infty h_{s_{k-1}}(s_k) \cdot P\left(\sigma_{t+k-1} = s_{k-1} | r_t = \rho, \sigma_t = s\right) ds_{k-1}.
\end{aligned}$$

Substituting this expression in (10) and repeating the same analysis for $P\left(\sigma_{t+k-1} = s_{k-1} | r_t = \rho, \sigma_t = s\right)$ we find, after k steps, the formula (9) and therefore the theorem. QED

Examples 2.2 To illustrate the use of 2.1 we look at some examples.

(i) **Classical GARCH(1, 1):** We take

$$\varphi(r, \sigma) = \left(a_0 + a_1 r^2 + b_1 \sigma^2\right)^{1/2}, \tag{11}$$

and ε_t iid, $\varepsilon_t \sim f$. We leave f unspecified, apart form requiring that it has mean 0 and variance 1. We can easily compute the kernel $h_u(s)$ in terms of f: if $\varepsilon \sim f$, then the pdf of $\varphi(u\varepsilon, u)$ is:

$$\frac{d}{ds} P\left(\left(a_0 + a_1 u^2 \varepsilon^2 + b_1 u^2\right)^{1/2} < s\right)$$

which is equal to 0 if $s \leq \sqrt{a_0 + b_1 u^2}$, and equals to

$$\begin{aligned}
& \tfrac{d}{ds}\left(\int_{-(s^2 - a_0 - b_1 u^2/a_1 u^2)^{1/2}}^{(s^2 - a_0 - b_1 u^2/a_1 u^2)^{1/2}} f(y)\ dy\right) = \\
& = \tfrac{1}{2} s \left(a_1 u^2 (s^2 - a_0 - b_1 u^2)\right)^{-1/2} \textstyle\sum_{\pm} f\left(\pm\left(\frac{s^2 - a_0 - b_1 u^2}{a_1 u^2}\right)^{1/2}\right)
\end{aligned}$$

if $s > \sqrt{a_0 + b_1 u^2}$. If f is symmetric, which is often, if not always, the case in applications, this simplifies to:

$$2s\left(a_1 u^2 (s^2 - a_0 - b_1 u^2)\right)^{-1/2} f\left(\left(\frac{s^2 - a_0 - b_1 u^2}{a_1 u^2}\right)^{1/2}\right) \chi_{\{s > \sqrt{a_0 + b_1 u^2}\}}, \tag{12}$$

χ_A being the indicator function of a set A. Popular choices for f are the standard Gaussian distribution

$$f(x) = \frac{1}{2\pi} e^{-x^2/2}$$

and the Student distribution with $n > 2$ degrees of freedom:

$$f(x) = \left(\frac{1}{n-2} \right)^{1/2} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \left(1 + \frac{x^2}{n-2} \right)^{-(n+1)/2}.$$

The latter is nowadays often used to model the empirically observed fat tails of the daily innovations.

These formulae, together with (9), can easily be implemented on a computer, and used for fast numerical computation of the probability densities (5), as an alternative to Monte-Carlo: cf. [5]

(ii) **EGARCH(1, 1):** Nelson's exponential GARCH or EGARCH model is given by:

$$\left(1 + \sum_{j=1}^{p} \beta_j L^j \right) \log \sigma_{t+1}^2 = \omega + \left(1 + \sum_{i=1}^{q} \alpha_i L^i \right) g(\varepsilon_t)$$

with

$$g(\varepsilon) = \gamma \left(|\varepsilon| - E(|\varepsilon|) \right) + \theta\varepsilon$$

and L the usual lag or back-shift operator: see Nelson [18] or for example Bollerslev, Engle and Nelson [3]. If $q = 0$ and $p = 1$, this enters into the present paper's framework, with φ given by

$$\varphi(r, \sigma) = \sigma^\beta e^{(\omega + g(r/\sigma))/2}.$$

It is again a simple matter to compute the kernel $h_{s'}(s)$, and we only record the result: if we put $C = \omega - E(|\varepsilon|) = \omega - \int |y| f(y) dy$, then, assuming $\gamma, \theta > 0$,

$$\begin{aligned} h_{s'}(s) &= \frac{2}{(\theta + \gamma)x} f\left(\frac{2(\log s - \beta \log s') - C}{\theta + \gamma} \right) \chi_{[e^C (s')^\beta, \infty)}(s) \\ &+ \frac{2}{(\theta - \gamma)x} f\left(\frac{2(\log s - \beta \log s') - C}{|\theta - \gamma|} \right) \chi_{[0, e^C (s')^\beta)}(s) \end{aligned}$$

Similar computations can be done for the other asymmetric GARCH models mentioned in [3] and for those studied by Dingh and Granger [8] and by Diebolt and Guégan [6], [7].

(iii) **Moments:** As a further illustration of our formula (8), we show how it can be used to compute the moments, or more generally the conditional expectation of $g(r_{t+k})$ for reasonable g (meaning the integrals below should converge absolutely). In fact,

$$\begin{aligned} & E\left(g(r_{t+k}|r_t = \rho, \sigma_t = s\right) \\ = & \int_{R^k_{\geq 0}} E\left(h(s_k\varepsilon)\right) \left(\Pi_{\nu=2}^k h_{s_{\nu-1}}(s_\nu)\right) \delta_{\varphi(\rho,s)}(s_1) ds_1 \cdots ds_k, \end{aligned}$$

ε being a random variable with pdf f and thus for $g(x) = x^n$ we find for the n-th moment of r_{t+k}:

$$E(r^n_{t+k}|r_t = \rho, \sigma_t = s) = \mu_{n,f} \int_{R^k_{\geq 0}} s_k^n \left(\Pi_{\nu=2}^k h_{s_{\nu-1}}(s_\nu)\right) \delta_{\varphi(\rho,s)}(s_1) ds_1 \cdots ds_k, \quad (13)$$

where $\mu_{n,f} = \int x^n f(x) dx$ is the n-th moment of f. Note that this formula only involves the kernels $h_{s'}(s)$. In the case of a classical GARCH(1, 1) these moments can also be calculated recursively, using the special form of the φ (the odd ones will all be 0). Note, however, that such a procedure won't work for a general GARCH(1, 1) of the form (3).

We end this section by describing the version of formula (9) when both φ and the pdf of ε_t depend explicitly on t. In fact, in that case we simply define the kernel $h^t_u(s)$ by

$$\varphi(u\varepsilon_t, u) \sim h^t_u(\cdot) \quad \text{if} \quad \varepsilon_t \simeq f$$

and replace $h_{s_{j-1}}(s_j)$ in (9) by $h^{t+j-1}_{s_{j-1}}(s_j)$ and $s_k f(x/s_k)$ by $s_k f^{t+k}(x/s_k)$.

Letting φ depend on time might have some relevance for practical modelling purposes, but it is not clear whether one really would want to use models with time-dependent distributions for the innovations ε_t, in view of the numerous identification and estimation problems this would entail.

3 Multiple period returns and prices

The profit & loss function for a multiple period time window $[t, t+k]$ involves the k-period return

$$r_{t+k,t} = \log(P_{t+k}/P_t) = r_{t+1} + \cdots + r_{t+k} \quad (14)$$

rather than r_{t+k}. In this section we will derive an integral formula for $r_{t+k,t}$. We will follow a slightly different approach, by considering the two-component Markov process $Z_t = (r_t, \sigma_t)$. Let P_{z_0} be the probability conditional to $Z_t = z_0 = (\rho_0, s_0)$. Then we can write for the joint pdf as:

$$P_{z_0}\left((Z_{t+1}, \cdots, Z_{t+k}) = (z_1, \cdots, z_k)\right) =$$
$$\Pi_{j=1}^k P_{z_0}\left(Z_{t+j} = z_j | (Z_{t+1}, \cdots, Z_{t+j-1}) = (z_1, \cdots, z_{j-1})\right) =$$
$$\Pi_{j=1}^k P_{z_0}\left(Z_{t+j} = z_j | Z_{t+j-1} = z_{j-1}\right),$$

where in the last line we used the Markov property. It follows that the joint pdf of $(r_{t+1}, \sigma_{t+1}) = ((x_1, s_1), \cdots, (r_{t+k} = x_k, \sigma_{t+k} = s_k))$, conditional to $(r_t, \sigma_t) = (\rho_0, s_0)$ can be written (somewhat formally) as

$$\Pi_{j=1}^k \frac{1}{s_j} f\left(\frac{x_j}{s_j}\right) \delta(s_j - \varphi(x_{j-1}, s_{j-1})), \tag{15}$$

$\delta(s - v)$ being the Dirac delta function and $s_0 = \varphi(\rho_0, s_0)$. The conditional pdf of $r_{t+k,t} = r_{t+1} + \cdots + r_{t+k} = x$ is found by integrating (15) against $\delta(x - (x_1 + \cdots + x_k))$. We can evaluate the $s_1, \cdots, s_k$ - integrals involving the delta -functions and obtain the following result:

Theorem 3.1 *Inductively define functions $\hat{s_j} = \hat{s_j}(x_1, \cdots, x_{j-1})$ by:*

$$\begin{aligned} \hat{s}_1 &= \varphi(\rho_0, s_0) \\ \hat{s}_j &= \varphi(x_{j-1}, \hat{s}_j) \end{aligned}$$

Then

$$P(r_{t+k,t} = x | r_t = \rho_0, \sigma_t = s_0) =$$
$$\int \cdots \int \frac{1}{\hat{s}_k} f\left(\frac{x - (x_1 + \cdots + x_{k-1})}{\hat{s}_k}\right) \Pi_{j=1}^{k-1} \frac{1}{\hat{s}_j} f\left(\frac{x_j}{\hat{s}_j}\right) dx_1 \cdots dx_{k-1}. \tag{16}$$

Remark 3.2 It is possible to rederive theorem 2.1 along these lines, by integrating (15) over everything except x_k and making suitable changes of variables (this is easy if φ is one-to-one). Also note that (16) does not have anymore the nice structure of a k-fold operator product, which will make it's asymptotics harder to analyse in section 6 below.

4 Asymptotics of Laplace Transforms

In this section we prove the following technical result on asymptotics of Laplace integrals which we use in the remaining sections :

Lemma 4.1 Let $\alpha > 0$, $s > 0, c > 0$ and $\beta \in R$. Then we have an asymptotic series

$$\int_0^\infty x^{-\beta} e^{-cx^{-\alpha}} e^{-sx} dx \simeq \left(\frac{s}{\alpha}\right)^{\frac{\beta-(\alpha/2)-1}{\alpha+1}} e^{-(\alpha+1)c^{1/(\alpha+1)}(\frac{s}{\alpha})^{\alpha/\alpha+1}} \sum_{j=0}^\infty C_j s^{-j\alpha/(\alpha+1)}, \tag{17}$$

with $C_0 = \sqrt{2\pi/\alpha(\alpha+1)}c^{(1-2\beta)/2(\alpha+1)}$.

Remark 4.2 We will in the following for simplicitly only keep the main term of the various asymptotics, leaving the full series to the reader, and thus simply write the conclusion of lemma 4.1 as:

$$\int_0^\infty x^{-\beta} e^{-sx-x^{-\alpha}} dx \simeq \sqrt{\frac{2\pi}{\alpha(\alpha+1}} \left(\frac{s}{\alpha}\right)^{\frac{\beta-(\alpha/2)-1}{\alpha+1}} e^{-(\alpha+1)(\frac{s}{\alpha})^{\alpha/\alpha+1)}}. \tag{18}$$

In fact, in practical applications to VaR one will often contend oneself with the principal term of the various asymptotic series we will derive below.

Proof of lemma 4.1: It suffices to prove (17) for $c = 1$, by a simple scaling argument. Note that we cannot directly apply Watson's lemma, since all the derivatives of $x^{-\beta}\exp(-sx - x^{-\alpha})$ are 0 in 0. We will first split the integral in two, as follows:

$$\begin{aligned} \int_0^\infty x^{-\beta} e^{-sx-x^{-\alpha}} dx &= \int_0^{s^{-1/(1+\alpha)}} + \int_{s^{-1/(1+\alpha)}}^\infty \\ &= I + II, \end{aligned} \tag{19}$$

noting that $sx = x^{-\alpha}$ precisely when $x = s^{-1/(\alpha+1)}$, and then analyze the two parts separately, using Laplace's method (or complex stationary phase, of one likes). We start with the second integral, II. Making the change of variables $x = s^{-1/(1+\alpha)}y$, we get

$$II = s^{\frac{\beta-1}{\alpha+1}} \int_1^\infty e^{-s^{\alpha/(\alpha+1)}(y+y^{-\alpha})} y^{-\beta} dy,$$

which, apart from the fore-factor, is a classical Laplace integral of the form

$$\int_1^\infty e^{-\lambda\varphi(y)} a(y) dy.$$

The main contribution to the asymptotics comes from the absolute minimum of the phase function $\varphi(y) = y + y^{-\alpha}$ on $[1, \infty)$ and/or from the boundary point $y = 1$. It is easily seen that $\varphi(y)$ has an absolute minimum on $R_{>0}$ at $y = y_c = \alpha^{1/(\alpha+1)}$. We distinguish three cases:

(i) $\alpha > 1$. In this case, $y_c \in (1, \infty)$ and we get a contribution

$$e^{-\lambda\varphi(y_c)}\left(\left(\frac{2\pi}{\lambda}\right)^{1/2}\frac{a(y_c)}{\varphi''(y_c)^{1/2}}+O\left(\lambda^{-3/2}\right)\right)$$

where the O-term stands for a complete asymptotic series in powers $\lambda^{-(1/2)-j}$. Computing $\varphi(y_c) = \alpha^{1/(\alpha+1)} + \alpha^{-\alpha/(\alpha+1)} = (\alpha+1)\alpha^{-\alpha/(\alpha+1)}$ and $\varphi''(y_c) = \alpha(\alpha+1)/\alpha^{(\alpha+2)/(\alpha+1)}$ and remembering the factor in front and the fact that $\lambda = s^{\alpha/(\alpha+1)}$, we find the following contribution to II:

$$\exp\left((\alpha+1)\left(\frac{s}{\alpha}\right)^{\alpha/(\alpha+1}\right)\left(\sqrt{2\pi/\alpha(\alpha+1)}\left(\frac{s}{\alpha}\right)^{\frac{\beta-(\alpha/2)-1}{\alpha+1}}+\cdots\right), \qquad (20)$$

the dots indicating lower order terms. We have to compare this with the contribution from the boundary point $y_c = 1$, which is

$$e^{-2s^{\alpha/(\alpha+1}}\left((1-\alpha)^{-1}s^{\frac{\beta-\alpha-1}{\alpha+1}}+\cdots\right). \qquad (21)$$

However, these will all be dominated by (20), as follows from the following elementary observation:

For all $\alpha > 0$:

$$(\alpha+1)\alpha^{-\alpha/(\alpha+1)} \leq 2 \qquad (22)$$

with equality iff $\alpha = 1$.

To prove (22) we have to show that

$$\log(\alpha+1) - \left(\frac{\alpha}{\alpha+1}\right)\log\alpha \leq \log 2$$

for $\alpha > 0$. Now the derivative of the left hand side equals $-\frac{\log\alpha}{(\alpha+1)^2}$ which is 0 iff $\alpha = 1$ and which is > 0 (< 0) if $\alpha < 1$ ($\alpha > 1$). Hence the right hand side has an absolute maximum in $\alpha = 1$, which equals $\log 2$. QED

We continue with the proof of lemma 4.1. We consider the two remaining cases for II:

(ii) $\alpha = 1$: The minimum y_c coincides with the boundary point, and we obtain 1/2 times (20).

(iii) $\alpha < 1$: In this case $y_c < 1$ and the asymptotics of II will be given by (21), since only $y = 1$ will contribute.

We will now repeat the analysis for the first integral in (19), I. We make the substitutions $x = s^{-1/(\alpha+1)}u^{-1}$ and find that

$$I = s^{(\beta-1)/(\alpha+1)} \int_1^\infty e^{-s^{\alpha/(\alpha+1)}(u^\alpha + u^{-1})} u^{\beta-2} du.$$

Now the phase function $\varphi(u) = u^\alpha + u^{-1}$ will have an absolute minimum in $u = u_c = \alpha^{-1/(\alpha+1)}$ and we compute, as before, that $\varphi(u_c) = (\alpha+1)\alpha^{-\alpha/(\alpha+1)}$ and that

$$\begin{aligned} \varphi''(u_c) &= \alpha(\alpha-1)\alpha^{-\left(\frac{\alpha-2}{\alpha+1}\right)} + 2\alpha^{\frac{3}{\alpha+1}} \\ &= \alpha^{2/(\alpha+1)} \left(\alpha(\alpha+1)\right) \alpha^{-\alpha/(\alpha+1)}. \end{aligned}$$

We now consider the same three cases as for II:

(i') $\alpha > 1$: Since $u_c < 1$, the only contribution to the asymptotics will come from the boundary point $u = 1$, which will give (21).

(ii') $\alpha = 1$: $u_c = 1$ and we get 1/2 times (20), as before.

(iii') $\alpha < 1$. Now $u_c > 1$ will give a contribution to the asymptotics of I, which turns out to be the same as (20) (but with $\alpha < 1$, of course). By lemma (22) this contribution will win again from that coming from the boundary point.

It now suffices to add up the asymptotics of I and II and observe that, once more, by (22), in cases $(i)+(i')$ and $(iii)+(iii')$ the contribution of the interior minimum will dominate that of the boundary point. QED

5 Precise asymptotics for r_{t+k} given r_t

We will now analyze the $|x| \to \infty$ asymptotics of

$$p_k(x) = p_k(x; t, \rho_0, s_0) = P(r_{t+k} = x | r_t = \rho_0, \sigma_t = s_0)$$

for *fixed* k, $(r_t)_t$ being given by a classical GARCH(1, 1), with

$$\varphi(r, \sigma) = (a_0 + a_1 r^2 + b_1 \sigma^2)^{1/2} \tag{23}$$

and normally distributed ε_t, iid.

The asymptotics of $p_k(x)$ for arbitrary k will follow inductively from theorem 2.1. The main step is the following lemma. Recall the definition (7) of H and the formula (12) for the kernel in the case of a GARCH(1,1), where we take

$$f(x) = (2\pi)^{-1/2} e^{-x^2/2},$$

the standard Gaussian density.

Lemma 5.1 *Suppose that $v(s) \simeq C s^{\beta} e^{-cs^{\alpha}}$ for $0 < s \to \infty$, where $\beta \in R$, $c > 0$ and $\alpha > 0$, suppose that φ is given by* **??** *Then:*

$$Hv(s) \simeq C' s^{(2\beta-\alpha)/(\alpha+2)} e^{-c' s^{2\alpha/(\alpha+2)}}, \quad s \to \infty,$$

where

$$c' = \frac{1}{2}(\alpha+2)(\alpha a_1)^{-\frac{\alpha}{\alpha+2}} c^{\frac{2}{\alpha+2}}$$

and

$$C' = \frac{2Ce^{\frac{b_1}{2a_1}}}{\sqrt{\alpha+2}}(c\alpha a_1)^{-\frac{\beta+1}{\alpha+2}}.$$

Proof. We first treat the case of an $ARCH(1)$: $b_1 = 0$, which is computationally somewhat simpler. In this case

$$Hv(s) = \gamma(s) \int_0^{\infty} \frac{1}{t} e^{-(s^2-a_0)/2a_1 t^2} v(t) dt,$$

where

$$\gamma(s) = \frac{1}{\sqrt{2\pi}} \frac{2s}{\sqrt{a_1(s^2-a_0)}},$$

for $s^2 > a_0$, while $v(s) = 0$ for $s^2 \leq a_0$. Making the change of variables $z = 1/t^2$, we obtain for $s^2 > a_0$, putting $\tilde{s} = \frac{s^2-a_0}{2a_1}$:

$$Hv(s) = \frac{1}{2}\gamma(s) \int_0^{\infty} e^{-\tilde{s}z} \frac{1}{z} v\left(\frac{1}{\sqrt{z}}\right) dz.$$

The integral on the right hand side is the Laplace transform of $z^{-1}v(z^{-1/2})$ evaluated in $(s^2 - a_0)/2a_1$, whose large s-behavior is completely determined by the small z-behavior of

$$\frac{1}{z} v\left(\frac{1}{\sqrt{z}}\right) \simeq C z^{-(\beta/2)-1} e^{-cz^{-\alpha/2}}, \quad z \to 0,$$

by the hypothesis on v. Part (i) of the lemma now follows from lemma 4.1 and straightforward calculations. We use here that $\exp(-c'(s^2 - a_0)^{\alpha/\alpha+2}) \simeq \exp(-c' s^{2\alpha/\alpha+2})$ for $s \to \infty$, since $\alpha/\alpha + 2 < 1$ (this would in fact be false otherwise).

The argument for a $GARCH(1,1)$ ($b_1 \neq 0$) is slightly more involved. First, in that case,

$$Hv(s) = \frac{2se^{b_1/2a_1}}{\sqrt{2\pi a_1}} \int_0^{\sqrt{(s^2-a_0)/b_1}} \frac{1}{\sqrt{t^2(s^2-a_0-b_1t^2)}} e^{-(s^2-a_0)/2a_1t^2} v(t)dt$$

if $s^2 > a_0$ and $Hv(s) = 0$ otherwise. Note that the integral no longer extends over the whole of the positive reals, as it did for an $ARCH(1)$. Making the same change of variables $z = 1/t^2$ as before, we obtain that, putting $\gamma_1(s) = \frac{2se^{b_1/2a_1}}{\sqrt{2\pi a_1}}$

$$\begin{aligned} Hv(s) &= \frac{1}{2}\gamma_1(s) \int_{b_1/(s^2-a_0)}^{\infty} \sqrt{\frac{z}{s^2-a_0-b_1z^{-1}}} e^{-(\frac{s^2-a_0}{2a_1})z} \frac{1}{z} v(\frac{1}{\sqrt{z}})dz \\ &= \frac{1}{2}\gamma_1(s) \frac{1}{\sqrt{b_1}} \int_{\widetilde{s}^{-1}}^{\infty} \sqrt{\frac{z}{\widetilde{s}z-1}} e^{-(\frac{b_1}{2a_1})\widetilde{s}z} \frac{1}{z} v\left(\frac{1}{\sqrt{z}}\right) dz, \end{aligned}$$

where we put $\widetilde{s} = (s^2-a_0)/b_1$, for notational convenience. One easily see's that the main contribution to the $\widetilde{s} \to \infty$-behavior of this integral will again come from the asymptotics of $z^{-1}v(z^{-1/2})$ as $z \to 0$, which, as before, is given by $Cz^{-(\beta/2)-1}\exp(-cz^{-\alpha/2})$. We therefore have reduced the problem to the asymptotics of the following integral:

$$\begin{aligned} & \int_{\widetilde{s}^{-1}}^{\infty} \sqrt{\frac{z}{\widetilde{s}z-1}} e^{-b_1\widetilde{s}z/2a_1} z^{-(\beta/2-1)} e^{-cz^{-\alpha/2}} dz \qquad (24) \\ = {} & \widetilde{s}^{(\beta-1)/2} \int_1^{\infty} \sqrt{\frac{w}{w-1}} e^{-b_1w/2a_1} w^{-(\beta/2)-1} e^{-c\widetilde{s}^{\alpha/2}w^{-\alpha/2}} dw \\ = {} & \frac{2}{\alpha}\widetilde{s}^{(\beta-1)/2} \int_0^1 \left(\frac{1}{1-y^{2/\alpha}}\right)^{1/2} e^{-b_1y^{-2/\alpha}/2a_1} y^{(\beta/\alpha)-1} e^{-c\widetilde{s}^{\alpha/2}y} dy, \end{aligned}$$

where we subsequently made the changes of variables $w = \widetilde{s}z$ and $y = w^{-\alpha/2}$ (note that in the second integral the phase function with minus the big parameter in front has it's minimum in $w = \infty$). The integral on the right is again a Laplace transform, whose main order asymptotic behavior is equal to that of

$$\int_0^1 y^{(\beta/\alpha)-1} e^{-b_1y^{-2/\alpha}/2a_1} e^{-c\widetilde{s}^{\alpha/2}y} dy. \qquad (25)$$

We can replace the interval of integration by $[0, \infty)$, thereby making an error of the form $O\left(s^{power}\exp(-(\text{const})s^{\alpha}\right)$ which will be of lower order, since $2\alpha/(\alpha+2) < \alpha$ for $\alpha > 0$. The resulting integral is a Laplace transform of the kind studied in lemma 4.1, with c equal to $b_1/2a_1$, β replaced by $-(\beta/\alpha)+1$, α by $2/\alpha$ and s by $c\tilde{s}^{\alpha/2}$. After some calculations we find the asymptotics. QED

Remark 5.2 It is surprising that the only difference between an ARCH(1) and a GARCH(1,1), as concerns the asymptotics of lemma 5.1, is the constant in front, which simply get's multiplied by an $\exp(b_1/2a_1)$ in case of a GARCH.

Lemma 5.3 *Suppose that $v(s) \simeq Cs^\beta e^{-cs^\alpha}$ for $0 < s \to \infty$, where $\beta \in R$, $c > 0$ and $\alpha > 0$, suppose that φ is given by (**??**) Then:*

$$Fv(s) \simeq C'|s|^{(2\beta-\alpha)/(\alpha+2)} e^{-c' s^{2\alpha/(\alpha+2)}}, \quad s \to \infty,$$

where

$$c' = \frac{1}{2}(\alpha+2)c^{\frac{2}{\alpha+2}}(\alpha)^{-\frac{\alpha}{\alpha+2}}$$

and

$$C' = \frac{2C}{\sqrt{\alpha+2}}(c\alpha)^{-\frac{\beta+1}{\alpha+2}}.$$

Proof.

$$\begin{aligned} Fv(s) &= \frac{2}{\sqrt{2\pi}}\int_0^\infty \frac{e^{-\frac{x^2}{2s^2}}}{s} v(s)ds \\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty \sqrt{u}e^{-\frac{ux^2}{2}} v(\frac{1}{\sqrt{u}})\frac{1}{u^{3/2}}du \\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-\frac{ux^2}{2}}\frac{1}{u}v(\frac{1}{\sqrt{u}})du, \end{aligned}$$

making the change of variables $u = 1/s^2$. The integral on the right hand side is the Laplace transform of $u^{-1}v(u^{-1/2})$ evaluated in $\frac{x^2}{2}$, whose large u-behavior is completely determined by the small u-behavior of

$$\frac{1}{u}v\left(\frac{1}{\sqrt{u}}\right) \simeq Cu^{-(\beta/2)-1}e^{-cu^{-\alpha/2}}, \quad u \to 0,$$

by the hypothesis on v. The lemma now follows from lemma 4.1 and straightforward calculations.

We next derive the asymptotic behavior of $H^k(\delta(\varphi))$.

Lemma 5.4

$$\begin{aligned} q_k(s) &= H^k(\delta_{\varphi(\rho_0,s_0)})(s), k>1 \\ &\simeq C_k s^{-(1-1/k)} e^{-\frac{k}{2a_1\varphi_0^{2/k}} s^{2/k}}, s\to\infty, \end{aligned}$$

where

$$C_k = \frac{e^{\frac{kb_1}{2a_1}}}{\sqrt{2\pi a_1}}\sqrt{\frac{2^{k-1}}{k}}\varphi_0^{-1/k},$$

with

$$\varphi_0 = \varphi(\rho_0, s_0) = (a_0 + a_1\rho_0^2 + b_1 s_0^2)^{1/2}.$$

Proof.
We will show by induction that $q_k \simeq C_k s^{\beta_k}\exp(-c_k s^{\alpha_k})$, where $\beta_k = -(1-1/k)$, $c_k = \frac{k}{2a_1\varphi_0^{2/k}}$, $\alpha_k = \frac{2}{k}$ and C_k is given in the statement of the lemma. First, for $k=1$:

$$\begin{aligned} q_1(s) &= H(\delta_{\varphi(\rho_0,s_0)})(s), \\ &= \frac{1}{\sqrt{2\pi}}\frac{2s}{\sqrt{a_1\varphi_0^2(s^2-a_0-b_1\varphi_0^2)}} e^{-\frac{1}{2}\frac{s^2-a_0-b_1\varphi_0^2}{a_1\varphi_0^2}} \\ &\simeq \frac{1}{\sqrt{2\pi a_1}}\frac{1}{\varphi_0} e^{\frac{b_1}{2a_1}} e^{-\frac{s^2}{2a_1\varphi_0^2}}, s\to\infty, \end{aligned}$$

as required.
We now assume that the lemma is true for $k-1$. Since $q_k(s) = H(q_{k-1})(s)$, we have by lemma 5.1:

$$q_k(s) \simeq C' s^{(2\beta_{k-1}-\alpha_{k-1})/((\alpha_{k-1}+2))} e^{-c' s^{\frac{2\alpha_{k-1}}{\alpha_{k-1}+2}}}, \quad s\to\infty \tag{26}$$

Now $\frac{2\alpha_{k-1}}{\alpha_{k-1}+2} = 2/k = \alpha_k$ and, similary,

$$\frac{2\beta_{k-1}-\alpha_{k-1}}{\alpha_{k-1}+2} = -(1-1/k) = \beta_k,$$

using the expressions for β_{k-1} and α_{k-1}. From lemma 5.1 we get:

$$\begin{aligned} c' &= \frac{1}{2}(\alpha_{k-1}+2) c_{k-1}^{\frac{2}{\alpha_{k-1}+2}} (\alpha_{k-1}a_1)^{-\frac{\alpha_{k-1}}{\alpha_{k-1}+2}} \\ &= \frac{k}{2a_1\varphi_0^{2/k}} = c_k, \end{aligned}$$

after a computation. Finally, by lemma 5.1 again,

$$\begin{aligned} C' &= \frac{2e^{\frac{b_1}{2a_1}}}{\sqrt{\alpha_{k-1}+2}}(c_{k-1}\alpha_{k-1}a_1)^{-\frac{\beta_{k-1}+1}{\alpha_{k-1}+2}} \\ &= 2e^{\frac{b_1}{2a_1}}\sqrt{\frac{k-1}{2k}}\left(\frac{1}{\varphi_0^{2/(k-1)}}\right)^{-1/2k}C_{k-1} \\ &= e^{\frac{b_1}{2a_1}}\sqrt{\frac{2(k-1)}{k}}\varphi_0^{1/k(k-1)}C_{k-1}, \end{aligned}$$

with

$$C_1 = \frac{1}{\sqrt{2\pi}}\frac{1}{a_1}\frac{1}{\varphi_0}e^{\frac{b_1}{2a_1}}$$

and a simple induction allows us to verify the formula for C_k. QED

We can now state the main result of this section:

Theorem 5.5 *Let $(r_t)_t$ be a GARCH(1, 1) with φ given by (23) and independent, normally distributed ε_t with mean 0 and variance 1. Fix a time t and a time-horizon $t+k$ and suppose that $r_t = \rho_0$ and $\sigma_t = s_0$. Let*

$$c_k = \frac{1}{2}ka_1^{-(1-1/k)}(a_0 + a_1\rho_0^2 + b_1 s_0^2)^{-1/k}$$

and

$$C_k = \frac{e^{\frac{(k-1)b_1}{2a_1}}}{\sqrt{2\pi}}a_1^{-\frac{1}{2}(1-1/k)}\sqrt{\frac{2^{k-1}}{k}}\varphi_0^{-1/k}.$$

Then

$$p_{t,k}(x;\rho_0,s_0) \simeq C_k\frac{e^{-c_k|x|^{2/k}}}{|x|^{1-1/k}}, \quad x \to \pm\infty \tag{27}$$

Proof.
The proof relies on the previous lemma. With the same notations as before,

$$p_{t,k}(x;\rho_0,s_0) = F(H^{k-1}(\delta_{\varphi_0})) = F(q_k(x)).$$

For $k = 1$, we get:

$$F(\delta_{\varphi_0})(x) = \frac{1}{\varphi_0\sqrt{2\pi}}e^{-\frac{x^2}{2\varphi_0^2}}.$$

For $k > 1$ we use the lemma 5.3 and the lemma 5.2 with

$$v(s) = q_{k-1}(s) \simeq C_{k-1}s^{\beta_{k-1}}e^{-c_{k-1}s^{\alpha_{k-1}}}.$$

Thus,

$$p_k(x) = F(v)(x) \simeq C'|x|^{(2\beta_{k-1}-\alpha_{k-1})/(\alpha_{k-1}+2)} e^{-c' x^{\frac{2\alpha_{k-1}}{\alpha_{k-1}-2}}},$$

with C' and c' given in the lemma 5.2. After some computations we get:

$$c_k = \frac{k}{2} \frac{1}{a_1^{k-1/k}} \frac{1}{\varphi_0^{2/k}},$$

and

$$C_k = \frac{e^{\frac{(k-1)b_1}{2a_1}}}{\sqrt{2\pi}} a_1^{-\frac{1}{2}(1-1/k)} \sqrt{\frac{2^{k-1}}{k}} \varphi_0^{-1/k},$$

as required. QED

6 Asymptotics of multi-period returns

If r_t represents a one-period logarithmic return considering r_{t+k} by itself does not make sense. A financiallymore relevant quantity here will be the corresponding k-period return, over $[t, t+k]$, which is given by (14): $r_{t+k,t} = r_{t+1} + r_{t+2} + \cdots + r_{t+k}$. The main result of this section is that , qualitatively, the extreme values of $r_{t+k,t}$, conditional on given values for X_t and σ_t at time t, behave like those of r_{t+k}:

Theorem 6.1 *Let $(r_t)_t$ follow a classical GARCH(1, 1), with $\varphi(r, \sigma) = (a_0 + a_1 r^2 + b_1 \sigma^2)^{1/2}$ and standard normally distributed $(\varepsilon_t)_t$, where we moreover suppose that $b_1 > 0$. Let $r_{t+k,t}$ be defined by (14). Fix a k and let $\rho_0 \in R, s_0 > 0$. Then there exist constants $c_k, c'_k, C_k, C'_k > 0$, depending on k, a_0, a_1, b_1, ρ_0 and s_0 such that for $|x| \geq 1$,*

$$C'_k |x|^{-(1-1/k)} e^{-c_k |x|^{2/k}} \leq P(r_{t+k,t} = x | r_t = \rho_0, \sigma_t = s_0) \leq C_k |x|^{-(1-1/k)} e^{-c_k |x|^{2/k}} \tag{28}$$

Explicit values for the constants can be extracted from the proof below: we won't do that here. Also, the restriction to $b_1 > 0$ is probably technical.

Proof. The proof is based on formula (16) from section 3, which in our situation reads:

$$P(r_{t+k,t} = x | r_t = \rho_0, \sigma_t = s_0) = \tag{29}$$

$$\left(\frac{1}{2\pi)}\right)^{(k-1/2)} \int_R \cdots \int_R \Pi_{j=1}^{k-1} \frac{1}{\hat{s}_j} e^{-x_j^2/2\hat{s}_j^2} \frac{1}{\hat{s}_k} e^{-(x-(x_1+\cdots+x_{k-1}))^2/2\hat{s}_k^2} dx_1 \cdots dx_{k-1},$$

where the standard deviations $\hat{s}_j = \hat{s}_j(x_1, \cdots, x_{j-1})$ are defined inductively by

$$\hat{s}_1^2 = (a_0 + a_1\rho_0^2 + b_1 s_0^2)^{1/2}$$
$$\hat{s}_j^2 = (a_0 + a_1 x_{j-1}^2 + b_1 \hat{s}_{j-1}^2)^{1/2}.$$

It easily follows that

$$\hat{s}_j^2 = \sum_{\nu=1}^{j-1} a_1 b_1^{\nu-1} x_{j-\nu}^2 + e_\nu,$$

where $e_1 = \hat{s}_1^2$ and $e_k = a_0 + b_1 e_{k-1}$. We will in fact establish a slightly more general result, replacing the $\hat{s}_j^2$ in (29) by functions $L_{j-1} = L_{j-1}(x_1, \cdots, x_{j-1})$ which are affine in in $x_1^2, \cdots x_{j-1}^2$ (note the shift by 1 of the index w.r;t. the notation used for $\hat{s}_j$). Here

$$L_j(x_1, \cdots, x_j) = \gamma_0^{(j)} + \sum_{\nu=1}^{j} \gamma_\nu^{(j)} x_\nu^2 \tag{30}$$

For our proof to work we will have to impose the condition:

$$\gamma_\nu^{(j)} > 0, \quad 0 \leq \nu \leq j \tag{31}$$

Note that L_0 is thus just a strictly positive constant. The $\hat{s}_j^2$ coming from a GARCH(1, 1) with $b_1 > 0$ fall into this class; those coming from an ARCH unfortunately do not.

We will also put an adjustable multiplicative constant $\eta > 0$ in the exponent of the final factor of (29) and estimate the functions $q_k(x)$ defined by

$$q_k(x) = q_k(x; \eta, L_0, \cdots, L_{k-1}) = \tag{32}$$

$$\int_R \cdots \int_R \left(\Pi_{j=1}^{k-1} \frac{e^{-x_j^2/2L_{j-1}}}{\sqrt{2\pi L_{j-1}}} \right) \frac{1}{\sqrt{2\pi L_{k-1}}} e^{-\eta(x-(x_1+\cdots x_{k-1}))^2/2L_{k-1}} dx_1 \cdots x_{k-1}.$$

More precisely, we will prove the following inequalities, from which theorem 6.1 will be an immediate consequence:

Claim 6.2 *For given k, affine forms $L_0, \cdots L_{k-1}$ as in (30), satisfying (31) and given $\eta > 0$, there exist strictly positive constants c, c', C and C' such that*

$$C|x|^{-(1-1/k)} e^{-c|x|^{2/k}} \leq q_k(x) \leq C'|x|^{-(1-1/k)} e^{-c'|x|^{2/k}} \tag{33}$$

The constants c, c', C and C' can be chosen to depend locally uniformly on η and the coefficients of the L_j.

We turn to the proof of the claim, which will be by induction on k. The idea is to estimate $q_k(x)$ from above and from below by a Laplace transform of a q_{k-1} with slightly modified η and L 's, modulo a neglegible error, and then use lemma 4.1 again. To accomplish this, we will eliminate the x_1 from all factors under the integral sign of (32), except the first one. We will use the following elementary inequality:

Lemma 6.3 *For all ε with $0 < \varepsilon \leq 1$ and all $a, b \in R$ one has that*

$$C^-_{b,\varepsilon} e^{-(1+\varepsilon)a^2} \leq e^{-(a+b)^2} \leq C^+_{b,\varepsilon} e^{-(1-\varepsilon)a^2}, \tag{34}$$

where $C^-_{b,\varepsilon} = \exp(-(\varepsilon^{-1}+1)b^2)$ and $C^+_{b,\varepsilon} = \exp((\varepsilon^{-1}-1)b^2)$

Proof. To prove for example the upper bound, write $\exp((1-\varepsilon)a^2)\exp(-(a+b)^2) = \exp(-\varepsilon a^2 + 2ab + b^2)$ and maximize over a. The lower bound is proven in the same way.

It is clear that (33) holds for $k = 1$. Now suppose that it holds for $k-1$. We then have to prove it for k. We first establish the upper bound in (33). Apply the second inequality in (34) with $a = \sqrt{\eta}(x-(x_2+\cdots+x_k))/\sqrt{2L_{k-1}}$ and $b = \sqrt{\eta}x_1/\sqrt{2L_{k-1}}$. The constant $C^+_{b,\varepsilon}$ then becomes

$$C^+_{b,\varepsilon} = e^{(\varepsilon^{-1}-1)\eta x_1^2/2L^2_{k-1}} \leq e^{(\varepsilon^{-1}-1)\eta x_1^2/2\gamma_0^{(k-1)}},$$

and we thus see that it can be absorbed in the first factor in the integrand of (32), $\exp(-x_1^2/2L_0)$, provided ε is sufficiently close to 1. In fact, $C_{b,\varepsilon} < \exp x_1^2/4L_0$ if

$$(1 + \gamma_0^{(k-1)}/2\eta L_0)^{-1} < \varepsilon < 1$$

and thus, with such a choice of ε we have that

$$q_k(x) \leq \int_R \cdots \int_R \frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}} \cdot \left(\Pi_{j=2}^{k-1} \frac{e^{-x_j^2/2L_{j-1}}}{\sqrt{2\pi L_{j-1}}}\right) \cdot \tag{35}$$

$$\cdot \frac{1}{\sqrt{2\pi L_{k-1}}} e^{-(1-\varepsilon)\eta(x-(x_2+\cdots x_{k-1}))^2/2L_{k-1}} dx_1 \cdots x_{k-1}.$$

We now split this integral as

$$\int_{|x_1|\leq 1} dx_1 \frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}} \int_{R^{k-2}} (\cdots) + \int_{|x_1|>1} dy_1 \frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}} \int_{R^{k-2}} (\cdots) \tag{36}$$

$$= I + II$$

and estimate the two pieces seperately. We first show that I is of the same order as a suitable $q_{k-1}(x)$. In fact, if $|x_1| \leq 1$, then for $\nu \geq 1$,

$$\begin{aligned} L_j(x_1, \cdots x_j) &\leq (\gamma_0^{(j)} + \gamma_1^{(j)} + \gamma_2^{(j)} x_2^2 + \cdots + \gamma_j^{(j)} x_j^2) \\ &=: L_j^*(x_2, \cdots, x_j). \end{aligned}$$

One also has that

$$\frac{L_j^*}{L_j} \leq \max(1, (\gamma_0^{(j)} + \gamma_1^{(j)})/\gamma_0^{(j)}),$$

this without any restriction on $(x_1, \cdots, x_j)$. It follows that, for a suitable constant $C > 0$ (which we won't specify),

$$|I| \leq C \int_{|x_1| \leq 1} \int_R \cdots \int_R \frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}} \cdot \left(\Pi_{j=2}^{k-1} \frac{e^{-x_j^2/2L_{j-1}^*}}{\sqrt{2\pi L_{j-1}^*}} \right)$$

$$\cdot \frac{1}{\sqrt{2\pi L_{k-1}^*}} e^{-(1-\varepsilon)\eta(x-(x_2+\cdots x_{k-1}))^2/2L_{k-1}^*} dx_1 \cdots x_{k-1}.$$

We recognize the integral over $dx_2 \cdots dx_{k-1}$ as a constant times $q_{k-1}(x; (1-\varepsilon)\eta; L_2^*, \cdots, L_{k-1}^*)$, and therefore, by the induction hypothesis, for suitable constants c, C,

$$|I| \leq C|x|^{-(1-1/(k-1))} e^{-c|x|^{2/(k-1)}}, \tag{37}$$

which is of strictly lower order than the inequality we're trying to establish for $q_k(x)$.

We next turn to the integral II. If $|x_1| > 1$, then

$$\begin{aligned} L_j(x_1, \cdots, x_j) &\leq (\gamma_0^{(j)} + \gamma_1^{(j)}) x_1^2 + \gamma_2^{(j)} x_2^2 + \cdots + \gamma_j^{(j)} x_j^2 \\ &= x_1^2 \left(\gamma_0^{(j)} + \gamma_1^{(j)} + \gamma_2^{(j)} \frac{x_2^2}{x_1^2} + \cdots + \gamma_j^{(j)} \frac{x_j^2}{x_1^2} \right) \\ &=: x_1^2 \widetilde{L}_j \left(\frac{x_2}{x_1}, \cdots, \frac{x_j}{x_1} \right), \end{aligned}$$

the last equation defining $\widetilde{L}_j$. Similarly, for $|x| > 1$ we can estimate

$$L_j(x_1, \cdots, x_j) \geq \gamma_1^{(j)} x_1^2 + \gamma_2^{(j)} x_2^2 + \cdots + \gamma_j^{(j)} x_j^2$$

$$
\begin{aligned}
&= x_1^2\left(\gamma_1^{(j)}+\gamma_2^{(j)}\frac{x_2^2}{x_1^2}+\cdots+\gamma_j^{(j)}\frac{x_j^2}{x_1^2}\right)\\
&\geq cx_1^2\widetilde{L}_j\left(\frac{x_2}{x_1},\cdots,\frac{x_j}{x_1}\right), \qquad (38)\\
& \qquad (39)
\end{aligned}
$$

provided that

$$c \leq \frac{\gamma_1^{(j)}}{\gamma_0^{(j)}+\gamma_1^{(j)}}.$$

Note that to have (38) with a $c > 0$ we need here that $\gamma_1^{(j)} > 0$, which is insured by (31). Substituting these inequalities in (32), we find that for suitable $C > 0$,

$$II \leq \int_{|x_1|>1}\frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}}\left(\Pi_{j=2}^{k-1}\frac{e^{-x_j^2/2x_1^2\widetilde{L}_{j-1}}}{|x_1|\sqrt{2\pi\widetilde{L}_{j-1}}}\right)\cdot$$

$$\cdot\frac{1}{|x_1|\sqrt{2\pi\widetilde{L}_{k-1}}}e^{-(1-\varepsilon)\eta(x-(x_2+\cdots x_{k-1}))^2/2x_1^2\widetilde{L}_{k-1}}dx_1\cdots x_{k-1}.$$

If we change variables to $y_j := x_j/x_1$ for $2 \leq j \leq k-1$ we see that the previous inequality can be written as:

$$II \leq C\int_{|x_1|>1}\frac{1}{|x_1|}\frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}}q_{k-1}\left(\frac{x}{x_1};(1-\varepsilon)\eta,\widetilde{L}_2,\cdots,\widetilde{L}_{k-1}\right).$$

By the induction-hypothesis, the $q_{k-1}(x/x_1)$ under the integrand is less than or equal to

$$C\left(\frac{|x|}{|x_1|}\right)^{\beta}e^{-c(|x|/|x_1|)^{\alpha}}$$

with

$$\alpha = 2/(k-1), \quad \beta = -(1-1/(k-1)) \qquad (40)$$

and thus, after a rescaling,

$$II \leq |x|^{\beta}C\int_{|x_1|>1}|x_1|^{-\beta-1}e^{-c(|x|/|x_1|)^{\alpha}}e^{-x_1^2}dx_1$$

with different constants $c, C > 0$ (in which are absorbed L_0 and also the various factors of 2π). We assume now wlog that $x > 0$ and we write the integral as twice the integral over $[1,\infty)$. We again want to use lemma 4.1 and for this we rewrite our integral as a Laplace transform with big parameter, by introducing the new variable $z = x_1^{-\alpha}$. Then the right hand side of (41) is less or equal a constant times

$$x^\beta \int_0^1 z^{(\beta/\alpha)-1} e^{-z^{-2/\alpha}} e^{-cx^\alpha z} dz$$

and by lemma 4.1, with $s = cx^\alpha$ and α, β replaced by, respectively, $2/\alpha$ and $1-(\beta/\alpha)$, we find that

$$II \leq Cx^{(2\beta-\alpha)/(\alpha+2)} \exp(-cx^{2\alpha/(\alpha+2)}).$$

Since, by (40), the two exponents of x in this formula are, respectively, $-(1-1/k)$ and $2/k$, this proves the desired upper bound for II and thus for $q_k(x)$, remembering (36) and (37).

We next turn to the lower bound for q_k. By the first inequality of lemma (34), we see in the same way as before that

$$e^{-\eta(x-(x_1+\cdots+x_{k-1}))^2/2L_{k-1}} \geq C_{b,\varepsilon}^- e^{-(1+\varepsilon)\eta(x-(x_2+\cdots+x_{k-1}))^2/2L_{k-1}}$$

where

$$C_{b,\varepsilon}^- = e^{-\eta(1+\varepsilon^{-1})x_1^2/2L_{k-1}} \geq e^{-\eta(1+\varepsilon^{-1})x_1^2/2\gamma_0^{(k-1)}}.$$

We can combine $C_{b,\varepsilon}$ with the first factor of the integrand of the defining equation (32) of $q_k(x)$ into a factor $e^{-\kappa x_1^2}$. Doing so, and limiting the x_1-integration in (32) to $|x_1| > 1$, we find that

$$q_k(x) \geq \int_{|x_1|>1} \int_R \cdots \int_R \frac{e^{-\kappa x_1^2}}{\sqrt{2\pi L_0}} \cdot \left(\Pi_{j=2}^{k-1} \frac{e^{-x_j^2/2L_{j-1}}}{\sqrt{2\pi L_{j-1}}} \right) \tag{41}$$

$$\cdot \frac{1}{\sqrt{2\pi L_{k-1}}} e^{-\eta(1+\varepsilon)(x-(x_2+\cdots x_{k-1}))^2/2L_{k-1}} dx_1 \cdots x_{k-1}.$$

As before, we next get rid of the x_1 in the $L_1, \cdots, L_{k-1}$; first, if $j \geq 1$, then

$$\begin{aligned} L_j(x_1, \cdots x_j) &\geq x_1^2 \left(\gamma_1^{(j)} + \gamma_2^{(j)} \frac{x_2^2}{x_1^2} + \cdots + \gamma_j^{(j)} \frac{x_j^2}{x_1^2} \right) \\ &=: x_1^2 \hat{L}_j \left(\frac{x_2}{x_1}, \cdots, \frac{x_j}{x_1} \right). \end{aligned}$$

Next, if $|x_1| > 1$, then

$$\begin{aligned} L_j(x_1, \cdots, x_j) &\leq (\gamma_0^{(j)} + \gamma_1^{(j)}) x_1^2 + \cdots + \gamma_j^{(j)} x_j^2 \\ &\leq c x_1^2 \hat{L}_j \left(\frac{x_2}{x_1}, \cdots, \frac{x_j}{x_1} \right), \end{aligned} \tag{42}$$

provided that $c \geq (\gamma_0^{(j)} + \gamma_1^{(j)})/\gamma_1^{(j)}$; there exists such a (finite) c since $\gamma_1^{(j)} > 0$ by (31). Substituting these inequalities in (41) and making the same change of variables $y_j = x_j/x_1$ as before ($j \geq 2$) one finds that, for a suitable constant $C > 0$,

$$q_k(x) \geq C \int_{|x_1|>1} \frac{e^{-\kappa x_1^2}}{|x_1|} q_{k-1}\left(\frac{x}{x_1}; (1+\varepsilon)\eta, \hat{L}_2, \cdots, \hat{L}_{k-1}, \eta(1+\varepsilon)\right).$$

Using the induction hypothesis and lemma 4.1, we find the required lower bound for $q_k(x)$. QED

7 Application to VaR

The asymptotics which we established are relevant for risk analysis. We illustrate this with a multi-period VaR-estimate based on theorem 6.1. We first note that the upper bound from that theorem implies that

$$\text{Prob}\,(r_{t+k,t} < -x) \leq C_k \int_x^\infty y^{-(1-1/k)} e^{-c_k y^{2/k}} dy \quad (x \geq 0).$$

The right hand side can be evaluated in terms of (one of) the incomplete , -function(s),

$$\gamma_{z,\infty}(x) := \int_x^\infty t^{z-1} e^{-t} dt,$$

leading to

$$\text{Prob}\,(r_{t+k,t} < -x) \leq \frac{1}{2} k c_k^{-1/2} C_k \gamma_{1/2,\infty}\left(c_k x^{2/k}\right) \quad (x \geq 0) \tag{43}$$

together with a similar lower bound, with c_k and C_k replaced by c_k' and C_k', respectively. Note that all these can be expressed in terms of a single special function, $\gamma_{1/2,\infty}$.

Next, for a given random variable X we denote by F_X it's cumulative distribution function and by $q_\alpha(X)$ it's α-th lower quantile:

$$q_\alpha^-(X) = \inf\{y : F_X(y) = \alpha\}, \quad \alpha \in [0,1]$$

and, more generally, for any non-negative non-decreasing function F we introduce

$$q_\alpha^-(F) = \inf\{y : F(y) = \alpha\},$$

so that $q_\alpha(X)$ is the same as $q_\alpha(F_X)$ (this abuse of notation won't cause confusion). We note the following trivial observation, whose proof is left to the reader:

Lemma 7.1 *Let F and G be non-decreasing continuous functions defined on a semi-infinite interval $(-\infty, a)$, such that $F \leq G$ there. Let α be in the range of G. Then*

$$q_\alpha^-(G) \leq q_\alpha^-(F).$$

We next recall that the k-period Value at Risk $VaR_{1-\alpha}(k)$ with confidence $1-\alpha$ is defined by

$$\begin{aligned} VaR_{1-\alpha}(k) &= \sup\{V \geq 0 : \mathrm{Prob(P_{t+k} - P_t < -V) = \alpha}\} \\ &= -q_\alpha^-(P_{t+k} - P_t) \\ &\simeq -q_\alpha(r_{t+k,t})P_t, \end{aligned}$$

where we made the usual approximation $e^r - 1 \simeq r$ for r small (this could easily have been circumvented, at the cost of complicating the formula). Applying the lemma with $F = F_{r_{t,t+k}}$ and G equal the right hand side of (43), both with domain $(-\infty, 0)$, and observing that $q_\alpha(G)$ is simply $G^{-1}(\alpha)$, we easily find that

$$VaR_{1-\alpha}(k) \leq \left(\frac{1}{c_k}\gamma_{1/2,\infty}^{-1}\left(\frac{2\alpha\sqrt{c_k}}{kC_k}\right)\right)^{k/2} P_t,$$

and a similar lower bound, with the constants replaced by the primed ones. This is the estimate which should replace the RiskMetrics proposal (2). Although more complicated than the latter one, it is quite explicit and can be easily evaluated, either numerically or asymptotically for small α. It is clear that for practical applications a good control of the constants is essential. As already stated, we wil return to that in a companion paper.

References

[1] Artzner, Ph., Delbaen, F., Eber, J.-M., Heath, D., 1999. Coherent measures of risk, Mathematical Finance, 9 (3), 203 - 228.

[2] Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity, J. of Econometrics 31, 307 - 325.

[3] Bollerslev, T., Engle, R. F., Nelson, D. B., 1994. ARCH models, in: Engle, R. F., McFadden, D., eds., Handbook of Econometrics, vol. 4, North Holland, Amsterdam.

[4] Brummelhuis, R. G. M., Guégan, D., 2000. Extreme values of conditional distributions of GARCH(1, 1) processes, prépublication 00.08, Université de Reims, Dépt. de math., submitted.

[5] Brummelhuis, R. G. M. , Guégan, D. and LaDoucette, S. In preparation.

[6] Diebolt, J., Guégan, D. , 1991, Le modèle β-ARCH, Comptes Rendues Acad. Sciences Paris, Série I, 312, 625 - 630.

[7] Diebolt, J., Guégan, D. , 1993, Tail behavior of the stationary density of general non linear autoregressive process of order one, J. Appl. Prob., 30, 315 - 329.

[8] Dingh, Z., Granger, C.W.J., 1996, Modelling volatility persistence of speculative returns: a new approach, Journal of Econometrics, 73, 185 - 215.

[9] Dowd, K.,1998, *Beyond Value at Risk*, Wiley Frontiers in Finance.

[10] Embrechts, P., Mikosch, T. Klüppelberg, C., 1997. *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

[11] Engle, R. F., 1982. Autoregressive conditional heteroskedacity with estimates of the variance of United Kingdom inflation, Econometrica 50, 907 - 1007.

[12] Engle, R. F., Lilian, D. F., Robbins, R. P., 1987. Estimating time varying risk premia in the term structure, Econometrica 55, 391 - 407.

[13] Frey, R., McNeil, A., 1998, Estimation of tail-related risk measures for heteroscedastic financial time series : an extreme value approach, to appear in J. Empirical Finance.

[14] Guégan, D., 1994. *Séries chronologiques non linéaires à temps discrèt*, Economica, coll. Statistique mathématique et probabilité, Paris.

[15] Guégan, D., 2000. A new model: the k-factor GIGARCH process, Journal of Signal Processing 4, 265 - 271.

[16] Jorion, P., 1997. Value at Risk : The New Benchmark for Controlling Market Risk, McGraw-Hill.

[17] Kesten, H., 1973. Random difference equations and renewal theory for products of random matrices, Acta Math. 131, 207 - 248.

[18] Nelson, D. B., 1990. Conditional heteroscedasticity in asset returns: a new approach, Econometrica, 59, 347 - 370.

[19] Mikosch, Th., Starica, C., 1998. Limit theory for the sample autocorrelations and extremes of a GARCH(1, 1) process, preprint.

[20] RiskMetrics - Technical Document, 1996, 4th edition, J. P. Morgan.