Ann. Inst. Statist. Math.
Vol. 47, No. 3, 551–579 (1995)

LOCAL ASYMPTOTIC NORMALITY OF MULTIVARIATE ARMA PROCESSES WITH A LINEAR TREND

BERNARD GAREL[1] AND MARC HALLIN[2]

[1]*Institut National Polytechnique, ENSEEIHT, 2, rue Camichel, F-31071 Toulouse Cédex, France*
[2]*Institut de Statistique, Université Libre de Bruxelles, Campus de la Plaine CP210, Boulevard du Triomphe, B-1050 Bruxelles, Belgium*

(Received November 16, 1993; revised October 12, 1994)

Abstract. The local asymptotic normality (LAN) property is established for multivariate ARMA models with a linear trend or, equivalently, for multivariate general linear models with ARMA error term. In contrast with earlier univariate results, the central sequence here is correlogram-based, i.e. expressed in terms of a generalized concept of residual cross-covariance function.

Key words and phrases: Multivariate linear model, multivariate ARMA process, local asymptotic normality.

1. Introduction

Multivariate time series models with a linear trend or, equivalently, general linear models with ARMA error term constitute one of the most fundamental statistical models whenever observations are made sequentially through time. Models of this type actually dominate in such fields of applications as econometrics.

Due to this considerable importance for applications, problems of statistical inference connected with such models have been intensively discussed: see e.g. Judge *et al.* (1985) for a review in the econometric context. Attention however has been concentrated almost exclusively on least squares and Gaussian likelihood methods (least square estimation; Gaussian likehood ratio, Wald or Lagrange multiplier tests; classical correlogram analysis, ...; see e.g. Dufour and King (1991)), not so much because of the plausibility of Gaussian assumptions as because of the commonly accepted opinion that Gaussian methods yield good asymptotic results, in a sense which by the way is not always that clear.

Now, if asymptotic performances are to be assessed, and if asymptotically optimal inference procedures are to be derived, the key result is the local asymptotic normality property (LAN; see LeCam (1960, 1986); Strasser (1985); LeCam and Yang (1990)), the usefulness of which has been clearly established in univariate time series problems: see Swensen (1985) for the AR case (with a linear trend component) and application to Durbin-Watson tests; Kreiss (1987), where the mixed ARMA case (without trend) is considered, with applications to adaptive estimation; Kreiss (1990*a*, 1990*b*) for applications to hypothesis testing in AR processes (without trend); Hallin and Puri (1994), where a rank-based LAN result (i.e., with a central sequence measurable with respect to the vector of residual ranks) allows for a general theory of rank tests for time-series analysis. Another approach to the asymptotic theory of (univariate and multivariate) ARMA processes has been developed, quite successfully and thoroughly (higher-order results also are derived), by Taniguchi; see e.g. Taniguchi (1983), and the many references in Taniguchi (1991). This approach, based on asymptotic expansion techniques, mainly concentrates on point estimation problems in Gaussian ARMA processes.

A multivariate version of Swensen (1985) and Kreiss (1987)'s univariate LAN results is established here, in the most general situation of a multivariate ARMA process with linear trend component. Contrary to these two references, however, our LAN property is stated under correlogram-based form, i.e. with a central sequence which is measurable with respect to (a generalized version of) residual cross-covariance matrices. Such a form is much closer to time-series analysis practice, where correlograms are the main and most familiar tool; it should therefore be intuitively more appealing to practitioners than an equivalent but more abstractly formulated result.

Section 2 is devoted to the notation and technical assumptions. Our main LAN result is stated in Section 3 under three distinct forms. The technical tool is a slightly generalized version of a lemma (Lemma 2.3) due to Swensen (1985), itself relying on a martingale central-limit theorem of LeCam (preliminary version of his 1986 book, Chapter 10). The proofs are concentrated in Section 4.

Boldface denote vectors and matrices; primes indicate transposes; tr $\boldsymbol{A}$, vec $\boldsymbol{A}$, and $\boldsymbol{A}\otimes\boldsymbol{B}$ as usual stand for the trace of $\boldsymbol{A}$, the vector resulting from stacking $\boldsymbol{A}$'s columns on top of each other, and the Kronecker product of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively.

2. The problem

2.1 *Notation and main assumptions*

The model to be considered throughout the paper is the multivariate linear model

$$\boldsymbol{Y}^{(N)} = \boldsymbol{X}^{(N)}\boldsymbol{\beta} + \boldsymbol{U}^{(N)}, \tag{2.1}$$

where

$$\boldsymbol{X}^{(N)} = \begin{pmatrix} x^{(N)}_{-p+1,1} & x^{(N)}_{-p+1,2} & \cdots & x^{(N)}_{-p+1,n} \\ \vdots & \vdots & & \\ x^{(N)}_{0,1} & x^{(N)}_{0,2} & & \vdots \\ x^{(N)}_{1,1} & x^{(N)}_{1,2} & & \\ \vdots & \vdots & & \\ x^{(N)}_{N,1} & x^{(N)}_{N,2} & \cdots & x^{(N)}_{N,n} \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}'^{(N)}_{-p+1} \\ \vdots \\ \boldsymbol{x}'^{(N)}_{N} \end{pmatrix} \tag{2.2}$$

denotes an $(N+p) \times n$ matrix of constants (regressors), with rows $\boldsymbol{x}_t^{\prime(N)}$, $t = -p+1, \ldots, N$, $\boldsymbol{\beta} = (\beta_{ij})$ is an $n \times m$ matrix of regression parameters,

$$\boldsymbol{U}^{(N)} = \begin{pmatrix} U_{-p+1,1}^{(N)} & U_{-p+1,2}^{(N)} & \cdots & U_{-p+1,m}^{(N)} \\ \vdots & \vdots & & \vdots \\ U_{N,1}^{(N)} & U_{N,2}^{(N)} & \cdots & U_{N,m}^{(N)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{U}_{-p+1}^{\prime(N)} \\ \vdots \\ \boldsymbol{U}_t^{\prime(N)} \\ \vdots \\ \boldsymbol{U}_N^{\prime(N)} \end{pmatrix} \tag{2.3}$$

is a nonobservable $(N+p) \times m$ matrix of random errors, and

$$\boldsymbol{Y}^{(N)} = \begin{pmatrix} Y_{-p+1,1}^{(N)} & Y_{-p+1,2}^{(N)} & \cdots & Y_{-p+1,m}^{(N)} \\ \vdots & \vdots & & \vdots \\ Y_{N,1}^{(N)} & Y_{N,2}^{(N)} & \cdots & Y_{N,m}^{(N)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Y}_{-p+1}^{\prime(N)} \\ \vdots \\ \boldsymbol{Y}_t^{\prime(N)} \\ \vdots \\ \boldsymbol{Y}_N^{\prime(N)} \end{pmatrix} \tag{2.4}$$

is the $(N+p) \times m$ matrix of observations. $\boldsymbol{Y}_t^{(N)}$ is thus the m-variate observation made at time t, and satisfies

$$\boldsymbol{Y}_t^{\prime(N)} = \boldsymbol{x}_t^{\prime(N)} \boldsymbol{\beta} + \boldsymbol{U}_t^{\prime(N)}, \qquad t = -p+1, \ldots, N. \tag{2.5}$$

Instead of the classical assumption that $(\boldsymbol{U}_{-p+1}^{(N)}, \ldots, \boldsymbol{U}_N^{(N)})$ is (Gaussian) white noise, we assume that it constitutes a finite realization of some solution of the multivariate stochastic difference equation (ARMA model)

$$\boldsymbol{U}_t - \sum_{i=1}^{p} \boldsymbol{A}_i \boldsymbol{U}_{t-i} = \boldsymbol{\varepsilon}_t + \sum_{i=1}^{q} \boldsymbol{B}_i \boldsymbol{\varepsilon}_{t-i}, \qquad t \in \mathbb{Z}, \tag{2.6}$$

where $\boldsymbol{A}_i$, $i = 1, \ldots, p$ and $\boldsymbol{B}_i$, $i = 1, \ldots, q$ are $m \times m$ real matrices and $\boldsymbol{\varepsilon}_t$ denotes an m-variate white noise, i.e. a process of independently, identically distributed random variables with density f, mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$.

We do not require however (2.6) to be an ARMA model of orders p and q: p and q here are upper bounds for the orders of ARMA dependence, in the sense that

$$|\boldsymbol{A}_{p_1}| \neq 0, \quad \boldsymbol{A}_{p_1+1} = \cdots = \boldsymbol{A}_p = \mathbf{0}; \qquad |\boldsymbol{B}_{q_1}| \neq 0, \quad \boldsymbol{B}_{q_1+1} = \cdots = \boldsymbol{B}_q = \mathbf{0}$$

for some specified values of p_1 and q_1, $0 \leq p_1 \leq p$, $0 \leq q_1 \leq q$.

Further assumptions of course are needed in order to obtain asymptotic results. They can be divided into three groups: Assumptions (A1) and (A2) deal with the ARMA model (2.6); Assumptions (B1) and (B2) are related with the asymptotic behavior of the regression constants; Assumptions (C1)–(C4) with the regularity conditions to be satisfied by the noise density f.

Assumptions (A1) and (A2) are the usual assumptions on identifiable, causal and invertible ARMA models.

(A1) The roots of the determinantal equations $|\boldsymbol{A}(z)| = 0$ and $|\boldsymbol{B}(z)| = 0$, $z \in \mathbb{C}$, with

$$\boldsymbol{A}(z) = \boldsymbol{I} - \sum_{i=1}^{p_1} \boldsymbol{A}_i z^i \quad \text{and} \quad \boldsymbol{B}(z) = \boldsymbol{I} + \sum_{i=1}^{q_1} \boldsymbol{B}_i z^i$$

lie outside the unit disk, so that (2.6) is causal and invertible.
(A2) The greater common left divisor of $\boldsymbol{A}(z)$ and $\boldsymbol{B}(z)$ is the $m \times m$ identity matrix.

Assumptions (B1) and (B2) are essentially equivalent to Grenander's classical conditions (Hannan (1970), p. 77). Consider the $n \times n$ cross-product matrices $\boldsymbol{C}_i^{(N)} = (N-i)^{-1} \sum_{t=i+1}^{N} \boldsymbol{x}_t^{(N)} \boldsymbol{x}_{t-i}^{\prime(N)}$, $i = 0, 1, \ldots, N-1$: $\boldsymbol{C}_i^{(N)}$ constitutes a lagged version of $\boldsymbol{C}_0^{(N)} = N^{-1} \sum_{t=1}^{N} \boldsymbol{x}_t^{(N)} \boldsymbol{x}_t^{\prime(N)}$, the diagonal elements of which can be assumed to be strictly positive; let $\boldsymbol{D}^{(N)}$ denote the diagonal matrix with elements $(\boldsymbol{C}_0^{(N)})_{jj}$, $j = 1, \ldots, n$.

(B1) Defining $\boldsymbol{R}_i^{(N)} = (\boldsymbol{D}^{(N)})^{-1/2} \boldsymbol{C}_i^{(N)} (\boldsymbol{D}^{(N)})^{-1/2}$, $\lim_{N\to\infty} \boldsymbol{R}_i^{(N)} = \boldsymbol{R}_i$, where $\boldsymbol{R}_0$ is positive definite and thus factorizes into $\boldsymbol{R}_0 = (\boldsymbol{K}\boldsymbol{K}')^{-1}$, with $\boldsymbol{K}$ a full rank symmetric $n \times n$ matrix. Put $\boldsymbol{K}^{(N)} = (\boldsymbol{D}^{(N)})^{-1/2} \boldsymbol{K}$.
(B2) The classical Noether conditions hold for all j: letting $\bar{x}_j^{(N)} = N^{-1} \sum_{t=1}^{N} x_{t,j}$,

$$\text{(2.7)} \qquad \lim_{N\to\infty} \max_{1 \le t \le N} \left\{ (x_{t,j}^{(N)} - \bar{x}_j^{(N)})^2 \Big/ \sum_{t=1}^{N} (x_{t,j}^{(N)} - \bar{x}_j^{(N)})^2 \right\} = 0, \qquad j = 1, \ldots, n.$$

Note that (B1) and (B2) jointly imply that $\sup_i \|\boldsymbol{R}_i\| < \infty$ ($\|\cdot\|$ denotes the spectral norm).

The regularity Assumptions (C1)–(C5) on innovation densities are as follows.

(C1) f is a nowhere vanishing continuous density (with respect to the Lebesgue measure μ on $\mathbb{R}^m$), with $\int \boldsymbol{x} f(\boldsymbol{x}) d\mu = \boldsymbol{0}$ and $\int \boldsymbol{x}\boldsymbol{x}' f(\boldsymbol{x}) d\mu = \boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ is positive definite, with diagonal elements $\boldsymbol{\Sigma}_{ii} = \sigma_i^2$, $i = 1, \ldots, m$.
(C2) There exists a square integrable random vector $\boldsymbol{D}f^{1/2}$ such that for all $\boldsymbol{0} \neq \boldsymbol{h} \to \boldsymbol{0}$

$$\text{(2.8)} \qquad (\boldsymbol{h}'\boldsymbol{h})^{-1} \int [f^{1/2}(\boldsymbol{x}+\boldsymbol{h}) - f^{1/2}(\boldsymbol{x}) - \boldsymbol{h}'\boldsymbol{D}f^{1/2}(\boldsymbol{x})]^2 d\mu \to 0,$$

i.e., $f^{1/2}$ is mean-square differentiable, with mean square gradient $\boldsymbol{D}f^{1/2}$.
(C3) Letting $\boldsymbol{\varphi}(\boldsymbol{x}) = (\varphi_1(\boldsymbol{x}), \ldots, \varphi_m(\boldsymbol{x}))' = -2\boldsymbol{D}f^{1/2}(\boldsymbol{x})/f^{1/2}(\boldsymbol{x})$, $\int [\varphi_i(\boldsymbol{x})]^4 f(\boldsymbol{x}) d\mu < \infty$, $i = 1, \ldots, m$.

Assumptions (C1) and (C3) guarantee the finiteness of second order moments and the (generalized) *Fisher information matrix* $\boldsymbol{\mathcal{I}}(f) = \int \boldsymbol{\varphi}(\boldsymbol{x})\boldsymbol{\varphi}'(\boldsymbol{x})f(\boldsymbol{x})d\mu$, which reduces to the covariance matrix of $-\boldsymbol{\varphi} = \mathbf{grad}(\log f)$ whenever f is derivable in the usual sense (except perhaps at a finite number of points). As can easily be checked, $\int \boldsymbol{\varphi}(\boldsymbol{x})f(\boldsymbol{x})d\mu = \mathbf{0}$ and $\int \boldsymbol{x}\boldsymbol{\varphi}'(\boldsymbol{x})f(\boldsymbol{x})d\mu = \boldsymbol{I}$. As in the one-dimensional case, $\boldsymbol{\Sigma}$ can be specified up to a positive factor only. Assumption (C2) is a multivariate version of the more familiar, one-dimensional, quadratic mean differentiability condition. It may seem difficult to be checked for: let us show that it is strictly equivalent to the existence of "partial" quadratic mean derivatives, the existence of which is easier to verify.

LEMMA 2.1. *A square-integrable function $g : \mathbb{R}^m \to \mathbb{R}$ is mean-square differentiable (i.e. there exists a square-integrable vector $\boldsymbol{D}g$ such that $(\boldsymbol{h}'\boldsymbol{h})^{-1} \int [g(\boldsymbol{x}+\boldsymbol{h}) - g(\boldsymbol{x}) - \boldsymbol{h}'\boldsymbol{D}g(\boldsymbol{x})]^2 d\mu \to 0$ for all $\boldsymbol{h} \to \mathbf{0}$, $\boldsymbol{h} \neq \mathbf{0}$) iff it admits partial quadratic mean derivatives (i.e. iff there exist m square-integrable functions $D_i g$, $i = 1, \ldots, m$, such that*

$$h^{-2} \int [g(x_1, \ldots, x_i + h, \ldots, x_m) - g(\boldsymbol{x}) - hD_i g(\boldsymbol{x})]^2 d\mu \to 0 \tag{2.9}$$

for all $h \to 0$, $h \neq 0$). Moreover, $\boldsymbol{D}g$ can be chosen as $(D_1 g, \ldots, D_m g)'$.

Note that such a property does not hold for the classical differentiability concept; for a proof, see Subsection 4.1. Another set of sufficient, simpler but slightly tighter, assumptions results from replacing (C2) with

(C2$'$) The gradient $\mathbf{grad}(f)$ exists at all but a finite number of points,

and defining $\boldsymbol{\varphi}$ in (C3) as $\boldsymbol{\varphi} = -\mathbf{grad}(f)/f$. (C1) and (C2$'$) indeed jointly imply that $f^{1/2}$ admits partial derivatives at all but a finite number of points; these partial derivatives in view of (C3) are square integrable. It follows (Malliavin (1982), pp. 123–124) that $f^{1/2}$ admits partial quadratic mean derivatives, hence (Lemma 2.1) that it is mean-square differentiable. Associated with mean-square differentiability, we also need the following result (cf. Subsection 4.1 for a proof).

LEMMA 2.2. *Assume that* (C1)–(C3) *hold. Then*

(i) *for all $\boldsymbol{h} \in \mathbb{R}^m$, $\int_{\mathbb{R}^m} [f^{1/2}(\boldsymbol{x}+\boldsymbol{h}) - f^{1/2}(\boldsymbol{x}) - \boldsymbol{h}'\boldsymbol{D}f^{1/2}(\boldsymbol{x})]^2 d\mu \leq \boldsymbol{h}'\boldsymbol{\mathcal{I}}(f)\boldsymbol{h}$.*

(ii) *for all $C > 0$ and all sequence $(\boldsymbol{h}_1^{(n)}, \boldsymbol{h}_2^{(n)})$ such that $\sup_n \|\boldsymbol{h}_1^{(n)}\| < \infty$ and $\boldsymbol{h}_2^{(n)} \to \mathbf{0}$, $n \to \infty$, $\boldsymbol{h}_1^{(n)} + \boldsymbol{h}_2^{(n)} \neq \mathbf{0}$,*

$$\lim_{n\to\infty} \sup_{\|\boldsymbol{h}\|\leq C} \int \|n^{-1/2}\boldsymbol{h} + \boldsymbol{h}^{(n)}\|^{-2} [f^{1/2}(\boldsymbol{x} + n^{-1/2}\boldsymbol{h} + \boldsymbol{h}^{(n)}) - f^{1/2}(\boldsymbol{x}) - (n^{-1/2}\boldsymbol{h} + \boldsymbol{h}^{(n)})'\boldsymbol{D}f^{1/2}(\boldsymbol{x})]^2 d\mu = 0.$$

Finally, we need an assumption on the joint distribution of "initial" (unobserved) values of the noise and the observations at time $-p+1, \ldots, -1, 0$, and another one guaranteeing that their asymptotic influence is asymptotically negligible.

(C4) The joint distribution of $(\boldsymbol{\varepsilon}_{-q+1}, \ldots, \boldsymbol{\varepsilon}_0, \boldsymbol{Y}_{-p+1}^{(N)}, \ldots, \boldsymbol{Y}_0^{(N)})$ admits a nowhere vanishing density $f_0(\cdot; \boldsymbol{\beta}, \boldsymbol{A}, \boldsymbol{B})$. If $\boldsymbol{A}^{(N)}(L) = \boldsymbol{I} - \sum_{i=1}^{p_1} \boldsymbol{A}_i^{(N)} L^i$, $\boldsymbol{B}^{(N)}(L) = \boldsymbol{I} + \sum_{i=1}^{p_1} \boldsymbol{B}_i^{(N)} L^i$ and $\boldsymbol{\beta}^{(N)}$ are such that $\boldsymbol{A}_i^{(N)} \to \boldsymbol{A}_i$, $\boldsymbol{B}_i^{(N)} \to \boldsymbol{B}_i$, $\boldsymbol{\beta}^{(N)} \to \boldsymbol{\beta}$ for $N \to \infty$, then

$$f_0(\boldsymbol{\varepsilon}_{-q+1}, \ldots, \boldsymbol{\varepsilon}_0, \boldsymbol{Y}_{-p+1}^{(N)}, \ldots, \boldsymbol{Y}_0^{(N)}; \boldsymbol{\beta}^{(N)}, \boldsymbol{A}^{(N)}, \boldsymbol{B}^{(N)}) \\ - f_0(\boldsymbol{\varepsilon}_{-q+1}, \ldots, \boldsymbol{\varepsilon}_0, \boldsymbol{Y}_{-p+1}^{(N)}, \ldots, \boldsymbol{Y}_0^{(N)}; \boldsymbol{\beta}, \boldsymbol{A}, \boldsymbol{B})$$

is $o_P(1)$, as $N \to \infty$, for $\boldsymbol{Y}^{(N)}$ satisfying (2.1) and (2.6).

(C5) The score function $\boldsymbol{\varphi}$ is piecewise Lipschitz, i.e. there exists a finite, measurable partition of $\mathbb{R}^m$ into J nonoverlapping subsets I_j, $j = 1, \ldots, J$ such that $\|\boldsymbol{\varphi}(\boldsymbol{x}) - \boldsymbol{\varphi}(\boldsymbol{y})\| \leq K\|\boldsymbol{x} - \boldsymbol{y}\|$ for all $\boldsymbol{x}$, $\boldsymbol{y}$ in I_j, $j = 1, \ldots, J$.

This latter assumption is weaker than Kreiss ((1987), p. 118)'s global Lipschitz assumption; its univariate version would be satisfied, e.g. by double exponential densities, which are excluded under Kreiss' setting.

The error process $\boldsymbol{U}_t^{(N)}$ is not required to be a realization of the *stationary* solution of (2.6) which, under (A1), exists and is unique. However, (C4) implies that the set of initial values determining $\boldsymbol{U}_t^{(N)}$ is bounded in probability, which, jointly with (A1), entails that $\boldsymbol{U}_t^{(N)}$ is a.s. asymptotically stationary, with a cross-covariance function $E(\boldsymbol{U}_t^{(N)} \boldsymbol{U}_{t-i}^{\prime(N)})$ converging, as $(t, N) \to \infty$, to the cross-covariance function $\boldsymbol{\Gamma}_i^U$ of (2.6)'s stationary solution. This possible, transient nonstationarity of $\boldsymbol{U}_t^{(N)}$ however has no influence upon asymptotic results and, for simplicity, $\boldsymbol{U}_t^{(N)}$ in the sequel is treated as if it were the stationary solution of (2.6).

2.2 *Swensen's lemma*

In order to prove LAN, we shall use a slight modification of a lemma of Swensen (1985), itself relying on LeCam (preliminary draft of his 1986 book, Chapter 10.5) and McLeish (1974). The definition of LAN considered in Swensen (1985) indeed is too weak for several statistical applications, where the stronger form given, e.g. in LeCam ((1986), pp. 272–274), LeCam and Yang ((1990), Section 5) or Fabian and Hannan (1982) is required. This latter form is the one considered below.

Denote by $P_1^{(N)}$ and $P_0^{(N)}$ two sequences of probability measures on measurable spaces $(\mathcal{X}^{(N)}, \mathcal{A}^{(N)})$. For all N, let $\mathcal{A}_t^{(N)} \subset \mathcal{A}_{t+1}^{(N)}$ be a filtration such that $\mathcal{A}_N^{(N)} = \mathcal{A}^{(N)}$, and denote by $P_{1,t}^{(N)}$ and $P_{0,t}^{(N)}$ the restrictions to $\mathcal{A}_t^{(N)}$ of $P_1^{(N)}$ and $P_0^{(N)}$, respectively. Assuming that $P_{1,t}^{(N)}$ is absolutely continuous (on $\mathcal{A}_t^{(N)}$) with respect to $P_{0,t}^{(N)}$, let $\alpha_0^{(N)} = 1$, $\alpha_t^{(N)} = dP_{1,t}^{(N)}/dP_{0,t}^{(N)}$ and $\xi_t^{(N)} = (\alpha_t^{(N)}/\alpha_{t-1}^{(N)})^{1/2} - 1$.

LEMMA 2.3. (Swensen (1985)) *Assume that the random variables* $\zeta_t^{(N)}$ *satisfy the following conditions (all convergences are in* $P_0^{(N)}$*-probability, as* $N \to \infty$*;*

expectations also are taken with respect to $P_0^{(N)}$*):* (i) $E\sum_{t=1}^N(\zeta_t^{(N)}-\xi_t^{(N)})^2 \to 0$; (ii) $\sup_N E\sum_{t=1}^N(\zeta_t^{(N)})^2 < \infty$; (iii) $\max_{1\le t\le N}|\zeta_t^{(N)}| \to 0$; (iv) $\sum_{t=1}^N(\zeta_t^{(N)})^2 - (\tau^{(N)})^2/4 \to 0$ *for some nonrandom sequence* $(\tau^{(N)},\ N\in\mathbb{N})$ *such that* $\sup_N(\tau^{(N)})^2 < \infty$; (v) $\sum_{t=1}^N E\{(\zeta_t^{(N)})^2 I[|\zeta_t^{(N)}| > 1/2] \mid \mathcal{A}_{t-1}^{(N)}\} \to 0$; (vi) $E[\zeta_t^{(N)} \mid \mathcal{A}_{t-1}^{(N)}] = 0$; (vii) $\sum_{t=1}^N E[(\xi_t^{(N)})^2 + 2\xi_t^{(N)} \mid \mathcal{A}_{t-1}^{(N)}] \to 0$. *Then, under* $P_0^{(N)}$*, as* $N\to\infty$,

$$\Lambda^{(N)} = \log(dP_1^{(N)}/dP_0^{(N)}) = 2\sum_{t=1}^N \zeta_t^{(N)} - (\tau^{(N)})^2/2 + o_P(1), \tag{2.10}$$

and the distribution of $[\Lambda^{(N)} + (\tau^{(N)})^2/2]/\tau^{(N)}$ *is asymptotically standard normal.*

The only difference with Swensen's original result is that the sequence $\tau^{(N)}$ there is a constant sequence. It is easy to see, however, that Swensen's proof still holds, with very minor modifications, in the case of a bounded sequence; and so does also LeCam's aforementioned result. In the sequel, $(\mathcal{X}^{(N)}, \mathcal{A}^{(N)})$ is $\mathbb{R}^{Nm}$, along with the corresponding Borel σ-field $\mathcal{B}^{Nm}$, and the filtration $\mathcal{A}_t^{(N)} = \mathcal{B}^{tm}$ associated with $\mathbb{R}^{Nm}$'s components one through tm, $t = 1,\ldots,N$.

3. Local asymptotic normality

3.1 *Further preparation*

Denote by $\boldsymbol{A}^{(N)}(L) = \boldsymbol{I} - \sum_{i=1}^p \boldsymbol{A}_i^{(N)}L^i$ and $\boldsymbol{B}^{(N)}(L) = \boldsymbol{I} + \sum_{i=1}^q \boldsymbol{B}_i^{(N)}L^i$, $N = 1,2,\ldots$, two sequences of difference operators satisfying (A1) (with orders p_2 and q_2 instead of p_1 and q_1, $p_1 \le p_2 \le p$ and $q_1 \le q_2 \le q$). Let $\boldsymbol{H}_u^{(N)}$, $u = 0,1,\ldots$ be the Green's matrices associated with $\boldsymbol{B}^{(N)}(L)$, so that $\boldsymbol{B}^{(N)}(L)\boldsymbol{\psi}_t = \boldsymbol{\eta}_t$, $t\ge 1$ iff

$$\begin{aligned}\boldsymbol{\psi}_t = \sum_{i=0}^{t-1}\boldsymbol{H}_i^{(N)}\boldsymbol{\eta}_{t-i} + (\boldsymbol{H}_t^{(N)}\cdots\boldsymbol{H}_{t+q_2-1}^{(N)})\\ \cdot\begin{pmatrix}\boldsymbol{I} & \boldsymbol{B}_1^{(N)} & \cdots & \boldsymbol{B}_{q_2-1}^{(N)}\\ \vdots & & & \vdots\\ & & & \boldsymbol{B}_1^{(N)}\\ \boldsymbol{0} & & & \boldsymbol{I}\end{pmatrix}\begin{pmatrix}\boldsymbol{\psi}_0\\ \boldsymbol{\psi}_{-1}\\ \vdots\\ \boldsymbol{\psi}_{-q_2+1}\end{pmatrix}, \quad t\ge 1\end{aligned} \tag{3.1}$$

(see e.g. Hallin (1984 or 1986)). Finally, denote by $\boldsymbol{\beta}^{(N)}$ a sequence of regression coefficient matrices. Define

$$(3.2)\quad e_t^{(N)} = \begin{cases} e_t^0 & t = 0, -1, \dots, -q+1 \\ \sum_{i=0}^{t-1} \boldsymbol{H}_i^{(N)}[\boldsymbol{A}^{(N)}(L)(\boldsymbol{Y}_{t-i}^{(N)} - \boldsymbol{\beta}'^{(N)}\boldsymbol{x}_{t-i}^{(N)})] & \\ \quad +(\boldsymbol{H}_t^{(N)} \cdots \boldsymbol{H}_{t+q_2-1}^{(N)}) & \\ \quad \cdot \begin{pmatrix} \boldsymbol{I} & \cdots & \boldsymbol{B}_{q_2-1}^{(N)} \\ & & \vdots \\ \boldsymbol{0} & & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} e_0^0 \\ e_{-1}^0 \\ \vdots \\ e_{-q_2+1}^0 \end{pmatrix} & t \geq 1, \end{cases}$$

where $e'^0 = (e_0'^0, e_{-1}'^0, \dots, e_{-q+1}'^0)$ is a q-tuple, independent of $\{\varepsilon_t, t \geq 1\}$, of i.i.d. vectors with density f. Obviously, if $e_t^0 = \varepsilon_t$, $t = 0, \dots, -q+1$ and $\boldsymbol{Y}_t^{(N)}$ satisfies (2.5) and (2.6), with $\boldsymbol{A}^{(N)}(L)$, $\boldsymbol{B}^{(N)}(L)$ and $\boldsymbol{\beta}^{(N)}$ instead of $\boldsymbol{A}(L)$, $\boldsymbol{B}(L)$ and $\boldsymbol{\beta}$, then $e_t^{(N)} = \varepsilon_t$ for all t. Similarly, let

$$(3.3)\quad e_t^{(0)} = \begin{cases} e_t^0 & t = 0, -1, \dots, -q+1 \\ \sum_{i=0}^{t-1} \boldsymbol{H}_i[\boldsymbol{A}(L)(\boldsymbol{Y}_{t-i}^{(N)} - \boldsymbol{\beta}'\boldsymbol{x}_{t-i}^{(N)})] & \\ \quad +(\boldsymbol{H}_t \cdots \boldsymbol{H}_{t+q_1-1}) & \\ \quad \cdot \begin{pmatrix} \boldsymbol{I} & \cdots & \boldsymbol{B}_{q_1-1} \\ & & \vdots \\ \boldsymbol{0} & & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} e_0^0 \\ \vdots \\ e_{-q_1+1}^0 \end{pmatrix} & t \geq 1, \end{cases}$$

where the Green's matrices associated with $\boldsymbol{B}(L)$ are denoted by $\boldsymbol{H}_u$. Further, denote by $H_f^{(N)}(\boldsymbol{\theta})$ the hypothesis under which the observation $\boldsymbol{Y}^{(N)}$ has been generated by (2.1) and (2.6) with parameter $\boldsymbol{\theta} = ((\mathrm{vec}\,\boldsymbol{\beta}')'(\mathrm{vec}\,\boldsymbol{A}_1)' \cdots (\mathrm{vec}\,\boldsymbol{A}_{p_1})'\boldsymbol{0}'_{(p-p_1)m^2\times 1}(\mathrm{vec}\,\boldsymbol{B}_1)' \cdots (\mathrm{vec}\,\boldsymbol{B}_{q_1})'\boldsymbol{0}'_{(q-q_1)m^2\times 1})'$ and innovation density f. Particularizing the sequences $\boldsymbol{\beta}^{(N)}$, $\boldsymbol{A}^{(N)}(L)$ and $\boldsymbol{B}^{(N)}(L)$, let $\boldsymbol{\beta}^{(N)} = \boldsymbol{\beta} + N^{-1/2}\boldsymbol{K}^{(N)}\boldsymbol{b}^{(N)}$,

$$(3.4)\quad \boldsymbol{A}_i^{(N)} = \begin{cases} \boldsymbol{A}_i + N^{-1/2}\boldsymbol{\gamma}_i^{(N)} & i = 1, \dots, p_1 \\ N^{-1/2}\boldsymbol{\gamma}_i^{(N)} & i = p_1+1, \dots, p, \end{cases}$$

$$(3.5)\quad \boldsymbol{B}_i^{(N)} = \begin{cases} \boldsymbol{B}_i + N^{-1/2}\boldsymbol{\delta}_i^{(N)} & i = 1, \dots, q_1 \\ N^{-1/2}\boldsymbol{\delta}_i^{(N)} & i = q_1+1, \dots, q \end{cases}$$

and $\boldsymbol{t}^{(N)} = ((\mathrm{vec}\,\boldsymbol{b}^{(N)})'(\mathrm{vec}\,\boldsymbol{\gamma}_1^{(N)})' \cdots (\mathrm{vec}\,\boldsymbol{\gamma}_p^{(N)})'(\mathrm{vec}\,\boldsymbol{\delta}_1^{(N)})' \cdots (\mathrm{vec}\,\boldsymbol{\delta}_q^{(N)})')'$, so that

$$\boldsymbol{\theta}^{(N)} = \boldsymbol{\theta} + N^{-1/2}\begin{pmatrix} \boldsymbol{K}^{(N)} \otimes \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \boldsymbol{t}^{(N)}.$$

The sequences $\boldsymbol{b}^{(N)}$, $\boldsymbol{\gamma}_i^{(N)}$, $\boldsymbol{\delta}_i^{(N)}$ are assumed to be such that $\sup_N(\boldsymbol{t}'^{(N)}\boldsymbol{t}^{(N)}) < \infty$.

Writing exact likelihood functions for the observed series $\boldsymbol{Y}^{(N)}$, as always in time-series analysis, involves unpleasant starting values problems. The likelihoods we are using here actually are the joint likelihoods of $(\boldsymbol{\varepsilon}_{-q+1}, \ldots, \boldsymbol{\varepsilon}_0, \boldsymbol{Y}^{(N)})$, from which exact likelihoods could be obtained on averaging the starting values out. As in Kreiss (1987), it follows from (C4) and (C5) below that these starting values have no influence on asymptotic results, so that their presence safely can be ignored; for a formal proof of this fact in the univariate case, see Hallin and Puri (1994), Lemma 4.2. Note that Swensen (1985), who deals with AR processes only, skips the problem by considering conditional likelihoods. The logarithm of the likelihood ratio for $H_f^{(N)}(\boldsymbol{\theta})$ against $H_f^{(N)}(\boldsymbol{\theta}^{(N)})$ then takes the form (o_P's are taken under $H_f^{(N)}(\boldsymbol{\theta})$)

$$\Lambda^{(N)}(\boldsymbol{Y}^{(N)}) = \sum_{t=1}^{N} \log[f(\boldsymbol{e}_t^{(N)})/f(\boldsymbol{e}_t^{(0)})] + o_P(1), \qquad N \to \infty, \tag{3.6}$$

where $\boldsymbol{e}_t^{(N)}$ and $\boldsymbol{e}_t^{(0)}$ are computed from (3.2) and (3.3), respectively, with arbitrary starting values $\boldsymbol{e}_t^0$ $(t = 0, -1, \ldots, -q+1)$; in order to fix the ideas, put e.g. $\boldsymbol{e}_0^0 = \cdots = \boldsymbol{e}_{-q+1}^0 = \mathbf{0}$.

3.2 *Main result*

A basic statistical tool in the traditional analysis of (explicity or implicity Gaussian) multiple time series is the observed (residual) cross-covariance function, i.e. the family of matrices

$$\boldsymbol{\Gamma}_i^{(N)} = (N-i)^{-1} \sum_{t=i+1}^{N} \boldsymbol{e}_t^{(0)} \boldsymbol{e}_{t-i}'^{(0)}, \qquad i = 1, \ldots, N-1. \tag{3.7}$$

This definition however has to be adapted to the present non Gaussian context: the matrices

$$\boldsymbol{\Gamma}_{i;f}^{(N)} = (N-i)^{-1} \sum_{t=i+1}^{N} \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \boldsymbol{e}_{t-i}'^{(0)}, \qquad i = 1, \ldots, N-1 \tag{3.8}$$

will be shown to play the same role and to admit the same interpretation, under innovation densities f, as classical cross-covariance matrices under Gaussian densities. Accordingly, we define $\boldsymbol{\Gamma}_{i;f}^{(N)}$ as the *f-cross covariance matrix of lag i*. Note that the past and future in (3.8) do not play symmetric roles, as they do in (3.7); this is in accordance with the fact that Gaussian ARMA processes are the only time-reversible ones (see Weiss (1975); Hallin *et al.* (1988)). Of course, in case of a Gaussian f, $\boldsymbol{\varphi}(\boldsymbol{e}) = \boldsymbol{e}$, so that (3.7) and (3.8) coincide.

In the nonserial part of the problem, the main role will be played by statistics of the form

$$\boldsymbol{T}_{i;f}^{(N)} = (N-i)^{-1} \sum_{t=i+1}^{N} \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \boldsymbol{x}_{t-i}'^{(N)} \boldsymbol{K}^{(N)}, \qquad i = 0, 1, \ldots, N-1. \tag{3.9}$$

Still because of the fact that $\boldsymbol{\varphi}(\boldsymbol{e}) = \boldsymbol{e}$ for Gaussian densities, the elements of $\boldsymbol{T}_{0;f}^{(N)}$ can be interpreted as generalized version of the numerators of classical t-statistics, those of $\boldsymbol{T}_{i;f}^{(N)}$, $i > 0$, as lagged versions thereof.

Since cross-covariances are such a familiar tool to most time-series analysts, it is desirable that the LAN result be expressed in terms of the matrices (3.8). Let $\boldsymbol{G}_u$ denote the Green's matrices associated with $\boldsymbol{A}(L)$. The following sequences will be used in relation with $\boldsymbol{\gamma}^{(N)}$ and $\boldsymbol{\delta}^{(N)}$:

$$(3.10)\quad \boldsymbol{d}_i^{(N)} = \sum_{j=1}^{\min(p,i)} \sum_{k=0}^{i-j} \sum_{\ell=0}^{\min(q_1,i-j-k)} \boldsymbol{H}_k \boldsymbol{\gamma}_j^{(N)} \boldsymbol{G}_{i-j-k-\ell} \boldsymbol{B}_\ell + \sum_{j=1}^{\min(q,i)} \boldsymbol{H}_{i-j} \boldsymbol{\delta}_j^{(N)},$$

with the convention $\boldsymbol{B}_0 = \boldsymbol{I}$. Finally, define

$$(3.11)\quad \zeta_t^{(N)} = \frac{1}{2} N^{-1/2} \sum_{j=0}^{t-1} \Bigg\{ \boldsymbol{H}_j \Bigg[\sum_{i=1}^{p} \boldsymbol{\gamma}_i^{(N)} (\boldsymbol{Y}_{t-i-j}^{(N)} - \boldsymbol{\beta}' \boldsymbol{x}_{t-i-j}^{(N)}) + \sum_{i=1}^{q} \boldsymbol{\delta}_i^{(N)} \boldsymbol{e}_{t-i-j}^{(0)} + \boldsymbol{A}(L) \boldsymbol{b}'^{(N)} \boldsymbol{K}'^{(N)} \boldsymbol{x}_{t-j}^{(N)} \Bigg] \Bigg\}' \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})$$

and, letting $\boldsymbol{h}_i = \boldsymbol{H}_i - \sum_{j=1}^{\min(p_1,i)} \boldsymbol{H}_{i-j} \boldsymbol{A}_j$,

$$(3.12)\quad (\tau^{(N)})^2 = \operatorname{tr} \Bigg\{ \boldsymbol{\mathcal{I}}(f) \sum_{i=1}^{\infty} \Bigg[\boldsymbol{d}_i^{(N)} \boldsymbol{\Sigma} \boldsymbol{d}_i'^{(N)} + \sum_{j=0}^{\infty} \boldsymbol{h}_j \boldsymbol{b}'^{(N)} \boldsymbol{K}' \boldsymbol{R}_{|i-j|} \boldsymbol{K} \boldsymbol{b}^{(N)} \boldsymbol{h}_j' \Bigg] \Bigg\}.$$

We now may state the LAN property which is the main result of this paper. This property is stated under three distinct forms. The first one (LAN 1) is reminiscent from Swensen's univariate result; it guarantees the existence of locally asymptotically optimal statistical procedures (e.g., the existence of *locally asymptotically maximin* or *most stringent* tests: see LeCam ((1986), Chapter 11)) but provides little information on their implementation, and is mainly related to the method of proof, based on Lemma 2.3. The second form (LAN 2) gives more information on the particular structure of the log-likelihood ratio, and gives an explicit expression for the central sequence $\boldsymbol{\Delta}^{(N)}(\boldsymbol{\theta})$, from which locally asymptotically optimal procedures can be easily derived (LeCam (1986), Chapter 11). In the third form (LAN 3), the central sequence $\boldsymbol{\Delta}^{(N)}(\boldsymbol{\theta})$ is expressed, as desired, in terms of the generalized residual cross-covariance matrices (3.8).

PROPOSITION 3.1. *Assume that* (A1), (A2), (B1), (B2), (C1)–(C4) *hold. Under* $H_f^{(N)}(\boldsymbol{\theta})$, *as* $N \to \infty$, (LAN 1)

$$(3.13)\quad \Lambda^{(N)}(\boldsymbol{Y}^{(N)}) = 2 \sum_{t=1}^{N} \zeta_t^{(N)} - \frac{1}{2} (\tau^{(N)})^2 + o_P(1),$$

and $(\Lambda^{(N)} + (\tau^{(N)})^2/2)/\tau^{(N)}$ *is asymptotically standard normal*; (LAN 2)

$$\Lambda^{(N)}(\boldsymbol{Y}^{(N)}) = \boldsymbol{t}'^{(N)}\boldsymbol{\Delta}^{(N)}(\boldsymbol{\theta}) - \frac{1}{2}\boldsymbol{t}'^{(N)}\boldsymbol{\Gamma}^{\boldsymbol{\Delta}}(\boldsymbol{\theta})\boldsymbol{t}^{(N)} + o_P(1), \tag{3.14}$$

with

(3.15)

$$\boldsymbol{\Delta}^{(N)}(\boldsymbol{\theta}) = \begin{pmatrix} \sum_{i=0}^{N-1} (\boldsymbol{I} \otimes \boldsymbol{h}_i)'(N-i)^{1/2} \operatorname{vec} \boldsymbol{T}_{i;f}^{(N)} \\ \sum_{i=1}^{N-1}\sum_{j=0}^{i-1}\sum_{k=0}^{\min(q_1,i-j-1)} [(\boldsymbol{G}_{i-j-k-1}\boldsymbol{B}_k) \otimes \boldsymbol{H}_j'](N-i)^{1/2} \operatorname{vec} \boldsymbol{\Gamma}_{i;f}^{(N)} \\ \vdots \\ \sum_{i=p}^{N-1}\sum_{j=0}^{i-p}\sum_{k=0}^{\min(q_1,i-j-p)} [(\boldsymbol{G}_{i-j-k-p}\boldsymbol{B}_k) \otimes \boldsymbol{H}_j'](N-i)^{1/2} \operatorname{vec} \boldsymbol{\Gamma}_{i;f}^{(N)} \\ \sum_{i=1}^{N-1} (\boldsymbol{I} \otimes \boldsymbol{H}_{i-1}')(N-i)^{1/2} \operatorname{vec} \boldsymbol{\Gamma}_{i;f}^{(N)} \\ \vdots \\ \sum_{i=q}^{N-1} (\boldsymbol{I} \otimes \boldsymbol{H}_{i-q}')(N-i)^{1/2} \operatorname{vec} \boldsymbol{\Gamma}_{i;f}^{(N)} \end{pmatrix}$$

and

$$\boldsymbol{\Gamma}^{\boldsymbol{\Delta}}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{\Gamma}_I^{\boldsymbol{\Delta}} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{\Gamma}_{II}^{\gamma\gamma}(i,i')) & (\boldsymbol{\Gamma}_{II}^{\gamma\delta}(i,i')) \\ \boldsymbol{0} & (\boldsymbol{\Gamma}_{II}^{\gamma\delta}(i,i')) & (\boldsymbol{\Gamma}_{II}^{\delta\delta}(i,i')) \end{pmatrix}, \tag{3.16}$$

where

$$\boldsymbol{\Gamma}_I^{\boldsymbol{\Delta}} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} (\boldsymbol{K}'\boldsymbol{R}_{|i-j|}\boldsymbol{K}) \otimes (\boldsymbol{h}_i'\boldsymbol{\mathcal{I}}(f)\boldsymbol{h}_j), \tag{3.17}$$

$$\begin{aligned} \boldsymbol{\Gamma}_{II}^{\gamma\gamma}(i,i') = \sum_{j=\max(i,i')}^{\infty} & \left[\sum_{k=0}^{j-i}\sum_{\ell=0}^{\min(q_1,j-i-k)} (\boldsymbol{G}_{j-i-k-\ell}\boldsymbol{B}_\ell) \otimes \boldsymbol{H}_k'\right] \\ & \times (\boldsymbol{\Sigma} \otimes \boldsymbol{\mathcal{I}}(f)) \left[\sum_{k=0}^{j-i'}\sum_{\ell=0}^{\min(q_1,j-i'-k)} (\boldsymbol{G}_{j-i'-k-\ell}\boldsymbol{B}_\ell) \otimes \boldsymbol{H}_k'\right], \\ & \qquad i,i' = 1,\ldots,p, \end{aligned} \tag{3.18}$$

$$\begin{aligned} \boldsymbol{\Gamma}_{II}^{\gamma\delta}(i,i') = \sum_{j=\max(i,i')}^{\infty} & \left[\sum_{k=0}^{j-i}\sum_{\ell=0}^{\min(q_1,j-i-k)} (\boldsymbol{G}_{j-i-k-\ell}\boldsymbol{B}_\ell) \otimes \boldsymbol{H}_k'\right] \\ & \times (\boldsymbol{\Sigma} \otimes \boldsymbol{\mathcal{I}}(f))[\boldsymbol{I} \otimes \boldsymbol{H}_{j-i'}], \quad i = 1,\ldots,p, \quad i' = 1,\ldots,q, \end{aligned}$$

$$\boldsymbol{\Gamma}_{II}^{\delta\delta}(i,i') = \sum_{j=\max(i,i')}^{\infty} [\boldsymbol{I} \otimes \boldsymbol{H}_{j-i}](\boldsymbol{\Sigma} \otimes \boldsymbol{\mathcal{I}}(f))[\boldsymbol{I} \otimes \boldsymbol{H}_{j-i'}], \qquad i,i' = 1,\ldots,q; \tag{3.19}$$

moreover, $\boldsymbol{\Delta}^{(N)}(\boldsymbol{\theta})$ *is asymptotically normal, with mean* $\mathbf{0}$ *and covariance matrix* $\boldsymbol{\Gamma}^{\Delta}(\boldsymbol{\theta})$; (*LAN 3*)

$$
\begin{aligned}
(3.20)\quad \Lambda^{(N)}(\boldsymbol{Y}^{(N)}) = & \sum_{i=0}^{N-1} (N-i)^{1/2} \operatorname{tr}[\boldsymbol{T}_{i;f}^{(N)} \boldsymbol{b}^{(N)} \boldsymbol{h}_i'] \\
& + \sum_{i=1}^{N-1} (N-i)^{1/2} \operatorname{tr}[\boldsymbol{\Gamma}_{i;f}^{(N)} \boldsymbol{d}_i'^{(N)}] - \frac{1}{2}(\tau^{(N)})^2 + o_P(1),
\end{aligned}
$$

and $(\Lambda^{(N)} + (\tau^{(N)})^2/2)/\tau^{(N)}$ *is asymptotically standard normal.*

Assumption (C5) so far has not been used. Whenever it holds, the influence of starting values upon $\sum_{t=1}^{N} \zeta_t^{(N)}$, $\boldsymbol{\Delta}^{(N)}(\boldsymbol{\theta})$ or $(N-i)^{1/2}\boldsymbol{T}_{i;f}^{(N)}$ is asymptotically negligible (cf. the remark at the end of Section 3.1).

4. Proofs

4.1 *Lemmas* 2.1 *and* 2.2

PROOF OF LEMMA 2.1. Denote by $L^2(\mathbb{R}^m)$ the set of square-integrable functions from $\mathbb{R}^m$ to $\mathbb{R}$. Assume that $g \in L^2(\mathbb{R}^n)$ admits partial derivatives in the mean square sense (2.9). For all $\boldsymbol{h} \in \mathbb{R}^m$, $\boldsymbol{x} \in \mathbb{R}^m$,

$$
\begin{aligned}
g(\boldsymbol{x}+\boldsymbol{h}) - g(\boldsymbol{x}) &= g\left(\boldsymbol{x} + \sum_{i=1}^{m} h_i \boldsymbol{e}_i\right) - g(\boldsymbol{x}) \\
&= \sum_{j=1}^{m} \left[g\left(\boldsymbol{x} + \sum_{i=0}^{j} h_i \boldsymbol{e}_i\right) - g\left(\boldsymbol{x} + \sum_{i=0}^{j-1} h_i \boldsymbol{e}_i\right)\right],
\end{aligned}
$$

where $h_0 = 0$, $\boldsymbol{e}_0 = \mathbf{0}$ and $\boldsymbol{e}_i$ denotes the i-th unit vector. Letting $\boldsymbol{D}g = (D_1 g, \ldots, D_m g)'$,

$$
\begin{aligned}
&(\boldsymbol{h}'\boldsymbol{h})^{-1} \int [g(\boldsymbol{x}+\boldsymbol{h}) - g(\boldsymbol{x}) - \boldsymbol{h}'\boldsymbol{D}g(\boldsymbol{x})]^2 d\mu \\
&= (\boldsymbol{h}'\boldsymbol{h})^{-1} \int \left\{ \sum_{j=1}^{m} \left[g\left(\boldsymbol{x} + \sum_{i=0}^{j} h_i \boldsymbol{e}_i\right) - g\left(\boldsymbol{x} + \sum_{i=0}^{j-1} h_i \boldsymbol{e}_i\right) \right.\right. \\
&\qquad \left. - h_j D_j g\left(\boldsymbol{x} + \sum_{i=0}^{j-1} h_i \boldsymbol{e}_i\right)\right] \\
&\qquad \left. + \sum_{j=1}^{m} \left[h_j D_j g\left(\boldsymbol{x} + \sum_{i=0}^{j-1} h_i \boldsymbol{e}_i\right) - h_j D_j g(\boldsymbol{x})\right]\right\}^2 d\mu \\
&\leq C \sum_{j=1}^{m} h_j^2 (\boldsymbol{h}'\boldsymbol{h})^{-1} h_j^{-2} \int \left[g\left(\boldsymbol{x} + \sum_{i=0}^{j} h_i \boldsymbol{e}_i\right) - g\left(\boldsymbol{x} + \sum_{i=0}^{j-1} h_i \boldsymbol{e}_i\right)\right.
\end{aligned}
$$

$$-h_jD_jg\left(\boldsymbol{x}+\sum_{i=0}^{j-1}h_i\boldsymbol{e}_i\right)\Bigg]^2 d\mu$$

$$+C\sum_{j=1}^{m}h_j^2(\boldsymbol{h}'\boldsymbol{h})^{-1}\int\left[D_jg\left(\boldsymbol{x}+\sum_{i=0}^{j-1}h_i\boldsymbol{e}_i\right)-D_jg(\boldsymbol{x})\right]^2 d\mu,$$

with $C > 0$. Now, the existence of mean square partial derivatives entails, for all $j \in \{1, \ldots, m\}$,

$$\lim_{h_j\to 0}(h_j)^{-2}\int\Bigg[g\left(\boldsymbol{x}+\sum_{i=0}^{j}h_i\boldsymbol{e}_i\right)-g\left(\boldsymbol{x}+\sum_{i=0}^{j-1}h_i\boldsymbol{e}_i\right)$$
$$-h_jD_jg\left(\boldsymbol{x}+\sum_{i=0}^{j-1}h_i\boldsymbol{e}_i\right)\Bigg]^2 d\mu = 0.$$

Due to the quadratic mean continuity of $D_jg \in L^2(\mathbb{R}^m)$,

$$\lim_{\boldsymbol{h}\to 0}\int\left[D_jg\left(\boldsymbol{x}+\sum_{i=0}^{j-1}h_i\boldsymbol{e}_i\right)-D_jg(\boldsymbol{x})\right]^2 d\mu = 0.$$

The mean square differentiability of g then follows from the fact that $h_j^2/\boldsymbol{h}'\boldsymbol{h} \leq 1$, $j \in \{1, \ldots, m\}$. The converse is straightforward.

PROOF OF LEMMA 2.2. (i) Since $\int[f^{1/2}(\boldsymbol{x}+\boldsymbol{h})-f^{1/2}(\boldsymbol{x})-\boldsymbol{h}'\boldsymbol{D}f^{1/2}(\boldsymbol{x})]^2 d\mu$ is bounded by

$$2\int[f^{1/2}(\boldsymbol{x}+\boldsymbol{h})-f^{1/2}(\boldsymbol{x})]^2 d\mu + 2\int[\boldsymbol{h}'\boldsymbol{D}f^{1/2}(\boldsymbol{x})]^2 d\mu$$
$$\leq 2\int\left[\int_0^1 \boldsymbol{h}'\boldsymbol{D}f^{1/2}(\boldsymbol{x}+\lambda\boldsymbol{h})d\lambda\right]^2 d\mu + \frac{1}{2}\boldsymbol{h}'\boldsymbol{\mathcal{I}}(f)\boldsymbol{h},$$

the result follows from the fact that

$$\int\left[\int_0^1 \boldsymbol{h}'\boldsymbol{D}f^{1/2}(\boldsymbol{x}+\lambda\boldsymbol{h})d\lambda\right]^2 d\mu \leq \int\int_0^1[\boldsymbol{h}'\boldsymbol{D}f^{1/2}(\boldsymbol{x}+\lambda\boldsymbol{h})]^2 d\lambda d\mu$$
$$= \int[\boldsymbol{h}'\boldsymbol{D}f^{1/2}(\boldsymbol{x})]^2 d\mu = \frac{1}{4}\boldsymbol{h}'\boldsymbol{\mathcal{I}}(f)\boldsymbol{h};$$

(ii) is an immediate consequence of (C2) and the definition of mean square differentiability.

4.2 *Some preliminary results*

The following technical lemmas will be needed in the proof of Proposition 3.1.

LEMMA 4.1. *For all* $t = 1, \dots, N$,

$$e_t^{(N)} - e_t^{(0)} = \sum_{i=0}^{t-1} H_i^{(N)} \{ (A^{(N)}(L) - A(L))(Y_{t-i}^{(N)} - \beta' x_{t-i}^{(N)}) - (B^{(N)}(L) - B(L)) e_{t-i}^{(0)} \} + \sum_{i=0}^{t-1} H_i^{(N)} A^{(N)}(L)(\beta' - \beta'^{(N)}) x_{t-i}^{(N)}. \tag{4.1}$$

LEMMA 4.2. *Let* $B^0(L)$ *satisfy Assumption* (A1). *There exists* $\eta > 0$ *such that* $\max_{1 \le i \le q_2} \|B_i - B_i^0\| < \eta$ *implies that* $B(L)$ *also satisfies* (A1).

LEMMA 4.3. *Let* $B_i^{(N)} = B_i + N^{-1/2} \delta_i^{(N)}$, *where* $\sup_N \|\delta_i^{(N)}\| < \infty$. *Then* $\sum_{i=0}^{\infty} \|H_i^{(N)} - H_i\|$ *is* $O(N^{-1/2})$, *as* $N \to \infty$.

PROOFS OF LEMMAS 4.1, 4.2 AND 4.3. Except for the presence of additional terms due to the trend component, the proofs essentially consist in adapting the corresponding univariate proofs of Kreiss (1987), where we refer to for details.

LEMMA 4.4. *Let* $(M_i, i \in \mathbb{N})$ *denote a sequence of* $m \times m$ *real matrices such that* $\sum_{i=0}^{\infty} \|M_i\| < \infty$ *and* $(V_i^{(N)}, i = 1, \dots, N; N \in \mathbb{N})$ *a triangular array of* $m \times 1$ *real vectors such that* $\lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} V_{i-k}^{(N)} V_{i-\ell}'^{(N)} = R(|k-\ell|)$, $k, \ell \in \mathbb{N}$, *with* $\sup_{i \in \mathbb{N}} \|R(i)\| < \infty$. *Then*

$$\lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} \sum_{k,\ell=0}^{i-1} M_k b'^{(N)} V_{i-k}^{(N)} V_{i-\ell}'^{(N)} b^{(N)} M_\ell' - \sum_{k,\ell=0}^{\infty} M_k b'^{(N)} R(|k-\ell|) b^{(N)} M_\ell' = \mathbf{0},$$

for all sequence $(b^{(N)}, N \in \mathbb{N})$ *of real* $m \times m$ *matrices such that* $\sup_N \|b^{(N)}\| < \infty$.

PROOF. The proof is elementary, and is left to the reader.

4.3 *Proof of Proposition* 3.1 (*LAN* 1)

The proof of the (LAN 1) part of Proposition 3.1 consists in showing that conditions (i)–(vii) in Swensen's Lemma (Lemma 2.3) are satisfied.

LEMMA 4.5. (Swensen's condition (i)) *Under* $H_f^{(N)}(\theta)$,

$$\lim_{N\to\infty} \sum_{t=1}^{N} E\left[\frac{1}{2}(e_t^{(0)} - e_t^{(N)})' \varphi(e_t^{(0)}) - \zeta_t^{(N)}\right]^2 = 0 \tag{4.2}$$

and

$$\lim_{N\to\infty}\sum_{t=1}^{N} E\left[\xi_t^{(N)} - \frac{1}{2}(\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)})'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\right]^2 = 0; \tag{4.3}$$

here $\xi_t^{(N)} = [f(\boldsymbol{e}_t^{(N)})/f(\boldsymbol{e}_t^{(0)})]^{1/2} - 1$, *and* $\zeta_t^{(N)}$ *is defined in* (3.11).

PROOF. Letting $\boldsymbol{\gamma}^{(N)}(L) = \sum_{i=1}^{p_2}\boldsymbol{\gamma}_i^{(N)}L^i$ and $\boldsymbol{\delta}^{(N)}(L) = \sum_{i=1}^{q_2}\boldsymbol{\delta}_i^{(N)}L^i$, define $\boldsymbol{Z}_t^{(N)} = 0$, $t \le 0$ and

$$\boldsymbol{Z}_t^{(N)} = \boldsymbol{\gamma}^{(N)}(L)(\boldsymbol{Y}_t^{(N)} - \boldsymbol{\beta}'\boldsymbol{x}_t^{(N)}) + \boldsymbol{\delta}^{(N)}(L)\boldsymbol{e}_t^{(0)}. \tag{4.4}$$

Under $H_f^{(N)}(\boldsymbol{\theta})$, $\boldsymbol{Z}_t^{(N)}$ for all $t = 1, \dots, N$ coincides with $\boldsymbol{\gamma}^{(N)}(L)\boldsymbol{U}_t^{(N)} + \boldsymbol{\delta}^{(N)}(L)\boldsymbol{\varepsilon}_t$, itself an asymptotically (as $t \to \infty$) stationary ARMA process. $\boldsymbol{Z}_t^{(N)}$ thus is also asymptotically stationary, uniformly in N, in the sense that there exist strictly stationary processes $\bar{\boldsymbol{Z}}_t^{(N)}$, $N = 1, \dots$ such that for all $\varepsilon > 0$ there is a $T : E[\|\bar{\boldsymbol{Z}}_t^{(N)} - \boldsymbol{Z}_t^{(N)}\|] < \varepsilon$, $N \ge t \ge T$ (this follows from the boundedness of the sequences $\boldsymbol{\gamma}_i^{(N)}$ and $\boldsymbol{\delta}_i^{(N)}$). The corresponding (asymptotic) cross-covariance matrices are $E[\boldsymbol{Z}_t^{(N)}\boldsymbol{Z}_{t-k}'^{(N)}] = E[\bar{\boldsymbol{Z}}_t^{(N)}\bar{\boldsymbol{Z}}_{t-k}'^{(N)}] + o(1)$ for $t \to \infty$, uniformly in N, so that

$$\sum_{i,j=0}^{\infty} \boldsymbol{H}_i E[\boldsymbol{Z}_t^{(N)}\boldsymbol{Z}_{t-|i-j|}'^{(N)}]\boldsymbol{H}_j' = \sum_{i=1}^{\infty} \boldsymbol{d}_i^{(N)}\boldsymbol{\Sigma}\boldsymbol{d}_i'^{(N)} + o(1). \tag{4.5}$$

Here again, the possible transient nonstationarity of $\boldsymbol{Z}_t^{(N)}$ has no influence upon asymptotic results, and $\boldsymbol{Z}_t^{(N)}$ in the sequel is treated as if it were stationary. With this notation,

$$\begin{aligned}
&\frac{1}{2}(\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)})'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \\
&\quad = \zeta_t^{(N)} + \frac{1}{2}N^{-1/2}\left[\sum_{i=1}^{t-1}(\boldsymbol{H}_i^{(N)} - \boldsymbol{H}_i)\boldsymbol{Z}_{t-i}^{(N)}\right]'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \\
&\qquad + \frac{1}{2}N^{-1/2}\left[\sum_{i=1}^{t-1}(\boldsymbol{H}_i^{(N)} - \boldsymbol{H}_i)\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right]'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \\
&\qquad - \frac{1}{2}N^{-1}\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i^{(N)}\boldsymbol{\gamma}^{(N)}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right]'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \\
&\quad = \zeta_t^{(N)} + \frac{1}{2}a_t^{(N)} + \frac{1}{2}b_t^{(N)} + \frac{1}{2}c_t^{(N)},
\end{aligned} \tag{4.6}$$

say. In order to prove (4.2), it is sufficient to show that $\sum_{t=1}^{N} E(a_t^{(N)})^2$, $\sum_{t=1}^{N} E(b_t^{(N)})^2$ and $\sum_{t=1}^{N} E(c_t^{(N)})^2$ converge to 0, under $H_f^{(N)}(\boldsymbol{\theta})$, as $N \to \infty$.

For $a_t^{(N)}$, we have

$$(4.7)\sum_{t=1}^{N} E(a_t^{(N)})^2$$
$$= N^{-1}\sum_{t=1}^{N}\mathrm{tr}\left\{\boldsymbol{\mathcal{I}}(f)\left[\sum_{i,j=1}^{t-1}(\boldsymbol{H}_i^{(N)}-\boldsymbol{H}_i)E(\boldsymbol{Z}_{t-i}^{(N)}\boldsymbol{Z}_{t-j}^{\prime(N)})(\boldsymbol{H}_j^{(N)}-\boldsymbol{H}_j)'\right]\right\}.$$

Due to (4.5), there exists a constant C_1 such that $\|E[\boldsymbol{Z}_{t-i}^{(N)}\boldsymbol{Z}_{t-j}^{\prime(N)}]\| \le C_1$, uniformly in N, t and $|i-j|$. Accordingly, for some other constant C_2,

$$\left\|\sum_{i,j=1}^{t-1}(\boldsymbol{H}_i^{(N)}-\boldsymbol{H}_i)E[\boldsymbol{Z}_{t-i}^{(N)}\boldsymbol{Z}_{t-j}^{\prime(N)}](\boldsymbol{H}_j^{(N)}-\boldsymbol{H}_j)'\right\|$$
$$\le C_2\left(\sum_{i=1}^{t-1}\|\boldsymbol{H}_i^{(N)}-\boldsymbol{H}_i\|\right)^2 < C_2\left(\sum_{i=1}^{\infty}\|\boldsymbol{H}_i^{(N)}-\boldsymbol{H}_i\|\right)^2.$$

This latter quantity, according to Lemma 4.3, converges to 0 as $N\to\infty$, which, along with (4.7), yields the desired result. Next, let us consider

$$(4.8)\quad \sum_{t=1}^{N} E(b_t^{(N)})^2$$
$$= N^{-1}\sum_{t=1}^{N}\mathrm{tr}\left\{\boldsymbol{\mathcal{I}}(f)\left[\sum_{i=1}^{t-1}(\boldsymbol{H}_i^{(N)}-\boldsymbol{H}_i)\boldsymbol{A}(L)\boldsymbol{b}^{\prime(N)}\boldsymbol{K}^{\prime(N)}\boldsymbol{x}_{t-i}^{(N)}\right]\right.$$
$$\left.\times\left[\sum_{i=1}^{t-1}(\boldsymbol{H}_i^{(N)}-\boldsymbol{H}_i)\boldsymbol{A}(L)\boldsymbol{b}^{\prime(N)}\boldsymbol{K}^{\prime(N)}\boldsymbol{x}_{t-i}^{(N)}\right]'\right\}.$$

Here,

$$\|\boldsymbol{K}^{\prime(N)}\boldsymbol{x}_{t-i}^{(N)}\| = [\boldsymbol{x}_{t-i}^{\prime(N)}(\boldsymbol{D}^{(N)})^{-1/2}\boldsymbol{K}\boldsymbol{K}'(\boldsymbol{D}^{(N)})^{-1/2}\boldsymbol{x}_{t-i}^{(N)}]^{1/2}$$
$$\le C_3N^{1/2}\|\boldsymbol{R}_0^{-1}\|\left[\sum_{i=1}^{n}\left[(x_{t,i}^{(N)})^2\Big/\sum_{t=1}^{N}(x_{t,i}^{(N)})^2\right]\right]^{1/2},$$

a quantity which, from Assumption (B2), is $o(N^{1/2})$, as $N\to\infty$, uniformly in t. This, along with the fact that (Lemma 4.3) $N^{1/2}\sum_{i=1}^{\infty}\|\boldsymbol{H}_i^{(N)}-\boldsymbol{H}_i\|$ remains bounded as $N\to\infty$, implies that $\|\sum_{i=1}^{t-1}(\boldsymbol{H}_i^{(N)}-\boldsymbol{H}_i)\boldsymbol{A}(L)\boldsymbol{b}^{\prime(N)}\boldsymbol{K}^{\prime(N)}\boldsymbol{x}_{t-i}^{(N)}\|$ is $o(1)$ as $N\to\infty$, uniformly in t, which in turn implies that (4.8) converges to zero, as was to be shown. The convergence to zero of

$$(4.9)\quad \sum_{t=1}^{N} E(c_t^{(N)})^2$$
$$= N^{-2}\sum_{t=1}^{N}E\left\{\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i^{(N)}\boldsymbol{\gamma}^{(N)}(L)\boldsymbol{b}^{\prime(N)}\boldsymbol{K}^{\prime(N)}\boldsymbol{x}_{t-i}^{(N)}\right]'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\right\}^2$$

follows along the same lines.
(ii) We now turn to the proof of (4.3). The constants $C_1, C_2, \ldots$ used here are not related to those appearing in other parts of the paper; $I[A]$ stands for the indicator of A. For any positive C_1,

$$
\begin{aligned}
(4.10)\qquad &\sum_{t=1}^{N} E\left[\xi_t^{(N)} - \frac{1}{2}(\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)})'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\right]^2 \\
&= \sum_{t=1}^{N} E\Bigg\{ I\left[\left\|\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right\|^2 \le C_1\right] \\
&\qquad \cdot \left[\xi_t^{(N)} - \frac{1}{2}(\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)})'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\right]^2\Bigg\} \\
&\quad + \sum_{t=1}^{N} E\Bigg\{ I\left[\left\|\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right\|^2 > C_1\right] \\
&\qquad \cdot \left[\xi_t^{(N)} - \frac{1}{2}(\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)})'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\right]^2\Bigg\} \\
&= A_1 + A_2, \qquad \text{say.}
\end{aligned}
$$

Still from Lemma 4.1,

$$
\begin{aligned}
(4.11)\quad (\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)}) &= N^{-1/2}\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)} + N^{-1/2}\sum_{i=1}^{t-1} (\boldsymbol{H}_i^{(N)} - \boldsymbol{H}_i)\boldsymbol{Z}_{t-i}^{(N)} \\
&\quad + N^{-1/2}\sum_{i=0}^{t-1} \boldsymbol{H}_i^{(N)} \boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)} \\
&\quad - N^{-1}\sum_{i=0}^{t-1} \boldsymbol{H}_i^{(N)} \boldsymbol{\gamma}^{(N)}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}.
\end{aligned}
$$

From arguments similar to those used in the proof of (4.2), it follows that, under $H_f^{(N)}(\boldsymbol{\theta})$, uniformly in t, as $N \to \infty$,

$$
\begin{aligned}
&(\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)}) I\left[\left\|\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right\|^2 \le C_1\right] \\
&= N^{-1/2}\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)} I\left[\left\|\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right\|^2 \le C_1\right] + o(1),
\end{aligned}
$$

a quantity a.s. of the form $N^{-1/2}\boldsymbol{h}_1^{(N)} + \boldsymbol{h}_2^{(N)}$ appearing in Lemma 2.2 (ii). Accordingly,

$$
A_1 = \sum_{t=1}^{N} E\Bigg\{ I\left[\left\|\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right\|^2 \le C_1\right]
$$

$$
\begin{aligned}
&\qquad\cdot E\left\{\left[\xi_t^{(N)}-\frac{1}{2}(\boldsymbol{e}_t^{(0)}-\boldsymbol{e}_t^{(N)})'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\right]^2 \,\Big|\, \boldsymbol{e}_{t-1}^{(0)}\boldsymbol{e}_{t-2}^{(0)},\dots\right\}\Bigg\}\\
&=\sum_{t=1}^{N}E\Bigg\{I\left[\left\|\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right\|^2\le C_1\right]\\
&\qquad\cdot\int\Bigg[\frac{f^{1/2}(\boldsymbol{e}-\boldsymbol{e}_t^{(0)}+\boldsymbol{e}_t^{(N)})-f^{1/2}(\boldsymbol{e})}{f^{1/2}(\boldsymbol{e})}\\
&\qquad\qquad-\frac{1}{2}(\boldsymbol{e}_t^{(0)}-\boldsymbol{e}_t^{(N)})'\boldsymbol{\varphi}(\boldsymbol{e})\Bigg]^2 f(\boldsymbol{e})d\boldsymbol{e}\Bigg\}\\
&\le\sum_{t=1}^{N}E[\|\boldsymbol{e}_t^{(0)}-\boldsymbol{e}_t^{(N)}\|^2]C^{(N)}(C_1),
\end{aligned}
$$

for all $C_1>0$, where, according to Lemma 2.2 (ii),

$$
\begin{aligned}
C^{(N)}(C_1)=\sup_{\|\boldsymbol{h}_1^{(N)}\|\le C_1}&\int\left\|N^{-1/2}\boldsymbol{h}+\boldsymbol{e}_t^{(0)}-\boldsymbol{e}_t^{(N)}-N^{-1/2}\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right\|^{-2}\\
&\times\Bigg[f^{1/2}\left(\boldsymbol{e}-N^{-1/2}\boldsymbol{h}_1^{(N)}-\boldsymbol{e}_t^{(0)}+\boldsymbol{e}_t^{(N)}+N^{-1/2}\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right)\\
&\quad-f^{1/2}(\boldsymbol{e})-\Bigg(N^{-1/2}\boldsymbol{h}_1^{(N)}+\boldsymbol{e}_t^{(0)}-\boldsymbol{e}_t^{(N)}\\
&\qquad\qquad-N^{-1/2}\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\Bigg)'\boldsymbol{D}f^{1/2}(\boldsymbol{e})\Bigg]^2 d\mu
\end{aligned}
$$

converges to zero as $N\to\infty$, for fixed C_1. In order to establish that A_1 converges to zero, it is thus sufficient to show that $\sum_{t=1}^{N}E\|\boldsymbol{e}_t^{(0)}-\boldsymbol{e}_t^{(N)}\|^2$ remains bounded as $N\to\infty$. Using the triangular and Cauchy-Schwarz inequalities, we readily obtain that

$$
\begin{aligned}
&\sum_{t=1}^{N}E\|\boldsymbol{e}_t^{(0)}-\boldsymbol{e}_t^{(N)}\|^2\\
&\quad\le 5\Bigg\{N^{-1}\sum_{t=1}^{N}E\left\|\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right\|^2\\
&\qquad+N^{-1}\sum_{t=1}^{N}E\left\|\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right\|^2\\
&\qquad+N^{-1}\sum_{t=1}^{N}E\left\|\sum_{i=1}^{t-1}(\boldsymbol{H}_i^{(N)}-\boldsymbol{H}_i)\boldsymbol{Z}_{t-i}^{(N)}\right\|^2
\end{aligned}
$$

$$+ N^{-1}\sum_{t=1}^{N} E\left\|\sum_{i=1}^{t-1}(\boldsymbol{H}_i^{(N)} - \boldsymbol{H}_i)\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right\|^2$$

$$+ N^{-2}\sum_{t=1}^{N} E\left\|\sum_{i=0}^{t-1} \boldsymbol{H}_i^{(N)}\boldsymbol{\gamma}^{(N)}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right\|^2\Bigg\}.$$

Each of these five sums can be shown to converge. Considering for instance the first of them, $N^{-1}\sum_{t=1}^{N} E\|\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\|^2 - E\|\sum_{i=0}^{\infty}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\|^2$ is $o(1)$ as $N\to\infty$, due to the fact that $\sum_{i=0}^{\infty}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}$ is second-order stationary. The remaining terms are handled similarly.

Finally, consider the second term A_2 in (4.10); applying Lemma 2.2 (i), we obtain

$$A_2 = \sum_{t=1}^{N} E\Bigg\{I\left[\left\|\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right\|^2 > C_1\right]$$
$$\cdot\int[f^{1/2}(\boldsymbol{e} - \boldsymbol{e}_t^{(0)} + \boldsymbol{e}_t^{(N)}) - f^{1/2}(\boldsymbol{e})$$
$$- (\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)})'\boldsymbol{D}f^{1/2}(\boldsymbol{e})]^2 f(\boldsymbol{e})d\boldsymbol{e}\Bigg\}$$
$$\leq \sum_{t=1}^{N} E\left\{I\left[\left\|\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right\|^2 > C_1\right](\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)})'\boldsymbol{\mathcal{I}}(f)(\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)})\right\}.$$

Because of the L_1-convergence, for all t, of $\|\sum_{k=0}^{j}\boldsymbol{H}_k\boldsymbol{Z}_{t-k}^{(N)}\|^2 - \|\sum_{k=0}^{\infty}\boldsymbol{H}_k\boldsymbol{Z}_{t-k}^{(N)}\|^2$ to zero, the triangular array $\|\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\|^2$, $N\in\mathbb{Z}$, is uniformly integrable. Since $\sum_{t=1}^{N} E\|\boldsymbol{e}_t^{(0)} - \boldsymbol{e}_t^{(N)}\|^2$ is bounded as $N\to\infty$, it follows that A_2 can be made arbitrarily small for large values of C_1.

LEMMA 4.6. (Swensen's condition(ii)) *Still under $H_f^{(N)}(\boldsymbol{\theta})$, with $\tau^{(N)}$ given in* (3.12), $\lim_{N\to\infty}[\sum_{t=1}^{N} E[\zeta_t^{(N)}]^2 - (\tau^{(N)}/2)^2] = 0$.

PROOF. From (3.11), $(\zeta_t^{(N)})^2$ decomposes into three terms:

$$[\zeta_t^{(N)}]^2 = \frac{1}{4N}\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right]'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)})\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right] \tag{4.12}$$
$$+ \frac{1}{2N}\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right]'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)})$$
$$\cdot\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right]$$

$$+\frac{1}{4N}\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right]'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)})$$
$$\cdot\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right].$$

Term number one has expectation

$$\frac{1}{4N}\operatorname{tr}\left\{\boldsymbol{\mathcal{I}}(f)E\left\{\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right]\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right]'\right\}\right\} \tag{4.13}$$
$$=\frac{1}{4N}\operatorname{tr}\left\{\boldsymbol{\mathcal{I}}(f)\sum_{i,j=0}^{t-1}\boldsymbol{H}_iE[\bar{\boldsymbol{Z}}_t^{(N)}\bar{\boldsymbol{Z}}_{t-|i-j|}'^{(N)}]\boldsymbol{H}_j'\right\}+o(1).$$

The expectation of the second term is zero, while, for the third one, we obtain

$$\frac{1}{4N}\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right]'\boldsymbol{\mathcal{I}}(f)\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right] \tag{4.14}$$
$$=\frac{1}{4N}\operatorname{tr}\left\{\boldsymbol{\mathcal{I}}(f)\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right]\right.$$
$$\left.\cdot\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right]'\right\}.$$

Summing (4.13) and (4.14) from 1 to N, letting $N\to\infty$ and applying Lemma 4.4 yields the desired result.

LEMMA 4.7. (Swensen's condition (iii)) *Under* $H_f^{(N)}(\boldsymbol{\theta})$, *as* $N\to\infty$, $\max_{1\le t\le N}|\zeta_t^{(N)}|=o_P(1)$.

PROOF. From (4.4) and the definition of $\zeta_t^{(N)}$,

$$\max_{1\le t\le N}|\zeta_t^{(N)}| \tag{4.15}$$
$$\le\frac{m}{2}\max_{1\le t\le N}\max_{1\le k\le m}N^{-1/2}\left|\left(\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right)_k\varphi_k(\boldsymbol{e}_t^{(0)})\right|$$
$$+\frac{m}{2}\max_{1\le t\le N}\max_{1\le k\le m}N^{-1/2}\left|\left(\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right)_k\varphi_k(\boldsymbol{e}_t^{(0)})\right|$$
$$=A_1^{(N)}+A_2^{(N)},\qquad\text{say.}$$

For any $\delta > 0$, any $k = 1, \ldots, m$,

$$
\begin{aligned}
(4.16)\ & P\left\{ N^{-1/2} \max_{1\le t\le N} \left| \sum_{i=0}^{t-1} \left(\boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right)_k \varphi_k(\boldsymbol{e}_t^{(0)}) \right| > \delta \right\} \\
&= P\Bigg\{ N^{-1} \sum_{t=1}^{N} \left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)} \right)_k^2 \varphi_k^2(\boldsymbol{e}_t^{(0)}) \\
&\qquad \cdot I\left[N^{-1/2} \left| \left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)} \right)_k \varphi_k(\boldsymbol{e}_t^{(0)}) \right| > \delta \right] > \delta^2 \Bigg\} \\
&\le N^{-1}\delta^{-2} \sum_{t=1}^{N} E\Bigg\{ \left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)} \right)_k^2 \varphi_k^2(\boldsymbol{e}_t^{(0)}) \\
&\qquad \cdot I\left[N^{-1/2} \left| \left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)} \right)_k \varphi_k(\boldsymbol{e}_t^{(0)}) \right| > \delta \right] > \delta^2 \Bigg\}.
\end{aligned}
$$

Since $\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}$ converges, in L^2, as $t \to \infty$ and uniformly in N, to a stationary process, the sequences $(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)})_k^2 \varphi_k^2(\boldsymbol{e}_t^{(0)})$ are uniformly integrable, i.e., for any $\varepsilon > 0$ there exists $N(\varepsilon)$ such that

$$
\begin{aligned}
(4.17)\qquad & E\Bigg\{ \left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)} \right)_k^2 \varphi_k^2(\boldsymbol{e}_t^{(0)}) \\
&\qquad \cdot I\left[\left| \left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)} \right)_k \varphi_k(\boldsymbol{e}_t^{(0)}) \right| > N^{1/2}\delta \right] \Bigg\} < \varepsilon,
\end{aligned}
$$

for all $N > N(\varepsilon)$, which implies that (4.16) and $A_1^{(N)}$ are $o_P(1)$ as $N \to \infty$. As for $A_2^{(N)}$,

$$
\begin{aligned}
& P\left[N^{-1/2} \max_{1\le t\le N} \left| \left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{A}(L) \boldsymbol{b}'^{(N)} \boldsymbol{K}'^{(N)} \boldsymbol{x}_{t-i}^{(N)} \right)_k \varphi_k(\boldsymbol{e}_t^{(0)}) \right| > \delta \right] \\
&\quad \le N^{-1}\delta^{-2} \sum_{t=1}^{N} \left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{A}(L) \boldsymbol{b}'^{(N)} \boldsymbol{K}'^{(N)} \boldsymbol{x}_{t-i}^{(N)} \right)_k^2 \\
&\quad E\left\{ \varphi_k^2(\boldsymbol{e}_t^{(0)}) I\left[\left| \left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{A}(L) \boldsymbol{b}'^{(N)} \boldsymbol{K}'^{(N)} \boldsymbol{x}_{t-i}^{(N)} \right)_k \varphi_k(\boldsymbol{e}_t^{(0)}) \right| > N^{1/2}\delta \right] \right\};
\end{aligned}
$$

due to the Noether condition, hence the convergence of $N^{-1}\sum_{t=1}^{N} \cdot \sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{A}(L) \boldsymbol{b}'^{(N)} \boldsymbol{K}'^{(N)} \boldsymbol{x}_{t-i}^{(N)}$, the stationarity of $\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})$ and the existence of a finite information matrix $\boldsymbol{\mathcal{I}}(f)$, $A_2^{(N)}$ also is $o_P(1)$, which completes the proof of this lemma.

LEMMA 4.8. (Swensen's condition (iv)) *Still under* $H_f^{(N)}(\boldsymbol{\theta})$, *as* $N \to \infty$,

$$\sum_{t=1}^{N} (\zeta_t^{(N)})^2 - (\tau^{(N)}/2)^2 = o_P(1). \tag{4.18}$$

PROOF. From (3.11) and (4.4),

$$
\begin{aligned}
\sum_{t=1}^{N} (\zeta_t^{(N)})^2 \quad & \\
= \frac{1}{4} N^{-1} \sum_{t=1}^{N} & \left[\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right]' \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)}) \left[\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right] \\
& + \frac{1}{4} N^{-1} \sum_{t=1}^{N} \left[\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{A}(L) \boldsymbol{b}'^{(N)} \boldsymbol{K}'^{(N)} \boldsymbol{x}_{t-i}^{(N)}\right]' \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)}) \\
& \qquad \cdot \left[\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{A}(L) \boldsymbol{b}'^{(N)} \boldsymbol{K}'^{(N)} \boldsymbol{x}_{t-i}^{(N)}\right] \\
& + \frac{1}{2} N^{-1/2} \sum_{t=1}^{N} \left[\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{A}(L) \boldsymbol{b}'^{(N)} \boldsymbol{K}'^{(N)} \boldsymbol{x}_{t-i}^{(N)}\right]' \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)}) \\
& \qquad \cdot \left[\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right].
\end{aligned}
\tag{4.19}
$$

As already mentioned, the sums $\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}$ converge in L^2, uniformly in N, as $t \to \infty$, to $\sum_{i=0}^{\infty} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}$, say, so that

$$
\begin{aligned}
N^{-1} \sum_{t=1}^{N} & \left[\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right]' \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)}) \left[\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right] \\
& = N^{-1} \sum_{t=1}^{N} \left[\boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)}) \sum_{i=0}^{\infty} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right]^2 + o_P(1).
\end{aligned}
$$

Now, applying, e.g., Hannan ((1970), Theorem 2, p. 203), and due to the fact that the sequences $\|\boldsymbol{\gamma}_i^{(N)}\|$ and $\|\boldsymbol{\delta}_i^{(N)}\|$ are eventually bounded for all i as $N \to \infty$,

$$N^{-1} \sum_{t=1}^{N} \left[\boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)}) \sum_{i=0}^{\infty} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right]^2 - E\left\{\left[\boldsymbol{\varphi}'(\boldsymbol{e}_1^{(0)}) \sum_{i=0}^{\infty} \boldsymbol{H}_i \boldsymbol{Z}_{1-i}^{(N)}\right]^2\right\}$$

converges a.s. to zero. Since

$$E\left\{\left[\boldsymbol{\varphi}'(\boldsymbol{e}_1^{(0)}) \sum_{i=0}^{\infty} \boldsymbol{H}_i \boldsymbol{Z}_{1-i}^{(N)}\right]^2\right\} = \operatorname{tr}\left\{\boldsymbol{\mathcal{I}}(f) \sum_{i,j=0}^{\infty} \boldsymbol{H}_i E[\boldsymbol{Z}_1^{(N)} \boldsymbol{Z}_{1-|i-j|}'^{(N)}] \boldsymbol{H}_j'\right\},$$

we obtain

$$(4.20)\quad N^{-1}\sum_{t=1}^{N}\left[\boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)})\sum_{i=0}^{\infty}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right]^2-\operatorname{tr}\left\{\boldsymbol{\mathcal{I}}(f)\sum_{i=1}^{\infty}\boldsymbol{d}_i^{(N)}\boldsymbol{\Sigma}\boldsymbol{d}_i'^{(N)}\right\}=o_P(1).$$

Next consider the second term in the right-hand side of (4.19); let

$$(4.21)\quad \begin{aligned}\boldsymbol{\alpha}_t^{(N)}&=N^{-1/2}\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)},\\ B^{(N)}&=\left\{\operatorname{tr}\sum_{t=1}^{N}\boldsymbol{\mathcal{I}}(f)\boldsymbol{\alpha}_t^{(N)}\boldsymbol{\alpha}_t'^{(N)}\right\}^{1/2}.\end{aligned}$$

Obviously, $\sum_{t=1}^{N}\boldsymbol{\alpha}_t'^{(N)}\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})/B^{(N)}$ has mean zero and variance one, and

$$\begin{aligned}&\sum_{t=1}^{N}E[(\boldsymbol{\alpha}_t'^{(N)}\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})/B^{(N)})I[|\boldsymbol{\alpha}_t'^{(N)}\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})|>B^{(N)}\varepsilon]]\\ &\qquad\le(B^{(N)})^{-2}\sum_{1\le k,\ell\le m}\sum_{t=1}^{N}|(\boldsymbol{\alpha}_t^{(N)})_k(\boldsymbol{\alpha}_t^{(N)})_\ell||\boldsymbol{\mathcal{I}}(f)_{k,\ell}|=o(1).\end{aligned}$$

It follows from the Noether condition (2.7) that $\max_{1\le t\le N}\|\boldsymbol{\alpha}_t^{(N)}\|$ is $o(1)$, as $N\to\infty$. The sum $\sum_{t=1}^{N}\boldsymbol{\alpha}_t'^{(N)}\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})/B^{(N)}$ thus satisfies the Lindeberg condition (see e.g. Shiryayev ((1984), p. 326)), and accordingly is asymptotically standard normal. It then follows from relative stability (Gnedenko and Kolmogorov (1968), p. 143) that

$$\begin{aligned}(4.22)\quad &N^{-1}\sum_{t=1}^{N}\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right]'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)})\\ &\qquad\cdot\left[\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{A}(L)\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-i}^{(N)}\right]\\ &\qquad-\operatorname{tr}\boldsymbol{\mathcal{I}}(f)\sum_{i,j=0}^{\infty}\sum_{k,\ell=1}^{p_1}\boldsymbol{H}_i\boldsymbol{A}_k\boldsymbol{b}'^{(N)}\boldsymbol{K}'\boldsymbol{R}_{|i-j+k-\ell|}\boldsymbol{K}\boldsymbol{b}'^{(N)}\boldsymbol{A}_\ell'\boldsymbol{H}_j'=o_P(1)\end{aligned}$$

as $N\to\infty$. Finally, in order for the third term in the right-hand side of (4.19) to be $o_P(1)$, it is sufficient that $\lim_{N\to\infty}N^{-1}E\{\sum_{t=1}^{N}\boldsymbol{\alpha}_t'^{(N)}\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)})\cdot\sum_{i=0}^{\infty}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\}^2=0$. This latter quantity decomposes into two parts:

$$\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\alpha}_t'^{(N)}E\left[\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)})\left(\sum_{i=1}^{\infty}\boldsymbol{d}_i^{(N)}\boldsymbol{\Sigma}\boldsymbol{d}_i'^{(N)}\right)\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\boldsymbol{\varphi}'(\boldsymbol{e}_t^{(0)})\right]\boldsymbol{\alpha}_t^{(N)}$$

which, due to (C3) and the fact that $\max_{1\leq t\leq N} \|\boldsymbol{\alpha}_t^{(N)}\| \to 0$ as $N \to \infty$, is $o(1)$, and

$$(4.23)\quad \frac{2}{N} \sum_{1\leq t_1<t_2\leq n} E\Bigg\{ \boldsymbol{\alpha}_{t_1}^{\prime(N)} \varphi(\boldsymbol{e}_{t_1}^{(0)})\varphi'(\boldsymbol{e}_{t_1}^{(0)}) \cdot \left(\sum_{i=0}^{\infty} \boldsymbol{H}_i \boldsymbol{Z}_{t_1-i}^{(N)}\right)\left(\sum_{i=0}^{\infty} \boldsymbol{H}_i \boldsymbol{Z}_{t_2-i}^{(N)}\right)' \varphi(\boldsymbol{e}_{t_2}^{(0)})\varphi'(\boldsymbol{e}_{t_2}^{(0)})\boldsymbol{\alpha}_{t_2}^{(N)}\Bigg\}.$$

Since $\boldsymbol{Z}_t^{(N)}$ is ARMA, $\sum_{i=0}^{\infty} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}$ is a linear process of the form $\sum_{i=0}^{\infty} \boldsymbol{\nu}_i^{(N)} \boldsymbol{\varepsilon}_{t-i}$, where $\sup_N \|\boldsymbol{\nu}_i^{(N)}\| \leq K_1\gamma^i$ for some $K_1 > 0$ and $0 < \gamma < 1$. Therefore, (4.23) is bounded by

$$\frac{2K}{N(1-\gamma^2)} \|\boldsymbol{\Sigma}\| \|\boldsymbol{\mathcal{I}}(f)\| \sup_t \|\boldsymbol{\alpha}_t^{(N)}\|^2 \sum_{i=1}^{N-1} (N-i)\gamma^i \leq K_2 \sup_t \|\boldsymbol{\alpha}_t^{(N)}\|^2 = o(1).$$

The desired result follows on piecing together (4.20) and (4.22).

LEMMA 4.9. (Swensen's condition (v)) *The sequence* $\sum_{t=1}^{N} (\zeta_t^{(N)})^2$ *is uniformly integrable, and condition* (v) *in Lemma* 2.3 *is satisfied.*

PROOF. Since

$$(\zeta_t^{(N)})^2 \leq N^{-1} \left[\left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right)' \varphi(\boldsymbol{e}_t^{(0)})\right]^2 + N^{-1} \left[\left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{A}(L) \boldsymbol{b}^{\prime(N)} \boldsymbol{K}^{\prime(N)} \boldsymbol{x}_{t-i}^{(N)}\right)' \varphi(\boldsymbol{e}_t^{(0)})\right]^2,$$

the uniform integrability of $\sum_{t=1}^{N} (\zeta_t^{(N)})^2$ follows from that of

$$(4.24)\quad N^{-1} \sum_{t=1}^{N} \left[\left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right)' \varphi(\boldsymbol{e}_t^{(0)})\right]^2$$

and (see (4.21))

$$(4.25)\quad \sum_{t=1}^{N} \left[\boldsymbol{\alpha}_t^{\prime(N)} \varphi(\boldsymbol{e}_t^{(0)})\right]^2.$$

The L^1 convergence to zero of

$$N^{-1} \sum_{t=1}^{N} \left[\left(\sum_{i=0}^{t-1} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right)' \varphi(\boldsymbol{e}_t^{(0)})\right]^2 - N^{-1} \sum_{t=1}^{N} \left[\left(\sum_{i=0}^{\infty} \boldsymbol{H}_i \boldsymbol{Z}_{t-i}^{(N)}\right)' \varphi(\boldsymbol{e}_t^{(0)})\right]^2$$

and the ergodic theorem entail the convergence to zero of

$$N^{-1}\sum_{t=1}^{N}\left[\left(\sum_{i=0}^{t-1}\boldsymbol{H}_i\boldsymbol{Z}_{t-i}^{(N)}\right)'\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})\right]^2-\operatorname{tr}\left\{\boldsymbol{\mathcal{I}}(f)\sum_{i=1}^{\infty}\boldsymbol{d}_i^{(N)}\boldsymbol{\Sigma}\boldsymbol{d}_i'^{(N)}\right\},$$

and hence, since $\operatorname{tr}\{\boldsymbol{\mathcal{I}}(f)\sum_{i=1}^{\infty}\boldsymbol{d}_i^{(N)}\boldsymbol{\Sigma}\boldsymbol{d}_i'^{(N)}\}$ is a bounded sequence, the uniform integrability of (4.24). As for (4.25), it has been shown in the proof of Lemma 4.8 that

$$\sum_{t=1}^{N}[\boldsymbol{\alpha}_t'^{(N)}\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})]^2-\operatorname{tr}\left\{\boldsymbol{\mathcal{I}}(f)\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{h}_j\boldsymbol{b}'^{(N)}\boldsymbol{K}'\boldsymbol{R}_{|i-j|}\boldsymbol{K}\boldsymbol{b}^{(N)}\boldsymbol{h}_j'\right\}=o_P(1);$$

hence, still as $N\to\infty$,

$$E\left\{\sum_{t=1}^{N}[\boldsymbol{\alpha}_t'^{(N)}\boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})]^2\right\}-\operatorname{tr}\left\{\boldsymbol{\mathcal{I}}(f)\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\boldsymbol{h}_j\boldsymbol{b}'^{(N)}\boldsymbol{K}'\boldsymbol{R}_{|i-j|}\boldsymbol{K}\boldsymbol{b}^{(N)}\boldsymbol{h}_j'\right\}=o(1).$$

The uniform integrability of (4.25) then follows from (an obvious modification of) Shiryayev ((1984), Theorem 3, p. 255). Finally condition (v) of Lemma 2.3 results from the same argument as in Swensen ((1985), p. 67)'s proof of (1.6).

LEMMA 4.10. (Swensen's conditions (vi) and (vii)) *For any N and t, under $H_f^{(N)}(\boldsymbol{\theta})$, $E[\zeta_t^{(N)}\mid \boldsymbol{Y}_1^{(N)},\dots,\boldsymbol{Y}_{t-1}^{(N)}]=0$, and, as $N\to\infty$, $\sum_{j=1}^{N}E[(\xi_t^{(N)})^2+2\xi_t^{(N)}\mid \boldsymbol{Y}_1^{(N)},\dots,\boldsymbol{Y}_{t-1}^{(N)}]=O_P(1)$.*

PROOF. The lemma is a straightforward consequence of the fact that $E(\boldsymbol{\varphi}(\boldsymbol{\varepsilon}_t))=\boldsymbol{0}$, and the mutual absolute continuity of all measures considered here.

The (LAN 1) part of Proposition 3.1 follows from Lemmas 4.5 through 4.10 and Swensen's Lemma 2.3.

4.4 *Proof of Proposition* 3.1 (*LAN* 2 *and LAN* 3)

The proof of the (LAN 2) and (LAN 3) parts of Proposition 3.1 consists in rewriting the quadratic approximation (3.13) under the more appropriate forms (3.14) and (3.20).

LEMMA 4.11. *Under $H_f^{(N)}(\boldsymbol{\theta})$, as $N\to\infty$, $2\sum_{t=1}^{N}\zeta_t^{(N)}=\boldsymbol{t}'^{(N)}\boldsymbol{\Delta}^{(N)}(\boldsymbol{\theta})+o_P(1)$.*

PROOF. From the definition (3.11),

$$2\sum_{t=1}^{N}\zeta_t^{(N)}=N^{-1/2}\sum_{t=1}^{N}\sum_{j=0}^{t-1}\left\{\boldsymbol{H}_j\boldsymbol{b}'^{(N)}\boldsymbol{K}'^{(N)}\boldsymbol{x}_{t-j}^{(N)}\right.$$

$$- \sum_{\ell=1}^{p_1} \boldsymbol{H}_j \boldsymbol{A}_\ell \boldsymbol{b}'^{(N)} \boldsymbol{K}'^{(N)} \boldsymbol{x}_{t-j-\ell}^{(N)} \Bigg\}' \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})$$

$$+ N^{-1/2} \sum_{t=1}^{N} \sum_{j=0}^{t-1} \sum_{\ell=1}^{p} \sum_{u=0}^{t-j-\ell-1} \sum_{k=0}^{q_1} \{\boldsymbol{H}_j \boldsymbol{\gamma}_\ell^{(N)} \boldsymbol{G}_u \boldsymbol{B}_k \boldsymbol{e}_{t-j-k-\ell-u}^{(0)}\}' \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)})$$

$$+ N^{-1/2} \sum_{t=1}^{N} \sum_{j=0}^{t-1} \sum_{\ell=1}^{q} \{\boldsymbol{H}_j \boldsymbol{\delta}_\ell^{(N)} \boldsymbol{e}_{t-j-\ell}^{(0)}\}' \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) + o_P(1)$$

$$= N^{-1/2} \sum_{i=0}^{N-1} \operatorname{tr} \left\{ \sum_{t=i+1}^{N} \boldsymbol{\varphi}(\boldsymbol{e}_t^{(N)}) \boldsymbol{x}_{t-i}'^{(N)} \boldsymbol{K}^{(N)} \boldsymbol{b}^{(N)} \boldsymbol{h}_i' \right\}$$

$$+ N^{-1/2} \sum_{\ell=1}^{p}$$

$$\cdot \sum_{i=\ell}^{N-1} \sum_{j=0}^{i-\ell} \sum_{k=0}^{\min(q_1, i-j-\ell)} \operatorname{tr} \left\{ \sum_{t=i+1}^{N} \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \boldsymbol{e}_{t-i}'^{(0)} \boldsymbol{B}_k' \boldsymbol{G}_{i-j-k-\ell}' \boldsymbol{\gamma}_\ell'^{(N)} \boldsymbol{H}_j' \right\}$$

$$+ N^{-1/2} \sum_{\ell=1}^{q} \sum_{i=\ell}^{N-1} \operatorname{tr} \left\{ \sum_{t=i+1}^{N} \boldsymbol{\varphi}(\boldsymbol{e}_t^{(0)}) \boldsymbol{e}_{t-i}'^{(0)} \boldsymbol{\delta}_\ell'^{(N)} \boldsymbol{H}_{i-\ell}' \right\} + o_P(1);$$

using the fact that $\operatorname{tr}(\boldsymbol{ABC}) = \operatorname{tr}(\boldsymbol{BCA})$, $\operatorname{tr}(\boldsymbol{AB}) = (\operatorname{vec} \boldsymbol{A}')'(\operatorname{vec} \boldsymbol{B})$ and $\operatorname{vec}(\boldsymbol{ABC}) = (\boldsymbol{C}' \otimes \boldsymbol{A})(\operatorname{vec} \boldsymbol{B})$, this latter quantity takes the form

$$\sum_{i=0}^{N-1} \operatorname{tr}\{\boldsymbol{b}^{(N)} \boldsymbol{h}_i' [(N-i)^{1/2} \boldsymbol{T}_{i;f}^{(N)}]\}$$

$$+ \sum_{\ell=1}^{p} \sum_{i=\ell}^{N-1} \sum_{j=0}^{i-\ell} \sum_{k=0}^{\min(q_1, i-j-\ell)} \operatorname{tr}\{\boldsymbol{\gamma}'^{(N)} \boldsymbol{H}_j' [(N-i)^{1/2} \boldsymbol{\Gamma}_{i;f}^{(N)}] \boldsymbol{B}_k' \boldsymbol{G}_{i-j-k-\ell}'\}$$

$$+ \sum_{\ell=1}^{q} \sum_{i=\ell}^{N-1} \operatorname{tr}\{\boldsymbol{\delta}_\ell'^{(N)} \boldsymbol{H}_{i-\ell}' [(N-i)^{1/2} \boldsymbol{\Gamma}_{i;f}^{(N)}]\} + o_P(1)$$

$$+ \sum_{\ell=1}^{p} (\operatorname{vec} \boldsymbol{\gamma}_\ell^{(N)})' \sum_{i=\ell}^{N-1} \sum_{j=0}^{i=\ell} \sum_{k=0}^{\min(q_1, i-j-\ell)} \operatorname{vec}[\boldsymbol{H}_j' (N-i)^{1/2} \boldsymbol{\Gamma}_{i;f}^{(N)} \boldsymbol{B}_k' \boldsymbol{G}_{i-j-k-\ell}']$$

$$+ \sum_{\ell=1}^{p} (\operatorname{vec} \boldsymbol{\delta}_\ell^{(N)})' \sum_{i=\ell}^{N-1} \operatorname{vec}[\boldsymbol{H}_{i-\ell}' (N-i)^{1/2} \boldsymbol{\Gamma}_{i;f}^{(N)}] + o_P(1)$$

$$+ \sum_{\ell=1}^{p} (\operatorname{vec} \boldsymbol{\gamma}_\ell^{(N)})'$$

$$\cdot \sum_{i=1}^{N-1} \sum_{j=0}^{i-\ell} \sum_{k=0}^{\min(q_1, i-j-\ell)} [(\boldsymbol{G}_{i-j-k-\ell} \boldsymbol{B}_k) \otimes \boldsymbol{H}_j'] (N-i)^{1/2} (\operatorname{vec} \boldsymbol{\Gamma}_{i;f}^{(N)})$$

$$+\sum_{\ell=1}^{q}\sum_{i=\ell}^{N-1}(\boldsymbol{I}\otimes\boldsymbol{H}'_{i-\ell})(N-i)^{1/2}(\operatorname{vec}\boldsymbol{\Gamma}_{i;f}^{(N)})+o_P(1).$$

LEMMA 4.12. *For all integers ℓ', ℓ'', the $m(\ell'+\ell'')$-tuple*

$$\begin{aligned}\boldsymbol{V}_{\ell',\ell''}^{(N)}=(N^{1/2}\operatorname{vec}\boldsymbol{T}_{0;f}^{(N)}\quad&(N-1)^{1/2}\operatorname{vec}\boldsymbol{T}_{1;f}^{(N)}\quad\cdots\\ \cdots\quad(N-\ell')^{1/2}\operatorname{vec}\boldsymbol{T}_{\ell'-1;f}^{(N)}\quad&(N-1)^{1/2}\operatorname{vec}\boldsymbol{\Gamma}_{1;f}^{(N)}\quad\cdots\\ &\cdots\quad(N-\ell'')^{1/2}\operatorname{vec}\boldsymbol{\Gamma}_{\ell'';f}^{(N)})\end{aligned}$$

is asymptotically multinormal, under $H_f^{(N)}(\boldsymbol{\theta})$, as $N\to\infty$, with mean $\boldsymbol{0}$ and covariance matrix

$$\begin{pmatrix}\begin{pmatrix}\boldsymbol{I}\otimes\mathcal{I}(f) & (\boldsymbol{K}'\boldsymbol{R}_1\boldsymbol{K})\otimes\mathcal{I}(f) & \cdots & (\boldsymbol{K}'\boldsymbol{R}_{\ell'-1}\boldsymbol{K})\otimes\mathcal{I}(f)\\ (\boldsymbol{K}'\boldsymbol{R}_1\boldsymbol{K})\otimes\mathcal{I}(f) & \boldsymbol{I}\otimes\mathcal{I}(f) & & \vdots\\ \vdots & & \ddots & \\ (\boldsymbol{K}'\boldsymbol{R}_{\ell'-1}\boldsymbol{K})\otimes\mathcal{I}(f) & \cdots & & \boldsymbol{I}\otimes\mathcal{I}(f)\end{pmatrix} & \boldsymbol{0}_{(m\ell')\times(m\ell'')}\\ \boldsymbol{0}_{(m\ell'')\times(m\ell')} & \begin{pmatrix}\boldsymbol{\Sigma}\otimes\mathcal{I}(f)\cdots\boldsymbol{0}\\ \ddots\\ \boldsymbol{0}\cdots\boldsymbol{\Sigma}\otimes\mathcal{I}(f)\end{pmatrix}\end{pmatrix}.$$

PROOF. The proof follows from a standard application, to arbitrary linear combinations, of the classical Hoeffding-Robbins central-limit result for ℓ''-dependent sequences: see Hallin and Puri ((1994), Section 3.1) for a univariate example. The details are left to the reader.

LEMMA 4.13. *The central sequence $\boldsymbol{\Delta}^{(N)}(\boldsymbol{\theta})$ under $H_f^{(N)}(\boldsymbol{\theta})$ is asymptotically multinormal, as $N\to\infty$, with mean $\boldsymbol{0}$ and covariance matrix $\boldsymbol{\Gamma}^{\boldsymbol{\Delta}}(\boldsymbol{\theta})$ given in* (3.16).

COROLLARY 4.1. $\boldsymbol{t}'^{(N)}\boldsymbol{\Gamma}^{\boldsymbol{\Delta}}(\boldsymbol{\theta})\boldsymbol{t}^{(N)}=\frac{1}{2}(\tau^{(N)})^2+o(1)$, *as $N\to\infty$.*

PROOF. Up to a remainder term the components of which are $o_P(1)$ as $\min(\ell',\ell'')\to\infty$, uniformly in (N), the central sequences $\boldsymbol{\Delta}^{(N)}(\boldsymbol{\theta})$ are linear combinations of the vectors $\boldsymbol{V}_{\ell',\ell''}^{(N)}$. For fixed ℓ',ℓ'' Lemma 4.12 implies asymptotic normality of such linear combinations, with mean zero and covariance matrix $\boldsymbol{\Gamma}_{\ell',\ell''}^{\boldsymbol{\Delta}}(\boldsymbol{\theta})$ converging to $\boldsymbol{\Gamma}^{\boldsymbol{\Delta}}(\boldsymbol{\theta})$ as $\min(\ell',\ell'')\to\infty$. The result then follows from a classical result on triangular arrays (Brockwell and Davis (1987), Proposition 6.3.9). The left upper corner of $\boldsymbol{\Gamma}_{\ell',\ell''}^{\boldsymbol{\Delta}}(\boldsymbol{\theta})$, for example, is $\boldsymbol{\Gamma}_{I;\ell',\ell''}^{\boldsymbol{\Delta}}(\boldsymbol{\theta})=\sum_{i,j=0}^{\min(\ell',\ell'')}(\boldsymbol{I}\otimes\boldsymbol{h}_i)'[(\boldsymbol{K}'\boldsymbol{R}_{|i-j|}\boldsymbol{K})\otimes\mathcal{I}(f)](\boldsymbol{I}\otimes\boldsymbol{h}_j)$ which, applying twice the identity $(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D})=(\boldsymbol{AC})\otimes(\boldsymbol{BD})$ reduces to $\boldsymbol{\Gamma}_{I;\ell',\ell''}^{\boldsymbol{\Delta}}(\boldsymbol{\theta})=\sum_{i,j=0}^{\min(\ell',\ell'')}(\boldsymbol{K}'\boldsymbol{R}_{|i-j|}\boldsymbol{K})\otimes(\boldsymbol{h}_i'\mathcal{I}(f)\boldsymbol{h}_j)$ and converges to $\boldsymbol{\Gamma}^{\boldsymbol{\Delta}}(\boldsymbol{\theta})$ given in (3.16). The corollary then follows from the (LAN 1) part of Proposition 3.1, which completes the proof of the (LAN 2) part of Proposition 3.1.

As for the (LAN 3) part, from the proof of Lemma 4.11, we have, permuting $\sum_i$ and $\sum_\ell$,

$$
\begin{aligned}
2\sum_{t=1}^{N}\zeta_t^{(N)} = &\sum_{i=0}^{N-1}(N-i)^{1/2}\,\mathrm{tr}[\boldsymbol{b}^{(N)}\boldsymbol{h}_i'\boldsymbol{T}_{i;f}^{(N)}] \\
&+\sum_{i=1}^{N-1}(N-i)^{1/2} \\
&\quad\cdot\sum_{\ell=1}^{\min(i,p)}\sum_{j=0}^{i-\ell}\sum_{k=1}^{\min(q_1,i-j-\ell)}\mathrm{tr}[\boldsymbol{\Gamma}_{i;f}^{(N)}\boldsymbol{B}_k'\boldsymbol{G}_{i-j-k-\ell}'\boldsymbol{\gamma}_\ell'^{(N)}\boldsymbol{H}_j'] \\
&+\sum_{i=1}^{N-1}(N-i)^{1/2}\sum_{\ell=1}^{\min(i,q)}\mathrm{tr}[\boldsymbol{\Gamma}_{i;f}^{(N)}\boldsymbol{\delta}_\ell^{(N)}\boldsymbol{H}_{i-\ell}']+o_P(1);
\end{aligned}
$$

(3.20) then readily follows.

References

Brockwell, P. J. and Davis, R. A. (1987). *Time Series: Theory and Methods*, Springer, New York.

Dufour, J.-M. and King, M. (1991). Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR (1) errors, *J. Econometrics*, **47**, 115–143.

Fabian, V. and Hannan, J. (1982). On estimation and adaptive estimation for locally asymptotically normal families, *Z. Wahrsch. Verw. Gebiete*, **59**, 459–478.

Garel, B. (1989). The asymptotic distribution of the likelihood ratio for MA processes with a regression trend, *Statist. Decisions*, **7**, 167–184.

Gnedenko, B. V. and Kolmogorov, A. N. (1968). *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading, Massachusetts.

Hallin, M. (1984). Spectral factorization of nonstationary moving average processes, *Ann. Statist.*, **12**, 172–192.

Hallin, M. (1986). Nonstationary q-dependent processes and time-varying moving-average models: invertibility properties and the forecasting problem, *Adv. in Appl. Probab.*, **18**, 170–210.

Hallin, M. and Puri, M. L. (1988). Optimal rank-based procedures for time-series analysis: testing an *ARMA* model against other *ARMA* models, *Ann. Statist.*, **16**, 402–432.

Hallin, M. and Puri, M. L. (1991). Time series analysis via rank-order theory: signed rank tests for *ARMA* models, *J. Multivariate Anal.*, **39**, 1–29.

Hallin, M. and Puri, M. L. (1994). Aligned rank tests for linear models with autocorrelated error terms, *J. Multivariate Anal.*, **50**, 175–237.

Hallin, M., Lefèvre, Cl. and Puri, M. L. (1988). On time-reversibility and the uniqueness of moving average representations for non-Gaussian stationary time-series, *Biometrika*, **75**, 170–171.

Hannan, E. J. (1970). *Multiple Time Series*, Wiley, New York.

Judge, G. G., Griffiths, W. E., Hill, R. C., Lütkepohl, H. and Lee, T. C. (1985). *The Theory and Practice of Econometrics*, 2nd ed., Wiley, New York.

Kreiss, J.-P. (1987). On adaptive estimation in stationary *ARMA* processes, *Ann. Statist.*, **15**, 112–133.

Kreiss, J.-P. (1990*a*). Testing linear hypotheses in autoregressions, *Ann. Statist.*, **18**, 1470–1482.

Kreiss, J.-P. (1990*b*). Local asymptotic normality for autoregression with infinite order, *J. Statist. Plann. Inference*, **26**, 185–219.

LeCam, L. (1960). Locally asymptotically normal families of distributions, *Univ. Calif. Pub. Statist.*, **3**, 37–98.

LeCam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*, Springer, New York.

LeCam, L. and Yang, Gr. L. (1990). *Asymptotics in Statistics*, Springer, New York.

Malliavin, P. (1982). *Intégration et Probabilités, Analyse de Fourier et Analyse Spectrale*, Masson, Paris.

McLeish, D. L. (1974). Dependent central limit theorems and invariance principles, *Ann. Probab.*, **2**, 620–628.

Shiriyayev, A. N. (1984). *Probability*, Springer, New York.

Strasser, H. (1985). *Mathematical Theory of Statistics*, de Gruyter, Berlin.

Swensen, A. R. (1985). The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend, *J. Multivariate Anal.*, **16**, 54–70.

Taniguchi, M. (1983). On the second order efficiency of estimates of Gaussian *ARMA* processes, *Ann. Statist.*, **11**, 157–169.

Taniguchi, M. (1991). Higher order asymptotic theory for time series analysis, *Lecture Notes in Statist.*, Springer, New York.

Weiss, G. (1975). Time series reversibility of linear stochastic processes, *J. Appl. Probab.*, **12**, 831–836.