Econometric Theory, **14**, 1998, 70–86. Printed in the United States of America.

STRONG CONSISTENCY OF ESTIMATORS FOR MULTIVARIATE ARCH MODELS

THIERRY JEANTHEAU
Université de Marne-la-Vallée

This paper deals with the asymptotic properties of quasi-maximum likelihood estimators for multivariate heteroskedastic models. For a general model, we give conditions under which strong consistency can be obtained; unlike in the current literature, the assumptions on the existence of moments of the error term are weak, and no study of the various derivatives of the likelihood is required. Then, for a particular model, the multivariate GARCH model with constant correlation, we describe the set of parameters where these conditions hold.

1. INTRODUCTION

As a result of the paper by Mandelbrot (1963), we know that for certain time series, and especially economic and financial time series, the conditional variance is not constant over time. Therefore, several models trying to take into account this particular behavior have been introduced. The most successful ones are undoubtedly the autoregressive conditional heteroskedastic (ARCH) model, introduced by Engle (1982), and some of its derivative models (GARCH, GARCH-M, EGARCH, etc.). The implementation of these parametric models is relatively simple. And, from a practical point of view, it is well known now how to identify, estimate, and test this kind of model (for a description of these methods and some empirical evidence, see the survey of Bollerslev, Chou, and Kroner, 1992).

From a theoretical point of view, however, the problem of statistical inference for these models remains partially open. Indeed, Weiss (1986) gave the first proof of consistency and asymptotic normality of the maximum likelihood estimator for univariate ARCH model but under strong conditions on the existence of the moments of the error term. On the other hand, a paper of Nelson (1990) showed that for a particular model, the GARCH(1,1) model, the process can be strictly stationary with infinite second moment, and the empirical studies (see Bollerslev et al., 1992) suggest that some financial time series seem to be characterized by such a process. Therefore, statistical inference under weak conditions on the existence of moments is needed.

I thank Laure Elie for all the help and encouragement I received while working on this paper. Address correspondence to: Thierry Jeantheau, Université de Marne-la-vallee, Equipe d'Analyse et de Mathematiques Apliquees, 2 rue de la butte verte, 93166 Noisy-le-grand Cedex, France.

  0266-4666/98 $9.50

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

Lumsdaine (1991) gave such a proof of consistency and asymptotic normality for the univariate GARCH(1,1) model, for the set of parameters where there exists a strictly stationary solution (including the IGARCH case). Lee and Hansen (1994) obtained the same result under weaker conditions. However, only a local consistency is obtained, and the technical proofs seem difficult to extend to other models. For the same models, Elie and Jeantheau (1995) gave a proof of global consistency, which is less technical and therefore can be applied to more complex univariate models, for instance, to a GARCH(p,q) model.

In this paper, we focus our attention on multivariate models. In the literature, several formulations of conditionally heteroskedastic multivariate models have been introduced (see, e.g., Bollerslev, 1987; Bollerslev, Engle, and Wooldridge, 1988; Engle and Kroner, 1995). But the complexity of the equations involved in all these models makes the study of the likelihood difficult, and the method of the proofs of Lumsdaine (1991) or Lee and Hansen (1994) cannot be used. The aim of this paper is to show that the methodology used in Elie and Jeantheau (1995) can be applied to multivariate models because the consistency of the quasi-maximum likelihood estimator is obtained without any study of the various derivatives of the log-likelihood function, and under weak conditions on the existence of moments.

The paper is organized as follows. Section 2 presents a general multivariate heteroskedastic model, the estimation method, and the assumptions under which we can derive the strong consistency of the estimator. For this purpose, we recall a theorem of consistency (see Pfanzagl, 1969). Section 3 presents a particular example of this model, the multivariate GARCH model with constant correlation, introduced by Bollerslev (1987). For this model, we study more particularly the stationarity and the identifiability, and then we are able to describe the set of parameters where the quasi-maximum likelihood estimator is strongly consistent. Section 4 concludes. Mathematical proofs are deferred to the Appendix.

2. PARAMETER ESTIMATION FOR MULTIVARIATE HETEROSKEDASTIC MODEL

2.1. The Model and Assumptions

Let $(\Omega, \mathcal{A}, \mathcal{P})$ be a probability space, $\{Y_t, t \in \mathbb{Z}\}$ an $\mathbb{R}^d$-valued process, and θ a parameter in $\Theta \subset \mathbb{R}^s$. We say that Y_t is a multivariate autoregressive process with conditionally heteroskedastic errors if, for all $t \in \mathbb{Z}$, we have

$$Y_t = \Phi_\theta(\underline{Y}_{t-1}) + \Delta_\theta(\underline{Y}_{t-1})\eta_t, \tag{1}$$

where

$\underline{Y}_{t-1} = (Y_{t-1}, Y_{t-2}, \dots)$,
Φ_θ is a measurable function from $(\mathbb{R}^d)^{\mathbb{N}} \to \mathbb{R}^d$,
Δ_θ is a measurable function from $(\mathbb{R}^d)^{\mathbb{N}} \to \mathbb{R}^d \otimes \mathbb{R}^d$,
$\{\eta_t, t \in \mathbb{Z}\}$ is a doubly infinite sequence of independent and identically distributed (i.i.d.) $\mathbb{R}^d$ random variables defined on Ω, with mean 0 and covariance matrix Γ such that

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

$$\Gamma = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1d} \\ \rho_{12} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_{(d-1)d} \\ \rho_{1d} & \cdots & \rho_{(d-1)d} & 1 \end{pmatrix},$$

η_t is independent of the σ-field I_{t-1} generated by $\{Y_{t-1}, Y_{t-2}, \ldots\} = \underline{Y}_{t-1}$, and η_t is I_t-measurable.

Note that the system extends infinitely far into the past. From (1), we remark that

$$E(Y_t/I_{t-1}) = \Phi_\theta(\underline{Y}_{t-1}) \quad \text{and} \quad \text{Var}(Y_t/I_{t-1}) = \Delta_\theta(\underline{Y}_{t-1})\,\Gamma\,\Delta_\theta(\underline{Y}_{t-1})'.$$

Therefore, $\Phi_\theta(\underline{Y}_{t-1})$ is the prediction of Y_t when its past is known, and $\Delta_\theta(\underline{Y}_{t-1})\eta_t$ is the error term. For simplicity, we will denote

$$\Phi_{t,\theta} = \Phi_\theta(\underline{Y}_{t-1}) \quad \text{and} \quad \Delta_{t,\theta} = \Delta_\theta(\underline{Y}_{t-1}),$$

and the conditional covariance matrix of the error term

$$H_{t,\theta} = \text{Var}(Y_t/I_{t-1}) = \Delta_{t,\theta}\,\Gamma\,\Delta'_{t,\theta}.$$

2.2. Main Result

Let the true value of the parameter θ_0 be in Θ. For ease of reference, we list the assumptions under which a strongly consistent estimator is obtained. (Note that "a.s." in Assumption A5 is defined as "almost surely.")

Assumption A0 (Compactness). Θ is compact.

Assumption A1 (Ergodicity). $\forall\theta_0 \in \Theta$, model (1) admits a unique strictly stationary and ergodic solution Y_t, following a stationary law P_{θ_0}.

Assumption A2 (Lower bound for the determinant of the conditional covariance matrix). There exists a deterministic constant $c > 0$ such that $\forall t$, $\forall\theta \in \Theta$, $\det(H_{t,\theta}) \geq c$.

Assumption A3 (Logarithmic moment). $\forall\theta_0 \in \Theta$, $E_{\theta_0}(|\log(\det H_{t,\theta_0})|) < \infty$.

Assumption A4 (Identifiability). The functions Φ and H are such that

$$\forall\theta \in \Theta, \forall\theta_0 \in \Theta, \quad \left.\begin{array}{c} \Phi_{t,\theta} = \Phi_{t,\theta_0} \quad P_{\theta_0}\text{ a.s.} \\ \text{and} \\ H_{t,\theta} = H_{t,\theta_0} \quad P_{\theta_0}\text{ a.s.} \end{array}\right\} \Rightarrow \theta = \theta_0.$$

Assumption A5 (Continuity). The functions Φ and H are continuous functions of the parameter θ.

Let us remark that we do not assume that the random variables η_t are Gaussian. Therefore, we use the quasi-maximum likelihood estimation method: We consider a function $F_T(\theta)$ that would be the conditional (on $\underline{Y}_0$) log likelihood of the sample $(Y_1, \ldots, Y_T)$ if the random variables (η_t) were Gaussian, that is to say

$$F_T(\theta) = F_T(\underline{Y}_T, \theta) = \frac{1}{T}\sum_{t=1}^{T} f(\underline{Y}_t, \theta)$$

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

with

$$f(\underline{Y}_t,\theta) = \log(\det H_{t,\theta}) + (Y_t - \Phi_{t,\theta})' H_{t,\theta}^{-1}(Y_t - \Phi_{t,\theta}). \tag{2}$$

Then we define our estimator $\hat{\theta}_T$ as any solution of the equation

$$\hat{\theta}_T = \arg\inf_{\theta} F_T(\underline{Y}_T,\theta). \tag{3}$$

We will refer to $F_T(\theta)$ as a contrast process and to $\hat{\theta}_T$ as a minimum contrast estimator (see Dacunha-Castelle and Duflo, 1983). It is the exact maximum conditional likelihood estimator if $(\eta_t, t \in \mathbb{Z})$ are Gaussian.

We can now give the main result of this paper.

THEOREM 2.1. *Under Assumptions* A0–A5, *the minimum contrast estimator for our model is strongly consistent, that is to say*

$$\hat{\theta}_T \xrightarrow{T\to\infty} \theta_0 \quad P_{\theta_0} \text{ a.s.}$$

This result will be proved in the next section. Let us remark that this is obtained under weak conditions; Assumptions A0, A1, A4, and A5 are common to get the consistency of a minimum contrast estimator, and A3 is a weak condition on the existence of moment. Assumption A2 is a crucial assumption in our proof.

2.3. Proof of the Strong Consistency

For model (1), the consistency of the estimator relies on the following theorem (a proof can be found in Pfanzagl, 1969, for i.i.d. data and can be immediately generalized for strictly stationary and ergodic data). Set $x^- = \inf(x,0)$.

THEOREM 2.2. *On* $(\Omega,\mathcal{A},P)$, *let* $\{Y_t, t \in Z\}$ *be a strictly stationary and ergodic process*, θ *a parameter in* Θ, *and* $F_T(\underline{Y}_T,\theta)$ *be a contrast process such that*

$$F_T(\underline{Y}_T,\theta) = T^{-1}\sum_{t=1}^{T} f(\underline{Y}_t,\theta),$$

where f is a measurable function with real values and continuous in θ. *Let* $B(\theta,\rho)$ *be the ball of center* θ *and radius* ρ, *and* $f_*(\theta,\rho) = \inf\{f(\underline{Y}_t,\theta'), \theta' \in B(\theta,\rho)\}$. *Make the following suppositions.*

*Hypothesis H*0. Θ *is compact.*
*Hypothesis H*1. *The function* $F(\theta_0,\theta)$, *defined for all* $\theta \in \Theta$ *by* $F(\theta_0,\theta) = E_{\theta_0}(f(\underline{Y}_1,\theta))$, *has a unique finite minimum at* θ_0.
*Hypothesis H*2. $\forall\theta \in \Theta$, $E_{\theta_0}(f_*^-(\theta,\rho)) > -\infty$.

Then, the minimum contrast estimator $\hat{\theta}_T$ *associated to* $F_T(\theta)$ *converges* P_{θ_0} *a.s. to* θ_0 *when* $T \to \infty$.

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

In the current literature on ARCH-type models, the consistency of estimators relies on theorems similar to the one given in Andrews (1987). In this case, the proof is often very complex and technical (see Lumsdaine, 1991; Lee and Hansen, 1994). The use of Theorem 2.2 allows us to simplify this proof. Indeed, notice that it is not assumed that the contrast function $F(\theta_0,\theta)$ is finite everywhere. Therefore, we can still deal with models where $E_{\theta_0}(f(\underline{Y}_1,\theta))$ is not always finite or where the finiteness is difficult to check, such as in most ARCH-type models. Furthermore, no assumption is made on the behavior of the derivatives of the contrast process.

The condition of moment induced by Theorem 2.2 is given by the finiteness of $F(\theta_0,\theta_0)$, and therefore we must have $E_{\theta_0}|f(\underline{Y}_1,\theta_0)| < +\infty$. This condition is fulfilled by the conditionally heteroskedastic model under the weak (logarithmic) moment condition A3. Therefore, in the univariate case, it is possible to use Theorem 2.2 to deal with the IGARCH model (see Elie and Jeantheau, 1995).

By Theorem 2.2, we get the strong consistency of our estimator (Theorem 2.1) if Assumptions A1–A5 imply H1 and H2, when $f(\underline{Y}_t,\theta)$ is given by (2). First, we remark that, under A2, $H_{t,\theta}$ is a positive definite matrix, as is $H_{t,\theta}^{-1}$, and, thus, $F_T(\underline{Y}_T,\theta)$ is always greater than $\log(c)$. Therefore, H2 holds obviously, and this remark is also important for the proof of the next proposition, which shows that H1 holds.

PROPOSITION 2.1. *Under Assumptions* A1–A4, $F_T(\underline{Y}_T,\theta)$ *converges*, *when* $T \to \infty$, P_{θ_0} *a.s. to* $F(\theta_0,\theta) = E_{\theta_0}(f(\underline{Y}_1,\theta))$, *and this function has a unique finite minimum in* θ_0.

Proof. See Appendix.

Therefore, the strong consistency of our estimator for the model (1), as given in Theorem 2.1, follows from Proposition 2.1 and Theorem 2.2.

3. APPLICATION TO THE MULTIVARIATE GARCH MODEL WITH CONSTANT CORRELATION

For particular models, however, the verification of Assumptions A1–A5 may be difficult. In the univariate case, it is possible to prove that these assumptions hold for a GARCH(1,1) model for the set of parameters Θ where the condition of strict stationarity given by Nelson (1990) is satisfied (including the IGARCH(1,1) case) and where the conditional variance is bounded from below. These results can be extended to strictly stationary GARCH(p,q) models, under some additional conditions of identifiability (see Elie and Jeantheau, 1995). The aim of this section is to show how this can be achieved for a multivariate GARCH model with constant correlation. We first introduce this model.

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

3.1. The Model

For simplicity, we will only deal with a case where Φ_θ is equal to a constant $\beta \in \mathbb{R}^d$. Therefore, we assume that

$$Y_t = \beta + \varepsilon_t, \tag{4}$$

where the error term ε_t is defined as follows (see Bollerslev, 1987).

DEFINITION 3.1. *A sequence* $\{\varepsilon_t, t \in \mathbb{Z}\}$ *of random variables with values in* $\mathbb{R}^d$ *follows a multivariate* GARCH(p,q) *process with constant correlation if*

$$\varepsilon_t = \Delta_t(\theta)\eta_t,$$

where $\Delta_t(\theta)$ *is a diagonal matrix and the elements of the diagonal* $\Delta_{t,ii}(\theta)$ *satisfy, for all* i, *the following relation*:

$$\begin{pmatrix} (\Delta_{t,11}(\theta))^2 \\ \vdots \\ (\Delta_{t,dd}(\theta))^2 \end{pmatrix} = W + \sum_{i=1}^{q} A_i \begin{pmatrix} \varepsilon_{t-i,1}^2 \\ \vdots \\ \varepsilon_{t-i,d}^2 \end{pmatrix} + \sum_{i=1}^{p} B_i \begin{pmatrix} (\Delta_{t-i,11}(\theta))^2 \\ \vdots \\ (\Delta_{t-i,dd}(\theta))^2 \end{pmatrix}, \tag{5}$$

where $W \in \mathbb{R}^d$, A_i *and* $B_i \in \mathbb{R}^d \times \mathbb{R}^d$, *and we assume that all coefficients of these matrices are positive. Furthermore, we have the following conditions on* η_t:

1. $\{\eta_t, t \in \mathbb{Z}\}$ *is a sequence of i.i.d.* $\mathbb{R}^d$*-valued random variables, with mean* 0 *and covariance matrix* Γ *such that*

$$\Gamma = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1d} \\ \rho_{12} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_{(d-1)d} \\ \rho_{1d} & \cdots & \rho_{(d-1)d} & 1 \end{pmatrix},$$

2. η_t *is independent of the* σ*-field* I_{t-1} *generated by* $\{Y_{t-1}, Y_{t-2}, \ldots\}$, *and*
3. *the law of* η_t *is such that there is no quadratic form* q *for which* $q(\eta_t) = \delta$ *a.s., with* $\delta \in \mathbb{R}$.

From this definition, we see that the conditional covariance matrix of ε_t, denoted by $H_t(\theta)$, is such that

$$\begin{cases} H_{t,ii}(\theta) = (\Delta_{t,ii}(\theta))^2 \\ H_{t,ij}(\theta) = \rho_{ij}\Delta_{t,ii}(\theta)\Delta_{t,jj}(\theta) \quad \text{for } i \neq j. \end{cases}$$

Therefore, the conditional correlation between $\varepsilon_{t,i}$ and $\varepsilon_{t,j}$ is constant. This assumption may seem restrictive, but some empirical studies have found it to be reasonable (see Bollerslev et al., 1992).

The multivariate GARCH model with constant correlation is an example of the model (1) described in the previous section. We just add Assumption A3 about the

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

law of η_t; this will be important for ensuring the identifiability of this model. The parameter θ of our model that we have to estimate is the $d(1 + 1 + dq + dp + (d-1)/2)$ vector constructed with β, W, A_i, B_i, and ρ_{ij}. In the following sections, we will derive conditions on this model under which Assumptions A1–A5 hold, and therefore under which the minimum contrast estimator is strongly consistent.

3.2. Stationarity (Assumption A1)

First, we must find the set of parameters Θ such that the solution of the GARCH model verifies Assumption A1 of strict stationarity and ergodicity. The weak stationarity of univariate ARCH and GARCH models can be found in the original papers of Engle (1982) and Bollerslev (1986), and one can derive from this the analogous results for multivariate models. The strict stationarity is more complex to obtain; for the GARCH(1,1) model, it is possible to give an explicit necessary and sufficient condition on the set of parameters (see Nelson, 1990) to get the result, but for the univariate GARCH(p,q) model, the condition is formulated with the Lyapunov exponent of a matrix associated to θ (see Bougerol and Picard, 1992), and a generalization of this result to multivariate GARCH models seems difficult.

However, it is possible to give a condition under which there exists a weakly stationary solution (see Bollerslev, 1987; Engle and Kroner, 1995). Therefore, for our model, we give this condition (denoted by B1), and then we prove that the solution is also strictly stationary and ergodic and thus that Assumption A1 holds. Let us remark that, for our purpose, B1 is too strong because it implies that the solution has a finite second moment, whereas the method of our proof requires only a finite logarithmic moment.

Let us denote as Id the $d \times d$ identity matrix.

PROPOSITION 3.1. *We will make the following assumption.*

Assumption B1. θ is such that $\det(\mathrm{Id} - \sum_{i=1}^{n}(A_i + B_i)\lambda^i)$ has its roots outside the unit circle.

Under Assumption B1, the multivariate GARCH(p,q) *model with constant correlation has a weakly stationary solution. Moreover, this solution is unique and is also strictly stationary and ergodic.*

Proof. See Appendix.

It is important to remark that one can derive the strict stationarity and the ergodicity because the solution of this model has an explicit expression in terms of the sequence $\{\eta_t, t \in \mathbb{Z}\}$; this is a specific feature of the multivariate GARCH model with constant correlation.

3.3. Identifiability (Assumption A4)

The aim of this section is to check Assumption A4. Using the backshift operator notation L, we can rewrite (5) as

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

$$P(L)\begin{pmatrix} H_{t,11}(\theta) \\ \vdots \\ H_{t,dd}(\theta) \end{pmatrix} = W + Q(L)\begin{pmatrix} \varepsilon_{t,1}^2 \\ \vdots \\ \varepsilon_{t,d}^2 \end{pmatrix}, \tag{6}$$

with P and Q two matrices with polynomial coefficients such that $P(L) = \text{Id} - \sum_{i=1}^{p} B_i L^i$ and $Q(L) = \sum_{i=1}^{q} A_i L^i$. If we multiply (6) by a matrix with polynomial coefficients $R(L) \neq \text{Id}$, we get another formulation of the multivariate GARCH model with the same solution; therefore, to be identifiable, this formulation must be minimal in a certain sense. Before giving a definition of the term *minimal*, we recall in the next section some properties of matrices with polynomial coefficients.

3.3.1. Definition and properties related to matrices with polynomial coefficients. The results and the proofs of this section can be found in Kailath (1980). We denote by *MP* the set of matrices with polynomial coefficients. Let us recall that a square matrix $M(L) \in MP$ is unimodular if its determinant is nonequal to 0 and independent of the lag operator L. Because of Theorem 3.1, we may define a greatest common left divisor (see Kailath,1980).

THEOREM 3.1. *Let* $A, B \in MP$ *such that* $\det A \not\equiv 0$ *and* $\det B \not\equiv 0$; *there exists* $D \in MP$ *such that*

every left divisor of D is also a left divisor of A and B, and
every left divisor of A and B is also a left divisor of D.

The matrix D is called the greatest common left divisor of A and B. Moreover, we have the Bezout equality

$\exists (U,V) \in MP^2 / D = AU + BV.$

It is important to note that the greatest common left divisor is not unique, but if D' is another greatest common left divisor of A and B, then there exists a unimodular matrix W such that $D' = DW$. Therefore, we say that two matrices of *MP* are coprime if any of their greatest common left divisor is unimodular.

In the univariate case, it is usual to assume that the two polynomials involved in the equation of the model are coprime to get the identifiability. But, because the greatest common left divisor is not unique for polynomial matrices, this assumption is not sufficient in the multivariate case. Therefore, we introduce here the notion of "column-reduced" matrix, which will be useful in what follows. Let $M(L)$ be a polynomial matrix and d_{ij} the degree of the polynomial $M_{ij}(L) = \sum_{l=0}^{d_{ij}} a_{ij,l} L^l$. We define $d_j(M)$ and the matrix M^{rc} by

$$d_j(M) = \sup_i d_{ij} \quad \text{and} \quad M_{ij}^{rc} = a_{ij,d_j}.$$

Then, we can introduce the following definition.

DEFINITION 3.2. *A polynomial matrix M is column reduced if the determinant of* M^{rc} *is nonequal to* 0.

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

3.3.2. Minimal multivariate GARCH formulation. We will give in this section a condition for the identifiability of the multivariate model (5). Let us give a first lemma.

LEMMA 3.1. *If U is a $d \times d$ matrix and V an I_{t-1}-measurable vector, we have*

$$U \begin{pmatrix} \varepsilon_{1,t}^2 \\ \vdots \\ \varepsilon_{d,t}^2 \end{pmatrix} = V \Rightarrow U = 0 \quad \text{and} \quad V = 0.$$

Proof. See Appendix.

We can now prove the following proposition.

PROPOSITION 3.2. *Let (P_1, Q_1) be a couple of polynomial matrices such that*

the model (5) *has a weakly stationary solution,*
$\det P_1 \not\equiv 0$ *and* $\det Q_1 \not\equiv 0$, *and*
P_1 *and* Q_1 *are coprime.*

Then, if ε_t is also the solution of a model written with the polynomial matrices (P_2, Q_2), there exists a polynomial matrix M such that

$$P_2 = M \; P_1 \quad \text{and} \quad Q_2 = M \; Q_1.$$

Proof. See Appendix.

As indicated previously, the condition "P_1 and Q_1 are coprime" is not sufficient for identifiability of the model. Indeed, it is possible to find a unimodular matrix M such that $M \neq \text{Id}$, $d_j(MP) = d_j(P)$, $d_j(MQ) = d_j(Q)$, and $M(0) = \text{Id}$. Therefore, the multivariate GARCH model formulated with (MP, MQ) has the same solution (it is the only way to find a couple of polynomial matrices with this property). The following definition gives an additional assumption to get rid of this case.

DEFINITION 3.3. *We say that the formulation of a multivariate GARCH (p,q) model is minimal if*

$$P(L) \begin{pmatrix} H_{t,11}(\theta) \\ \vdots \\ H_{t,dd}(\theta) \end{pmatrix} = W + Q(L) \begin{pmatrix} \varepsilon_{t,1}^2 \\ \vdots \\ \varepsilon_{t,d}^2 \end{pmatrix},$$

with P and Q satisfying

1. $P(0) = \text{Id}$ *and* $Q(0) = 0$.
2. $\det P \not\equiv 0$ *and* $\det Q \not\equiv 0$.
3. *P and Q are coprime.*
4. $\forall j$, $1 \leq j \leq d$, $d_j(P) = d_j \leq p$ *and* $d_j(Q) = d_j \leq q$.
5. *P or Q is column reduced.*

With this last condition, we can prove the following result, which is a justification of the term *minimal*.

PROPOSITION 3.3. *Let* (P_1, Q_1) *define a minimal formulation of a multivariate* GARCH(p,q) *model, such that there exists a weakly stationary solution denoted* ε_t; *then, if* ε_t *is also the solution of another model written with* (P_2, Q_2), *there exists* j, *such that* $d_j(P_2) > d_j(P_1)$ *or* $d_j(Q_2) > d_j(Q_1)$.

Proof. See Appendix.

According to the results of the previous section, let us introduce the following assumption.

Assumption B2. The formulation at θ_0 of the multivariate GARCH(p,q) model with constant correlation is minimal.

The following proposition proves that, under this assumption, the model is identifiable.

PROPOSITION 3.4. *Let* $\theta \in \Theta$. *Under Assumptions* B1 *and* B2, *and if* $H_{t,\theta}$ *is the weakly stationary solution of the multivariate GARCH model with constant correlation*, *we have*

$$\forall \theta \in \Theta, \forall \theta_0 \in \Theta, \qquad \left.\begin{array}{cc} \beta = \beta_0 & P_{\theta_0}\ a.s. \\ & and \\ H_{t,\theta} = H_{t,\theta_0} & P_{\theta_0}\ a.s. \end{array}\right\} \Rightarrow \theta = \theta_0.$$

Proof. See Appendix.

3.4. Consistency of the Minimum Contrast Estimator

We can now give a set of sufficient conditions to get the consistency of the minimum contrast estimator. Let us introduce the following additional assumptions.

Assumption B0. Θ is compact.

Assumption B3. There exist two strictly positive constants c_1 and c_2 such that all the elements of W are greater than $c_1^{1/d}$ and $\det \Omega \geq c_2$.

THEOREM 3.2. *Under Assumptions* B0–B3, *the minimum contrast estimator for a multivariate* GARCH(p,q) *with constant correlation* (*see* (4)) *is strongly consistent*.

Proof. See Appendix.

The stationarity and the identifiability of the model have been verified in the previous sections. It is not difficult to see that Assumption B3 implies that the determinant of the conditional covariance matrix has a lower bound. Last, because under B1 our solution is weakly stationary, we have a finite second moment; therefore, Assumption A3 of finite logarithmic moment is easily verified.

4. CONCLUSION

This paper has shown a way to give a proof of the consistency of the quasi-maximum likelihood estimator for multivariate GARCH models. However, many questions remain open. First, we verified the necessary assumptions for only one particular model, the GARCH model with constant correlation, for which we are able to describe a set of parameters where the consistency holds; it would be useful to apply it to other multivariate models. It would be also very important to complete the asymptotic theory of these models by giving a proof of the asymptotic normality of the estimators. However, it seems difficult to give such a proof without a deep study of the contrast process $F_T(\theta)$ and its derivatives; unfortunately, the expressions involved in these models are very cumbersome.

REFERENCES

Andrews, D.W.K. (1987) Consistency in nonlinear econometric model: A generic uniform law of large number. *Econometrica* 55, 1465–1471.

Bollerslev, T.P. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 309–328.

Bollerslev, T.P. (1987) Modeling the coherence in short run nominal exchange rates: A multivariate generalized ARCH approach. *Review of Economics and Statistics* 72, 498–505.

Bollerslev, T.P., R.Y. Chou, & K.F. Kroner (1992) ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics* 31, 309–328.

Bollerslev, T.P., R.F. Engle, & J.M. Wooldridge (1988) A capital asset pricing model with time varying covariances. *Journal of Political Economy* 96, 116–131.

Bougerol, P. & N. Picard (1992) Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* 52, 115–127.

Dacunha-Castelle, D. & M. Duflo (1983) *Probabilités et statistiques; Problèmes à temps mobile.* Paris: Masson.

Elie, L. & T. Jeantheau (1995) Consistance dans les modèles hétéroscedastiques. *Comptes Rendus de l'Académie des Sciences Série I* 320, 1255–1258.

Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* 50, 987–1008.

Engle, R.F. & K.F. Kroner (1995) Multivariate simultaneous generalized ARCH. *Econometric Theory* 11, 122–150.

Kailath, T. (1980) *Linear Systems.* Englewood Cliffs, New Jersey: Prentice-Hall.

Lee, S.W. & B.E. Hansen (1994) Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory* 10, 29–52.

Lumsdaine, R.L. (1991) Asymptotic Properties of the Maximum Likelihood Estimator in GARCH(1.1) and IGARCH(1,1) models. Manuscript, Princeton University and National Bureau of Economic Research.

Mandelbrot, B. (1963) The variation of certain speculative prices. *Journal of Business* 36, 394–419.

Nelson, D. (1990) Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* 6, 318–334.

Pfanzagl, J. (1969) On the measurability and consistency of minimum contrast estimates. *Metrika* 14, 249–272.

Weiss, A.A. (1986) Asymptotic theory for ARCH models: Stability, estimation and testing. *Econometric Theory* 2, 107–131.

APPENDIX

Proof of Proposition 2.1. Assumption A2 implies that $F_T(\underline{Y}_T,\theta) \geq \log(c)$, and A1 and the ergodic theorem yield

$$F(\theta_0,\theta) = E_{\theta_0}(f(\underline{Y}_1,\theta)) \quad \text{if } E_{\theta_0}(f^+(\underline{Y}_1,\theta)) < \infty,$$
$$= +\infty \quad \text{if not.}$$

Because

$$F(\theta_0,\theta_0) = E_{\theta_0}(\log(\det H_{1,\theta_0})) + E_{\theta_0}((Y_1 - \Phi_{1,\theta_0})' H_{1,\theta_0}^{-1}(Y_1 - \Phi_{1,\theta_0})),$$

by A3, the first term is finite, the second term is equal to d, and $F(\theta_0,\theta_0)$ is finite. Now, using A1, we deduce that $F_T(Y,\theta)$ converges P_{θ_0} a.s. to a function $F(\theta_0,\theta)$ such that

$$F(\theta_0,\theta) - F(\theta_0,\theta_0) = E_{\theta_0}(\log(\det H_{1,\theta}) - \log(\det H_{1,\theta_0}))$$
$$+ E_{\theta_0}((Y_1 - \Phi_{1,\theta})' H_{1,\theta}^{-1}(Y_t - \Phi_{1,\theta}) + d).$$

We can write

$$Y_1 - \Phi_{1,\theta} = (Y_1 - \Phi_{1,\theta_0}) + (\Phi_{1,\theta_0} - \Phi_{1,\theta})$$
$$= \Delta_{1,\theta_0}\eta_1 + (\Phi_{1,\theta_0} - \Phi_{1,\theta}).$$

Because $H_{1,\theta}^{-1}$ is a positive definite matrix, there exists a matrix M such that $H_{1,\theta}^{-1} = M'M$ and

$$(Y_1 - \Phi_{1,\theta})' H_{1,\theta}^{-1}(Y_1 - \Phi_{1,\theta}) = (\Delta_{1,\theta_0}\eta_1 + (\Phi_{1,\theta_0} - \Phi_{1,\theta}))' M'M(\Delta_{1,\theta_0}\eta_1 + (\Phi_{1,\theta_0} - \Phi_{1,\theta}))$$
$$= (A\eta_1 + B)'(A\eta_1 + B),$$

where the two random variables A and B are I_0-measurable and independent of η_1. Therefore,

$$E_{\theta_0}((Y_1 - \Phi_{1,\theta})' H_{1,\theta}^{-1}(Y_1 - \Phi_{1,\theta})) = \int dP_{A,B}(a,b) E_{\theta_0}((a\eta_1 + b)'(a\eta_1 + b))$$
$$= \int dP_{A,B}(a,b) E_{\theta_0}(\mathrm{tr}(a'\Omega a) + b'b)$$
$$= E_{\theta_0}(\mathrm{tr}(A'\Omega A) + B'B).$$

That is to say,

$$E_{\theta_0}((Y_1 - \Phi_{1,\theta})' H_{1,\theta}^{-1}(Y_1 - \Phi_{1,\theta})) = E_{\theta_0}(\mathrm{tr}(H_{1,\theta_0} H_{1,\theta}^{-1}))$$
$$+ E_{\theta_0}((\Phi_{1,\theta_0} - \Phi_{1,\theta})' H_{1,\theta}^{-1}(\Phi_{1,\theta_0} - \Phi_{1,\theta})).$$

Thus,

$$F(\theta_0,\theta) - F(\theta_0,\theta_0) \geq E_{\theta_0}(\log(\det H_{1,\theta}) - \log(\det H_{1,\theta_0})) + E_{\theta_0}(\mathrm{tr}(H_{1,\theta_0} H_{1,\theta}^{-1}) - d),$$

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

the equality holds if and only if $\Phi_{1,\theta_0} = \Phi_{1,\theta}$ a.s. In this case, we obtain

$$F(\theta_0,\theta) - F(\theta_0,\theta_0) = E_{\theta_0}(-\log(\det H_{1,\theta_0}H_{1,\theta}^{-1}) + \operatorname{tr}(H_{1,\theta_0}H_{1,\theta}^{-1}) - d).$$

Because H_{1,θ_0} and $H_{1,\theta}^{-1}$ are positive definite matrix, the d eigenvalues λ_i of $H_{1,\theta_0}H_{1,\theta}^{-1}$ are positive and

$$F(\theta_0,\theta) - F(\theta_0,\theta_0) = E_{\theta_0}\left(\sum_{i=1}^{d} - \log\lambda_i + \lambda_i - 1\right).$$

Using the inequality $x - 1 \geq \log x$ for $x > 0$, we see that $F(\theta_0,\theta) - F(\theta_0,\theta_0) > 0$. Furthermore, the equality holds if and only if all the eigenvalues are equal to 1, i.e., when $H_{1,\theta_0}H_{1,\theta}^{-1} = \mathrm{Id}$ a.s. Therefore, according to the assumption of identifiability A4, we have $F(\theta_0,\theta) = F(\theta_0,\theta_0)$ if and only if $\theta = \theta_0$. ■

Proof of Proposition 3.1. First, we prove that there exists a stationary process $H_t(\theta)$ that satisfies (5). Denoting $A(\eta_t^2)$ the matrix where the lth column of A is multiplied by $\eta_{t,l}^2$, this equation may be formulated as

$$\begin{pmatrix} H_{t,11}(\theta) \\ \vdots \\ H_{t,dd}(\theta) \end{pmatrix} = W + \sum_{i=1}^{n} (A_i(\eta_t^2) + B_i)\begin{pmatrix} H_{t-i,11}(\theta) \\ \vdots \\ H_{t-i,dd}(\theta) \end{pmatrix},$$

where $n = \sup(p,q)$, $A_i(\eta_t^2) = 0$ for $i > q$ and $B_i = 0$ for $i > p$.

Let us consider $V_{t,\theta}' = (H_{t,11}(\theta),\ldots,H_{t,dd}(\theta),H_{t-1,11}(\theta),\ldots,H_{t-n+1,dd}(\theta))$. We have

$$V_{t,\theta} = \begin{pmatrix} A_1(\eta_{t-1}^2) + B_1 & A_2(\eta_{t-2}^2) + B_2 & \cdots & \cdots & A_n(\eta_{t-n}^2) + B_n \\ \mathrm{Id} & & & & 0 \\ & \mathrm{Id} & & & \vdots \\ & & \ddots & & \vdots \\ & & & \mathrm{Id} & 0 \end{pmatrix} V_{t-1,\theta} + \begin{pmatrix} W \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

We get

$$V_{t,\theta} = F(\zeta_t)V_{t-1,\theta} + G, \qquad \textbf{(7)}$$

with $\zeta_t' = (\eta_{t-1},\eta_{t-2},\ldots,\eta_{t-n})$ and $G' = (W,0,\ldots,0)$. The term $F(\zeta_t)$ is the preceding matrix. Therefore,

$$V_{t,\theta} = F(\zeta_t)\cdots F(\zeta_{t-k+1})V_{t-k,\theta} + \sum_{i=0}^{k-1} F(\zeta_t)\cdots F(\zeta_{t-i+1})G.$$

Let us first prove that the second term here converges in L^1 and a.s. when $k \to \infty$. Indeed, it is a series of positive terms and $E(F(\zeta_t)\cdots F(\zeta_{t-i+1})G) = F^iG$, with

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

$$F = \begin{pmatrix} A_1 + B_1 & A_2 + B_2 & \cdots & A_n + B_n \\ \mathrm{Id} & & & 0 \\ & \ddots & & \vdots \\ & & \mathrm{Id} & 0 \end{pmatrix}.$$

Thus, $\det(\lambda \mathrm{Id} - F)$ and $\det(\mathrm{Id} - \sum_{i=1}^{n}(A_i + B_i)\lambda^{-i})$ has the same roots. Under B1, it implies that the series converges in L^1 and therefore also a.s. Set

$$\overline{V}_{t,\theta} = \sum_{i=0}^{\infty} F(\zeta_t) \cdots F(\zeta_{t-i+1})G.$$

By independence of the random variables η_t, the process (ζ_t) and, consequently, the process $(\overline{V}_{t,\theta})$ are strictly stationary and ergodic. Moreover, because

$$\overline{V}_{t,\theta} = F(\zeta_t)\overline{V}_{t-1} + G,$$

$\overline{V}_{t,\theta}$ is a strictly stationary solution of (7) and $\overline{V}_{t,\theta}$ is in L^1. Therefore, $\bar{\varepsilon}_{t,i} = \overline{H}_{t,ii}^{1/2}\eta_{t,i}$ is a strictly stationary multivariate GARCH(p,q) process, and is also in L^2.

Let us now prove that this solution is unique. Indeed, if ε_t is another solution, then $V_{t,\theta}$ satisfies

$$V_{t,\theta} = F(\zeta_t) \cdots F(\zeta_{t-k+1})V_{t-k,\theta} + \sum_{i=0}^{k-1} F(\zeta_t) \cdots F(\zeta_{t-i+1})B.$$

The first term converges to 0 in L^1 because

$$F(\zeta_t) \cdots F(\zeta_{t-k+1})V_{t-k,\theta} = F^k E(V_{t-k,\theta}) = c_1 F^k \xrightarrow{k\to\infty} 0,$$

where $c_1 > 0$. Therefore, $\overline{V}_{t,\theta} = V_{t,\theta}$ and, as a consequence, $\bar{\varepsilon}_t = \varepsilon_t$. ■

Proof of Lemma 3.1. We have $\varepsilon_{t,i}^2 = H_{t,ii}\eta_{t,i}^2$, and the first line of this equation is

$$\sum_{i=1}^{d} U_{1i}H_{t,ii}\eta_{t-1,i}^2 = V_1. \tag{8}$$

Set $W_i = U_{1i}H_{t,ii}$ and let μ be the measure of $(W_1,\ldots,W_d,V_1)$. Because η_t is independent of $(W_1,\ldots,W_d,V_1)$, we get, by Fubini's theorem,

$$P\left(\sum_{i=1}^{d} W_i\eta_{t,i}^2 = V_1\right) = \int P\left(\sum_{i=1}^{d} w_i\eta_{t,i}^2 = v_1\right) d\mu(w_1,\ldots,w_d,v_1).$$

Because the left term is equal to 1, then

$$P\left(\sum_{i=1}^{d} w_i\eta_{t,i}^2 = v_1\right) = 1 \ \mu \text{ a.s.}$$

But we have assumed that there is no quadratic form such that $q(\eta_t) = c$. Hence, $w_1 = \cdots = w_d = v_1 = 0$, μ a.s., and then $W_1 = \cdots = W_d = V_1 = 0$, P a.s. Because $H_{t,ii} > 0$, we have, for all i, $U_{1i} = 0$ and $V_1 = 0$. It is also true for the other elements of V. ■

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

Proof of Proposition 3.2. Because ε_t is a stationary solution of both models, it implies that, for $i = 1$ or 2, $\det P_i(L)$ has its roots outside the unit circle. Therefore, $\det P_i(L)$ is a rational fraction and we have

$$\det P_i(L) = \sum_{j=0}^{\infty} c_j L^j \quad \text{for} \quad |L| \leq 1,$$

where c_j converges to 0 with exponential rate. Let us denote by $\widetilde{P}_i(L)$ the cofactor matrix of $P_i(L)$. We have

$$P_i(L)^{-1} = \frac{\widetilde{P}_i(L)}{\det P_i(L)},$$

and $P_i(L)^{-1}Q_i(L)$ can be developed in a series with coefficients converging to 0 with exponential rate. Furthermore, for $i = 1,2$,

$$\begin{pmatrix} H_{t,11}(\theta) \\ \vdots \\ H_{t,dd}(\theta) \end{pmatrix} = P_i(1)^{-1}W + P_i(L)^{-1}Q_i(L)\begin{pmatrix} \varepsilon_{t,1}^2 \\ \vdots \\ \varepsilon_{t,d}^2 \end{pmatrix},$$

and it implies that

$$(P_1(1)^{-1} - P_2(1)^{-1})W + (P_1(L)^{-1}Q_1(L) - P_2(L)^{-1}Q_2(L))\begin{pmatrix} \varepsilon_{t,1}^2 \\ \vdots \\ \varepsilon_{t,d}^2 \end{pmatrix} = 0. \tag{9}$$

Set

$$(P_1(L)^{-1}Q_1(L) - P_2(L)^{-1}Q_2(L)) = \sum_{j=1}^{\infty} \Phi_j L^j,$$

where Φ_j are $d \times d$ matrices. Conditioning (9) by I_{t-j_0-1}, where $j_0 = \inf\{j/\Phi_j \neq 0\}$, we get

$$\Phi_{j_0}\begin{pmatrix} \varepsilon_{t-j_0,1}^2 \\ \vdots \\ \varepsilon_{t-j_0,d}^2 \end{pmatrix} = V,$$

where V is I_{t-j_0-1}-measurable. By Lemma 3.1, it implies that Φ_{j_0} and V are equal to 0. Thus,

$$P_1(L)^{-1}Q_1(L) = P_2(L)^{-1}Q_2(L). \tag{10}$$

Because P_1 and Q_1 are coprime, the Bezout equality (see Theorem 3.1) and (10) yield

$$\begin{aligned} U + P_1^{-1}Q_1V = P_1^{-1} \Rightarrow U + P_2^{-1}Q_2V = P_1^{-1} \Rightarrow P_2U + Q_2V \\ = P_2P_1^{-1} \Rightarrow P_2 = (P_2U + Q_2V)P_1. \end{aligned}$$

Therefore, $P_2 = MP_1$, with $M = P_2U + Q_2V$. Using (10), we get also $Q_2 = MQ_1$. ■

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

Proof of Proposition 3.3. Proposition 3.2 implies that $P_2 = MP_1$ and $Q_2 = MQ_1$. If P_1 is column reduced, because $(MP_1)(0) = \mathrm{Id}$, we have $M(0) = \mathrm{Id}$ and M must be equal to $\mathrm{Id} + LR(L)$, where $R(L)$ is a polynomial matrix. Let us calculate the supremum of the degrees of each column j of MP_1: they can be equal to $d_j(P_1)$ only if $RP_1^{rc} = 0$. Because P_1 is column reduced, P_1^{rc} is a full rank matrix, and this implies that $R(L) = 0$. We can make the same demonstration if Q_1 is column reduced. ■

Proof of Proposition 3.4. First, let us remark that we have already $\beta_0 = \beta$ and that $H_{t,\theta_0} = H_{t,\theta}$ obviously implies that $\rho_{0ij} = \rho_{ij}$. Furthermore, (5) yields

$$\begin{pmatrix} (H_{t,11}(\theta))^2 \\ \vdots \\ (H_{t,dd}(\theta))^2 \end{pmatrix} = W + \sum_{i=1}^{q} A_i \begin{pmatrix} \varepsilon_{t-i,1}^2 \\ \vdots \\ \varepsilon_{t-i,d}^2 \end{pmatrix} + \sum_{i=1}^{p} B_i \begin{pmatrix} (H_{t-i,11}(\theta))^2 \\ \vdots \\ (H_{t-i,dd}(\theta))^2 \end{pmatrix},$$

and, with obvious notations,

$$\begin{pmatrix} (H_{t,11}(\theta_0))^2 \\ \vdots \\ (H_{t,dd}(\theta_0))^2 \end{pmatrix} = W_0 + \sum_{i=1}^{q} A_{0i} \begin{pmatrix} \varepsilon_{t-i,1}^2 \\ \vdots \\ \varepsilon_{t-i,d}^2 \end{pmatrix} + \sum_{i=1}^{p} B_{0i} \begin{pmatrix} (H_{t-i,11}(\theta_0))^2 \\ \vdots \\ (H_{t-i,dd}(\theta_0))^2 \end{pmatrix}.$$

Then, if $H_{t,\theta_0} = H_{t,\theta}$, we get

$$V + \sum_{i=1}^{q} M_i \begin{pmatrix} \varepsilon_{t-i,1}^2 \\ \vdots \\ \varepsilon_{t-i,d}^2 \end{pmatrix} + \sum_{i=1}^{p} M_{q+i} \begin{pmatrix} (H_{t-i,11}(\theta))^2 \\ \vdots \\ (H_{t-i,dd}(\theta))^2 \end{pmatrix} = 0, \tag{11}$$

where V is a d-vector and M_i are $d \times d$ matrices. We must prove that all these terms are equal to 0. First, (11) yields

$$M_1 \begin{pmatrix} \varepsilon_{t-1,1}^2 \\ \vdots \\ \varepsilon_{t-1,d}^2 \end{pmatrix} = U,$$

where U is an I_{t-2}-measurable vector. Therefore, from Lemma 3.1, this implies that both M_1 and U are equal to 0.

Now, because $M_1 = 0$, we have

$$M_{q+1} \begin{pmatrix} (H_{t-1,11}(\theta))^2 \\ \vdots \\ (H_{t-1,dd}(\theta))^2 \end{pmatrix} = -V - \sum_{i=2}^{q} M_i \begin{pmatrix} \varepsilon_{t-i,1}^2 \\ \vdots \\ \varepsilon_{t-i,d}^2 \end{pmatrix} - \sum_{i=2}^{p} M_{q+i} \begin{pmatrix} (H_{t-i,11}(\theta))^2 \\ \vdots \\ (H_{t-i,dd}(\theta))^2 \end{pmatrix}. \tag{12}$$

Let us suppose that P is column reduced; if $M_{q+1} \neq 0$, because P^{rc} is a full rank matrix, we have $M_{q+1}P^{rc} \neq 0$. Therefore, from Proposition 3.3, the left term of (12) must have a formulation with at least one column j with $d_j(P)$ lags, which is in contradiction with the right term of this equation, which has only $d_j(P) - 1$ lags. Therefore, we must have

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press

$M_{q+1} = 0$. The same demonstration holds if Q is column reduced. To end the proof, we iterate the same demonstration for $M_2, M_{q+2}, M_3, \ldots, M_{q+p}$ and then show that V is also equal to 0. ■

Proof of Theorem 3.2. We must verify that A1–A5 hold; we proved A1 and A4 in previous sections, and A5 holds obviously. Furthermore, we have $H_{t,\theta} = \Delta_{t,\theta}\Omega\Delta'_{t,\theta}$. Under B3, the elements of W are greater than $c_1^{1/d}$, and, because the elements of A_i and B_i are positive, it is the same for $H_{ii,t}$. Therefore,

$$\det(\Delta_{t,\theta}) \geq c_1^{1/2}.$$

Because we assumed that $\det\Omega \geq c_2$, we have $\det(H_{t,\theta}) \geq c_1 c_2$, and A2 holds.

For A3, we have

$$\begin{aligned} E_{\theta_0}(\log\det H_{t,\theta_0}) &= E_{\theta_0}(\log\det(\Delta_{t,\theta_0}\Omega\Delta_{t,\theta_0})') \\ &= \log(\det\Omega) + E_{\theta_0}(\log\det(\Delta_{t,\theta_0}\Delta_{t,\theta_0})') \\ &= \log(\det\Omega) + \sum_{i=1}^{d} E_{\theta_0}(\log H_{t,ii}(\theta_0)). \end{aligned}$$

Under B1, we know that $E_{\theta_0}(H_{t,ii}) < \infty$. By Jensen's inequality, we get $E_{\theta_0}((\log\det H_{t,\theta_0})^+) < +\infty$. Furthermore, because $\det(H_{t,\theta_0}) \geq c_1 c_2$, we also have $E_{\theta_0}(|\log(\det H_{t,\theta_0})|) < \infty$. ■

https://doi.org/10.1017/S0266466698141038 Published online by Cambridge University Press