Technical appendix to "Asymptotic properties of QML estimators for VARMA models with time-dependent coefficients, Parts 1 and 2"

Abdelkamel Alj [*], Rajae Azrak [†] and Guy Mélard[‡].

July 13, 2015

Abstract

This technical appendix contains proofs for the asymptotic properties of quasi-maximum likelihood (QML) estimators for vector autoregressive moving average (VARMA) models in the case where the coefficients depend on time instead of being constant. We refer to the main theorems of the paper "Asymptotic properties of QML estimators for VARMA models with time-dependent coefficients, Part 1" (Alj, Ley and Mélard, 2015b) and of another paper in preparation "Asymptotic properties of QML estimators for VARMA models with time-dependent coefficients, Part 2" (Alj, Azrak and Mélard, 2015a). In the latter paper, the coefficients depend on time but also on the number of observations.

Version 10.0

[*]Université Moulay Ismail, Faculté des Sciences juridiques, économiques et sociales, B.P. 3102 Toulal, Meknès, Morocco. This work was started while at Dept of Mathematics, Université libre de Bruxelles CP112, Blvd du Triomphe, B-1050 Bruxelles, Belgium (e-mail: abdelkamel.alj@gmail.com).

[†]Université Mohammed V, Faculté des Sciences juridiques, économiques et sociales, Rabat, Morocco, (e-mail: razrak96@yahoo.fr).

[‡]ECARES, Université libre de Bruxelles CP114/4, Avenue Franklin Roosevelt 50, B-1050 Bruxelles, Belgium (e-mail: gmelard@ulb.ac.be). Corresponding author.

1 Notations

In the proofs below, we frequently use the following notations:

- det(.): the determinant of a matrix.
- $(.)^{-1}$: the inverse of a matrix.
- $(.)^T$: the transpose of a matrix.
- vec(.): the vec-operator, written as vec transforms matrix into a vector, by stacking all the columns of this matrix one underneath the other.
- $\otimes$: the Kronecker product.
- tr(.): the trace of a square matrix and is defined to be the sum of the elements on the main diagonal (the diagonal from the upper left to the lower right).
- $\|.\|_F$: the Schur or Frobenius norm.
- I_m: identity matrix of order m.
- $O(\frac{1}{n})$: of order $\frac{1}{n}$.
- $|.|$: the absolute value.

This is the technical appendix to two papers Alj, Ley and Mélard (2015b) and Alj, Azrak and Mélard (2015a). Both papers will be referred to as "the main paper". The notations are essentially the same.[1] The superscript $^{(n)}$ can be omitted for Alj, Ley and Mélard (2015b).

We have a model depending on a m-dimensional vector of parameters of interest θ for the r-dimensional variable $x_t^{(n)}, t = 1, \ldots, n$. θ^0 is the true value of θ. Denote $F_{t-1}^{(n)}$ the sigma-algebra generated by $\{x_s^{(n)} : s \leq t-1\}$. In the context of the two papers $F_t^{(n)}$ is as a matter of fact generated by $\{\varepsilon_s : s \leq t\}$, which do not depend on n, so that $F_t^{(n)}$ can ne written F_t. Denote the conditional expectation and the residual by

$$\hat{x}_{t|t-1}^{(n)} = E_\theta\left(x_t^{(n)}/F_{t-1}^{(n)}\right), \qquad e_t^{(n)}(\theta) = x_t^{(n)} - \hat{x}_{t|t-1}^{(n)}. \tag{1.1}$$

Denote $\Sigma_t^{(n)}(\theta) = g_t^{(n)}(\theta)\Sigma g_t^{(n)T}(\theta)$ the residual covariance matrix, where the elements of the covariance matrix Σ are nuisance parameters and $g_t^{(n)}(\theta)$ is a $r \times r$ matrix. For $\theta = \theta^0$, we suppose

$$e_t^{(n)}(\theta^0) = g_t^{(n)}\epsilon_t, \tag{1.2}$$

[1]The main difference is that a superscript $^{(n)}$ is added to many symbols for the case of the second paper, Alj, Azrak and Mélard (2015a), n being the series length.

where $g_t^{(n)} = g_t^{(n)}(\theta^0)$ and the ϵ_t are independent random variables with mean 0 and covariance matrix Σ. We will also denote $\Sigma_t^{(n)} = \Sigma_t^{(n)}(\theta^0) = g_t^{(n)} \Sigma g_t^{(n)T}$, not to be confused with Σ. Note that

$$\|\Sigma_t^{(n)}\|_F \leq \|g_t^{(n)}\|_F^2 \|\Sigma\|_F, \tag{1.3}$$

using Lemma 3.4 below.

The model for $x_t^{(n)}$ can be written

$$x_t^{(n)} = e_t^{(n)}(\theta) + \sum_{k=1}^{t-1} \psi_{tk}^{(n)}(\theta) e_{t-k}^{(n)}(\theta). \tag{1.4}$$

We can also write

$$e_t^{(n)}(\theta) = \sum_{k=0}^{t-1} \psi_{t0k}^{(n)}(\theta, \theta^0) g_{t-k}^{(n)} \epsilon_{t-k}. \tag{1.5}$$

Denoting

$$\frac{\partial e_t^{(n)}(\theta)}{\partial \theta_i} = \sum_{k=1}^{t-1} \psi_{tik}^{(n)}(\theta, \theta^0) g_{t-k}^{(n)} \epsilon_{t-k}, \tag{1.6}$$

$$\frac{\partial^2 e_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j} = \sum_{k=1}^{t-1} \psi_{tijk}^{(n)}(\theta, \theta^0) g_{t-k}^{(n)} \epsilon_{t-k}, \tag{1.7}$$

$$\frac{\partial^3 e_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j \partial \theta_l} = \sum_{k=1}^{t-1} \psi_{tijlk}^{(n)}(\theta, \theta^0) g_{t-k}^{(n)} \epsilon_{t-k}, \tag{1.8}$$

for $i, j, l = 1, ..., m$, where the coefficients are defined in the main paper. We will use these coefficients for $\theta = \theta^0$ and they will be denoted respectively by

$$\psi_{tk}^{(n)} = \psi_{tk}^{(n)}(\theta^0), \psi_{tik}^{(n)} = \psi_{tik}^{(n)}(\theta^0, \theta^0), \psi_{tijk}^{(n)} = \psi_{tijk}^{(n)}(\theta^0, \theta^0), \psi_{tijlk}^{(n)} = \psi_{tijlk}^{(n)}(\theta^0, \theta^0). \tag{1.9}$$

Because of (1.2),

$$\psi_{t00}^{(n)} = 1 \text{ and } \psi_{t0k}^{(n)} = 0 \text{ for } k > 0. \tag{1.10}$$

2 Completion of the proof of Theorem 3.1

Remark 2.1 *In the proof of Theorem 3.1 of the main paper we proceed like in Azrak and Mélard (2006). The idea is to check the four assumptions of Theorem 2.1.*

Lemma 2.1 *Under the assumptions of Theorem 3.1, Assumption $H_{2.1}$ of Theorem 2.1-2.2 is true.*

Proof. We have to show that $E_{\theta_0}(\partial \alpha_t^{(n)}(\theta)/\partial \theta_i) \leq C_1$, $i = 1, \ldots, m$, where

$$\alpha_t^{(n)}(\theta) = \log\left\{\det\left(\Sigma_t^{(n)}(\theta)\right)\right\} + e_t^{(n)T}(\theta) \Sigma_t^{(n)-1}(\theta) e_t^{(n)}(\theta). \tag{2.1}$$

Its first derivative is given in Lemma 4.4. Then, the proof is a direct application of Lemma 4.11. □

Lemma 2.2 *Under the assumptions of Theorem 3.1, Assumption* $H_{2.2}$ *of Theorem 2.1-2.2 is true.*

Proof. We have to show that there is a constant C_2 such that

$$E_{\theta^0}\left|\frac{\partial^2\alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j}-E_\theta\left[\frac{\partial^2\alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j}/F_{t-1}\right]\right|^2\leq C_2 \tag{2.2}$$

Note that by using Lemma 4.5, we have

$$\begin{aligned}
&E_{\theta^0}\left|\frac{\partial^2\alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j}-E_\theta\left[\frac{\partial^2\alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j}/F_{t-1}\right]\right|^2\\
&=E_{\theta^0}\left|2\left(e_t^{(n)T}(\theta)\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_j}\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_i}\right)+2\left(e_t^{(n)T}(\theta)\Sigma_t^{(n)-1}(\theta)\frac{\partial^2e_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j}\right)\right.\\
&+2\left(e_t^{(n)T}(\theta)\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i}\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_j}\right)+\left(e_t^{(n)T}(\theta)\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i\partial\theta_j}e_t^{(n)}(\theta)\right)\\
&\left.-\operatorname{tr}\left(\Sigma_t^{(n)}(\theta)\frac{\partial^2\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i\partial\theta_j}\right)\right|^2,\quad i,j=1,...,m.
\end{aligned}\tag{2.3}$$

Then, the result is a direct consequence of Lemma 4.12. □

Lemma 2.3 *The expression after the limit on the left hand side of Assumption* $H_{2.4}$ *in Theorem 2.1 is bounded by*

$$\frac{1}{n}\left|\sum_{t=1}^{n}\left\{\frac{\partial^3\alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_l}\right\}_{\theta=\theta^0}\right|,\qquad\text{for}\quad i,j,l=1,...,m.$$

Proof. By using the mean value theorem we can write, for $i,j=1,...,m$,

$$\left\{\sum_{t=1}^{n}\frac{\partial^2\alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j}\right\}_{\theta=\theta^*}-\left\{\sum_{t=1}^{n}\frac{\partial^2\alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j}\right\}_{\theta=\theta^0}=\sum_{l=1}^{m}(\theta_l^*-\theta_l^0)\left\{\sum_{t=1}^{n}\frac{\partial^3\alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_l}\right\}_{\theta=\theta_{ijl}^{(n)**}},$$

$\theta_{ijl}^{(n)**}$ being a point on the line joining θ^* and θ^0, with $\|\theta^*-\theta^0\|<\Delta$. Taking care that Δ will tend to 0, we have

$$(n\Delta)^{-1}\left|\sum_{t=1}^{n}\left(\left\{\frac{\partial^2\alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j}\right\}_{\theta=\theta^*}-\left\{\frac{\partial^2\alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j}\right\}_{\theta=\theta^0}\right)\right|=$$

$$
\begin{aligned}
&= (n\Delta)^{-1}\left|\sum_l (\theta_l^* - \theta_l^0)\left\{\sum_{t=1}^n \frac{\partial^3 \alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_l}\right\}_{\theta=\theta_{ijl}^{(n)**}}\right| \\
&\leq (n\Delta)^{-1}\|\theta^* - \theta^0\|\left|\left\{\sum_{t=1}^n \frac{\partial^3 \alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_l}\right\}_{\theta=\theta_{ijl}^{(n)**}}\right| \\
&\leq \frac{1}{n}\left|\left\{\sum_{t=1}^n \frac{\partial^3 \alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_l}\right\}_{\theta=\theta_{ijl}^{(n)**}}\right| \\
&\leq \frac{1}{n}\left|\left\{\sum_{t=1}^n \frac{\partial^3 \alpha_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_l}\right\}_{\theta=\theta^0}\right|,
\end{aligned}
$$

by continuity when $\|\theta^* - \theta^0\| \to 0$. □

Remark 2.2 *This proof is identical to a part of the proof of Theorem 2 in Azrak and Mélard (2006, p. 323).*

Lemma 2.4 *The expression $\partial^3\alpha_t^{(n)}(\theta)/\{\partial\theta_i\partial\theta_j\partial\theta_l\}$, $i, j, l = 1, ..., m$, which arises in Lemma 2.3 can be written*

$$
\Phi_{1t}^{(n)}(\theta) + \Phi_{2t}^{(n)}(\theta) + \Psi_{1t}^{(n)}(\theta) + \Psi_{2t}^{(n)}(\theta) + \Psi_{3t}^{(n)}(\theta), \tag{2.4}
$$

where

$$
\begin{aligned}
\Phi_{1t}^{(n)}(\theta) &= \mathrm{tr}\left(\frac{\partial^2\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i\partial\theta_j}\frac{\partial\Sigma_t^{(n)}(\theta)}{\partial\theta_l}\right) + \mathrm{tr}\left(\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_j}\frac{\partial^2\Sigma_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_l}\right) \\
&\quad + \mathrm{tr}\left(\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i}\frac{\partial^2\Sigma_t^{(n)}(\theta)}{\partial\theta_j\partial\theta_l}\right) + \mathrm{tr}\left(\Sigma_t^{(n)-1}(\theta)\frac{\partial^3\Sigma_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_l}\right),
\end{aligned} \tag{2.5}
$$

$$
\Phi_{2t}^{(n)}(\theta) = e_t^{(n)T}(\theta)\frac{\partial^3\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_l}e_t^{(n)}(\theta), \tag{2.6}
$$

$$
\begin{aligned}
\Psi_{1t}^{(n)}(\theta) &= 2e_t^{(n)T}(\theta)\frac{\partial^2\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i\partial\theta_j}\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_l} + 2e_t^{(n)T}(\theta)\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_j}\frac{\partial^2 e_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_l} \\
&\quad + 2e_t^{(n)T}(\theta)\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i}\frac{\partial^2 e_t^{(n)}(\theta)}{\partial\theta_j\partial\theta_l} + 2e_t^{(n)T}(\theta)\Sigma_t^{(n)-1}(\theta)\frac{\partial^3 e_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j\partial\theta_l} \\
&\quad + 2e_t^{(n)T}(\theta)\frac{\partial^2\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i\partial\theta_l}\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_j} + 2e_t^{(n)T}(\theta)\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_l}\frac{\partial^2 e_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j} \\
&\quad + 2e_t^{(n)T}(\theta)\frac{\partial^2\Sigma_t^{(n)-1}(\theta)}{\partial\theta_j\partial\theta_l}\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_i},
\end{aligned} \tag{2.7}
$$

$$\Psi_{2t}^{(n)}(\theta) = 2\left\{ \frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_j} \frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i} \frac{\partial e_t^{(n)}(\theta)}{\partial\theta_l} + \frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i} \frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_j} \frac{\partial e_t^{(n)}(\theta)}{\partial\theta_l} \right.$$
$$\left. + \frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i} \frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_l} \frac{\partial e_t^{(n)}(\theta)}{\partial\theta_j} \right\}, \tag{2.8}$$

and

$$\Psi_{3t}^{(n)}(\theta) = 2\left\{ \frac{\partial^2 e_t^{(n)T}(\theta)}{\partial\theta_i\partial\theta_j} \Sigma_t^{(n)-1}(\theta) \frac{\partial e_t^{(n)}(\theta)}{\partial\theta_l} + \frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_j} \Sigma_t^{(n)-1}(\theta) \frac{\partial^2 e_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_l} \right.$$
$$\left. + \frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i} \Sigma_t^{(n)-1}(\theta) \frac{\partial^2 e_t^{(n)}(\theta)}{\partial\theta_j\partial\theta_l} \right\}. \tag{2.9}$$

Proof. The proof of (2.4)-(2.9) is direct. □

3 Technical Lemmas

Lemma 3.1 *Let $x = \{x_i\}$ represent an $1\times m$ vector of m variables. For any invertible matrix $A(x)$ and continuously differentiable at every x, we have*

$$\frac{\partial \log\{\det(A(x))\}}{\partial x_i} = \operatorname{tr}\left(A^{-1}(x)\frac{\partial A(x)}{\partial x_i}\right),$$

$i = 1,\ldots,m$, and the first derivative of $A^{-1}(x)$ is

$$\frac{\partial A^{-1}(x)}{\partial x_i} = -A^{-1}(x)\frac{\partial A(x)}{\partial x_i}A^{-1}(x).$$

For $i, j = 1,\ldots,m$, the second derivative is

$$\frac{\partial^2 A^{-1}(x)}{\partial x_i\partial x_j} = -A^{-1}(x)\frac{\partial^2 A(x)}{\partial x_i\partial x_j}A^{-1}(x) + A^{-1}(x)\frac{\partial A(x)}{\partial x_i}A^{-1}(x)\frac{\partial A(x)}{\partial x_j}A^{-1}(x)$$
$$+A^{-1}(x)\frac{\partial A(x)}{\partial x_j}A^{-1}(x)\frac{\partial A(x)}{\partial x_i}A^{-1}(x).$$

For the proof, see Harville (1997, pp. 305-309).

Lemma 3.2 *For all real numbers $a_1, a_2, ..., a_n$, n a strictly positive integer, and for all integer powers $p \geq 1$ we have*

$$\left|\sum_{i=1}^{n} a_i\right|^p \leq n^{p-1}\sum_{i=1}^{n}|a_i|^p .$$

See Steele (2004, p. 36).

Lemma 3.3

a) *For any column vector a we have*

$$\operatorname{vec}(a^T) = \operatorname{vec}(a) = a.$$

b) *The basic connection between the vec operator and the Kronecker product is*

$$\operatorname{vec}(ab^T) = b \otimes a,$$

for any two column vectors a and b.

c) *The basic connection between the trace and the vec operator is*

$$\operatorname{tr}(BA) = \operatorname{tr}(AB) = \operatorname{vec}(A^T)^T \operatorname{vec}(B).$$

d) *Let A, B, C and D be four matrices such that AC and BD exist.*

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD),$$

and the transpose of a Kronecker product is

$$(A \otimes B)^T = A^T \otimes B^T.$$

e) *Let A, B and C be three matrices such that the matrix product ABC is defined. Then,*

$$\operatorname{vec}(ABC) = (C^T \otimes A) \operatorname{vec}(B).$$

f) *Let A, B, C and D be four matrices such that the matrix product $ABCD$ is defined and square. Then,*

$$\operatorname{tr}(ABCD) = \operatorname{vec}(A^T)^T (D^T \otimes B) \operatorname{vec}(C).$$

g) *Let A and B be respectively $m \times n$ and $p \times q$ matrices, then $B \otimes A = K_{p,m}(A \otimes B)K_{n,q}$, where e.g. $K_{n,q}$ is an $nq \times nq$ commutation matrix, which is an orthogonal matrix such that for any $n \times q$ matrix C we have $K_{n,q} \operatorname{vec}(C) = \operatorname{vec}(C^T)$, and noting that $K_{1,q} = K_{q,1} = I_q$ and that $K_{n,q}K_{q,n} = I_{nq}$.*

These propositions are well known [see, e.g., Magnus and Neudecker (2007, pp. 32-36 and pp. 54-55)].

Lemma 3.4 *For any two real matrices A and B such that the product $A^T B$ exists, we have*

$$\operatorname{tr}^2(A^T B) \leq \operatorname{tr}(A^T A) \operatorname{tr}(B^T B),$$

and it is equivalent to

$$\operatorname{tr}(A^T B) \leq \|A\|_F \|B\|_F ,$$

with equality if and only if one of the matrices A and B is a multiple of the other. In addition we have

$$\left\|A^T B\right\|_F \leq \|A\|_F \|B\|_F ,$$

and also

$$\|\operatorname{vec}(A)\|_F = \|A\|_F .$$

Finally, using the definition of $K_{n,q}$ in Lemma 3.3 g), $\|K_{n,q}\|_F = \sqrt{nq}$ and $\|A \otimes B\|_F = \|A\|_F \|B\|_F$.

See Magnus and Neudecker (2007, p. 228) and Seber (2008, p. 235).

4 Key Lemmas

Lemma 4.1 *We have for any symmetric $r \times r$ matrix Ω*

$$e_t^{(n)T}(\theta)\Omega\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_i} = \frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i}\Omega e_t^{(n)}(\theta), \quad i = 1, ..., m \tag{4.1}$$

and, for $i, j = 1, ..., m$,

$$E_\theta\left(e_t^{(n)T}(\theta)\Omega\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_i}/F_{t-1}\right) = 0, \tag{4.2}$$

$$E_\theta\left(e_t^{(n)T}(\theta)\Omega\frac{\partial^2 e_t^{(n)}(\theta)}{\partial\theta_i\partial\theta_j}/F_{t-1}\right) = 0. \tag{4.3}$$

Furthermore

$$E_\theta\left(\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i}\Omega\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_j}/F_{t-1}\right) = \frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i}\Omega\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_j}. \tag{4.4}$$

The proof will be given in Appendix 5.1.

Lemma 4.2 *Let ϵ_t be a sequence of independent random vectors with zero mean, covariance matrix Σ and finite fourth-order moments. Then, for $\ell \neq 0$,*

$$E\left(\operatorname{vec}(\epsilon_t\epsilon_{t-\ell}^T)\operatorname{vec}(\epsilon_t\epsilon_{t-\ell}^T)^T\right) = \Sigma \otimes \Sigma, \tag{4.5}$$

$$E\left(\operatorname{vec}(\epsilon_t\epsilon_{t-\ell}^T)\operatorname{vec}(\epsilon_{t-\ell}\epsilon_t^T)^T\right) = K_{r,r}(\Sigma \otimes \Sigma), \tag{4.6}$$

where the $r^2 \times r^2$ matrix $K_{r,r}$ is the commutation matrix that was defined in Lemma 3.3. Let

$$\kappa_t =^{def} E\left(\operatorname{vec}(\epsilon_t\epsilon_t^T)\operatorname{vec}(\epsilon_t\epsilon_t^T)^T\right) = E\left((\epsilon_t\epsilon_t^T)\otimes(\epsilon_t\epsilon_t^T)\right), \tag{4.7}$$

which depends on t, in general. If the ϵ_t are all Gaussian, then κ_t does not depend on t and

$$\kappa_t = \kappa =^{def} \operatorname{vec}(\Sigma)\operatorname{vec}(\Sigma)^T + (I_{r^2} + K_{r,r})(\Sigma\otimes\Sigma). \tag{4.8}$$

The proof will be given in Appendix 5.2.

Remark 4.1 *The following lemma will be used several times in rather different contexts so we use more general notations than needed: $\partial^{(q)}e_t^{(n)}(\theta)$ can denote $e_t^{(n)}(\theta)$ for $q=0$, a first-order derivative $\partial e_t^{(n)}(\theta)/\partial\theta_i$ if $q=1$, a second-order derivative $\partial^2 e_t^{(n)}(\theta)/(\partial\theta_i\partial\theta_j)$ if $q=2$, or even a third-order derivative $\partial^3 e_t^{(n)}(\theta)/(\partial\theta_i\partial\theta_j\partial\theta_l)$ if $q=3$, $i,j,l=1,\ldots,m$. In (4.9) below, $\psi_{tk}^{(n)(q)}(\theta,\theta^0)$ will denote the coefficients $\psi_{t0k}^{(n)}(\theta,\theta^0)$ in (1.5), for $q=0$, the coefficients $\psi_{tik}^{(n)}(\theta,\theta^0)$ in (1.6), for $q=1$, the coefficients $\psi_{tijk}^{(n)}(\theta,\theta^0)$ in (1.7), for $q=2$, the coefficients $\psi_{tijlk}^{(n)}(\theta,\theta^0)$ in (1.8), for $q=3$. Note also that $\psi_{t0}^{(n)(q)}(\theta,\theta^0)=0$ for $q=1,2,3$, according to (1.6)-(1.8). See Remark 4.2 for hints on applications of that lemma.*

Lemma 4.3 *Let $\partial^{(q)}$ denote a q-th order derivative with respect to some component of θ, $S_t^{(n)}(\theta)$ be a square symmetric matrix, let q_1, q_2, q_3, q_4, and d be positive integers, ν_1 and ν_2 be integers. Let*

$$\partial_\theta^{(q)}e_t^{(n)}(\theta) = \sum_{k=0}^{t-1}\psi_{tk}^{(n)(q)}(\theta,\theta^0)g_{t-k}^{(n)}\epsilon_{t-k}, \tag{4.9}$$

where the $\psi_{tk}^{(n)(q)}(\theta,\theta^0)$ and $g_t^{(n)}$ are matrix coefficients and the ϵ_t are independent random vectors with zero mean, covariance matrix Σ and finite fourth-order moments as in Lemma 4.2. We denote $\psi_{tk}^{(n)(q)} = \psi_{tk}^{(n)(q)}(\theta^0,\theta^0)$. Then

$$E_{\theta^0}\left[\partial_\theta^{(q_1)}e_t^{(n)T}(\theta)S_t^{(n)}(\theta)\partial_\theta^{(q_2)}e_{t+d}^{(n)}(\theta)\right] = \sum_{k=0}^{t-1}\operatorname{tr}\{g_{t-k}^{(n)T}\psi_{tk}^{(n)(q_1)T}S_t^{(n)}\psi_{t+d,k+d}^{(n)(q_2)}g_{t-k}^{(n)}\Sigma\}, \tag{4.10}$$

where $S_t^{(n)} = S_t^{(n)}(\theta^0)$, and, with $\tilde\delta = 0$ or 1,

$$E_{\theta^0}\Big[E_\theta\left\{\partial_\theta^{(q_1)}e_t^{(n)T}(\theta)S_t^{(n)}(\theta)\partial_\theta^{(q_2)}e_t^{(n)}(\theta) - \tilde\delta E_\theta\left(\partial_\theta^{(q_1)}e_t^{(n)T}(\theta)S_t^{(n)}(\theta)\partial_\theta^{(q_2)}e_t^{(n)}(\theta)\right)/F_{t-\nu_1}\right\}.$$
$$E_\theta\left\{\partial_\theta^{(q_3)}e_{t+d}^{(n)T}(\theta)S_{t+d}^{(n)}(\theta)\partial_\theta^{(q_4)}e_{t+d}^{(n)}(\theta) - \tilde\delta E_\theta\left(\partial_\theta^{(q_3)}e_{t+d}^{(n)T}(\theta)S_{t+d}^{(n)}(\theta)\partial_\theta^{(q_4)}e_{t+d}^{(n)}(\theta)\right)/F_{t-\nu_2}\right\}\Big] \tag{4.11}$$

$$= \sum_{k=\max(\nu_1,\nu_2,0)}^{t-1} M_{t0kk}^{(n)(q_2,q_1)T} \Xi_{t-k} M_{tdkk}^{(n)(q_3,q_4)} \tag{4.12}$$

$$+ \sum_{k_1,k_2=\max(\nu_1,\nu_2,0)}^{t-1}\sum^{t-1} M_{t0k_2k_1}^{(n)(q_2,q_1)T} K_{r,r}(\Sigma \otimes \Sigma) M_{tdk_1k_2}^{(n)(q_3,q_4)} \tag{4.13}$$

$$+ \sum_{k_1,k_2=\max(\nu_1,\nu_2,0)}^{t-1}\sum^{t-1} M_{t0k_2k_1}^{(n)(q_2,q_1)T} (\Sigma \otimes \Sigma) M_{tdk_2k_1}^{(n)(q_3,q_4)} \tag{4.14}$$

$$+(1-\tilde{\delta}) \left\{ \sum_{k_1=\max(\nu_1,0)}^{t-1} M_{t0k_1k_1}^{(n)(q_2,q_1)T} \right\} (\operatorname{vec}\Sigma)(\operatorname{vec}\Sigma)^T \left\{ \sum_{k_3=\max(\nu_2,-d)}^{t-1} M_{tdk_3k_3}^{(n)(q_3,q_4)} \right\}. \tag{4.15}$$

where we have the $r^2 \times 1$ vector

$$M_{tfk'k''}^{(n)(q',q'')} = \operatorname{vec}(g_{t-k'}^{(n)T} \psi_{t+f,k'+f}^{(n)(q')T} S_{t+f}^{(n)} \psi_{t+f,k''+f}^{(n)(q'')} g_{t-k''}^{(n)}), \tag{4.16}$$

with $q', q'' = q_1, q_2, q_3, q_4$, and $k', k'' = k, k_1, k_2$, and the $r^2 \times r^2$ matrix

$$\Xi_t = \kappa_t - \{K_{r,r}(\Sigma \otimes \Sigma) + (\Sigma \otimes \Sigma) + (\operatorname{vec}\Sigma)(\operatorname{vec}\Sigma)^T\}. \tag{4.17}$$

In the special case where $q_1 = q_2 = q_3 = q_4 = 0$, $d = 0$, and $\nu_1 = \nu_2 = 0$, (4.12)-(4.15) *simplify to*

$$M_{t000}^{(n)(0,0)T} \{\kappa_t - (\operatorname{vec}\Sigma)(\operatorname{vec}\Sigma)^T\} M_{t000}^{(n)(0,0)}, \tag{4.18}$$

where $M_{t000}^{(n)(0,0)} = \operatorname{vec}(g_t^{(n)T} S_t^{(n)} g_t^{(n)})$.

The proof will be given in Appendix 5.3.

Remark 4.2 *We will use the first part of this Lemma in Lemmas 4.7, 4.8 and 4.22. We will use the second part of this Lemma in Lemmas 4.7, 4.9, 4.11, 4.13, 4.15, 4.16, 4.18, and 4.20. Lemma 4.7 will be used to deliver upper bounds for general q_i's when $d = 0$ and $\nu_1 = \nu_2 = \nu$, and will no longer be discussed here. For the first part of the present Lemma, we use $q_1 = q_2 = 1$ and also $q_1 = q_2 = 2$ in Lemma 4.8. In general we have $q_1 = q_3 = 1$ and $q_2 = q_4 = 1$ and the component of θ involved in the derivative will be denoted respectively θ_i and θ_j except in Lemmas 4.15, 4.16, and 4.20. In Lemma 4.15, $q_1 = q_2 = q_3 = q_4 = 0$ and no derivative of $e_t(\theta)$ is involved. In Lemma 4.16, $q_1 = q_3 = 0$ and $q_2 = q_4 = 1, 2$ or 3. In Lemma 4.20 where $q_1 = q_3 = 2$ and $q_2 = q_4 = 1$ and the component of θ involved in the derivative will be denoted θ_i, θ_j for the former, and θ_l, for the latter. In all cases except Lemma 4.9, we have $d = 0$. In Lemma 4.9, we use $\nu_1 = 1$ and $\nu_2 = 1 - d$. In Lemmas 4.16, 4.18, and 4.20 we have an arbitrary $\nu > 0$ and $d = 0$ so $\nu_1 = \nu_2 = \nu$, and the summations start at ν. Otherwise $\nu_1 = \nu_2 = 1$, in Lemmas 4.11 and 4.13, or $\nu_1 = \nu_2 = 0$, in Lemma 4.15. In Lemmas 4.9, 4.13, and 4.20, $S_t^{(n)} = \Sigma^{(n)-1}(\theta)$ in Lemma 4.18 it is $\partial_{\theta_l}\Sigma^{(n)-1}(\theta)$ and in Lemma 4.16, it is either $\Sigma^{(n)-1}(\theta)$ or a first- or second-order derivative of it. In*

Lemma 4.15, it is a third derivative of $\Sigma^{(n)-1}(\theta)$*, whereas in Lemmas 4.8 and 4.11, it is* I_r*, and in Lemma 4.22, it is a product* $K_{ti}^{(n)}K_{t+d,i}^{(n)T}$*, where* $K_{ti}^{(n)}$ *is defined in* (4.39)*. In Lemmas 4.11 and 4.13,* $\tilde{\delta} = 0$ *and in the other cases we have* $\tilde{\delta} = 1$ *so that the term* (4.15) *cancels.*

Lemma 4.4 *Let* $\alpha_t^{(n)}(\theta)$ *as defined in Assumption* $H_{2.1}$ *of the main paper. Then*

$$\frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_i} = a_1^{(n)i}(\theta) + a_2^{(n)i}(\theta) + a_3^{(n)i}(\theta), \tag{4.19}$$

where

$$a_1^{(n)i}(\theta) = \mathrm{tr}\left(\Sigma_t^{(n)-1}(\theta)\frac{\partial \Sigma_t^{(n)}(\theta)}{\partial \theta_i}\right), \tag{4.20}$$

$$a_2^{(n)i}(\theta) = 2e_t^{(n)T}(\theta)\{\Sigma_t^{(n)}(\theta)\}^{-1}\frac{\partial e_t^{(n)}(\theta)}{\partial \theta_i}, \tag{4.21}$$

$$a_3^{(n)i}(\theta) = e_t^{(n)T}(\theta)\frac{\partial \Sigma_t^{(n)-1}(\theta)}{\partial \theta_i}e_t^{(n)}(\theta). \tag{4.22}$$

We have for all t *that* $\{\partial \alpha_t^{(n)}(\theta)/\partial \theta_i, F_t\}$*,* $i = 1, ..., m$*, is such that*

$$E_\theta\left(\frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_i}/F_{t-1}\right) = 0. \tag{4.23}$$

The proof will be given in Appendix 5.4.

Lemma 4.5 *We have for all* t *that*

$$\left\{\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j} - E_\theta\left(\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j}/F_{t-1}\right), F_t\right\} \tag{4.24}$$

$i, j = 1, ..., m$*, is such that*

$$E_\theta\left(\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j} - E_\theta\left(\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j}/F_{t-1}\right)/F_{t-1}\right) = 0. \tag{4.25}$$

Furthermore

$$E\left(\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j}/F_{t-1}\right) = 2\frac{\partial e_t^{(n)T}(\theta)}{\partial \theta_i}\Sigma_t^{(n)-1}(\theta)\frac{\partial e_t^{(n)}(\theta)}{\partial \theta_j} + \mathrm{tr}\left[\Sigma_t^{(n)-1}(\theta)\frac{\partial \Sigma_t^{(n)}(\theta)}{\partial \theta_i}\Sigma_t^{(n)-1}(\theta)\frac{\partial \Sigma_t^{(n)}(\theta)}{\partial \theta_j}\right]. \tag{4.26}$$

The proof will be given in Appendix 5.5.

Lemma 4.6 *Under the assumptions of Theorem 3.1 of the main paper,*

$$E_{\theta^0}\left(\left\{e_t^{(n)T}(\theta)e_t^{(n)}(\theta)\right\}^4\right) \le m_1^4 M_1. \tag{4.27}$$

Furthermore

$$E_{\theta^0}\left(e_t^{(n)T}(\theta)e_t^{(n)}(\theta)\right)^2 \le m_1^2 M_1^{1/2}, \quad \text{and} \quad |E_{\theta^0}\left(e_t^{(n)T}(\theta)e_t^{(n)}(\theta)\right)| \le r m_1 \|\Sigma\|_F. \tag{4.28}$$

The proof will be given in Appendix 5.6.

Lemma 4.7 *With the notations of Lemma 4.3 for $d = 0$, $q_1 = q_3$, $q_2 = q_4$, and $\nu_1 = \nu_2 = \nu \ge 1$ under the assumptions of Theorem 3.1, denoting $P(\nu)$ a polynomial in ν such that $P(1) = 1$, and assuming that*

$$\sum_{k=\nu}^{t-1} \left\|\psi_{tk}^{(n)(q_i)}\right\|_F^2 < N_i' P(\nu)\Phi^{\nu-1}, \quad i = 1, 2, \tag{4.29}$$

$$\sum_{k=\nu}^{t-1} \left\|\psi_{tk}^{(n)(q_i)}\right\|_F^4 < N_i'' P(\nu)\Phi^{\nu-1}, \quad i = 1, 2, \tag{4.30}$$

$$\|S_t\|_F < \widetilde{m}, \tag{4.31}$$

for all t, then (4.10) *is bounded from above by*

$$m_1 \widetilde{m} (N_1' N_2')^{1/2} \|\Sigma\|_F, \tag{4.32}$$

and (4.11) *is bounded from above by*

$$m_1^2 \widetilde{m}^2 \left\{ M_3 (N_1'' N_2'')^{1/2} P(\nu)\Phi^{\nu-1} + (r + 2 - \tilde{\delta}) N_1' N_2' P^2(\nu)\Phi^{2(\nu-1)} \|\Sigma\|_F^2 \right\}. \tag{4.33}$$

In the case where $q_1 = q_2 = q_3 = q_4 = 0$ and $\nu_1 = \nu_2 = 0$, (4.11) *is bounded by*

$$m_1^2 \widetilde{m}^2 M_3. \tag{4.34}$$

The proof will be given in Appendix 5.7.

Lemma 4.8 *Under the assumptions of Theorem 3.1 and $i, j = 1, ..., m$:*

$$\left| E_{\theta^0}\left(\frac{\partial e_t^{(n)T}(\theta)}{\partial \theta_i} \frac{\partial e_t^{(n)}(\theta)}{\partial \theta_i}\right)\right| \le m_1 N_1 \|\Sigma\|_F, \tag{4.35}$$

$$\left| E_{\theta^0}\left(\frac{\partial^2 e_t^{(n)T}(\theta)}{\partial \theta_i \partial \theta_j} \frac{\partial^2 e_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j}\right)\right| \le m_1 N_3 \|\Sigma\|_F. \tag{4.36}$$

The proof will be given in Appendix 5.8.

Lemma 4.9 *We have*

$$\mathrm{cov}_{\theta^0}\left(\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i}\Sigma_t^{(n)-1}(\theta)\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_j}, \frac{\partial e_{t+d}^{(n)T}(\theta)}{\partial\theta_i}\Sigma_{t+d}^{(n)-1}(\theta)\frac{\partial e_{t+d}^{(n)}(\theta)}{\partial\theta_j}\right)$$

$$\begin{aligned} &= \sum_{k=1}^{t-1} M_{t0kk}^{(n)jiT}\Xi_{t-k}M_{tdkk}^{(n)ij} \\ &\quad + \sum_{k_1=1}^{t-1}\sum_{k_2=1}^{t-1} M_{t0k_2k_1}^{(n)jiT}K_{r,r}(\Sigma\otimes\Sigma)M_{tdk_1k_2}^{(n)ij} \\ &\quad + \sum_{k_1=1}^{t-1}\sum_{k_2=1}^{t-1} M_{t0k_2k_1}^{(n)jiT}(\Sigma\otimes\Sigma)M_{tdk_2k_1}^{(n)ij}. \end{aligned} \tag{4.37}$$

for $i, j = 1, ..., m,$ *with* $k', k'' = k, k_1, k_2,$

$$M_{tfk'k''}^{(n)ij} = \mathrm{vec}(g_{t-k'}^{(n)T}\psi_{t+f,i,k'+f}^{(n)T}\Sigma_{t+f}^{(n)-1}\psi_{t+f,j,k''+f}^{(n)}g_{t-k''}^{(n)}).$$

The proof will be given in Appendix 5.9.

Lemma 4.10 *We have*

$$E_{\theta^0}\left(\frac{\partial\alpha_t^{(n)}(\theta)}{\partial\theta_i}\frac{\partial\alpha_t^{(n)}(\theta)}{\partial\theta_j}/F_{t-1}\right) - E_{\theta^0}\left(\frac{\partial\alpha_t^{(n)}(\theta)}{\partial\theta_i}\frac{\partial\alpha_t^{(n)}(\theta)}{\partial\theta_j}\right) =$$

$$\begin{aligned} =\ & 4\left[\left\{\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_j}\Sigma_t^{(n)-1}(\theta)\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_i}\right\}_{\theta=\theta^0} - E_{\theta^0}\left(\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_j}\Sigma_t^{(n)-1}(\theta)\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_i}\right)\right] \\ & +2\left[\left\{\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i}\right\}_{\theta=\theta^0} - E_{\theta^0}\left(\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i}\right)\right]K_{tj}^{(n)} \\ & +2\left[\left\{\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_j}\right\}_{\theta=\theta^0} - E_{\theta^0}\left(\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_j}\right)\right]K_{ti}^{(n)}. \end{aligned} \tag{4.38}$$

with

$$K_{ti}^{(n)} = \Sigma_t^{(n)-1}g_t^{(n)}E\left(\epsilon_t^{\otimes 3}\right)\left(g_t^{(n)T}\otimes g_t^{(n)T}\right)\mathrm{vec}\left(\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_i}\right)_{\theta=\theta^0} \tag{4.39}$$

and

$$E\left(\epsilon_t^{\otimes 3}\right) = E\left(\epsilon_t\epsilon_t^T\otimes\epsilon_t^T\right). \tag{4.40}$$

The proof of this Lemma will be given in Appendix 5.10.

Lemma 4.11 *Under the assumptions of Theorem 3.1 of the main paper, the assumption $H_{2.1}$ of the main paper is satisfied, which means that there exists a positive constant C_1 such that for all t*

$$E_{\theta^0}\left\{\left|\frac{\partial \alpha_t^{(n)}(\theta)}{\partial \theta_i}\right|^4\right\} < C_1.$$

The proof of this Lemma will be given in Appendix 5.11.

Lemma 4.12 *Let us consider*

$$E_{\theta^0}\left|\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j} - E_\theta\left[\frac{\partial^2 \alpha_t^{(n)}(\theta)}{\partial \theta_i \partial \theta_j}/F_{t-1}\right]\right|^2,$$

as defined in the left hand side of (2.2). *Then, under the assumptions of Theorem 3.1 of the main paper, that expression is bounded by $C_2 = \max(C_2^{(1)}, C_2^{(2)}, C_2^{(3)})$ where $C_2^{(1)}, C_2^{(2)}$ and $C_2^{(3)}$ are given by (5.43), (5.44) and (5.45) respectively.*

The proof of this Lemma will be given in Appendix 5.12.

Lemma 4.13 *Under the assumptions of Theorem 3.1 of the main paper, the assumptions i and ii of Lemma A.1 of the main paper are satisfied for $Z_{tij}^{(n)}$ defined by*

$$Z_{tij}^{(n)} = \left\{\frac{\partial e_t^{(n)T}(\theta)}{\partial \theta_i}\Sigma_t^{(n)-1}(\theta)\frac{\partial e_t^{(n)}(\theta)}{\partial \theta_j}\right\}_{\theta=\theta^0} - E_{\theta^0}\left(\frac{\partial e_t^{(n)T}(\theta)}{\partial \theta_i}\Sigma_t^{(n)-1}(\theta)\frac{\partial e_t^{(n)}(\theta)}{\partial \theta_j}\right), \tag{4.41}$$

$i, j = 1, ..., m$.

The proof of this Lemma will be given in Appendix 5.13.

Lemma 4.14 *Let $\Phi_{1t}^{(n)}(\theta)$ defined by (2.5). Then under the assumptions $H_{3.3}$-$H_{3.5}$ of the main paper, we have $\Phi_{1t}^{(n)}(\theta^0) \leq \tilde{\Phi}_1$ with $\tilde{\Phi}_1 = K_5^{1/2}K_1^{1/2} + 2K_4^{1/2}K_2^{1/2} + m_2^{1/2}K_3^{1/2}$.*

The proof of this Lemma will be given in Appendix 5.14.

Lemma 4.15 *Let $\Phi_{2t}^{(n)}(\theta)$ defined by (2.6). Defining*

$$Z_t^{(n)} = \Phi_{2t}^{(n)}(\theta^0) - E_{\theta^0}\Phi_{2t}^{(n)}(\theta), \tag{4.42}$$

under the assumptions $H_{3.3}$-$H_{3.5}$ of the main paper, there exists a constant C_3 such that $E(Z_t^{(n)2}) < C_3$.

The proof of this Lemma will be given in Appendix 5.15.

Lemma 4.16 *Using the notation* $\partial^{(q)}e_t^{(n)}(\theta)$ *introduced in* (4.9) *in Lemma 4.3 and* (1.5)*, we consider* $W_t^{(n)(q)}$ *defined by*

$$W_t^{(n)(q)} = \left(e_t^{(n)T}(\theta)S_t^{(n)}(\theta)\partial^{(q)}e_t^{(n)}(\theta)\right)_{\theta=\theta^0} - E_{\theta^0}\left[e_t^{(n)T}(\theta)S_t^{(n)}(\theta)\partial^{(q)}e_t^{(n)}(\theta)\right], \quad (4.43)$$

where $S_t^{(n)}(\theta)$ *is a square symmetric matrix. Then we have*

$$\begin{aligned} E\{E(W_t^{(n)(q)}/F_{t-\nu})^2\} &= \sum_{k=\nu}^{t-1} M_{tkk}^{(n)(q0)T}\Xi_{t-k}M_{tkk}^{(n)(0q)} \\ &+ \sum_{k_1=\nu}^{t-1}\sum_{k_2=\nu}^{t-1} M_{tk_1k_2}^{(n)(q0)T}K_{r,r}(\Sigma\otimes\Sigma)M_{tk_1k_2}^{(n)(0q)} \\ &+ \sum_{k_1=\nu}^{t-1}\sum_{k_2=\nu}^{t-1} M_{tk_1k_2}^{(n)(q0)T}(\Sigma\otimes\Sigma)M_{tk_2k_1}^{(n)(0q)}, \end{aligned} \quad (4.44)$$

where $M_{tk'k''}^{(n)(0q)}$ *and* $M_{tk'k''}^{(n)(q0)}$ *are defined by*

$$M_{tk'k''}^{(n)(0q)} = \mathrm{vec}(g_{t-k'}^{(n)T}\psi_{t0k'}^{(n)T}S_t^{(n)}\psi_{tk''}^{(n)(q)}g_{t-k''}^{(n)}), \quad M_{tk'k''}^{(n)(q0)} = \mathrm{vec}(g_{t-k'}^{(n)T}\psi_{tk'}^{(n)(q)T}S_t^{(n)}\psi_{t0k''}^{(n)}g_{t-k''}^{(n)}),$$

where $q', q'' = q, 0$*, and* $k', k'' = k, k_1, k_2$*.*

The proof of this Lemma will be given in Appendix 5.16.

Lemma 4.17 *We consider* $W_t^{(n)(q)}$ *defined in (4.43). Then, under the assumptions* $H_{3.3}$*-*$H_{3.5}$ *of the main paper,* $\{W_t^{(n)(q)}, F_t\}$ *satisfies the two conditions i and ii of Definition A.1 and the assumptions of Lemma A.2 in the main paper.*

The proof of this Lemma will be given in Appendix 5.17.

Lemma 4.18 *We consider* $X_t^{(n)ilj}$ *defined by*

$$X_t^{(n)ilj} = \left(\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i}\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_l}\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_j}\right)_{\theta=\theta^0} - E_{\theta^0}\left[\frac{\partial e_t^{(n)T}(\theta)}{\partial\theta_i}\frac{\partial\Sigma_t^{(n)-1}(\theta)}{\partial\theta_l}\frac{\partial e_t^{(n)}(\theta)}{\partial\theta_j}\right], \quad (4.45)$$

then we have

$$\begin{aligned} E\{E(X_t^{(n)ilj}/F_{t-\nu})^2\} &= \sum_{k=\nu}^{t-1} M_{tkk}^{(n)jliT}\Xi_{t-k}M_{tkk}^{(n)ilj} \\ &+ \sum_{k_1=\nu}^{t-1}\sum_{k_2=\nu}^{t-1} M_{tk_2k_1}^{(n)jliT}K_{r,r}(\Sigma\otimes\Sigma)M_{tk_1k_2}^{(n)ilj} \\ &+ \sum_{k_1=\nu}^{t-1}\sum_{k_2=\nu}^{t-1} M_{tk_2k_1}^{(n)jliT}(\Sigma\otimes\Sigma)M_{tk_2k_1}^{(n)ilj}, \end{aligned} \quad (4.46)$$

where $M^{(n)ilj}_{tk_1k_2}$ *is defined by*

$$M^{(n)ilj}_{tk'k''} = \mathrm{vec}\left(g^{(n)T}_{t-k'}\psi^{(n)T}_{tik'}\left\{\frac{\partial\Sigma^{(n)-1}_t(\theta)}{\partial\theta_l}\right\}_{\theta=\theta^0}\psi^{(n)}_{tjk''}g^{(n)}_{t-k''}\right), \tag{4.47}$$

with $k', k'' = k, k_1, k_2$.

The proof will be given in Appendix 5.18.

Lemma 4.19 *We consider* $X^{(n)ilj}_t$ *defined in (4.45). Then, under the assumptions* $H_{3.3}$*-*$H_{3.5}$ *of the main paper,* $\{X^{(n)ilj}_t, F_t\}$ *satisfies the two conditions i and ii of Definition A.1 and the assumptions of Lemma A.2 in the main paper.*

The proof of this Lemma will be given in Appendix 5.19.

Lemma 4.20 *We consider* $Y^{(n)ijl}_t$ *defined by*

$$Y^{(n)ijl}_t = \left(\frac{\partial^2 e^{(n)T}_t(\theta)}{\partial\theta_i\partial\theta_j}\Sigma^{(n)-1}_t(\theta)\frac{\partial e^{(n)}_t(\theta)}{\partial\theta_l}\right)_{\theta=\theta^0} - E_{\theta^0}\left[\frac{\partial^2 e^{(n)T}_t(\theta)}{\partial\theta_i\partial\theta_j}\Sigma^{(n)-1}_t(\theta)\frac{\partial e^{(n)}_t(\theta)}{\partial\theta_l}\right], \tag{4.48}$$

then we have

$$\begin{aligned} E\{E(Y^{(n)ijl}_t/F_{t-\nu})^2\} &= \sum_{k=\nu}^{t-1} M^{(n)l,ijT}_{tkk}\Xi_{t-k}M^{(n)ij,l}_{tkk} \\ &+ \sum_{k_1=\nu}^{t-1}\sum_{k_2=\nu}^{t-1} M^{(n)l,ijT}_{tk_2k_1}K_{r,r}(\Sigma\otimes\Sigma)M^{(n)ij,l}_{tk_1k_2} \\ &+ \sum_{k_1=\nu}^{t-1}\sum_{k_2=\nu}^{t-1} M^{(n)l,ijT}_{tk_2k_1}(\Sigma\otimes\Sigma)M^{(n)ij,l}_{tk_2k_1}, \end{aligned} \tag{4.49}$$

with $M^{(n)ij,l}_{tk_1k_2}$ *and* $M^{(n)l,ij}_{tk_1k_2}$ *are defined by*

$$M^{(n)ij,l}_{tk'k''} = \mathrm{vec}\left(g^{(n)T}_{t-k'}\psi^{(n)T}_{tijk'}\Sigma^{(n)-1}_t\psi^{(n)}_{tlk''}g^{(n)}_{t-k''}\right), M^{(n)l,ij}_{tk'k''} = \mathrm{vec}\left(g^{(n)T}_{t-k'}\psi^{(n)T}_{tlk'}\Sigma^{(n)-1}_t\psi^{(n)}_{tijk''}g^{(n)}_{t-k''}\right), \tag{4.50}$$

with $k', k'' = k, k_1, k_2$.

The proof will be given in Appendix 5.20.

Lemma 4.21 *We consider* $Y^{(n)ijl}_t$ *defined in (4.48). Then, under the assumptions* $H_{3.3}$*-*$H_{3.5}$ *of the main paper,* $\{Y^{(n)ijl}_t, F_t\}$ *satisfies the two conditions i and ii of Definition A.1 and the assumptions of Lemma A.2 in the main paper.*

The proof of this Lemma will be given in Appendix 5.21.

Lemma 4.22 *Under the assumptions of Theorem 3.1 of the main paper, the two assumptions i and ii of Lemma A.1 of the main paper are satisfied for $\tilde{Z}_{tij}^{(n)}$ defined by*

$$\tilde{Z}_{tij}^{(n)} = \left(\left\{ \frac{\partial e_t^{(n)T}(\theta)}{\partial \theta_j} \right\}_{\theta=\theta^0} - E_{\theta^0} \left(\frac{\partial e_t^{(n)T}(\theta)}{\partial \theta_j} \right) \right) K_{ti}^{(n)}, \tag{4.51}$$

$i, j = 1, ..., m$, where $K_{ti}^{(n)}$ is defined by (4.39).

The proof of this Lemma will be given in Appendix 5.22.

5 Proof of Lemmas

For the proofs, we will use simplified notations for the partial derivatives, like ∂_i instead of $\partial/\partial\theta_i$ and ∂_{ij} instead of $\partial^2/\partial\theta_i\partial\theta_j$. To simplify the notations, all suerscripts $^{(n)}$ are omitted in the proofs. Anyway, in the assumptions of the main paper Alj et al. (2015a) which involve upper bounds, these upper bounds are uniform in n so that mentioning n is superfluous.

5.1 Proof of Lemma 4.1

i. (4.1) holds because the left hand side of the equality is a scalar and the matrix Ω is symmetric.

ii. We can show (4.2) by using the definition (1.1), noting that x_t does not depend on θ and that $\partial_i \hat{x}_{t/t-1}(\theta) \in F_{t-1}$:

$$\begin{aligned} E_\theta \left(e_t^T(\theta)\Omega\partial_i e_t(\theta)/F_{t-1} \right) &= E_\theta \left(e_t^T(\theta)\Omega\partial_i\{x_t - \hat{x}_{t/t-1}(\theta)\}/F_{t-1} \right) \\ &= -E_\theta \left(e_t^T(\theta)\Omega\partial_i\hat{x}_{t/t-1}(\theta)/F_{t-1} \right) \\ &= -E_\theta \left(e_t^T(\theta)/F_{t-1} \right) \Omega\partial_i\hat{x}_{t/t-1}(\theta) = 0. \end{aligned}$$

The proof is similar for (4.3).

iii. For (4.4) we have like in the proof of (4.2)

$$\begin{aligned} E_\theta \left(\partial_i e_t^T(\theta)\Omega\partial_j e_t(\theta)/F_{t-1} \right) &= E_\theta \left(\partial_i\{x_t - \hat{x}_{t/t-1}(\theta)\}^T\Omega\partial_j\{x_t - \hat{x}_{t/t-1}(\theta)\}/F_{t-1} \right) \\ &= E_\theta \left(\partial_i\hat{x}_{t/t-1}^T(\theta)\Omega\partial_j\hat{x}_{t/t-1}(\theta)/F_{t-1} \right) \\ &= \partial_i\hat{x}_{t/t-1}^T(\theta)\Omega\partial_j\hat{x}_{t/t-1}(\theta) \\ &= \partial_i e_t^T(\theta)\Omega\partial_j e_t(\theta). \qquad \square \end{aligned}$$

5.2 Proof of Lemma 4.2

The parts (4.5) and (4.6) of the lemma are a special case of the following results: if X and Y are independent random vectors, of respective dimensions n and q, with vector mean 0 and respective covariance matrices Σ_X and Σ_Y, then

$$\begin{aligned} E\left(\operatorname{vec}(XY^T)\operatorname{vec}(XY^T)^T\right) &= (\Sigma_Y \otimes \Sigma_X), \\ E\left(\operatorname{vec}(YX^T)\operatorname{vec}(XY^T)^T\right) &= K_{n,q}(\Sigma_Y \otimes \Sigma_X). \end{aligned}$$

We prove the first assertion by using Lemma 3.3 b) and d)

$$\operatorname{vec}(XY^T)\operatorname{vec}(XY^T)^T = (Y \otimes X)(Y^T \otimes X^T) = (YY^T) \otimes (XX^T),$$

hence the expectation is $\Sigma_Y \otimes \Sigma_X$. Similarly, using also Lemma 3.3 g) and d),

$$\begin{aligned} \operatorname{vec}(YX^T)\operatorname{vec}(XY^T)^T &= (X \otimes Y)(Y^T \otimes X^T) \\ &= K_{n,q}(Y \otimes X)(Y^T \otimes X^T) \\ &= K_{n,q}(YY^T) \otimes (XX^T), \end{aligned}$$

hence its expectation is $K_{n,q}(\Sigma_Y \otimes \Sigma_X)$. Finally (4.8) is a consequence of Kollo and von Rosen (2005, p. 207). □

5.3 Proof of Lemma 4.3

For the first part, the left hand side of (4.10) is a scalar, so it is equal to its trace. Replacing the derivatives using (4.9) leads to

$$\begin{aligned} &\partial_\theta^{(q_1)} e_t^T(\theta) S_t(\theta) \partial_\theta^{(q_2)} e_{t+d}(\theta) \\ = &\sum_{k_1=0}^{t-1} \sum_{k_2=0}^{t+d-1} \operatorname{tr}\{\epsilon_{t-k_1}^T g_{t-k_1}^T \psi_{tk_1}^{(q_1)T}(\theta,\theta^0) S_t(\theta) \psi_{t+d,k_2}^{(q_2)}(\theta,\theta^0) g_{t+d-k_2}\epsilon_{t+d-k_2}\} \\ = &\sum_{k_1=0}^{t-1} \sum_{k_2=0}^{t+d-1} \operatorname{tr}\{g_{t-k_1}^T \psi_{tk_1}^{(q_1)T}(\theta,\theta^0) S_t(\theta) \psi_{t+d,k_2}^{(q_2)}(\theta,\theta^0) g_{t+d-k_2}\epsilon_{t+d-k_2}\epsilon_{t-k_1}^T\}, \quad (5.1) \end{aligned}$$

by using Lemma 3.3 c) so taking the expectation for $\theta = \theta^0$ leaves a simple sum with $k_2 - d = k_1$ since the ϵ_t are independent. Because $E(\epsilon_t \epsilon_t^T) = \Sigma$, for all t, this leads to (4.10).

For the second part, for the first factor of (4.11), we have

$$\begin{aligned} &\partial_\theta^{(q_1)} e_t^T(\theta) S_t(\theta) \partial_\theta^{(q_2)} e_t(\theta) - \tilde{\delta} E_\theta[\partial_\theta^{(q_1)} e_t^T(\theta) S_t(\theta) \partial_\theta^{(q_2)} e_t(\theta)] \quad (5.2) \\ = &\sum_{k_1=0}^{t-1} \sum_{k_2=0}^{t-1} \epsilon_{t-k_1}^T g_{t-k_1}^T \psi_{tk_1}^{(q_1)T}(\theta,\theta^0) S_t(\theta) \psi_{tk_2}^{(q_2)}(\theta,\theta^0) g_{t-k_2}\epsilon_{t-k_2} \quad (5.3) \\ - &\tilde{\delta} E_\theta \left\{ \sum_{k_1=0}^{t-1} \sum_{k_2=0}^{t-1} \epsilon_{t-k_1}^T g_{t-k_1}^T \psi_{tk_1}^{(q_1)T}(\theta,\theta^0) S_t(\theta) \psi_{tk_2}^{(q_2)}(\theta,\theta^0) g_{t-k_2}\epsilon_{t-k_2} \right\}. \end{aligned}$$

We apply the expectation with respect to $F_{t-\nu_1}$ to (5.2), which consists in replacing the lower summation indices from 0 to $\max(\nu_1, 0)$. Taking the trace, and using Lemma

3.3 c) , the first term (5.3) becomes

$$\sum_{k_1,k_2=\max(\nu_1,0)}^{t-1}\sum^{t-1} \operatorname{tr}\left\{g_{t-k_1}^T\psi_{tk_1}^{(q_1)T}(\theta,\theta^0)S_t(\theta)\psi_{tk_2}^{(q_2)}(\theta,\theta^0)g_{t-k_2}\epsilon_{t-k_2}\epsilon_{t-k_1}^T\right\}$$

$$= \sum_{k_1,k_2=\max(\nu_1,0)}^{t-1}\sum^{t-1} \left\{\operatorname{vec}\left(g_{t-k_1}^T\psi_{tk_1}^{(q_1)T}(\theta,\theta^0)S_t(\theta)\psi_{tk_2}^{(q_2)}(\theta,\theta^0)g_{t-k_2}\right)^T\right\}^T \operatorname{vec}(\epsilon_{t-k_2}\epsilon_{t-k_1}^T)$$

$$= \sum_{k_1,k_2=\max(\nu_1,0)}^{t-1}\sum^{t-1} \left\{\operatorname{vec}\left(g_{t-k_2}^T\psi_{tk_2}^{(q_2)T}(\theta,\theta^0)S_t(\theta)\psi_{tk_1}^{(q_1)}(\theta,\theta^0)g_{t-k_1}\right)\right\}^T \operatorname{vec}(\epsilon_{t-k_2}\epsilon_{t-k_1}^T). \tag{5.4}$$

Similarly, since

$$\partial_\theta^{(q)}e_{t+d}(\theta) = \sum_{k=0}^{t+d-1}\psi_{t+d,k}^{(q)}(\theta,\theta_0)g_{t+d-k}\epsilon_{t+d-k}, \tag{5.5}$$

the second factor of (4.11) includes

$$\partial_\theta^{(q_3)}e_{t+d}^T(\theta)S_{t+d}(\theta)\partial_\theta^{(q_4)}e_{t+d}(\theta)$$

$$= \sum_{k_3=0}^{t+d-1}\sum_{k_4=0}^{t+d-1} \operatorname{tr}\{\epsilon_{t+d-k_3}^T g_{t+d-k_3}^T\psi_{t+d,k_3}^{(q_3)T}(\theta,\theta^0)S_{t+d}(\theta)\psi_{t+d,k_4}^{(q_4)}(\theta,\theta^0)g_{t+d-k_4}\epsilon_{t+d-k_4}\}$$

$$= \sum_{k_3=-d}^{t-1}\sum_{k_4=-d}^{t-1} \operatorname{tr}\{\epsilon_{t-k_4}\epsilon_{t-k_3}^T g_{t-k_3}^T\psi_{t+d,k_3+d}^{(q_3)T}(\theta,\theta^0)S_{t+d}(\theta)\psi_{t+d,k_4+d}^{(q_4)}(\theta,\theta^0)g_{t-k_4}\}$$

$$= \sum_{k_3=-d}^{t-1}\sum_{k_4=-d}^{t-1} \{\operatorname{vec}(\epsilon_{t-k_3}\epsilon_{t-k_4}^T)\}^T \operatorname{vec}\left(g_{t-k_3}^T\psi_{t+d,k_3+d}^{(q_3)T}(\theta,\theta^0)S_{t+d}(\theta)\psi_{t+d,k_4+d}^{(q_4)}(\theta,\theta^0)g_{t-k_4}\right),$$

where we have changed the summation indices k_3 and k_4 by subtracting d and moved factors in the opposite way. Taking the conditional expectation with respect to $F_{t-\nu_2}$, the second factor of (4.11) has lower summation indices equal to $\max(\nu_2, -d)$ and is written

$$\sum_{k_3,k_4=\max(\nu_2,-d)}^{t-1}\sum^{t-1} [\operatorname{vec}(\epsilon_{t-k_4}\epsilon_{t-k_3}^T) - \tilde{\delta}E\{\operatorname{vec}(\epsilon_{t-k_4}\epsilon_{t-k_3}^T)\}]^T M_{tdk_3k_4}^{(q_3,q_4)}(\theta,\theta^0), \tag{5.6}$$

where the $r^2 \times 1$ matrix $M_{tdk_3k_4}^{(q_3,q_4)}(\theta,\theta^0)$ is a special case of

$$M_{tfk'k''}^{(q',q'')}(\theta,\theta^0) = \operatorname{vec}\left(g_{t-k'}^T\psi_{t+f,k'+f}^{(q')T}(\theta,\theta^0)S_{t+f}(\theta)\psi_{t+f,k''+f}^{(q'')}(\theta,\theta^0)g_{t-k''}\right), \tag{5.7}$$

for $f = d$. Similarly, for the first factor of (4.11), (5.4) can be written

$$\sum_{k_1,k_2=\max(\nu_1,0)}^{t-1}\sum^{t-1} M_{t0k_2k_1}^{(q_2,q_1)T}(\theta,\theta^0)\operatorname{vec}(\epsilon_{t-k_2}\epsilon_{t-k_1}^T),$$

where the $r^2 \times 1$ matrix $M^{(q_2,q_1)}_{t0k_2k_1}(\theta,\theta^0)$ is as defined in (5.7) for $f=0$. Hence (5.2) is equal to

$$\sum_{k_1,k_2=\max(\nu_1,0)}^{t-1}\sum^{t-1} M^{(q_2,q_1)T}_{t0k_2k_1}(\theta,\theta^0)[\mathrm{vec}(\epsilon_{t-k_2}\epsilon^T_{t-k_1}) - \tilde{\delta}E\{\mathrm{vec}(\epsilon_{t-k_2}\epsilon^T_{t-k_1})\}], \tag{5.8}$$

a special case of (5.7) but for $f=0$. Note that q', q'', k', k'' are replaced by q_2, q_1, k_2, k_1, in that order. Taking the expectation of the product of (5.8) by (5.6) for $\theta=\theta^0$, with the notation $M^{(q',q'')}_{tfk'k''} = M^{(q',q'')}_{tfk'k''}(\theta^0,\theta^0)$, using again Lemma 3.3 c), (4.11) is equal to

$$\sum_{k_1,k_2=\max(\nu_1,0)}^{t-1}\sum^{t-1}\ \sum_{k_3,k_4=\max(\nu_2,-d)}^{t-1}\sum^{t-1} M^{(q_2,q_1)T}_{t0k_2k_1}\Omega_t^{k_1,k_2,k_3,k_4} M^{(q_3,q_4)}_{tdk_3k_4}, \tag{5.9}$$

where $M^{(q_2,q_1)}_{t0k_2k_1}$ and $M^{(q_3,q_4)}_{tdk_3k_4}$ comply with the definition of $M^{(q',q'')}_{tfk'k''}$ in (4.16), for $f=0$ and $f=d$, respectively, and

$$\begin{aligned}\Omega_t^{k_1,k_2,k_3,k_4} &= E\Big[\Big(\mathrm{vec}(\epsilon_{t-k_2}\epsilon^T_{t-k_1}) - \tilde{\delta}E\{\mathrm{vec}(\epsilon_{t-k_2}\epsilon^T_{t-k_1})\}\Big)\,. \\ &\qquad \Big(\mathrm{vec}(\epsilon_{t-k_4}\epsilon^T_{t-k_3}) - \tilde{\delta}E\{\mathrm{vec}(\epsilon_{t-k_4}\epsilon^T_{t-k_3})\}\Big)^T\Big] \\ &= E\left\{\mathrm{vec}(\epsilon_{t-k_2}\epsilon^T_{t-k_1})\,\mathrm{vec}(\epsilon_{t-k_4}\epsilon^T_{t-k_3})^T\right\} \\ &\qquad -\tilde{\delta}E\left\{\mathrm{vec}(\epsilon_{t-k_2}\epsilon^T_{t-k_1})\right\}E\left\{\mathrm{vec}(\epsilon_{t-k_4}\epsilon^T_{t-k_3})^T\right\},\end{aligned} \tag{5.10}$$

since $2\tilde{\delta} - \tilde{\delta}^2 = \tilde{\delta}$, given that $\tilde{\delta} = 0$ or 1. We need to distinguish four cases for the first term of $\Omega_t^{k_1,k_2,k_3,k_4}$: (i) $k_1=k_2=k_3=k_4$, (ii) $k_1=k_3$ and $k_2=k_4$, with $k_1 \neq k_2$, (iii) $k_1=k_4$ and $k_2=k_3$, with $k_1\neq k_2$, (iv) $k_1=k_2$ and $k_3=k_4$, with $k_1\neq k_3$, and in the other cases, it is equal to 0, because the ϵ_t are independent with mean 0. The second term of $\Omega_t^{k_1,k_2,k_3,k_4}$ is equal to 0 for $k_1\neq k_2$ and $k_3\neq k_4$ and is equal to $(\mathrm{vec}\,\Sigma).(\mathrm{vec}\,\Sigma)^T$, otherwise. Hence, given Lemma 4.2, (4.11) becomes (assuming

$d \geq 0$):

$$\sum_{k=\max(\nu_1,\nu_2,0)}^{t-1} M_{t0kk}^{(q_2,q_1)T} \kappa_{t-k} M_{tdkk}^{(q_3,q_4)} \tag{5.11}$$

$$+ \sum_{\substack{k_1,k_2=\max(\nu_1,\nu_2,0)\\k_1\neq k_2}}^{t-1}\sum^{t-1} M_{t0k_2k_1}^{(q_2,q_1)T} K_{r,r}(\Sigma \otimes \Sigma) M_{tdk_1k_2}^{(q_3,q_4)}$$

$$+ \sum_{\substack{k_1,k_2=\max(\nu_1,\nu_2,0)\\k_1\neq k_2}}^{t-1}\sum^{t-1} M_{t0k_2k_1}^{(q_2,q_1)T} (\Sigma \otimes \Sigma) M_{tdk_2k_1}^{(q_3,q_4)}$$

$$+ \sum_{\substack{k_1,k_3=\max(\nu_1,\nu_2,0)\\k_1\neq k_3}}^{t-1}\sum^{t-1} M_{t0k_1k_1}^{(q_2,q_1)T} (\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)^T M_{tdk_3k_3}^{(q_3,q_4)}$$

$$-\tilde{\delta}\left\{\sum_{k_1=\max(\nu_1,0)}^{t-1} M_{t0k_1k_1}^{(q_2,q_1)T}\right\}(\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)^T\left\{\sum_{k_3=\max(\nu_2,-d)}^{t-1} M_{tdk_3k_3}^{(q_3,q_4)}\right\}. \tag{5.12}$$

The terms (5.11) and (5.12) are nearly in final form. Completing the three double sums for equal indices and compensating yields for the remaining terms

$$\sum_{k_1,k_2=\max(\nu_1,\nu_2,0)}^{t-1}\sum^{t-1} M_{t0k_2k_1}^{(q_2,q_1)T} K_{r,r}(\Sigma \otimes \Sigma) M_{tdk_1k_2}^{(q_3,q_4)}$$

$$+ \sum_{k_1,k_2=\max(\nu_1,\nu_2,0)}^{t-1}\sum^{t-1} M_{t0k_2k_1}^{(q_2,q_1)T} (\Sigma \otimes \Sigma) M_{tdk_2k_1}^{(q_3,q_4)}$$

$$+\left\{\sum_{k_1=\max(\nu_1,0)}^{t-1} M_{t0k_1k_1}^{(q_2,q_1)T}\right\}\times(\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)^T\times\left\{\sum_{k_3=\max(\nu_2,-d)}^{t-1} M_{tdk_3k_3}^{(q_3,q_4)}\right\} \tag{5.13}$$

$$- \sum_{k=\max(\nu_1,\nu_2,0)}^{t-1} M_{t0kk}^{(q_2,q_1)}\{K_{r,r}(\Sigma \otimes \Sigma) + (\Sigma \otimes \Sigma) + (\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)^T\} M_{tdkk}^{(q_3,q_4)}. \tag{5.14}$$

The sum (5.14) can be combined with (5.11), using the definition (4.17), giving (4.12), whereas (5.13) combines with (5.12) to give (4.15). The two other double sums correspond to (4.13) and (4.14).

There remains to check (4.18) when $q_1 = q_2 = q_3 = q_4 = 0$, $d = 0$, and $\nu_1 = \nu_2 = 0$. Given (1.10), there is only case (i) for $\Omega_t^{k_1,k_2,k_3,k_4} \neq 0$, when $k_1 = k_2 = k_3 = k_4 = 0$ and then (5.10) is then equal to $\kappa_t - \tilde{\delta}(\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)^T$. □

5.4 Proof of Lemma 4.4

By using Lemma 3.1, the derivative of $\alpha_t(\theta)$ is given by:

$$\partial_i \alpha_t(\theta) = \mathrm{tr}\left(\Sigma_t^{-1}(\theta)\partial_i \Sigma_t(\theta)\right) + 2e_t^T(\theta)\Sigma_t^{-1}\partial_i e_t(\theta) + e_t^T(\theta)\partial_i \Sigma_t^{-1}(\theta)e_t(\theta). \quad (5.15)$$

Hence taking the conditional expectation and using Lemma 4.1

$$E_\theta\{\partial_i \alpha_t(\theta)/F_{t-1}\} = \mathrm{tr}\left(\Sigma_t^{-1}(\theta)\partial_i \Sigma_t(\theta)\right) - E_\theta\{e_t^T(\theta)\partial_i \Sigma_t^{-1}(\theta)e_t(\theta)/F_{t-1}\}. \quad (5.16)$$

Moreover the second term is equal to

$$\begin{aligned}
&E_\theta[\mathrm{tr}(e_t(\theta)e_t^T(\theta)\partial_i \Sigma_t^{-1}(\theta))/F_{t-1}] \\
&= \mathrm{tr}[E_\theta(e_t(\theta)e_t^T(\theta)/F_{t-1})\partial_i \Sigma_t^{-1}(\theta)] \\
&= \mathrm{tr}[\Sigma_t(\theta)\partial_i \Sigma_t^{-1}(\theta)] \quad (5.17)\\
&= \mathrm{tr}[\Sigma_t(\theta)\Sigma_t^{-1}(\theta)\partial_i \Sigma_t(\theta)\Sigma_t^{-1}(\theta)] = \mathrm{tr}[\Sigma_t^{-1}(\theta)\partial_i \Sigma_t(\theta)],
\end{aligned}$$

proving (4.23). □

5.5 Proof of Lemma 4.5

From (5.15) we have, for $i, j = 1, ..., m$,

$$\begin{aligned}
\partial_{ij}\alpha_t(\theta) &= \partial_i \,\mathrm{tr}\left(\Sigma_t^{-1}(\theta)\partial_j \Sigma_t(\theta)\right) + 2\partial_i\left(e_t^T(\theta)\Sigma_t^{-1}\partial_j e_t(\theta)\right) \\
&\quad + \partial_i\left(e_t^T(\theta)\partial_j \Sigma_t^{-1}(\theta)e_t(\theta)\right).
\end{aligned}$$

The first two terms can be written respectively

$$\partial_i \,\mathrm{tr}\left(\Sigma_t^{-1}(\theta)\partial_j \Sigma_t(\theta)\right) = \mathrm{tr}\left(\partial_i \Sigma_t^{-1}(\theta)\partial_j \Sigma_t(\theta)\right) + \mathrm{tr}\left(\Sigma_t^{-1}(\theta)\partial_{ij}\Sigma_t(\theta)\right),$$

and

$$\begin{aligned}
\partial_i\left(e_t^T(\theta)\Sigma_t^{-1}\partial_j e_t(\theta)\right) &= \partial_i e_t(\theta)\Sigma_t^{-1}(\theta)\partial_j e_t(\theta) + e_t^T(\theta)\partial_i \Sigma_t^{-1}(\theta)\partial_j e_t(\theta) \\
&\quad + e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_{ij} e_t(\theta).
\end{aligned}$$

Furthermore, by using Lemma 4.1, we have

$$\partial_i\left(e_t^T(\theta)\partial_j \Sigma_t^{-1}(\theta)e_t(\theta)\right) = 2e_t^T(\theta)\partial_j \Sigma_t^{-1}(\theta)\partial_i e_t(\theta) + e_t^T(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)e_t(\theta).$$

Consequently, we get

$$\begin{aligned}
\partial_{ij}\alpha_t(\theta) &= \mathrm{tr}\left(\partial_i \Sigma_t^{-1}(\theta)\partial_j \Sigma_t(\theta)\right) + \mathrm{tr}\left(\Sigma_t^{-1}(\theta)\partial_{ij}\Sigma_t(\theta)\right) \\
&\quad + 2\partial_i e_t(\theta)\Sigma_t^{-1}(\theta)\partial_j e_t(\theta) + 2e_t^T(\theta)\partial_i \Sigma_t^{-1}(\theta)\partial_j e_t(\theta) \\
&\quad + 2e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_{ij} e_t(\theta) \\
&\quad + 2e_t^T(\theta)\partial_j \Sigma_t^{-1}(\theta)\partial_i e_t(\theta) + e_t^T(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)e_t(\theta), \quad (5.18)
\end{aligned}$$

We take the conditional expectation with respect to F_{t-1}. It can be omitted from the third term, by using Lemma 4.1. The next three terms vanish and the conditional expectation of the last term, $E_\theta(e_t^T(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)e_t(\theta)/F_{t-1})$, is equal to $\mathrm{tr}[\Sigma_t(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)]$, by proceeding like in (5.17). Consequently we have:

$$\begin{aligned} E_\theta(\partial_{ij}\alpha_t(\theta)/F_{t-1}) &= 2\partial_i e_t(\theta)\Sigma_t^{-1}(\theta)\partial_j e_t(\theta) + \mathrm{tr}\left(\partial_i\Sigma_t^{-1}(\theta)\partial_j\Sigma_t(\theta)\right) \\ &\quad + \mathrm{tr}\left(\Sigma_t^{-1}(\theta)\partial_{ij}\Sigma_t(\theta)\right) + \mathrm{tr}[\Sigma_t(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)]. \end{aligned} \tag{5.19}$$

Using Lemma 3.1, the last term of (5.19) can be written

$$\mathrm{tr}\left(\Sigma_t(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)\right) = -\,\mathrm{tr}\left[\partial_{ij}\Sigma_t(\theta)\Sigma_t^{-1}(\theta)\right] + 2\,\mathrm{tr}\left[\partial_i\Sigma_t(\theta)\Sigma_t^{-1}(\theta)\partial_j\Sigma_t(\theta)\Sigma_t^{-1}(\theta)\right]. \tag{5.20}$$

In addition, using the same Lemma 3.1, the second term of (5.19) equals

$$\mathrm{tr}\left(\partial_i\Sigma_t^{-1}(\theta)\partial_j\Sigma_t(\theta)\right) = -\,\mathrm{tr}\left[\partial_i\Sigma_t(\theta)\Sigma_t^{-1}(\theta)\partial_j\Sigma_t(\theta)\Sigma_t^{-1}(\theta)\right],$$

so the sum of the last three terms of (5.19) is equal to $\mathrm{tr}\left[\partial_i\Sigma_t(\theta)\Sigma_t^{-1}(\theta)\partial_j\Sigma_t(\theta)\Sigma_t^{-1}(\theta)\right]$, proving (4.26). Now, subtracting (5.19) from (5.18), we obtain after some cancellations

$$\begin{aligned} \partial_{ij}\alpha_t(\theta) - E_\theta[\partial_{ij}\alpha_t(\theta)/F_{t-1}] &= 2e_t^T(\theta)\partial_j\Sigma_t^{-1}(\theta)\partial_i e_t(\theta) + 2e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_{ij}e_t(\theta) \\ &+ 2e_t^T(\theta)\partial_i\Sigma_t^{-1}(\theta)\partial_j e_t(\theta) + e_t^T(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)e_t(\theta) \\ &- \mathrm{tr}(\Sigma_t\partial_{ij}\Sigma_t^{-1}(\theta)). \end{aligned}$$

Taking the conditional expectation of this expression with respect to F_{t-1} and applying again Lemma 3.5 achieves the proof. □

5.6 Proof of Lemma 4.6

We have to show that

$$E_{\theta^0}\left(\left[e_t^T(\theta)e_t(\theta)\right]^4\right) \le m_1^4 M_1.$$

Since $e_t(\theta^0) = g_t\epsilon_t$, we have

$$\begin{aligned} \left(e_t^T(\theta^0)e_t(\theta^0)\right)^4 &= [\mathrm{tr}\left(\epsilon_t^T g_t^T g_t\epsilon_t\right)]^4 = \left[\mathrm{tr}\left(g_t^T g_t\epsilon_t\epsilon_t^T\right)\right]^4 \\ &\le \left\|g_t^T g_t\right\|_F^4\left(\epsilon_t^T\epsilon_t\right)^4 = \|g_t\|_F^8\left(\epsilon_t^T\epsilon_t\right)^4, \end{aligned}$$

using Lemma 3.4 twice, then

$$E_{\theta^0}\left(e_t^T(\theta)e_t(\theta)\right)^4 \le \|g_t\|_F^8\, E\left[\left(\epsilon_t^T\epsilon_t\right)^4\right].$$

Consequently, since $H_{3.4}$ implies that $E(\epsilon_t^T\epsilon_t)^4 \le M_1$, and $H_{3.5}$, that $\|g_t\|_F^2 \le m_1$,

$$E_{\theta^0}\left(e_t^T(\theta)e_t(\theta)\right)^4 \le m_1^4 M_1.$$

Similarly $E_{\theta^0}\left(e_t^T(\theta)e_t(\theta)\right)^2 \le \|g_t\|_F^4\, E\left(\epsilon_t^T\epsilon_t\right)^2 \le m_1^2 M_1^{1/2}$, using Cauchy-Schwarz inequality. Finally $|E_{\theta^0}\left(e_t^T(\theta)e_t(\theta)\right)| = \mathrm{tr}(\Sigma_t) \le r m_1\|\Sigma\|_F$. □

5.7 Proof of Lemma 4.7

By the first part of Lemma 4.3 and Lemma 3.4

$$\begin{aligned}\left|E_{\theta^0}\left[\partial_\theta^{(q_1)}e_t^T(\theta)S_t(\theta)\partial_\theta^{(q_2)}e_t(\theta)\right]\right| &\le \sum_{k=1}^{t-1}\|g_{t-k}\|_F^2\|\psi_{tk}^{(q_1)}\|_F\|S_t\|_F\|\psi_{tk}^{(q_2)}\|_F\|\Sigma\|_F \\ &\le m_1\widetilde{m}\|\Sigma\|_F\left\{\sum_{k=1}^{t-1}\|\psi_{tk}^{(q_1)}\|_F^2\times\sum_{k=1}^{t-1}\|\psi_{tk}^{(q_2)}\|_F^2\right\}^{1/2} \\ &\le m_1\widetilde{m}(N_1'N_2')^{1/2}\|\Sigma\|_F,\end{aligned}$$

using (4.29), since $P(1)=1$, (4.31), Cauchy-Schwartz inequality and assumption $H_{3.5}$ of Theorem 3.1 in the main paper. This proves (4.32). By the second part of Lemma 4.3, Lemma 3.4 and given the definition of $M_{tfk'k''}^{(q',q'')}$ for $f=0$ in (4.16),

$$\begin{aligned}\left\|M_{t0k'k''}^{(q',q'')}\right\|_F &= \|g_{t-k'}\|_F\left\|\psi_{t,k'}^{(q')}\right\|_F\|S_t\|_F\left\|\psi_{t,k''}^{(q'')}\right\|_F\|g_{t-k''}\|_F \\ &\le m_1\widetilde{m}\left\|\psi_{t,k'}^{(q')}\right\|_F\left\|\psi_{t,k''}^{(q'')}\right\|_F, \qquad (5.21)\end{aligned}$$

for all t,k',k'', and using (4.31) and assumption $H_{3.5}$ of Theorem 3.1 in the main paper Using $H_{3.4}$, we have $\|\Xi_{t-k}\|_F\le M_3$. Hence, given (5.21),

$$\begin{aligned}\left|\sum_{k=\nu}^{t-1}M_{t0kk}^{(q_2,q_1)T}\Xi_{t-k}M_{t0kk}^{(q_1,q_2)}\right| &\le m_1^2\widetilde{m}^2M_3\sum_{k=\nu}^{t-1}\left\|\psi_{tk}^{(q_1)}\right\|_F^2\left\|\psi_{tk}^{(q_2)}\right\|_F^2 \\ &\le m_1^2\widetilde{m}^2M_3(N_1''N_2'')^{1/2}P(\nu)\Phi^{\nu-1}, \qquad (5.22)\end{aligned}$$

using Cauchy-Schwarz inequality and (4.30). This is the first term of (4.33). Given the restrictions in the statement, using again (5.21), the term (4.13) can be bounded in absolute value by

$$\begin{aligned}&\sum_{k_1,k_2=\nu}^{t-1}\sum^{t-1}\left\|M_{t0k_2k_1}^{(q_2,q_1)T}\right\|_F\|K_{r,r}(\Sigma\otimes\Sigma)\|_F\left\|M_{t0k_1k_2}^{(q_1,q_2)}\right\|_F \\ \le\ & m_1^2\widetilde{m}^2r\,\|\Sigma\|_F^2\sum_{k_1=\nu}^{t-1}\left\|\psi_{tk_1}^{(q_1)}\right\|_F^2\sum_{k_2=\nu}^{t-1}\left\|\psi_{tk_2}^{(q_2)}\right\|_F^2 \\ \le\ & m_1^2\widetilde{m}^2rN_1'N_2'P^2(\nu)\Phi^{2(\nu-1)}\,\|\Sigma\|_F^2, \qquad (5.23)\end{aligned}$$

using now (4.29) and $\|K_{r,r}\|_F=r$. This gives the first part of the second term of (4.33). Still given the restrictions, the term (4.14) can be rewritten, using (5.7) and Lemma 3.3 f)

$$\sum_{k_1,k_2=\nu}^{t-1}\sum^{t-1}\operatorname{tr}\left(M_{t0k_2k_1}^{(q_2,q_1)}\Sigma M_{t0k_1k_2}^{(q_1,q_2)T}\Sigma\right) \qquad (5.24)$$

whose absolute value can be bounded similarly by

$$m_1^2\widetilde{m}^2 \left\|\Sigma\right\|_F^2 \sum_{k_1=\nu}^{t-1} \left\|\psi_{tk_1}^{(q_1)}\right\|_F^2 \sum_{k_2=\nu}^{t-1} \left\|\psi_{tk_2}^{(q_2)}\right\|_F^2 \leq m_1^2\widetilde{m}^2 N_1' N_2' P^2(\nu)\Phi^{2(\nu-1)} \left\|\Sigma\right\|_F^2 . \quad (5.25)$$

Finally the term (4.15) is also bounded in absolute value by

$$|1-\tilde{\delta}| m_1^2\widetilde{m}^2 N_1' N_2' P^2(\nu)\Phi^{2(\nu-1)} \left\|\Sigma\right\|_F^2 . \quad (5.26)$$

To obtain the second part of the second term of (4.33), the last two bounds can be put together noting that $\tilde{\delta} \in \{0,1\}$ implies that $1+|1-\tilde{\delta}| = 2-\tilde{\delta}$. In order to prove (4.34), we use the second part of Lemma 4.3 in the case where $q_1 = q_2 = q_3 = q_4 = 0$ and $\nu_1 = \nu_2 = 0$, more specifically (4.18). The upper bound (5.21) of $\|M_{t000}^{(0,0)}\|_F = \mathrm{vec}(g_t^T S_t g_t)$ becomes $m_1\widetilde{m}$. An upper bound of $\|\kappa_t - \tilde{\delta}(\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)^T\|_F$ based on assumption $H_{3.4}$ of Theorem 3.1 in the main paper is M_3. Consequently, in that case, the upper bound (4.33) simplifies to $m_1^2\widetilde{m}^2 M_3$. □

5.8 Proof of Lemma 4.8

We have to obtain an upper bound to the absolute value of $E_{\theta^0}\left\{\partial_i e_t^T(\theta)\partial_i e_t(\theta)\right\}$ and of $E_{\theta^0}\left\{\partial_{ij} e_t^T(\theta)\partial_{ij} e_t(\theta)\right\}$. From (1.6) and (1.7) we can write the expressions in the left hand side of (4.35) and (4.36) as special cases of the left hand side of (4.10), so we can use the first part of Lemma 4.3 with $q_1 = q_2 = 1$, and $q_1 = q_2 = 2$, respectively. Also $S(\theta)$ is equal to the identity matrix. We use the first part of Lemma 4.7 respectively with $N_1' = N_2' = N_1$ and $N_1' = N_2' = N_3$, as defined in assumption $H_{3.2}$ of Theorem 3.1 in the main paper. Hence

$$|E_{\theta^0}\left\{\partial_i e_t^T(\theta)\partial_i e_t(\theta)\right\}| \quad \leq m_1 N_1 \left\|\Sigma\right\|_F , |E_{\theta^0}\left\{\partial_{ij} e_t^T(\theta)\partial_{ij} e_t(\theta)\right\}| \quad \leq m_1 N_3 \left\|\Sigma\right\|_F ,$$

using also $H_{3.5}$. □

5.9 Proof of Lemma 4.9

We have to compute

$$\mathrm{cov}_{\theta^0}\left(\partial_i e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_j e_t(\theta), \partial_i e_{t+d}^T(\theta)\Sigma_{t+d}^{-1}(\theta)\partial_j e_{t+d}(\theta)\right).$$

The proof is a direct consequence of the second part of Lemma 4.3 using $q_1 = q_2 = q_3 = q_4 = 1$, and $\partial_\theta^{(q')} e_t(\theta) = \partial_i e_t(\theta)$ for $q' = q_1 = q_3$ and $\partial_\theta^{(q'')} e_t(\theta) = \partial_j e_t(\theta)$ for $q'' = q_2 = q_4$. Indeed the covariance can be put under the form of (4.11) provided $\tilde{\delta} = 1$, $S(\theta) = \Sigma^{-1}(\theta)$, and $\nu_1 = 1$ and $\nu_2 = 1-d$, so that the two conditional expectations can be replaced by ordinary expectations. Note that all the lower summation limits of (4.12)-(4.15) will be 1. Then, $M_{tfk'k''}^{(q',q'')}$ becomes

$$M_{tfk'k''}^{ij} = g_{t-k'}^T \psi_{t+f,i,k'+f}^T \Sigma_{t+f}^{-1} \psi_{t+f,j,k''+f} g_{t-k''},$$

noting that $\Sigma_{t+f}^{-1} = \{g_{t+f}\Sigma g_{t+f}^T\}^{-1}$, which completes the proof. □

5.10 Proof of Lemma 4.10

We have to show (4.38) with K_{ti} defined by (4.39) and $E(\epsilon_t^{\otimes 3})$ defined by (4.40). Using (5.15), the first derivative of $\alpha_t(\theta)$ with respect to θ_i is equal to $\sum_{l=1}^{3} a_l^i(\theta)$using (4.20)-(4.22), hence

$$\partial_i \alpha_t(\theta).\partial_j \alpha_t(\theta) = \sum_{l_1=1}^{3}\sum_{l_2=1}^{3} a_{l_1}^i(\theta) a_{l_2}^j(\theta). \tag{5.27}$$

We take the expectation and the conditional expectation of (5.27) term by term. We need to evaluate

$$A_{l_1 l_2}^{ij} = E_{\theta^0}\left(a_{l_1}^i(\theta) a_{l_2}^j(\theta)/F_{t-1}\right) - E_{\theta^0}\left(a_{l_1}^i(\theta) a_{l_2}^j(\theta)\right),$$

for $l_1, l_2 = 1, 2, 3$ and the left hand side of (4.38) is

$$\sum_{l_1=1}^{3}\sum_{l_2=1}^{3} A_{l_1 l_2}^{ij}. \tag{5.28}$$

By symmetry we can limit ourselves to $l_1 \leq l_2$. First note that $a_1^i(\theta)a_1^j(\theta)$ is non random, hence $A_{11}^{ij} = 0$.

Using the fact that, for $\theta = \theta^0$, $e_t(\theta)$ and $\partial_i e_t(\theta)$ are independent

$$\begin{aligned} E_{\theta^0}\left(a_1^i(\theta)a_2^j(\theta)/F_{t-1}\right) &= E_{\theta^0}\left(\text{tr}\left({\Sigma_t}^{-1}(\theta)\partial_i\Sigma_t(\theta)\right).2e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_i e_t(\theta)/F_{t-1}\right) \\ &= 0. \end{aligned} \tag{5.29}$$

Hence $E_{\theta^0}\left(a_1^i(\theta)a_2^j(\theta)\right) = 0$ and $A_{12}^{ij} = 0$.

To compute A_{13}^{ij}, we have

$$\begin{aligned} & E_{\theta^0}\left(a_1^i(\theta)a_3^j(\theta)/F_{t-1}\right) \\ = \;& E_{\theta^0}\left(\text{tr}\left[\Sigma_t^{-1}(\theta)\partial_i\Sigma_t(\theta)\right] e_t^T(\theta)\partial_j\Sigma_t^{-1}(\theta)e_t(\theta)/F_{t-1}\right) \\ = \;& \text{tr}\left[\Sigma_t^{-1}\left\{\partial_i\Sigma_t(\theta)\right\}_{\theta=\theta^0}\right] \text{tr}\left[\left\{\partial_j\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0} E_{\theta^0}\left(e_t(\theta)e_t^T(\theta)/F_{t-1}\right)\right] \\ = \;& -\text{tr}\left[\Sigma_t^{-1}\left\{\partial_i\Sigma_t(\theta)\right\}_{\theta=\theta^0}\right] \text{tr}\left[\Sigma_t^{-1}\left\{\partial_j\Sigma_t(\theta)\right\}_{\theta=\theta^0}\right], \end{aligned} \tag{5.30}$$

using Lemma 3.1, then $E_{\theta^0}\left(a_1^i(\theta)a_3^j(\theta)\right) = (5.30)$ and $A_{13}^{ij} = 0$.

To compute A_{22}^{ij}, we deduce from Lemma 4.1,

$$\begin{aligned} & E_{\theta^0}\left(a_2^i(\theta)a_2^j(\theta)/F_{t-1}\right) \\ = \;& E_{\theta^0}\left(2e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_i e_t(\theta)2e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_j e_t(\theta)/F_{t-1}\right) \\ = \;& 4E_{\theta^0}\left(e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_i e_t(\theta)\partial_j e_t^T(\theta)\Sigma_t^{-1}(\theta)e_t(\theta)/F_{t-1}\right) \\ = \;& 4E_{\theta^0}\left(\text{tr}\left(\Sigma_t^{-1}(\theta)\partial_i e_t(\theta)\partial_j e_t^T(\theta)\Sigma_t^{-1}(\theta)e_t(\theta)e_t^T(\theta)\right)/F_{t-1}\right) \\ = \;& 4\,\text{tr}\left[\Sigma_t^{-1}\left\{\partial_i e_t(\theta)\partial_j e_t^T(\theta)\right\}_{\theta=\theta^0}\Sigma_t^{-1}E_{\theta^0}\left(e_t(\theta)e_t^T(\theta)\right)\right] \\ = \;& 4\,\text{tr}\left[\Sigma_t^{-1}\left\{\partial_i e_t(\theta)\partial_j e_t^T(\theta)\right\}_{\theta=\theta^0}\right] \\ = \;& 4\left\{\partial_j e_t^T(\theta)\right\}_{\theta=\theta^0}\Sigma_t^{-1}\left\{\partial_i e_t(\theta)\right\}_{\theta=\theta^0}. \end{aligned} \tag{5.31}$$

since $e_t(\theta)$ and $\partial_l e_t(\theta)$ with $l = i, j$ are independent for $\theta = \theta^0$. Then $E_{\theta^0}\left(a_2^i(\theta)a_2^j(\theta)\right)$ is equal to

$$4E_{\theta^0}\left(\partial_j e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_i e_t(\theta)\right), \tag{5.32}$$

which differs from (5.31). The resulting difference A_{22}^{ij} is the first term in the right hand side of (4.38).

Now consider A_{23}^{ij}

$$E_{\theta^0}\left(a_2^i(\theta)a_3^j(\theta)/F_{t-1}\right) = E_{\theta^0}\left(2e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_i e_t(\theta)e_t^T(\theta)\partial_j\Sigma_t^{-1}(\theta)e_t(\theta)/F_{t-1}\right),$$

we have from (1.2) $e_t(\theta^0) = g_t\epsilon_t$ then,

$$\begin{aligned} & E_{\theta^0}\left(a_2^i(\theta)a_3^j(\theta)/F_{t-1}\right) \\ = \; & 2E\left(\operatorname{tr}\left(\epsilon_t^T g_t^T \Sigma_t^{-1}\left\{\partial_i e_t(\theta)\right\}_{\theta=\theta^0} \epsilon_t^T g_t^T \left\{\partial_j \Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0} g_t\epsilon_t\right)/F_{t-1}\right) \\ = \; & 2E\left(\operatorname{tr}\left(g_t^T \Sigma_t^{-1}\left\{\partial_i e_t(\theta)\right\}_{\theta=\theta^0} \epsilon_t^T g_t^T \left\{\partial_j \Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0} g_t\epsilon_t\epsilon_t^T\right)/F_{t-1}\right), \end{aligned} \tag{5.33}$$

thus, using Lemma 3.3 f) and then e),

$$\begin{aligned} (5.33) = \; & 2E\Big[\operatorname{vec}\left(\left\{\partial_i e_t^T(\theta)\right\}_{\theta=\theta^0}\Sigma_t^{-1}g_t\right)^T\left(\epsilon_t\epsilon_t^T\otimes\epsilon_t^T\right) \\ & \times \operatorname{vec}\left(g_t^T\left\{\partial_j\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}g_t\Big]/F_{t-1}\right), \\ = \; & 2\left\{\partial_i e_t^T(\theta)\right\}_{\theta=\theta^0}\Sigma_t^{-1}g_t E\left[(\epsilon_t\epsilon_t^T)\otimes\epsilon_t^T/F_{t-1}\right]\operatorname{vec}\left(g_t^T\left\{\partial_j\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}g_t\right) \\ = \; & 2\left\{\partial_i e_t^T(\theta)\right\}_{\theta=\theta^0}\Sigma_t^{-1}g_t E\left(\epsilon_t^{\otimes 3}\right)\left(g_t^T\otimes g_t^T\right)\operatorname{vec}(\left\{\partial_j\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}), \end{aligned}$$

since the argument of the first vec factor is a $r\times 1$ vector and $E\left[(\epsilon_t\epsilon_t^T)\otimes\epsilon_t^T/F_{t-1}\right]$ is equal to $E\left[(\epsilon_t\epsilon_t^T)\otimes\epsilon_t^T\right]$, given by (4.40). Consequently

$$E_{\theta^0}\left(a_2^i(\theta)a_3^j(\theta)/F_{t-1}\right) = 2\left\{\partial_i e_t^T(\theta)\right\}_{\theta=\theta^0}K_{tj}, \tag{5.34}$$

with K_{tj} is defined by (4.39):

$$K_{tj} = \Sigma_t^{-1}g_t E\left(\epsilon_t^{\otimes 3}\right)\left(g_t^T\otimes g_t^T\right)\operatorname{vec}\left[\left\{\partial_j\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right].$$

Similarly

$$E_{\theta^0}\left(a_2^i(\theta)a_3^j(\theta)\right) = 2E_{\theta^0}\left(\partial_i e_t^T(\theta)\right)K_{tj}, \tag{5.35}$$

and consequently

$$A_{23}^{ij} = 2\left[\left\{\partial_i e_t^T(\theta)\right\}_{\theta=\theta^0} - E_{\theta^0}\left(\partial_i e_t^T(\theta)\right)\right]K_{tj}. \tag{5.36}$$

By symmetry

$$A_{32}^{ij} = 2\left[\left\{\partial_j e_t^T(\theta)\right\}_{\theta=\theta^0} - E_{\theta^0}\left(\partial_j e_t^T(\theta)\right)\right]K_{ti}. \tag{5.37}$$

Finally, for computing A_{33}^{ij}, consider

$$E_{\theta^0}\left(a_3^i(\theta)a_3^j(\theta)/F_{t-1}\right) = E_{\theta^0}\left(e_t^T(\theta)\partial_i\Sigma_t^{-1}(\theta)e_t(\theta)e_t^T(\theta)\partial_j\Sigma_t^{-1}(\theta)e_t(\theta)/F_{t-1}\right).$$

We have for $\theta = \theta^0$ and by using Lemma 3.3 c)

$$\begin{aligned}
&= E\left(\mathrm{tr}\left(g_t^T \left\{\partial_i \Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0} g_t \epsilon_t \epsilon_t^T\right)\right. \\
&\quad \left. \times\ \mathrm{tr}\left(\epsilon_t \epsilon_t^T g_t^T \left\{\partial_j \Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0} g_t\right) / F_{t-1}\right) \\
&= \mathrm{vec}\left(g_t^T \left\{\partial_i \Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0} g_t\right)^T \\
&\quad E\left(\mathrm{vec}\left(\epsilon_t^T \epsilon_t\right) \times \mathrm{vec}\left(\epsilon_t \epsilon_t^T\right)^T / F_{t-1}\right) \mathrm{vec}\left(g_t^T \left\{\partial_j \Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0} g_t\right) \\
&= \mathrm{vec}\left(g_t^T \left\{\partial_i \Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0} g_t\right)^T \kappa_t \,\mathrm{vec}\left(g_t^T \left\{\partial_j \Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0} g_t\right).
\end{aligned}$$

Moreover

$$E_{\theta^0}\left(a_3^i(\theta) a_3^j(\theta)\right) = E_{\theta^0}\left(a_3^i(\theta) a_3^j(\theta) / F_{t-1}\right), \tag{5.38}$$

because of (4.7) and therefore $A_{33} = 0$.

Finally, from (5.27)-(5.38), we have

$$\begin{aligned}
& E_{\theta^0}\left(\partial_i \alpha_t(\theta) \partial_j \alpha_t(\theta) / F_{t-1}\right) - E_{\theta^0}\left(\partial_i \alpha_t(\theta) \partial_j \alpha_t(\theta)\right) \\
&= 4\left[\partial_j e_t^T(\theta) \Sigma_t^{-1}(\theta) \partial_i e_t(\theta) - E_{\theta^0}\left(\partial_j e_t^T(\theta) \Sigma_t^{-1}(\theta) \partial_i e_t(\theta)\right)\right] \\
&\quad +2\left[\left\{\partial_i e_t^T(\theta)\right\}_{\theta=\theta^0} - E_{\theta^0}\left(\partial_i e_t^T(\theta)\right)\right] \Sigma_t^{-1} g_t E\left(\epsilon_t^{\otimes 3}\right) \\
&\quad \times \left(g_t^T \otimes g_t^T\right) \mathrm{vec}\left(\left\{\partial_j \Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right) \\
&\quad +2\left[\left\{\partial_j e_t^T(\theta)\right\}_{\theta=\theta^0} - E_{\theta^0}\left(\partial_j e_t^T(\theta)\right)\right] \Sigma_t^{-1} g_t E\left(\epsilon_t^{\otimes 3}\right) \\
&\quad \times \left(g_t^T \otimes g_t^T\right) \mathrm{vec}\left(\left\{\partial_i \Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right),
\end{aligned}$$

which corresponds to (4.38). □

5.11 Proof of Lemma 4.11

Since $\partial_i \alpha_t(\theta)$ is decomposed in a sum of three terms a_ℓ^i, $\ell = 1, 2, 3$, using the notations (4.20-4.22), we can use Lemma 3.2 by taking $p = 4$ and $n = 3$:

$$\left|a_1^i(\theta) + a_2^i(\theta) + a_3^i(\theta)\right|^4 \leq 3^3\left(\left|a_1^i(\theta)\right|^4 + \left|a_2^i(\theta)\right|^4 + \left|a_3^i(\theta)\right|^4\right). \tag{5.39}$$

We need to show that

$$E_{\theta^0}\left\{\left|a_1^i(\theta) + a_2^i(\theta) + a_3^i(\theta)\right|\right\}^4 \leq C_1 = 3^4 \max(C_1^{(1)}, C_1^{(2)}, C_1^{(3)}),$$

where $C_1^{(1)}, C_1^{(2)}$ and $C_1^{(3)}$ will be given by the upper bounds in the right hand sides of (5.40), (5.41) and (5.42) respectively. Recall that $\Sigma_t(\theta)$ is symmetric, hence its inverse and derivatives. From (4.20) and the results of Lemma 3.4 we have

$$\begin{aligned}
\left|a_1^i(\theta)\right|^4 &= \left|\mathrm{tr}\left(\Sigma_t^{-1}(\theta) \partial_i \Sigma_t(\theta)\right)\right|^4 \\
&\leq \mathrm{tr}^2\left(\left\{\Sigma_t^{-1}(\theta)\right\}^T \Sigma_t^{-1}(\theta)\right) \mathrm{tr}^2\left(\left\{\partial_i \Sigma_t(\theta)\right\}^T \partial_i \Sigma_t(\theta)\right) \\
&\leq \left\|\Sigma_t^{-1}(\theta)\right\|_F^4 \left\|\partial_i \Sigma_t(\theta)\right\|_F^4.
\end{aligned}$$

Therefore, replacing θ by θ^0 and under $H_{3.3}$ and $H_{3.5}$ in the main paper

$$E_{\theta^0}\left(\left|a_1^i(\theta)\right|^4\right) = \left|a_1^i(\theta)\right|^4 \leq \left\|\Sigma_t^{-1}\right\|_F^4 \left\|\{\partial_i\Sigma_t(\theta)\}_{\theta=\theta^0}\right\|_F^4 \leq C_1^{(1)} = m_2^2K_1^2. \qquad (5.40)$$

From (4.21) we have

$$\begin{aligned}\left|a_2^i(\theta)\right|^4 &= \left|\mathrm{tr}\left(2e_t^T(\theta)\Sigma_t{}^{-1}(\theta)\partial_i e_t(\theta)\right)\right|^4 \\ &= 2^4\left|\mathrm{tr}\left(\Sigma_t^{-1}(\theta)\partial_i e_t(\theta)e_t^T(\theta)\right)\right|^4.\end{aligned}$$

By using Lemma 3.4

$$\begin{aligned}\left|a_2^i(\theta)\right|^4 &\leq 2^4\,\mathrm{tr}^2\left(\left\{\Sigma_t^{-1}(\theta)\right\}^T\Sigma_t^{-1}(\theta)\right)\mathrm{tr}^2\left(e_t(\theta)\partial_i e_t^T(\theta)\partial_i e_t(\theta)e_t^T(\theta)\right) \\ &\leq 2^4\left\|\Sigma_t^{-1}(\theta)\right\|_F^4 \mathrm{tr}^2\left(\partial_i e_t^T(\theta)\partial_i e_t(\theta)e_t^T(\theta)e_t(\theta)\right) \\ &\leq 2^4\left\|\Sigma_t^{-1}(\theta)\right\|_F^4 \mathrm{tr}\left(\partial_i e_t^T(\theta)\partial_i e_t(\theta)\partial_i e_t^T(\theta)\partial_i e_t(\theta)\right)\times \\ &\qquad \times\mathrm{tr}\left(e_t^T(\theta)e_t(\theta)e_t^T(\theta)e_t(\theta)\right) \\ &= 2^4\left\|\Sigma_t^{-1}(\theta)\right\|_F^4\left(\partial_i e_t^T(\theta)\partial_i e_t(\theta)\right)^2\left(e_t^T(\theta)e_t(\theta)\right)^2.\end{aligned}$$

Hence, using arguments like in Lemma 4.1, we have for $\theta = \theta^0$, that the derivatives $\partial_i e_t(\theta)$ are independent of $e_t(\theta)$ for $i = 1, ..., m$ and $t = 1, ..., n$. Then under $H_{3.5}$

$$E_{\theta^0}\left(\left|a_2^i(\theta)\right|^4\right) \leq 2^4 m_2^2 E_{\theta^0}\left(\partial_i e_t^T(\theta)\partial_i e_t(\theta)\right)^2 E_{\theta^0}\left(e_t^T(\theta)e_t(\theta)\right)^2,$$

The factor $E_{\theta^0}\left(\partial_i e_t^T(\theta)\partial_i e_t(\theta)\right)^2$ can be put in the form of (4.11) in Lemma 4.3 provided $\nu_1 = \nu_2 = 1$, $d = 0$, $\tilde{\delta} = 0$, $q_1 = q_2 = q_3 = q_4 = 1$, and $S_t(\theta)$ is the identity matrix. So, using the second part of Lemma 4.7, with $\nu = 1$, $S(\theta) = I_r$, hence $\widetilde{m} = r^2$, and also Lemma 4.6

$$E_{\theta^0}\left(\left|a_2^i(\theta)\right|^4\right) \leq C_1^{(2)} = 2^4r^2m_1^4m_2^2M_1^{1/2}\left\{M_3N_2 + (r+2)N_1^2\left\|\Sigma\right\|_F^2\right\}. \quad (5.41)$$

From (4.22) we have

$$\begin{aligned}\left|a_3^i(\theta)\right|^4 &= \left|\mathrm{tr}\left(e_t^T(\theta)\partial_i\Sigma_t^{-1}(\theta)e_t(\theta)\right)\right|^4 \\ &= \left|\mathrm{tr}^2\left(\partial_i\Sigma_t^{-1}(\theta)e_t(\theta)e_t^T(\theta)\right)\right|^2.\end{aligned}$$

By Lemma 3.4

$$\begin{aligned}\left|a_3^i(\theta)\right|^4 &\leq \mathrm{tr}^2\left(\left\{\partial_i\Sigma_t^{-1}(\theta)\right\}^2\right)\mathrm{tr}^2\left(\left[e_t^T(\theta)e_t(\theta)e_t^T(\theta)e_t(\theta)\right]\right) \\ &\leq \left\|\partial_i\Sigma_t^{-1}(\theta)\right\|_F^4\left(e_t^T(\theta)e_t(\theta)\right)^4,\end{aligned}$$

hence

$$E_{\theta^0}\left(\left|a_3^i(\theta)\right|^4\right) \leq \left\|\left\{\partial_i\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right\|_F^4 E_{\theta^0}\left\{\left(e_t^T(\theta)e_t(\theta)\right)^4\right\},$$

and, under $H_{3.3}$ and using Lemma 4.6

$$E_{\theta^0}\left(\left|a_3^i(\theta)\right|^4\right) \leq C_1^{(3)} = K_4^2 m_1^4 M_1. \tag{5.42}$$

In conclusion, from (5.40), (5.41) and (5.42), the assumption $H_{2.1}$ is satisfied by taking

$$C_1 = 3^4 \max(C_1^{(1)}, C_1^{(2)}, C_1^{(3)}),$$

which completes the proof. □

5.12 Proof of Lemma 4.12

To check that (2.3) is bounded by a constant C_2. We have assumed that the parameters in the autoregressive and moving average coefficients, so in $e_t(\theta)$ are functionally independent than those involved in the covariance matrix so in $g_t(\theta)$. Assume they are ordered so that the first s parameters are involved in $e_t(\theta)$ and the remaining $m-s$ are involved in $g_t(\theta)$, thus in $\Sigma_t(\theta)$. For $i,j=1,\dots,s$, both θ_i and θ_j are involved in $e_t(\theta)$ but not in $\Sigma_t(\theta)$. Then (2.3) becomes

$$E_{\theta^0}\left|2\left(e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_{ij}e_t(\theta)\right)\right|^2$$

which equals to

$$4E_{\theta^0}\left|\operatorname{tr}\left(e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_{ij}e_t(\theta)\right)\right|^2 = 4E_{\theta^0}\left|\operatorname{tr}\left(\Sigma_t^{-1}(\theta)\partial_{ij}e_t(\theta)e_t^T(\theta)\right)\right|^2.$$

Using Lemma 3.4 and the fact that $e_t(\theta)$ and $\partial_{ij}e_t(\theta)$ are independent

$$\begin{aligned}
(2.3) &\leq 4\operatorname{tr}\left(\Sigma_t^{-1}\Sigma_t^{-1}\right)E_{\theta^0}\left|\operatorname{tr}\left(e_t(\theta)\partial_{ij}e_t^T(\theta)\partial_{ij}e_t(\theta)e_t^T(\theta)\right)\right| \\
&\leq 4\operatorname{tr}\left(\Sigma_t^{-1}\Sigma_t^{-1}\right)E_{\theta^0}\left|\operatorname{tr}\left(\partial_{ij}e_t^T(\theta)\partial_{ij}e_t(\theta)e_t^T(\theta)e_t(\theta)\right)\right| \\
&\leq 4\left\|\Sigma_t^{-1}\right\|_F^2 E_{\theta^0}\left(\partial_{ij}e_t^T(\theta)\partial_{ij}e_t(\theta)\right)E_{\theta^0}\left(e_t^T(\theta)e_t(\theta)\right).
\end{aligned}$$

Thus, under $H_{3.5}$, using Lemmas 4.6 and 4.8 together

$$(2.3) \leq C_2^{(1)} = 4rm_1^2 m_2 N_3 \|\Sigma\|_F^2. \tag{5.43}$$

Now, suppose that $i \leq s$ and $j > s$ so that θ_i is involved in $e_t(\theta)$ whereas θ_i is involved in $\Sigma_t(\theta)$. Therefore (2.3) becomes

$$E_{\theta^0}\left|\left(e_t^T(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)e_t(\theta)\right) - \operatorname{tr}\left(\Sigma_t\left\{\partial_{ij}\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right)\right|^2$$

which, using Lemmas 3.2 and 3.4, can be bounded by

$$\begin{aligned}
& 2E_{\theta^0}\left(e_t^T(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)e_t(\theta)\right)^2 + 2\left[\operatorname{tr}\left(\Sigma_t\left\{\partial_{ij}\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right)\right]^2 \\
\leq\; & 2E_{\theta^0}\left[\operatorname{tr}\left(\partial_{ij}\Sigma_t^{-1}(\theta)e_t(\theta)e_t^T(\theta)\right)^2\right] \\
& \quad +2\operatorname{tr}\left(\Sigma_t\Sigma_t\right)\operatorname{tr}\left(\left\{\partial_{ij}\Sigma_t^{-1}(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right) \\
\leq\; & 2\operatorname{tr}\left(\left\{\partial_{ij}\Sigma_t^{-1}(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right)E_{\theta^0}\left[\operatorname{tr}\left(e_t(\theta)e_t^T(\theta)\right)^2\right] \\
& \quad +2\left\|\Sigma_t\right\|_F^2\left\|\left\{\partial_{ij}\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right\|_F^2 \\
\leq\; & 2\left\|\left\{\partial_{ij}\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right\|_F^2\left[E_{\theta^0}\left(e_t^T(\theta)e_t(\theta)\right)^2 + \left\|\Sigma_t\right\|_F^2\right].
\end{aligned}$$

In addition using (1.3), Lemma 4.6 and under $H_{3.3}$-$H_{3.5}$

$$(2.3) \leq C_2^{(2)} = 2K_5 m_1^2 (M_1^{1/2} + \|\Sigma\|_F^2). \tag{5.44}$$

Finally, suppose that $i, j > s$ so that θ_i and θ_j are both involved in $\Sigma_t(\theta)$. Then

$$E_{\theta^0} \left| 2 \left(e_t^T(\theta) \partial_j \Sigma_t^{-1}(\theta) \partial_i e_t(\theta) \right) \right|^2 ,$$

and it is equal to

$$4E_{\theta^0} \left| \text{tr} \left(e_t^T(\theta) \partial_j \Sigma_t^{-1}(\theta) \partial_i e_t(\theta) \right) \right|^2 \quad = \quad 4E_{\theta^0} \left| \text{tr} \left(\partial_j \Sigma_t^{-1}(\theta) \partial_i e_t(\theta) e_t^T(\theta) \right) \right|^2 .$$

Thus, by using Lemma 3.4,

$$\begin{aligned}
(2.3) \quad &\leq \quad 4 \, \text{tr} \left(\left\{ \partial_j \Sigma_t^{-1}(\theta) \right\}_{\theta=\theta^0} \left\{ \partial_j \Sigma_t^{-1}(\theta) \right\}_{\theta=\theta^0} \right) E_{\theta^0} \left| \text{tr} \left(e_t(\theta) \partial_i e_t^T(\theta) \partial_i e_t(\theta) e_t^T(\theta) \right) \right| \\
&\leq \quad 4 \left\| \left\{ \partial_j \Sigma_t^{-1}(\theta) \right\}_{\theta=\theta^0} \right\|_F^2 E_{\theta^0} \left| \text{tr} \left(\partial_i e_t^T(\theta) \partial_i e_t(\theta) e_t^T(\theta) e_t(\theta) \right) \right| \\
&\leq \quad 4 \left\| \left\{ \partial_j \Sigma_t^{-1}(\theta) \right\}_{\theta=\theta^0} \right\|_F^2 E_{\theta^0} \left(\partial_i e_t^T(\theta) \partial_i e_t(\theta) \right) E_{\theta^0} \left(e_t^T(\theta) e_t(\theta) \right) .
\end{aligned}$$

Therefore, using Lemmas 4.6 and 4.8 and under $H_{3.3}$,

$$(2.3) \leq C_2^{(3)} = 4r K_4 m_1^2 N_1 \|\Sigma\|_F^2. \tag{5.45}$$

Consequently, using (5.43)-(5.45), (2.3) is bounded by $C_2 = \max(C_2^{(1)}, C_2^{(2)}, C_2^{(3)})$. □

5.13 Proof of Lemma 4.13

Let us check the conditions of Lemma A.1 in the main paper with $\delta = 1$. For condition i, we have to show that

$$E \left(Z_{tij}^2 \right) \quad = \quad E_{\theta^0} \left[\partial_i e_t^T(\theta) \Sigma_t^{-1}(\theta) \partial_j e_t(\theta) - E_{\theta^0} \left(\partial_i e_t^T(\theta) \Sigma_t^{-1}(\theta) \partial_j e_t(\theta) \right) \right]^2$$

is finite. That expression is bounded by

$$E_{\theta^0} \left(\partial_i e_t^T(\theta) \Sigma_t^{-1}(\theta) \partial_j e_t(\theta) \right)^2 . \tag{5.46}$$

We can use the second part of Lemma 4.3 to evaluate that expression with $q_1 = q_2 = q_3 = q_4 = 1$, $\tilde{\delta} = 0$, $d = 0$, $S_t(\theta) = \Sigma_t^{-1}(\theta)$, and $\nu_1 = \nu_2 = 1$. Then using Lemma 4.7 with $\widetilde{m} = m_2$, $N_1' = N_2' = N_1$, and $N_1'' = N_2'' = N_2$, given the assumptions $H_{3.2}$, $H_{3.4}$ and $H_{3.5}$ of Theorem 3.1 in the main paper, the upper bound of (5.46) is

$$E_{\theta^0} \left(\partial_i e_t^T(\theta) \Sigma_t^{-1}(\theta) \partial_j e_t(\theta) \right)^2 \leq m_1^2 m_2^2 \{ M_3 N_2 + (r^2 + 2) N_1^2 \, \|\Sigma\|_F^2 \}. \tag{5.47}$$

To check condition ii, let us show that

$$\begin{aligned}
E \left(\frac{1}{n} \sum_{t=1}^{n} Z_{tij} \right)^2 \quad &= \quad \frac{1}{n^2} \sum_{t=1}^{n} \sum_{t'=1}^{n} E \left(Z_{tij} Z_{t'ij} \right) \qquad (5.48) \\
&= \quad O \left(\frac{1}{n} \right) .
\end{aligned}$$

But (5.48) is equal to

$$\frac{1}{n^2}\sum_{t=1}^{n} E\left(Z_{tij}^2\right) + \frac{2}{n^2}\sum_{t=1}^{n}\sum_{t'=t+1}^{n} E\left(Z_{tij}Z_{t'ij}\right). \tag{5.49}$$

The first term can be bounded by $1/n$ times (5.47). The second term can be written as twice

$$\frac{1}{n^2}\sum_{d=1}^{n-1}\sum_{t=1}^{n-d} \mathrm{cov}_{\theta^0}\left(\partial_i e_t^T(\theta)\Sigma_t^{-1}(\theta)\partial_j e_t(\theta), \partial_i e_{t+d}^T(\theta)\Sigma_{t+d}^{-1}(\theta)\partial_j e_{t+d}(\theta)\right). \tag{5.50}$$

Then by using Lemma 4.9 and given the second part of assumption $H_{3.7}$ of Theorem 3.1 in the main paper, we have (5.50) $= O(1/n)$. Hence the two conditions of Lemma A.1 are completed. □

5.14 Proof of Lemma 4.14

Using the expression of $\Phi_{1t}(\theta^0)$ in (2.5) and Lemma 3.4

$$\begin{aligned}
\left|\Phi_{1t}(\theta^0)\right| \leq\ & \mathrm{tr}\left[\left\{\partial_{ij}\Sigma_t^{-1}(\theta)\partial_{ij}\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right]^{1/2} \mathrm{tr}\left[\{\partial_l\Sigma_t(\theta)\partial_l\Sigma_t(\theta)\}_{\theta=\theta^0}\right]^{1/2} \\
& + \mathrm{tr}\left[\left\{\partial_j\Sigma_t^{-1}(\theta)\partial_j\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right]^{1/2} \mathrm{tr}\left[\{\partial_{il}\Sigma_t(\theta)\partial_{il}\Sigma_t(\theta)\}_{\theta=\theta^0}\right]^{1/2} \\
& + \mathrm{tr}\left[\left\{\partial_i\Sigma_t^{-1}(\theta)\partial_i\Sigma_t^{-1}(\theta)\right\}_{\theta=\theta^0}\right]^{1/2} \mathrm{tr}\left[\{\partial_{jl}\Sigma_t(\theta)\partial_{jl}\Sigma_t(\theta)\}_{\theta=\theta^0}\right]^{1/2} \\
& + \mathrm{tr}\left[\Sigma_t^{-1}\Sigma_t^{-1}\right]^{1/2} \mathrm{tr}\left[\{\partial_{ijl}\Sigma_t(\theta)\partial_{ijl}\Sigma_t(\theta)\}_{\theta=\theta^0}\right]^{1/2} \\
\leq\ & \tilde{\Phi}_1,
\end{aligned}$$

where, using K_1, K_2, K_3, K_4, and K_5 as defined in assumption $H_{3.5}$ of Theorem 3.1 in the main theorem, and m_2 in assumption $H_{3.5}$,

$$\tilde{\Phi}_1 = K_5^{1/2}K_1^{1/2} + 2K_4^{1/2}K_2^{1/2} + m_2^{1/2}K_3^{1/2}. \tag{5.51}$$

□

5.15 Proof of Lemma 4.15

$E(Z_t^2)$ is a special case of (4.11) in the second part of Lemma 4.3 with $\nu_1 = \nu_2 = 0$, $\tilde{\delta} = 1$, $d = 0$, $q_1 = q_2 = q_3 = q_4 = 0$, and $S(\theta) = \partial_{ijl}\Sigma_t^{-1}(\theta)$. The latter has an upper bound K_3 when $\theta = \theta^0$ by assumption $H_{3.3}$ of Theorem 3.1 in the main paper. Using Lemma 4.7 for that case and (4.34), under the assumptions of Theorem 3.1 in the main paper, we have $C_3 = m_1^2 K_3^2 M_3$. □

5.16 Proof of Lemma 4.16

We use $W_t^{(q)}$ defined by (4.43). Then, we use Lemma 4.3, with $\nu_1 = \nu_2 = \nu$, $d = 0$, $\tilde{\delta} = 1$, $q_1 = q_3 = 0$ and each of the alternative combinations (i) $q_2 = q_4 = 1$ and $S_{t,ij}(\theta) = \partial_{ij}\Sigma_t^{-1}(\theta)$, (ii) $q_2 = q_4 = 2$ and $S_{t,i}(\theta) = \partial_i\Sigma_t^{-1}(\theta)$, and (iii) $q_2 = q_4 = 3$, $S_t(\theta) = \Sigma_t^{-1}(\theta)$, respectively, giving the result.

□

5.17 Proof of Lemma 4.17

To prove (i), an upper bound of $E\{E(W_t^{(q)}/F_{t-\nu})^2\}$ given in Lemma 4.16 is obtained using Lemma 4.7 where $N_2' = N_1$, for $q = 1$, $N_2' = N_3$, for $q = 2$, $N_2' = N_5$, for $q = 3$, $N_2'' = N_2$ for $q = 1$, $N_2'' = N_4$, for $q = 2$, $\widetilde{m} = K_5$, for $q = 1$, $\widetilde{m} = K_4$, for $q = 2$, $\widetilde{m} = m_2$, for $q = 3$, where N_1, N_2, N_3, N_4, N_5 are defined in assumption $H_{3.2}$ of Theorem 3.1 in the main paper, K_4, K_5, in assumption $H_{3.3}$, m_2 in assumption $H_{3.5}$. Note that neither N_1' or N_1'', nor N_2'' for $q = 3$ needs to be specified. Hence the upper bound $\psi_\nu c_t$ has the form $m_1 K_4^2 \left\{M_3 N_2 \Phi^{\nu-1} + (r+1) N_1^2 \Phi^{2(\nu-1)} \|\Sigma\|_F^2\right\}$, i.e. a constant times a function going exponentially to zero when $\nu \to \infty$. The conditions (i) and (ii) of Definition A.1 and the assumptions of Lemma A.2, both in the main paper, are thus verified. □

5.18 Proof of Lemma 4.18

We consider X_t^{ilj} defined in (4.45). We use Lemma 4.3 with $\nu_1 = \nu_2 = \nu$, $d = 0$, $\tilde{\delta} = 1$, $q_1 = q_2 = q_3 = q_4 = 1$ and $S_t(\theta) = \partial_l \Sigma_t^{-1}(\theta)$, giving the result. □

5.19 Proof of Lemma 4.19

For $\nu \geq 1$, we have to show that there exist sequences $\psi_\nu \to 0$ as $\nu \to \infty$, and c_t such that (i) $E\{E(X_t^{ilj}|F_{t-\nu})^2\} \leq \psi_\nu c_t$, and (ii) $E(X_t^{ilj} - E(X_t^{ilj}|F_{t+\nu}))^2 \leq \psi_{\nu+1} c_t$. Since X_t^{ilj} involves only ϵ_{t-k} for $k > 0$, $E(X_t^{ilj}|F_{t+\nu}) = X_t^{ilj}$, the left hand side of (ii) is 0. To prove (i), an upper bound of $E\{E(X_t^{ilj}/F_{t-\nu})^2\}$ given in Lemma 4.18 is obtained using Lemma 4.7 where $N_1' = N_3$, $N_2' = N_1$, $N_1'' = N_4$, $N_2'' = N_2$, $\widetilde{m} = m_2$, and N_1, N_2, N_3, N_4 are defined in assumption $H_{3.2}$ of Theorem 3.1 in the main paper, and m_2, similarly defined in assumption $H_{3.5}$. Hence the upper bound $\psi_\nu c_t$ has the form $m_1 K_4^2 \left\{M_3 N_2 \Phi^{\nu-1} + (r+1) N_1^2 \Phi^{2(\nu-1)} \|\Sigma\|_F^2\right\}$, i.e. a constant times a function going exponentially to zero when $\nu \to \infty$. The conditions (i) and (ii) of Definition A.1 and the assumptions of Lemma A.2, both in the main paper, are thus verified. □

5.20 Proof of Lemma 4.20

We consider Y_t^{ijl} defined in (4.48). Like in the proof of Lemma 4.18, we use Lemma 4.3 with $\nu_1 = \nu_2 = \nu$, $d = 0$, $\tilde{\delta} = 1$, but this time with $q_1 = q_3 = 2$ and $q_2 = q_4 = 1$, using components θ_i and θ_j of θ in the former case and components θ_l of θ in the latter case, and $S_t(\theta) = \Sigma_t^{-1}(\theta)$. This gives directly the result. □

5.21 Proof of Lemma 4.21

For $\nu \geq 1$, we have to show that there exist sequences $\psi_\nu \to 0$ as $\nu \to \infty$, and c_t such that (i) $E\{E(Y_t^{ijl}|F_{t-\nu})^2\} \leq \psi_\nu c_t$, and (ii) $E(Y_t^{ijl} - E(Y_t^{ijl}|F_{t+\nu}))^2 \leq \psi_{\nu+1} c_t$. Since Y_t^{ijl} involves only ϵ_{t-k} for $k > 0$, $E(Y_t^{ijl}|F_{t+\nu}) = Y_t^{ijl}$, the left hand side of (ii) is 0. To prove (i), an upper bound of $E\{E(Y_t^{ijl}/F_{t-\nu})^2\}$ given in Lemma 4.18 is obtained using Lemma 4.7 where $N_1' = N_2' = N_1$, $N_1'' = N_2'' = N_2$ and N_1 and

N_2 are defined in assumption $H_{3.2}$ of Theorem 3.1 in the main paper, and $\widetilde{m}$ is K_4, similarly defined in assumption $H_{3.3}$. Hence the upper bound $\psi_\nu c_t$ has the form $m_1 K_4^2 \left\{ M_3 N_2 \Phi^{\nu-1} + (r+1) N_1^2 \Phi^{2(\nu-1)} \|\Sigma\|_F^2 \right\}$, i.e. a constant times a function going exponentially to zero when $\nu \to \infty$. The conditions (i) and (ii) of Definition A.1 and the assumptions of Lemma A.2, both in the main paper, are thus verified. □

5.22 Proof of Lemma 4.22

Note that $\widetilde{Z}_{tij}$ has appeared as A_{32}^{ij} in (5.37). We have to check the conditions of Lemma A.1 in the main paper, with $i, j = 1, ..., m$,

i. $E\left(\widetilde{Z}_{tij}^2\right)$ is bounded,

ii. $E\left(\frac{1}{n}\sum_{t=1}^n \widetilde{Z}_{tij}\right)^2 = O(1/n)$.

We start by the first condition: $E(\widetilde{Z}_{tij})^2 = E(\partial_j e_t^T(\theta) K_{ti})^2$ can be evaluated using the first part of Lemma 4.3, giving

$$\sum_{k=1}^{t-1} \operatorname{tr}\left[g_{t-k}^T \psi_{tjk}^T K_{ti} K_{ti}^T \psi_{tjk} g_t \Sigma \right]$$

and this can be bounded, using Lemma 4.7 by $m_1 N_1 \|K_{ti}\|_F^2 \|\Sigma\|_F$. Moreover

$$K_{ti} = \Sigma_t^{-1} g_t E\left(\epsilon_t^{\otimes 3}\right) \left(g_t^T \otimes g_t^T\right) \operatorname{vec}\left(\partial_i \Sigma_t^{-1}(\theta)\right)_{\theta=\theta^0},$$

and $E\left(\epsilon_t^{\otimes 3}\right) = E\left(\epsilon_t \epsilon_t^T \otimes \epsilon_t^T\right)$. Then, under $H_{3.3}$-$H_{3.5}$

$$\|K_{ti}\|_F^2 \leq m_2 m_1^3 M_2^2 K_4. \tag{5.52}$$

Hence i is proved. To prove ii, in a manner analogous to while checking Lemma 4.13,

$$E\left(\frac{1}{n}\sum_{t=1}^n \widetilde{Z}_{tij}\right)^2 = \frac{1}{n^2}\sum_{t=1}^n E\left(\widetilde{Z}_{tij}^2\right) + \frac{2}{n^2}\sum_{t=1}^n \sum_{t'=t+1}^n E\left(\widetilde{Z}_{tij}\widetilde{Z}_{t'ij}\right). \tag{5.53}$$

The first term can be bounded by $1/n$ times the bound obtained in condition i. The second term can be written as twice

$$\frac{1}{n^2}\sum_{d=1}^{n-1}\sum_{t=1}^{n-d} \operatorname{cov}_{\theta^0}\left(\partial_j e_t^T(\theta) K_{ti}, \partial_j e_{t+d}^T(\theta) K_{t+d,i}\right), \tag{5.54}$$

and this can be bounded, using the first part of Lemma 4.3 with $S_t(\theta) = K_{ti}K_{t+d,i}^T$, (5.52), and independence among the ϵ_t's and Lemma 3.4, by

$$\begin{aligned}
& \frac{1}{n^2}\sum_{d=1}^{n-1}\sum_{t=1}^{n-d}\sum_{k=1}^{t-1} \operatorname{tr}\left[g_{t-k}^T \psi_{t+d,j,k+d}^T K_{ti} K_{t+d,i}^T \psi_{tjk} g_{t-k}\Sigma)\right] \\
\leq \; & \frac{1}{n^2}\left\|\Sigma\right\|_F \sum_{d=1}^{n-1}\sum_{t=1}^{n-d} \left\|K_{ti}\right\|_F \left\|K_{t+d,i}^T\right\|_F \sum_{k=1}^{t-1} \left\|g_{t-k}\right\|_F^2 \left\|\psi_{tjk}\right\|_F \left\|\psi_{t+d,j,k+d}\right\|_F \\
\leq \; & \left\|\Sigma\right\|_F m_2 m_1^3 M_2^2 K_4 \frac{1}{n^2}\sum_{d=1}^{n-1}\sum_{t=1}^{n-d}\sum_{k=1}^{t-1} \left\|g_{t-k}\right\|_F^2 \left\|\psi_{tik}\right\|_F \left\|\psi_{t+d,i,k+d}\right\|_F \\
= \; & O(1/n),
\end{aligned}$$

using assumption $H_{3.5}$ and the first part of assumption $H_{3.7}$ of Theorem 3.1 in the main paper. □

References

[1] Alj, A., Azrak, R., Mélard, G. (2015a) Asymptotic properties of QML estimators for VARMA models with time-dependent coefficients, Part 2, in preparation.

[2] Alj, A., Ley, C., Mélard, G. (2015b) Asymptotic properties of QML estimators for VARMA models with time-dependent coefficients, Part 1, submitted.

[3] Azrak, R. and Mélard, G. (2006) Asymptotic properties of quasi-likelihood estimators for ARMA models with time-dependent coefficients, *Statistical Inference for Stochastic Processes* 9, 279-330.

[4] Harville, D. A. (1997) *Matrix algebra from a statistician's perspective*, Springer-Verlag, New York.

[5] Kollo, T. and von Rosen, D. (2005) *Advanced multivariate statistics with matrices.*, Springer Verlag, New York.

[6] Magnus, J. R. and Neudecker, H. (1988) *Matrix differential calculus with applications in statistics and econometrics*, Wiley, New York.

[7] Seber, A. F. (2008) *A matrix handbook for statisticians*, Wiley, New York.

[8] Steele, J. M. (2004) *The Cauchy-Schwarz master class: An introduction to the art of mathematical inequalities*, Cambridge University Press and the Mathematical Association of America, Cambridge UK and Washington DC.