

ON THE STATIONARITY OF DYNAMIC CONDITIONAL CORRELATION MODELS

JEAN-DAVID FERMANIAN
Crest-Ensaë

HASSAN MALONGO
Amundi & Univ. Paris Dauphine

We provide conditions for the existence and the uniqueness of strictly stationary solutions of the usual Dynamic Conditional Correlation GARCH models (DCC-GARCH). The proof is based on Tweedie's (1988) criteria, after having rewritten DCC-GARCH models as nonlinear Markov chains. We also study the existence of their moments and discuss the tightness of our sufficient conditions.

1. INTRODUCTION

1.1. The problem

In multivariate extensions of GARCH models, modelers are faced with the problem of correlations (between asset returns, in most applications). The simplest idea is to assume that these correlations are constant in time, and constitute only an additional matrix of parameters. This has provided the class of Constant Conditional Correlations models (CCC), first introduced by Bollerslev (1990). Since CCC models can be seen as the components of first-order Markov processes, once such models are rewritten in an extended vector space, it is relatively easy to prove the existence of strictly stationary and explicit solutions, even if the latter are analytically complex: see classical textbooks, for instance, Francq and Zakoïan (2010).

It rapidly became apparent that the assumption of constant correlations is too strong. It does not correspond to economic intuition or many empirical features: see the recent paper by Otranto and Bauwens (2013) and the numerous references therein, for instance. Therefore, Engle (2002) proposed extending CCC specifications by adding particular dynamics on the (conditional) correlation matrices of returns, denoted here by (R_t) . To ensure modelers are dealing with true correlation matrices, he introduced a nonlinear transform: there exists a sequence of variance-covariance matrices (Q_t) such that $R_t = \text{diag}(Q_t)^{-1/2} Q_t \text{diag}(Q_t)^{-1/2}$, and

We thank C. Francq and J.-M. Zakoïan for their valuable remarks and discussions. Moreover, we are grateful to Eric Renault and two anonymous referees, who have proposed a number of ways of improving the article. Finally, the authors thank the Labex "Ecodec" for its support. Address correspondence to Jean-David Fermanian, 3 avenue Pierre Larousse, 92245 Malakoff cedex, France, e-mail: jean-david.fermanian@ensae.fr.

(Q_t) -dynamics are specified instead of (R_t) -dynamics directly, contrary to other authors (Tse and Tsui, 2002 or Pelletier, 2006, for instance). This nonlinear transform ensures that R_t is always a correlation matrix, i.e., positive semidefinite with ones on its main diagonal. Nonetheless, it considerably complicates the work of stating DCC model stationarity conditions. Indeed, analytically tractable solutions to such processes no longer exist. This explains why the existence and uniqueness of DCC model stationarity solutions have not yet been established in the literature, nor have the finiteness of their moments. Particularly, this implies that theoretically sound statistical inference procedures do not yet exist, as noted in Caporin and McAleer (2013).

Despite their theoretical shortcomings, DCC models have been used intensively among academics and practitioners. Besides numerous applied works, several extensions of the baseline DCC representation have been proposed in the literature: inclusion of asymmetries (Cappiello, Engle, and Sheppard, 2006), volatility thresholds (Kasch and Caporin, 2013), macro-variables (Otranto and Bauwens, 2013), univariate switching regime probabilities (Pelletier, 2006, Billio and Caporin, 2005, Fermanian and Malongo, 2013), among others. Other authors have revisited the DCC parameterization itself: Billio, Caporin, and Gobbo (2006), Franses and Hafner (2009), etc. Therefore, there is an urgent need for new theoretical results concerning the seminal DCC model itself.

Usually in econometrics, proving the existence of stationary solutions is the first step towards developing a full asymptotic theory (consistency/asymptotic normality of QML estimates typically, as in Comte and Lieberman (2003) in the case of multivariate GARCH models), because laws of large numbers and some CLTs are easily obtained in this case. In the GARCH literature, this essential task has been fulfilled by Bougerol and Picard (1992) for univariate GARCH models, by Ling and McAleer (2003) for multivariate ARMA-GARCH models, and by Boussama et al. (2011) for BEKK models. In the case of DCC models, a key-stone is missing: a theory for inference has been proposed by Engle and Sheppard (2001), but their two stage estimation procedure is contingent on the underlying DCC process being strictly stationary and ergodic (see their Assumption A.2). The goal of this paper is to fill this gap.

After introducing some notations, we define DCC models at the beginning of Section 2. They will be rewritten as “almost linear” Markov chains in Subsection 2.2. The existence of strong and weak stationary solutions is stated in Subsection 3.1. Subsection 3.2 exhibits sufficient conditions to get their uniqueness. We discuss the tightness of our technical conditions from a qualitative standpoint in Section 4. Proof of the propositions and theorems are detailed in the appendices.

1.2. Notations

Consider an (n, m) matrix $M = [m_{ij}]_{1 \leq i \leq n, 1 \leq j \leq m}$.

- $M \geq 0$ (resp. $M > 0$) means that all elements of M are nonnegative (resp. strictly positive), and $|M| = [|m_{ij}|]_{1 \leq i \leq n, 1 \leq j \leq m}$.

- If $n = m$, let the diagonal matrix $diag(M) = [m_{ij}\mathbf{1}(i = j)]_{1 \leq i \leq m, 1 \leq j \leq m}$ and the vector $Vecd(M) = [m_{ii}]_{1 \leq i \leq m}$ in \mathbb{R}^m .
- If $n = m$ and M is symmetric, $Vech(M)$ denotes the $m(m + 1)/2 =: m^*$ column vector whose components are read from M column-wise and without redundancy. To be formal, $Vech(M) = [\tilde{m}_k]_{1 \leq k \leq m^*}$, where $\tilde{m}_k = m_{ij}$ for the unique couple of indices (i, j) in $\{1, \dots, m\}^2$, $i \geq j$ such that $[m + (m - 1) + \dots + (m - j + 2)]^+ + (i - j + 1) = k$. This defines a one-to-one mapping ϕ between the indices $k \in \{1, \dots, m^*\}$ and the pairs (i, j) , $i \geq j$, $1 \leq i, j \leq m$, i.e., $(i, j) = (\phi_1(k), \phi_2(k)) =: \phi(k)$.
- \otimes denotes the usual Kronecker product, and $M^{\otimes p} = M \otimes \dots \otimes M$ (p times). \odot denotes the element-by-element product. If v is a vector in \mathbb{R}^n , then $v \odot M = [v_i m_{ij}]_{1 \leq i \leq n, 1 \leq j \leq m}$.
- We will consider several matrix norms, particularly

$$\|M\|_{\max} = \max_{1 \leq i \leq n, 1 \leq j \leq m} |m_{ij}|,$$

and the spectral norm, defined for any squared matrix by

$$\|M\|_s = \sup \left\{ \sqrt{\lambda} \mid \lambda \text{ is an eigenvalue of } M'M \right\} = \sup_{\mathbf{x}} \frac{\|M\mathbf{x}\|_2}{\|\mathbf{x}\|_2}.$$

Besides, we will consider any norm \mathcal{N} for vectors, that is not the Euclidian norm $\|\cdot\|_2$. Then, we can define the norm $\|\cdot\|_{\mathcal{N}}$ for matrices by setting $\|M\|_{\mathcal{N}} = \sup_{\mathbf{x}} \mathcal{N}(M\mathbf{x})/\mathcal{N}(\mathbf{x})$. Note that $\|M\|_{\infty} = \max_i \sum_j |m_{ij}|$ when $\mathcal{N}(\mathbf{x}) = \|\mathbf{x}\|_{\infty} = \max_i |x_i|$.

- $\rho(M)$ denotes the spectral radius of the squared matrix M , i.e., the largest of the modulus of M 's eigenvalues. If M is positive semidefinite, then $\rho(M) = \|M\|_s$ and its smallest eigenvalue is denoted by $\lambda_1(M)$.
- For any column vector $z_t \in \mathbb{R}^m$, we denote $z_t = (z_{1,t}, \dots, z_{m,t})'$ and $\vec{z}_t := (z_{1,t}^2, \dots, z_{m,t}^2)'$.
- e denotes a vector of ones, the dimension of which will be implicit. 0_m (resp. I_m) denotes the $m \times m$ matrix of zeros (resp. identity matrix). When the dimension of an identity matrix is not specified, it will be denoted by Id .
- If M depends on $\mathbf{x} \in A$, then $\sup_{\mathbf{x} \in A} M(\mathbf{x})$ is the matrix $[\sup_{\mathbf{x} \in A} m_{ij}(\mathbf{x})]$.

2. DYNAMIC CONDITIONAL CORRELATION MODELS

2.1. The classical DCC specification

Let us reiterate here the standard DCC model, as introduced in Engle (2002). Consider a stochastic process $(y_t)_{t \in \mathbb{Z}}$ in \mathbb{R}^m , typically a vector of m asset returns. The sigma field generated by the past information of this process until up to and including time $t - 1$ is denoted by \mathcal{I}_{t-1} .

Modeling the expected returns of financial series is a problem per se, that has generated a huge amount of literature. In this paper, our focus will be on the dynamics of the conditional variance-covariance of y_t instead. Therefore, following current practice, we will assume we can remove the conditional means of our returns. Let $\mu_t(\theta) = E[y_t | \mathcal{I}_{t-1}] =: E_{t-1}[y_t]$ be the conditional mean vector of y_t . It depends on a vector of parameters $\theta \in \Theta$. We define a “detrended” series $(z_t)_{t \in \mathbb{Z}}$ by

$$y_t = \mu_t(\theta) + z_t, \quad E_{t-1}[z_t] = 0.$$

For convenience, the conditional mean $\mu_t(\theta)$ is assumed to be measurable with respect to $\sigma(z_{t-1}, z_{t-2}, \dots)$. Therefore, $\mathcal{I}_t = \sigma(y_t, y_{t-1}, \dots) = \sigma(z_t, z_{t-1}, \dots)$.

Let us denote by H_t the variance-covariance matrix of the t -observations, conditionally on \mathcal{I}_{t-1} : $Var(y_t | \mathcal{I}_{t-1}) = Var(z_t | \mathcal{I}_{t-1}) := H_t$. As usual with DCC-type models, we split the variance-covariance matrix H_t between volatility terms on one side (in D_t), and correlation coefficients on the other side (in R_t):

$$H_t = D_t^{1/2} R_t D_t^{1/2}, \quad D_t = \text{diag}(h_{1,t}, \dots, h_{m,t}), \tag{1}$$

where $h_{k,t}$ denotes the “instantaneous variance” of the return $y_{k,t}$ (or $z_{k,t}$, equivalently), conditionally on \mathcal{I}_{t-1} . We assume GARCH-type models on every margin, but with potential cross-effects between all these volatilities:

$$Vecd(D_t) = V_0 + \sum_{i=1}^r A_i \cdot Vecd(D_{t-i}) + \sum_{j=1}^s B_j \cdot \vec{z}_{t-j}, \tag{2}$$

for some deterministic nonnegative matrices $(A_i)_{i=1, \dots, r}$ and $(B_j)_{j=1, \dots, s}$, and for a positive vector V_0 in \mathbb{R}^m . We will set $A_i := [a_{k,l}^{(i)}]_{1 \leq k, l \leq m}$, $i = 1, \dots, r$, and $B_j := [b_{k,l}^{(j)}]_{1 \leq k, l \leq m}$, $j = 1, \dots, s$.

Let us introduce the vector of so-called “standardized residuals” $\varepsilon_t := D_t^{-1/2} z_t$. Obviously, $E_{t-1}[\varepsilon_t] = 0$ and $E_{t-1}[\varepsilon_t \varepsilon_t'] = R_t$. We impose that $M^{1/2}$ is positive definite for any positive definite matrix M . In this case, the square root of R_t is uniquely defined: see Serre (2010), Theorem 6.1. This will be our convention throughout the article.

The dynamics of correlations are given by the traditional Dynamic Conditional Correlation specification:

$$R_t = \text{diag}(Q_t)^{-\frac{1}{2}} Q_t \text{diag}(Q_t)^{-\frac{1}{2}}, \tag{3}$$

where the sequence of matrices $(Q_t)_{t \in \mathbb{Z}}$ satisfies

$$Q_t = W_0 + \sum_{k=1}^{\nu} M_k Q_{t-k} M_k' + \sum_{l=1}^{\mu} N_l \varepsilon_{t-l} \varepsilon_{t-l}' N_l', \tag{4}$$

for some deterministic matrices $(M_k)_{k=1, \dots, \nu}$ and $(N_l)_{l=1, \dots, \mu}$, and for a positive definite constant matrix W_0 . Obviously, when such a sequence $(Q_t)_{t \geq -\nu}$ is

initialized with ν nonnegative definite (possibly null) matrices, every $Q_t, t \geq 0$, will be definite positive. In Theorem 1, we prove that a “doubly infinite” stationary sequence $(Q_t)_{t \in \mathbb{Z}}$ of definite positive matrices exists and which satisfies (4).

We will set $M_k := [m_{p,q}^{(k)}]_{1 \leq p,q \leq m}, k = 1, \dots, \nu$, and $N_l := [n_{p,q}^{(l)}]_{1 \leq p,q \leq m}, l = 1, \dots, \mu$. In practice, the positive matrix W_0 (or the constant vector $Vech(W_0)$ in \mathbb{R}^{m^*} equivalently) is a parameter that has to be estimated, most often during the first stage.

Aielli (2013) noticed that the estimation of the unknown matrix W_0 is not straightforward, because it cannot be deduced trivially from the unconditional correlation between the standardized residuals ε_t . Therefore, he introduced a new variety of DCC-GARCH models (called cDCC), where (4) is replaced by

$$Q_t = W_0 + \sum_{k=1}^{\nu} M_k Q_{t-k} M_k' + \sum_{l=1}^{\mu} N_l \text{diag}(Q_{t-l})^{-1/2} \varepsilon_{t-l} \varepsilon_{t-l}' \text{diag}(Q_{t-l})^{-1/2} N_l'. \tag{5}$$

Under this new assumption, cDCC can be seen as a particular BEKK model (Engle and Kroner, 1995). Therefore, Aielli obtained the existence of strictly and/or weakly stationary solutions, applying the conditions of Boussama, Fuchs, and Stelzer (2011) on BEKK processes. Actually, Aielli’s model (5) is a smart but not intuitive “ad hoc” specification. Its main justification appears as essentially technical, to avoid the nonlinear feature of Engle’s original DCC model (4). Under the standard latter specification, DCC models can no longer be rewritten as BEKK models and other techniques have to be found. In this paper, we obtain similar results to Aielli (2013), but by keeping the original specification of DCC models and without relying on another surrounding family of processes.

2.2. DCC as Markov chains

Actually, it is possible to rewrite the previous DCC model as a Markov chain, that looks like an AR(1) process. This rewrite will become a crucial tool when studying stationary solutions hereafter. Set

$$X_t := (X_t^{(1)}, X_t^{(2)}, X_t^{(3)}, X_t^{(4)})', \tag{6}$$

where

$$\begin{aligned} X_t^{(1)} &:= (\text{Vecd}(D_t), \dots, \text{Vecd}(D_{t-r+1}))', \\ X_t^{(2)} &:= (\vec{z}_t, \dots, \vec{z}_{t-s+1})', \\ X_t^{(3)} &:= (\text{Vech}(Q_t), \dots, \text{Vech}(Q_{t-\nu+1}))', \text{ and} \\ X_t^{(4)} &:= (\text{Vech}(\varepsilon_t \varepsilon_t'), \dots, \text{Vech}(\varepsilon_{t-\mu+1} \varepsilon_{t-\mu+1}'))'. \end{aligned}$$

The dimensions of the four previous random vectors are $rm, sm, \nu m^*$, and μm^* , respectively. Their sum, the dimension of X_t , is denoted by d . With simple block

matrix calculations, random matrices (T_t) and a vector process (ζ_t) exist, such that the dynamics of X_t , any solution of the DCC model, may be rewritten as

$$X_t = T_t \cdot X_{t-1} + \zeta_t, \tag{7}$$

for any t . We will write the block matrix $T_t := [T_{ij,t}]_{1 \leq i, j \leq 4}$ with convenient random matrices $T_{ij,t}$.

Knowing (7), the underlying process (X_t) can be seen as a vectorial autoregressive of order one, but with random matrix coefficients (T_t). Let us detail the AR(1) form of (7):

- set $T_{1k,t} = 0$ when $k = 3, 4$,

$$T_{11,t} := \begin{bmatrix} A_1 & A_2 & \cdots & \cdots & A_r \\ I_m & 0_m & \cdots & \cdots & 0_m \\ 0_m & I_m & 0_m & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_m & \cdots & 0_m & I_m & 0_m \end{bmatrix}, \text{ and } T_{12,t} := \begin{bmatrix} B_1 & B_2 & \cdots & \cdots & B_s \\ 0_m & \cdots & \cdots & \cdots & 0_m \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 0_m & \cdots & \cdots & \cdots & 0_m \end{bmatrix}.$$

- We deduce from Equation (2) that

$$D_t \vec{\varepsilon}_t = \vec{\varepsilon}_t \odot \text{Vecd}(D_t) = \vec{z}_t = \vec{\varepsilon}_t \odot V_0 + \sum_{i=1}^r \vec{\varepsilon}_t \odot A_i \cdot \text{Vecd}(D_{t-i}) + \sum_{j=1}^s \vec{\varepsilon}_t \odot B_j \cdot \vec{z}_{t-j}. \tag{8}$$

Let us set $T_{23,t} = T_{24,t} = 0$,

$$T_{21,t} := \begin{bmatrix} \vec{\varepsilon}_t \odot A_1 & \vec{\varepsilon}_t \odot A_2 & \cdots & \cdots & \vec{\varepsilon}_t \odot A_r \\ 0_m & \cdots & \cdots & \cdots & 0_m \\ \vdots & & & & \vdots \\ 0_m & \cdots & \cdots & \cdots & 0_m \end{bmatrix}, \text{ and}$$

$$T_{22,t} := \begin{bmatrix} \vec{\varepsilon}_t \odot B_1 & \vec{\varepsilon}_t \odot B_2 & \cdots & \cdots & \vec{\varepsilon}_t \odot B_s \\ I_m & 0_m & \cdots & \cdots & 0_m \\ 0_m & I_m & 0_m & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_m & \cdots & 0_m & I_m & 0_m \end{bmatrix}.$$

- Clearly, matrices $\tilde{M}_k, k = 1, \dots, \nu$, exist such that

$$\text{Vec}(M_k Q_{t-k} M_k') = \tilde{M}_k \cdot \text{Vec}(Q_{t-k}).$$

Similarly, matrices $\tilde{N}_l, l = 1, \dots, \mu$, exist such that

$$\text{Vec}(N_l \varepsilon_{t-l} \varepsilon_{t-l}' N_l') = \tilde{N}_l \cdot \text{Vec}(\varepsilon_{t-l} \cdot \varepsilon_{t-l}').$$

It is possible to explicitly write the previous matrices \tilde{M}_k and \tilde{N}_l 's. Indeed, with the notations of Subsection 1.2, $\tilde{M}_k = [\tilde{m}_{u,v}^{(k)}]_{1 \leq u,v \leq m^*}$ where

$$\tilde{m}_{u,v}^{(k)} = m_{\phi_1(u), \phi_1(v)}^{(k)} m_{\phi_2(u), \phi_2(v)}^{(k)}.$$

Similarly, $\tilde{N}_l = [\tilde{n}_{u,v}^{(l)}]_{1 \leq u,v \leq m^*}$ and $\tilde{n}_{u,v}^{(l)} = n_{\phi_1(u), \phi_1(v)}^{(l)} m_{\phi_2(u), \phi_2(v)}^{(l)}$. Then, set $T_{31,t} = T_{32,t} = 0$,

$$T_{33,t} := \begin{bmatrix} \tilde{M}_1 & \tilde{M}_2 & \cdots & \cdots & \tilde{M}_\nu \\ I_{m^*} & 0_{m^*} & \cdots & \cdots & 0_{m^*} \\ 0_{m^*} & I_{m^*} & 0_{m^*} & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_{m^*} & \cdots & 0_{m^*} & I_{m^*} & 0_{m^*} \end{bmatrix}, \text{ and } T_{34,t} := \begin{bmatrix} \tilde{N}_1 & \tilde{N}_2 & \cdots & \cdots & \tilde{N}_\mu \\ 0_{m^*} & \cdots & \cdots & \cdots & 0_{m^*} \\ \vdots & & & & \vdots \\ 0_{m^*} & \cdots & \cdots & \cdots & 0_{m^*} \end{bmatrix}.$$

- $T_{4k,t} = 0, k = 1, 2, 3$, and define the $\mu m^* \times \mu m^*$ matrix

$$T_{44,t} := \begin{bmatrix} 0_{m^*} & 0_{m^*} & \cdots & \cdots & 0_{m^*} \\ I_{m^*} & 0_{m^*} & \cdots & \cdots & 0_{m^*} \\ 0_{m^*} & I_{m^*} & 0_{m^*} & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_{m^*} & \cdots & 0_{m^*} & I_{m^*} & 0_{m^*} \end{bmatrix}.$$

Moreover, rewrite $\zeta_t = (\zeta_t^{(1)}, \zeta_t^{(2)}, \zeta_t^{(3)}, \zeta_t^{(4)})$, where, with obvious sizes, these vectors are

$$\begin{aligned} \zeta_t^{(1)} &= (V_0, 0_{m^*}, \dots, 0_{m^*})', \quad \zeta_t^{(2)} = (\tilde{\varepsilon}_t \odot V_0, 0_{m^*}, \dots, 0_{m^*})', \\ \zeta_t^{(3)} &= (Vech(W_0), 0_{m^*}, \dots, 0_{m^*})', \text{ and } \zeta_t^{(4)} = (Vech(\varepsilon_t \varepsilon_t'), 0_{m^*}, \dots, 0_{m^*})'. \end{aligned}$$

Since D_t, Q_t , and R_t are \mathcal{I}_{t-1} -measurable, it is easy to see that the filtration induced by the observations is the natural filtration of (X_t) : $\sigma(X_t, X_{t-1}, \dots) = \mathcal{I}_t$ for all t . From now on, we consider (\mathcal{I}_t) as the filtration that is generated by (X_t) . This means that $E_{t-1}[Z] = E[Z|\mathcal{I}_{t-1}] = E[Z|X_{t-1}, X_{t-2}, \dots]$, for any random vector Z . And a process (Z_t) is said to be (one-order, implicitly) \mathcal{I} -Markov if the law of Z_t given \mathcal{I}_{t-1} is the law of Z_t given X_{t-1} .

Intuitively, the sequence (X_t) is \mathcal{I} -Markov because it is the case for the processes (ζ_t) and (T_t) themselves. To prove this formally, we need an assumption concerning the data generating process (DGP) of (z_t) .

Let us define the t -vector of innovations by

$$\eta_t := R_t^{-1/2} \varepsilon_t = R_t^{-1/2} D_t^{-1/2} z_t. \tag{9}$$

Note that $E_{t-1}[\eta_t] = 0$ and $E_{t-1}[\eta_t \eta_t'] = I_m$ by construction. The definition of these innovations implies that, for every $t, \sigma(\eta_j, j \leq t) \subset \sigma(\varepsilon_j, j \leq t) \subset \mathcal{I}_t$. Nonetheless, we will not establish whether there are equalities between the latter

filtrations. Technically speaking, this would be equivalent to stating the invertibility of the underlying process.

Assumption A0. $(\eta_t)_{t \in \mathbb{Z}}$ possesses the Markov property with respect to the filtration \mathcal{I} . In particular, $E[\eta_t | \mathcal{I}_{t-1}] = E[\eta_t | X_{t-1}]$ for every t .

Obviously, the latter assumption is satisfied if $(\eta_t)_{t \in \mathbb{Z}}$ is a sequence of identically distributed and mutually independent random vectors, with $E[\eta_t] = 0$ and $E[\eta_t \eta_t'] = I_m$. For instance, if the random vectors η_t are standardized Gaussian and mutually independent, the process $(z_t)_{t \in \mathbb{Z}}$ is conditionally Gaussian, a standard case in practice.

PROPOSITION 1. *Under A0, the process (X_t) is Markov of order one with respect to its natural filtration.*

See the proof in the appendix.

Remark 1. It would be possible to define a slightly different DGP for the DCC model above: consider a \mathcal{I} -martingale difference (i.i.d., for instance) sequence $(e_t)_{t \in \mathbb{Z}}$ and set $z_t = H_t^{1/2} e_t$, $E_{t-1}[e_t] = 0$, $E_{t-1}[e_t e_t'] = I_m$. In the latter case, the process (η_t) above would be defined by $\eta_t = R_t^{-1/2} D_t^{-1/2} H_t^{1/2} e_t$. Then, it is easy to check that (η_t) is \mathcal{I} -Markov and is a martingale difference. In other words, A0 would apply in such circumstances.

3. STATIONARITY OF DCC MODELS

3.1. Existence of stationary DCC solutions

The AR dynamics of X_t were defined above thanks to T_t and ζ_t , which will be stochastic only through ε_t , i.e., through the t -innovation η_t and the \mathcal{I}_{t-1} -measurable matrix R_t . This creates a major difficulty in proving the existence of stationary solutions. In particular, this means that T_t depends on some components of X_t . Therefore, it will be difficult to find explicit expressions like $X_t = f(\eta_t, \eta_{t-1}, \dots)$ for some deterministic and measurable function f , because the link between T_t and the past innovations (or observations) is highly nonlinear.

To obtain the existence of stationary solutions in the previous DCC model, we will invoke Tweedie’s (1988) criterion. The latter result will provide the existence of an invariant probability measure for the Markov chain defined by (7). This technique has already been used in several papers in econometrics, notably Ling and McAleer (2003) or Ling (1999).

To get the stationarity conditions of (z_t) , we have to control the magnitude of the random matrix T_t , which depends on the random variables ε_{kt}^2 , $k = 1, \dots, m$. The mean of the latter variables is one, but they are not independent. This is in contrast with Ling and McAleer (2003). Moreover, unfortunately, the joint law of $\vec{\varepsilon}_t$ is a function of R_t , i.e., a function of X_{t-1} . That is why we need the following conditions:

Assumption E1. For some $p \geq 1$, $E[\|\eta_t\|^{2p}] < \infty$ and $\rho(T^*) < 1$, where $T^* := \sup_{\mathbf{x} \in \mathbb{R}^d} E[|T_t^{\otimes p}| | X_{t-1} = \mathbf{x}]$.

We reiterate that T_t depends on $\vec{\varepsilon}_t$, that $\varepsilon_t = R_t^{1/2} \eta_t$, and that the components of η_t are uncorrelated. As such, the coefficients of T^* are finite because all of the coefficients of R_t are less than one (in absolute values). When there are no correlation dynamics, the matrices M_k and N_l are zero and we recover CCC models. In the latter case, our Assumption E1 is reduced to the main assumption of Ling and McAleer (Theorem 2.2) that was stated for vectorial ARMA-GARCH models.

Assumption E2. The law of η_t given that $X_{t-1} = \mathbf{x}$ is absolutely continuous with respect to the Lebesgue measure, and its density is denoted by $f_{\eta_t}(\cdot | \mathbf{x})$, for every $\mathbf{x} \in \mathbb{R}^d$ and t . The function $\mathbf{x} \mapsto f_{\eta_t}(\eta | \mathbf{x})$ is continuous for every $\eta \in \mathbb{R}^m$ and t . There exists an integrable function H s.t. $\sup_t \sup_{\mathbf{x} \in \mathbb{R}^d} f_{\eta_t}(\eta | \mathbf{x}) \leq H(\eta)$ for every $\eta \in \mathbb{R}^m$. Moreover, $\sup_t E[\|\eta_t\|^{2p} | X_{t-1} = \mathbf{x}] \leq \bar{h}(\|\mathbf{x}\|)$, for some function \bar{h} that satisfies $\lim_{v \rightarrow +\infty} \bar{h}(v)/v^\gamma = 0$ for every $\gamma > 0$.

The latter technical assumption is trivially satisfied when (η_t) is an i.i.d. sequence of random vectors s.t. $E[\|\eta_t\|^{2p}] < +\infty$. Otherwise, E2 provides some constraints insofar as the law of η_t depends on the past values of the DCC process. Similar conditions may appear in the literature about the nonparametric estimation of conditional expectations. However, most of them relate to the boundedness of \bar{h} and/or its derivatives (as in Assumption 3 in Newey 1997, for instance), or to the moments of \bar{h} (as Assumption 1 in Donald et al. 2003, for instance). Clearly, E2 is weaker than such assumptions.

THEOREM 1. *Under the assumptions A0, E1, and E2, the process (z_t, D_t, R_t) as defined by Equations (1), (2), (3), and (4), possesses a strictly stationary solution. The latter process is measurable with respect to the σ -field \mathcal{I} induced by the observations. Moreover, the $2p$ -th moments of a solution (z_t) are finite.*

Example 1

In practice and for the sake of parsimony, it is usual to assume diagonal-type DCC models, where all the parameter matrices are diagonal, assuming no “cross-effects” in terms of volatilities and/or correlations. This means the nonnegative real numbers $a_u^{(i)}, b_u^{(j)}, m_u^{(k)}$, and $n_u^{(l)}, u = 1, \dots, m$, exist such that

$$A_i = \text{diag} \left(a_1^{(i)}, \dots, a_m^{(i)} \right), i = 1, \dots, r, B_j = \text{diag} \left(b_1^{(j)}, \dots, b_m^{(j)} \right), j = 1, \dots, s,$$

$$M_k = \text{diag} \left(m_1^{(k)}, \dots, m_m^{(k)} \right), k = 1, \dots, \nu, N_l = \text{diag} \left(n_1^{(l)}, \dots, n_m^{(l)} \right), l = 1, \dots, \mu.$$

The associated matrices \tilde{M}_k and \tilde{N}_l are also diagonal. Set $\tilde{M}_k = \text{diag} \left(\tilde{m}_l^{(k)} \right)_{1 \leq l \leq m^*}$, and check that $\tilde{m}_l^{(k)} = m_{\phi_1^{(l)}}^{(k)} m_{\phi_2^{(l)}}^{(k)}$. Now, let us specify the previous Assumption E1 when $p = 1$.

Since $E[\vec{\varepsilon}_{kt} | X_{t-1} = \mathbf{x}] = 1$ for every index k , T^* is simply $|T_t|$, replacing $\vec{\varepsilon}_t$ by one. Denote by P^* the characteristic polynomial of T^* ,

i.e., $P^*(\lambda) = \text{Det}(T^* - \lambda Id)$. It can be seen easily that two polynomials P_1^* and P_2^* s.t. $P^*(\lambda) = P_1^*(\lambda)P_2^*(\lambda)$ exist. Here, P_1^* denotes the characteristic polynomial of the block-matrix $[[T_{ij,t}]_{1, \leq i, j \leq 2}]$, replacing \vec{e}_t by one. P_2^* is the characteristic polynomial of the previous matrix $|T_{33,t}|$. Tedious, but relatively uncomplicated, algebraic calculations provide

$$P_1^*(\lambda) = \pm \lambda^{\pi_1} \prod_{k=1}^m \left(\sum_{i=1}^r a_k^{(i)} \lambda^{r+s-i} + \sum_{j=1}^s b_k^{(j)} \lambda^{r+s-j} - \lambda^{r+s} \right),$$

$$P_2^*(\lambda) = \pm \lambda^{\pi_2} \prod_{l=1}^{m^*} \left(\sum_{k=1}^v \tilde{m}_l^{(k)} \lambda^{v-k} - \lambda^v \right),$$

for some integers π_1 and π_2 . Let λ_0 be a nonzero root of P^* . If λ_0 is a root of P_1^* then there exists an index $k \in \{1, \dots, m\}$ such that $\sum_{i=1}^r a_k^{(i)} \lambda_0^{r+s-i} + \sum_{j=1}^s b_k^{(j)} \lambda_0^{r+s-j} = \lambda_0^{r+s}$. If $|\lambda_0| \geq 1$, this implies $1 \leq \sum_{i=1}^r a_k^{(i)} + \sum_{j=1}^s b_k^{(j)}$.

On the other side and similarly, if λ_0 is a root of P_2^* and if $|\lambda_0| \geq 1$, then there exists $l \in \{1, \dots, m^*\}$ s.t. $1 \leq \sum_{k=1}^v |\tilde{m}_l^{(k)}|$. In other words, a sufficient condition to fulfill Assumption E1 is

$$\sup_{k=1, \dots, m} \sum_{i=1}^r a_k^{(i)} + \sum_{j=1}^s b_k^{(j)} < 1, \text{ and } \sup_{l=1, \dots, m^*} \sum_{k=1}^v |\tilde{m}_l^{(k)}| < 1. \tag{10}$$

Nonetheless, to apply Theorem 1 in the general case, it may be hard to check the condition on the spectral radius of T^* . This is due to the analytical complexity of $T_t^{\otimes p}$, $p > 1$, or to the calculation of its eigenvalues, even when $p = 1$. In the next theorem, we provide more explicit conditions in the case $p = 1$, i.e., so that the second-order moments of (z_t) are finite. These conditions ensure that E1 will be satisfied. In other words, the conditions will be stronger than E1, but they may be more practical. Indeed, it is often important to obtain sufficient conditions that can be written explicitly in terms of the model parameters, for instance, for inference purposes (e.g., the optimization stage to get QML estimates).

Let us consider \mathcal{N} (resp. \mathcal{N}^*) an arbitrary norm for vectors in \mathbb{R}^m (resp. \mathbb{R}^{m^*}). Denote by $\|\cdot\|_{\mathcal{N}}$ and $\|\cdot\|_{\mathcal{N}^*}$ the associated norms for matrices.

THEOREM 2. *If*

$$\sum_{i=1}^r \|A_i\|_{\mathcal{N}} + \sum_{j=1}^s \|B_j\|_{\mathcal{N}} < 1, \text{ and} \tag{11}$$

$$\sum_{k=1}^v \|\tilde{M}_k\|_{\mathcal{N}^*} < 1, \tag{12}$$

then Assumption E1 is satisfied with $p = 1$.

Note that the conditions of Theorem 2 do not depend on the matrices N_l , $l = 1, \dots, \mu$, which is a relatively unexpected result. Once they are satisfied, and under A0 and E2, Theorem 1 applies.

By choosing \mathcal{N} as the maximum norm for vectors, it can easily be checked that $\|A_i\|_{\mathcal{N}} = \sup_{p=1, \dots, m} \sum_{q=1}^m a_{p,q}^{(i)}$, and similarly with the matrices B_j . Alternatively, we can choose $\mathcal{N}(\mathbf{x}) = \|\mathbf{x}\|_2$, that induces the spectral norm $\|A_i\|_{\mathcal{N}} = \|A_i\|_s$. Obviously, we can choose these norms for \mathcal{N}^* and the matrices \tilde{M}_k .

It is often of value to assume that the Markov chain is initialized at $t = 0$ by drawing X_0 following its stationary law. Introducing the filtration $\mathcal{I}_t^* := \sigma(X_0, z_1, \dots, z_t)$, we can easily see that the DCC solution is now measurable with respect to the σ -field induced by the innovations and the initial value, because, whenever $t > 0$,

$$\sigma(X_0, z_1, \dots, z_t) = \sigma(X_0, \varepsilon_1, \dots, \varepsilon_t) = \sigma(X_0, \eta_1, \dots, \eta_t).$$

Example 1 (Continued)

Consider a diagonal-type DCC model and maximum norms for vectors. In this case, the condition (11) becomes

$$\sum_{i=1}^r \sup_{l=1, \dots, m} a_l^{(i)} + \sum_{j=1}^s \sup_{l=1, \dots, m} b_l^{(j)} < 1,$$

and the condition (12) is $\sum_{k=1}^v \sup_{l=1, \dots, m^*} |\tilde{m}_l^{(k)}| < 1$. These two conditions are stronger than (10), as expected.

Example 2

To reduce the number of free parameters even further, scalar-DCC models are often introduced. In this case, all the unknown matrices are simply products of a scalar and an identity matrix:

$$A_i = a^{(i)} I_m, i = 1, \dots, r, B_j = b^{(j)} I_m, j = 1, \dots, s, \\ M_k = m^{(k)} I_m, k = 1, \dots, v, N_l = n^{(l)} I_m, l = 1, \dots, \mu.$$

Such models are very popular, because they allow the number of free parameters to be drastically reduced. With obvious notations, the conditions of Theorem 1 and 2 are the same as above:

$$\sum_{i=1}^r a^{(i)} + \sum_{j=1}^s b^{(j)} < 1, \text{ and } \sum_{k=1}^v |m^{(k)}|^2 < 1. \tag{13}$$

In passing, we recover the usual (second-order and strict) conditions of stationarity for GARCH-type models:

$$0 \leq a^{(i)}, b^{(j)} \leq 1, \text{ and } \sum_{i=1}^r a^{(i)} + \sum_{j=1}^s b^{(j)} < 1.$$

3.2. Uniqueness of stationary DCC solutions

Even if stationary solutions of the DCC model do exist, we are not initially sure *a priori* that they are unique. Besides its theoretical interest, this problem has

practical implications. For instance, for any process, the convergence of simulated trajectories towards the same stationary law, independently of the initialization stage, is a desirable feature. Moreover, the uniqueness of invariant measures of a Markov process implies the ergodicity of the stationary solution (see Douc et al. 2014, Corollary 7.17). This is particularly important for inference purposes. Indeed, the estimation of DCC models is typically based on M-estimates (Quasi Maximum Likelihood, for instance). These techniques rely heavily on uniform Laws of Large Numbers, that are most often deduced from the ergodicity of the process. The conditions for identifiability and consistency rely on some expectations with respect to the underlying invariant measure of the given stationary process. If, for a given set of parameters, several invariant measures exist, then it becomes difficult to check such conditions. Finally, with several underlying invariant measures, we cannot exclude the possibility of switches from one stationary trajectory to another, disturbing the econometric analysis (stationarity tests, statistical uncertainty around estimates, etc.).

Unfortunately, this uniqueness is not given “for free” by Tweedie’s Lemma A.1. Moreover, the usual arguments concerning the uniqueness of stationary GARCH-type solutions do not apply here. Indeed, under the Markov-chain specification given by Equation (7), the matrix T_t is itself a function of the random vector X_t through the $\vec{\varepsilon}_t$ factors. This is a major difference with the CCC case, and we need to find another strategy. In this section, we provide some uniqueness results under some more or less restrictive assumptions.

Now, we will consider only stationary solutions of the DCC model, as given in Section 3.1. We know that such solutions exist under the (sufficient) conditions of Theorem 1 or 2, but it is not necessary to impose such conditions from the outset. Obviously, we will need other technical assumptions.

Assumption U0. The sequence of innovations $(\eta_t)_{t \in \mathbb{Z}}$ is highly stationary and ergodic.

Assumption U1. $\|T_{33}\|_s < 1$.

The matrix T_{33} has been introduced in Subsection 2.2, under the name $T_{33,t}$. Since $T_{33,t}$ does not depend on time, we have removed the index t here.

Thanks to the latter assumptions, we will be able to bound $\|Q_t\|_{\max}$ from above by a stationary process (q_t) , and from below by a constant. Moreover, $\lambda_1(Q_t)$ will be bounded from below. These tools will be crucial in proving the uniqueness of stationary DCC solutions. To do so, let us introduce some intermediate quantities.

- The process (q_t) , defined by

$$q_t := \frac{\|Vech(W_0)\|_2}{1 - \|T_{33}\|_s} + \sqrt{\frac{m^3(m+1)}{2}} \sum_{l=1}^{\mu} \|\tilde{N}_l\|_s \xi_{t-l},$$

where $\xi_t := \sum_{k=0}^{+\infty} \|T_{33}\|_s^k \|\eta_{t-k}\|_2^2$.

- The constants $C_\lambda := \lambda_1(W_0)$ and $C_q := \min_{i=1,\dots,m}(W_0)_{ii}$.
- The constants

$$C_\lambda^* := \frac{\lambda_1(W_0)}{1 - \sum_{k=1}^v (m^{(k)})^2}, \text{ and } C_q^* := \frac{\min_{i=1,\dots,m}(W_0)_{ii}}{1 - \sum_{k=1}^v (m^{(k)})^2}.$$

- $\kappa = \max(v, \mu)$ and, for every $j = 1, \dots, \kappa$, set

$$\beta_{j,t} := \mathbf{1}(j \leq v) \|M_j\|_s^2 + \mathbf{1}(j \leq \mu) \|N_j\|_s^2 \frac{4(2m+1)m^{1/2}}{\sqrt{C_\lambda^* C_q}} \|\eta_t\|_2^2 \sqrt{q_t}.$$

Let N_t^* be the (κ, κ) -squared random matrix

$$N_t^* := \begin{bmatrix} \beta_{1,t} & \beta_{2,t} & \cdots & \cdots & \beta_{\kappa,t} \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}.$$

Note that the sequences (ξ_t) , (q_t) , and (N_t^*) are stationary and ergodic because any ξ_t , q_t , or N_t^* is a measurable function of the innovations (η_t) that are stationary and ergodic under Assumption U0.

Assumption U2. $E[\ln^+ \|N_t^*\|] < \infty$ and the top Lyapunov exponent of the sequence (N_t^*) , defined by $\gamma_N := \lim_{t \rightarrow +\infty} t^{-1} E[\ln(\|N_1^* N_2^* \dots N_t^*\|)]$, is strictly negative.

Such conditions are standard in the GARCH literature (see Francq and Zakoian, 2010, Section 2.2.2, for instance). Note that $\gamma_N \leq E[\ln \|N_1^*\|]$ for any norm $\|\cdot\|$.

Actually, the technical assumptions U1 and U2 above will ensure the uniqueness of (ε_t) , (Q_t) , and (R_t) only. To get the uniqueness of (D_t) and then of (z_t) itself, we need a last assumption: with the notations of Subsection 2.2, set

$$\bar{T}_t := \begin{bmatrix} T_{11,t} & T_{12,t} \\ T_{21,t} & T_{22,t} \end{bmatrix}, \text{ and } \bar{T}^* = E[\bar{T}_t].$$

Note that \bar{T}^* does not depend on any particular sequence (ε_t) nor t , because $E[\varepsilon_{kt}^2] = 1$ for every k .

Assumption U3. $\rho(\bar{T}^*) < 1$

THEOREM 3. *Under A0 and U0–U3, a strictly stationary solution of the DCC model is unique and ergodic, given a sequence (η_t) .*

The latter result can be strengthened in the following particular case, which is commonly encountered in the literature.

Assumption U4. The underlying DCC model is “partially” scalar, i.e., scalars $m^{(k)}$ exist such that $M_k = m^{(k)}I_m$ for all $k = 1, \dots, v$. Moreover, $\rho(M^*) < 1$ by setting

$$M^* := \begin{bmatrix} (m^{(1)})^2 & (m^{(2)})^2 & \dots & \dots & (m^{(v)})^2 \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}.$$

Obviously, U4 is not mandatory to get our uniqueness result, even if it allows the technical condition U2 to be weakened most often, by lowering the $\beta_{j,t}$ terms. In every case, this “partially” scalar case encompasses the common practice of scalar DCC (or scalar multivariate GARCH) models.

COROLLARY 1. Under A0 and U0–U4, a strictly stationary solution of the DCC model is unique and ergodic, given a sequence (η_t) , replacing C_λ (resp. C_q) by C_λ^* (resp. C_q^*) in U2.

Example 2 (Continued)

In the case of scalar DCC models of order one, it is easy to specify the conditions above. Here, $r = s = v = \mu = 1$,

$$A_1 = a^{(1)}I_m, \quad B_1 = b^{(1)}I_m, \quad M_1 = m^{(1)}I_m, \quad N_1 = n^{(1)}I_m.$$

Assumptions U1 and U4 are equivalent and mean $|m^{(1)}| < 1$. Assumption U4 is fulfilled if $E[\ln \|N_1^*\|_{\max}] < 0$, or if

$$E \left[\ln \left((m^{(1)})^2 + (n^{(1)})^2 \frac{4(2m+1)m^{1/2}}{\sqrt{C_\lambda} C_q} \|\eta_t\|_2^2 \sqrt{q_t} \right) \right] < 0. \tag{14}$$

This expectation could be easily evaluated by simulation, by noting that η_t and q_t are independent. Finally,

$$\bar{T}^* = \begin{bmatrix} a^{(1)} & b^{(1)} \\ a^{(1)} & b^{(1)} \end{bmatrix} \otimes I_m.$$

Through elementary algebra, it can be checked that the characteristic function of \bar{T}^* is the function $x \mapsto (-x)^m (a^{(1)} + b^{(1)} - x)^m$. Then Assumption U3 means $a^{(1)} + b^{(1)} < 1$. Therefore, as expected, the conditions required for stationary DCC solutions to be unique are more demanding than for them to simply exist, due to U2. Generally, the latter condition will be fulfilled more easily if $\sup_t \|N_t\|_s$ is “a lot smaller” than one, if m is not “too large”, and if the tails of η_t are not “too heavy”.

4. DISCUSSION AND PRACTICAL CONSIDERATIONS

Now, let us discuss the sufficient conditions to obtain the existence of stationary DCC solutions, as given in Theorems 1 and 2. The most explicit ones are (11) and (12). Since scalar DCC models are by far the most commonly used models in the literature, we focus on the conditions of the previous examples 1 and 2 above, particularly (13). In the latter case, the conditions on the coefficients of the volatility process are those generally applied in the univariate GARCH literature. It can be proved they are necessary and sufficient for the existence of second-order and strictly stationary GARCH solutions (see Francq and Zakoian, 2010, Theorem 2.6 and Remark 2.6). More interestingly, we can check empirically how tight the (now) new conditions on the coefficients of the (Q_t) process are. In other words, in Example 2, is the constraint $\sum_{k=1}^p |m^{(k)}|^2 < 1$ close to a necessary and sufficient condition for generating stationary trajectories of (z_t) , (R_t) and/or (Q_t) s?

For illustrative purposes, we have considered a very simple bivariate scalar DCC model of order one, given by

$$\begin{aligned}
 h_{k,t} &= v_0 + a h_{k,t-1} + b \varepsilon_{k,t}^2, \quad k = 1, 2, \\
 v_0 &= 1/4, \quad a = 0.8, \quad b = 0.1, \\
 Q_t &= W_0 + (m^{(1)})^2 Q_{t-1} + (n^{(1)})^2 \varepsilon_{t-1} \varepsilon'_{t-1}, \\
 W_0 &= I_2/2 + ee'/2, \quad e = [1, 1]',
 \end{aligned}$$

with our notations. The latter process is generated by i.i.d. innovations (η_t) that are independent standard bivariate Gaussian vectors. We initialize the process at $t = 0$ with $Q_0 = R_0 = I_2$ and $h_{k,0} = 1/2, k = 1, 2$.

In the paper, we have stated theoretically that the values of the coefficients of the matrices $N_l, l = 1, \dots, \mu$ (or the coefficients $n^{(l)}$ in the scalar case) do not matter to obtain the existence of stationary DCC solutions. We have verified this fairly counter-intuitive fact empirically: with the model above, different coefficients $n^{(1)}$ do not seem to modify the shape of the trajectories we generate, independently of the other parameters. Therefore, our experiments will lead with a fixed value $n^{(1)} = \sqrt{3}$. Note that this value is larger than one and this could be seen “naively” as a source of nonstationarity.

As expected, the value one is key for $m^{(1)}$. When the latter is very close to one but less than one ($(m^{(1)})^2 = 0.999$ in our case), we check that the simulated trajectories of (z_t) , (Q_t) , and (R_t) look stationary, once the influences of the starting values have been forgotten (broadly speaking when $t \geq 4000$): see Figure 1 (solid lines). On the other side, when this auto-regressive parameter is larger than one, even by a small amount ($(m^{(1)})^2 = 1.001$ in our case), we observe that the (Q_t) trajectories explode: see Figure 1 (dashed lines). Apparently, this is not the case for the (R_t) correlation coefficients. They tend towards some constant levels (Figure 2), but differ from one experiment to another one. This phenomenon is a consequence of the normalization stage (3), but it seems difficult

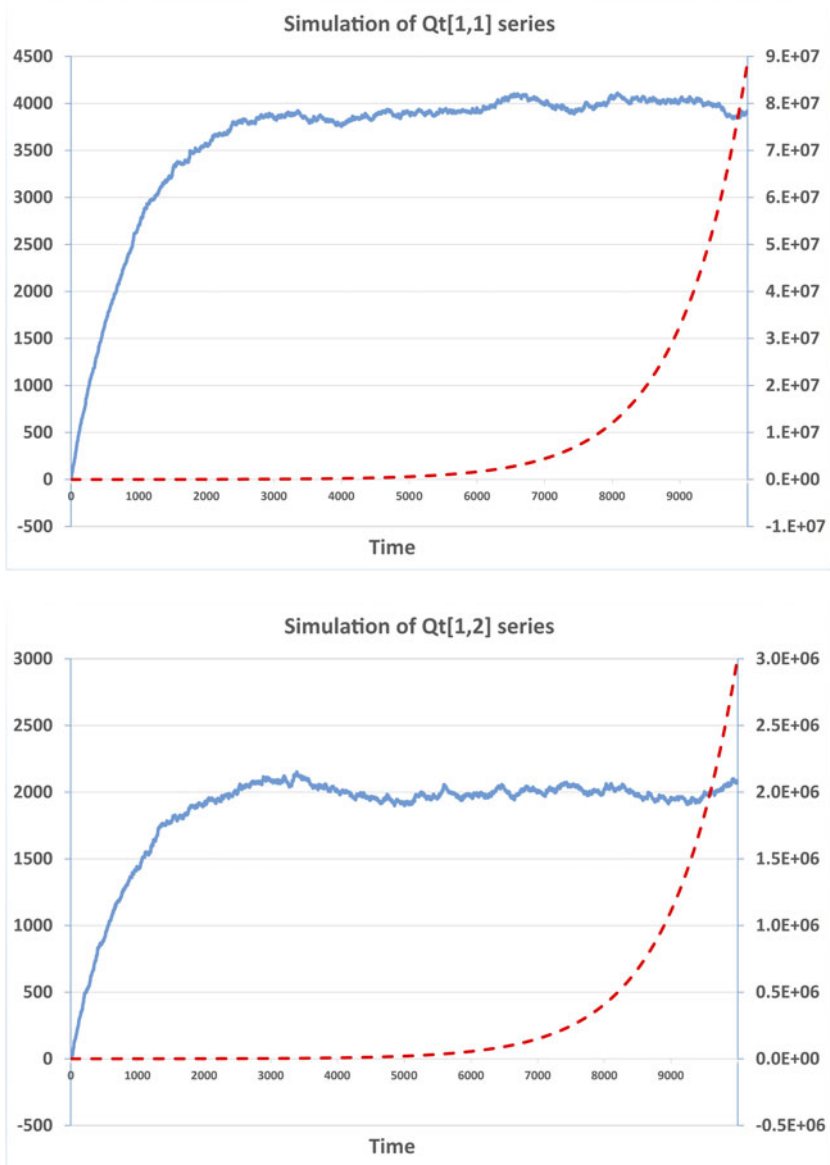


FIGURE 1. Some simulated trajectories of $Q_t[1,1]$ (top) and $Q_t[1,2]$ (bottom) when $m^{(1)} = \sqrt{0.999}$ (solid line, left axis) or $m^{(1)} = \sqrt{1.001}$ (dashed line, right axis).

to maintain that feature will happen for almost every trajectory and any DCC model. Indeed, by managing increasingly (very) high numbers with the (Q_t) process, we cannot exclude the possibility of spurious unexpected (R_t) behaviors.

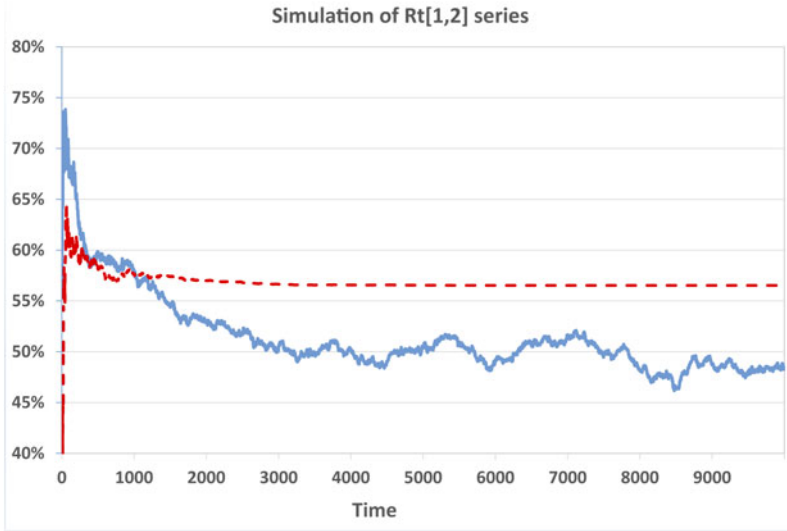


FIGURE 2. Some simulated trajectories of $R_t[1,2]$ when $m^{(1)} = \sqrt{0.999}$ (solid line) or $m^{(1)} = \sqrt{1.001}$ (dashed line).

Moreover, we have checked that the (z_t) trajectories that we generated in this case do not seem to exhibit nonstationary patterns visually. This is logical because our DCC model tends towards a CCC model in such situations. Therefore, it is likely that modelers will be able to manage (i.e., evaluate and simulate numerically) DCC trajectories in practice, even if (12) or its generalizations are not satisfied. Actually, this task remains feasible as long as the numerical values of (Q_t) are manageable (i.e., not too large) by our software. This is the case when the number of dates is not too large.

We have replaced Gaussian innovations (η_t) by fat-tailed random vectors, to check to what extent this may be a source of instability. The η_{kt} components, $k = 1, 2$ and $t = 0, \dots, T$, have been drawn following mutually independent standardized Student laws with ν degrees of freedom, $\nu > 0$. We reiterate here that the η_t -moments of order ν or higher do not exist. As long as $\nu \geq 2$ and $m^{(1)} < 1$, we do not observe explosive patterns for (z_t) . The components of R_t appear to become stationary, even if the decrease in initial value effects is very slow when ν is close to two. On the contrary, when $\nu < 2$, the processes (Q_t) and (R_t) are highly unstable. The former exhibit very spiky trajectories, while the latter often tend to be attracted by the 1 or (-1) area. Besides, some (z_t) trajectories reach very high and unrealistic values (10^{80} , for instance). In every case, when $m^{(1)} \geq 1$, the (Q_t) trajectories explode and the (R_t) ones tend to constant values that depend on each experiment. In such cases, the return process (z_t) may reach huge values, but only when $\nu \leq 2$, apparently. This analysis illustrates the necessity of considering innovations with finite second-order moments, and the fact that higher-order moments are not mandatory to obtain stationary solutions.

Concerning the sufficient conditions that guarantee that stationary DCC solutions are unique, it is more difficult to evaluate their tightness because they are more intricate and involve too many model characteristics. Nonetheless, with the simple scalar DCC model used in Example 2, we observe that the key condition (14) will be more demanding when the number m of underlyings increases. On the contrary, smaller values $|n^{(1)}|$ would help. And the partially scalar case induces significantly less demanding conditions than the general case, because the values of the constants C_λ^* and C_q^* in the denominator are a lot higher than C_λ and C_q , respectively. The effect of $|m^{(1)}|$ is ambiguous because it appears in several quantities, especially T_{33} and C_λ^* .

REFERENCES

- Aielli, G.P. (2013) Dynamic Conditional Correlation: on properties and estimation. *Journal of Business & Economic Statistics* 31, 282–299.
- Billio, M. & Caporin, M. (2005) Multivariate Markov switching dynamic conditional correlation GARCH representations for contagion analysis. *Statistical Methods & Applications* 14, 145–161.
- Billio, M., M. Caporin & M. Gobbo (2006) Flexible dynamic conditional correlation multivariate GARCH for asset allocation. *Applied Financial Economics Letters* 2, 123–130.
- Bollerslev, T. (1990) Modeling the Coherence in Short Run Nominal Exchange Rates: A Multivariate generalized ARCH Model. *The Review of Economics and Statistics* 72, 498–505.
- Bougerol, P. & Picard, N. (1992) Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* 52, 1714–1729.
- Boussama, F., F. Fuchs & R. Stelzer (2011) Stationarity and Geometric Ergodicity of BEKK Multivariate GARCH Models. *Stochastic Processes and their Applications* 121, 2331–2360.
- Cappiello L., R.F. Engle & K. Sheppard (2006) Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics* 4, 537–572.
- Caporin, M. & M. McAleer (2013) Ten Things You Should Know about the Dynamic Conditional Correlation Representation. *Econometrics* 1, 115–126.
- Comte, F. & O. Lieberman (2003) Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis* 84, 61–84.
- Donald, S.G., Imbens, G.W. & W.K. Newey (2003) Empirical likelihood estimation and consistent tests with conditional moment restrictions. *Journal of Econometrics* 117, 55–93.
- Douc, R., E. Moulines & D. Stoffer (2014) *Nonlinear Time Series*. Chapman & Hall.
- Engle, R.F. & F.K. Kroner (1995) Multivariate simultaneous generalized ARCH. *Econometric Theory* 11, 122–150.
- Engle, R.F. & K. Sheppard (2001) Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Working Paper 2001-15, University of California at San Diego.
- Engle, R.F. (2002) Dynamic conditional correlation: a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics* 20, 339–350.
- Fermanian, J.-D. & H. Malongo (2013) On the link between instantaneous volatilities, switching regime probabilities and correlation dynamics. Working paper Crest.
- Franco, C. & J.-M. Zakořan (2010) *GARCH models*. Wiley.
- Franses, P.H. & C.M. Hafner (2009) A generalized dynamic conditional correlation model: Simulation and application to many assets. *Econometric Reviews* 28, 612–631.
- Higham, N.J. (2008) *Functions of Matrices*. SIAM.
- Kasch, M. & M. Caporin (2013) Volatility threshold dynamic conditional correlations: An international analysis. *Journal of Financial Econometrics* 11, 706–742.
- Ling, S. (1999) On the probabilistic properties of a double threshold ARMA conditional heteroskedastic model. *Journal of Applied Probability* 36, 688–705.

- Ling, S. & M. McAleer (2003) Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory* 19, 280–310.
- Lütkepohl, H. (1996) *Handbook of Matrices*. Wiley.
- Newey, W.K. (1997) Convergence rates and asymptotic normality for series estimators. *Journal of Econometrics* 79, 147–168.
- Otranto, E. & L. Bauwens (2015) Modeling the dependence of conditional correlations on volatility. *Journal of Business & Economic Statistics*. Available online.
- Pelletier, D. (2006) Regime Switching for Dynamic Correlations. *Journal of Econometrics* 131(1–2), 445–473.
- Serre, D. (2010) *Matrices: Theory and Applications*. GTM 216, Springer.
- Tse, Y.K. & A.K.C. Tsui (2002) A multivariate GARCH model with time-varying correlations. *Journal of Business and Economic Statistics* 20, 351–362.
- Tweedie, R.L. (1988) Invariant measure for Markov chains with no irreducibility assumptions. *Journal of Applied Probability* 25A, 275–285.

APPENDIX A. Technical lemmas

We recall Tweedie’s criterion, a key tool to prove the existence of an invariant probability measure for a Markov chain. This result has a remarkable advantage: contrary to more commonly used techniques (based on some Lyapunov–Foster conditions, for example), it is not necessary to state the irreducibility of the underlying Markov chain, to obtain the existence of stationary solutions. Technically speaking, proving the irreducibility of such a nonlinear Markov chain is a very challenging task in general.

Let $(X_t)_{t=1,2,\dots}$ be a temporally homogeneous Markov chain with a locally compact separable metric state space (S, \mathcal{B}) . The transition probability is $P(x, A) = P(X_t \in A | X_{t-1} = x)$, where $x \in S$ and $A \in \mathcal{B}$. Tweedie’s (1988) Theorem 2 provides:

LEMMA A.1. *Suppose that (X_t) is a Feller chain, i.e., for each bounded continuous function h on S , the function of \mathbf{x} given by $E[h(X_t) | X_{t-1} = \mathbf{x}]$ is also continuous.*

1. *If there exists, for some compact set $A \in \mathcal{B}$, a nonnegative function g and $\varepsilon > 0$ satisfying*

$$\int_{A^c} P(x, dy)g(y) \leq g(x) - \varepsilon, \quad x \in A^c, \quad (\text{A.1})$$

then there exists a σ -finite invariant measure μ for P with $0 < \mu(A) < \infty$.

2. *Furthermore, if*

$$\int_A \mu(dx) \left[\int_{A^c} P(x, dy)g(y) \right] < \infty, \quad (\text{A.2})$$

then μ is finite and hence $\pi = \mu / \mu(S)$ is an invariant probability measure.

3. *Furthermore, if*

$$\int_{A^c} P(x, dy)g(y) \leq g(x) - f(x), \quad x \in A^c, \quad (\text{A.3})$$

then μ admits a finite f -moment; that is, $\int_S \mu(dy)f(y) < \infty$.

The following Lemma is our version of Lemma A.2 in Ling and McAleer (2003). Therefore, its proof is omitted.

LEMMA A.2. For a given squared matrix T , if $\rho(|T|) < 1$, then there exists a vector $M > 0$ such that $(Id - |T'|)M > 0$.

APPENDIX B. Proof of Proposition 1

Note that ε_t (or $\vec{\varepsilon}_t$, or even $Vech(\varepsilon_t \varepsilon_t')$) is a function of the couple (R_t, η_t) only. Due to (3) and (4), R_t is a deterministic function of X_{t-1} . Since η_t is Markov with respect to \mathcal{I} , the law of ε_t knowing \mathcal{I}_{t-1} is the law of ε_t knowing X_{t-1} merely. The same assertion applies with T_t, ζ_t , or X_t itself, instead of ε_t .

In other words, the nonlinearity of the DCC model comes mainly from $\vec{\varepsilon}_t$ in T_t . But there exist constant matrices (of zeros and ones) F and G such that (7) can be rewritten

$$X_t = (F \cdot \vec{\varepsilon}_t^*) \odot T_o X_{t-1} + (G \cdot Vech(\varepsilon_t \varepsilon_t')^*) \odot \zeta_o, \tag{B.1}$$

where T_o (resp. ζ_o) is the T_t matrix (resp. ζ_t vector) when $\varepsilon_t = 1$, $\vec{\varepsilon}_t^* := [\vec{\varepsilon}_t', 1]'$ and $Vech(\varepsilon_t \varepsilon_t')^* := [Vech(\varepsilon_t \varepsilon_t')', 1]'$. Since $\varepsilon_t = R_t^{1/2} \eta_t$ and since R_t is a measurable function of X_{t-1} , then X_t is clearly a function of X_{t-1} and of the innovation η_t only, that are Markov. These arguments prove the Markovian structure of the (X_t) process under A0. ■

APPENDIX C. Proof of Theorem 1

First, let us check that (X_t) is a Feller chain in a convenient space, to be able to apply Lemma A.1 afterwards. Let h be a bounded and continuous function on \mathbb{R}^d . Clearly,

$$\begin{aligned} E[h(X_t) | X_{t-1} = \mathbf{x}] &= E[h(T_t \mathbf{x} + \zeta_t) | X_{t-1} = \mathbf{x}] \\ &= E[h(\psi_1(\varepsilon_t \varepsilon_t') \mathbf{x} + \psi_2(\varepsilon_t \varepsilon_t')) | X_{t-1} = \mathbf{x}], \end{aligned}$$

for some continuous transforms ψ_1 and ψ_2 . Note that $\varepsilon_t = R_t^{1/2} \eta_t$ and that $R_t^{1/2}$ is a continuous function of X_{t-1} . Indeed, $R_t \mapsto R_t^{1/2}$ is continuous (see, e.g., Proposition 6.3 in Serre (2010)), and $X_{t-1} \mapsto R_t$ is continuous by construction. Then,

$$E[h(X_t) | X_{t-1} = \mathbf{x}] = E\left[h \circ \tilde{\psi}(\mathbf{x}, \eta_t) | X_{t-1} = \mathbf{x}\right] = \int h \circ \tilde{\psi}(\mathbf{x}, \eta) f_{\eta_t}(\eta | \mathbf{x}) d\eta,$$

for some continuous transform $\tilde{\psi}$. Now, consider a sequence of vectors (\mathbf{x}_n) that tends to \mathbf{x} when $n \rightarrow \infty$. Since h is bounded and since the sequence $(h \circ \tilde{\psi}(\mathbf{x}_n, \eta) f_{\eta_t}(\eta | \mathbf{x}_n))_n$ is convergent for every η , we can apply the dominated convergence theorem under E2. We deduce that $\mathbf{x} \mapsto E[h(X_t) | X_{t-1} = \mathbf{x}]$ is continuous and then (X_t) is Feller.

Note that the vector X_t belongs to the metric space \mathbb{R}^d , endowed with the usual topology. Since we impose that the matrices Q_t will be positive definite, X_t will live in a subspace of \mathbb{R}^d , where $X_t^{(3)}$ will gather only the components of definite positive matrices. It is easy to check that this subspace is separable and locally compact. Therefore, the assumptions of Lemma A.1 are satisfied.

Second, set $g(\mathbf{x}) = 1 + |\mathbf{x}|^{\otimes p} M$, for an arbitrary positive vector M , that will be chosen after. Let us check that the latter function can be invoked as in Lemma A.1. Clearly,

$$E[g(X_t) | X_{t-1} = \mathbf{x}] = 1 + E\left[|(T_t \mathbf{x} + \zeta_t)|^{\otimes p} M | X_{t-1} = \mathbf{x}\right].$$

By expanding the Kronecker products, we can check that $(T_t \mathbf{x} + \zeta_t)^{\otimes p} = (T_t \mathbf{x})^{\otimes p} + R(\mathbf{x})$, with

$$\|R(\mathbf{x})\| \leq C_0 \left(\|\zeta_t\| \cdot \|(T_t \mathbf{x})^{\otimes (p-1)}\| + \dots + \|\zeta_t\|^{p-1} \cdot \|(T_t \mathbf{x})\| + \|\zeta_t\|^p \right),$$

for some positive constant C_0 and any multiplicative matrix norm $\|\cdot\|$.

Note that $(T_t \mathbf{x})^{\otimes k} = T_t^{\otimes k} \cdot \mathbf{x}^{\otimes k}$. Recall that T_t is a function of $\vec{\varepsilon}_t$, i.e., of ε_t . Then, its conditional law depends on R_t , i.e., it is a function of X_{t-1} . We deduce

$$\begin{aligned} E \left[|(T_t \mathbf{x})^{\otimes p}| \mid X_{t-1} = \mathbf{x} \right]' M &\leq |\mathbf{x}^{\otimes p}|' E \left[|T_t^{\otimes p}|' \mid X_{t-1} = \mathbf{x} \right] M \\ &\leq |\mathbf{x}^{\otimes p}|' \left(\sup_{\mathbf{x} \in \mathbb{R}^d} E \left[|T_t^{\otimes p}|' \mid X_{t-1} = \mathbf{x} \right] \right) M \\ &\leq |\mathbf{x}^{\otimes p}|' (T^*)' M. \end{aligned}$$

Now, choose M as provided by Lemma A.2, when the matrix T in this lemma is replaced by T^* .

Moreover, $\varepsilon_t = R_t^{1/2} \eta_t$, and the (positive definite) matrix $R_t^{1/2}$ can be chosen so that all its coefficients are less than $m^{1/2}$ (diagonalize this matrix on an orthonormal basis and invoke Cauchy–Schwartz inequality). This implies that constants α_k exist such that $\|Vech(\varepsilon_t \varepsilon_t')^{\otimes k}\| \leq \alpha_k \|Vech(\eta_t \eta_t')^{\otimes k}\|$ when $k \leq p$. Since $E[\|\eta_t\|^2 \mid X_{t-1} = \mathbf{x}] < \bar{h}(\|\mathbf{x}\|)$ by assumption, some constants $c_{k,l}$ such that $E_{t-1}[\|\zeta_t\|^k \cdot \|\vec{\varepsilon}_t\|^l] < c_{k,l} \bar{h}(\|\mathbf{x}\|)^{(k+l)/p}$ for any couple (k, l) , $k + l \leq p$ exist. We deduce the boundedness of $E[T_t^{\otimes k} \mid X_{t-1} = \mathbf{x}]$, $k \leq p$, and

$$E[\|R(\mathbf{x})\| \mid X_{t-1} = \mathbf{x}] \leq C_1 \left(\bar{h}(\|\mathbf{x}\|)^{1/p} \|\mathbf{x}^{\otimes (p-1)}\| + \dots + \bar{h}(\|\mathbf{x}\|)^{(p-1)/p} \|\mathbf{x}\| + \bar{h}(\|\mathbf{x}\|)^p \right),$$

for some positive constant C_1 . Applying E2, we have obtained

$$\begin{aligned} E[g(X_t) \mid X_{t-1} = \mathbf{x}] &\leq 1 + |\mathbf{x}^{\otimes p}|' (T^*)' M + O \left(\sum_{k=0}^{p-1} \|\mathbf{x}^{\otimes k}\| \cdot \|\mathbf{x}\|^a \right) \\ &\leq g(\mathbf{x}) - |\mathbf{x}^{\otimes p}|' (Id - (T^*)') M + O \left(\sum_{k=0}^{p-1} \|\mathbf{x}^{\otimes k}\| \cdot \|\mathbf{x}\|^a \right), \quad (\mathbf{C.1}) \end{aligned}$$

for every constant $a > 0$. By Lemma A.2, $(Id - (T^*)')M$ is strictly positive. Then, a positive constant c_0 exists such that

$$|\mathbf{x}^{\otimes p}|' (Id - (T^*)') M \geq c_0 \sum_{j=1}^d |x_j|^p,$$

for every d -dimensional vector \mathbf{x} . Set $N(\mathbf{x}) := \sum_{j=1}^d |x_j|^p$. By a similar reasoning, a positive constant c_1 exists such that $g(\mathbf{x}) \geq c_1 N(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^d$. Moreover, by applying Hölder’s inequality, we have

$$\sum_{i_1, \dots, i_k} |x_{i_1} \cdots x_{i_k}| = \left(\sum_{i=1}^d |x_i| \right)^k \leq \left(\sum_{i=1}^d |x_i|^p \right)^{k/p} = d^{k/p} N(\mathbf{x})^{k/p},$$

for every $k \leq p$. Then, a positive constant c_2 exists such that

- $g(\mathbf{x}) \leq 1 + \|M\| \sum_{i_1, \dots, i_p} |x_{i_1} \cdots x_{i_p}| \leq 1 + c_2 N(\mathbf{x})$, and
- every “residual” term $\|\mathbf{x}^{\otimes k}\|$ is bounded above by (a scalar times) $N(\mathbf{x})^{k/p}$, when $k < p$.

Therefore, this provides

$$E[g(X_t) | X_{t-1} = \mathbf{x}] \leq g(\mathbf{x}) \left[1 - c_0 \frac{N(\mathbf{x})}{g(\mathbf{x})} + O\left(\sup_{k=0, \dots, p-1} \frac{N(\mathbf{x})^{(k+a)/p}}{g(\mathbf{x})} \right) \right] \\ \leq g(\mathbf{x}) \left[1 - \frac{c_0 N(\mathbf{x})}{1 + c_2 N(\mathbf{x})} + O\left(\sup_{k=0, \dots, p-1} \frac{N(\mathbf{x})^{(k+a)/p}}{c_1 N(\mathbf{x})} \right) \right].$$

Let us define the set $A := \{\mathbf{x} \in \mathbb{R}^d | N(\mathbf{x}) \leq \Delta\}$, for some $\Delta > 1$. When Δ is sufficiently large, we obtain, for any $\mathbf{x} \notin A$ and a power a s.t. $0 < a < 1/p$,

$$0 \leq E[g(X_t) | X_{t-1} = \mathbf{x}] \leq g(\mathbf{x}) \left[1 - \frac{c_0}{2c_2} + O\left(\frac{\Delta^{a-1/p}}{c_1} \right) \right] < g(\mathbf{x}) \left[1 - \frac{c_0}{3c_2} \right]. \tag{C.2}$$

Since $g(\mathbf{x}) \geq 1$, it follows that $E[g(X_t) | X_{t-1} = \mathbf{x}] \leq g(\mathbf{x}) - \varepsilon$ for some $\varepsilon > 0$. This proves Equation (A.1) in Lemma A.1. Therefore, there exists a σ -finite invariant measure μ for the Markov chain (X_t) , and $0 < \mu(A) < \infty$.

For any $\mathbf{x} \in A$, Equation (C.1) provides

$$E[g(X_t) | X_{t-1} = \mathbf{x}] \leq g(\mathbf{x}) + O\left(\sum_{k=0}^{p-1} \|\mathbf{x}^{\otimes k}\| \cdot \|\mathbf{x}\|^a \right) \leq c_3 \Delta^{(1+a/p)}$$

for some constant c_3 that does not depend on \mathbf{x} . Then,

$$\int_A \mu(dx) \left[\int_{A^c} P(x, dy) g(y) \right] \leq \int_A \mu(dx) E[g(X_t) | X_{t-1} = \mathbf{x}] \leq C \Delta^{1+a/p} \mu(A) < \infty.$$

We deduce that μ is finite and hence $\pi = \mu/\mu(\mathbb{R}^d)$ is an invariant probability measure of (X_t) . This implies that a strictly stationary solution satisfying (7) exists, still denoted by (X_t) .

Third, by invoking Equation (C.2), we get (A.3) in Lemma A.1 with $f(\mathbf{x}) = \beta g(\mathbf{x})$, for some $\beta \in (0, 1)$. Since $g(\mathbf{x}) \geq c_1 N(\mathbf{x})$, we obtain

$$E_\pi [N(X_t)] < \infty. \tag{C.3}$$

In particular, invoking Hölder’s inequality, this implies that $E_\pi [z_{it}^{2k}] < \infty$, for every $i = 1, \dots, m$ and every $k \leq p$. ■

Remark C.1. Equation (C.3) provides a lot more than only the finiteness of z ’s moments. Overall, it means that

$$E_\pi \left[\sum_{i=1}^m h_{it}^p \right] < \infty, \quad E_\pi \left[\sum_{i=1}^m z_{it}^{2p} \right] < \infty, \\ E_\pi \left[\sum_{i,j=1}^m |Q_{ij,t}|^p \right] < \infty, \quad \text{and} \quad E_\pi \left[\sum_{i=1}^m |\varepsilon_{it}|^{2p} \right] < \infty.$$

APPENDIX D. Proof of Theorem 2

Let us consider λ , a nonzero eigenvalue of T^* , when $p = 1$. We can easily check that this matrix is simply T_t , by replacing \bar{e}_t by one, and replacing the coefficients of the matrices \tilde{M}_k and \tilde{N}_l by their absolute values (remember that the matrices A_i and B_j are already nonnegative). Let $\mathbf{v} = (\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}, \mathbf{v}^{(4)})$ be the associated eigenvector, where the dimensions of the subvectors $\mathbf{v}^{(k)}$, $k = 1, \dots, 4$ are consistent with those of X_t in (6). We can further split the latter subvectors, so that they comply with the matrices A_i, B_j, \tilde{M}_k , and \tilde{N}_l . With obvious vector sizes, we will denote $\mathbf{v}^{(1)} = (\mathbf{v}_1^{(1)}, \dots, \mathbf{v}_r^{(1)})$, $\mathbf{v}^{(2)} = (\mathbf{v}_1^{(2)}, \dots, \mathbf{v}_s^{(2)})$, $\mathbf{v}^{(3)} = (\mathbf{v}_1^{(3)}, \dots, \mathbf{v}_v^{(3)})$ and $\mathbf{v}^{(4)} = (\mathbf{v}_1^{(4)}, \dots, \mathbf{v}_\mu^{(4)})$.

By simple block-matrix calculations, the relation $T^*\mathbf{v} = \lambda\mathbf{v}$ implies

$$\begin{aligned} \mathbf{v}_1^{(1)} &= \mathbf{v}_1^{(2)} = \sum_{i=1}^r \frac{A_i \mathbf{v}_1^{(1)}}{\lambda^i} + \sum_{j=1}^s \frac{B_j \mathbf{v}_1^{(2)}}{\lambda^j}, \\ \mathbf{v}_1^{(3)} &= \sum_{k=1}^v \frac{\tilde{M}_k \mathbf{v}_1^{(3)}}{\lambda^k} + \sum_{l=1}^\mu \frac{\tilde{N}_l \mathbf{v}_1^{(4)}}{\lambda^l}, \text{ and } \mathbf{v}_1^{(4)} = 0. \end{aligned}$$

Note that $\mathbf{v}_i^{(1)} = \mathbf{v}_1^{(1)} / \lambda^i$ and $\mathbf{v}_j^{(2)} = \mathbf{v}_1^{(2)} / \lambda^j$ for every i and j . Moreover, $\mathbf{v}_k^{(3)} = \mathbf{v}_1^{(3)} / \lambda^k$ and $\mathbf{v}_l^{(4)} = 0$ for every k and l .

If $\lambda \geq 1$, then

$$\begin{aligned} \mathcal{N}(\mathbf{v}_1^{(1)}) &\leq \sum_{i=1}^r \|A_i\|_{\mathcal{N}} \frac{\mathcal{N}(\mathbf{v}_1^{(1)})}{|\lambda|^i} + \sum_{j=1}^s \|B_j\|_{\mathcal{N}} \frac{\mathcal{N}(\mathbf{v}_1^{(1)})}{|\lambda|^j} \\ &\leq \mathcal{N}(\mathbf{v}_1^{(1)}) \left(\sum_{i=1}^r \|A_i\|_{\mathcal{N}} + \sum_{j=1}^s \|B_j\|_{\mathcal{N}} \right). \end{aligned}$$

Similarly,

$$\mathcal{N}^*(\mathbf{v}_1^{(3)}) \leq \sum_{k=1}^v \|\tilde{M}_k\|_{\mathcal{N}^*} \frac{\mathcal{N}^*(\mathbf{v}_1^{(3)})}{|\lambda|^k} \leq \mathcal{N}^*(\mathbf{v}_1^{(3)}) \sum_{k=1}^v \|\tilde{M}_k\|_{\mathcal{N}^*}.$$

Since $\mathbf{v} \neq 0$, we obtain

$$1 \leq \sum_{i=1}^r \|A_i\|_{\mathcal{N}} + \sum_{j=1}^s \|B_j\|_{\mathcal{N}}, \text{ or } 1 \leq \sum_{k=1}^v \|\tilde{M}_k\|_{\mathcal{N}^*}.$$

This proves the result. ■

APPENDIX E. Proof of Theorem 3

Suppose that two strongly stationary solutions (X_t) and (\tilde{X}_t) exist. Since both satisfy Equation (7), with obvious notations, we can write for every t

$$X_t = T_t \cdot X_{t-1} + \zeta_t, \text{ and } \tilde{X}_t = \tilde{T}_t \cdot \tilde{X}_{t-1} + \tilde{\zeta}_t.$$

Note that the difference between T_t and \tilde{T}_t is only due to the (a priori different) factors ε_t and $\tilde{\varepsilon}_t$. We want to prove that, for every t , almost certainly $X_t = \tilde{X}_t$.

The problem will be solved if we prove the uniqueness of the process $(X_t^{(3)}, X_t^{(4)})$, given by subvectors of (X_t) . For the moment, assume it has been proved. Recall that

$$X_t^{(3)} := (\text{Vech}(Q_t), \dots, \text{Vech}(Q_{t-\nu+1}))', \text{ and}$$

$$X_t^{(4)} := (\text{Vech}(\varepsilon_t \varepsilon_t'), \dots, \text{Vech}(\varepsilon_{t-\mu+1} \varepsilon'_{t-\mu+1}))'.$$

(R_t) is therefore unique, due to (3). Moreover, the sequence of random matrices (T_t) and of noises (ζ_t) are also unique, similarly to the CCC case. Now, let us prove the uniqueness of $Y_t := (X_t^{(1)}, X_t^{(2)})$, knowing (η_t) . This would imply the uniqueness of the instantaneous volatility process (D_t) and of the return process (z_t) themselves. With our notations, we have

$$Y_t = \tilde{T}_t Y_{t-1} + \tilde{\zeta}_t, \text{ and } \tilde{Y}_t = \tilde{T}_t \tilde{Y}_{t-1} + \tilde{\zeta}_t,$$

for every t , by setting $\tilde{\zeta}_t = (\zeta_t^{(1)}, \zeta_t^{(2)})$. The arguments are then standard: for instance, see Theorem 2.4's proof in Francq and Zakoian (2010).

We recall the reasoning briefly, to get the uniqueness of (Y_t) . Note that

$$Y_t - \tilde{Y}_t = \tilde{T}_t \tilde{T}_{t-1} \dots \tilde{T}_{t-p} \cdot (Y_{t-p-1} - \tilde{Y}_{t-p-1}),$$

whenever $p > 1$. Since the sequences (Y_t) and (\tilde{Y}_t) are stationary, it is sufficient to prove that $\|\tilde{T}_t \tilde{T}_{t-1} \dots \tilde{T}_{t-p}\|$ tends to zero a.e. when p tends to the infinity, for any matrix norm. This is the case under Assumption U3 because, for every sequence (ε_t) ,

$$E[\ln \|\tilde{T}_t \tilde{T}_{t-1} \dots \tilde{T}_1\|_1] \leq \ln E[\|\tilde{T}_t \tilde{T}_{t-1} \dots \tilde{T}_1\|_1] = \ln \|\tilde{T}_t^*\|^t,$$

by invoking Jensen's inequality, the stationarity of (\tilde{T}_t) , and by noting that all the coefficients of the matrices \tilde{T}_t are nonnegative. It is wellknown that $\lim_{t \rightarrow \infty} t^{-1} \ln(\|A^t\|) = \ln \rho(A)$, for any squared matrix A . Apply this result with $A = \tilde{T}_t^*$. Therefore $\gamma_T = \lim_{t \rightarrow \infty} t^{-1} E[\ln \|\tilde{T}_t \tilde{T}_{t-1} \dots \tilde{T}_1\|_1]$, the top Lyapunov exponent of the sequence of random matrices (\tilde{T}_t) , is strictly negative under Assumption U3. Since the sequence of matrices (\tilde{T}_t) is strictly stationary under U0, we get $\lim_{p \rightarrow +\infty} \|\tilde{T}_t \tilde{T}_{t-1} \dots \tilde{T}_{t-p}\|_1 = 0$ with probability one (Theorem 2.3 in Francq and Zakoian, 2010). This provides the uniqueness of the processes (D_t) and (z_t) , once we assume the uniqueness of the processes (Q_t) and (ε_t) .

Now, let us prove the uniqueness of $(X_t^{(3)}, X_t^{(4)})$ or, in other terms, of (Q_t, ε_t) . This task is clearly more tricky, because we will have to deal with the nonlinear feature of the DCC specification. Here, the convenient matrix norm will be the spectral norm $\|\cdot\|_s$. Consider two stationary solutions (Q_t, ε_t) and $(\tilde{Q}_t, \tilde{\varepsilon}_t)$. Since the spectral norm is submultiplicative, we deduce from (4) that

$$\|Q_t - \tilde{Q}_t\|_s \leq \sum_{k=1}^{\nu} \|M_k\|_s^2 \|Q_{t-k} - \tilde{Q}_{t-k}\|_s$$

$$+ \sum_{l=1}^{\mu} \|N_l\|_s^2 \|\varepsilon_{t-l} \varepsilon'_{t-l} - \tilde{\varepsilon}_{t-l} \tilde{\varepsilon}'_{t-l}\|_s. \tag{E.1}$$

The key point will be to bound from above the terms $\|\varepsilon_{t-l}\varepsilon'_{t-l} - \tilde{\varepsilon}_{t-l}\tilde{\varepsilon}'_{t-l}\|_s$ by a function of $\|Q_{t-l} - \tilde{Q}_{t-l}\|_s$. To lighten the indices, we assume $l = 0$. Clearly, we have

$$\begin{aligned} \|\varepsilon_t\varepsilon'_t - \tilde{\varepsilon}_t\tilde{\varepsilon}'_t\|_s &= \left\| R_t^{1/2}\eta_t\eta'_tR_t^{1/2} - \tilde{R}_t^{1/2}\eta_t\eta'_t\tilde{R}_t^{1/2} \right\|_s \\ &\leq \left\| \left(R_t^{1/2} - \tilde{R}_t^{1/2} \right) \eta_t\eta'_tR_t^{1/2} \right\|_s + \left\| \tilde{R}_t^{1/2}\eta_t\eta'_t \left(R_t^{1/2} - \tilde{R}_t^{1/2} \right) \right\|_s \\ &\leq \left\| R_t^{1/2} - \tilde{R}_t^{1/2} \right\|_s \|\eta_t\eta'_t\|_s \left\| R_t^{1/2} \right\|_s + \left\| \tilde{R}_t^{1/2} \right\|_s \|\eta_t\eta'_t\|_s \left\| R_t^{1/2} - \tilde{R}_t^{1/2} \right\|_s. \end{aligned}$$

Since the rank of $\eta_t\eta'_t$ is one, $\|\eta_t\eta'_t\|_s = Tr(\eta_t\eta'_t) = \|\eta_t\|_2^2$. Moreover,

$$\left\| R_t^{1/2} \right\|_s = \rho(R_t)^{1/2} \leq Tr(R_t)^{1/2} = \sqrt{m}.$$

We deduce

$$\|\varepsilon_t\varepsilon'_t - \tilde{\varepsilon}_t\tilde{\varepsilon}'_t\|_s \leq 2m^{1/2}\|\eta_t\|_2^2 \left\| R_t^{1/2} - \tilde{R}_t^{1/2} \right\|_s. \tag{E.2}$$

Since the spectral norm is unitarily invariant, Theorem 6.2 in Hingham (2008) provides

$$\left\| R_t^{1/2} - \tilde{R}_t^{1/2} \right\|_s \leq \frac{1}{\lambda_1(R_t)^{1/2} + \lambda_1(\tilde{R}_t)^{1/2}} \|R_t - \tilde{R}_t\|_s. \tag{E.3}$$

Note that, for any t ,

$$\begin{aligned} \lambda_1(R_t) &= \min_{\mathbf{x}} \frac{\mathbf{x}'R_t\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \min_{\mathbf{x}} \frac{\mathbf{x}'diag(Q_t)^{-1/2}Q_tdiag(Q_t)^{-1/2}\mathbf{x}}{\mathbf{x}'\mathbf{x}} \\ &\geq \min_{\mathbf{y}} \frac{\mathbf{y}'Q_t\mathbf{y}}{\mathbf{y}'\mathbf{y}} \min_{\mathbf{x}} \frac{\|diag(Q_t)^{-1/2}\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} \\ &\geq \lambda_1(Q_t) \min_i \frac{1}{q_{ii,t}} \geq \frac{C_\lambda}{\|Q_t\|_{\max}}, \end{aligned}$$

invoking Lemma E.2. Since the same inequality applies with $\lambda_1(\tilde{R}_t)$, we get

$$\frac{1}{\lambda_1(R_t)^{1/2} + \lambda_1(\tilde{R}_t)^{1/2}} \leq \frac{\|Q_t\|_{\max}^{1/2} + \|\tilde{Q}_t\|_{\max}^{1/2}}{\sqrt{C_\lambda}}. \tag{E.4}$$

Moreover,

$$\begin{aligned} R_t - \tilde{R}_t &= \left(diag(Q_t)^{-1/2} - diag(\tilde{Q}_t)^{-1/2} \right) Q_tdiag(Q_t)^{-1/2} \\ &\quad + diag(\tilde{Q}_t)^{-1/2} (Q_t - \tilde{Q}_t)diag(Q_t)^{-1/2} \\ &\quad + diag(\tilde{Q}_t)^{-1/2} \tilde{Q}_t \left(diag(Q_t)^{-1/2} - diag(\tilde{Q}_t)^{-1/2} \right) := \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3. \end{aligned}$$

Note that $\mathcal{R}_1 = \left[(q_{ii,t} - \tilde{q}_{ii,t})q_{ij,t}q_{jj,t}^{-1/2}q_{ii,t}^{-1/2} / (q_{ii,t}^{1/2} + \tilde{q}_{ii,t}^{1/2}) \right]_{1 \leq i, j \leq m}$ and $|q_{ij,t}| \leq \sqrt{q_{ii,t}}\sqrt{q_{jj,t}}$ (Cauchy–Schwartz). Since $\|A\|_{\max} \leq \|A\|_s \leq m\|A\|_{\max}$, we get

$$\|\mathcal{R}_1\|_{\max} \leq C_q^{-1} \|diag(q_{ii,t} - \tilde{q}_{ii,t})\|_{\max} \leq C_q^{-1} \|Q_t - \tilde{Q}_t\|_{\max} \leq C_q^{-1} \|Q_t - \tilde{Q}_t\|_s,$$

and $\|\mathcal{R}_1\|_s \leq mC_q^{-1} \|Q_t - \tilde{Q}_t\|_s$. Similarly, $\|\mathcal{R}_3\|_s \leq mC_q^{-1} \|Q_t - \tilde{Q}_t\|_s$. By Lemma E.2, we obtain

$$\begin{aligned} \|\mathcal{R}_2\|_s &= \left\| \text{diag}(\tilde{Q}_t)^{-1/2} (Q_t - \tilde{Q}_t) \text{diag}(Q_t)^{-1/2} \right\|_s \\ &\leq \left\| \text{diag}(\tilde{Q}_t)^{-1/2} \right\|_s \left\| \text{diag}(Q_t)^{-1/2} \right\|_s \|Q_t - \tilde{Q}_t\|_s \\ &\leq \frac{1}{\sqrt{\min_i q_{ii,t}}} \frac{1}{\sqrt{\min_i \tilde{q}_{ii,t}}} \|Q_t - \tilde{Q}_t\|_s \leq \frac{1}{C_q} \|Q_t - \tilde{Q}_t\|_s. \end{aligned}$$

Globally, we get

$$\|R_t - \tilde{R}_t\|_s \leq \frac{2m+1}{C_q} \|Q_t - \tilde{Q}_t\|_s \tag{E.5}$$

everywhere. Recalling (E.2)–(E.5), we deduce

$$\|\varepsilon_t \varepsilon'_t - \tilde{\varepsilon}_t \tilde{\varepsilon}'_t\|_s \leq \frac{2m^{1/2} \|\eta_t\|_2^2}{\sqrt{C_\lambda}} \cdot \frac{2m+1}{C_q} \left(\|Q_t\|_{\max}^{1/2} + \|\tilde{Q}_t\|_{\max}^{1/2} \right) \|Q_t - \tilde{Q}_t\|_s. \tag{E.6}$$

Set $v_t := \|Q_t - \tilde{Q}_t\|_s$. By using the previous inequality and the notation of Lemma E.1, we obtain

$$v_t \leq \sum_{k=1}^{\nu} \|M_k\|_s^2 v_{t-k} + \sum_{l=1}^{\mu} \|N_l\|_s^2 \frac{4m^{1/2}(2m+1)}{\sqrt{C_\lambda} C_q} \|\eta_{t-l}\|_2^2 \sqrt{q_t} v_{t-l} := \sum_{j=1}^{\kappa} \beta_{j,t} v_{t-j}, \tag{E.7}$$

for all t and with our notations.

Setting $\vec{v}_t := [v_t, v_{t-1}, \dots, v_{t-\kappa+1}]'$, we get

$$0 \leq \vec{v}_t \leq N_t^* \vec{v}_{t-1} \leq \dots \leq N_t^* N_{t-1}^* \dots N_{t-p}^* \vec{v}_{t-p-1},$$

for any positive integer p . By the stationarity of the (Q_t) and (\tilde{Q}_t) trajectories, the norm of \vec{v}_t is bounded by a constant that is independent of t . Moreover, under the assumptions U0 and U2, $\|N_t^* N_{t-1}^* \dots N_{t-p}^*\|_s$ tends to zero a.e. when $p \rightarrow +\infty$ and for any fixed t (see Francq and Zakoian 2010, Theorem 2.3). We deduce $v_t \rightarrow 0$ a.e. when $t \rightarrow \infty$, because (v_t) can be initialized arbitrarily far in the past. This implies that $Q_t = \tilde{Q}_t$ a.e. Therefore, $R_t = \tilde{R}_t$ a.e. and $\varepsilon_t = \tilde{\varepsilon}_t$ a.e., knowing (η_t) . This concludes the proof of uniqueness. The ergodicity of the (now unique) DCC solution is a consequence of Corollary 7.17 in Douc et al. (2014). ■

LEMMA E.1. *Under assumptions U0 and U1, for almost every trajectory of a solution (Q_t) of the DCC model, we have*

$$\|Q_t\|_{\max} \leq \frac{\|Vech(W_0)\|_2}{1 - \|T_{33}\|_s} + \sqrt{\frac{m^3(m+1)}{2}} \sum_{l=1}^{\mu} \|\tilde{N}_l\|_s \xi_{t-l} := q_t,$$

where $\xi_t := \sum_{k=0}^{+\infty} \|T_{33}\|_s^k \|\eta_{t-k}\|_2^2$.

If these innovations $|\eta_t|$ are bounded from above by a positive constant C_η a.e., then the latter inequality is simply

$$\|Q_t\|_{\max} \leq \frac{\|Vech(W_0)\|_2 + \sqrt{\frac{m^3(m+1)}{2}} \sum_{l=1}^\mu \|\tilde{N}_l\|_s C_\eta^2}{1 - \|T_{33}\|_s}.$$

Proof of Lemma E.1. For any t , $\|Vech(\varepsilon_t \varepsilon'_t)\|_s = \|Vech(\varepsilon_t \varepsilon'_t)\|_2 \leq \sqrt{m(m+1)} \|\varepsilon_t\|_\infty^2 / \sqrt{2}$. Moreover, since $\|\mathbf{x}\|_s = \|\mathbf{x}\|_2$ for any vector \mathbf{x} and $\|A\|_{\max} \leq \|A\|_s$ for any matrix A (Lütkepohl, 1996, p. 111), we get

$$\begin{aligned} \|\varepsilon_t\|_\infty \leq \|\varepsilon_t\|_s &\leq \|R_t^{1/2} \eta_t\|_s \leq \|R_t^{1/2}\|_s \cdot \|\eta_t\|_s \\ &\leq \|R_t\|_s^{1/2} \|\eta_t\|_2 \leq \sqrt{m} \|\eta_t\|_2. \end{aligned}$$

This proves the inequality $\|Vech(\varepsilon_{t-l} \varepsilon'_{t-l})\|_s \leq \sqrt{m^3(m+1)} \|\eta_{t-l}\|_2^2 / \sqrt{2}$, for every t and l .

With the notations of Subsection 2.2, consider the dynamics of the random vector $X_t^{(3)} := (Vech(Q_t), \dots, Vech(Q_{t-\nu+1}))'$. Clearly, $X_t^{(3)} = T_{33} X_{t-1}^{(3)} + \pi_t$, where

$$\pi_t := Vech(W_0) + \sum_{l=1}^\mu \tilde{N}_l Vech(\varepsilon_{t-l} \varepsilon'_{t-l}).$$

We deduce from U1 that

$$\begin{aligned} \|Q_t\|_{\max} &\leq \|X_t^{(3)}\|_{\max} \leq \|X_t^{(3)}\|_s \leq \sum_{k=0}^{+\infty} \|T_{33}\|_s^k \|\pi_{t-k}\|_s \\ &\leq \sum_{k=0}^{+\infty} \|T_{33}\|_s^k \left\{ \|Vech(W_0)\|_s + \sum_{l=1}^\mu \|\tilde{N}_l\|_s \cdot \|Vech(\varepsilon_{t-k-l} \varepsilon'_{t-l})\|_s \right\} \\ &\leq \frac{\|Vech(W_0)\|_s}{1 - \|T_{33}\|_s} + \sum_{k=0}^{+\infty} \|T_{33}\|_s^k \sum_{l=1}^\mu \|\tilde{N}_l\|_s \cdot \sqrt{\frac{m^3(m+1)}{2}} \|\eta_{t-k-l}\|_2^2 \\ &\leq \frac{\|Vech(W_0)\|_s}{1 - \|T_{33}\|_s} + \sqrt{\frac{m^3(m+1)}{2}} \sum_{l=1}^\mu \|\tilde{N}_l\|_s \xi_{t-l} := q_t. \end{aligned}$$

Since the spectral norm of $Vech(W_0)$ is its Euclidian norm, as for any vector, we obtain the result. ■

LEMMA E.2. *Under Assumption U1, for almost every trajectory of a solution (Q_t) of the DCC model, we have $\lambda_1(Q_t) \geq C_\lambda$ and $\min_{i=1, \dots, m} q_{ii,t} \geq C_q$, where $C_\lambda = \lambda_1(W_0)$ and $C_q := \min_{i=1, \dots, m} (W_0)_{ii}$. In addition, if we assume U4, then $\lambda_1(Q_t) \geq C_\lambda^*$ and $\min_{i=1, \dots, m} q_{ii,t} \geq C_q^*$, with*

$$C_\lambda^* := \frac{\lambda_1(W_0)}{1 - \sum_{k=1}^\nu (m(k))^2} \text{ and } C_q^* := \frac{\min_{i=1, \dots, m} (W_0)_{ii}}{1 - \sum_{k=1}^\nu (m(k))^2}.$$

Proof of Lemma E.2. Is it known that, for any two positive definite matrices A and B , $\lambda_1(A+B) \geq \lambda_1(A) + \lambda_1(B)$ (Weyl's Theorem. See Lütkepohl, 1996, p. 75). In our case, we deduce that $\lambda_1(Q_t) \geq \lambda_1(W_0)$ everywhere, due to Equation (4).

We can improve this lower bound in the particular case of “partially” scalar DCC models. Indeed, in this case, we have

$$\lambda_1(Q_t) \geq \lambda_1(W_0) + \sum_{k=1}^v \lambda_1 \left((m^{(k)})^2 Q_{t-k} \right) \geq \lambda_1(W_0) + \sum_{k=1}^v (m^{(k)})^2 \lambda_1(Q_{t-k}). \tag{E.8}$$

Introduce the random vector $\vec{\lambda}_t := (\lambda_1(Q_t), \dots, \lambda_1(Q_{t-v+1}))'$ and $\vec{\lambda}_W := (\lambda_1(W_0), 0, \dots, 0)'$. Because of (E.8), we have $\vec{\lambda}_t \geq M^* \vec{\lambda}_{t-1} + \vec{\lambda}_W$ for every t . Under Assumption U4, it is easy to check that $\sum_{k=0}^{+\infty} (M^*)^k$ is absolutely convergent and that

$$\vec{\lambda}_t \geq \sum_{k=0}^{+\infty} (M^*)^k \vec{\lambda}_W := \vec{\lambda}_\infty,$$

for every t . Obviously, $M^* \vec{\lambda}_\infty + \vec{\lambda}_W = \vec{\lambda}_\infty$. Due to the definition of M^* , this implies that all the components of $\vec{\lambda}_\infty$ are the same, i.e., a real number λ_∞ exists such that $\vec{\lambda}_\infty = \lambda_\infty e$, $e \in \mathbb{R}^v$. Taking the first component of the vectorial equation $\lambda_\infty M^* e + \vec{\lambda}_W = \lambda_\infty e$ provides $\lambda_\infty \sum_{k=1}^v (m^{(k)})^2 + \lambda_1(W_0) = \lambda_\infty$. This states the lower bound of $\lambda_1(Q_t)$ under U4.

Consider a fixed index $i = 1, \dots, m$. The reasoning for the sequence $(q_{ii,t})_t$ is similar, because

$$q_{ii,t} \geq (W_0)_{ii} + \sum_{k=1}^v (m^{(k)})^2 q_{ii,t-k},$$

for all t , as this inequality is playing the same role as (E.8). This implies the desired result. ■