ON THE STATIONARITY OF DYNAMIC CONDITIONAL CORRELATION MODELS

JEAN-DAVID FERMANIAN
Crest-Ensae*

HASSAN MALONGO
Amundi & Paris Dauphine**

May 21, 2014

Abstract

We provide conditions for the existence and the unicity 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. Moreover, we study the existence of their finite moments.

Key words and phrases: Multivariate dynamic models, conditional correlations, stationarity, DCC.

*3 avenue Pierre Larousse, 92245 Malakoff cedex, France; jean-david.fermanian@ensae.fr
**Univ. Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75116 Paris, France; hassan.malongo@amundi.com
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 written as first-order Markov processes, it is relatively easy to prove the existence of strictly stationary and explicit solutions, even if the latter ones are analytically complex: see classical textbooks, for instance Francq and Zakoïan (2010).

It appeared rapidly that the assumption of constant correlations is too strong. It does not correspond to economic intuition and to many empirical features: see the recent paper of Otranto and Bauwens (2013) and the numerous references therein, for instance. Therefore, Engle (2002) has proposed to extend CCC specifications by adding a particular dynamics on the (conditional) correlation matrices of returns, denoted here by \( R_t \). To insure 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 \( (Q_t) \)-dynamics are specified instead of \( (R_t) \) dynamics directly, contrary to other authors (Tse and Tsui, 2002 or Pelletier, 2004, for instance). This nonlinear transform insures that \( R_t \) is always a correlation matrix, i.e. positive semidefinite with ones on its main diagonal. Nonetheless, it complicates a lot the work of stating stationarity conditions of DCC models. Indeed, analytically tractable solutions of such processes do not exist anymore. This explains why the existence of stationarity solutions of DCC models and their unicity have not been established in the literature yet, nor the finiteness of their moments. Particularly, this implies that theoretically sound statistical inference procedures do not exist yet, as noticed in Caporin
and McAleer (2013).

Despite their theoretical insufficiencies, DCC models have been used intensively among academics and practitioners. Beside numerous applied works, several extensions of the baseline DCC representation have been proposed in the literature: inclusion of asymmetries (Cappiello, Engle, and Sheppard, 2006), of volatility thresholds (Kasch and Caporin, 2013), of macro-variables (Otranto and Bauwens, 2013), of 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), because law of large numbers (potentially uniform) and some CLTs are obtained easily in this case. In the GARCH literature, this essential task has been fulfilled notably by Bougerol and Picard (1992) for univariate GARCH models, by Ling & McAleer (2003) for multivariate ARMA-GARCH models, by Boussama et al. (2011) for BEKK models, notably. In the case of DCC models, a stone is missing because a theory for inference has been proposed by Engle and Sheppard (2001), but their two stage estimation procedure applies under the condition that the underlying DCC process is strictly stationary and ergodic (see their Assumption A.2). The goal of this paper is to fill this gap, focusing on the stationarity problem.

After having introduced 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 unicity. The proofs are gathered 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 \(\text{diag}(M) = [m_{ij}]_{1 \leq i \leq m, 1 \leq j \leq m}\) and the vector \(\text{Vecd}(M) = [m_{ii}]_{1 \leq i \leq m}\) in \(\mathbb{R}^m\).

- If \(n = m\) and \(M\) is symmetric, \(\text{Vech}(M)\) denotes the \(m(m+1)/2 := m^*\) column vector whose components are read from \(M\) column-wise and without redundancy. To be formal, \(\text{Vech}(M) = [\tilde{m}_k]_{1 \leq k \leq m^*}\), where \(\tilde{m}_k = m_{ij}\) for the unique couple of indices \((i, j)\) in \(\{1, \ldots, m\}^2\), \(i \geq j\) such that \([m+(m-1)+\ldots+(m-j+2)]^+ + (i-j+1) = k\). This defines a one-to-one mapping \(\phi\) between the indices \(k \in \{1, \ldots, m^*\}\) and the pairs \((i, j)\), \(i \geq j\), \(1 \leq i, j \leq m\), i.e. \((i, j) = (\phi_1(k), \phi_2(k)) = \phi(k)\).

- \(\rho(M)\) 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 its smallest eigenvalue is denoted by \(\lambda_1(M)\).

- \(\otimes\) denotes the usual Kronecker product, and \(M^\otimes p = M \otimes \ldots \otimes M\) (\(p\) times). \(\odot\) denotes the element-by-element product. If \(v\) is a vector in \(\mathbb{R}^n\), then \(v \odot M = [v m_{ij}]_{1 \leq i \leq n, 1 \leq j \leq m}\).

- We will consider several matrix norms, particularly \(\|M\|_\infty = \max_{1 \leq i \leq n, 1 \leq j \leq m} |m_{ij}|\), and the spectral norm defined for any squared matrix by

  \[\|M\|_s = \sup \{ \sqrt{\lambda} \mid \lambda \text{ is an eigenvalue of } M'M \} = \sup_x \frac{\|Mx\|_2}{\|x\|_2} .\]

  More generally, for any norm \(N\) for vectors, we can define a norm \(\|\cdot\|_N\) for matrices by setting \(\|M\|_N = \sup x N(Mx)/N(x)\).

- For any column vector \(z_t \in \mathbb{R}^m\), we denote \(z_t = (z_{1,t}, \ldots, z_{m,t})'\) and \(\bar{z}_t := (z_{1,t}^2, \ldots, z_{m,t}^2)'\).
• $e$ denotes a vector of ones, whose dimension 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 $x \in A$, than $\sup_{x \in A} M(x)$ is the matrix $[\sup_{x \in A} m_{ij}(x)]$.

2 Dynamic Conditional Correlation models

2.1 The classical DCC specification

Let us recall 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 (but including) time $t - 1$ is denoted by $\mathcal{I}_{t-1}$.

Modeling the expected returns of the 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 $y_t$. 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 w.r.t. $\sigma(z_t, z_{t-1}, \ldots)$. Therefore, $\mathcal{I}_t = \sigma(y_t, y_{t-1}, \ldots) = \sigma(z_t, z_{t-1}, \ldots)$.

Assume we can write $z_t = H_t^{1/2} e_t$, where

• $(e_t)_{t \in \mathbb{Z}}$ is a standard white noise in $\mathbb{R}^m$: $E[e_t] = 0$, $Var(e_t) = I_m$ and the vectors $e_t$ are mutually independent.

• $H_t = [h_{ij,t}]_{1 \leq i,j \leq m}$ is 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 $R_t$):

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

where $h_{k,t} := h_{kk,t}$ denotes the “instantaneous variance” of the return $y_{k,t}$ (or $z_{k,t}$, equivalently), conditionally on $I_{t-1}$. In the literature, this conditional volatility of the asset $k$ knowing $I_{t-1}$ is often denoted by $\sigma_{kk,t}$ instead of $h_{k,t}^{1/2}$.

We assume GARCH-type models on every margin, but with cross-effects between all these volatilities potentially:

$$\text{Vecd}(D_t) = V_0 + \sum_{i=1}^{r} A_i \text{Vecd}(D_{t-i}) + \sum_{j=1}^{s} B_j \cdot \hat{z}_t - j,$$

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

Let us introduce the so-called “standardized residuals” $\tilde{\varepsilon}_t := D_t^{-1/2} \cdot \varepsilon_t$. The dynamics of correlations are given by the traditional Dynamic Conditional Correlation specification:

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

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

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

for some deterministic matrices $(M_k)_{k=1,...,\nu}$ and $(N_l)_{l=1,...,\mu}$, and for a positive definite constant matrix $W_0$. Note that $Q_t$ will be definite positive too. We will set $M_k := [m_{p,q}^{(k)}]_{1 \leq p,q \leq m}$, $k = 1,...,\nu$, and $N_l := [n_{p,q}^{(l)}]_{1 \leq p,q \leq m}$, $l = 1,...,\mu$. In practice, the positive matrix $W_0$ (or the constant vector $\text{Vech}(W_0)$ in $\mathbb{R}^m$).
equivalently) is a parameter to be estimated, often in a first stage.

Since \( E_{t-1}[\varepsilon_t \varepsilon_t'] = R_t \), there exists a sequence of independent random vectors \((\eta_t)_{t \in \mathbb{Z}} \) in \( \mathbb{R}^m \) such that

\[
\varepsilon_t = R_t^{1/2} \eta_t,
\]

with \( E_{t-1}[\eta_t] = 0 \) and \( E_{t-1}[\eta_t \eta_t'] = I_m \). It can be imposed that \( R_t^{1/2} \) is positive definite. 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.

Note that the processes \((D_t)\), \((Q_t)\) and then \((R_t)\) are \( \mathcal{I}_{t-1} \) measurable, if they exist. Moreover, the definition of the 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.

Aielli (2013) has 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 \( \varepsilon_t \). Therefore, he introduced a new variety of DCC-GARCH models (called cDCC), where (4) is replaced by

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

(6)

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 (6) is a smart but not intuitive and ”ad-hoc” specification. Its main justification appears as essentially technical, to avoid the nonlinear feature of the original Engle’s DCC model (4). Under the latter usual specification, DCC models cannot be rewritten as BEKK models anymore and other techniques have to be found. In this paper, we obtain the same type of results as Aielli (2013), but by keeping the original specification of DCC models and without relying on another encom-
passing 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 rewriting will become a crucial tool to study of stationary solutions hereafter. Set

\[ X_t := (X_t^{(1)}, X_t^{(2)}, X_t^{(3)}, X_t^{(4)})', \]  

where

\[ X_t^{(1)} := (\text{vec}(D_t), \ldots, \text{vec}(D_{t-r+1}))', \]

\[ X_t^{(2)} := (\tilde{z}_t, \ldots, \tilde{z}_{t-s+1})', \]

\[ X_t^{(3)} := (\text{vech}(Q_t), \ldots, \text{vech}(Q_{t-\nu+1}))', \] and

\[ X_t^{(4)} := (\text{vech}(\varepsilon_t\varepsilon_t'), \ldots, \text{vech}(\varepsilon_{t-\mu+1}\varepsilon_{t-\mu+1}'))'. \]

The dimensions of the four previous random vectors are \( rm, sm, \nu m^* \) and \( \mu m^* \) respectively. Their sum, the dimension of \( X_t \), is denoted by \( d \). With simple block matrix calculations, there exist random matrices \( (T_t) \) and a vector process \( (\zeta_t) \) such that the dynamics of \( X_t \), any solution of the DCC model, may be rewritten as

\[ X_t = T_t X_{t-1} + \zeta_t, \]  

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 (8), the underlying process \( (X_t) \) can be seen as a vectorial autoregressive of order one, but with random matrix-coefficients \( (T_t) \). Actually, \( T_t \) and \( \zeta_t \) will be stochastic only through \( \varepsilon_t \), i.e. through the \( t \)-innovation \( \eta_t \) and the \( \mathcal{I}_{t-1} \)-measurable matrix \( R_t \). This creates a major difficulty to prove 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
\( X_t = f(\eta_t, \eta_{t-1}, \ldots) \) for some deterministic measurable function \( f \), because the link between \( T_t \) and the past innovations (or observations) is highly nonlinear.

Let us detail the AR(1) form of (8):

- set \( T_{1k,t} = 0 \) when \( k = 3, 4 \),

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

- We deduce from Equation (2) that

\[
D_t \tilde{\varepsilon}_t = \tilde{\varepsilon}_t \odot \text{vecd}(D_t) = \tilde{z}_t = \tilde{\varepsilon}_t \odot V_0 + \sum_{i=1}^r \tilde{\varepsilon}_t \odot A_i \text{vecd}(D_{t-i}) + \sum_{j=1}^s \tilde{\varepsilon}_t \odot B_j \tilde{z}_{t-j}. \tag{9}
\]

Let us set \( T_{23,t} = T_{24,t} = 0 \),

\[
T_{21,t} := \begin{bmatrix} \tilde{\varepsilon}_t \odot A_1 & \tilde{\varepsilon}_t \odot A_2 & \cdots & \cdots & \tilde{\varepsilon}_t \odot A_r \\ 0_m & \cdots & \cdots & \cdots & 0_m \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_m & \cdots & \cdots & \cdots & 0_m \end{bmatrix}, \text{ and }
T_{22,t} := \begin{bmatrix} \tilde{\varepsilon}_t \odot B_1 & \tilde{\varepsilon}_t \odot B_2 & \cdots & \cdots & \tilde{\varepsilon}_t \odot B_s \\ I_m & 0_m & \cdots & \cdots & 0_m \\ 0_m & I_m & 0_m & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_m & \cdots & 0_m & I_m & 0_m \end{bmatrix}
\]
Clearly, there exist matrices $\tilde{M}_k$, $k = 1, \ldots, \nu$, such that

$$Vech(M_kQ_{t-k}M_k') = \tilde{M}_kVech(Q_{t-k}).$$

Similarly, there exists matrices $\tilde{N}_l$, $l = 1, \ldots, \mu$, such that

$$Vech(N_l\varepsilon_{t-l}\varepsilon_{t-l}'N_l') = \tilde{N}_lVech(\varepsilon_{t-l}\varepsilon_{t-l}').$$

It is possible to write explicitly the previous matrices $\tilde{M}_k$ and $\tilde{N}_l$. Indeed, with the notations of Subsection 1.2, $\tilde{M}_k = [\tilde{m}_{\phi_1(u),\phi_2(v)}(k)_{u,v}]_{1 \leq u,v \leq m}$ where

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

Then, set $T_{31,t} = T_{32,t} = 0$,

$$T_{33,t} := \begin{bmatrix}
M_1 & M_2 & \cdots & \cdots & M_{\nu} \\
I_{m^*} & 0_{m^*} & \cdots & \cdots & 0_{m^*} \\
0_{m^*} & I_{m^*} & 0_{m^*} & \cdots & \cdots \\
\vdots & \vdots & \ddots & \cdots & \cdots \\
0_{m^*} & \cdots & 0_{m^*} & I_{m^*} & 0_{m^*}
\end{bmatrix},$$

and $T_{34,t} := \begin{bmatrix}
N_1 & N_2 & \cdots & \cdots & N_{\mu} \\
0_{m^*} & \cdots & \cdots & \cdots & 0_{m^*} \\
\vdots & \vdots & \cdots & \cdots & \cdots \\
0_{m^*} & \cdots & \cdots & \cdots & 0_{m^*}
\end{bmatrix}$.

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

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

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

$$\zeta_t^{(1)} = (V_0, 0_m, \ldots, 0_m)'$$,
$$\zeta_t^{(2)} = (\tilde{\varepsilon}_t \odot V_0, 0_m, \ldots, 0_m)'$$,
$$\zeta_t^{(3)} = (\text{Vech}(W_0), 0_{m^*}, \ldots, 0_{m^*})'$$, and
$$\zeta_t^{(4)} = (\text{Vech}(\varepsilon_t \varepsilon_t'), 0_{m^*}, \ldots, 0_{m^*})'$$.

Intuitively, the model \((X_t)\) is \(I\)-Markovian because it is the case for the process \((\zeta_t)\) and \((T_t)\) themselves. Indeed, \(\varepsilon_t\) (or \(\tilde{\varepsilon}_t\), or even \(\text{Vech}(\varepsilon_t \varepsilon_t')\)) is a function of the couple \((R_t, \eta_t)\) only. Due to (3) and (4), \(R_t\) is a deterministic function of \(X_{t-1}\). Since \(\eta_t\) is independent of \(I_{t-1}\), the law of \(\varepsilon_t\) knowing \(I_{t-1}\) is just the law of \(\varepsilon_t\) knowing \(X_{t-1}\). The same assertion applies with \(T_t\), \(\zeta_t\), or with \(X_t\) itself, instead of \(\varepsilon_t\).

In other words, the non-linearity of the DCC model is coming mainly from \(\tilde{\varepsilon}_t\) in \(T_t\). But there exist constant matrices (of zeros and ones) \(F\) and \(G\) such that (8) can be rewritten

$$X_t = (\tilde{\varepsilon}_t \otimes F) \odot T_o X_{t-1} + (\text{Vech}(\varepsilon_t \varepsilon_t') \cdot G) \odot \zeta_o,$$

where \(T_o\) (resp. \(\zeta_o\)) is the \(T_t\) matrix (resp. \(\zeta_t\) vector) when \(\varepsilon_t = 1\). 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 \(\eta_t\) only. These arguments prove the Markovian structure of the \((X_t)\) process.

### 3 Stationarity of DCC models

#### 3.1 Existence of stationary DCC solutions

To obtain the existence of stationary solutions of 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 (8). 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 \(T_t\), that depends on the random variables \(\varepsilon_{kt}, k = 1, \ldots, m\). The latter variables have a variance one, but they are not independent. This is in contrast with Ling and McAleer (2003). Moreover, unfortunately, the joint law of \(\varepsilon_t\) is a function of \(R_t\), i.e. a function of \(X_{t-1}\). That is why we need the following condition.

**Assumption E1:** For some \(p \geq 1\), \(E[\|\eta_t\|^{2p}] < \infty\) and \(\rho(T^*) < 1\), where

\[
T^* := \sup_{x \in \mathbb{R}^d} E[\|T_t^p\| | X_{t-1} = x].
\]

Recall that \(T_t\) depends on \(\varepsilon_t\), that \(\varepsilon_t = R_t^{1/2} \eta_t\), and that the components of \(\eta_t\) are uncorrelated. Then, all the coefficients of \(T^*\) are finite because all 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), stated for Vector ARMA-GARCH models.

**Theorem 1** Under Assumption E1, 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 w.r.t. the \(\sigma\)-field \(\mathcal{I}\) induced by the observations. Moreover, the solution \((z_t)\) is second-order stationary and the \(2p\)-th moments of \(z_t\) are finite.

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

\[
A_i = diag(a_1^{(i)}, \ldots, a_m^{(i)}), \ i = 1, \ldots, r, \ B_j = diag(b_1^{(j)}, \ldots, b_m^{(j)}), \ j = 1, \ldots, s,
\]

\[
M_k = diag(m_1^{(k)}, \ldots, m_m^{(k)}), \ k = 1, \ldots, \nu, \ N_l = diag(n_1^{(l)}, \ldots, n_m^{(l)}), \ l = 1, \ldots, \mu.
\]
The associated matrices $\tilde{M}_k$ and $\tilde{N}_l$ are diagonal too. Set $\tilde{M}_k = \text{diag}(\tilde{m}_l^{(k)})_{1 \leq l \leq m^*}$, 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 $E[\varepsilon_{kt} | X_{t-1} = x] = 1$ for every index $k$, $T^*$ is simply $|T_k|$, replacing $\varepsilon_t$ by one. Denote by $P^*$ the characteristic polynomial of $T^*$, i.e. $P^*(\lambda) = \text{Ker}(T^* - \lambda \text{Id})$. It can be seen easily that there exist two polynomials $P^*_1$ and $P^*_2$ s.t. $P^*(\lambda) = P^*_1(\lambda)P^*_2(\lambda)$. Here, $P^*_1$ denotes the characteristic polynomial of the block-matrix $||T_{ij,t}||_{1 \leq i,j \leq 2}$, replacing $\varepsilon_t$ by one. And $P^*_2$ is the characteristic polynomial of the previous matrix $|T_{33,t}|$. Tedious (but not so difficult) algebraic calculations provide

$$P^*_1(\lambda) = \pm \lambda^{\pi_1} \prod_{k=1}^m \left( \sum_{i=1}^r a_k^{(i)} \lambda^{r+s-i} + \sum_{j=1}^s b_k^{(j)} \lambda^{r+s-j} - \lambda^{r+s} \right),$$

$$P^*_2(\lambda) = \pm \lambda^{\pi_2} \prod_{l=1}^{m^*} \left( \sum_{k=1}^{\nu} \tilde{m}_l^{(k)} \lambda^{\nu-k} - \lambda^{\nu} \right),$$

for some integers $\pi_1$ and $\pi_2$. Let $\lambda_0$ be a non-zero root of $P^*$, If $\lambda_0$ is a root of $P^*_1$ then there exists an index $k \in \{1, \ldots, m\}$ 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 $\lambda_0$ is a root of $P^*_2$ and if $|\lambda_0| \geq 1$, then there exists $l \in \{1, \ldots, m^*\}$ s.t. $1 \leq \sum_{k=1}^\nu |\tilde{m}_l^{(k)}|$. In other words, a sufficient condition to fulfill Assumption E1 is

$$\sup_{k=1,\ldots,m} \sum_{i=1}^r a_k^{(i)} + \sum_{j=1}^s b_k^{(j)} < 1, \quad \text{and} \quad \sup_{l=1,\ldots,m^*} \sum_{k=1}^\nu |\tilde{m}_l^{(k)}| < 1. \quad (11)$$

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, if only the second-order moments of $z_t$ are of interest ($p = 1$). These conditions insure that
E1 will be satisfied. In other words, they 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 purpose (the optimization stage to get QML estimates, e.g.).

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

**Theorem 2** If
\[
\sum_{i=1}^{r} \| A_i \|_{\mathcal{N}} + \sum_{j=1}^{s} \| B_j \|_{\mathcal{N}} < 1, \quad \text{and}
\]
\[
\sum_{k=1}^{\nu} \| \tilde{M}_k \|_{\mathcal{N}^\ast} < 1,
\]
then Assumption E1 is satisfied with $p = 1$, and Theorem 1 applies.

Note that the conditions of Theorem 2 do not depend on the sequence of matrices $(N_l), l = 1, \ldots, \mu$.

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

It is often fruitful to assume that the Markov chain is initialized at $t = 0$ by drawing $X_0$ in the stationary law. Introducing the filtration $\mathcal{I}_t := \sigma(X_0, z_1, \ldots, z_t)$, we check easily that the DCC solution is now measurable w.r.t. the $\sigma$-field induced by the innovations and the initial value, because, for all $t > 0$,
\[
\sigma(X_0, z_1, \ldots, z_t) = \sigma(X_0, \varepsilon_1, \ldots, \varepsilon_t) = \sigma(X_0, \eta_1, \ldots, \eta_t).
\]

**Example 1 (Continued):** Consider a diagonal-type DCC model and max-
minimum norms for vectors. In this case, the condition (12) becomes

\[ \sum_{i=1}^{r} \sup_{l=1,\ldots,m} a_{l}^{(i)} + \sum_{j=1}^{s} \sup_{l=1,\ldots,m} b_{l}^{(j)} < 1, \]

and the condition (13) is \( \sum_{k=1}^{\nu} \sup_{l=1,\ldots,m} |\tilde{m}_{l}^{(k)}| < 1 \). These two conditions are stronger than (11), as expected.

**Example 2:** To reduce even more the number of free parameters, scalar-DCC models are often assumed. In this case, all the unknown matrices are simply the product of a scalar and an identity matrix:

- \( A_{i} = a^{(i)} I_{m}, i = 1,\ldots,r \)
- \( B_{j} = b^{(j)} I_{m}, j = 1,\ldots,s \)
- \( M_{k} = m^{(k)} I_{m}, k = 1,\ldots,\nu \)
- \( N_{l} = n^{(l)} I_{m}, l = 1,\ldots,\mu \)

Such models are very popular, because they allow a dramatic reduction of the number of free parameters. With obvious notations, the conditions of Theorem 1 and 2 are the same:

\[ \sum_{i=1}^{r} a^{(i)} + \sum_{j=1}^{s} b^{(j)} < 1, \quad \text{and} \quad \sum_{k=1}^{\nu} |m^{(k)}|^{2} < 1, \]

(14)

In passing, we recover the usual condition of stationarity of GARCH-type models:

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

### 3.2 Unicity of stationary DCC solutions

Even if there exist stationary solutions of the DCC model, we are not insured *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 unicity of invariant measures
is important for inference purpose. Indeed, the estimation of DCC models is based on M-estimates typically (Quasi Maximum Likelihood, for instance). These techniques rely strongly on uniform Law of Large Numbers, obtained through the ergodicity of the model trajectories most often. The conditions for identifiability and consistency rely on some expectations w.r.t. the underlying invariant measure of the given stationary process. If there are several possible measures, for a given set of parameters, it becomes difficult to check such conditions. Finally, with several underlying invariant measures, we cannot exclude the possibility of switches from a stationary trajectory to another one, disturbing the econometric analysis (stationarity tests, statistical uncertainty around estimates, etc).

Unfortunately, this unicity is not given “for free” by Tweedie’s Lemma 6. Moreover, the usual arguments concerning the unicity of stationary GARCH-type solutions do not apply here. Indeed, under the Markov-chain form given by Equation (8), the matrix $T_t$ is itself a function of the random vector $X_t$ through the $\varepsilon_t$ factors. This is a major difference with the CCC case, notably. That is why we need to find another strategy. Now, we provide some unicity 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 this is not mandatory a priori.

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.

Assumption U2: The underlying DCC model is “partially” scalar, i.e. there exist scalars $m^{(k)}$ such that $M_k = m^{(k)} I_n$ for all $k = 1, \ldots, \nu$. Moreover,
\( \rho(M^*) < 1 \) by setting

\[
M^* := \begin{bmatrix}
(m^{(1)})^2 & (m^{(2)})^2 & \cdots & (m^{(\nu)})^2 \\
1 & 0 & \cdots & 0 \\
0 & 1 & 0 & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 1 & 0
\end{bmatrix}.
\]

Actually, U2 will not be mandatory to get our unicity result, even if allows a weakening of the other technical conditions. In every case, this “partially” scalar case is in line with the common practice of scalar DCC (or scalar multivariate GARCH) models.

Thanks to the latter assumptions, we will be able to bound \( \|Q_t\|_\infty \) 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 to prove the unicity of stationary DCC solutions.

**Lemma 3** Under Assumption U1, for almost every trajectory of a solution \((Q_t)\) of the DCC model, we have

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

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

The sequences \((\xi_t)\) and \((q_t)\) are stationary and ergodic because any \( \xi_t \) or \( q_t \) is a measurable function of the innovations \((\eta_t)\) that are i.i.d.

If these innovations \(|\eta_t|\) are bounded from above by a positive constant \( C_\eta \) a.e., then the latter inequality is simply

\[
\|Q_t\|_\infty \leq \frac{\|\text{Vech}(W_0)\|_s + \sqrt{\frac{m^3(m+1)}{2}} \sum_{l=1}^{\mu} \|\tilde{N}_l\|_s C_\eta^2}{1 - \|T_{33}\|_s}.
\]
Lemma 4 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 \(\min_{i=1,\ldots,m} q_{ii,t} \geq C_q\), where \(C_\lambda = \lambda_1(W_0)\) and \(C_q := \min_{i=1,\ldots,m}(W_0)_{ii}\). In addition, if we assume U2, then we can set

\[
C_\lambda = \frac{\lambda_1(W_0)}{1 - \sum_{k=1}^p (m(k))^2} \quad \text{and} \quad C_q = \frac{\min_{i=1,\ldots,m}(W_0)_{ii}}{1 - \sum_{k=1}^p (m(k))^2}.
\]

The proofs of these lemmas are postponed to the end of the appendix.

Let \(\kappa = \max(\nu, \mu)\) and, for every \(j = 1, \ldots, \kappa\), set

\[
\beta_{j,t} := 1(j \leq \nu)\|M_j\|^2_s + 1(j \leq \mu)\|N_j\|^2_s 4(2m + 1) m^{1/2} \sqrt{C_\lambda C_q} \|\eta_t\|^2_s \sqrt{q_t}.
\]

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

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

Clearly, the sequence of matrices \((N_t^*)\) is stationary and ergodic because any \(N_t^*\) is a measurable function of the innovations \((\eta_t)\).

Assumption U3: \(E[\ln^+ \|N_t^*\|] < \infty\) and the top Lyapunov exponent of the sequence \((N_t^*)\), defined by \(\gamma_N := \lim_{t \to +\infty} t^{-1} E[\ln(\|N_1^* N_2^* \cdots N_t^*\|)]\), is strictly negative.

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

Actually, the technical assumptions U1-U3 above will insure the unicity of \((\varepsilon_t)\), \((Q_t)\) and \((R_t)\) only. To get the unicity 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 $(\varepsilon_t)$ nor $t$, because $E[\varepsilon_{kt}^2] = 1$ for every $k$.

**Assumption U4**: $\rho(\bar{T}^*) < 1$.

**Theorem 5** Under the assumptions of Lemmas 3 and 4, and under U3-U4, a strictly stationary solution of the DCC model is unique, given the sequence $(\eta_t)$.

**Example 2 (Continued)**: In the case of scalar DCC models of order one, it is easy to specify the conditions above. Here, $r = s = \nu = \mu = 1$,

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

Assumptions U1 and U2 are equivalent and mean $|m^{(1)}| < 1$. Assumption U3 is fulfilled if $E[\ln \| N^*_1 \|_\infty] < 0$, or if

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

This expectation could be evaluated by simulation easily, by noting that $\eta_t$ 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 checked that the characteristic function of $\bar{T}^*$ is the function $x \mapsto (x)^m(a^{(1)} + b^{(1)} - x)^m$. Then Assumption U4 means $a^{(1)} + b^{(1)} < 1$. Therefore, as expected, the sufficient conditions for the unicity of stationary DCC solutions are more demanding than for their existence, due to U3. 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 $\eta_t$ are not "too heavy".

**Acknowledgments**

We thank C. Francq and J.-M. Zakoïan for fruitful remarks and discussions. Moreover, we are grateful to the co-Editor and two anonymous referee, who have proposed a lot of ways of improving the article. Finally, the authors thank the Labex "Ecodec" for its support.

**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, e.g.), 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 $(X_t)_{t=1,2,...}$ be a temporally homogeneous Markov chain with a locally compact completely 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 $A \in \mathcal{B}$. Theorem 2 of Tweedie (1988) provides:

**Lemma 6** Suppose that $(X_t)$ is a Feller chain, i.e. for each bounded continuous function $h$ on $S$, the function of $x$ given by $E[h(X_{t-1}) | X_{t-1} = x]$ is also continuous.

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

$$\int_{A^c} P(x,dy)g(y) \leq g(x) - \varepsilon, \quad x \in A^c,$$  \hspace{1cm} (15)
then there exists a σ–finite invariant measure \( \mu \) for \( P \) with \( 0 < \mu(A) < \infty \).

2. Furthermore, if
\[
\int_A \mu(dx) \left[ \int_{A^c} P(x, dy) g(y) \right] < \infty,
\]
then \( \mu \) is finite and hence \( \pi = \mu/\mu(S) \) is an invariant probability measure.

3. Furthermore, if
\[
\int_{A^c} P(x, dy) g(y) \leq g(x) - f(x), \quad x \in A^c,
\]
then \( \mu \) 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 7** For a given squared matrix \( T \), if \( \rho(\|T\|) < 1 \), then there exists a vector \( M > 0 \) such that \((Id - |T|') M > 0\).

### B Proof of Theorem 1:

First, let us check that \((X_t)\) is a Feller chain. Let \( h \) be a bounded and continuous function on \( \mathbb{R}^d \). Clearly,
\[
E[h(X_t) \mid X_{t-1} = x] = E[h(T_t x + \zeta_t) \mid X_{t-1} = x]
= E[h(\psi_1(\varepsilon_t \varepsilon_t') x + \psi_2(\varepsilon_t \varepsilon_t')) \mid X_{t-1} = x],
\]
for some continuous transforms \( \psi_1 \) and \( \psi_2 \). 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 Proposition 6.3 in Serre (2010), e.g.) and invoke \( X_{t-1} \mapsto R_t \) is continuous too by construction. Then,
\[
E[h(X_t) \mid X_{t-1} = x] = E[h \circ \tilde{\psi}(x, \eta_t) \mid X_{t-1} = x] = \int h \circ \tilde{\psi}(x, \eta) dP_\eta(\eta),
\]
for some continuous transform $\tilde{\psi}$. Now, consider a sequence of vectors $(x_n)$ that tends to $x$ when $n \to \infty$. Since $h$ is bounded and since the sequence $(h \circ \tilde{\psi}(x_n, \eta))_n$ is convergent for every $\eta$, we can apply the dominated convergence theorem. We deduce that $x \mapsto E[h(X_t) \mid X_{t-1} = x]$ is continuous and $(X_t)$ is Feller.

Second, set $g(x) = 1 + |x^{\otimes p}|'M$, for an arbitrary positive vector $M$, that will be chosen after. Clearly,

$$E[g(X_t) \mid X_{t-1} = x] = 1 + E \left[ |(T_t x + \zeta_t)^{\otimes p}|' \mid X_{t-1} = x \right] M.$$ 

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

$$\|R(x)\| \leq \alpha_0 \left( \|\zeta_t\| \cdot \|(T_t x)^{(p-1)}\| + \ldots + \|\zeta_t\|^{p-1} \cdot \|(T_t x)\| + \|\zeta_t\|^p \right),$$

for some positive constant $\alpha_0$ and any multiplicative matrix norm $\|\cdot\|$.

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

$$E[||(T_t x)^{\otimes p}| \mid X_{t-1} = x]'M \leq |x^{\otimes p}|' \left( \sup_{x \in \mathbb{R}^d} E[|(T_t x)^{\otimes p}|' \mid X_{t-1} = x] \right) M$$

$$\leq |x^{\otimes p}|' \left( \sup_{x \in \mathbb{R}^d} E[|(T_t x)^{\otimes p}|' \mid X_{t-1} = x] \right) M$$

Now, choose $M$ as provided by Lemma 7, 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 in an orthonormal basis and invoke Cauchy-Schwartz inequality). This implies there exist constants $\alpha_k$ such that $\|\text{Vech}(\varepsilon_\ell \varepsilon_\ell')^{\otimes k}\| \leq \alpha_k \|\text{Vech}(\eta_\ell \eta_\ell')^{\otimes k}\|$ when $k \leq$
Since $E[\|\eta\|^{2p}] < \infty$ by assumption, there exist some constants $c_{k,l}$ such that $E_{t-1}[\|\zeta_t\|^k \cdot \|\varepsilon_t\|^l] < c_{k,l}$ for any couple $(k,l)$, $k + l \leq p$. We deduce the boundedness of $T^k_t$, $k \leq p$, and

$$E[\|R(x)\| | X_{t-1} = x] \leq \alpha_1 \left(\|x^{(p-1)}\| + \ldots + \|x\| + 1\right),$$

for some positive constant $\alpha_1$. We have obtained

$$E[g(X_t) | X_{t-1} = x] \leq 1 + \|x^{(p-1)}(T^*)' M + O \left(\sum_{k=0}^{p-1} \|x^{(k)}\|\right) \leq g(x) - \|x^{(p-1)}(Id - (T^*)') M + O \left(\sum_{k=0}^{p-1} \|x^{(k)}\|\right).$$

(18)

By Lemma 7, $(Id - (T^*)') M$ is strictly positive. Then, there exists a positive constant $c_0$ such that

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

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

$$\sum_{i_1, \ldots, i_k} |x_{i_1} \cdot \ldots \cdot x_{i_k}| = \left(\sum_{i=1}^{d} |x_i|\right)^k \leq \left(\sum_{i=1}^{d} |x_i|^p\right)^{k/p} d^k,$$

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

- $g(x) \leq 1 + \|M\| \sum_{i_1, \ldots, i_p} |x_{i_1} \cdot \ldots \cdot x_{i_p}| \leq 1 + c_2 N(x)$, and

- every “residual” term $\|x^{(k)}\|$ is bounded above by (a scalar times) $N(x)^{k/p}$, when $k < p$. 

23
Therefore, this provides
\[
E[g(X_t) \mid X_{t-1} = \mathbf{x}] \leq g(\mathbf{x}) \left[ 1 - c_0 \frac{N(\mathbf{x})}{g(\mathbf{x})} + O \left( \sup_{k=0, \ldots, p-1} \frac{N(\mathbf{x})^{k/p}}{g(\mathbf{x})} \right) \right]
\]
\[
\leq g(\mathbf{x}) \left[ 1 - \frac{c_0 N(\mathbf{x})}{1 + c_2 N(\mathbf{x})} + O \left( \sup_{k=0, \ldots, p-1} \frac{N(\mathbf{x})^{k/p}}{c_1 N(\mathbf{x})} \right) \right].
\]

Let us define the set \( A := \{ \mathbf{x} \in \mathbb{R}^d \mid N(\mathbf{x}) \leq \Delta \} \), for some \( \Delta > 1 \). When \( \Delta \) is sufficiently large, we obtain, for any \( \mathbf{x} \notin A \),
\[
0 \leq E[g(X_t) \mid X_{t-1} = \mathbf{x}] \leq g(\mathbf{x}) \left[ 1 - \frac{c_0}{2c_2} + O \left( \frac{\Delta^{-1/p}}{c_1} \right) \right] < g(\mathbf{x}) \left[ 1 - \frac{c_0}{3c_2} \right].
\]

(19)

Since \( g(\mathbf{x}) \geq 1 \), it follows that \( E[g(X_t) \mid X_{t-1} = \mathbf{x}] \leq g(\mathbf{x}) - \varepsilon \) for some \( \varepsilon > 0 \). This proves Equation (15) in Lemma 6. Therefore, there exists a \( \sigma \)-finite invariant measure \( \mu \) for the Markov chain \( (X_t) \), and \( 0 < \mu(A) < \infty \).

For any \( \mathbf{x} \in A \), Equation (18) provides
\[
E[g(X_t) \mid X_{t-1} = \mathbf{x}] \leq g(\mathbf{x}) + O \left( \sum_{k=0}^{p-1} \| x^{\otimes k} \| \right) \leq C \Delta^p
\]
for some constant \( C \) that does not depend on \( \mathbf{x} \). Then,
\[
\int_A \mu(dx) \left[ \int_A P(x, dy) g(y) \right] \leq \int_A \mu(dx) E[g(X_t) \mid X_{t-1} = \mathbf{x}] \leq C \Delta^p \mu(A) < \infty.
\]

We deduce that \( \mu \) is finite and hence \( \pi = \mu / \mu(\mathbb{R}^d) \) is an invariant probability measure of \( (X_t) \). This implies there exists a strictly stationary solution satisfying (8), still denoted by \( X_t \).

Third, by invoking Equation (19), we get (17) in Lemma 6 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
\[
E_\pi [N(X_t)] < \infty.
\]

(20)

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

**Remark 8** Equation (20) provides a lot more than only the finiteness of $z$’s moments. Globally, this means that

$$E_\pi \left[ \sum_{i=1}^{m} h_{it}^p \right] < \infty, \quad E_\pi \left[ \sum_{i=1}^{m} z_{it}^{2p} \right] < \infty,$$

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

### Proof of Theorem 2:

Let us consider $\lambda$, a non zero eigenvalue of $T^*$, when $p = 1$. We check easily that this matrix is simply $T_t$, but replacing $\vec{\varepsilon}_t$ by one, and the coefficients of the matrices $\tilde{M}_k$ and $\tilde{N}_l$ by their absolute values (remind that the matrices $A_i$ and $B_j$ are already nonnegative). Let $v = (v^{(1)}, v^{(2)}, v^{(3)}, v^{(4)})$ be the associated eigenvector, where the dimensions of the subvectors $v^{(k)}$, $k = 1, \ldots, 4$ are consistent with those of $X_t$ in (7). We can split even more the latter subvectors, to be conformable with the matrices $A_i$, $B_j$, $\tilde{M}_k$ and $\tilde{N}_l$. With obvious vector sizes, we will denote $v^{(1)} = (v^{(1)}_1, \ldots, v^{(1)}_r)$, $v^{(2)} = (v^{(2)}_1, \ldots, v^{(2)}_s)$, $v^{(3)} = (v^{(3)}_1, \ldots, v^{(3)}_\nu)$ and $v^{(4)} = (v^{(4)}_1, \ldots, v^{(4)}_\mu)$.

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

$$v^{(1)}_1 = v^{(1)}_2 = \sum_{i=1}^{r} \frac{A_i v^{(1)}_1}{\lambda^i} + \sum_{j=1}^{s} \frac{B_j v^{(2)}_1}{\lambda^j},$$

$$v^{(3)}_1 = \sum_{k=1}^{\nu} \frac{|\tilde{M}_k| v^{(1)}_1}{\lambda^k} + \sum_{l=1}^{\mu} \frac{|\tilde{N}_l| v^{(4)}_1}{\lambda^l}, \quad \text{and } v^{(4)}_1 = 0.$$

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

25
If $\lambda \geq 1$, then

$$
\mathcal{N}(v^{(1)}_1) \leq \sum_{i=1}^{r} \lVert A_i \rVert_{\mathcal{N}} \frac{\mathcal{N}(\lVert v^{(1)}_1 \rVert)}{\lvert \lambda \rvert} + \sum_{j=1}^{s} \lVert B_j \rVert_{\mathcal{N}} \frac{\mathcal{N}(v^{(1)}_1)}{\lvert \lambda \rvert^n} \\
\leq \mathcal{N}(v^{(1)}_1) \left( \sum_{i=1}^{r} \lVert A_i \rVert_{\mathcal{N}} + \sum_{j=1}^{s} \lVert B_j \rVert_{\mathcal{N}} \right).
$$

Similarly,

$$
\mathcal{N}(v^{(3)}_1) \leq \sum_{k=1}^{\nu} \lVert \tilde{M}_k \rVert_{\mathcal{N}^*} \frac{\mathcal{N}^*(\lVert v^{(3)}_1 \rVert)}{\lvert \lambda \rvert^k} \leq \mathcal{N}^*(v^{(3)}_1) \sum_{k=1}^{\nu} \lVert \tilde{M}_k \rVert_{\mathcal{N}^*}.
$$

Since $\nu \neq 0$, we obtain

$$
1 \leq \sum_{i=1}^{r} \lVert A_i \rVert_{\mathcal{N}} + \sum_{j=1}^{s} \lVert B_j \rVert_{\mathcal{N}} \text{ or } 1 \leq \sum_{k=1}^{\nu} \lVert \tilde{M}_k \rVert_{\mathcal{N}^*}.
$$

This proves the result. ■

**D Proof of Theorem 5:**

Suppose there exist two strongly stationary solutions $(X_t)$ and $(\tilde{X}_t)$. Since both of them satisfy Equation (8), with obvious notations, we can write for every $t$

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

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

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

$$
X^{(3)}_t := (Vech(Q_t), \ldots, Vech(Q_{t-\nu+1}))', \text{ and }
$$
\[X_t^{(4)} := (Vech(\varepsilon_t \varepsilon_t'), \ldots, Vech(\varepsilon_{t-\mu+1} \varepsilon_{t-\mu+1}'))'.\]

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

\[Y_t = \bar{T}_t Y_{t-1} + \bar{\zeta}_t, \text{ and } \bar{Y}_t = \bar{T}_t \bar{Y}_{t-1} + \bar{\zeta}_t,
\]

for every \(t\), by setting \(\bar{\zeta}_t = (\zeta_t^{(1)}, \zeta_t^{(2)})\). The arguments are then standard: for instance, see Theorem 2.4’s proof in Francq and Zakoian (2010). To get the unicity of \((Y_t)\), it is sufficient to assume that the top Lyapunov exponent \(\gamma_T\) of the sequence of random matrices \((\bar{T}_t)\) is strictly negative. This is the case under Assumption U4 because, for every sequence \((\varepsilon_t)\),

\[E[\ln \|\bar{T}_t \bar{T}_{t-1} \ldots \bar{T}_1\|_1] \leq \ln E[\|\bar{T}_t \bar{T}_{t-1} \ldots \bar{T}_1\|_1] \leq \ln \|\bar{T}_*\|^t_1,
\]

by invoking the matrix norm \(\|A\|_1 := \sum_{i,j} |a_{ij}|\). The fist inequality is due to Jensen’s inequality. The second one is a consequence of the conditional independence between all the r.v. \(\varepsilon_t, \ldots, \varepsilon_1\). Indeed, every term of the random matrix \(\bar{T}_t\), say the \((i,j)\)-th, is the product of a random variable \(\varepsilon_{k_{ij},t}^{2\alpha_{ij}}\) and a deterministic term \(b_{ij}\), where \(\alpha_{ij} \in \{0,1\}\) and \(k_{ij}\) is an index between 1 and \(m\). Denote by \(b_{ij}\) the \((i,j)\)-th term of the matrix \(\bar{T}_*\). Actually, \(\|\bar{T}_t \bar{T}_{t-1} \ldots \bar{T}_1\|_1\) is a sum of terms like

\[\varepsilon_{k_{i_1,j_1},t}^{2\alpha_{i_1,j_1}} \varepsilon_{k_{i_2,j_2},t}^{2\alpha_{i_2,j_2}} \ldots \varepsilon_{k_{i_t,j_t},t}^{2\alpha_{i_t,j_t}} |b_{i_1,j_1} \ldots b_{i_t,j_t}|,
\]

over some collection of indices \(i_1, j_1, \ldots, i_t, j_t\). The expectation of this term is simply \(|b_{i_1,j_1} \ldots b_{i_t,j_t}|\). By collecting all the latter terms, we get \(\|\bar{T}_*\|^t_1\). We
deduce there exists a constant \( C \) s.t.

\[
E[\ln \|\bar{T}_t \bar{T}_{t-1} \ldots \bar{T}_1\|_1] \leq \ln \left\{ C \| (\bar{T}^*)^t \|_{s*} \right\} \leq \ln \left( C \| \bar{T}^* \|_{s*} \right).
\]

Therefore, since \( \gamma_T = \lim_{t \to \infty} t^{-1} E[\ln \|\bar{T}_t \bar{T}_{t-1} \ldots \bar{T}_1\|_1] \), \( \gamma_T \) is strictly negative under Assumption U4, providing the unicity of the processes \((D_t)\) and \((z_t)\) (once we assume the unicity of the processes \((Q_t)\) and \((\varepsilon_t)\)).

Now, let us prove the unicity of \((X_t^{(3)}, X_t^{(4)})\) or, in other terms, of \((Q_t, \varepsilon_t)\). This 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 \((Q_t, \varepsilon_t)\) and \((\tilde{Q}_t, \tilde{\varepsilon}_t)\).

Since the spectral norm is sub-multiplicative, we deduce from (4) that

\[
\|Q_t - \tilde{Q}_t\|_s \leq \sum_{k=1}^{\nu} \|M_k\|^2 \|Q_{t-k} - \tilde{Q}_{t-k}\|_s + \sum_{l=1}^{\mu} \|N_l\|^2 \|\varepsilon_{t-l} \tilde{\varepsilon}_{t-l} - \tilde{\varepsilon}_{t-l} \tilde{\varepsilon}_{t-l}\|_s. \tag{21}
\]

The key point will be to bound from above the terms \(\|\varepsilon_{t-l} \tilde{\varepsilon}_{t-l} - \tilde{\varepsilon}_{t-l} \tilde{\varepsilon}_{t-l}\|_s\) by a function of \(\|Q_{t-l} - \tilde{Q}_{t-l}\|_s\). To lighten the indices, we assume \(l = 0\). Clearly, we have

\[
\|\varepsilon_t \tilde{\varepsilon}_t - \tilde{\varepsilon}_t \tilde{\varepsilon}_t\|_s = \|\tilde{R}^{1/2}_t \eta \tilde{\eta}_t \tilde{R}^{1/2}_t - \tilde{R}^{1/2}_t \eta \tilde{\eta}_t \tilde{R}^{1/2}_t\|_s \\
\leq \|(\tilde{R}^{1/2}_t - \tilde{R}^{1/2}_t) \eta \tilde{\eta}_t \tilde{R}^{1/2}_t\|_s + \|\tilde{R}^{1/2}_t \eta \tilde{\eta}_t (\tilde{R}^{1/2}_t - \tilde{R}^{1/2}_t)\|_s \\
\leq \|\tilde{R}^{1/2}_t - \tilde{R}^{1/2}_t\|_s \|\eta \tilde{\eta}_t\|_s \|\tilde{R}^{1/2}_t\|_s + \|\tilde{R}^{1/2}_t\|_s \|\eta \tilde{\eta}_t\|_s \|\tilde{R}^{1/2}_t - \tilde{R}^{1/2}_t\|_s.
\]

Since the rank of \(\eta \tilde{\eta}_t\) is one, \(\|\eta \tilde{\eta}_t\|_s = Tr(\eta \tilde{\eta}_t) = \|\eta\|_2^2\). Moreover,

\[
\|\tilde{R}^{1/2}_t\|_s = \rho(\tilde{R}_t)^{1/2} \leq Tr(\tilde{R}_t)^{1/2} = \sqrt{m}.
\]

We deduce

\[
\|\varepsilon_t \tilde{\varepsilon}_t - \tilde{\varepsilon}_t \tilde{\varepsilon}_t\|_s \leq 2m^{1/2} \|\eta\|_2 \|\tilde{R}^{1/2}_t - \tilde{R}^{1/2}_t\|_s. \tag{22}
\]
Since the spectral norm is unitarily invariant, Theorem 6.2 in Hingham (2008) provides
\[ \|R_t^{1/2} - \tilde{R}_t^{1/2}\|_s \leq \frac{1}{\lambda_1(R_t)^{1/2} + \lambda_1(\tilde{R}_t)^{1/2}} \|R_t - \tilde{R}_t\|_s. \] (23)

Note that, for any \( t \),
\[ \lambda_1(R_t) = \min_x x' R_t x = \min_x x' \text{diag}(Q_t)^{-1/2} Q_t \text{diag}(Q_t)^{-1/2} x \]
\[ \geq \min_y y' Q_t y \min_x \|\text{diag}(Q_t)^{-1/2} x\|_2 \]
\[ \geq \lambda_1(Q_t) \min_i \frac{1}{q_{ii,t}} \geq \frac{C}{\|Q_t\|_\infty}, \]
invoking Lemma 4. Since the same inequality applies with \( \lambda_1(\tilde{R}_t) \), we get
\[ \frac{1}{\lambda_1(R_t)^{1/2} + \lambda_1(\tilde{R}_t)^{1/2}} \leq \frac{\|Q_t\|_\infty^{1/2} + \|\tilde{Q}_t\|_\infty^{1/2}}{\sqrt{C}}. \] (24)

Moreover,
\[ R_t - \tilde{R}_t = (\text{diag}(Q_t)^{-1/2} - \text{diag}(\tilde{Q}_t)^{-1/2}) Q_t \text{diag}(Q_t)^{-1/2} \]
\[ + \text{diag}(\tilde{Q}_t)^{-1/2} (Q_t - \tilde{Q}_t) \text{diag}(Q_t)^{-1/2} \]
\[ + \text{diag}(\tilde{Q}_t)^{-1/2} \tilde{Q}_t (\text{diag}(Q_t)^{-1/2} - \text{diag}(\tilde{Q}_t)^{-1/2}) =: \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3. \]

Note that \( \mathcal{R}_1 = [(q_{ii,t} - \tilde{q}_{ii,t}) q_{ij,t} q_{ji,t}^{-1/2} q_{ij,t}^{-1/2} q_{ji,t}^{-1/2} + \tilde{q}_{ii,t}^{-1/2}]/(\sqrt{q_{ii,t}} + \sqrt{\tilde{q}_{ii,t}}) \leq \lambda_i \) and \( \|q_{ij,t}\| \leq \sqrt{q_{ij,t}} + \sqrt{\tilde{q}_{ij,t}} \) (Cauchy-Schwartz). Since \( \|A\|_\infty \leq \|A\|_s \leq m \|A\|_\infty \), we get
\[ \|\mathcal{R}_1\|_\infty \leq C_q^{-1} \|\text{diag}(q_{ii,t} - \tilde{q}_{ii,t})\|_\infty \leq C_q^{-1} \|Q_t - \tilde{Q}_t\|_\infty \leq C_q^{-1} \|Q_t - \tilde{Q}_t\|_s, \]
and \( \|\mathcal{R}_1\|_s \leq m C_q^{-1} \|Q_t - \tilde{Q}_t\|_s \). Similarly, \( \|\mathcal{R}_3\|_s \leq m C_q^{-1} \|Q_t - \tilde{Q}_t\|_s \). By
Lemma 4, we obtain

\[ \|R_2\|_s = \|\text{diag}(\tilde{Q}_t)^{-1/2}(Q_t - \tilde{Q}_t)\text{diag}(Q_t)^{-1/2}\|_s \]
\[ \leq \|\text{diag}(\tilde{Q}_t)^{-1/2}\|_s \|\text{diag}(Q_t)^{-1/2}\|_s \|Q_t - \tilde{Q}_t\|_s \]
\[ \leq \frac{1}{\sqrt{\min_i q_{ii,t}}} \frac{1}{\sqrt{\min_i \tilde{q}_{ii,t}}} \|Q_t - \tilde{Q}_t\|_s \leq \frac{1}{C_q} \|Q_t - \tilde{Q}_t\|_s. \]

Globally, we get

\[ \|R_t - \tilde{R}_t\|_s \leq \frac{2m + 1}{C_q} \|Q_t - \tilde{Q}_t\|_s \] (25)

everywhere. Recalling (22), (23) (24) and (25), we deduce

\[ \|\varepsilon_t' - \tilde{\varepsilon}_t'\|_s \leq \frac{2m^{1/2}\|\eta_t\|_2^2}{\sqrt{C_\lambda}} \frac{2m + 1}{C_q} \left( \|Q_t\|_\infty^{1/2} + \|\tilde{Q}_t\|_\infty^{1/2} \right) \|Q_t - \tilde{Q}_t\|_s. \] (26)

Set \( v_t := \|Q_t - \tilde{Q}_t\|_s \). By using the previous inequality and the notation of Lemma 3, we obtain

\[ v_t \leq \sum_{k=1}^\mu \|M_k\|^2 v_{t-k} + \sum_{l=1}^\mu \|N_l\|^2 \frac{4m^{1/2}(2m + 1)}{\sqrt{C_\lambda} C_q} \|\eta_{t-l}\|^2 \sqrt{q_{t-l}} \varepsilon_{t-l} := \sum_{j=1}^\kappa \beta_{j,t} v_{t-j}, \] (27)

for all \( t \) and with our notations.

Under Assumption U3, the product \( \|N_t\|_s \|N_{t-1}\|_s \cdots \|N_{t-p}\|_s \) tends to zero a.e. when \( p \to +\infty \) and any fixed \( t \) (see Bougerol and Picard, 1992, for instance). We deduce \( v_t \to 0 \) a.e. when \( t \to \infty \). This implies that \( Q_t = \tilde{Q}_t \) a.e. because \( (v_t) \) can be initialized arbitrarily far in the past. Therefore, \( R_t = \tilde{R}_t \) a.e. and \( \varepsilon_t = \tilde{\varepsilon}_t \) a.e., knowing \( (\eta_t) \). This concludes the proof. ■

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

\[ \pi_t := \text{Vech}(W_0) + \sum_{l=1}^\mu \tilde{N}_l \text{Vech}(\varepsilon_{t-l}' \varepsilon_{t-l-1}). \]
Under Assumption U1, the sum $\sum_{k=0}^{+\infty} T_{33}^k \pi_{t-k}$ is absolutely convergent a.e., and then $X_{t}^{(3)} = \sum_{k=0}^{+\infty} T_{33}^k \pi_{t-k}$.

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

$$
\|\varepsilon_t\|_\infty \leq \|\varepsilon_t\|_s \leq \|R_t^{1/2}\|_{\infty, s} \leq \|R_t^{1/2}\|_{s} \|\eta_t\|_s
\leq \|R_t^{1/2}\|_{s} \|\eta_t\|_2 \leq \sqrt{m} \|\eta_t\|_2.
$$

This proves the inequality $\|Vech(\varepsilon_t \varepsilon_t')\|_s \leq \sqrt{m^3(m+1)m} \|\eta_t\|_2^2 / \sqrt{2}$, for every $t$ and $l$. We deduce

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

**Proof of Lemma 4:** Is it known that, for any two squared 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 obviously 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

$$
\lambda_1(Q_t) \geq \lambda_1(W_0) + \sum_{k=1}^{\nu} \lambda_1(m(\nu)Q_{t-k}) \geq \lambda_1(W_0) + \sum_{k=1}^{\nu} (m(\nu))^2 \lambda_1(Q_{t-k}). \quad (28)
$$

Introduce the random vector $\vec{\lambda}_t := (\lambda_1(Q_t), \ldots, \lambda_1(Q_{t-\nu+1}))'$ and $\vec{\lambda}_W := (\lambda_1(W_0), 0, \ldots, 0)'$. 31
Because of (28), we have $\tilde{\lambda}_t \geq M^*\tilde{\lambda}_{t-1} + \tilde{\lambda}_W$ for every $t$. Under Assumption U2, it is easy to check that $\sum_{k=0}^{+\infty}(M^*)^k$ is absolutely convergent and that

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

for every $t$. Obviously, $M^*\lambda^*_{\infty} + \lambda^*_W = \lambda^*_{\infty}$. Due to the definition of $M^*$, this implies that all the components of $\lambda^*_{\infty}$ are the same, i.e. there exists a real number $\lambda^*_{\infty}$ such that $\lambda^*_{\infty} = \lambda^*_{\infty} e, e \in \mathbb{R}^\nu$. Taking the first component of the vectorial equation $\lambda^*_{\infty} M^* e + \lambda^*_W = \lambda^*_{\infty} e$ provides $\lambda^*_{\infty} \sum_{k=1}^{\nu}(m(k))^2 + \lambda^*_1(W_0) = \lambda^*_{\infty}$. This proves the lower bound of $\lambda^*_1(Q_t)$ under U2.

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

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

for all $t$, this inequality playing the same role as (28). This implies the desired result. ■

REFERENCES


