Virtual Historical Simulation for estimating the conditional VaR of large portfolios

CHRISTIAN FRANCQ[*] AND JEAN-MICHEL ZAKOÏAN[†]

Abstract

In order to estimate the conditional risk of a portfolio's return, two strategies can be advocated. A multivariate strategy requires estimating a dynamic model for the vector of risk factors, which is often challenging, when at all possible, for large portfolios. A univariate approach based on a dynamic model for the portfolio's return seems more attractive. However, when the combination of the individual returns is time varying, the portfolio's return series is typically non stationary which may invalidate statistical inference. An alternative approach consists in reconstituting a "virtual portfolio", whose returns are built using the current composition of the portfolio. This paper establishes the asymptotic properties of this method, which we call Virtual Historical Simulation (VHS). Numerical illustrations on simulated and real data are provided.

Keywords: Confidence Intervals for VaR, Multivariate GARCH, Estimation risk, Dynamic Portfolio.

1 Introduction

The quantitative standards laid down under Basel Accord II and III allow banks to develop internal models for setting aside capital. Methods that incorporate time dependence to quantify market risks are able to use knowledge of the conditional distribution. In particular, the conditional Value-at-Risk (VaR) of financial returns, with a given confidence level α (typically, $\alpha = 1\%$ or 5%) is nothing else, from a statistical point of view, than the opposite of the α-quantile of the conditional

[*]CREST and University Lille 3, BP 60149, 59653 Villeneuve d'Ascq cedex, France. E-Mail: christian.francq@univ-lille3.fr

[†]Corresponding author: Jean-Michel Zakoïan, University Lille 3 and CREST, 5 Avenue Henri Le Chatelier, 91120 Palaiseau, France. E-mail: zakoian@ensae.fr. The second author gratefully thanks the IDR ACP "Régulation et risques systémiques" for financial support.

distribution of the portfolio returns. Estimating conditional quantiles, or more generally conditional risk measures, of a time series of financial returns is thus crucial for risk management.

It is also essential, for risk management purposes, to be able to evaluate the accuracy of such estimators of conditional risks. Uncertainty implied by statistical procedures in the implementation of risk measures may lead to false security in financial markets (see e.g. Farkas, Fringuellotti and Tunaru (2016) and the references therein). Estimation risk thus needs to be accounted for, in addition to market risk. However, drawing - for instance - confidence intervals (CI) for the conditional Value-at-Risk (VaR) is generally challenging for two main reasons. Firstly, because the stochastic nature of the conditional VaR does not allow in general to reduce the problem to the estimation of a parameter. Deriving CIs for a stochastic process is obviously more intricate than for a parameter. Secondly, quantiles being obtained as the solutions of optimization problems based on non-smooth functions, establishing asymptotic properties of conditional VaR estimators may become a difficult task.

Increasing attention has been directed in the recent econometric literature to the inference of risk measures in dynamic risk models. Francq and Zakoïan (2015) derived asymptotic confidence intervals for the conditional VaR of a series of financial returns driven by a parametric dynamic model. Robust backtesting procedures were developed by Escanciano and Olmo (2010, 2011), and Gouriéroux and Zakoïan (2013) studied the effect of estimation on the coverage probabilities. Several articles proposed resampling methods: among others, Christoffersen and Gonçalves (2005) and Spierdijk (2016) considered using bootstrap procedures for constructing confidence intervals for VaR; Hurlin, Laurent, Quaedvlieg and Smeekes (2017) proposed boostrap-based comparison tests of two conditional risk measures. See Nieto and Ruiz (2016) for an extensive survey of the methods for constructing and evaluating VaR forecasts that have been proposed in the literature.

Most existing studies on risk measure inference focus on the risk of a single financial asset. The aim of the present article is to estimate conditional VaR's for *portfolios* of financial assets. From a statistical point of view, the extension is far from trivial. First, because evaluating the quantile of a linear combination of variables may require knowledge of the complete joint distribution of such variables. When the object of interest is a *conditional* quantile, this approach requires specifying a dynamic model for the vector of returns of the assets involved in the portfolio. Second, portfolios compositions are generally time-varying, in particular if the agents adopt a mean-variance approach which, in a dynamic framework, requires specifying the first two conditional moments. This typically entails non-stationarity of the portfolio's return time series, as we shall see in more detail.

A natural approach for obtaining the VaR of a portfolio relies on specifying a multivariate GARCH model for the vector of underlying asset returns. This approach was investigated by Rombouts and Verbeek (2009), and its asymptotic properties–with or without sphericity of the innovations vector– were derived by Francq and Zakoïan (2017). As noted by Rombouts and Verbeek the advantage of multivariate approaches is to "take into account the dynamic interrelationships between the portfolio components, while the model underlying the VaR calculations is independent of the portfolio composition". On the other hand, for large portfolios multivariate approaches often become untractable due to the well-known dimensionality curse. In this paper, we study the properties of an alternative *univariate* procedure relying on "virtual portolios" constructed with the *current composition* of the portfolio. The series of virtual returns thus obtained allows to estimate the conditional VAR of the portfolio. By circumventing the inherent non stationarity of the observed portfolio's time series, this procedure–which we call Virtual Historical Simulation (VHS)–is amenable to asymptotic statistical inference. From a numerical point of view, it allows to avoid difficulties caused by the dimensionality curse in estimation of multivariate volatility models for vectors of asset returns.

The paper is organized as follows. Section 2 presents a general multivariate model and multivariate approaches for estimating the conditional VaR of a portfolio whose composition at the current date may depend on the historical prices. Section 3 studies univariate methods. We first consider a "naive" approach in which a standard volatility model is directly fitted to the portfolio returns series. Then we derive in Section 4 the asymptotic properties of the VHS procedure under general assumptions. Section 5 presents some numerical illustrations based on Monte Carlo experiments.Proofs are collected in the Appendix.

2 General setup

2.1 Model and dynamic portfolio

Let $\boldsymbol{p}_t = (p_{1t}, \dots, p_{mt})'$ denote the vector of prices of m assets at time t. Let $\boldsymbol{y}_t = (y_{1t}, \dots, y_{mt})'$ denote the corresponding vector of log-returns, with $y_{it} = \log(p_{it}/p_{i,t-1})$ for $i = 1, \dots, m$. We assume throughout that the vector of log-returns follow a general multivariate model of the form

$$\boldsymbol{y}_t = \boldsymbol{m}_t(\boldsymbol{\vartheta}_0) + \boldsymbol{\epsilon}_t, \qquad \boldsymbol{\epsilon}_t = \boldsymbol{\Sigma}_t(\boldsymbol{\vartheta}_0)\boldsymbol{\eta}_t, \tag{2.1}$$

where $(\boldsymbol{\eta}_t)$ is a sequence of independent and identically distributed (iid) $\mathbb{R}^m$-valued variables with zero mean and identity covariance matrix; the $m \times m$ non-singular matrix $\boldsymbol{\Sigma}_t(\boldsymbol{\vartheta}_0)$ and the $m \times 1$

vector $\boldsymbol{m}_t(\boldsymbol{\vartheta}_0)$ are specified as functions parameterized by a d-dimensional parameter $\boldsymbol{\vartheta}_0$ of the past values of $\boldsymbol{y}_t$:

$$\boldsymbol{m}_t(\boldsymbol{\vartheta}_0) = \boldsymbol{m}(\boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \dots, \boldsymbol{\vartheta}_0), \qquad \boldsymbol{\Sigma}_t(\boldsymbol{\vartheta}_0) = \boldsymbol{\Sigma}(\boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \dots, \boldsymbol{\vartheta}_0). \tag{2.2}$$

Let V_t denote the value at time t of a portfolio composed of $\mu_{i,t-1}$ units of asset i, for $i = 1, \dots, m$:

$$V_0 = \sum_{i=1}^{m} \mu_i p_{i0}, \quad V_t = \sum_{i=1}^{m} \mu_{i,t-1} p_{it}, \quad \text{for } t \geq 1 \tag{2.3}$$

where the $\mu_{i,t-1}$ are measurable functions of the prices up to time $t-1$, and the μ_i are constants. The return of the portfolio over the period $[t-1, t]$ is, for $t \geq 1$, assuming that $V_{t-1} \neq 0$,

$$\frac{V_t}{V_{t-1}} - 1 = \sum_{i=1}^{m} a_{i,t-1} e^{y_{it}} - 1 \approx \sum_{i=1}^{m} a_{i,t-1} y_{it} + a_{0,t-1}$$

where

$$a_{i,t-1} = \frac{\mu_{i,t-1} p_{i,t-1}}{\sum_{j=1}^{m} \mu_{j,t-2} p_{j,t-1}}, \quad i = 1, \dots, m \quad \text{and} \quad a_{0,t-1} = -1 + \sum_{i=1}^{m} a_{i,t-1}.$$

We assume that, at date t, the investor may rebalance his portfolio under a "self-financing" constraint.

SF: The portfolio is rebalanced in such a way that $\sum_{i=1}^{m} \mu_{i,t-1} p_{it} = \sum_{i=1}^{m} \mu_{i,t} p_{it}$.

In other words, the value at time t of the portfolio bought at time $t-1$ equals the value at time t of the portfolio bought at time t. An obvious consequence of the self-financing assumption **SF**, is that the change of value of the portfolio between $t-1$ and t is only due to the change of value of the underlying assets:

$$V_t - V_{t-1} = \sum_{i=1}^{m} \mu_{i,t-1}(p_{i,t} - p_{i,t-1}).$$

Another consequence is that the weights $a_{i,t-1}$ sum up to 1, that is $a_{0,t-1} = 0$. Thus, under **SF** we have $\frac{V_t}{V_{t-1}} - 1 \approx r_t$, where

$$r_t = \sum_{i=1}^{m} a_{i,t-1} y_{it} = \mathbf{a}'_{t-1} \boldsymbol{y}_t, \qquad a_{i,t-1} = \frac{\mu_{i,t-1} p_{i,t-1}}{\sum_{j=1}^{m} \mu_{j,t-1} p_{j,t-1}}, \tag{2.4}$$

for $i = 1, \dots, m$, and $\mathbf{a}_{t-1} = (a_{1,t-1}, \dots, a_{m,t-1})'$. A portfolio is usually called *crystallized* when the number of units of each asset is time independent, that is for each $i = 1, \dots, m$, $\mu_{i,t-1} = \mu_i$ for all t. We will call *static* a portfolio with fixed proportion in value of each return, that is for each $i = 1, \dots, m$, $a_{i,t-1} = a_i$ for all t.

2.2 Conditional VaR of a dynamic portfolio

The *conditional* VaR of the portfolio's return process (r_t) at risk level $\alpha \in (0,1)$, denoted by $\mathrm{VaR}_{t-1}^{(\alpha)}(r)$, is defined by

$$P_{t-1}\left[r_t < -\mathrm{VaR}_{t-1}^{(\alpha)}(r)\right] = \alpha, \tag{2.5}$$

where P_{t-1} denotes the historical distribution conditional on $\{\boldsymbol{p}_u, u < t\}$. In view of (2.1) and (2.4), the portfolio's return satisfies

$$r_t = \mathbf{a}_{t-1}'\boldsymbol{m}_t(\boldsymbol{\vartheta}_0) + \mathbf{a}_{t-1}'\boldsymbol{\Sigma}_t(\boldsymbol{\vartheta}_0)\boldsymbol{\eta}_t, \tag{2.6}$$

from which it follows that its conditional VaR at level α is given by[1]

$$\mathrm{VaR}_{t-1}^{(\alpha)}(r) = -\mathbf{a}_{t-1}'\boldsymbol{m}_t(\boldsymbol{\vartheta}_0) + \mathrm{VaR}_{t-1}^{(\alpha)}\left(\mathbf{a}_{t-1}'\boldsymbol{\Sigma}_t(\boldsymbol{\vartheta}_0)\boldsymbol{\eta}_t\right). \tag{2.7}$$

The VaR formula can be simplified if we assume that the errors $\boldsymbol{\eta}_t$ have a spherical distribution, that is, for any non-random vector $\boldsymbol{\lambda} \in \mathbb{R}^m$, $\boldsymbol{\lambda}'\boldsymbol{\eta}_t \stackrel{d}{=} \|\boldsymbol{\lambda}\|\eta_{1t}$, where $\|\cdot\|$ denotes the euclidian norm on $\mathbb{R}^m$, η_{it} denotes the i-th component of $\boldsymbol{\eta}_t$, and $\stackrel{d}{=}$ stands for the equality in distribution. Under the sphericity assumption we have

$$\mathrm{VaR}_{t-1}^{(\alpha)}(r) = -\mathbf{a}_{t-1}'\boldsymbol{m}_t(\boldsymbol{\vartheta}_0) + \left\|\mathbf{a}_{t-1}'\boldsymbol{\Sigma}_t(\boldsymbol{\vartheta}_0)\right\| \mathrm{VaR}^{(\alpha)}(\eta), \tag{2.8}$$

where $\mathrm{VaR}^{(\alpha)}(\eta)$ is the (marginal) VaR of η_{1t}.

2.3 Multivariate approaches

Multivariate approaches require specifying the first two conditional moments in (2.2) of the vector of individual returns. While the conditional mean is generally modelled using a small-order AR process, there are plenty of GARCH-type specifications for the conditional variance. See for instance Bauwens, Laurent and Rombouts (2006), Francq and Zakoïan (2010, Chapter 11) or Bauwens, Hafner and Laurent (2012) for presentations of the most commonly used specifications.

2.3.1 Conditional VaR estimation under conditional ellipticity

Under the sphericity assumption, a natural strategy for estimating the conditional VaR of a portfolio is to estimate $\boldsymbol{\vartheta}_0$ by some consistent estimator $\widehat{\boldsymbol{\vartheta}}_n$ in a first step, to extract the residuals and to estimate $\mathrm{VaR}^{(\alpha)}(\eta)$ in a second step.

[1] The presence of the sign "$-$" in this formula comes from the fact that the VaR is defined in terms of returns instead of loss variables.

An estimator of the conditional VaR at level α accounting for the conditional ellipticity is thus

$$\widehat{\mathrm{VaR}}^{(\alpha)}_{S,t-1}(r) = -\mathbf{a}'_{t-1}\widetilde{\boldsymbol{m}}_t(\widehat{\boldsymbol{\vartheta}}_n) + \|\mathbf{a}'_{t-1}\widetilde{\boldsymbol{\Sigma}}_t(\widehat{\boldsymbol{\vartheta}}_n)\|\xi_{n,1-2\alpha}, \tag{2.9}$$

where $\xi_{n,1-2\alpha}$ is the empirical $(1-2\alpha)$-quantile of the residuals $\widehat{\boldsymbol{\eta}}_t = \widetilde{\boldsymbol{\Sigma}}_t^{-1}(\widehat{\boldsymbol{\vartheta}}_n)\{\boldsymbol{y}_t - \widetilde{\boldsymbol{m}}_t(\widehat{\boldsymbol{\vartheta}}_n)\}$. Here $\widetilde{\boldsymbol{m}}_t(\widehat{\boldsymbol{\vartheta}}_n)$ and $\widetilde{\boldsymbol{\Sigma}}_t(\widehat{\boldsymbol{\vartheta}}_n)$ denote the estimated conditional mean and variance of $\boldsymbol{y}_t$ based on initial values $\widetilde{\boldsymbol{y}}_i$ for $i \leq 0$.

Francq and Zakoian (2017) derived, under appropriate assumptions, the asymptotic joint distribution of $\widehat{\boldsymbol{\vartheta}}_n$ and $\xi_{n,1-2\alpha}$ from which confidence intervals for the VaR can be deduced.

2.4 Conditional VaR estimation without the sphericity assumption

The Filtered Historical Simulation (FHS) approach (see Barone-Adesi, Giannopoulos and Vosper (1999), Mancini and Trojani (2011) and the references therein) does not require any symmetry assumption. It relies on estimating the conditional quantile of a linear combination of the components of the innovation, where the coefficients depend on both the model parameter and the past returns. Indeed, the conditional VaR of the portfolio return is

$$\mathrm{VaR}^{(\alpha)}_{t-1}(r) = \mathrm{VaR}^{(\alpha)}_{t-1}\left\{\boldsymbol{b}_t(\boldsymbol{\vartheta}_0) + \boldsymbol{c}'_t(\boldsymbol{\vartheta}_0)\boldsymbol{\eta}_t\right\}$$

where $\boldsymbol{b}_t(\boldsymbol{\vartheta}) = \mathbf{a}'_{t-1}\boldsymbol{m}_t(\boldsymbol{\vartheta})$ and $\boldsymbol{c}'_t(\boldsymbol{\vartheta}) = \mathbf{a}'_{t-1}\boldsymbol{\Sigma}_t(\boldsymbol{\vartheta})$. A natural estimator is thus

$$\widehat{\mathrm{VaR}}^{(\alpha)}_{FHS,t-1}(r) = -q_\alpha\left(\{\boldsymbol{b}_t(\widehat{\boldsymbol{\vartheta}}_n) + \boldsymbol{c}'_t(\widehat{\boldsymbol{\vartheta}}_n)\widehat{\boldsymbol{\eta}}_s, \quad 1 \leq s \leq n\}\right).$$

Based on this estimator and asymptotic arguments, Francq and Zakoian (2017) proposed CIs for the conditional VaR at time t of the portfolio return.

3 Univariate approaches

An obvious alternative to the multivariate approaches is to estimate a univariate GARCH model on the series of portfolio returns. We will see that this approach, which can be called "naive", is actually invalid in general, due to the fact that the return's portfolio is a time-varying combination of the individual returns. Instead, a reasonable approach consists in reconstituting a "virtual portfolio", whose returns are built using the current composition of the portfolio. We start by considering the naive approach.

3.1 Invalidity of the naive univariate approach

For simplicity, we consider a crystallized portfolio, with weight μ_i and initial price p_{i0} for the asset $i \in \{1, \ldots, m\}$. The composition $\boldsymbol{a}_{t-1}$ of such a portfolio is non stationary, in general. Indeed, we have

$$\log(a_{i,t}/a_{j,t}) = \log(\mu_i p_{i,0}/\mu_j p_{j,0}) + \sum_{k=1}^{t} D_{i,j,k}, \quad D_{i,j,k} = y_{i,k} - y_{j,k},$$

and $(\sum_{k=1}^{t} D_{i,j,k})_{t \geq 1}$ is a non stationary integrated process of order 1 under general assumptions.[2] More precisely, the following lemma shows that, with probability tending to one, the composition $\boldsymbol{a}_{t-1}$ of the portfolio converges to the set of the vectors $\boldsymbol{e}_i$ of the canonical basis (corresponding to single-asset portfolios): $P(\boldsymbol{a}_{t-1} \in \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}) \to 1$ as $t \to \infty$.

Lemma 3.1. *Consider a process $(D_k)_{k \geq 1}$. Assume that there exist real sequences $a_n > 0$ and b_n, both tending to zero, such that*

$$Z_n := a_n \sum_{k=1}^{n} D_k + b_n \xrightarrow{\mathcal{L}} Z \text{ as } n \to \infty, \tag{3.1}$$

for some random variable Z, whose cdf is continuous at 0 and such that $p = P(Z > 0) \in (0, 1)$. For any $c > 0$, we have $P(\sum_{k=1}^{n} D_k > c) \to p$ and $P(\sum_{k=1}^{n} D_k < -c) \to 1 - p$ as $n \to \infty$.

Note that a generalized central limit theorem of the form (3.1) holds for any iid sequence (D_k) whenever the distribution of D_k belongs to the domain of attraction of Z, which then follows a stable distribution. If the assumptions of Lemma 3.1 hold with $D_k = D_{i,j,k}$ for any pair (i, j), with $i \neq j$, then all the ratios $a_{i,t}/a_{j,t}$ are arbitrarily close to either 1 or 0 with probability tending to 1 as $t \to \infty$. In that case, the composition $\boldsymbol{a}_{t-1}$ tends to be totally undiversified, but is not always close to the same single-asset composition $\boldsymbol{e}_i$. If the dynamics of the individual returns ϵ_{it} are not identical, the dynamics of the return r_t will be time-varying, and the naive method based on a fixed stationary GARCH model is likely to produce poor results.

Simulation experiments reported in Section 5 confirm that for crystallized portfolios, the naive approach behaves badly due to the non stationarity of the univariate returns r_t. Of course, for static portfolios the non stationarity issue vanishes, but such portfolios may be considered as artificial. The next section studies a remedy to the non stationarity issue, while keeping the univariate framework.

[2] By the Chung-Fuks theorem, this is the case when $\boldsymbol{y}_t$ is iid with zero mean and a non-singular covariance matrix $\boldsymbol{\Sigma}$. The non stationarity of the process also holds, for instance, if the sequence $(D_{i,j,k})_k$ is mixing and nondegenerated.

3.2 The VHS univariate approach

At time t_0, given the current portfolio composition $\boldsymbol{a}_{t_0-1} = \boldsymbol{x}$, say, and the series of individual returns, we construct a series of virtual returns

$$r_t^*(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{y}_t, \qquad t \in \mathbb{Z}.$$

Note that, in general, $r_t^*(\boldsymbol{x}) \neq r_t$ because the composition of the (non virtual) portfolio is time varying ($\boldsymbol{a}_{t-1} \neq \boldsymbol{x}$, in general, for $t \neq t_0$). Given the stationarity of $(\boldsymbol{y}_t)$, it is clear that the series of virtual returns $\{r_t^*(\boldsymbol{x})\}$ is also stationary, with conditional moments

$$E_{t-1}\{r_t^*(\boldsymbol{x})\} =: \mu_t(\boldsymbol{x}), \qquad \mathrm{var}_{t-1}\{r_t^*(\boldsymbol{x})\} =: \sigma_t^2(\boldsymbol{x}),$$

where $E_{t-1}(X) = E(X \mid r_s^*(\boldsymbol{x}), s < t)$ for any variable X, and the variance is defined accordingly. Thus, $r_t^*(\boldsymbol{x})$ follows a model of the form

$$r_t^*(\boldsymbol{x}) = \mu_t(\boldsymbol{x}) + \sigma_t(\boldsymbol{x})u_t, \quad \text{where } E_{t-1}(u_t) = 0 \text{ and } \mathrm{var}_{t-1}(u_t) = 1. \tag{3.2}$$

Noting that $r_{t_0} = r_{t_0}^*(\boldsymbol{a}_{t_0-1})$, the conditional VaR at time t_0 thus satisfies

$$\mathrm{VaR}_{t_0-1}^{*(\alpha)}(r_{t_0}) = -\mu_{t_0}(\boldsymbol{a}_{t_0-1}) + \sigma_{t_0}(\boldsymbol{a}_{t_0-1})\mathrm{VaR}_{t_0-1}^{*(\alpha)}(u_{t_0}) \tag{3.3}$$

where $\mathrm{VaR}_{t-1}^{*(\alpha)}(X)$ is the VaR of X at level α conditional on $(r_s^*(\boldsymbol{x}), s < t)$.

Note that the martingale difference (u_t) may not be iid, as the following example illustrates.

Example 3.1. Consider the bivariate ARCH(1) process, defined as the stationary non anticipative solution of the model

$$\boldsymbol{y}_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \boldsymbol{\Sigma}_t \boldsymbol{\eta}_t, \quad \boldsymbol{\Sigma}_t = \begin{pmatrix} \sigma_{1t}^2 := \omega_1 + \alpha_{11} y_{1,t-1}^2 + \alpha_{12} y_{2,t-1}^2 & 0 \\ 0 & \sigma_{2t}^2 := \omega_2 + \alpha_{21} y_{1,t-1}^2 \end{pmatrix},$$

where $\boldsymbol{\eta}_t$ iid $(\boldsymbol{0}, \boldsymbol{I})$, and assuming that the components η_{1t} and η_{2t} are independent. Let the return of the portfolio which is fully invested in the first asset, that is, $r_t = (1,0)\boldsymbol{y}_t = y_{1t}$. Denote by $\mathcal{F}_{1t}$ the σ-field generated by $\{y_{1u}, u \leq t\}$. We have $E(r_t|\mathcal{F}_{1,t-1}) = 0$ and

$$\begin{aligned} E(r_t^2|\mathcal{F}_{1,t-1}) &= E(\sigma_{1t}^2|\mathcal{F}_{1,t-1}) = \omega_1 + \alpha_{11} y_{1,t-1}^2 + \alpha_{12} E(y_{2,t-1}^2|\mathcal{F}_{1,t-1}) \\ &= \omega_1 + \alpha_{11} y_{1,t-1}^2 + \alpha_{12}\sigma_{2,t-1}^2 E(\eta_{2,t-1}^2|\mathcal{F}_{1,t-1}) \\ &= \omega_1 + \alpha_{11} y_{1,t-1}^2 + \alpha_{12}\sigma_{2,t-1}^2 := \sigma_t^2. \end{aligned}$$

It follows that (r_t) satisfies the model $r_t = \sigma_t u_t$, where

$$u_t = \frac{\sigma_{1t}}{\sigma_t}\eta_{1t} = \left(1 + \frac{\alpha_{12}\sigma^2_{2,t-1}(\eta^2_{2,t-1} - 1)}{\omega_1 + \alpha_{11}y^2_{1,t-1} + \alpha_{12}\sigma^2_{2,t-1}}\right)^{1/2}\eta_{1t}.$$

It is then clear that $(u_t, \mathcal{F}_{1,t})$ is a martingale difference but (u_t) is generally not iid (except when $\alpha_{12} = 0$ or $\eta^2_{2,t}$ is degenerated).

Even in the simple previous example, the conditional quantile $\mathrm{VaR}^{*(\alpha)}_{t_0-1}(u_{t_0})$ popping up in (3.3) cannot be explicitly computed. Whether or not this quantity could be estimated nonparametrically is beyond the scope of this paper. Instead, we consider a "hybrid" VaR defined by

$$\mathrm{VaR}^{(\alpha)}_{H,t_0-1}(r_{t_0}) = -\mu_{t_0}(\boldsymbol{a}_{t_0-1}) + \sigma_{t_0}(\boldsymbol{a}_{t_0-1})\mathrm{VaR}^{(\alpha)}(u) \tag{3.4}$$

where $\mathrm{VaR}^{(\alpha)}(u)$ is the marginal VaR of u_t at level α. An estimator of $\mathrm{VaR}^{(\alpha)}_{H,t_0-1}(r)$ is obtained as follows: given $\boldsymbol{a}_{t_0-1} = \boldsymbol{x}$,

Step 1: Compute the virtual historical returns $r^*_t(\boldsymbol{x})$ for $t = 1, \dots, n$.

Step 2: Estimate $\mu_t(\boldsymbol{x})$ and $\sigma_t(\boldsymbol{x})$. Denote by $\hat{\mu}_t(\boldsymbol{x})$ and $\hat{\sigma}_t(\boldsymbol{x})$ the resulting estimators, and by $\hat{u}_t = \{r^*_t(\boldsymbol{x}) - \hat{\mu}_t(\boldsymbol{x})\}/\hat{\sigma}_t(\boldsymbol{x})$ the residuals.

Step 3: Compute the α-quantile $\xi_{n,\alpha}$ of $\{\hat{u}_s, 1 \leq s \leq n\}$ and define an estimator of $\mathrm{VaR}^{(\alpha)}_{t_0-1}(r)$ as

$$\widehat{\mathrm{VaR}}^{(\alpha)}_{VHS,t_0-1}(r) = -\hat{\mu}_{t_0}(\boldsymbol{x}) - \hat{\sigma}_{t_0}(\boldsymbol{x})\xi_{n,\alpha}. \tag{3.5}$$

This procedure is particularly appropriate for large portfolios, when the large dimension of the vector of underlying assets precludes–or at least formidably complicates– estimation of multivariate volatility models. Moreover, the following example shows that for large portfolios a univariate GARCH model is a reasonable assumption for the virtual returns.

Example 3.2. Suppose that m is large and that the vector of log-returns is driven by a vector $\boldsymbol{f}_t$ of K factors (with $K \ll m$) as

$$\boldsymbol{y}_t = \boldsymbol{\beta}\boldsymbol{f}_t + \boldsymbol{\epsilon}_t$$

where $\boldsymbol{\beta}$ is a $m \times K$ matrix, $\mathrm{var}_{t-1}(\boldsymbol{f}_t) = \boldsymbol{F}_t$ is a full-rank matrix and $\mathrm{var}_{t-1}(\boldsymbol{\epsilon}_t) = \boldsymbol{\Sigma}_t$ is a diagonal matrix. With a composition fixed to $\boldsymbol{x}$, the virtual portfolio's returns thus satisfy

$$r^*_t = \boldsymbol{x}'\boldsymbol{\beta}\boldsymbol{f}_t + \boldsymbol{x}'\boldsymbol{\epsilon}_t.$$

Suppose that the portfolio is well-diversified so $x_i = O(1/m)$ for $i = 1, \dots, m$. The conditional variance of the error term is given by $\mathrm{var}_{t-1}(\boldsymbol{x}'\boldsymbol{\epsilon}_t) = \boldsymbol{x}'\boldsymbol{\Sigma}^2_t\boldsymbol{x}$ and is of order $1/m$ in probability as

m goes to infinity, while $\mathrm{var}_{t-1}(\boldsymbol{x}'\boldsymbol{\beta}\boldsymbol{f}_t) = \boldsymbol{x}'\boldsymbol{\beta}\boldsymbol{F}_t\boldsymbol{\beta}'\boldsymbol{x} = O_P(1)$ and does not vanish as m increases under appropriate assumptions. [3] It follows that $r_t^* \approx \boldsymbol{x}'\boldsymbol{\beta}\boldsymbol{f}_t$. If now $K = 1$ and the (real-valued) factor f_t is the solution of a GARCH model, the process $\boldsymbol{x}'\boldsymbol{\beta}f_t$ will follow the same model up to a change of scale. It is therefore natural to fit a GARCH model for the virtual returns under these assumptions when m is large.

4 Asymptotic properties of the VHS approach

To obtain asymptotic properties of the VHS procedure, we make the following parametric assumptions on Model (3.2). For simplicity, we consider the model without conditional mean, that is $\mu_t(\boldsymbol{x}) = 0$. For some (known) function $\sigma : \mathbb{R}^\infty \times \Theta \to (0, \infty)$, let

$$\sigma_t(\boldsymbol{x}; \boldsymbol{\theta}) = \sigma(r_{t-1}^*(\boldsymbol{x}), r_{t-2}^*(\boldsymbol{x}), \dots; \boldsymbol{\theta}), \tag{4.1}$$

where $\boldsymbol{\theta}_0 = \boldsymbol{\theta}_0(\boldsymbol{x})$ is the true value of the finite dimensional parameter $\boldsymbol{\theta}$, belonging to some compact set Θ. To alleviate notations, we will denote the virtual returns by $\epsilon_t := r_t^*(\boldsymbol{x})$. Model (3.2) thus reduces to

$$\epsilon_t = \sigma_t u_t, \quad \sigma_t = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots; \boldsymbol{\theta}). \tag{4.2}$$

Recall that iidness of the sequence (u_t) is not a natural assumption in our framework. To study the asymptotic properties of the VHS estimator, we introduce the following additional assumptions. Let $D_t(\boldsymbol{\theta}) = \sigma_t^{-1}(\boldsymbol{\theta})\partial\sigma_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$.

A1: The sequence (u_t) is stationary and ergodic, with $E|u_t|^{4+\nu} < \infty$ for some $\nu > 0$, and mixing coefficients $\{\alpha(h)\}_{h\geq 0}$ satisfying

$$\sum_{h=1}^{\infty} h^{r^*}\alpha^{(\nu-\epsilon)/(4+\nu-\epsilon)}(h) < \infty \text{ for some } r^* > \frac{\kappa(4+2\nu)}{\nu - \kappa(4+2\nu)} \text{ and some } \epsilon \in (0, \nu).$$

Denoting by $\mathcal{F}_{t-1}$ the sigma-field generated by $\{u_u, u < t\}$, suppose that $E(u_t \mid \mathcal{F}_{t-1}) = 0$ and $E(u_t^2 \mid \mathcal{F}_{t-1}) = 1$. Assume that the conditional distribution of u_t given $\mathcal{F}_{t-1}$ has a density f_{t-1} such that $f_{t-1}(\xi_\alpha) > 0$ a.s. and $E\sup_{\xi\in V(\xi_\alpha)} f_{t-1}^4(\xi) < \infty$ for some neighborhood $V(\xi_\alpha)$ of ξ_α. Assume also that this density is continuous at ξ_α uniformly in $\mathcal{F}_{t-1}$, in the sense that for sufficiently small $\varepsilon > 0$, there exists a stationary and ergodic sequence (K_t) such that

[3] For instance if the matrix $\boldsymbol{\beta}$ does not contain too many many zeroes or, more precisely, if at least one column $\boldsymbol{\beta}_j$ of $\boldsymbol{\beta}$ is such that $\liminf |\boldsymbol{x}'\boldsymbol{\beta}_j| > 0$ as $m \to \infty$.

$K_{t-1} \in \mathcal{F}_{t-1}$ and

$$\sup_{x\in[\xi_\alpha-\varepsilon,\xi_\alpha+\varepsilon]} |f_{t-1}(x) - f_{t-1}(\xi_\alpha)| \leq K_{t-1}\varepsilon$$

with $EK_t^4 < \infty$ a.s.

A2: (ϵ_t) is a strictly stationary and ergodic solution of (4.2), and there exists $s > 0$ such that $E|\epsilon_1|^s < \infty$.

A3: Assume that there exists a sequence D_{t,T_n} such that $D_t = D_{t,T_n} + \widetilde{D}_{t,T_n}$, where $T_n \to \infty$ and $T_n = O(n^\kappa)$ for some $\kappa \in [0, \nu/4(2+\nu))$ (with $E|u_t|^{4+\nu} < \infty$) and D_{t,T_n} is measurable with respect to $u_{t-1}, \ldots, u_{t-T_n}$, and for any $r \geq 0$

$$E\|D_t\|^r < \infty, \qquad \sup_{n\geq 1} E\|D_{t,T_n}\|^r < \infty, \qquad \widetilde{D}_{t,T_n} = o_P(1) \text{ as } n \to \infty.$$

A4: For some $\underline{\omega} > 0$, almost surely, $\sigma_t(\boldsymbol{\theta}) \in (\underline{\omega}, \infty]$ for any $\boldsymbol{\theta} \in \Theta$. Moreover, for $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta$, we have $\sigma_t(\boldsymbol{\theta}_1) = \sigma_t(\boldsymbol{\theta}_2)$ *a.s.* if and only if $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$.

A5: There exist a random variable C_1 measurable with respect to $\{\epsilon_u, u < 0\}$ and a constant $\rho \in (0,1)$ such that $\sup_{\boldsymbol{\theta}\in\Theta} |\sigma_t(\boldsymbol{\theta}) - \widetilde{\sigma}_t(\boldsymbol{\theta})| \leq C_1\rho^t$.

A6: The function $\boldsymbol{\theta} \mapsto \sigma(x_1, x_2, \ldots; \boldsymbol{\theta})$ has continuous second-order derivatives, and

$$\sup_{\boldsymbol{\theta}\in\Theta} \left\| \frac{\partial\sigma_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} - \frac{\partial\widetilde{\sigma}_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} \right\| + \left\| \frac{\partial^2\sigma_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} - \frac{\partial^2\widetilde{\sigma}_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \right\| \leq C_1\rho^t,$$

where C_1 and ρ are as in **A5**.

A7: There exists a neighborhood $V(\boldsymbol{\theta}_0)$ of $\boldsymbol{\theta}_0$ such that

$$\sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)} \left\| \frac{1}{\sigma_t(\boldsymbol{\theta})} \frac{\partial\sigma_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} \right\|^4, \qquad \sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)} \left\| \frac{1}{\sigma_t(\boldsymbol{\theta})} \frac{\partial^2\sigma_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \right\|^2, \qquad \sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)} \left| \frac{\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta})} \right|^{2\delta}$$

have finite expectations.

For particular volatility models, some of the assumptions can be simplified as the following lemma shows.

Lemma 4.1. *For the standard GARCH(1,1) model*

$$\epsilon_t = \sigma_t u_t, \quad \sigma_t^2 = \omega_0 + \alpha_0\epsilon_{t-1}^2 + \beta_0\sigma_{t-1}^2, \qquad (u_t) \overset{iid}{\sim} (0,1), \quad \omega_0 > 0, \alpha_0 > 0, \beta_0 > 0, \tag{4.3}$$

Assumptions **A1-A7** *reduce to: i)* $E\log(\alpha_0\eta_t^2 + \beta_0) < 0$*; ii)* η_t^2 *has a non-degenerate distribution with* $E\eta_t^4 < \infty$*; iv)* $\Theta = \{(\omega, \alpha, \beta)\}$ *is a compact subset of* $(0,\infty)^3$ *such that, for all* $\theta \in \Theta$*,* $\omega > \underline{\omega}$ *for some* $\underline{\omega} > 0$ *and* $\beta < 1$*.*

We are now in a position to state our main result.

Theorem 4.1. *Assume* $\xi_\alpha < 0$. *Let* **A1-A7** *hold. Then*

$$\begin{pmatrix} \sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) \\ \sqrt{n}(\xi_\alpha - \xi_{n,\alpha}) \end{pmatrix} \overset{\mathcal{L}}{\to} \mathcal{N}(0, \Sigma_\alpha), \qquad \Sigma_\alpha = \begin{pmatrix} J^{-1}\boldsymbol{S}^{11}J^{-1} & \Lambda_\alpha \\ \Lambda_\alpha' & \zeta_\alpha \end{pmatrix},$$

where $J = E(D_t D_t')$ *with* $D_t = D_t(\boldsymbol{\theta}_0)$, *and*

$$\begin{aligned}
\Lambda_\alpha &= \frac{\xi_\alpha}{Ef_{t-1}(\xi_\alpha)} J^{-1}\boldsymbol{S}^{11}J^{-1}\Omega + \frac{1}{2Ef_{t-1}(\xi_\alpha)} J^{-1}\boldsymbol{S}_\alpha^{12}, \qquad \Omega = E\{f_{t-1}(\xi_\alpha)D_t\}, \\
\zeta_\alpha &= \frac{1}{\{Ef_{t-1}(\xi_\alpha)\}^2} \left(\xi_\alpha^2 \Omega' J^{-1}\boldsymbol{S}^{11}J^{-1}\Omega + \xi_\alpha \Omega' J^{-1}\boldsymbol{S}_\alpha^{12} + \boldsymbol{S}_\alpha^{22}\right), \\
\boldsymbol{S}^{11} &= E\left[\left\{E\left(u_t^4 \mid \mathcal{F}_{t-1}\right) - 1\right\} D_t D_t'\right], \qquad \boldsymbol{S}_\alpha^{22} = \sum_{h=-\infty}^{\infty} cov\left(\mathbf{1}_{\{u_t < \xi_\alpha\}}, \mathbf{1}_{\{u_{t-h} < \xi_\alpha\}}\right), \\
\boldsymbol{S}_\alpha^{12} &= \boldsymbol{S}_\alpha^{21'} = \sum_{h=0}^{\infty} cov\left\{(u_t^2 - 1)D_t, \mathbf{1}_{\{u_{t+h} < \xi_\alpha\}}\right\}.
\end{aligned}$$

Let $\widehat{\boldsymbol{\Sigma}}_\alpha(\boldsymbol{x})$ denote a consistent estimator of $\boldsymbol{\Sigma}_\alpha(\boldsymbol{x})$. By the delta method, an approximate $(1-\alpha_0)\%$ confidence interval (CI) for $\mathrm{VaR}_t(\alpha)$ has bounds given by

$$\sigma_t\{\boldsymbol{x}; \hat{\boldsymbol{\theta}}_n(\boldsymbol{x})\}\xi_{n,1-2\alpha}^{|u|}(\boldsymbol{x}) \pm \frac{1}{\sqrt{n}}\Phi^{-1}(1-\alpha_0/2)\left\{\boldsymbol{\delta}_{t-1}'(\boldsymbol{x})\widehat{\boldsymbol{\Sigma}}_\alpha(\boldsymbol{x})\boldsymbol{\delta}_{t-1}(\boldsymbol{x})\right\}^{1/2}, \tag{4.4}$$

where

$$\boldsymbol{\delta}_{t-1}'(\boldsymbol{x}) = \left(\frac{\partial \sigma_t(\boldsymbol{x}; \hat{\boldsymbol{\theta}}_n(\boldsymbol{x}))}{\partial \boldsymbol{\theta}} \xi_{n,1-2\alpha}^{|u|}(\boldsymbol{x}) \qquad \sigma_t\{\boldsymbol{x}; \hat{\boldsymbol{\theta}}_n(\boldsymbol{x})\} \right).$$

5 Numerical illustrations

The first part of the section illustrates the invalidity of the naive approach. The second part presents a selection of Monte-Carlo experiments aiming at studying the performance of the previous approaches in finite sample. Real data examples will be presented in the third part.[4]

5.1 Non stationarity of the portfolio's return

For simplicity, we consider a crystallized equally weighted portfolio of 3 assets (of initial price $p_{i0} = 1000$) $V_t = \sum_{i=1}^3 p_{it}$. Thus, the return portfolio composition is time varying, with coefficients

[4] The code and data used in the paper are available on the web site
http://perso.univ-lille3.fr/~cfrancq/Christian-Francq/VaRPortfolio.html

$\mathbf{a}_{t-1} = (a_{1,t-1}, a_{2,t-1}, a_{3,t-1})'$ and $a_{i,t-1} = p_{i,t-1}/\sum_{j=1}^{3} p_{j,t-1}$. Assume that the vector of the log-returns is iid, centered, with variance $\mathrm{Var}(\boldsymbol{\epsilon}_t) = \boldsymbol{\Sigma}^2 = \boldsymbol{DRD}$, with

$$\boldsymbol{D} = \begin{pmatrix} 0.01 & 0 & 0 \\ 0 & 0.02 & 0 \\ 0 & 0 & 0.04 \end{pmatrix}, \quad \boldsymbol{R} = \begin{pmatrix} 1 & -0.855 & 0.855 \\ -0.855 & 1 & -0.810 \\ 0.855 & -0.810 & 1 \end{pmatrix}.$$

The composition $\boldsymbol{a}_{t-1}$ of the portfolio is plotted in Figure 2. As we have seen in Section 3.1, this vector is non stationary. More precisely, by the Chung-Fuks theorem, with increasing probability, the composition $\boldsymbol{a}_{t-1}$ of the portfolio is arbitrarily close to one of the three single-asset portfolios $(1,0,0), (0,1,0)$ and $(0,0,1)$.

It is thus non surprising to see that the univariate return series r_t plotted in Figure 1 exhibits some nonstationarity features, in particular marginal heteroscedasticity. The increased variance in the second part of the sample reflects the fact that the portfolio tends to be less and less diversified (see Figure 2).

However, because the series also presents conditional heteroscedasticity, we fitted a GARCH(1,1) model which corresponds to common practice. The parameters of this model are estimated online, starting from $t = 200$. We have $\widehat{\mathrm{VaR}}^{(\alpha)}_{FHS,t-1}(r) = -q_\alpha\left(\{\boldsymbol{a}'_{t-1}\boldsymbol{\epsilon}_1, \dots, \boldsymbol{a}'_{t-1}\boldsymbol{\epsilon}_{t-1}\}\right)$. These empirical quantiles were computed starting from $t = 150$. The spherical method, based on the estimation of $\boldsymbol{\Sigma}$, was computed on the same range of observations. Figure **??** displays the sample paths of the true conditional VaR as well as the 3 estimated VaRs. It can be seen that the spherical method converges faster to the true value than the FHS method. On the other hand, the univariate method fails to converge to the theoretical conditional VaR. This can be explained by the difference between the information sets (point iii) in Section **??**), and also by the non stationarity of the univariate series of portfolio returns. appropriate for this non stationary series. Figure **??** provides another simulation, including the VHS method. It is seen that, contrary to Figure **??**, the VHS estimator behaves like the spherical estimator. This is not surprising since the returns are iid in this setting.

5.2 Monte-Carlo experiments

A Monte Carlo study was conducted in order to compare the multivariate and univariate approaches in finite sample. We simulated m-multivariate factor models, with two GARCH factors of the form

$$f_{1t} = \sigma_{1t}\eta_{1t}, \quad f_{2t} = \sigma_{2t}\eta_{2t},$$

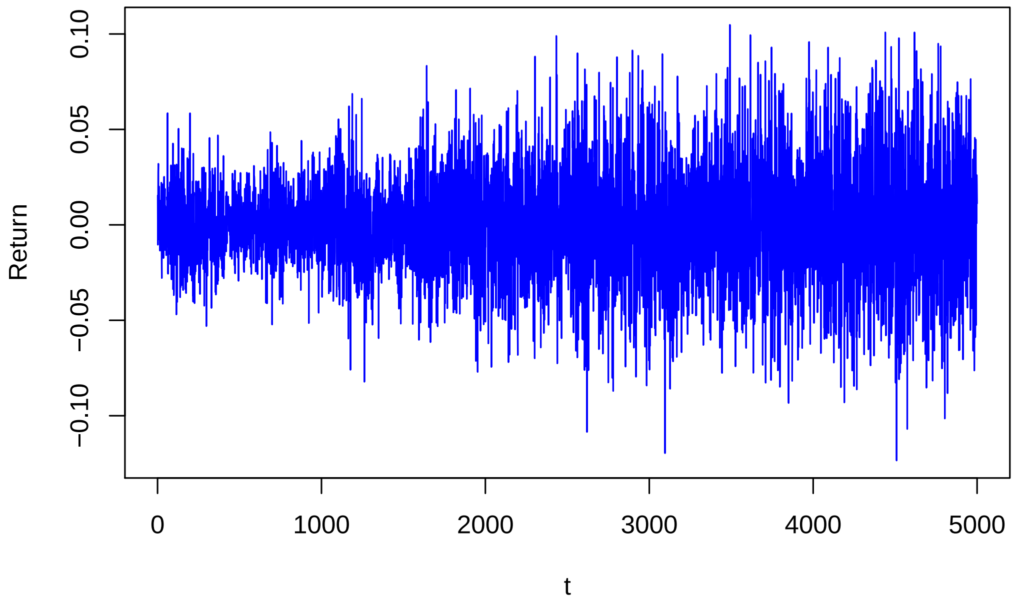

Figure 1: Returns of the crystallized portfolio.

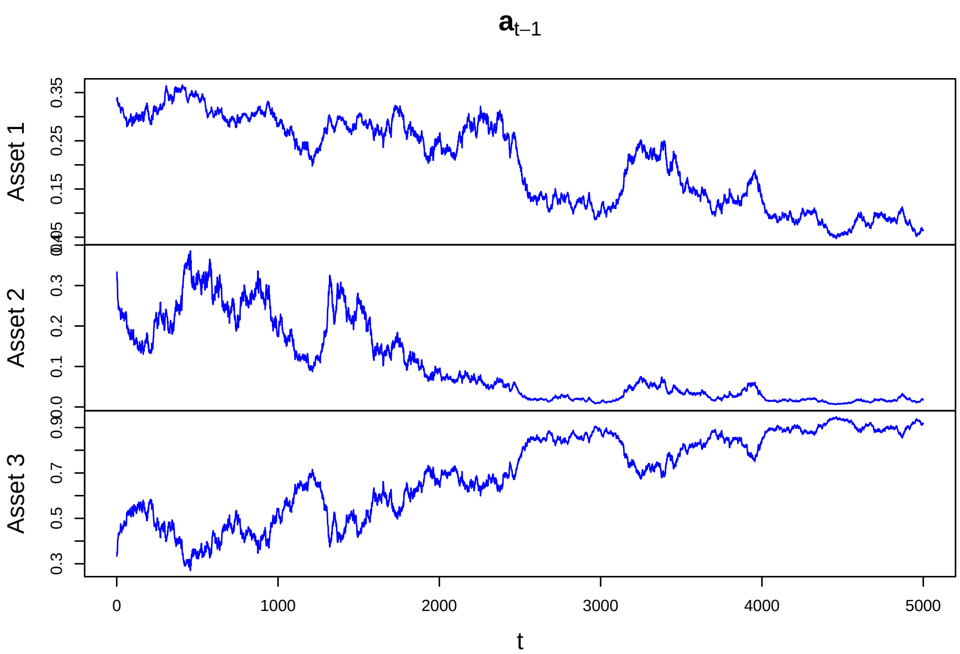

Figure 2: Time-varying composition of the crystallized portfolio.

where $(\eta_{1t})_t$ and $(\eta_{2t})_t$ are two independent sequences of iid $\mathcal{N}(0,1)$-distributed random variables. The volatilities follow standard GARCH$(1,1)$ equations of the form

$$\sigma_{it}^2 = \omega_i + \alpha_i f_{i,t-1}^2 + \beta_i \sigma_{i-1,t}^2.$$

We took $(\omega_1, \alpha_1, \beta_1) = (1, 0.09, 0.87)$ and $(\omega_2, \alpha_2, \beta_2) = (0.1, 0.7, 0.01)$, so the dynamics of the two factors be quite distinct. The even and odd components of our simulated factor model are respectively of the form

$$\epsilon_{2k,t} = f_{2t} + e_{2k,t}, \qquad \epsilon_{2k+1,t} = f_{1t} + e_{2k+1,t},$$

where $(e_{kt})_t$, for $k = 1, \ldots, m$, are idiosyncratic independent iid noises with law $\mathcal{N}(0, 0.1^2)$. To obtain a graphical comparison of the VaR estimates, we first simulated a trajectory of size $1,100$ of the factor model with $m = 4$. A crystalized portfolio of composition $(1/m, \ldots, 1/m)$ at time $t = 1$ has been considered. The four competing estimators of the 5% VaR_{t-1} at time $t = 1001$ were estimated on the basis on the first $1,000$ simulated values $\epsilon_1, \ldots, \epsilon_{t-1}$. Then, VaR at time $t = 1,002$ was estimated based on the past $1,000$ simulations $\epsilon_2, \ldots, \epsilon_{t-1}$. We continued until we obtained the last VaR estimations at time $t = 1,100$. Figure 3 shows that the estimates obtained by the Spherical, FHS and VHS methods are very close (actually, they are not distinguishable on the figure), whereas the estimates obtained by the naive method behave differently. This can be explained by the fact that the portfolio is crystallized but not static. In other words, even if the portfolio is constituted of an equal quantity of the m simulated assets, the return r_t is not a fixed average of the individuals returns ϵ_{kt} (see Figure 4).

In this first graphical illustration, the number of estimated VaRs is not sufficient to compare the methods by using formal backtests. We thus considered the same framework of GARCH estimations on rolling windows of length $1,000$, but the methods have been backtested of a longer period of length $2,000$. Moreover, in order to obtain a clearcut comparison between the naive method and the VHS method, the composition of the portfolio has to be highly time-varying. We thus simulated portfolios whose composition alternates as follows: we take an equal proportion of the returns of the even assets $\epsilon_{2k,t}$ during a period of length 100, and then we switch to an equal proportion of the odd assets $\epsilon_{2k+1,t}$ during another period of length 100. Table 1 summarizes the results of the 4 VaR estimation methods for $m = 2$ or $m = 4$. This simulation exercise is intensive since 2000 DCC-GARCH models must be estimated for each of the two multivariate methods, and 2000 univariate GARCH(1,1) models must be estimated for each of the univariate methods. The spherical and FHS methods become rapidly too time consuming when the number m of returns increases, because multivariate m-GARCH models have to be estimated. Interestingly, the numerical complexity of the univariate methods does not increase much with m, so that Table 1 reports results on portfolios of $m = 8$ and $m = 100$ assets for the univariate methods only.

Viol gives the relative frequency of violations (in %), the columns LRuc, LRind and LRcc give respectively the p-values of the the unconditional coverage test that the probability of violation is equal to the nominal 5% level, the independence test that the violations are independent and the conditional coverage test of Christoffersen (2003). Conclusions drawn from those backtests, which solely focus on the violations, are that all methods are validated on these experiments. It is necessary to introduce alternative statistics, related to the amount of violation, to compare the different approaches. The next column $\overline{\text{VaR}}$ provides the average VaR, while the column AV displays the average amount of violation, and the column ES gives the expected shorfall, that is the average loss when the VaR is violated: for each estimator $\widehat{VaR}_t$ of the conditional VaR, let

$$\text{AV} = \frac{\sum_{t=1}^n -(\epsilon_t + \widehat{VaR}_t)\mathbb{1}_{\epsilon_t < -\widehat{VaR}_t}}{\sum_{t=1}^n \mathbb{1}_{\epsilon_t < -\widehat{VaR}_t}}, \qquad \text{ES} = \frac{\sum_{t=1}^n -\epsilon_t \mathbb{1}_{\epsilon_t < -\widehat{VaR}_t}}{\sum_{t=1}^n \mathbb{1}_{\epsilon_t < -\widehat{VaR}_t}}.$$

These statistics clearly show that the naive approach is inefficient compared to its competitors. With this method, the amount of violation tends to be higher whatever the size m of the portfolio. For these statistics AV and ES, the VHS approach appears comparable to the multivariate methods when comparison is possible, that is when m is not too large. Alternative comparisons are provided by considering the loss function

$$\text{Loss} = \frac{1}{n}\sum_{t=1}^n -(\epsilon_t + \widehat{VaR}_t)(\alpha - \mathbb{1}_{\epsilon_t < -\widehat{VaR}_t}).$$

The last column of Table 1 reports, for each of the three non-naive methods, p-values of the Diebold-Mariano (1995) test for the null that the naive method produces the same loss against the alternative that it induces higher loss. The null is rejected in each situation, leading to the same conclusion as before: the naive method is outperformed by its three competitors when m is small, and is outperformed by the other univariate method when m is large.

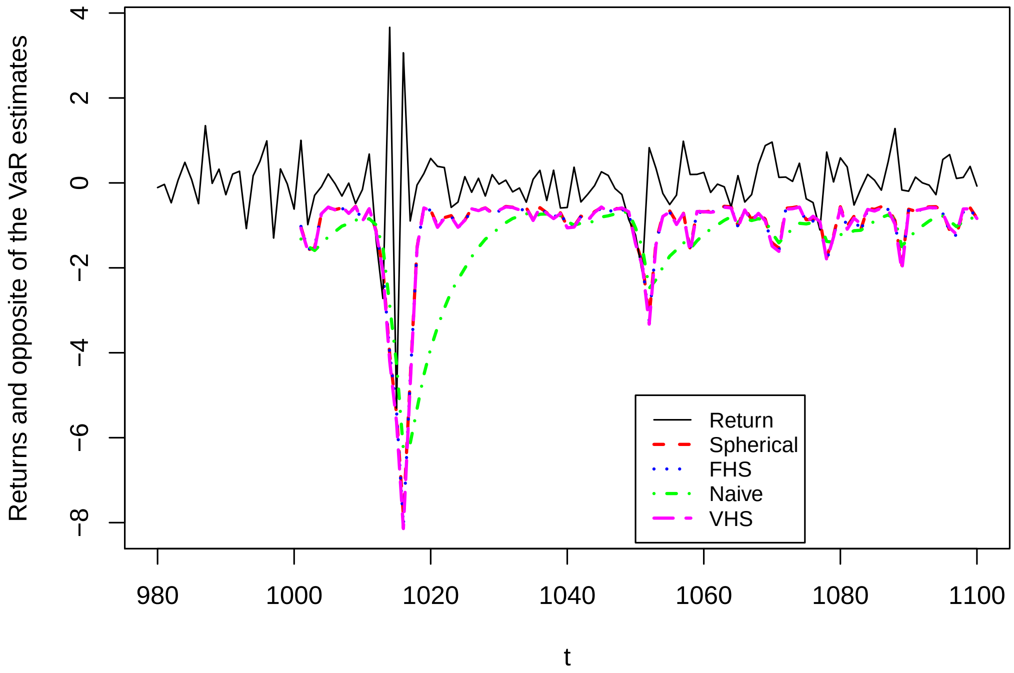

Figure 3: Comparison of 4 VaR at the horizon 1 for a crystallized portfolio of $m = 4$ simulated assets.

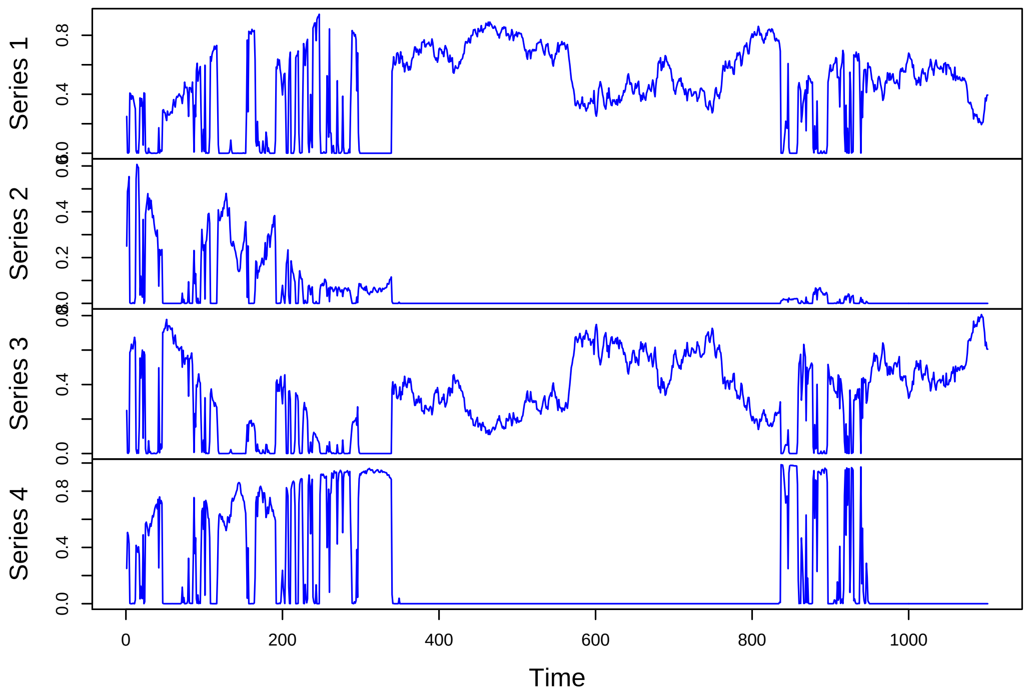

Figure 4: Time-varying composition of the return of the crystallized portfolio as function of the returns of the individual assets.

Table 1: Backtests of the 5%-VaR estimates

		Viol	LRuc	LRind	LRcc	$\overline{\text{VaR}}$	AV	ES	Loss	DM
$m = 2$	Naive	5.20	68.34	79.24	88.89	4.64	1.87	6.56	0.33	-
	VHS	5.55	26.71	72.63	50.81	4.26	1.07	5.30	0.27	5.e-10
	Spherical	5.30	54.19	86.71	81.87	4.28	1.10	5.49	0.27	1.e-09
	FHS	5.60	22.67	90.67	47.82	4.25	1.10	5.33	0.27	3.e-09
$m = 4$	Naive	5.55	26.71	12.95	17.12	4.51	1.90	6.08	0.33	-
	VHS	4.35	17.30	90.93	39.26	4.60	1.18	5.45	0.28	1.e-07
	Spherical	5.20	68.34	5.95	15.59	4.36	1.19	5.61	0.28	2.e-08
	FHS	4.30	14.15	87.19	33.50	4.60	1.19	5.55	0.28	2.e-07
$m = 8$	Naive	5.10	83.79	56.34	82.87	4.87	1.61	6.10	0.33	-
	VHS	5.50	31.23	69.02	55.44	4.44	1.05	5.38	0.28	1.e-08
$m = 100$	Naive	4.90	83.69	6.93	18.8	4.53	2.16	7.65	0.34	-
	VHS	5.25	61.07	81.42	85.46	4.56	1.09	5.61	0.29	6.e-09

Appendices

A Proofs

A.1 Proof of Lemma 3.1

We have

$$P\left(\sum_{k=1}^{n} D_k > c\right) = P(Z_n > c_n) = P(Z_n > 0) - \text{sign}(c_n)P(Z_n \in (0, c_n])$$

with $c_n = a_n c + b_n$ and obvious notation. We have $P(Z_n > 0) \to p$ and, for any $\varepsilon > 0$, there exists $c_\varepsilon > 0$ such that $\lim_{n\to\infty} P(Z_n \in (0, c_n]) \leq \lim_{n\to\infty} P(Z_n \in [-c_\varepsilon, c_\varepsilon]) \leq \varepsilon$. □

A.2 Proof of Lemma 4.1.

Letting $a(z) = \alpha_0 z^2 + \beta_0$, the volatility of a GARCH(1,1) model can be written as

$$\sigma_t^2(\boldsymbol{\theta}_0) = \sigma_{t,T_n}^2 + \widetilde{\sigma}_{t,T_n}^2, \qquad \sigma_{t,T_n}^2 = \omega_0 \left\{1 + \sum_{k=1}^{T_n} \prod_{i=1}^{k} a(u_{t-i})\right\}.$$

Note that, under the strict stationarity condition $E \log a(u_1) < 0$, we have

$$\widetilde{\sigma}_{t,T_n}^2 = \omega_0 \sum_{k=T_n+1}^{\infty} \prod_{i=1}^{k} a(u_{t-i}) \to 0 \text{ a.s. when } T_n \to \infty \tag{A.1}$$

We also set $\epsilon_t^2 = \epsilon_{t,T_n}^2 + \widetilde{\epsilon}_{t,T_n}^2$, where $\epsilon_{t,T_n} = u_t \sigma_{t,T_n}$ is a measurable function of $u_t, \ldots, u_{t-T_n}$. We thus have

$$\sigma_{t,T_n}^2 = \sum_{i=0}^{T_n-2} \beta_0^i \left(\omega_0 + \alpha_0 \epsilon_{t-i-1,T_n-i-1}^2\right) + \beta_0^{T_n-1} \sigma_{t-T_n+1,1}^2.$$

The first and second components of D_t are bounded, and thus can be handled easily. The last component of D_t has the form

$${}_\beta\sigma_t^2 := \frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2(\boldsymbol{\theta}_0)}{\partial \beta} = {}_\beta\sigma_{t,T_n}^2 + {}_\beta\widetilde{\sigma}_{t,T_n}^2, \qquad {}_\beta\sigma_{t,T_n}^2 = \frac{\sum_{i=1}^{T_n-2} i\beta_0^{i-1}\left(\omega_0 + \alpha_0 \epsilon_{t-i,T_n-i}^2\right)}{\sigma_{t,T_n}^2}.$$

Note that ${}_\beta\sigma_{t,T_n}^2$ is a measurable function of $u_{t-1}, \ldots, u_{t-T_n}$ and, using the inequality $x/(1+x) \leq x^s$ for any $x \geq 0$ and any $s \in (0, 1)$, we have

$${}_\beta\sigma_{t,T_n}^2 \leq \frac{1}{(1-\beta_0)^2} + \sum_{i=1}^{T_n-2} \frac{i\beta_0^{i-1}\alpha_0\epsilon_{t-i,T_n-i}^2}{\omega_0 + \beta_0^i \alpha_0 \epsilon_{t-i,T_n-i}^2} \leq \frac{1}{(1-\beta_0)^2} + \frac{\alpha_0^s}{\beta_0 \omega_0^s} \sum_{i=1}^{T_n-2} i \left\{\beta_0^s\right\}^i \epsilon_{t-i,T_n-i}^{2s}.$$

Recall that the strictly stationary solution of a GARCH model satisfies $E|\epsilon_t|^{\nu_0} < \infty$ for some $\nu_0 > 0$. Therefore, for any $r \geq 1$, choosing $s > 0$ such that $E|\epsilon_t|^{2sr} < \infty$, the Hölder inequality shows that

$$\sup_n \left\|{}_\beta\sigma^2_{t,T_n}\right\|_r \leq \frac{1}{(1-\beta_0)^2} + \frac{\alpha_0^s}{\beta_0\omega_0^s}\left\|\epsilon_1^{2s}\right\|_r \sum_{i=1}^{\infty} i\left\{\beta_0^s\right\}^i < \infty, \quad \left\|{}_\beta\sigma_t^2\right\|_r < \infty.$$

Now note that

$$\begin{aligned}
{}_\beta\widetilde{\sigma}^2_{t,T_n} &= \frac{1}{\sigma^2_{t,T_n}}\frac{\partial\sigma_t^2(\boldsymbol{\theta}_0)}{\partial\beta} + \frac{\partial\sigma_t^2(\boldsymbol{\theta}_0)}{\partial\beta}\left(\frac{1}{\sigma_t^2(\boldsymbol{\theta}_0)} - \frac{1}{\sigma^2_{t,T_n}}\right) \\
&= \frac{1}{\sigma^2_{t,T_n}}\sum_{i=1}^{T_n-2} i\beta_0^{i-1}\left(\omega_0 + \alpha_0\epsilon^2_{t-i,T_n-i}\right) + \frac{1}{\sigma^2_{t,T_n}}\sum_{i=1}^{T_n-2} i\beta_0^{i-1}\alpha_0(\epsilon^2_{t-i} - \epsilon^2_{t-i,T_n-i}) \\
&\quad + \frac{1}{\sigma^2_{t,T_n}}\sum_{i=T_n-1}^{\infty} i\beta_0^{i-1}\left(\omega_0 + \alpha_0\epsilon^2_{t-i}\right) + {}_\beta\sigma_t^2\left(1 - \frac{\sigma_t^2(\boldsymbol{\theta}_0)}{\sigma^2_{t,T_n}}\right) \\
&= -\frac{\widetilde{\sigma}^2_{t,T_n}}{\sigma^2_{t,T_n}}{}_\beta\sigma_t^2 + \frac{\alpha_0\sum_{i=1}^{T_n-2} i\beta_0^{i-1}\widetilde{\epsilon}^2_{t-i,T_n-i}}{\sigma^2_{t,T_n}} + \frac{\sum_{i=T_n-1}^{\infty} i\beta_0^{i-1}\left(\omega_0 + \alpha_0\epsilon^2_{t-i}\right)}{\sigma^2_{t,T_n}}.
\end{aligned} \tag{A.2}$$

In view of (A.1), the first term of the right-hand side of the equality tends to zero in probability. Using Lemma 2.3 in Francq and Zakoïan (2010), the strict stationarity condition $E\log a(u_1) < 0$ entails the existence of $s \in (0,1)$ such that $\rho := Ea^s(u_1) < 1$. We then have $E\widetilde{\epsilon}^{2s}_{t,T_n} \leq K\rho^{T_n}$, which entails

$$E\left|\sum_{i=1}^{T_n-2} i\beta_0^{i-1}\widetilde{\epsilon}^2_{t-i,T_n-i}\right|^s \leq K\sum_{i=1}^{T_n-2} i^s\beta_0^{s(i-1)}\rho^{T_n-i} \to 0 \quad \text{as} \quad n \to \infty.$$

Noting that $E|X_n|^s \to 0$ for some $s > 0$ entails that $X_n \to 0$ in probability, we conclude that the second term of the right-hand side of the equality (A.2) tends to zero in probability. Let $s \in (0,1)$ such that $E|\epsilon_t|^{2s} < \infty$. We have

$$E\left|\sum_{i=T_n-1}^{\infty} i\beta_0^{i-1}\left(\omega_0 + \alpha_0\epsilon^2_{t-i}\right)\right|^s \leq \left(\omega_0^s + \alpha_0^s E|\epsilon_1|^{2s}\right)\sum_{i=T_n-1}^{\infty} i^s(\beta_0^s)^{i-1} \to 0$$

as $n \to \infty$. If follows that the third term of the right-hand side of the equality (A.2) tends to zero in probability. We thus have shown that ${}_\beta\sigma^2_{t,T_n}$ can be chosen as being the last component of D_{t,T_n}. As already argued, the two other components are handled more easily. □

A.3 Proof of Theorem 4.1

We start by showing the following lemma.

Lemma A.1. *Under* **A1** *and* **A3**, *we have*

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\begin{pmatrix}(u_t^2-1)D_t \\ \mathbf{1}_{\{u_t<\xi_\alpha\}} - \alpha\end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \boldsymbol{S}_\alpha), \quad \boldsymbol{S}_\alpha = \begin{pmatrix}\boldsymbol{S}^{11} & \boldsymbol{S}^{12}_\alpha \\ \boldsymbol{S}^{21}_\alpha & \boldsymbol{S}^{22}_\alpha\end{pmatrix}.$$

Proof. Let $c_0 \in \mathbb{R}$, $c_1 \in \mathbb{R}^m$, $c = (c_0, c_1')'$ and

$$x_t = (u_t^2 - 1)c_1' D_t + c_0(\mathbf{1}_{\{u_t < \xi_\alpha\}} - \alpha), \quad x_{t,n} = (u_t^2 - 1)c_1' D_{t,T_n} + c_0(\mathbf{1}_{\{u_t < \xi_\alpha\}} - \alpha).$$

We will apply the central limit theorem given in Francq and Zakoian (2005) (hereafter FZ) to the triangular array $(x_{t,n})$. Note that, by **A1** and **A3**, we have

$$\sup_n \|u_t^2 D_{t,T_n}\|_{2+\nu^*}^{2+\nu^*} \leq \|u_t^{4+2\nu^*}\|_p \left\| \sup_n \|D_{t,T_n}\|^{2+\nu^*} \right\|_q < \infty$$

if $0 < \nu^* < \nu/2$ and $p = q/(q-1) > 1$ is sufficiently small to satisfy $p(4+2\nu^*) < 4+\nu$. Therefore (1) in FZ is satisfied. Now, note that $\|\widetilde{D}_{t,T_n}\|^r \to 0$ in probability, and the sequence $\|\widetilde{D}_{t,T_n}\|^r$ is uniformly integrable because

$$\sup_n \|\widetilde{D}_{t,T_n}\|_r \leq \sup_n \|D_{t,T_n}\|_r + \|D_t\|_r < \infty.$$

From Theorem 3.5 in Billingsley (1999) it follows that $E\|\widetilde{D}_{t,T_n}\|^r \to 0$ as $n \to \infty$, for any $r \geq 1$. With $\nu^* = \nu/4$, we have

$$\begin{aligned} \mathrm{Var}\frac{1}{\sqrt{n}}\sum_{t=1}^n (x_t - x_{t,n}) &= E\left\{E(u_t^4 \mid \mathcal{F}_{t-1}) - 1\right\} c_1' \widetilde{D}_{t,T_n} \widetilde{D}_{t,T_n}' c_1 \\ &\leq \|E(u_t^4 \mid \mathcal{F}_{t-1}) - 1\|_{1+\nu^*} \|c_1' \widetilde{D}_{t,T_n}\|_{2(1+\nu^*)/\nu^*}^2 \to 0 \end{aligned} \tag{A.3}$$

as $n \to \infty$. Therefore

$$\lim_{n\to\infty} n^{-1}\mathrm{Var}\sum_{t=1}^n x_{t,n} = \lim_{n\to\infty} n^{-1}\mathrm{Var}\sum_{t=1}^n x_t \to \sigma^2 = c'\boldsymbol{S}_\alpha c.$$

Thus (2) in FZ is satisfied. Conditions (3) and (4) in FZ are satisfied if ν^* is chosen sufficiently close to $\nu/2$ and $\nu - \epsilon < 2\nu^* < \nu$. By Denoting by $\{\alpha_n(h), h \in \mathbb{N}\}$ the strong mixing coefficients of $\{x_{n,t})_t$, we have (4) of FZ. It follows that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^n x_{t,n} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2).$$

The conclusion then follows from (A.3) and the Cramér-Wold device. □

Now we turn to the proof of Theorem 4.1. Let $u_t(\boldsymbol{\theta}) = \epsilon_t/\sigma_t(\boldsymbol{\theta})$. Note that, by **A4** and **A5**, for n large enough

$$\left|\hat{u}_t - u_t(\hat{\boldsymbol{\theta}}_n)\right| = \left|\epsilon_t \frac{\sigma_t(\hat{\boldsymbol{\theta}}_n) - \widetilde{\sigma}_t(\hat{\boldsymbol{\theta}}_n)}{\widetilde{\sigma}_t(\hat{\boldsymbol{\theta}}_n)\sigma_t(\hat{\boldsymbol{\theta}}_n)}\right| \leq \frac{C}{\underline{\omega}}\rho^t u_t \sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)} \left|\frac{\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta})}\right|. \tag{A.4}$$

A Taylor expansion around $\boldsymbol{\theta}_0$ and **A4**, **A5** yield

$$\hat{u}_t = u_t - u_t D_t'(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) + r_{n,t}$$

with

$$r_{n,t} = \frac{1}{2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)' \frac{\partial^2 u_t(\boldsymbol{\theta}_{n,t})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) + \hat{u}_t - u_t(\hat{\boldsymbol{\theta}}_n),$$

where $D_t = D_t(\boldsymbol{\theta}_0)$ and $\boldsymbol{\theta}_{n,t}$ is between $\hat{\boldsymbol{\theta}}_n$ and $\boldsymbol{\theta}_0$. Following the approach of Knight (1998) and Koenker (2006) (see also Francq and Zakoian (2015)), we then obtain

$$\sqrt{n}(\xi_{n,\alpha} - \xi_\alpha) = \arg\min_{z \in \mathbb{R}} Q_n(z)$$

where

$$Q_n(z) \quad = \quad zX_n + I_n(z) + J_n(z) + K_n(z)$$

with

$$\begin{aligned}
X_n &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\mathbf{1}_{\{u_t < \xi_\alpha\}} - \alpha), \\
I_n(z) &= \sum_{t=1}^{n} \int_0^{z/\sqrt{n}} (\mathbf{1}_{\{u_t \leq \xi_\alpha + s\}} - \mathbf{1}_{\{u_t < \xi_\alpha\}}) ds, \\
J_n(z) &= \sum_{t=1}^{n} \int_0^{R_{t,n}/\sqrt{n}} \left(\mathbf{1}_{\{u_t - \xi_\alpha - z/\sqrt{n} \leq u\}} - \mathbf{1}_{\{u_t - \xi_\alpha - z/\sqrt{n} < 0\}} \right) du, \\
K_n(z) &= \sum_{t=1}^{n} \frac{R_{t,n}}{\sqrt{n}} \mathbf{1}^*_{\{u_t - \xi_\alpha \in (0, z/\sqrt{n})\}},
\end{aligned}$$

and $R_{t,n} = u_t D_t' \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) - \sqrt{n} r_{n,t}$. We will show that

$$Q_n(z) = \frac{z^2}{2} E f_0(\xi_\alpha) + z\{X_n + \xi_\alpha \Omega' \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)\} + O_P(1). \tag{A.5}$$

Noting that

$$\begin{aligned}
K_n(z) &= \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_t \mathbf{1}^*_{\{u_t - \xi_\alpha \in (0, z/\sqrt{n})\}} D_t' \right) \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \\
&\quad - \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)' \frac{1}{2n} \sum_{t=1}^{n} \frac{\partial^2 u_t(\boldsymbol{\theta}_{n,t})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \mathbf{1}^*_{\{u_t - \xi_\alpha \in (0, z/\sqrt{n})\}} \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \\
&\quad - \sum_{t=1}^{n} \left\{ \hat{u}_t - u_t(\hat{\boldsymbol{\theta}}_n) \right\} \mathbf{1}^*_{\{u_t - \xi_\alpha \in (0, z/\sqrt{n})\}} \\
:= &\; K_{n1}(z) + K_{n2}(z) + K_{n3}(z),
\end{aligned}$$

the proof of (A.5) will be divided in the following steps.

i) $K_{ni}(z) \to 0$ in probability as $n \to \infty$, for $i = 2, 3$.

ii) $K_{n1}(z) = z\xi_\alpha \Omega' \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) + o_P(1)$ in probability as $n \to \infty$.

iii) $J_n(z)$ does not depend on z asymptotically.

iv) $I_n(z) \to \frac{z^2}{2} E f_0(\xi_\alpha)$ in probability as $n \to \infty$.

To prove i) for $i = 2$, note that

$$\frac{\partial^2 u_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} = -u_t \frac{\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta})} \frac{1}{\sigma_t(\boldsymbol{\theta})} \frac{\partial^2 \sigma_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} + 2u_t \frac{\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta})} \frac{1}{\sigma_t^2(\boldsymbol{\theta})} \frac{\partial \sigma_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} \frac{\partial \sigma_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}.$$

In view of **A1** and the last part of **A7**, for $\boldsymbol{\theta} \in V(\boldsymbol{\theta}_0)$, $u_t \frac{\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta})}$ admits a moment larger than 2. The first part of **A7** and the Cauchy-Schwartz inequality then entail that

$$E \sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)} \left\| \frac{\partial^2 u_t(\boldsymbol{\theta}_{n,t})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \right\|^{1+\nu} < \infty. \tag{A.6}$$

for some $\nu > 0$. By the Hölder inequality, for $\boldsymbol{\theta}_{n,t} \in V(\boldsymbol{\theta}_0)$,

$$\begin{aligned} &\left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 u_t(\boldsymbol{\theta}_{n,t})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \mathbf{1}^*_{\{u_t - \xi_\alpha \in (0, z/\sqrt{n})\}} \right\| \\ \le\ & \left\{ \frac{1}{n} \sum_{t=1}^n \sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)} \left\| \frac{\partial^2 u_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \right\|^{1+\nu} \right\}^{1/(1+\nu)} \left\{ \frac{1}{n} \sum_{t=1}^n \mathbf{1}^*_{\{u_t - \xi_\alpha \in (0, z/\sqrt{n})\}} \right\}^{\nu/(1+\nu)} \end{aligned}$$

By (A.6) and the ergodic theorem, the limit of the first term of the latter product is almost surely finite. Letting $\nu_{t,n} = \mathbf{1}^*_{\{u_t - \xi_\alpha \in (0, z/\sqrt{n})\}}$ and $\overline{\nu}_{t,n} = \nu_{t,n} - E(\nu_{t,n} \mid \mathcal{F}_{t-1})$, we have

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{1}^*_{\{u_t - \xi_\alpha \in (0, z/\sqrt{n})\}} = \frac{1}{\sqrt{n}} \sum_{t=1}^n \overline{\nu}_{t,n} + \frac{1}{\sqrt{n}} \sum_{t=1}^n E(\nu_{t,n} \mid \mathcal{F}_{t-1}).$$

First note that

$$E(\nu_{t,n} \mid \mathcal{F}_{t-1}) = \int_{\xi_\alpha}^{\xi_\alpha + z/\sqrt{n}} f_{t-1}(x)dx = \frac{z}{\sqrt{n}} f_{t-1}(\xi_\alpha) + \frac{k_{t,n}}{\sqrt{n}}$$

where

$$|k_{t,n}| = \sqrt{n} \left| \int_{\xi_\alpha}^{\xi_\alpha + z/\sqrt{n}} \{f_{t-1}(x) - f_{t-1}(\xi_\alpha)\} dx \right| \le K_{t-1} \frac{z^2}{\sqrt{n}},$$

by **A1**. Now, note that we have $E \frac{1}{\sqrt{n}} \sum_{t=1}^n \overline{\nu}_{t,n} = 0$ and

$$\mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \overline{\nu}_{t,n} \right) = E\overline{\nu}_{1,n}^2 = \int_{\xi_\alpha}^{\xi_\alpha + z/\sqrt{n}} E\left\{ 1 - \frac{z}{\sqrt{n}} f_0(\xi_\alpha) - \frac{k_{t,n}}{\sqrt{n}} \right\}^2 f_0(x)dx \to 0,$$

using again **A1**. Moreover, almost surely

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n E(\nu_{t,n} \mid \mathcal{F}_{t-1}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \int_{\xi_\alpha}^{\xi_\alpha + z/\sqrt{n}} f_{t-1}(x)dx \to zEf_{t-1}(\xi_\alpha).$$

We thus have shown that $\sum_{t=1}^n \mathbf{1}^*_{\{u_t-\xi_\alpha\in(0,z/\sqrt{n})\}} = O_P(\sqrt{n})$. Thus, i) for $i=2$ is established. By the same arguments and (A.4), it can be shown that i) for $i=3$ holds.

Turning to ii), we have

$$\begin{aligned}
E\left(u_t\mathbf{1}^*_{\{u_t-\xi_\alpha\in(0,z/\sqrt{n})\}}D_t' \mid \mathcal{F}_{t-1}\right) &= \int_{\xi_\alpha}^{\xi_\alpha+z/\sqrt{n}} xf_{t-1}(x)dxD_t' \\
&= \int_{\xi_\alpha}^{\xi_\alpha+z/\sqrt{n}} xf_{t-1}(\xi_\alpha)dxD_t' + \int_{\xi_\alpha}^{\xi_\alpha+z/\sqrt{n}} x\{f_{t-1}(x)-f_{t-1}(\xi_\alpha)\}dxD_t' \\
&= \xi_\alpha f_{t-1}(\xi_\alpha)\frac{z}{\sqrt{n}}D_t' + \frac{k^*_{t,n}}{\sqrt{n}}D_t'
\end{aligned}$$

with

$$\begin{aligned}
\left|k^*_{t,n}\right| &= \sqrt{n}\left|f_{t-1}(\xi_\alpha)\frac{z^2}{2n} + \int_{\xi_\alpha}^{\xi_\alpha+z/\sqrt{n}} x\left\{f_{t-1}(x)-f_{t-1}(\xi_\alpha)\right\}dx\right| \\
&\le \frac{z^2}{\sqrt{n}}\left\{2K_{t-1}\xi_\alpha + \frac{f_{t-1}(\xi_\alpha)}{2} + \frac{|z|}{\sqrt{n}}K_{t-1}\right\}.
\end{aligned}$$

Denoting by d_t a generic element of D_t, we also have

$$\begin{aligned}
&EE\left[\left\{u_t\mathbf{1}^*_{\{u_t-\xi_\alpha\in(0,z/\sqrt{n})\}}d_t - E\left(u_t\mathbf{1}^*_{\{u_t-\xi_\alpha\in(0,z/\sqrt{n})\}}d_t \mid \mathcal{F}_{t-1}\right)\right\}^2 \mid \mathcal{F}_{t-1}\right] \\
&= \int_{\xi_\alpha}^{\xi_\alpha+z/\sqrt{n}} E\left(x-\xi_\alpha f_{t-1}(\xi_\alpha)\frac{z}{\sqrt{n}}+\frac{k^*_{t,n}}{\sqrt{n}}\right)^2 d_t^2 f_{t-1}(x)dx = o(1), \qquad \text{(A.7)}
\end{aligned}$$

as $n\to\infty$. To show that the expectation inside the latter integral is finite, we used in particular the fact that

$$E\sup_{x\in[\xi_\alpha,\xi_\alpha+z/\sqrt{n}]} d_t^2 f_{t-1}^2(\xi_\alpha)f_{t-1}(x) \le \sqrt{Ed_t^8 E\sup_{\xi\in V(\xi_\alpha)} f_{t-1}^4(\xi)} < \infty$$

for sufficiently large n under **A1** and **A7**. Hence, ii) is established.

To prove iii), write $J_n(z) = \sum_{t=1}^n J_{n,t}$. Write $r_{n,t} = r_{n,t}(\hat{\boldsymbol{\theta}}_n)$, $R_{n,t} = R_{n,t}(\hat{\boldsymbol{\theta}}_n)$, $J_{n,t} = J_{n,t}(\hat{\boldsymbol{\theta}}_n)$ and $J_n(z) = J_n(z,\hat{\boldsymbol{\theta}}_n)$. Let $(\boldsymbol{\theta}_n)$ be a deterministic sequence such that $\sqrt{n}(\boldsymbol{\theta}_n-\boldsymbol{\theta}_0) = O(1)$. By the change of variable $u = u_t v$, we have

$$\begin{aligned}
E\left(J_{n,t}(\boldsymbol{\theta}_n) \mid \mathcal{F}_{t-1}\right) &= \int_0^{D_t'(\boldsymbol{\theta}_n-\boldsymbol{\theta}_0)+o_P(n^{-1/2})} E\left(u_t\mathbf{1}^*_{\{u_t\in(\xi_\alpha+z/\sqrt{n},(\xi_\alpha+z/\sqrt{n})(1-v)^{-1})\}} \mid \mathcal{F}_{t-1}\right)dv \\
&= \frac{\xi_\alpha^2}{2}f_{t-1}(\xi_\alpha)(\boldsymbol{\theta}_n-\boldsymbol{\theta}_0)'D_tD_t'(\boldsymbol{\theta}_n-\boldsymbol{\theta}_0) + o_P(n^{-1}).
\end{aligned}$$

By the arguments used to show (A.7), we can show that

$$E\left\{J_{n,t}(\boldsymbol{\theta}_n) - (J_{n,t}(\boldsymbol{\theta}_n) \mid \mathcal{F}_{t-1})\right\}^2 = o(n^{-1}).$$

We thus have

$$J_n(z, \boldsymbol{\theta}_n) = \frac{\xi_\alpha^2}{2} \sqrt{n}(\boldsymbol{\theta}_n - \boldsymbol{\theta}_0)' E\{f_0(\xi_\alpha) D_1 D_1'\} \sqrt{n}(\boldsymbol{\theta}_n - \boldsymbol{\theta}_0) + o(1), \quad a.s.$$

It follows that $J_n(z, \boldsymbol{\theta}_n)$ does not depend of z asymptotically. Since this is true for any sequence such that $\sqrt{n}(\boldsymbol{\theta}_n - \boldsymbol{\theta}_0) = O(1)$, this also true almost surely for $J_n(z)$ and iii) is established.

By the previously used arguments, it can be shown that iv) holds which completes the proof of (A.5). By Lemma 2.2 in Davis et al. (1992) and convexity arguments, we can conclude that

$$\sqrt{n}(\xi_\alpha - \xi_{n,\alpha}) \quad = \quad \frac{\xi_\alpha}{E f_0(\xi_\alpha)} \Omega' \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) + \frac{1}{E f_0(\xi_\alpha)} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\mathbf{1}_{\{u_t < \xi_\alpha\}} - \alpha) + o_P(1).$$

We have the following Taylor expansion

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \quad = \quad \frac{-J^{-1}}{2\sqrt{n}} \sum_{t=1}^{n} (1 - u_t^2) D_t + o_P(1).$$

Hence

$$\mathrm{Cov}_{as} \left(\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0), \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\mathbf{1}_{\{u_t < \xi_\alpha\}} - \alpha) \right) \quad = \quad \frac{1}{2} p_\alpha J^{-1} \Omega$$

and thus

$$\mathrm{Var}_{as}\{\sqrt{n}(\xi_{n,\alpha} - \xi_\alpha)\} = \left\{ \xi_\alpha^2 \frac{\kappa_4 - 1}{4} + \frac{\xi_\alpha p_\alpha}{f(\xi_\alpha)} \right\} \Omega' J^{-1} \Omega + \frac{\alpha(1-\alpha)}{f^2(\xi_\alpha)},$$
$$\mathrm{Cov}_{as} \left(\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0), \sqrt{n}(\xi_\alpha - \xi_{n,\alpha}) \right) = \lambda_\alpha J^{-1} \Omega.$$

We have $\Omega' J^{-1} \Omega = 1$ and thus we obtain

$$\mathrm{Var}_{as}\{\sqrt{n}(\xi_\alpha - \xi_{n,\alpha})\} \quad = \quad \zeta_\alpha.$$

By the CLT for martingale differences, we get the announced result. □

References

Barone-Adesi, G., Giannopoulos, K., and L. Vosper (1999) VaR without correlations for nonlinear portfolios. *Journal of Futures Markets* 19, 583–602.

Bauwens, L., Hafner, C.M. and S. Laurent (2012) *Handbook of Volatility Models and Their Applications.* Wiley.

Bauwens, L., Laurent, S. and J.V.K. Rombouts (2006) Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21, 79–109.

Christoffersen, P.F. (2003) *Elements of financial risk management.* Academic Press, London.

Christoffersen, P.F. and S. Gonçalves (2005) Estimation Risk in Financial Risk Management. *Journal of Risk* 7, 1–28.

Diebold, F.X. and R.S. Mariano (1995) Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13, 253–263.

Escanciano, J.C., and J. Olmo (2010) Backtesting parametric VaR with estimation risk. *Journal of Business and Economic Statistics* 28, 36–51.

Escanciano, J.C. and J. Olmo (2011) Robust backtesting tests for value-at-risk models. *Journal of Financial Econometrics* 9, 132–161.

Farkas, W., Fringuellotti, F. and R. Tunaru (2016) Regulatory Capital Requirements: Saving Too Much for Rainy Days? Unpublished document.

Francq, C., and J.M. Zakoïan (2010) *GARCH models: structure, statistical inference and financial applications.* Chichester: John Wiley.

Francq, C. and J.M. Zakoïan (2015) Risk-parameter estimation in volatility models. *Journal of Econometrics* 184, 158–173.

Francq, C. and J.M. Zakoïan (2017): Estimation risk for the VaR of portfolios driven by semi-parametric multivariate models. Unpublished document. Available at http://ecares.ulb.ac.be/ecaresdocuments/seminars1617/francq.pdf

Gouriéroux, C. and J.M. Zakoïan (2013) Estimation adjusted VaR. *Econometric Theory* 29, 735–770.

Hurlin, C., Laurent, S., Quaedvlieg, R. and S. Smeekes (2017) Risk Measure Inference. *Journal of Business & Economic Statistics* 35, 499–512.

Mancini, L. and F. Trojani (2011) Robust Value-at-Risk prediction. *Journal of Financial Econometrics* 9, 281–313.

Nieto, M. R. and E. Ruiz (2016) Frontiers in VaR forecasting and backtesting. *International Journal of Forecasting* 32, 475–501.

Rombouts, J.V.K. and M. Verbeek (2009) Evaluating portfolio Value-at-Risk using semi-parametric GARCH models. *Quantitative Finance* 9, 737–745.

Santos, A.A.P., Nogales F.J. and E. Ruiz (2013) Comparing univariate and multivariate models to forecast portfolio Value-at-Risk. *Journal of Financial Econometrics* 11, 400–441.

Spierdijk, L. (2016) Confidence Intervals for ARMA-GARCH Value-at-Risk: The Case of Heavy Tails and Skewness. *Computational Statistics and Data Analysis* 100, 545–559.