



Munich Personal RePEc Archive

## **Variance targeting estimation of multivariate GARCH models**

Francq, Christian and Horvath, Lajos and Zakoian,  
Jean-Michel

CREST and University Lille 3, University of Utah, CREST and  
University Lille 3

6 August 2014

Online at <https://mpra.ub.uni-muenchen.de/57794/>  
MPRA Paper No. 57794, posted 06 Aug 2014 16:28 UTC

# Variance targeting estimation of multivariate GARCH models\*

CHRISTIAN FRANCO<sup>†</sup>, LAJOS HORVATH<sup>‡</sup> AND JEAN-MICHEL ZAKOÏAN<sup>§</sup>

## Abstract

We establish the strong consistency and the asymptotic normality of the variance-targeting estimator (VTE) of the parameters of the multivariate CCC-GARCH( $p, q$ ) processes. This method alleviates the numerical difficulties encountered in the maximization of the quasi likelihood by using an estimator of the unconditional variance. It is shown that the distribution of the VTE can be consistently estimated by a simple residual bootstrap technique. We also use the VTE for testing the model adequacy. A test statistic in the spirit of the score test is constructed, and its asymptotic properties are derived under the null assumption that the model is well specified. An extension of the VT method to asymmetric CCC-GARCH models incorporating leverage effects is studied. Numerical illustrations are provided and an empirical application based on daily exchange rates is proposed.

*JEL Classification.* C13 and C22

*Keywords.* Adequacy Test for CCC-GARCH models, Bootstrap, Leverage Effect, Quasi Maximum Likelihood Estimation, Variance Targeting Estimator.

---

\*C. Francq and J-M. Zakoïan gratefully acknowledge financial support of the Agence Nationale de la Recherche (ANR) through the ORA Program (Project PRAM ANR-10-ORAR-008-01) and of the Ecodec Labex. The research of Lajos Horváth is supported by NSF grant DMS 1305858.

<sup>†</sup>CREST and Université Lille 3, BP 60149, 59653 Villeneuve d'Ascq cedex, France. E-Mail: christian.francq@univ-lille3.fr

<sup>‡</sup>University of Utah, Department of Mathematics, 155 South 1400 East, Salt Lake City UT 84112-0090, USA. E-mail: horvath@math.utah.edu

<sup>§</sup>Corresponding author: Jean-Michel Zakoïan, CREST and Université Lille 3. Address: CREST, 15 boulevard Gabriel Péri, 92245 Malakoff Cedex, France. E-mail: zakoian@ensae.fr, Phone number: 33 141177725.

# 1 Introduction

Despite their formidable success in the empirical finance literature, multivariate conditionally heteroskedastic (GARCH) models are notoriously difficult to estimate. Even for a small number of series, commonly used specifications of the conditional covariance matrix contain a large number of parameters, except if some very restrictive and generally arbitrary conditions are set on the parameters. Moreover, the positive definiteness of the conditional covariance matrix may entail quite complicated constraints on such parameters. As a consequence, the Quasi-Maximum Likelihood (QML) method - arguably the most popular estimation method in the univariate GARCH setting - can be difficult to apply. For this reason, the variance targeting (VT) estimation method proposed by Engle and Mezrich (1996) has gained popularity in the recent financial econometrics literature.

VT is based on a reparameterization of the conditional variance matrix, in which the matrix of intercepts is replaced by the unconditional covariance matrix. A simple moment estimator of this matrix is used in the first-step while, conditioning on this estimate, the remaining parameters are estimated by QML in the second step. This method has been recommended in the textbooks of Hull (2003) and Christoffersen (2003). It has been widely discussed in the recent academic literature, for example by Boudt, Daníelsson and Laurent (2012), Laurent, Rombouts and Violante (2012). The asymptotic properties of the VT estimator (VTE) have been established for univariate GARCH( $p, q$ ) by Kristensen and Linton (2004) and by Francq, Horváth and Zakoian (2011). Hill and Renault (2012), and Vaynman and Beare (2013) studied the asymptotic behavior of the VTE in the presence of heavy tails. For the first time in a multivariate setting, the asymptotic properties of the VTE were recently established in the case of the BEKK-GARCH(1,1)<sup>1</sup> model by Pedersen and Rahbek (2013).

In this article, we study VT estimation of the CCC-GARCH( $p, q$ ) models<sup>2</sup>. This class is attractive for its tractability and ease of interpretation. In particular, conditions for the positive definiteness of the conditional variance and for the existence of strictly stationary solutions are simple and explicit, in contrast to most of its competitors<sup>3</sup>. Our first aim is to establish the consistency and asymptotic normality (CAN) of the VTE under similar assumptions as those used by Francq,

<sup>1</sup>The BEKK model was studied by Baba, Engle, Kraft and Kroner, in a preliminary version of Engle and Kroner (1995).

<sup>2</sup>Constant Conditional Correlations (CCC) models were introduced by Bollerslev (1990) and extended by Jeantheau (1998).

<sup>3</sup>For an overview on multivariate GARCH models, see for example the surveys by Bauwens, Laurent and Rombouts (2006), Silvennoinen and Teräsvirta (2009), and Chapter 11 in Francq and Zakoian (2010).

Horváth and Zakoian (2011) in the univariate case. We also propose a consistent residual bootstrap procedure for approximating the asymptotic distribution of the VTE. Our second aim is to use the VTE for testing the model adequacy. Even if the QML method for the whole set of parameters is not used, the derivatives of the quasi-likelihood with respect to the first components of the parameter (those which are estimated in the first step of the VT method) can be used to derive a test in the spirit of the score test. The VTE can indeed be viewed as a "constrained estimator", the estimation of the theoretical variance being forced to coincide with the empirical variance. An important difference with usual score tests is that our "constraint" is random, as it depends on the observations.

The paper is organized as follows. Section 2 introduces the model and some assumptions. Section 3 establishes the consistency and asymptotic normality of the VTE, as well as the validity of the residual bootstrap procedure. Section 4 considers an extension of the VT method to asymmetric CCC-GARCH models. Section 5 develops a bootstrap approximation for the distribution of the VTE. Section 6 defines an adequacy test statistic based on the VTE and derives its asymptotic distribution under the null assumption that the model is well specified. Numerical illustrations are given in Section 7. In particular, we propose an empirical study based on daily exchange rates. Proofs are displayed in Section 8. Section 9 concludes. Additional technicalities are provided in an appendix.

## 2 Model and notations

Let  $(\boldsymbol{\epsilon}_t)$ , with  $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \dots, \epsilon_{mt})'$ , be a vector process of dimension  $m$ . The process  $(\boldsymbol{\epsilon}_t)$  is called a CCC-GARCH( $p, q$ ) if it satisfies

$$\left\{ \begin{array}{l} \boldsymbol{\epsilon}_t = \mathbf{H}_t^{1/2} \boldsymbol{\eta}_t, \\ \mathbf{H}_t = \mathbf{D}_t \mathbf{R}_0 \mathbf{D}_t, \quad \mathbf{D}_t^2 = \text{diag}(\underline{\mathbf{h}}_t), \\ \underline{\mathbf{h}}_t - \mathbf{h}_0 = \sum_{i=1}^q \mathbf{A}_{0i} (\underline{\boldsymbol{\epsilon}}_{t-i} - \mathbf{h}_0) + \sum_{j=1}^p \mathbf{B}_{0j} (\underline{\mathbf{h}}_{t-j} - \mathbf{h}_0), \end{array} \right. \quad (2.1)$$

where  $\underline{\boldsymbol{\epsilon}}_t = (\epsilon_{1t}^2, \dots, \epsilon_{mt}^2)'$  and  $\mathbf{R}_0$  is a correlation matrix. The matrices  $\mathbf{A}_{0i}$  and  $\mathbf{B}_{0j}$  are matrices of size  $m \times m$  with positive coefficients and  $\mathbf{h}_0$  is a vector of dimension  $m$  such that  $\left\{ \mathbf{I}_m - \sum_{i=1}^r (\mathbf{A}_{0i} + \mathbf{B}_{0i}) \right\} \mathbf{h}_0$  has strictly positive coefficients (with  $r = \max\{p, q\}$  and usual conventions). The innovations  $(\boldsymbol{\eta}_t)$  are iid centered variables on  $\mathbb{R}^m$  with identity covariance matrix.

The CCC model was introduced by Bollerslev (1990) when the matrices  $\mathbf{A}_{0i}$  and  $\mathbf{B}_{0j}$  are diag-

onal. By contrast, in (2.1) the conditional variance  $h_{kk,t}$  of the  $k$ -th component of  $\epsilon_t$  depends not only on its past values but also on the past values of the other components. For this reason, model (2.1) is referred to as the *Extended CCC* model by He and Teräsvirta (2004).

In the latter reference, a sufficient condition for second-order and strict stationarity of a CCC-GARCH(1,1) is given. A sufficient condition for strict stationarity and the existence of fourth-order moments of the CCC-GARCH( $p, q$ ) is established in Aue et al. (2009). In particular, it is known that (2.1) admits a strict and second-order non anticipative stationary solution  $(\epsilon_t)$  when

$$\mathbf{A}: \text{the spectral radius of } \sum_{i=1}^q \mathbf{A}_{0i} + \sum_{j=1}^p \mathbf{B}_{0j} \text{ is strictly less than 1.}$$

Moreover, under this assumption, we have that  $E\mathbf{h}_t = \mathbf{h}_0$ .

Turning to estimation, the generic parameter value consists of the coefficients of the vector  $\mathbf{h}$  and the matrices  $\mathbf{A}_i$  and  $\mathbf{B}_j$  (corresponding to the true values  $\mathbf{h}_0, \mathbf{A}_{0i}$  and  $\mathbf{B}_{0j}$ , respectively), and the coefficients of the lower triangular part (excluding the diagonal) of the correlation matrix  $\mathbf{R} = (\rho_{ij})$ . We will distinguish the first  $m$  parameters, that is the components of  $\mathbf{h}$ , which will be estimated empirically, and the other parameters, which will be estimated via a QML optimization. Thus the total number of unknown parameters is

$$s_1 = s_0 + m, \quad s_0 = m^2(p + q) + \frac{m(m-1)}{2}.$$

The parameter vector is denoted  $\boldsymbol{\vartheta} = (\mathbf{h}', \boldsymbol{\theta}')$ , with

$$\boldsymbol{\theta} = (\theta_1, \dots, \theta_{s_0})' = (\boldsymbol{\alpha}'_1, \dots, \boldsymbol{\alpha}'_q, \boldsymbol{\beta}'_1, \dots, \boldsymbol{\beta}'_p, \boldsymbol{\rho}'_1)' := (\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\rho}')',$$

where  $\boldsymbol{\rho}' = (\rho_{21}, \dots, \rho_{m1}, \rho_{32}, \dots, \rho_{m2}, \dots, \rho_{m,m-1}) \in \mathbb{R}^{m(m-1)/2}$ ,  $\boldsymbol{\alpha}_i = \text{vec } \mathbf{A}_i \in \mathbb{R}^{m^2}$ ,  $i = 1, \dots, q$ , and  $\boldsymbol{\beta}_j = \text{vec } \mathbf{B}_j \in \mathbb{R}^{m^2}$ ,  $j = 1, \dots, p$ . The parameter space for  $\boldsymbol{\theta}$  is

$$\Theta \subset [0, \infty)^{m^2(p+q)} \times (-1, 1)^{m(m-1)/2}.$$

The true parameter value is denoted by

$$\boldsymbol{\vartheta}_0 = (\mathbf{h}'_0, \boldsymbol{\theta}'_0)', \quad \boldsymbol{\theta}_0 = (\boldsymbol{\alpha}'_{01}, \dots, \boldsymbol{\alpha}'_{0q}, \boldsymbol{\beta}'_{01}, \dots, \boldsymbol{\beta}'_{0p}, \boldsymbol{\rho}'_0)' = (\boldsymbol{\alpha}'_0, \boldsymbol{\beta}'_0, \boldsymbol{\rho}'_0)'.$$

### 3 Asymptotic Properties of the VTE of the CCC-GARCH

Let  $(\epsilon_1, \dots, \epsilon_n)$  be an observation of length  $n$  of the unique non anticipative and strictly stationary solution  $(\epsilon_t)$  of model (2.1). Conditionally on nonnegative initial values  $\epsilon_0, \dots, \epsilon_{1-q}, \tilde{\mathbf{h}}_0, \dots, \tilde{\mathbf{h}}_{1-p}$ ,

the Gaussian quasi-likelihood can be written as

$$L_n(\boldsymbol{\vartheta}) = L_n(\boldsymbol{\vartheta}; \boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_n) = \prod_{t=1}^n \frac{1}{(2\pi)^{m/2} |\widetilde{\mathbf{H}}_t|^{1/2}} \exp\left(-\frac{1}{2} \boldsymbol{\epsilon}_t' \widetilde{\mathbf{H}}_t^{-1} \boldsymbol{\epsilon}_t\right),$$

where the  $\widetilde{\mathbf{H}}_t$ 's are recursively defined, for  $t \geq 1$ , by

$$\begin{cases} \widetilde{\mathbf{H}}_t &= \widetilde{\mathbf{D}}_t \mathbf{R} \widetilde{\mathbf{D}}_t, \quad \widetilde{\mathbf{D}}_t = \{\text{diag}(\widetilde{\mathbf{h}}_t)\}^{1/2}, \\ \widetilde{\mathbf{h}}_t &= \widetilde{\mathbf{h}}_t(\boldsymbol{\vartheta}) = \mathbf{h} + \sum_{i=1}^q \mathbf{A}_i (\boldsymbol{\epsilon}_{t-i} - \mathbf{h}) + \sum_{j=1}^p \mathbf{B}_j (\widetilde{\mathbf{h}}_{t-j} - \mathbf{h}). \end{cases}$$

Note that, up to an unimportant additive constant,

$$-2 \log L_n(\boldsymbol{\vartheta}) = \sum_{t=1}^n \widetilde{\ell}_t, \quad \text{where } \widetilde{\ell}_t = \widetilde{\ell}_t(\mathbf{h}, \boldsymbol{\theta}) = \boldsymbol{\epsilon}_t' \widetilde{\mathbf{H}}_t^{-1} \boldsymbol{\epsilon}_t + \log |\widetilde{\mathbf{H}}_t|,$$

and  $|\widetilde{\mathbf{H}}_t|$  denotes the determinant of  $\widetilde{\mathbf{H}}_t$ . The VTE of the parameter  $\mathbf{h}_0$  is first defined by the empirical mean

$$\widehat{\mathbf{h}}_n = \frac{1}{n} \sum_{t=1}^n \boldsymbol{\epsilon}_t.$$

The VTE of the parameter  $\boldsymbol{\theta}_0$  is then defined by

$$\widehat{\boldsymbol{\theta}}_n = \arg \min_{\boldsymbol{\theta} \in \Theta} \widetilde{\mathcal{L}}_n(\boldsymbol{\theta}), \tag{3.1}$$

where

$$\widetilde{\mathcal{L}}_n(\boldsymbol{\theta}) = n^{-1} \sum_{t=1}^n \widetilde{\ell}_{t,n}, \quad \text{and } \widetilde{\ell}_{t,n} = \widetilde{\ell}_t(\widehat{\mathbf{h}}_n, \boldsymbol{\theta})$$

Let  $\widehat{\boldsymbol{\vartheta}}_n = (\widehat{\mathbf{h}}_n', \widehat{\boldsymbol{\theta}}_n')'$  be the VTE of  $\boldsymbol{\vartheta}_0$ .

If the parameter  $\mathbf{h}_0$  were known, the QMLE of the parameter  $\boldsymbol{\theta}_0$  would be defined by

$$\widehat{\boldsymbol{\theta}}_n^Q = \arg \min_{\boldsymbol{\theta} \in \Theta} \sum_{t=1}^n \widetilde{\ell}_t(\mathbf{h}_0, \boldsymbol{\theta}).$$

### 3.1 Consistency and asymptotic normality

Francq and Zakoïan (2012) provide conditions for the CAN of the QMLE of the whole parameter  $\boldsymbol{\vartheta}_0$ .

Their conditions can be adapted to our framework as follows. Let  $\mathcal{A}_\theta(z) = \sum_{i=1}^q \mathbf{A}_i z^i$  and  $\mathcal{B}_\theta(z) = \mathbf{I}_m - \sum_{j=1}^p \mathbf{B}_j z^j$ . By convention,  $\mathcal{A}_\theta(z) = \mathbf{0}$  if  $q = 0$  and  $\mathcal{B}_\theta(z) = \mathbf{I}_m$ , if  $p = 0$ .

**A1:**  $\boldsymbol{\theta}_0 \in \Theta$  and  $\Theta$  is compact.

**A2:**  $\forall \boldsymbol{\theta} \in \Theta, \quad |\mathcal{B}_\theta(z)| = 0 \Rightarrow |z| > 1$ .

**A3:** For  $i = 1, \dots, m$  the distribution of  $\tilde{\eta}_{it}$  is not concentrated on 2 points and  $P(\tilde{\eta}_{it} > 0) \in (0, 1)$ .

**A4:** If  $p > 0$ ,  $\mathcal{A}_{\theta_0}(z)$  and  $\mathcal{B}_{\theta_0}(z)$  are left coprime and  $[\mathbf{A}_{0q} \ \mathbf{B}_{0p}]$  has full rank  $m$ .

**A5:**  $\mathbf{R}$  is a positive definite correlation matrix for all  $\boldsymbol{\theta} \in \Theta$ .

**A6:**  $\boldsymbol{\theta}_0 \in \overset{\circ}{\Theta}$ , where  $\overset{\circ}{\Theta}$  is the interior of  $\Theta$ .

**A7:**  $E\|\boldsymbol{\eta}_t \boldsymbol{\eta}_t'\|^2 < \infty$ .

Under **A2** we can define  $\underline{\mathbf{h}}_t(\boldsymbol{\vartheta}) = \mathbf{h} + \mathcal{B}_{\boldsymbol{\theta}}^{-1}(L)\mathcal{A}_{\boldsymbol{\theta}}(L)(\underline{\boldsymbol{\epsilon}}_t - \mathbf{h})$ , where  $L$  is the usual backshift operator. Similarly,  $\ell_t$  is obtained by replacing  $\tilde{\mathbf{h}}_t(\boldsymbol{\vartheta})$  with  $\underline{\mathbf{h}}_t(\boldsymbol{\vartheta})$  in  $\tilde{\ell}_t$ .

The asymptotic behavior of  $\hat{\boldsymbol{\theta}}_n^Q$  is obviously similar to that of the full QMLE  $\hat{\boldsymbol{\vartheta}}_n^Q$ , as stated in the following theorem. The proof is omitted.

**Theorem 3.1** (CAN of the unfeasible QMLE of  $\boldsymbol{\theta}_0$ ). *Under Assumptions **A** and **A1-A5**, we have  $\hat{\boldsymbol{\theta}}_n^Q \rightarrow \boldsymbol{\theta}_0$  a.s. as  $n \rightarrow \infty$ . Under the additional assumptions **A6-A7**,  $\sqrt{n}(\hat{\boldsymbol{\theta}}_n^Q - \boldsymbol{\theta}_0)$  converges in distribution to  $\mathcal{N}(0, \mathbf{J}^{-1}\mathbf{I}\mathbf{J}^{-1})$ , where  $\mathbf{J}$  is a positive-definite matrix and  $\mathbf{I}$  is a semi positive-definite matrix, defined by*

$$\mathbf{I} = E\left(\frac{\partial \ell_t(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\theta}} \frac{\partial \ell_t(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\theta}'}\right), \quad \mathbf{J} = E\left(\frac{\partial^2 \ell_t(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right).$$

The spectral norm of a matrix  $\mathbf{A}$  is denoted by  $\|\mathbf{A}\|$ , and its  $L^p$ -norm is defined by  $\|\mathbf{A}\|_p = (E\|\mathbf{A}\|^p)^{1/p}$  for  $p \geq 1$ . We also need to introduce the notation  $\underline{\boldsymbol{\eta}}_t = (\eta_{1t}^2, \dots, \eta_{mt}^2)'$ . Aue et al. (2009) showed that there exists a stationary solution of (2.1) satisfying  $\|\boldsymbol{\epsilon}_t\|_4 < \infty$ , if

$$\|\boldsymbol{\eta}_1\|_4 < \infty \quad \text{and} \quad \sum_{i=1}^r \left\| \mathbf{A}_{0i} \text{diag}(\underline{\boldsymbol{\eta}}_1) + \mathbf{B}_{0i} \right\|_2 < 1. \quad (3.2)$$

Let

$$\mathbf{K} = E\left(\frac{\partial^2 \ell_t(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\theta} \partial \boldsymbol{h}'}\right), \quad \mathbf{G} = \begin{pmatrix} \mathbf{I}_m & 0 \\ -\mathbf{J}^{-1}\mathbf{K} & -\mathbf{J}^{-1} \end{pmatrix}$$

and

$$\boldsymbol{\Sigma}_X = \text{Var}(\mathbf{x}_t), \quad \mathbf{x}_t = \begin{pmatrix} \mathbf{C}(\boldsymbol{\theta}_0) \{\underline{\boldsymbol{\epsilon}}_t - \underline{\mathbf{h}}_t(\boldsymbol{\vartheta}_0)\} \\ \frac{\partial}{\partial \boldsymbol{\theta}} \ell_t(\boldsymbol{\vartheta}_0) \end{pmatrix},$$

where

$$\mathbf{C}(\boldsymbol{\theta}) = \left\{ \mathbf{I}_m - \sum_{i=1}^r (\mathbf{A}_i + \mathbf{B}_i) \right\}^{-1} \left( \mathbf{I}_m - \sum_{i=1}^p \mathbf{B}_i \right),$$

provided that  $\mathbf{I}_m - \sum_{i=1}^r (\mathbf{A}_i + \mathbf{B}_i)$  is non singular, which is the case at  $\boldsymbol{\theta}_0$  under Assumption **A**.

**Theorem 3.2** (CAN of the VTE). *Under the assumptions of Theorem 3.1,*

$$\widehat{\boldsymbol{\vartheta}}_n \rightarrow \boldsymbol{\vartheta}_0 \quad a.s.$$

as  $n \rightarrow \infty$ . *Under the additional assumptions in (3.2), we have*

$$\sqrt{n} \left( \widehat{\boldsymbol{\vartheta}}_n - \boldsymbol{\vartheta}_0 \right) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma} := \mathbf{G} \boldsymbol{\Sigma}_X \mathbf{G}'). \quad (3.3)$$

It is worth noting that in the univariate case ( $m = 1$ ), our conditions for CAN reduce to those used by Francq, Horváth and Zakoïan (2011) to establish the asymptotic properties of the VTE of the GARCH( $p, q$ ) coefficients. In particular, the observed process is (only) required to have finite fourth-order moments.

### 3.2 Estimating the asymptotic covariance matrix

We now discuss the estimation of the asymptotic covariance matrix  $\boldsymbol{\Sigma}$ . Let

$$\widehat{\boldsymbol{\Sigma}}_X = \frac{1}{n} \sum_{t=1}^n \widehat{\mathbf{x}}_t \widehat{\mathbf{x}}_t', \quad \widehat{\mathbf{x}}_t = \begin{pmatrix} \mathbf{C}(\widehat{\boldsymbol{\theta}}_n) \left\{ \boldsymbol{\varepsilon}_t - \widetilde{\mathbf{h}}_t(\widehat{\boldsymbol{\vartheta}}_n) \right\} \\ \frac{\partial}{\partial \boldsymbol{\theta}} \widetilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n) \end{pmatrix}.$$

The appendix describes a recursive algorithm to compute  $\frac{\partial}{\partial \boldsymbol{\theta}} \widetilde{\ell}_t(\boldsymbol{\vartheta})$  in the CCC-GARCH(1,1) case.

Let also

$$\widehat{\mathbf{J}} = \frac{1}{n} \sum_{t=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \widetilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n), \quad \widehat{\mathbf{K}} = \frac{1}{n} \sum_{t=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \mathbf{h}'} \widetilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n). \quad (3.4)$$

A consistent estimator for  $\boldsymbol{\Sigma}$  is then given by

$$\widehat{\boldsymbol{\Sigma}} = \widehat{\mathbf{G}} \widehat{\boldsymbol{\Sigma}}_X \widehat{\mathbf{G}}', \quad \widehat{\mathbf{G}} = \begin{pmatrix} \mathbf{I}_m & 0 \\ -\widehat{\mathbf{J}}^{-1} \widehat{\mathbf{K}} & -\widehat{\mathbf{J}}^{-1} \end{pmatrix}. \quad (3.5)$$

The computation of  $\widehat{\boldsymbol{\Sigma}}$  requires the evaluation of complicated first and second-order derivatives. More precisely, for  $\widehat{\boldsymbol{\Sigma}}_X$  one needs to compute  $\partial \widetilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n) / \partial \boldsymbol{\theta}$  for  $t = 1, \dots, n$ . These  $n$  vectors of derivatives cannot be evaluated numerically within a reasonable amount of time. In the Appendix we thus provide recursive formulas for a rapid computation of the first-order derivatives. Second order derivatives are also required for the computation of  $\widehat{\boldsymbol{\Sigma}}$ . Since each second-order derivative has to be evaluated only once, in view of

$$\widehat{\mathbf{J}} = \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \left\{ \frac{1}{n} \sum_{t=1}^n \widetilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n) \right\}, \quad \widehat{\mathbf{K}} = \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \mathbf{h}'} \left\{ \frac{1}{n} \sum_{t=1}^n \widetilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n) \right\},$$

these derivatives can be obtained numerically.



### 3.3 The diagonal case

When the matrices  $\mathbf{A}_{0i}$  and  $\mathbf{B}_{0j}$  are diagonal, as in the CCC model of Bollerslev (1990), the asymptotic normality in (3.3) does not hold because the true parameter value has components equal to zero, and thus **A6** is not satisfied. It is known that the asymptotic distribution of the QMLE for univariate GARCH models is different when the true parameter value is at the frontier of the parameter space (see Francq and Zakoian (2007)). Although similar results have not yet been established for multivariate GARCH, it is clear that the asymptotic distribution of  $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$  cannot be Gaussian when some components of  $\boldsymbol{\theta}_0$  are zero.

However, if these constraints ( $\mathbf{A}_{0i}$  and  $\mathbf{B}_{0j}$  diagonal) are known from the econometrician, the approach developed for the unconstrained model can be straightforwardly adapted. The VTE is defined as in (3.1) but with  $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\rho}')$  where the components of  $\boldsymbol{\alpha}$  and  $\boldsymbol{\beta}$  are now  $\alpha_i = \text{diag}(\mathbf{A}_i) \in \mathbb{R}^m$ ,  $i = 1, \dots, q$ ,  $\beta_j = \text{diag}(\mathbf{B}_j) \in \mathbb{R}^m$ ,  $j = 1, \dots, p$ . The parameter space  $\Theta$  is now a subset of  $[0, \infty)^{m(p+q)} \times (-1, 1)^{m(m-1)/2}$ . For  $\ell = 1, \dots, m$ , let  $\mathcal{A}_{\ell\boldsymbol{\theta}}(z) = \sum_{i=1}^q \alpha_{i\ell} z^i$  and  $\mathcal{B}_{\ell\boldsymbol{\theta}}(z) = \sum_{j=1}^p \beta_{j\ell} z^j$ . Assumptions **A2** and **A4** can be replaced by the simpler conditions

$$\mathbf{A2}': \quad \max_{1 \leq \ell \leq m} \sum_{j=1}^p \beta_{j\ell} < 1.$$

$$\mathbf{A4}': \quad \text{If } p > 0, \text{ for } \ell = 1, \dots, m, \mathcal{A}_{\ell\boldsymbol{\theta}_0}(z) \text{ and } \mathcal{B}_{\ell\boldsymbol{\theta}_0}(z) \text{ do not have common roots, and } \alpha_{0q\ell} + \beta_{0p\ell} \neq 0.$$

The individual components of  $\boldsymbol{\epsilon}_t$  are univariate GARCH  $(p, q)$  processes and the fourth-moment condition  $\|\boldsymbol{\epsilon}_t\|_4 < \infty$  reduces to

$$\text{For } \ell = 1, \dots, m, \text{ the spectral radius of } E(A_{\ell t} \otimes A_{\ell t}) \text{ is strictly less than 1,} \quad (3.6)$$

where, omitting the indices 0 for the true parameter values,

$$A_{\ell t} = \begin{pmatrix} \alpha_{1\ell} \eta_{\ell t}^2 + \beta_{1\ell} & \beta_{2\ell} & \cdots & \beta_{p-1,\ell} & \beta_{p\ell} & \alpha_{2\ell} & \cdots & \alpha_{q-1,\ell} & \alpha_{q\ell} \\ 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & \cdots & 0 \\ \eta_{\ell t}^2 & 0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

is a  $(p + q - 1) \times (p + q - 1)$  matrix (see Ling and McAleer, 2002)<sup>4</sup>. Although the dimensions of  $\theta$  and  $\vartheta$  are much smaller than in the general case, we keep the same notation for ease of exposition. We also use the same notations for the matrices involved in the asymptotic distribution of the VTE.

**Theorem 3.3** (CAN of the VTE in the diagonal case). *Under Assumptions **A1**, **A2'**, **A3**, **A4'**, **A5**, **A6** and (3.6), we have*

$$\sqrt{n} \left( \widehat{\vartheta}_n - \vartheta_0 \right) \xrightarrow{d} \mathcal{N}(0, \mathbf{G} \boldsymbol{\Sigma}_X \mathbf{G}').$$

## 4 Moments Targeting for the Asymmetric CCC-GARCH

Extension of the VT method to GARCH models allowing the leverage effect is not straightforward. Indeed, a key requirement of this method is the possibility to express the unconditional variance of the returns in terms of the volatility parameters, and to reparametrize the model with the unconditional variance as a new parameter. Without additional assumptions (for instance symmetry) on the errors distribution, this is generally not possible. Consider for example the univariate GJR GARCH(1,1) model

$$\epsilon_t = \sigma_t \eta_t, \quad \sigma_t^2 = \omega + \alpha_+ (\epsilon_{t-1}^+)^2 + \alpha_- (\epsilon_{t-1}^-)^2 + \beta \sigma_{t-1}^2,$$

where  $x^+ = \max(x, 0) = (-x)^-$  for any real number  $x$ , with  $\omega > 0, \alpha_+, \alpha_-, \beta \geq 0$ , with  $(\eta_t) \stackrel{iid}{\sim} (0, 1)$  and  $\eta_t$  independent of  $\{\epsilon_u, u < t\}$ . Then, straightforward calculation shows that if the variance of  $\epsilon_t$  exists, it is given by  $E\epsilon_t^2 = E\sigma_t^2 = \{1 - \alpha_+ E(\eta_t^+)^2 - \alpha_- E(\eta_t^-)^2 - \beta\}^{-1} \omega$ . If  $\eta_t$  is symmetrically distributed, then  $E\epsilon_t^2 = \{1 - \frac{1}{2}(\alpha_+ + \alpha_-) - \beta\}^{-1} \omega$  and we have the reparametrization

$$\sigma_t^2 - h_0 = \alpha_+ \left\{ (\epsilon_{t-1}^+)^2 - \frac{1}{2} h_0 \right\} + \alpha_- \left\{ (\epsilon_{t-1}^-)^2 - \frac{1}{2} h_0 \right\} + \beta (\sigma_{t-1}^2 - h_0).$$

However, the symmetry assumption is very strong. Another parametrization is

$$\sigma_t^2 = h_0 + \alpha_+ \{ (\epsilon_{t-1}^+)^2 - \underline{\epsilon}_0^+ \} + \alpha_- \{ (\epsilon_{t-1}^-)^2 - \underline{\epsilon}_0^- \} + \beta (\sigma_{t-1}^2 - h_0),$$

where, still assuming that the variance of  $\epsilon_t$  exists,  $\underline{\epsilon}_0^+ = E(\epsilon_t^+)^2$ ,  $\underline{\epsilon}_0^- = E(\epsilon_t^-)^2$ , and  $h_0 = \underline{\epsilon}_0^+ + \underline{\epsilon}_0^-$ . In this parametrization, the moments  $\underline{\epsilon}_0^+$  and  $\underline{\epsilon}_0^-$  can be "targeted" in a first step by the empirical means of  $(\epsilon_t^+)^2$  and  $(\epsilon_t^-)^2$ .

<sup>4</sup>A simple sufficient condition for (3.6) to hold is:  $\max_{1 \leq i \leq m} E(\eta_{it}^4) \left( \frac{\sum_{j=1}^q \alpha_{ij}}{1 - \sum_{j=1}^p \beta_{ij}} \right)^2 < 1$  (see Giraitis et al. (2006)).

We apply this approach for estimating an Asymmetric extension of the CCC-GARCH( $p, q$ ), the CCC-AGARCH( $p, q$ ) model defined by

$$\left\{ \begin{array}{l} \epsilon_t = \mathbf{H}_t^{1/2} \boldsymbol{\eta}_t, \\ \mathbf{H}_t = \mathbf{D}_t \mathbf{R}_0 \mathbf{D}_t, \quad \mathbf{D}_t^2 = \text{diag}(\mathbf{h}_t), \\ \mathbf{h}_t - \mathbf{h}_0 = \sum_{i=1}^q \mathbf{A}_{0i,+} (\boldsymbol{\epsilon}_{t-i}^+ - \boldsymbol{\epsilon}_{0+}) + \mathbf{A}_{0i,-} (\boldsymbol{\epsilon}_{t-i}^- - \boldsymbol{\epsilon}_{0-}) + \sum_{j=1}^p \mathbf{B}_{0j} (\mathbf{h}_{t-j} - \mathbf{h}_0), \end{array} \right.$$

where, using the notation  $x^+ = \max(x, 0) = (-x)^-$  for any real number  $x$ ,

$$\boldsymbol{\epsilon}_t^+ = \left( \{\epsilon_{1t}^+\}^2, \dots, \{\epsilon_{mt}^+\}^2 \right)', \quad \boldsymbol{\epsilon}_t^- = \left( \{\epsilon_{1t}^-\}^2, \dots, \{\epsilon_{mt}^-\}^2 \right)', \quad \boldsymbol{\epsilon}_{0+} = E\boldsymbol{\epsilon}_t^+, \quad \boldsymbol{\epsilon}_{0-} = E\boldsymbol{\epsilon}_t^-,$$

$\mathbf{h}_0 = \boldsymbol{\epsilon}_{0+} + \boldsymbol{\epsilon}_{0-}$  such that  $\mathbf{h}_0 - \sum_{i=1}^q (\mathbf{A}_{0i,+} \boldsymbol{\epsilon}_{0+} + \mathbf{A}_{0i,-} \boldsymbol{\epsilon}_{0-}) - \sum_{j=1}^p \mathbf{B}_{0j} \mathbf{h}_0$  has strictly positive coefficients, the matrices  $\mathbf{A}_{0i,+}$ ,  $\mathbf{A}_{0i,-}$  and  $\mathbf{B}_{0j}$  have size  $m \times m$  and positive coefficients.

The generic parameter value now consists of the coefficients of the vectors  $\boldsymbol{\epsilon}_+$ ,  $\boldsymbol{\epsilon}_-$ , and the matrices  $\mathbf{A}_{i*} := [\mathbf{A}_{i,+} \ \mathbf{A}_{i,-}]$  and  $\mathbf{B}_j$  (corresponding to the true values  $\boldsymbol{\epsilon}_{0+}, \boldsymbol{\epsilon}_{0-}, \mathbf{A}_{0i*} = [\mathbf{A}_{0i,+} \ \mathbf{A}_{0i,-}]$  and  $\mathbf{B}_{0j}$ , respectively), and the coefficients of the lower triangular part of  $\mathbf{R} = (\rho_{ij})$ . The parameter vector is denoted by  $\boldsymbol{\vartheta}_* = (\boldsymbol{\epsilon}'_*, \boldsymbol{\theta}'_*)'$ , where  $\boldsymbol{\epsilon}_* = (\boldsymbol{\epsilon}'_+, \boldsymbol{\epsilon}'_-)'$ ,  $\boldsymbol{\theta}_* = (\boldsymbol{\alpha}'_{1*}, \dots, \boldsymbol{\alpha}'_{q*}, \boldsymbol{\beta}'_1, \dots, \boldsymbol{\beta}'_p, \boldsymbol{\rho}'_1, \dots, \boldsymbol{\rho}'_p)'$  :=  $(\boldsymbol{\alpha}'_*, \boldsymbol{\beta}'_*, \boldsymbol{\rho}'_*)' \in \mathbb{R}^{s_0^*}$ , with  $\boldsymbol{\alpha}_{i*} = \text{vec } \mathbf{A}_{i*} \in \mathbb{R}^{2m^2}$ ,  $i = 1, \dots, q$  and  $s_0^* = s_0 + m^2 q$ .

The VTE of the parameter  $\boldsymbol{\epsilon}_{0*} = (\boldsymbol{\epsilon}'_{0+}, \boldsymbol{\epsilon}'_{0-})'$  is defined by

$$\widehat{\boldsymbol{\epsilon}}_{n*} = (\widehat{\boldsymbol{\epsilon}}'_{n+}, \widehat{\boldsymbol{\epsilon}}'_{n-})', \quad \widehat{\boldsymbol{\epsilon}}_{n+} = \frac{1}{n} \sum_{t=1}^n \boldsymbol{\epsilon}_t^+, \quad \widehat{\boldsymbol{\epsilon}}_{n-} = \frac{1}{n} \sum_{t=1}^n \boldsymbol{\epsilon}_t^-.$$

The VTE  $\boldsymbol{\epsilon}_{n*}$  of the parameter  $\boldsymbol{\theta}_{0*}$  is defined as in (3.1), but with  $\tilde{\ell}_{t,n} = \tilde{\ell}_t(\widehat{\boldsymbol{\epsilon}}_{n*}, \boldsymbol{\theta}_*)$  and  $\boldsymbol{\theta}_* \in \Theta_* \subset [0, \infty)^{m^2(p+2q)} \times (-1, 1)^{m(m-1)/2}$ . Let  $\widehat{\boldsymbol{\vartheta}}_{n*} = (\widehat{\boldsymbol{\epsilon}}'_{n*}, \widehat{\boldsymbol{\theta}}'_{n*})'$  be the VTE of  $\boldsymbol{\vartheta}_{0*}$ .

The next result provides the asymptotic distribution of  $\widehat{\boldsymbol{\vartheta}}_{n*}$ . To save space, notations and assumptions are deferred to the appendix.

**Theorem 4.1** (CAN of the VTE). *Under Assumptions A3, A5, A7 and B1-B4, the VTE of parameter  $\boldsymbol{\vartheta}_{0*}$  in the CCC-AGARCH( $p, q$ ) model satisfies*

$$\widehat{\boldsymbol{\vartheta}}_{n*} \rightarrow \boldsymbol{\vartheta}_{0*} \quad \text{a.s.} \quad \text{and} \quad \sqrt{n} \left( \widehat{\boldsymbol{\vartheta}}_{n*} - \boldsymbol{\vartheta}_{0*} \right) \xrightarrow{d} \mathcal{N}(0, \mathbf{G}_* \boldsymbol{\Sigma}_{X*} \mathbf{G}'_*) \quad \text{as } n \rightarrow \infty,$$

where  $\mathbf{G}_*$  and  $\boldsymbol{\Sigma}_{X*}$  are defined in Appendix 8.3.2.

## 5 Bootstrap approximation for the distribution of the VTE

It often happens, specially for statistics for which the asymptotic distribution is not easily estimated, that the bootstrap approximation gives a better insight of the actual distribution than the asymptotic theory. For this reason, a residual bootstrap procedure is proposed in this section.

Numerous bootstrap procedures have been used in time series analysis, but Shimizu (2013) notes that most articles investigating the validity of bootstrap techniques concern linear time series models (see Kreiss, Paparoditis and Politis (2011) and the references therein). Shimizu (2013) also points out that in empirical studies the limitation of the bootstrap tends to be underestimated. He studies several bootstrap procedures for univariate ARMA-GARCH models, and shows that they are not always consistent. We propose the following residual bootstrap procedure.

**Step 1** Compute the VTE  $\hat{\vartheta}_n$  and denote by  $\hat{F}_n$  the empirical distribution function of the rescaled residuals  $\mathbf{S}_n^{-1/2}(\hat{\boldsymbol{\eta}}_t - \boldsymbol{\mu}_n)$  for  $t = 1, \dots, n$ , where  $\hat{\boldsymbol{\eta}}_t = \tilde{\mathbf{H}}_t^{-1/2}(\hat{\boldsymbol{\vartheta}}_n)\boldsymbol{\epsilon}_t$ ,  $\boldsymbol{\mu}_n = n^{-1} \sum_{t=1}^n \hat{\boldsymbol{\eta}}_t$  and  $\mathbf{S}_n = n^{-1} \sum_{t=1}^n (\hat{\boldsymbol{\eta}}_t - \boldsymbol{\mu}_n)(\hat{\boldsymbol{\eta}}_t - \boldsymbol{\mu}_n)'$ . Note that almost surely,  $\mathbf{S}_n$  is invertible for  $n$  large enough.

**Step 2** Generate an iid bootstrap sample  $\boldsymbol{\eta}_1^*, \dots, \boldsymbol{\eta}_n^*$  with distribution  $\hat{F}_n$ .

To explain the last step, note that from (8.2) in the proof of Theorem 3.1 we have

$$\frac{\partial}{\partial \boldsymbol{\theta}} \tilde{\ell}_t(\boldsymbol{\vartheta}_0) = \tilde{\boldsymbol{\Delta}}_t(\boldsymbol{\vartheta}_0) \mathbf{V}_t + o(\rho^t), \quad a.s., \quad (5.1)$$

where  $\mathbf{V}_t = \text{vec}(\mathbf{I}_m - \mathbf{R}^{-1/2} \boldsymbol{\eta}_t \boldsymbol{\eta}_t' \mathbf{R}^{1/2})$  and  $\tilde{\boldsymbol{\Delta}}_t(\boldsymbol{\vartheta})$  is a measurable function of  $\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_t$ . More precisely (see the appendix) the line  $i \in \{1, \dots, (p+q)m^2\}$  of the  $s_0 \times m^2$  matrix  $\tilde{\boldsymbol{\Delta}}_t(\boldsymbol{\vartheta})$  is equal to the transpose of

$$2 \text{vec} \left( \tilde{\mathbf{D}}_t^{-1} \frac{\partial \tilde{\mathbf{D}}_t}{\partial \boldsymbol{\theta}_i} \right) = \text{vec} \left( \tilde{\mathbf{D}}_t^{-2} \frac{\partial \text{diag}(\tilde{\mathbf{h}}_t)}{\partial \boldsymbol{\theta}_i} \right),$$

and the line  $i \in \{(p+q)m^2 + 1, \dots, s_0\}$  is the transpose of  $\text{vec} \left( \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \boldsymbol{\theta}_i} \right)$ .

**Step 3** Define the bootstrapped estimator

$$\hat{\boldsymbol{\vartheta}}_n^* = \hat{\boldsymbol{\vartheta}}_n + \hat{\mathbf{G}} \frac{1}{n} \sum_{t=1}^n \mathbf{x}_{t,n}^*,$$

where

$$\mathbf{x}_{t,n}^* = \begin{pmatrix} \hat{\mathbf{C}} \{ \mathbf{U}_{t,n}^{*2} - \mathbf{I}_m \} \tilde{\mathbf{h}}_t(\hat{\boldsymbol{\vartheta}}_n) \\ \tilde{\boldsymbol{\Delta}}_t(\hat{\boldsymbol{\vartheta}}_n) \mathbf{V}_{t,n}^* \end{pmatrix},$$

with  $\widehat{\mathbf{C}} = \mathbf{C}(\widehat{\boldsymbol{\theta}}_n)$  and

$$\mathbf{U}_{t,n}^* = \text{diag}(\widehat{\mathbf{R}}^{1/2} \boldsymbol{\eta}_t^*), \quad \mathbf{V}_{t,n}^* = \text{vec}(\mathbf{I}_m - \widehat{\mathbf{R}}^{-1/2} \boldsymbol{\eta}_t^* \boldsymbol{\eta}_t^{*\prime} \widehat{\mathbf{R}}^{1/2}).$$

Note that the procedure does not require simulating and estimating multivariate GARCH models, which would be too time-consuming. The distribution of the VTE is mimicked by a Newton-Raphson type iteration. The following result shows the asymptotic validity of this procedure.

**Theorem 5.1** (Consistency of the bootstrap procedure). *Let the assumptions of Theorem 3.2 be satisfied. For almost all sequence  $(\boldsymbol{\epsilon}_t)$  satisfying (2.1), the bootstrap procedure is consistent in the sense that, given  $(\boldsymbol{\epsilon}_t)$ ,*

$$\sqrt{n} \left( \widehat{\boldsymbol{\vartheta}}_n^* - \widehat{\boldsymbol{\vartheta}}_n \right) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}) \quad \text{as } n \rightarrow \infty.$$

## 6 Adequacy test based on the VTE

In this section, we consider testing the validity of our model. We first introduce some additional notations. Let

$$\boldsymbol{\Sigma}_R = \mathbf{M} \boldsymbol{\Sigma}_Y \mathbf{M}', \quad \mathbf{M} = (\mathbf{H} - \mathbf{K}' \mathbf{J}^{-1} \mathbf{K}, -\mathbf{K}' \mathbf{J}^{-1}, \mathbf{I}_m), \quad \mathbf{H} = E \left( \frac{\partial^2 \ell_t(\boldsymbol{\vartheta}_0)}{\partial \mathbf{h} \partial \mathbf{h}'} \right),$$

and

$$\boldsymbol{\Sigma}_Y = \text{Var}(\mathbf{y}_t), \quad \mathbf{y}_t = \begin{pmatrix} \mathbf{x}_t \\ \frac{\partial}{\partial \mathbf{h}} \ell_t(\boldsymbol{\vartheta}_0) \end{pmatrix}.$$

In the spirit of the score test, our test is based on the derivative of the quasi-likelihood with respect to the first components of  $\boldsymbol{\vartheta}$ . If the model is correct, such derivatives should be small when evaluated at the VTE  $\widehat{\boldsymbol{\vartheta}}_n$ . More precisely, let the variance targeting test (VTT) statistic be defined as

$$R_n = \mathbf{T}_n' \widehat{\boldsymbol{\Sigma}}_R^{-1} \mathbf{T}_n, \quad \mathbf{T}_n = \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \mathbf{h}} \tilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n),$$

where  $\widehat{\boldsymbol{\Sigma}}_R$  is a consistent estimator of  $\boldsymbol{\Sigma}_R$  (for instance defined in (6.1) or (6.2) below), assuming it is invertible. The next result gives the asymptotic distribution of the statistic  $R_n$  under the null assumption

$$H_0 : \text{the CCC-GARCH}(p, q) \text{ model is correctly specified.}$$

**Theorem 6.1** (Asymptotic distribution of the adequacy test statistic). *Under  $H_0$  and the assumptions of Theorem 3.1, and if  $\widehat{\boldsymbol{\Sigma}}_R \rightarrow \boldsymbol{\Sigma}_R$  in probability, and  $\boldsymbol{\Sigma}_R$  is non singular, we have*

$$R_n \xrightarrow{d} \chi_m^2.$$

At the asymptotic level  $\alpha$ , the estimated CCC-GARCH( $p, q$ ) model can thus be rejected if  $R_n > \chi_m^2(1 - \alpha)$ , where  $\chi_m^2(1 - \alpha)$  is the  $(1 - \alpha)$ -quantile of the  $\chi^2$  distribution with  $m$  degrees of freedom. Comparing the proposed test with other adequacy tests (see references in the introduction) we note that, by using the VTE, we do not need to estimate the model by the full QML method.

## 6.1 Estimators of $\Sigma_R$

A natural estimator of  $\Sigma_R$  is obtained by replacing the moments involved in  $\mathbf{M}$  and  $\Sigma_Y$  by sample counterparts. Let

$$\widehat{\Sigma}_Y = \frac{1}{n} \sum_{t=1}^n \mathbf{r}_t \mathbf{r}_t', \quad \mathbf{r}_t = \begin{pmatrix} \mathbf{C}(\widehat{\boldsymbol{\theta}}_n) \{ \boldsymbol{\epsilon}_t - \widetilde{\mathbf{h}}_t(\widehat{\boldsymbol{\vartheta}}_n) \} \\ \frac{\partial}{\partial \boldsymbol{\theta}} \widetilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n) \\ \frac{\partial}{\partial \mathbf{h}} \widetilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n) \end{pmatrix}.$$

In addition to  $\widehat{\mathbf{J}}$  and  $\widehat{\mathbf{K}}$  defined in (3.4), introduce the matrices

$$\widehat{\mathbf{H}} = \frac{1}{n} \sum_{t=1}^n \frac{\partial^2}{\partial \mathbf{h} \partial \mathbf{h}'} \widetilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n)$$

and

$$\widehat{\mathbf{M}} = \begin{pmatrix} \widehat{\mathbf{H}} - \widehat{\mathbf{K}}' \widehat{\mathbf{J}}^{-1} \widehat{\mathbf{K}} & -\widehat{\mathbf{K}}' \widehat{\mathbf{J}}^{-1} & \mathbf{I}_m \end{pmatrix}.$$

We then define an estimator of  $\Sigma_R$  by

$$\widehat{\Sigma}_R = \widehat{\mathbf{M}} \widehat{\Sigma}_Y \widehat{\mathbf{M}}'. \quad (6.1)$$

An alternative estimator is obtained by adding a step to the bootstrap procedure given in Section 5. Similar to (5.1), we have

$$\frac{\partial}{\partial \mathbf{h}} \widetilde{\ell}_t(\boldsymbol{\vartheta}_0) = \widetilde{\boldsymbol{\Omega}}_t(\boldsymbol{\vartheta}_0) \mathbf{V}_t,$$

where  $\widetilde{\boldsymbol{\Omega}}_t(\boldsymbol{\vartheta})$  is a  $m \times m^2$  matrix, whose line  $i \in \{1, \dots, m\}$  is the transpose of  $2\text{vec}(\widetilde{\mathbf{D}}_t^{-1} \frac{\partial \widetilde{\mathbf{D}}_t}{\partial \boldsymbol{\vartheta}_i})$ . The additional step is the following.

**Step 4** Define the bootstrapped test statistic

$$\mathbf{T}_n^* = \widehat{\mathbf{M}} \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{r}_t^*, \quad \mathbf{r}_t^* = \begin{pmatrix} \widehat{\mathbf{C}} \{ \mathbf{U}_{t,n}^{*2} - \mathbf{I}_m \} \widetilde{\mathbf{h}}_t(\widehat{\boldsymbol{\vartheta}}_n) \\ \left( \begin{array}{c} \widetilde{\boldsymbol{\Delta}}_t(\widehat{\boldsymbol{\vartheta}}_n) \\ \widetilde{\boldsymbol{\Omega}}_t(\widehat{\boldsymbol{\vartheta}}_n) \end{array} \right) \mathbf{V}_{t,n}^* \end{pmatrix}.$$

The proof of the following result is identical to that of Theorem 5.1, and therefore it is omitted.

**Theorem 6.2** (Asymptotic validity of the bootstrap procedure). *Under the assumptions of Theorem 6.1, in particular  $H_0$ , and conditionally to almost all sequence  $(\epsilon_t)$  satisfying (2.1), we have*

$$\sqrt{n}\mathbf{T}_n^* \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}_R) \quad \text{as } n \rightarrow \infty.$$

The previous result shows that the matrix  $\boldsymbol{\Sigma}_R$  can be consistently estimated by

$$\widehat{\boldsymbol{\Sigma}}_R = E\left(\mathbf{T}_n^* \mathbf{T}_n^{*'} \mid \epsilon_1, \dots, \epsilon_n\right). \quad (6.2)$$

## 6.2 Univariate case

This test can in particular be employed in the univariate case ( $m = 1$ ). For simplicity, we focus on the GARCH(1,1) model, which is also the most widely used in practice. A simpler version of the test can be obtained in this case.

With  $\boldsymbol{\vartheta}_0 = (v_0, \alpha_0, \beta_0)'$  we have

$$\epsilon_t = \sqrt{h_t(\boldsymbol{\vartheta}_0)}\eta_t, \quad (\eta_t) \stackrel{iid}{\sim} (0, 1), \quad (6.3)$$

with, for  $v > 0$  and  $(\alpha, \beta)$  belonging to some compact parameter space  $\Theta \subset (0, \infty) \times (0, 1)$ ,

$$h_t(\boldsymbol{\vartheta}) = v + \alpha(\epsilon_{t-1}^2 - v) + \beta\{h_{t-1}(\boldsymbol{\vartheta}) - v\} = v + \sum_{i=0}^{\infty} \beta^i \alpha(\epsilon_{t-i-1}^2 - v).$$

The VTE of  $\boldsymbol{\vartheta}_0$  is  $\hat{\boldsymbol{\vartheta}}_n = (\frac{1}{n} \sum_{t=1}^n \epsilon_t^2, \hat{\alpha}_n, \hat{\beta}_n)'$  and we have

$$\mathbf{T}_n = \frac{1 - \hat{\alpha}_n - \hat{\beta}_n}{1 - \hat{\beta}_n} \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{\tilde{h}_t(\hat{\boldsymbol{\vartheta}}_n)} \left\{ 1 - \frac{\epsilon_t^2}{\tilde{h}_t(\hat{\boldsymbol{\vartheta}}_n)} \right\}.$$

**Theorem 6.3** (Adequacy test of the GARCH(1,1)). *Let  $(\epsilon_t)$  generated by (6.3) with  $(\alpha_0, \beta_0)$  belonging to the interior of  $\Theta$  and  $(\alpha_0 + \beta_0)^2 + (\kappa_4 - 1)\alpha_0^2 < 1$ , where  $\kappa_4 = E\eta_t^4$ . Assume that the distribution of  $\eta_t^2$  is not concentrated on a set of cardinality 2. Then  $\boldsymbol{\Sigma}_R$  is non singular and for any consistent estimator  $\widehat{\boldsymbol{\Sigma}}_R$  of  $\boldsymbol{\Sigma}_R$ , we have  $R_n \xrightarrow{d} \chi_1^2$ .*

We show in the proof that

$$\boldsymbol{\Sigma}_R = (\kappa_4 - 1)(\sigma^2 \delta + 1)\delta, \quad (6.4)$$

where  $\sigma^2 = \left(\frac{1 - \beta_0}{1 - \alpha_0 - \beta_0}\right)^2 E h_t^2$  and  $\delta = \mathbf{H} - \mathbf{K}' \mathbf{J}^{-1} \mathbf{K}$ . A consistent estimator of  $\boldsymbol{\Sigma}_R$  can be easily defined.

## 7 Numerical illustrations

### 7.1 Simulation experiments

To assess the performance of the adequacy test based on the VTE developed in Section 6, we first simulated  $N = 1,000$  independent trajectories of size  $n = 2,000$  and  $4,000$  of the univariate GARCH(1,1) model

$$\epsilon_t = \sqrt{h_t}\eta_t, \quad h_t = 0.03 + 0.09(\epsilon_{t-1}^2 - 0.03) + 0.89(h_{t-1} - 0.03).$$

The benchmark distribution for  $\eta_t$  is the standard Gaussian. It is however well known that the conditional distributions of the financial returns often exhibit non-normalities, in particular fat-tailedness and skewness. For these reasons, we considered the case where  $\eta_t$  follows the Asymmetric Exponential Power Distribution (AEPD) introduced by Zhu and Zinde-Walsh (2009). For the parameter of this distribution, we took the values estimated by Zhu and Zinde-Walsh on the S&P500 (namely  $\alpha = 0.4$ ,  $p_1 = 1.182$  and  $p_2 = 1.802$ , as in their Table 2), and the distribution has then been centered and reduced. We also tried other distributions for  $\eta_t$ , without qualitatively changing the outputs. For a GARCH(1,1), the 4th-order stationarity condition is  $\tau = (\alpha + \beta)^2 + (\kappa_4 - 1)\alpha^2 < 1$ . For both distributions this condition is satisfied: we get  $\tau = 0.9766$  for the Gaussian and  $\tau \simeq 0.985$  for the AEPD. In Table 1, the frequencies of rejection under the null assumption appear to be reasonably close to the nominal levels. Next, we simulated Markov-switching models of the form  $\epsilon_t = \sigma(\Delta_t)\eta_t$ , where  $(\eta_t) \stackrel{iid}{\sim} \mathcal{N}(0, 1)$  and  $(\Delta_t)$  is a two-state Markov chain independent of  $(\eta_t)$ . It should be noted that such models display similarities with the GARCH(1,1), in particular the absence of serial correlation for  $(\epsilon_t)$  together with the autocorrelation of  $(\epsilon_t^2)$ . Interestingly, the test strongly rejects the validity of the GARCH(1,1) model for these simulated Markov-switching models (see the right panel of Table 1).

Next, we turn to multivariate illustrations. Table 2 displays estimation results for 500 simulations of a bivariate CCC-GARCH(1,1) model. The two components of  $\boldsymbol{\eta}_t$  are independent and distributed according to the previous AEPD. The true parameter value is such that the matrix  $\mathbf{B}$  is diagonal, so the volatility of each component is related to its own lagged value, and to the lagged values of the squared observations of both components. However, in a first step, we do not take this information into account in the estimation. The results in Table 2 are in accordance with the consistency of the VTE, in particular the medians of the estimated parameters are very close to the true values. As expected, the accuracy increases as the sample size increases from  $n = 500$  to  $n = 4,000$ . Figure 1 shows non parametric estimators of the density of two components of  $\hat{\boldsymbol{\vartheta}}_n - \boldsymbol{\vartheta}_0$ . For the coefficient



Table 1: Empirical size and power of the univariate VTT for the null of a GARCH(1,1) model: over 1,000 independent replications, relative frequencies (in %) of rejection of the null, when the null is correct (models I and II) or when the null is incorrect (models III and IV).

Model	$n$	$\alpha$		Model	$n$	$\alpha$	
		1%	2.1			1%	66.4
I	2000	5%	6.2	III	2000	5%	83.2
		10%	9.7			10%	87.4
		1%	1.8			1%	66.4
II	2000	5%	5.8	IV	2000	5%	83.2
		10%	10.8			10%	87.4
		1%	1.0			1%	58.9
I	4000	5%	4.4	III	4000	5%	90.2
		10%	9.8			10%	92.3
		1%	0.5			1%	86.6
II	4000	5%	4.2	IV	4000	5%	93.2
		10%	8.4			10%	95.6

I: GARCH(1,1) with  $\boldsymbol{\vartheta}_0 = (0.03, 0.09, 0.89)$  and  $\eta_t \sim \mathcal{N}(0, 1)$

II: As model I, but  $\eta_t$  follows the AEPD of Zhu and Zinde-Walsh (2009)

III: Markov-switching model  $\epsilon_t = \sigma(\Delta_t)\eta_t$  where  $\eta_t \sim \mathcal{N}(0, 1)$ ,  $\sigma(1) = 1/200$ ,  $\sigma(2) = 3/200$ , and the Markov chain  $\Delta_t$  has transition probabilities  $p(1, 1) = p(2, 2) = 0.99$

IV: As model III, but with the transition probabilities  $p(1, 1) = p(2, 2) = 0.01$

$A(2, 2)$ , the estimated density is very close to a Gaussian, at least for  $n = 400$ . On the contrary, for the coefficient  $B(1, 2)$  a huge discrepancy is observed. This is not surprising as the true value of  $B(1, 2)$  is equal to zero. Thus  $\boldsymbol{\vartheta}_0$  does not belong to the interior of the parameter space and the VTE is not asymptotically Gaussian (see comments in Section 3.3).

Table 2: Sampling distribution of the VTE of  $\boldsymbol{\vartheta}_0$  over 500 replications for the CCC-GARCH(1,1) model with AEPD errors.

parameter	true val.	bias	RMSE	min	$Q_1$	$Q_2$	$Q_3$	max
$n = 500$								
$h$	10.00	0.08	2.17	6.32	8.65	9.85	10.95	23.10
	10.00	0.07	1.84	6.47	8.90	9.82	10.85	21.82
$\alpha$	0.10	0.00	0.05	0.00	0.06	0.09	0.13	0.26
	0.05	0.00	0.04	0.00	0.01	0.04	0.07	0.25
	0.05	0.01	0.05	0.00	0.02	0.05	0.08	0.23
	0.05	0.00	0.04	0.00	0.01	0.04	0.07	0.24
$\beta$	0.75	-0.11	0.22	0.00	0.59	0.70	0.77	0.94
	0.00	0.13	0.24	0.00	0.00	0.00	0.16	1.00
	0.00	0.11	0.21	0.00	0.00	0.00	0.16	0.86
	0.80	-0.17	0.32	0.00	0.55	0.74	0.81	0.97
$\rho$	0.80	0.00	0.02	0.74	0.79	0.80	0.82	0.86
$n = 4,000$								
$h$	10.00	0.04	0.72	8.44	9.54	9.98	10.48	13.01
	10.00	0.01	0.63	8.54	9.58	9.98	10.43	12.98
$\alpha$	0.10	0.00	0.02	0.06	0.09	0.10	0.12	0.17
	0.05	0.00	0.02	0.00	0.04	0.05	0.06	0.12
	0.05	0.00	0.02	0.00	0.04	0.05	0.06	0.10
	0.05	0.00	0.02	0.00	0.04	0.05	0.06	0.11
$\beta$	0.75	0.06	0.13	0.00	0.68	0.73	0.75	0.82
	0.00	0.08	0.17	0.00	0.00	0.00	0.11	0.70
	0.00	0.07	0.15	0.00	0.00	0.00	0.09	0.84
	0.80	0.10	0.21	0.00	0.68	0.77	0.80	0.87
$\rho$	0.80	0.00	0.01	0.78	0.80	0.80	0.80	0.82

RMSE is the Root Mean Square Error,  $Q_i$ ,  $i = 1, 3$ , denote the quartiles.

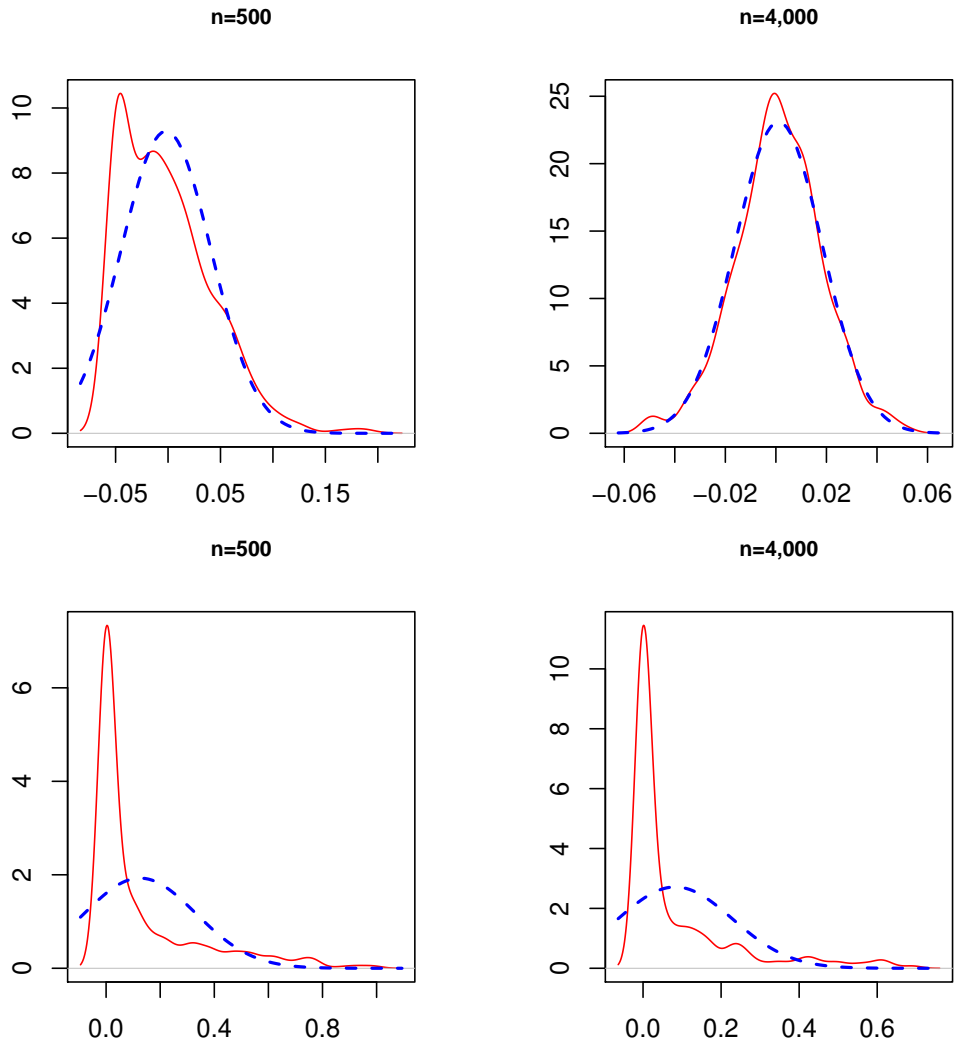


Figure 1: Kernel density estimator (in full line) of the distribution of the VTE errors for the estimation of  $\mathbf{A}(2,2)$  (top panel) and  $\mathbf{B}(2,1)$  (bottom panel), and gaussian density (in dotted line) with the same mean and variance.

In a second step, we conducted the same experiments but the model was now estimated under the assumption that the matrix  $\mathbf{B}$  was diagonal. The results displayed in Table 3 are, unsurprisingly, better than with the unrestricted estimator. Figure 2 shows that the estimated densities of the estimators over the 500 simulations are now close to a Gaussian density for  $n$  sufficiently large.

Next, we performed a Monte Carlo experiment with the aim to compare the empirical accuracies of the VTE and QMLE. We simulated a model close to that estimated on the real series of the next section. More precisely, we simulated 100 independent replications of a 3-dimensional CCC-GARCH(1,1) model, with diagonal matrices  $\mathbf{A} = \text{diag}(0.04, 0.04, 0.04)$  and

Table 3: As Table 2, but for the constrained VTE.

parameter	true val.	bias	RMSE	min	$Q_1$	$Q_2$	$Q_3$	max
$n = 500$								
$h$	10.00	-0.01	1.79	6.21	8.68	9.78	10.99	17.74
	10.00	-0.02	1.57	6.57	8.80	9.81	11.00	16.11
$\alpha$	0.10	0.00	0.05	0.00	0.06	0.09	0.13	0.24
	0.05	0.01	0.04	0.00	0.03	0.05	0.08	0.26
	0.05	0.01	0.04	0.00	0.03	0.05	0.08	0.20
	0.05	0.00	0.04	0.00	0.02	0.05	0.07	0.18
$\beta$	0.75	-0.03	0.11	0.00	0.66	0.73	0.78	0.96
	0.80	-0.05	0.14	0.00	0.71	0.78	0.84	0.99
$\rho$	0.80	0.00	0.02	0.74	0.79	0.80	0.81	0.85
$n = 4,000$								
$h$	10.00	-0.02	0.70	8.16	9.51	9.89	10.43	12.85
	10.00	-0.01	0.59	8.57	9.58	9.94	10.33	12.27
$\alpha$	0.10	0.00	0.02	0.05	0.09	0.10	0.11	0.15
	0.05	0.00	0.01	0.01	0.04	0.05	0.06	0.08
	0.05	0.00	0.02	0.01	0.04	0.05	0.06	0.12
	0.05	0.00	0.01	0.01	0.04	0.05	0.06	0.10
$\beta$	0.75	0.00	0.03	0.67	0.73	0.75	0.77	0.82
	0.80	0.00	0.03	0.70	0.78	0.80	0.82	0.87
$\rho$	0.80	0.00	0.01	0.78	0.80	0.80	0.81	0.82

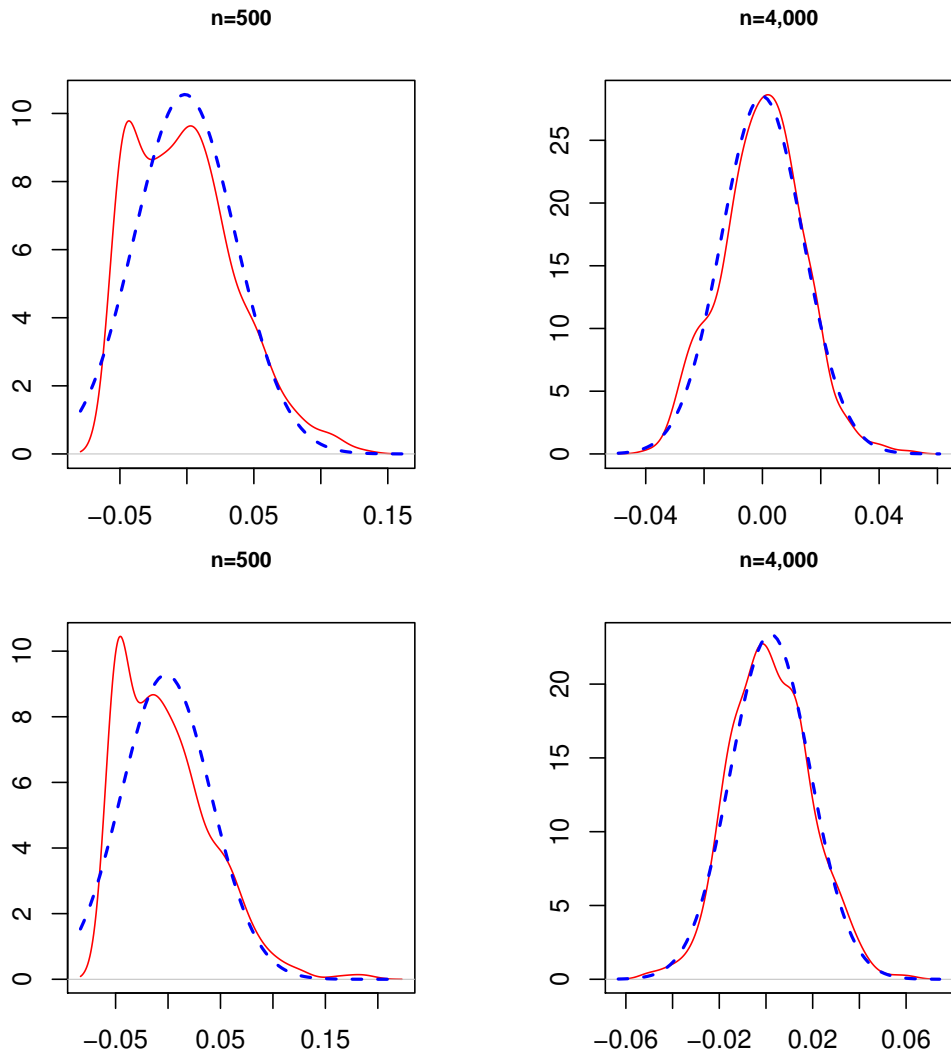


Figure 2: As Figure 1 but for the constrained VTE of  $\mathbf{A}(2,2)$  (top panel) and  $\mathbf{B}(2,2)$  (bottom panel).

$\mathbf{B} = \text{diag}(0.93, 0.93, 0.93)$ , with  $\boldsymbol{\omega} = (1/3, 1/3, 1/3)'$  and with  $\text{vech}(\mathbf{R}) = (0.8, 0.64, 0.8)'$ . As distribution for  $\boldsymbol{\eta}_t$ , we still took the AEPD of Zhu and Zinde-Walsh. Figure 3 displays the distribution of the estimation errors for simulations of length  $n = 5000$ . The upper-left, upper-right, bottom-left and bottom-right panels correspond respectively to the estimation errors for the 3 parameters involved in  $\mathbf{h}$ ,  $\mathbf{A}$ ,  $\mathbf{B}$  and  $\mathbf{R}$ . The distributions of the VTE and QMLE are very similar. For the simulation length  $n = 500$  (not presented here to save space), the distributions of the two estimations are also equivalent, but the estimators are of course less accurate.

To complete the section, we compared the computation time of the VTE and QMLE, on the same CCC-GARCH(1,1) models as in the previous experiment, but with  $m$  varying from 2 to 5. Table 4 shows that, as expected, the VTE is more rapid than the QMLE. The computation time increases rapidly with  $m$ , but the relative time-computation gain does not depend much on  $m$ , nor on  $n$ . When the number  $m$  of individual components is too large with respect to  $n$  (say,  $m \geq 5$

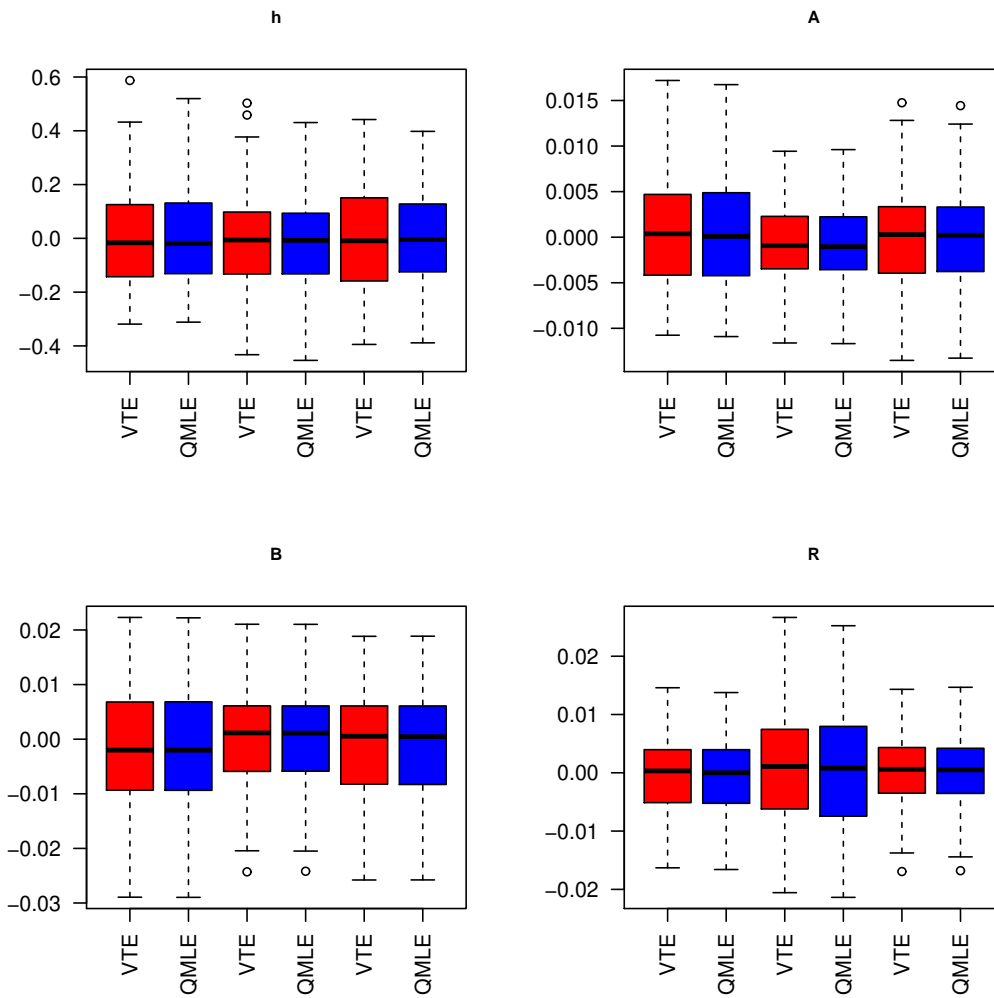


Figure 3: Boxplot of the estimation errors for the VTE and QMLE

Table 4: Seconds of CPU time for computing the VTE and QMLE

	$n = 500$				$n = 5000$			
	$m = 2$	$m = 3$	$m = 4$	$m = 5$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
VTE	2.92	7.44	16.66	24.38	33.93	97.12	205.50	186.91
QMLE	4.98	13.62	31.23	35.24	61.07	136.81	263.74	269.99

when  $n = 500$ ), both estimators give very inaccurate estimates, and the comparison of the time computations does not make sense anymore. Indeed, when the model contains too many parameters, the optimization routine (we used the function `nlmib` of R) generally cannot progress, and it returns (sometimes relatively rapidly) final estimates that are close to the initial values of the parameters.

## 7.2 An application to exchange rates

We consider log-returns series of the daily exchange rates of the American Dollar (USD), the Japanese Yen (JPY) and the British Pound (GBP) versus the Euro. The observations have been downloaded from the website of the European Central Bank<sup>5</sup>, and cover the period from January 5, 1999 to January 22, 2013, which corresponds to 3601 observations. A CCC-GARCH(1,1) model has been fitted by VTE to the trivariate series  $\epsilon_t = (\text{USD}_t, \text{JPY}_t, \text{GBP}_t)'$ .

With obvious notations (in particular the estimated standard deviations, obtained from the empirical estimator (3.5), are into brackets), the estimated parameters can be written as

$$\hat{\mathbf{h}} = \begin{pmatrix} 0.44 & (0.07) \\ 0.65 & (0.14) \\ 0.25 & (0.05) \end{pmatrix} \begin{matrix} \text{USD} \\ \text{JPY} \\ \text{GBP} \end{matrix}$$

$$\hat{\mathbf{A}} = \begin{pmatrix} 0.03 & (0.01) & 0.00 & (0.01) & 0.00 & (0.02) \\ 0.01 & (0.02) & 0.07 & (0.03) & 0.00 & (0.04) \\ 0.01 & (0.01) & 0.01 & (0.01) & 0.06 & (0.02) \end{pmatrix} \begin{matrix} \text{USD} \\ \text{JPY} \\ \text{GBP} \end{matrix}$$

USD                  JPY                  GBP

$$\hat{\mathbf{B}} = \begin{pmatrix} 0.94 & (0.01) & 0.00 & (0.04) & 0.02 & (0.10) \\ 0.00 & (0.03) & 0.86 & (0.00) & 0.14 & (0.15) \\ 0.01 & (0.02) & 0.00 & (0.03) & 0.87 & (0.01) \end{pmatrix} \begin{matrix} \text{USD} \\ \text{JPY} \\ \text{GBP} \end{matrix}$$

USD                  JPY                  GBP

<sup>5</sup><http://www.ecb.int/home/html/index.en.html>

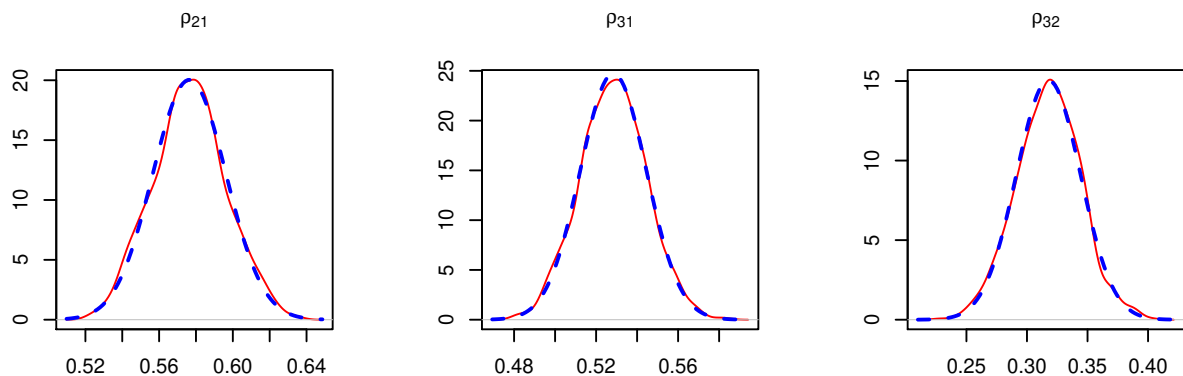


Figure 4: Kernel density estimators for the bootstrap distribution of the VTE of  $\rho_0$  (full line), and gaussian density (dotted line) with the same mean and variances.

and, introducing the half-vectorization operator<sup>6</sup>,

$$\text{vech}(\hat{\mathbf{R}}_n) = \begin{pmatrix} 0.58 & (0.02) \\ 0.53 & (0.01) \\ 0.32 & (0.02) \end{pmatrix} \begin{array}{l} \text{USD} - \text{JPY} \\ \text{USD} - \text{GBP} \\ \text{JPY} - \text{GBP} \end{array}$$

When using the empirical estimate of  $\Sigma_R$  defined in (6.1), the adequacy test statistic of Theorem 6.1 is equal to  $R_n = 0.838$ , corresponding to the  $p$ -value  $P(\chi_3^2 > 0.838) = 0.84$ . Using the bootstrap estimate of  $\Sigma_R$  defined in (6.2), with 1,000 bootstrap replications, we obtain  $R_n = 0.69$ , which gives the  $p$ -value 0.88. Therefore the CCC-GARCH(1,1) model is not rejected by the VTT.

The bootstrap estimates of the standard deviations are not given because, as expected, they are very similar to those displayed into brackets. Figure 4 displays the bootstrap approximation for the distributions of the estimators of the correlation matrix elements  $\rho_0$ . This figure clearly shows that the error terms of the GARCH model are significantly positively correlated. This is not surprising since the three exchange rates are against the same currency, and thus exhibit comovements. Indeed, when the euro depreciates, the depreciation is often with respect to the three other currencies.

Note that, none of the off-diagonal elements of the matrices  $\mathbf{A}_0$  and  $\mathbf{B}_0$  seems to be significantly non zero. However, one has to be cautious in the interpretation of  $t$ -test statistics because Theorem 3.2 requires that the parameter belong to the interior of the parameter space (see Francq and Zakoïan (2009) for testing zero coefficients in univariate GARCH models). However, the results of

<sup>6</sup> $\overline{\text{vech}(\mathbf{A})}$  of a symmetric  $n \times n$  matrix  $\mathbf{A}$  is the  $n(n+1)/2 \times 1$  column vector obtained by vectorizing only the lower triangular part of  $\mathbf{A}$ .



Section 3.3 allow to re-estimate the CCC-GARCH(1,1) model with the constraint that the matrices  $\mathbf{A}_0$  and  $\mathbf{B}_0$  be diagonal. The estimated parameters of the constrained model are

$$\hat{\mathbf{h}} = \begin{pmatrix} 0.44 & (0.058) \\ 0.65 & (0.142) \\ 0.25 & (0.067) \end{pmatrix} \begin{matrix} \text{USD} \\ \text{JPY} \\ \text{GBP} \end{matrix}$$

$$\hat{\mathbf{A}} = \begin{pmatrix} 0.025 & (0.003) & 0 & 0 \\ 0 & & 0.054 & (0.006) & 0 \\ 0 & & 0 & & 0.052 & (0.003) \end{pmatrix} \begin{matrix} \text{USD} \\ \text{JPY} \\ \text{GBP} \\ \text{USD} & \text{JPY} & \text{GBP} \end{matrix}$$

$$\hat{\mathbf{B}} = \begin{pmatrix} 0.966 & (0.001) & 0 & 0 \\ 0 & & 0.935 & (0.005) & 0 \\ 0 & & 0 & & 0.939 & (0.004) \end{pmatrix} \begin{matrix} \text{USD} \\ \text{JPY} \\ \text{GBP} \\ \text{USD} & \text{JPY} & \text{GBP} \end{matrix}$$

and

$$\text{vech}(\hat{\mathbf{R}}) = \begin{pmatrix} 0.595 & (0.020) \\ 0.530 & (0.016) \\ 0.320 & (0.023) \end{pmatrix} \begin{matrix} \text{USD} - \text{JPY} \\ \text{USD} - \text{GBP} \\ \text{JPY} - \text{GBP} \end{matrix}$$

The estimated standard deviations were obtained using Theorem 3.3. As expected, the estimated coefficients are very close to those of the unconstrained model and the accuracy increased. The opposite of the log-likelihood of the full model is 1659.186 and that of the constrained model is 1656.311 (up to some unimportant constant). If one uses a standard likelihood-ratio test, the difference does not appear to be significant, and the constrained model is not rejected at any reasonable significance level. However, the caveat concerning the reliability of asymptotic tests when the parameter may lie on the boundary of the parameter space applies.

The conclusion from this empirical study is that the relationships between the different exchange rates are mainly instantaneous. The conditional correlations are strongly positive while the spillover effects in the volatilities are weak.

## 8 Proofs

### 8.1 Proof of the consistency of $\vartheta_n$ in Theorem 3.2

The strong convergence of  $\widehat{\mathbf{h}}_n$  to  $\mathbf{h}_0$  is a direct consequence of Assumption **A** and the ergodic theorem. To complete the proof it suffices to show that

$$\sup_{\boldsymbol{\theta} \in \Theta} \left| \widetilde{\mathcal{L}}_n(\boldsymbol{\theta}) - n^{-1} \sum_{t=1}^n \widetilde{\ell}_t(\mathbf{h}_0, \boldsymbol{\theta}) \right| \rightarrow 0 \quad a.s. \quad (8.1)$$

The mean value theorem yields

$$\sup_{\boldsymbol{\theta} \in \Theta} \left| \ell_t(\widehat{\mathbf{h}}_n, \boldsymbol{\theta}) - \ell_t(\mathbf{h}_0, \boldsymbol{\theta}) \right| \leq \sup_{\boldsymbol{\vartheta} \in \Xi} \left| \frac{\partial}{\partial \mathbf{h}} \ell_t(\boldsymbol{\vartheta}) \right| |\widehat{\mathbf{h}}_n - \mathbf{h}_0|,$$

where  $\Xi = \Lambda \times \Theta$ , with  $\Lambda$  a compact subset of  $(0, \infty)^m$  whose interior contains  $\mathbf{h}_0$ . Moreover,

$$\sup_{\boldsymbol{\theta} \in \Theta} \left| \widetilde{\ell}_t(\widehat{\mathbf{h}}_n, \boldsymbol{\theta}) - \widetilde{\ell}_t(\mathbf{h}_0, \boldsymbol{\theta}) \right| \leq 2 \sup_{\boldsymbol{\vartheta} \in \Xi} \left| \widetilde{\ell}_t(\boldsymbol{\vartheta}) - \ell_t(\boldsymbol{\vartheta}) \right| + \sup_{\boldsymbol{\theta} \in \Theta} \left| \ell_t(\widehat{\mathbf{h}}_n, \boldsymbol{\theta}) - \ell_t(\mathbf{h}_0, \boldsymbol{\theta}) \right|.$$

In view of the consistency of  $\widehat{\mathbf{h}}_n$ , the convergence in (8.1) will thus follow from the Cesàro lemma and

$$\sup_{\boldsymbol{\vartheta} \in \Xi} \left| \widetilde{\ell}_t(\boldsymbol{\vartheta}) - \ell_t(\boldsymbol{\vartheta}) \right| \rightarrow 0 \quad a.s. \quad \text{and} \quad E \sup_{\boldsymbol{\vartheta} \in \Xi} \left| \frac{\partial}{\partial \mathbf{h}} \ell_t(\boldsymbol{\vartheta}) \right| < \infty,$$

The a.s. convergence was established in the proof of Theorem 11.7 in Francq and Zakoian (2010, page 295), and the latter inequality follows from formula (11.67) on page 300 of the previous reference.

### 8.2 Proof of the asymptotic normality of $\vartheta_n$ in Theorem 3.2

Introduce the martingale difference

$$\boldsymbol{\nu}_t = \underline{\boldsymbol{\epsilon}}_t - \underline{\mathbf{h}}_t = (\mathbf{U}_t^2 - \mathbf{I}_m) \underline{\mathbf{h}}_t,$$

where  $\mathbf{U}_t = \text{diag}(\mathbf{R}_0^{1/2} \boldsymbol{\eta}_t)$ . Replacing  $\underline{\mathbf{h}}_{t-j}$  by  $\underline{\boldsymbol{\epsilon}}_{t-j} - \boldsymbol{\nu}_{t-j}$ ,  $j = 0, \dots, p$ , in the last equation of (2.1), we get the ARMA equation

$$\underline{\boldsymbol{\epsilon}}_t - \mathbf{h}_0 = \sum_{i=1}^r (\mathbf{A}_{0i} + \mathbf{B}_{0i}) (\underline{\boldsymbol{\epsilon}}_{t-i} - \mathbf{h}_0) + \boldsymbol{\nu}_t - \sum_{i=1}^p \mathbf{B}_{0i} \boldsymbol{\nu}_{t-i}.$$

Taking the average of the two sides of the equality for  $t = 1, \dots, n$  leads to

$$\widehat{\mathbf{h}}_n - \mathbf{h}_0 = \left\{ \mathbf{I}_m - \sum_{i=1}^r (\mathbf{A}_{0i} + \mathbf{B}_{0i}) \right\}^{-1} \left( \mathbf{I}_m - \sum_{i=1}^p \mathbf{B}_{0i} \right) n^{-1} \sum_{t=1}^n \boldsymbol{\nu}_t + o(1), \quad a.s.$$

The central limit theorem (CLT) for square integrable stationary martingale differences (see Billingsley (1961)) entails that

$$\sqrt{n}(\widehat{\mathbf{h}}_n - \mathbf{h}_0) = \frac{\mathbf{C}}{\sqrt{n}} \sum_{t=1}^n (\mathbf{U}_t^2 - \mathbf{I}_m) \underline{\mathbf{h}}_t + o_P(1) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}_h),$$

where  $\mathbf{C} = \mathbf{C}(\boldsymbol{\theta}_0)$ , as defined before Theorem 3.2, and

$$\boldsymbol{\Sigma}_h = \mathbf{C} E \{ (\mathbf{U}_t^2 - \mathbf{I}_m) E(\underline{\mathbf{h}}_t \underline{\mathbf{h}}_t') (\mathbf{U}_t^2 - \mathbf{I}_m) \} \mathbf{C}'.$$

Noting that  $\tilde{\ell}_{t,n}(\boldsymbol{\theta}) = \tilde{\ell}_t(\widehat{\mathbf{h}}_n, \boldsymbol{\theta})$ , we have

$$\begin{aligned} \mathbf{0}_{s_0} &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \boldsymbol{\theta}} \tilde{\ell}_{t,n}(\widehat{\boldsymbol{\theta}}_n) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \boldsymbol{\theta}} \tilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \boldsymbol{\theta}} \tilde{\ell}_t(\boldsymbol{\vartheta}_0) + \left( \frac{1}{n} \sum_{t=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j} \tilde{\ell}_t(\bar{\boldsymbol{\vartheta}}_i) \right)_{s_0 \times s_1} \sqrt{n}(\widehat{\boldsymbol{\vartheta}}_n - \boldsymbol{\vartheta}_0) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \boldsymbol{\theta}} \tilde{\ell}_t(\boldsymbol{\vartheta}_0) + \mathbf{J}_n \sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) + \mathbf{K}_n \sqrt{n}(\widehat{\mathbf{h}}_n - \mathbf{h}_0), \end{aligned}$$

where, for some  $\bar{\boldsymbol{\vartheta}}_i$  between  $\widehat{\boldsymbol{\vartheta}}_n$  and  $\boldsymbol{\vartheta}_0$ ,

$$\mathbf{J}_n = \left( \frac{1}{n} \sum_{t=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j} \tilde{\ell}_t(\bar{\boldsymbol{\vartheta}}_i) \right)_{s_0 \times s_0}, \quad \mathbf{K}_n = \left( \frac{1}{n} \sum_{t=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta}_i \partial h_j} \tilde{\ell}_t(\bar{\boldsymbol{\vartheta}}_i) \right)_{s_0 \times m}.$$

In view of (11.69) and (11.70) in Francq and Zakoian (2010), and using the elementary equality  $\text{Trace}(A'B) = (\text{vec } A)' \text{vec } B$ , we also have

$$\frac{\partial}{\partial \boldsymbol{\theta}} \tilde{\ell}_t(\boldsymbol{\vartheta}_0) = \boldsymbol{\Delta}_t \mathbf{V}_t + o(\rho^t) \quad \text{a.s.}, \quad (8.2)$$

with some  $\rho \in [0, 1)$ ,

$$\mathbf{V}_t = \text{vec}(\mathbf{I}_m - \mathbf{R}^{-1/2} \boldsymbol{\eta}_t \boldsymbol{\eta}_t' \mathbf{R}^{1/2})$$

and  $\boldsymbol{\Delta}_t$  is a  $s_0 \times m^2$  matrix, which is measurable with respect to the sigma-field generated by  $\{\boldsymbol{\eta}_u, u < t\}$ . Because  $\mathbf{J}$  is positive-definite,  $\mathbf{J}_n$  is a.s. non-singular for sufficiently large  $n$ . It follows that, for  $n$  large enough

$$\sqrt{n}(\widehat{\boldsymbol{\vartheta}}_n - \boldsymbol{\vartheta}_0) = \begin{pmatrix} \mathbf{I}_m & 0 \\ -\mathbf{J}_n^{-1} \mathbf{K}_n & -\mathbf{J}_n^{-1} \end{pmatrix} \mathbf{X}_n$$

where

$$\mathbf{X}_n := \begin{pmatrix} \sqrt{n}(\widehat{\mathbf{h}}_n - \mathbf{h}_0) \\ \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \boldsymbol{\theta}} \tilde{\ell}_t(\boldsymbol{\vartheta}_0) \end{pmatrix} = \begin{pmatrix} \frac{\mathbf{C}}{\sqrt{n}} \sum_{t=1}^n (\mathbf{U}_t^2 - \mathbf{I}_m) \underline{\mathbf{h}}_t \\ \frac{1}{\sqrt{n}} \sum_{t=1}^n \boldsymbol{\Delta}_t \mathbf{V}_t \end{pmatrix} + o_P(1).$$

Using again the CLT for square integrable stationary martingale differences (see Billingsley (1961)), we have

$$\mathbf{X}_n \xrightarrow{d} \mathcal{N} \left\{ 0, \boldsymbol{\Sigma}_X = \begin{pmatrix} \boldsymbol{\Sigma}_h & \mathbf{L} \\ \mathbf{L}' & \mathbf{I} \end{pmatrix} \right\},$$

with  $\mathbf{L} = E\{\mathbf{C}(U_t^2 - I_m)\underline{\mathbf{h}}_t \mathbf{V}_t' \boldsymbol{\Delta}_t' \mathbf{C}'\}$ . The conclusion follows.

### 8.3 Assumptions and the proof of Theorem 4.1

#### 8.3.1 Notations and assumptions

Let  $\mathcal{A}_{\theta_*}^+(z) = \sum_{i=1}^q \mathbf{A}_{i,+} z^i$ ,  $\mathcal{A}_{\theta_*}^-(z) = \sum_{i=1}^q \mathbf{A}_{i,-} z^i$  and  $\mathcal{B}_{\theta_*}(z) = \mathbf{I}_m - \sum_{j=1}^p \mathbf{B}_j z^j$ . We refer to Proposition 3.1 in Francq and Zakoian (2012) for the definition of the matrix  $M_1(\mathcal{A}_{\theta_*}^+, \mathcal{A}_{\theta_*}^-, \mathcal{B}_{\theta_*})$  introduced below.

**B1:**  $\theta_{0*} \in \overset{\circ}{\Theta}_*$  and  $\Theta_*$  is compact.

**B2:**  $\forall \theta_* \in \Theta_*$ ,  $|\mathcal{B}_{\theta_*}(z)| = 0 \Rightarrow |z| > 1$ .

**B3:** If  $p > 0$ ,  $\mathcal{A}_{\theta_{0*}}^+(1) + \mathcal{A}_{\theta_{0*}}^-(1) \neq 0$ ,  $\mathcal{A}_{\theta_{0*}}^+(z)$ ,  $\mathcal{A}_{\theta_{0*}}^-(z)$  and  $\mathcal{B}_{\theta_{0*}}(z)$  are left coprime and the matrix  $M_1(\mathcal{A}_{\theta_{0*}}^+, \mathcal{A}_{\theta_{0*}}^-, \mathcal{B}_{\theta_{0*}})$  has full rank  $m$ .

**B4:**  $(\boldsymbol{\epsilon}_t)$  is a strictly stationary solution of Model (4) and satisfies  $\|\boldsymbol{\epsilon}_t\|_4 < \infty$ .

#### 8.3.2 Proof of Theorem 4.1

The consistency can be established by the arguments used in the proof of Theorem 3.2. We omit the details. For the asymptotic normality, introduce the martingale difference

$$\boldsymbol{\nu}_t^* = \begin{pmatrix} \boldsymbol{\nu}_t^+ \\ \boldsymbol{\nu}_t^- \end{pmatrix} = \underline{\boldsymbol{\epsilon}}_t^* - \boldsymbol{\Gamma} \underline{\mathbf{h}}_t = (\mathbf{U}_t^{(2)} - \boldsymbol{\Gamma}) \underline{\mathbf{h}}_t,$$

where

$$\underline{\boldsymbol{\epsilon}}_t^* = \begin{pmatrix} \boldsymbol{\epsilon}_t^+ \\ \boldsymbol{\epsilon}_t^- \end{pmatrix}, \quad \boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{\Gamma}_+ \\ \boldsymbol{\Gamma}_- \end{pmatrix} = E \mathbf{U}_t^{(2)}, \quad \mathbf{U}_t^{(2)} = \begin{pmatrix} (\mathbf{U}_t^+)^2 \\ (\mathbf{U}_t^-)^2 \end{pmatrix},$$

$$\mathbf{U}_t^+ = \text{diag}(\eta_{1t}^+, \dots, \eta_{mt}^+), \quad \mathbf{U}_t^- = \text{diag}(\eta_{1t}^-, \dots, \eta_{mt}^-).$$

We have

$$\begin{aligned}\underline{\boldsymbol{\epsilon}}_t^+ - \underline{\boldsymbol{\epsilon}}_{0+} &= \boldsymbol{\Gamma}_+(\underline{\boldsymbol{h}}_t - \underline{\boldsymbol{h}}_0) + \boldsymbol{\nu}_t^+ \\ &= \sum_{i=1}^r \boldsymbol{\Gamma}_+(\mathbf{A}_{0i,+} + \mathbf{B}_{0i}\boldsymbol{\Gamma}_+^{-1})(\underline{\boldsymbol{\epsilon}}_{t-i}^+ - \underline{\boldsymbol{\epsilon}}_{0+}) + \boldsymbol{\Gamma}_+\mathbf{A}_{0i,-}(\underline{\boldsymbol{\epsilon}}_{t-i}^- - \underline{\boldsymbol{\epsilon}}_{0-}) \\ &\quad + \boldsymbol{\nu}_t^+ - \sum_{j=1}^p \boldsymbol{\Gamma}_+\mathbf{B}_{0j}\boldsymbol{\Gamma}_+^{-1}\boldsymbol{\nu}_{t-j}^+, \end{aligned}$$

and a similar equation for  $\underline{\boldsymbol{\epsilon}}_t^- - \underline{\boldsymbol{\epsilon}}_{0-}$ . It follows that

$$\underline{\boldsymbol{\epsilon}}_t^* - \underline{\boldsymbol{\epsilon}}_{0*} = \sum_{i=1}^r \mathbf{C}_{0i,*}(\underline{\boldsymbol{\epsilon}}_{t-i}^* - \underline{\boldsymbol{\epsilon}}_{0*}) + \boldsymbol{\nu}_t^* - \sum_{j=1}^p \mathbf{B}_{0j,*}\boldsymbol{\nu}_{t-j}^*,$$

where

$$\mathbf{C}_{0i,*} = \begin{pmatrix} \boldsymbol{\Gamma}_+(\mathbf{A}_{0i,+} + \mathbf{B}_{0i}\boldsymbol{\Gamma}_+^{-1}) & \boldsymbol{\Gamma}_+\mathbf{A}_{0i,-} \\ \boldsymbol{\Gamma}_-\mathbf{A}_{0i,+} & \boldsymbol{\Gamma}_-(\mathbf{A}_{0i,-} + \mathbf{B}_{0i}\boldsymbol{\Gamma}_-^{-1}) \end{pmatrix}, \quad \mathbf{B}_{0j,*} = \begin{pmatrix} \mathbf{B}_{0j} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{0j} \end{pmatrix}.$$

Taking the average of the two sides of the equality for  $t = 1, \dots, n$  leads to

$$\widehat{\underline{\boldsymbol{\epsilon}}}_{n*} - \underline{\boldsymbol{\epsilon}}_{0*} = \frac{\mathbf{C}_*}{n} \sum_{t=1}^n \boldsymbol{\nu}_t^* + o(1), \quad a.s.$$

where

$$\mathbf{C}_* = \left\{ \mathbf{I}_{2m} - \sum_{i=1}^r \mathbf{C}_{0i,*} \right\}^{-1} \left( \mathbf{I}_{2m} - \sum_{j=1}^p \mathbf{B}_{0j,*} \right).$$

By arguments already used,

$$\sqrt{n}(\widehat{\underline{\boldsymbol{\epsilon}}}_{n*} - \underline{\boldsymbol{\epsilon}}_{0*}) = \frac{\mathbf{C}_*}{\sqrt{n}} \sum_{t=1}^n (\mathbf{U}_t^{(2)} - \boldsymbol{\Gamma})\underline{\boldsymbol{h}}_t + o_P(1) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}_h^*),$$

where

$$\boldsymbol{\Sigma}_h^* = \mathbf{C}_* E \left\{ (\mathbf{U}_t^{(2)} - \boldsymbol{\Gamma}) E(\underline{\boldsymbol{h}}_t \underline{\boldsymbol{h}}_t') (\mathbf{U}_t^{(2)} - \boldsymbol{\Gamma})' \right\} \mathbf{C}_*'.$$

The rest of the proof is similar to the proof of the asymptotic normality of  $\boldsymbol{\vartheta}_n$  in Theorem 3.2.

Therefore it is omitted. We will only define the matrices required to derive the asymptotic variance

in Theorem 4.1. Let

$$\mathbf{G}_* = \begin{pmatrix} \mathbf{I}_{2m} & \mathbf{0} \\ -\mathbf{J}_*^{-1}\mathbf{K}_* & -\mathbf{J}_*^{-1} \end{pmatrix}, \quad \mathbf{J}_* = E \left( \frac{\partial^2 \ell_t(\boldsymbol{\vartheta}_{0*})}{\partial \boldsymbol{\theta}_* \partial \boldsymbol{\theta}_*'} \right), \quad \mathbf{K}_* = E \left( \frac{\partial^2 \ell_t(\boldsymbol{\vartheta}_{0*})}{\partial \boldsymbol{\theta}_* \partial \underline{\boldsymbol{\epsilon}}_*'} \right),$$

$$\boldsymbol{\Sigma}_{X^*} = \text{Var}(\mathbf{x}_t^*), \quad \mathbf{x}_t^* = \begin{pmatrix} \mathbf{C}_* \{ \underline{\boldsymbol{\epsilon}}_t - \underline{\boldsymbol{h}}_t(\boldsymbol{\vartheta}_{0*}) \} \\ \frac{\partial}{\partial \boldsymbol{\theta}_*} \ell_t(\boldsymbol{\vartheta}_{0*}) \end{pmatrix},$$

#### 8.4 Proof of Theorem 5.1

Since  $\widehat{\mathbf{G}} \rightarrow \mathbf{G}$  and  $\boldsymbol{\Sigma} = \mathbf{G}\boldsymbol{\Sigma}_X\mathbf{G}'$ , it suffices to show that, conditionally on  $(\boldsymbol{\epsilon}_t)$ ,

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{x}_{t,n}^* \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}_X). \quad (8.3)$$

Note that, conditionally on  $(\boldsymbol{\epsilon}_t)$ , for each  $n$  the random vectors  $\mathbf{x}_{1,n}^*, \mathbf{x}_{2,n}^*, \dots$  are independent and centered, with finite second-order moments. From the Lindeberg CLT for triangular arrays and the Wold-Cramèr device, to prove (8.3) it suffices to show that for any  $\boldsymbol{\lambda} \in \mathbb{R}^{s_1}$

$$\frac{1}{n} \sum_{t=1}^n \text{Var} \boldsymbol{\lambda}' \mathbf{x}_{t,n}^* \rightarrow \boldsymbol{\lambda}' \boldsymbol{\Sigma}_X \boldsymbol{\lambda} \quad \text{as } n \rightarrow \infty, \quad (8.4)$$

and for all  $\varepsilon > 0$

$$\frac{1}{n} \sum_{t=1}^n E \boldsymbol{\lambda}' \mathbf{x}_{t,n}^* \mathbf{x}_{t,n}^{*'} \boldsymbol{\lambda} 1_{\{|\boldsymbol{\lambda}' \mathbf{x}_{t,n}^*| \geq \sqrt{n}\varepsilon\}} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (8.5)$$

Recall that we are reasoning conditionally on  $(\boldsymbol{\epsilon}_t)$ . Therefore, the moments involved in (8.4) and (8.5) are computed with respect to the distribution of the bootstrap sample, and the convergences must hold for almost all sequence  $(\boldsymbol{\epsilon}_t)$  satisfying (2.1). Let us give the arguments for the proof of (8.4) when  $\boldsymbol{\lambda} = \mathbf{e}_1 := (1, 0, \dots, 0)'$  and  $m = 1$ . In this case,  $\widehat{\mathbf{C}}$  and  $\widehat{\underline{\mathbf{h}}}_t^2(\widehat{\boldsymbol{\vartheta}}_n)$  are scalar and, given  $(\boldsymbol{\epsilon}_t)$ ,

$$\text{Var} \boldsymbol{\lambda}' \mathbf{x}_{t,n}^* = \widehat{\mathbf{C}}^2 \widehat{\underline{\mathbf{h}}}_t^2(\widehat{\boldsymbol{\vartheta}}_n) (\hat{\mu}_4 - 1),$$

where  $\hat{\mu}_4$  is the empirical fourth-order moment of  $\widehat{\boldsymbol{\eta}}_1, \dots, \widehat{\boldsymbol{\eta}}_n$  (centered and reduced). Using the consistency of  $\widehat{\boldsymbol{\vartheta}}_n$ , the smoothness of  $\boldsymbol{\vartheta} \mapsto \underline{\mathbf{h}}_t(\boldsymbol{\vartheta})$  and the asymptotic irrelevance of the initial values, *i.e.*

$$\sup_{\boldsymbol{\vartheta} \in \Xi} \left| \widehat{\underline{\mathbf{h}}}_t^2(\boldsymbol{\vartheta}) - \underline{\mathbf{h}}_t^2(\boldsymbol{\vartheta}) \right| \leq K \rho^t$$

for any compact set  $\Xi$ , some  $K > 0$  and  $\rho \in (0, 1)$ , it can be shown (see, for example, the proof of Theorem 2 in Francq and Zakoïan (2013)) that, for almost all sequence  $(\boldsymbol{\epsilon}_t)$  satisfying (2.1),  $\hat{\mu}_4 \rightarrow E\eta_1^4$  as  $n \rightarrow \infty$ ,

$$\widehat{\mathbf{C}}^2 \widehat{\underline{\mathbf{h}}}_t^2(\widehat{\boldsymbol{\vartheta}}_n) (\hat{\mu}_4 - 1) \rightarrow \mathbf{C}^2 \widetilde{\underline{\mathbf{h}}}_t^2(\boldsymbol{\vartheta}_0) (E\eta_1^4 - 1) \text{ for } t \text{ fixed and } n \rightarrow \infty,$$

and

$$\frac{1}{n} \sum_{t=1}^n \widehat{\mathbf{C}}^2 \widehat{\underline{\mathbf{h}}}_t^2(\widehat{\boldsymbol{\vartheta}}_n) (\hat{\mu}_4 - 1) \rightarrow \mathbf{C}^2 E \underline{\mathbf{h}}_t^2(\boldsymbol{\vartheta}_0) (E\eta_1^4 - 1) = \boldsymbol{\Sigma}_X(1, 1)$$

as  $n \rightarrow \infty$ . More generally, by the previous arguments we have

$$\text{Var} \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbf{x}_{t,n}^* = \frac{1}{n} \sum_{t=1}^n E \mathbf{x}_{t,n}^* \mathbf{x}_{t,n}^{*'} \rightarrow \boldsymbol{\Sigma}_X \quad \text{as } n \rightarrow \infty,$$

which entails (8.4).

Now we turn to the proof of (8.5). For notational simplicity, we still give the arguments in the case  $\boldsymbol{\lambda} = \mathbf{e}_1$  and  $m = 1$ . Since there exists  $\underline{c} > 0$  such that  $\widehat{\mathbf{C}}^2 \widetilde{\mathbf{h}}_t^2(\widehat{\boldsymbol{\vartheta}}_n) > \underline{c}$  for all sequence  $(\epsilon_t)$ , we have

$$E\boldsymbol{\lambda}' \mathbf{x}_{t,n}^* \mathbf{x}_{t,n}^{*'} \boldsymbol{\lambda} 1_{\{|\boldsymbol{\lambda}' \mathbf{x}_{t,n}^*| \geq \sqrt{n}\epsilon\}} \leq \widehat{\mathbf{C}}^2 \widetilde{\mathbf{h}}_t^2(\widehat{\boldsymbol{\vartheta}}_n) E|\boldsymbol{\eta}_1^{*2} - 1|^2 1_{\{|\boldsymbol{\eta}_1^{*2} - 1| \geq \sqrt{n}\epsilon/\underline{c}\}}. \quad (8.6)$$

Because  $E|\boldsymbol{\eta}_1^{*2} - 1|^2 < \infty$  and the event  $\{|\boldsymbol{\eta}_1^{*2} - 1| \geq \sqrt{n}\epsilon/\underline{c}\}$  decreases to  $\emptyset$  as  $n \rightarrow \infty$ , the dominated convergence theorem shows that the right-hand side of (8.6) tends almost surely to zero. We then obtain (8.5) by Cesàro's lemma. The result can be extended for  $m > 1$  and for other values of  $\boldsymbol{\lambda}$ , which completes the proof.

## 8.5 Proof of Theorem 6.1

Similar to (8.2) we have

$$\frac{\partial}{\partial \mathbf{h}} \tilde{\ell}_t(\boldsymbol{\vartheta}_0) = \boldsymbol{\Omega}_t \mathbf{V}_t + o(\rho^t) \quad \text{a.s.},$$

where  $\boldsymbol{\Omega}_t$  is a  $m \times m^2$  matrix, measurable with respect to the sigma-field generated by  $\{\eta_u, u < t\}$ .

A coordinate wise Taylor expansion yields

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \mathbf{h}} \tilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \mathbf{h}} \tilde{\ell}_t(\boldsymbol{\vartheta}_0) + \mathbf{H}_n \sqrt{n} (\widehat{\mathbf{h}}_n - \mathbf{h}_0) + \mathbf{K}'_n \sqrt{n} (\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0),$$

where  $\mathbf{K}_n$  is defined in the proof of (3.3) Section 8.2, and

$$\mathbf{H}_n = \left( \frac{1}{n} \sum_{t=1}^n \frac{\partial^2}{\partial h_i \partial h_j} \tilde{\ell}_t(\bar{\boldsymbol{\vartheta}}_i) \right)_{m \times m}.$$

Write

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \mathbf{h}} \tilde{\ell}_t(\widehat{\boldsymbol{\vartheta}}_n) = \left( \mathbf{H}_n - \mathbf{K}'_n \mathbf{J}_n^{-1} \mathbf{K}_n \quad -\mathbf{K}'_n \mathbf{J}_n^{-1} \quad \mathbf{I}_m \right) \mathbf{Y}_n,$$

where

$$\mathbf{Y}_n := \left( \begin{array}{c} \mathbf{X}_n \\ \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial}{\partial \mathbf{h}} \tilde{\ell}_t(\boldsymbol{\vartheta}_0) \end{array} \right) = \left( \begin{array}{c} \frac{\underline{c}}{\sqrt{n}} \sum_{t=1}^n (\mathbf{U}_t^2 - \mathbf{I}_m) \mathbf{h}_t \\ \frac{1}{\sqrt{n}} \sum_{t=1}^n \left( \begin{array}{c} \boldsymbol{\Delta}_t \\ \boldsymbol{\Omega}_t \end{array} \right) \mathbf{V}_t \end{array} \right) + o_P(1).$$

By (3.2) and the CLT for square integrable stationary martingale differences, we have  $\mathbf{Y}_n \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}_Y)$ . We thus have  $\mathbf{T}_n \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}_R)$ .

## 8.6 Proof of Theorem 6.3

The vector  $\mathbf{Y}_n$  introduced in the proof of Theorem 6.1 has the form

$$\mathbf{Y}_n = \frac{1}{\sqrt{n}} \sum_{t=1}^n (1 - \eta_t^2) \begin{pmatrix} -\frac{1-\beta_0}{1-\alpha_0-\beta_0} h_t \\ \frac{1}{h_t} \sum_{i \geq 0} \beta^i (\epsilon_{t-i-1}^2 - v_0) \\ \frac{1}{h_t} \sum_{i \geq 0} \beta^i (h_{t-i-1} - v_0) \\ \frac{1-\alpha_0-\beta_0}{1-\beta_0} \frac{1}{h_t} \end{pmatrix} + o_P(1).$$

It follows that

$$\Sigma_Y = (\kappa_4 - 1) \begin{pmatrix} \left(\frac{1-\beta_0}{1-\alpha_0-\beta_0}\right)^2 E h_t^2 & \mathbf{f}' \\ \mathbf{f} & \mathcal{J} \end{pmatrix},$$

where  $\mathbf{f}' = (0, 0, -1)$  and  $\mathcal{J} = E \left( \frac{1}{h_t^2} \frac{\partial h_t(\boldsymbol{\vartheta}_0)}{\partial(\alpha, \beta, v)'} \frac{\partial h_t(\boldsymbol{\vartheta}_0)}{\partial(\alpha, \beta, v)} \right)$ . In particular,

$$\mathcal{J}_{33} = E \left( \frac{1}{h_t^2} \right) \left( \frac{1-\alpha_0-\beta_0}{1-\beta_0} \right)^2.$$

Now we evaluate  $\Sigma_R$ . We have  $\mathbf{H} = \mathcal{J}_{33}$  and  $\mathbf{K}' = (\mathcal{J}_{31}, \mathcal{J}_{32})$ . With the notations introduced in (6.4), we get

$$\Sigma_Y = (\kappa_4 - 1) \begin{pmatrix} \sigma^2 & \mathbf{0}_{1 \times 2} & -1 \\ \mathbf{0}_{2 \times 1} & \mathbf{J} & \mathbf{K} \\ -1 & \mathbf{K}' & \mathbf{H} \end{pmatrix}$$

resulting in  $\mathbf{M}\Sigma_Y = (\kappa_4 - 1)(\sigma^2\delta - 1, \mathbf{0}_{1,2}, 2\delta)$  and thus (6.4) follows.

Now we show that  $\Sigma_Y$  is positive-definite. Note that  $\Sigma_Y = (\kappa_4 - 1)E(\mathbf{S}_t \mathbf{S}_t')$ , where

$$\mathbf{S}_t = \begin{pmatrix} -e^{-1} h_t \\ h_t^{-1} \frac{\partial h_t}{\partial(\alpha, \beta)'}(\boldsymbol{\vartheta}_0) \\ e h_t^{-1} \end{pmatrix}, \quad e = \frac{1-\alpha_0-\beta_0}{1-\beta_0},$$

which yields that  $\Sigma_Y$  is semi positive-definite. Now if  $\mathbf{c}'\Sigma_Y \mathbf{c} = 0$  for some  $\mathbf{c} = (c_1, c_2, c_3, c_4)' \in \mathbb{R}^4$  we have

$$c_1 h_t^2 + c_2 \frac{\partial h_t}{\partial \alpha}(\boldsymbol{\vartheta}_0) + c_3 \frac{\partial h_t}{\partial \beta}(\boldsymbol{\vartheta}_0) + c_4 = 0 \quad a.s.$$

Thus

$$x_{t-2} \eta_{t-1}^4 + y_{t-2} \eta_{t-1}^2 + z_{t-2} = 0$$

for some variables  $x_{t-2}, y_{t-2}, z_{t-2}$  belonging to the  $\sigma$ -field generated by  $\{\eta_t - i, i \geq 2\}$ . In view of the independence between  $\eta_{t-1}^2$  and this  $\sigma$ -field, by solving the latter equation we find that  $\eta_{t-1}^2$  takes, almost surely, at most two constant values. Thus, we have shown that  $\Sigma_Y$  is positive-definite if the law of  $\eta_t^2$  is not concentrated on two points. It follows that  $\Sigma_R > 0$ , because  $\mathbf{M} \neq \mathbf{0}_{1 \times 4}$ .



## 9 Conclusion

In the framework of CCC-GARCH models, we studied a method, already widely used by practitioners, consisting in estimating the unconditional variances of the individual returns in a first step, and estimating the remaining parameters by QML in a second step. The main motivation for this method is computational and our experiments showed that the reduction of computational time compared to the full QML can be effective. Interestingly, the accuracy loss entailed by the two-step procedure is often barely visible on the simulation experiments. We also showed that the method is not limited to the standard CCC model. Models incorporating leverage effects can also be estimated in a similar way: instead of targeting the variances, moments related to the signs of the returns are more naturally targeted. A question of interest for future research is whether dynamic correlation models (such as the DCC model) can be handled this way.

## A Appendix: First-order derivatives of the criterion

To avoid using numerical derivatives in the practical estimation of CCC-GARCH models, it is important to be able to compute explicitly (in a recursive way) the derivatives of the criterion. For the CCC-GARCH(1,1) model we have

$$\tilde{\mathbf{h}}_t = (\mathbf{I}_m - \mathbf{A} - \mathbf{B})\mathbf{h} + \mathbf{A}\boldsymbol{\epsilon}_{t-1} + \mathbf{B}\tilde{\mathbf{h}}_{t-1}.$$

Write  $\boldsymbol{\alpha} = \text{vec}(\mathbf{A}), \boldsymbol{\beta} = \text{vec}(\mathbf{B})$ . The derivatives of  $\tilde{\mathbf{h}}_t$  with respect to the parameters can be computed recursively as follows, for  $t \geq 1$

$$\begin{aligned} \frac{\partial \tilde{\mathbf{h}}_t}{\partial \mathbf{h}'} &= \mathbf{I}_m - \mathbf{A} - \mathbf{B} + \mathbf{B} \frac{\partial \tilde{\mathbf{h}}_{t-1}}{\partial \mathbf{h}'}, \\ \frac{\partial \tilde{\mathbf{h}}_t}{\partial \boldsymbol{\alpha}'} &= (\boldsymbol{\epsilon}_{t-1} - \mathbf{h})' \otimes \mathbf{I}_m + \mathbf{B} \frac{\partial \tilde{\mathbf{h}}_{t-1}}{\partial \boldsymbol{\alpha}'}, \\ \frac{\partial \tilde{\mathbf{h}}_t}{\partial \boldsymbol{\beta}'} &= (\tilde{\mathbf{h}}_{t-1} - \mathbf{h})' \otimes \mathbf{I}_m + \mathbf{B} \frac{\partial \tilde{\mathbf{h}}_{t-1}}{\partial \boldsymbol{\beta}'}, \end{aligned}$$

or equivalently, letting  $\boldsymbol{\phi}' = (\mathbf{h}', \boldsymbol{\alpha}', \boldsymbol{\beta}')$

$$\frac{\partial \tilde{\mathbf{h}}_t}{\partial \boldsymbol{\phi}'} = \left( \mathbf{I}_m - \mathbf{A} - \mathbf{B} \begin{pmatrix} \boldsymbol{\epsilon}_{t-1} - \mathbf{h} \\ \tilde{\mathbf{h}}_{t-1} - \mathbf{h} \end{pmatrix}' \otimes \mathbf{I}_m \right) + \mathbf{B} \frac{\partial \tilde{\mathbf{h}}_{t-1}}{\partial \boldsymbol{\phi}'},$$

with, for example,  $\frac{\partial \tilde{\mathbf{h}}_0}{\partial \boldsymbol{\phi}} = 0$ . For the CCC-GARCH( $p, q$ ) model we similarly have

$$\frac{\partial \tilde{\mathbf{h}}_t}{\partial \boldsymbol{\phi}'} = \left( \mathbf{I}_m - \sum_{i=1}^r (\mathbf{A}_i + \mathbf{B}_i) \begin{pmatrix} \boldsymbol{\epsilon}_{t-1} - \mathbf{h} \\ \vdots \\ \boldsymbol{\epsilon}_{t-q} - \mathbf{h} \\ \tilde{\mathbf{h}}_{t-1} - \mathbf{h} \\ \vdots \\ \tilde{\mathbf{h}}_{t-p} - \mathbf{h} \end{pmatrix}' \otimes \mathbf{I}_m + \sum_{j=1}^p \mathbf{B}_j \frac{\partial \tilde{\mathbf{h}}_{t-j}}{\partial \boldsymbol{\phi}'}, \right)$$

where  $\boldsymbol{\phi}' = (\mathbf{h}', \boldsymbol{\alpha}'_1, \dots, \boldsymbol{\alpha}'_q, \boldsymbol{\beta}'_1, \dots, \boldsymbol{\beta}'_p)$ .

Now let  $\tilde{\mathbf{D}}_t^2 = \text{diag}(\tilde{\mathbf{h}}_t)$  and let  $\mathbf{R}^{j\cdot}$  denote the  $j$ th row of  $\mathbf{R}^{-1}$ . We have (see for example (11.67)-(11.70) in Francq and Zakoian (2010))

$$\begin{aligned} \frac{\partial \tilde{\ell}_t(\boldsymbol{\vartheta})}{\partial \vartheta_i} &= -\text{Tr} \left\{ \left( \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t' \tilde{\mathbf{D}}_t^{-1} \mathbf{R}^{-1} + \mathbf{R}^{-1} \tilde{\mathbf{D}}_t^{-1} \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t' \right) \tilde{\mathbf{D}}_t^{-1} \frac{\partial \tilde{\mathbf{D}}_t}{\partial \vartheta_i} \tilde{\mathbf{D}}_t^{-1} \right\} \\ &\quad + 2\text{Tr} \left( \tilde{\mathbf{D}}_t^{-1} \frac{\partial \tilde{\mathbf{D}}_t}{\partial \vartheta_i} \right) \\ &= 2\text{Tr} \left\{ \left( \mathbf{I}_m - \tilde{\mathbf{D}}_t^{-1} \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t' \tilde{\mathbf{D}}_t^{-1} \mathbf{R}^{-1} \right) \tilde{\mathbf{D}}_t^{-1} \frac{\partial \tilde{\mathbf{D}}_t}{\partial \vartheta_i} \right\} \\ &= -\sum_{j=1}^m \left\{ 1 - \left( \mathbf{R}^{j\cdot} \tilde{\mathbf{D}}_t^{-1} \boldsymbol{\epsilon}_t \right) \frac{\epsilon_{jt}}{\{\tilde{h}_{jt}\}^{1/2}} \right\} \frac{1}{\tilde{h}_{jt}} \frac{\partial \tilde{h}_{jt}}{\partial \vartheta_i} \\ &:= \boldsymbol{\omega}'_t \frac{\partial \tilde{\mathbf{h}}_t}{\partial \vartheta_i}, \end{aligned}$$

for  $i = 1, \dots, s_2 = m + (p + q)m^2$ . It follows that

$$\frac{\partial \tilde{\ell}_t(\boldsymbol{\vartheta})}{\partial \boldsymbol{\phi}} = (\mathbf{I}_{s_2} \otimes \boldsymbol{\omega}'_t) \text{vec} \left( \frac{\partial \tilde{\mathbf{h}}_t}{\partial \boldsymbol{\phi}'} \right).$$

Similarly

$$\begin{aligned} \frac{\partial \tilde{\ell}_t(\boldsymbol{\vartheta})}{\partial \vartheta_i} &= -\text{Tr} \left( \mathbf{R}^{-1} \tilde{\mathbf{D}}_t^{-1} \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t' \tilde{\mathbf{D}}_t^{-1} \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \vartheta_i} \right) + \text{Tr} \left( \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \vartheta_i} \right) \\ &= \text{Tr} \left\{ \left( \mathbf{I}_m - \mathbf{R}^{-1} \tilde{\mathbf{D}}_t^{-1} \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t' \tilde{\mathbf{D}}_t^{-1} \right) \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \vartheta_i} \right\}, \end{aligned}$$

for  $i = s_2 + 1, \dots, s_1$ .

## References

**Aue, A., Hörmann, S., Horváth, L. and M. Reimherr** (2009) Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics* 37, 4046–4087.

- Bauwens, L., Laurent, S. and J.V.K. Rombouts** (2006) Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21, 79–109.
- Billingsley, P.** (1961) The Lindeberg-Levy theorem for martingales. *Proceedings of the American Mathematical Society* 12, 788–792.
- Bollerslev, T.** (1990) Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Review of Economics and Statistics* 72, 498–505.
- Boudt, K., Danielsson J. and S. Laurent** (2013) Robust forecasting of dynamic conditional correlation GARCH models. *International Journal of Forecasting* 29, 244–257.
- Christoffersen, P.F.** (2003) *Elements of financial risk management*. Academic Press, London.
- Engle, R.F. and K. Kroner** (1995) Multivariate simultaneous generalized ARCH. *Econometric Theory* 11, 122–150.
- Engle, R.F. and J. Mezrich** (1996) GARCH for groups. *Risk* 9, 36–40.
- Franco, C., Horváth, L. and J-M. Zakoïan** (2011) Merits and drawbacks of variance targeting in GARCH Models. *Journal of Financial Econometrics* 9, 619–656.
- Franco C. and J-M. Zakoïan** (2007) Quasi-maximum likelihood estimation in GARCH processes when some coefficients are equal to zero. *Stochastic Processes and their Applications* 117, 1265–1284.
- Franco, C. and J-M. Zakoïan** (2009) Testing the nullity of GARCH coefficients : correction of the standard tests and relative efficiency comparisons. *Journal of the American Statistical Association* 104, 313–324.
- Franco, C. and J-M. Zakoïan** (2010) *GARCH Models: Structure, Statistical Inference and Financial Applications*. Chichester: John Wiley.
- Franco, C. and J-M. Zakoïan** (2012) QML estimation of a class of multivariate asymmetric GARCH models. *Econometric Theory* 28, 179–206.
- Franco, C. and J-M. Zakoïan** (2013) Optimal predictions of powers of conditionally heteroscedastic processes. *J.R. Statist. Soc. B* 75, 345–367.
- Giraitis, L., Leipus, R. and D. Surgailis** Recent advances in ARCH modelling. *In: Teyssière, G. and Kirman, A. (Eds.): Long Memory in Economics*, 3–38. Springer, New-York.
- He, C. and T. Teräsvirta** (2004) An extended constant conditional correlation GARCH model and its fourth-moment structure. *Econometric Theory* 20, 904–926.
- Hill, J. and E. Renault** (2012) Variance Targeting for Heavy Tailed Time Series. Unpublished document.
- Hull, J.** 2003. *Options, Futures, and Other Derivatives*. Princeton Hall, New Jersey, USA.

- Jeantheau, T.** (1998) Strong consistency of estimators for multivariate ARCH models. *Econometric Theory* 14, 70–86.
- Kreiss, J.P., Paparoditis, E. and D.N. Politis** (2011) On the range of validity of the autoregressive sieve bootstrap. *The Annals of Statistics* 39, 2103–2130.
- Kristensen, D. and O. Linton** (2004) Consistent standard errors for target variance approach to GARCH estimation. *Econometric Theory* 20, 990–993.
- Laurent, S., Rombouts, J.V.K. and F. Violante** (2012) On the forecasting accuracy of multivariate GARCH models. *Journal of Applied Econometrics* 27, 934–955.
- Ling, S. and M. McAleer** (2002) Necessary and sufficient moment conditions for the GARCH( $r, s$ ) and asymmetric GARCH( $r, s$ ) models. *Econometric Theory* 18, 722–729.
- Pedersen, R.S. and A. Rahbek** (2013) Multivariate variance targeting in the BEKK-GARCH model. Forthcoming in *The Econometrics Journal*.
- Silvennoinen, A. and T. Teräsvirta** (2009) Multivariate GARCH models. *Handbook of Financial Time Series* T.G. Andersen, R.A. Davis, J-P. Kreiss and T. Mikosch, eds. New York: Springer.
- Shimizu, K.** (2013) The bootstrap does not always work for heteroscedastic models. *Statistics & Risk Modeling* 30, 189–204.
- Vaynman, I. and B. K. Beare** (2013) Stable limit theory for the variance targeting estimator. Unpublished document, University of California, San Diego.
- Zhu, D. and V. Zinde-Walsh** (2009) Properties and estimation of asymmetric exponential power distribution. *Journal of Econometrics* 148, 86–99.