Econometric Theory, **15**, 1999, 583–621. Printed in the United States of America.

ON ASYMPTOTIC INFERENCE IN COINTEGRATED TIME SERIES WITH FRACTIONALLY INTEGRATED ERRORS

P. JEGANATHAN
University of Michigan

Vector valued autoregressive models with fractionally integrated errors are considered. The possibility of the coefficient matrix of the model having eigenvalues with absolute values equal or close to unity is included. Quadratic approximation to the log-likelihood ratios in the vicinity of auxiliary estimators of the parameters is obtained and used to make a rough identification of the approximate unit eigenvalues, including complex ones, together with their multiplicities. Using the identification thus obtained, the stationary linear combinations (cointegrating relationships) and the trends that induce the nonstationarity are identified, and Wald-type inference procedures for the parameters associated with them are constructed. As in the situation in which the errors are independent and identically distributed (i.i.d.), the limiting behaviors are nonstandard in the sense that they are neither normal nor mixed normal. In addition, the ordinary least squares procedure, which works reasonably well in the i.i.d. errors case, becomes severely handicapped to adapt itself approximately to the underlying model structure, and hence its behavior is significantly inferior in many ways to the procedures obtained here.

1. INTRODUCTION

Consider the two-dimensional observations $(X_i, Y_i), i = 1, \ldots, n$, in which

$$\left.\begin{aligned} Y_i &= \beta X_i + u_i \\ X_i &= \gamma X_{i-1} + v_i \end{aligned}\right\}, \tag{1}$$

where the variables $U_i = (u_i, v_i)'$ form a suitable (approximately) stationary linear process with possibly involving unknown parameters. When the parameter γ is in a close but unknown vicinity (in a suitable sense that will become clear subsequently) of unity, both X_i and Y_i become integrated processes whereas their

This is a corrected version of Technical Report 277, 1996, Department of Statistics, University of Michigan. I express my thanks to Professor Peter Schmidt, Economics Department, Michigan State University, for bringing to my attention some recent works that led to Remark 2 of this paper. I also thank the Co-Editor, Professor K. Tanaka, and a referee for their careful reading of the paper and pointing out important corrections. Address correspondence to: P. Jeganathan, Department of Statistics, University of Michigan, Ann Arbor, MI 48105, USA; e-mail: jegan@umich.edu.

© 1999 Cambridge University Press 0266-4666/99 $9.50

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

linear combination is stationary. Here an integrated process roughly means a process that can be written in the form of a partial sum of an approximately stationary sequence. Models of this form are called cointegrated models. The primary aim in this context is the asymptotic inference problems with respect to the hypothesis of cointegration, that is, with respect to the hypothesis of near unity of the parameter γ (with the remaining parameters treated as unknown nuisance parameters) and similarly with respect to the cointegration parameter β that expresses the cointegration relationship.

Because, in (1), $Y_i = \gamma\beta X_{i-1} + u_i + \beta v_i$, it will be convenient to modify the system in the form

$$\left.\begin{aligned} Y_i &= \beta X_{i-1} + u_i \\ X_i &= \gamma X_{i-1} + v_i \end{aligned}\right\}, \tag{2}$$

where the present u_i takes the role of $u_i + \beta v_i$. The inference procedures obtained under this system can be transformed to those under (1), essentially through the one–one transformation $(\beta, \gamma) \mapsto (\beta\gamma, \gamma)$.

A more general and realistic situation would be to consider, which is the main purpose of this paper, the multivariate observations $Z_i, i = 1, \ldots, n$, such that

$$Z_i = AZ_{i-1} + U_i, \tag{3}$$

where U_i forms an approximately stationary linear process with possibly involving unknown parameters. If the parameter matrix A has at least one eigenvalue close to unity in absolute value and has no eigenvalues greater than unity in absolute values, then the model (3) may be said to be cointegrated. Then one first analyzes for the presence of cointegrating relationships, that is, for the presence of approximate unit eigenvalues, and then one isolates the cointegrating relationships and the parameters associated with them by decomposing the model in the form of a generalization of the system (1) or (2).

A solution to these asymptotic inference problems has been provided in Jeganathan (1997) under the situation in which the linear process has the representation

$$U_i = \sum_{j=0}^{\infty} c_j(\alpha)\mathcal{E}_{i-j} \tag{4}$$

with $\mathcal{E}_j$ independent and identically distributed (i.i.d.), such that it is invertible in the sense that

$$\mathcal{E}_i = \sum_{j=0}^{\infty} d_j(\alpha)U_{i-j}. \tag{5}$$

Here $c_j(\alpha)$ and $d_j(\alpha)$ are, in the specific context of (2), 2×2 matrices depending on the unknown parameter vector α. In addition it was assumed there that the coefficients $c_j(\alpha)$ and $d_j(\alpha)$ are absolutely summable in such a way that the partial sum process $n^{-1/2}\sum_{i=1}^{[nt]} U_i$ converges in law in an appropriate sense to a

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

Brownian motion. Because the literature on cointegrated models is rather extensive, we only refer to Jeganathan (1997), Johansen (1995), Phillips (1991), Reinsel (1993), Tanaka (1996), Toda and Phillips (1993), and the survey papers by, for instance, Stock (1994) and Watson (1994) for the detailed references for the earlier approaches and results in this direction.

In this paper we consider the situation in which the process U_i will be fractionally integrated in the sense that

$$(1-L)^{\alpha}U_i = \mathcal{E}_i,$$

where L is the back-shift operator defined by $LU_i = U_{i-1}$. (Although we assume for the most part that $\mathcal{E}_i$ are i.i.d., the extension, which will be needed in this paper itself, of the results to the situation where $\mathcal{E}_i$ themselves form an invertible linear process with absolutely summable coefficients will be indicated later; see Remark 4 of Section 2.3. In addition, the situation where the components of U_i will have different fractional integration coefficients will also be indicated; see Remark 2 of section 2.3.) Here, $|\alpha| < \frac{1}{2}$, a restriction that will be assumed to hold throughout what follows without further mention. (When $\alpha = 0$, it is to be understood that $d_0 = 1 = c_0$ and $d_j = 0 = c_j$ for $j \geq 1$.) The U_i still has the representation (4) and (5) but with

$$c_j(\alpha) = (\alpha)(1+\alpha)\dots(j-1+\alpha)/j! \sim \frac{1}{(\alpha-1)!\,j^{1-\alpha}}$$

and

$$d_j(\alpha) = (-\alpha)(1-\alpha)\dots(j-1-\alpha)/j! \sim \frac{1}{(-\alpha-1)!\,j^{1+\alpha}}$$

(see, e.g., Hosking, 1981; Granger and Joyeux, 1980). Here, when $-\frac{1}{2} < \alpha < 0$, d_j are not absolutely summable although c_j are, whereas when $0 < \alpha < \frac{1}{2}$, d_j are absolutely summable but c_j are not. In addition, in either one of these cases, the partial sum process $n^{-H}\sum_{i=1}^{[nt]} U_i$, where $H = \frac{1}{2} + \alpha$, converges weakly to what is called a fractional Brownian motion instead of the Brownian motion (see, e.g., Davydov, 1970; Gorodetskii, 1977; Taqqu, 1975). For a detailed treatment of the fractional Brownian motion, see, for instance, Mandelbrot and van Ness (1968) and Samorodnitsky and Taqqu (1994). For the significance and importance of the fractionally integrated models, see, for instance, the survey and review papers Robinson (1994a) and Baillie (1996); the issue of the *Journal of Econometrics* where Baillie (1996) appears is devoted to the topic related to the problem of the present paper.

The differences indicated previously induce significant differences in the asymptotic behaviors of the inference procedures described in Jeganathan (1997), though the construction and the form of the procedures themselves will be similar. The indicated procedures were constructed based on the methods described in LeCam (1986) and LeCam and Yang (1990) which involve the availability of

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

suitable auxiliary estimators of the parameters that are then modified or improved so as to adapt in the best possible way to the approximating structure of the likelihood ratios (of the observables) available in the vicinity of the given auxiliary estimators. For the purpose of convenience in this paper, such procedures will be referred to as approximate likelihood (Appx.L) based procedures.

It may be recalled that in the situation considered in Jeganathan (1997), the degree of approximate adaptability of the ordinary least squares (OLS) procedure and the Appx.L based procedures to the underlying model structure is comparable in many ways, for instance, having the same rate of convergence, among other similarities. On the other hand, in the present context, the OLS procedure will become severely handicapped to adapt itself, even approximately, to the underlying model and hence will be significantly inferior to those of Appx.L procedures.

In particular, under $\gamma = 1$, a result given in Sowell (1990) will imply that the OLS estimators of β and γ based on (1) converge in law at the rate $n^{\min\{1,1+2\alpha\}}$. On the other hand, we shall show that the Appx.L based procedures have the rate $n^{\max\{1,1+\alpha\}}$, the improvement being significant. Similarly, in Lee (1996) it is shown, assuming that the first component of U_i is with the fractional parameter α_1 but that of the second component is i.i.d., that is, with $\alpha_2 = 0$, the OLS estimator of β has the rate $n^{-1-\alpha_1}$, but the present Appx.L estimator will have the rate $n^{\max\{1+\alpha_2,1+\alpha_2-\alpha_1\}}$ (which reduces to $n^{\max\{1,1-\alpha_1\}}$ for the case $\alpha_2 = 0$ of Lee, 1996). Without going into the details, we also mention that similar difficulties arise for the OLS procedures treated in Cheung and Lai (1993) and Dolado and Marmol (1996) in the context of models similar to the present one. Major differences also arise in the structure of the limiting distributions. For instance, that of the OLS estimator in Sowell (1990) is concentrated either on the positive section or on the negative section of the real line, according to whether α is positive or negative, whereas such distortions do not arise in the Appx.L based procedures, though its form is not exactly the same as that of the corresponding $\alpha = 0$ situation in Jeganathan (1997).

The situation with respect to the case $\gamma = -1$ becomes more significant. According to a result obtained (under the Gaussian assumption) by Chan and Terrin (1995), the OLS estimators of β and γ will converge, when $0 < \alpha < \frac{1}{2}$, at the same rate n as that for $\alpha = 0$. We shall show that Appx.L based procedures will also converge at the rate n but over the entire range $-\frac{1}{2} < \alpha < \frac{1}{2}$. In addition, the structures of the limiting distributions of the inference procedures will remain the same as those of the $\alpha = 0$ situation of Jeganathan (1997), whereas this is not the case for OLS estimators. The underlying reason, which is different from that of Jeganathan (1997), is that when $\gamma = -1$, the (mis)behavior of the spectral density of the error process U_i near the origin, and the possible nonabsolute summability of the coefficients d_i's, plays no role in the asymptotic behaviors. The same type of phenomenon arises in the multivariate situation (3) with complex roots.

The situation with respect to $|\gamma| < 1$ is still more significant in the sense that a result of Chan and Terrin (1995) implies that the OLS estimators will be not even consistent, whereas we shall show that the Appx.L based procedures will con-

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

verge at the usual rate $n^{1/2}$, despite the fact that d_i's are not absolutely summable when $-\frac{1}{2} < \alpha < 0$. This later behavior is the same as that of the usual situation in which the coefficients c_j and d_j decay in j exponentially. Similar results have been established in Robinson (1994b) for the generalized least squares (and more general Whittle-type) estimators of the slope parameter of the usual regression models with errors of the present type but only with $0 < \alpha < \frac{1}{2}$; though the possibility of the regressors being stochastic is allowed, the present situation is not covered because the regressors are assumed to be independent of the errors. (Note that OLS is also usually interpreted as a pseudo MLE in the usual sense. It is shown in Pötscher, 1991 and Tanaka, 1996, Sect. 8.5, that an inconsistency phenomenon similar to that mentioned previously can arise with respect to pseudo MLE in general in several other related contexts also.)

The plan of the rest of the paper is as follows. In Section 2, we consider the system (2) and obtain the quadratic approximation to the log-likelihood ratios. Given this quadratic approximation, the construction of inference procedures will remain essentially the same as that described in Jeganathan (1997). Note however that obtaining auxiliary estimators is not as straightforward as in Jeganathan (1997) because OLS estimators cannot serve the purpose in the present context, for the reasons indicated previously. (The auxiliary estimators that we shall use are known as the conditional sum of squares estimators.) In Section 2.2 we then describe the limiting distributions of the procedures of Section 2.1, both under the null and under the alternatives. Section 2.3 contains remarks on some extensions or modifications. In Section 3, we consider the more general and realistic situation (3) and obtain the following, analogous to those in Jeganathan (1997): (i) we obtain the quadratic approximation to the log-likelihood ratios in the vicinity of an auxiliary estimator and suggest a procedure of roughly identifying the approximate unit eigenvalues. (ii) Using the identification thus obtained in the preceding first stage, we obtain in the second stage an equivalent triangular representation of the model (3) in the form of a generalization of (1) or (2), thereby isolating the cointegrating relationships and the nonstationary trends. (iii) It is shown further that the estimators obtained under the unrestricted system (3) can be used for the quadratic approximation and hence for the inference purposes of the parameters of the reduced triangular system of the second stage (ii). The limiting distributions of these procedures are derived in Section 3.2, some of the arguments of which rely on the corresponding ideas in Jeganathan (1997), which in turn used those in Chan and Wei (1988), Jeganathan (1991), and Tsay and Tiao (1991). To avoid interrupting the flow of the main arguments, some of the details of the proofs are collected in Section 4.

Because in Section 3 we are allowing all possible unit roots, together with their possible multiplicities, complete treatment of all the basic ideas and techniques in such general context becomes rather complicated, unfortunately. For this reason the simpler system of Section 2 will be dealt with in some detail, and then the ideas involved will be transformed to the more general case, mostly by analogy whenever possible.

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

Most of the notations are standard. By a vector we always mean a column vector. For a matrix B, its transpose will be denoted by B', and the operator $\mathrm{vec}(B)$ stands for the column vector obtained by stacking the transposed rows of B. Unless otherwise specified, the integrals, ordinary or stochastic, are always over the unit interval $[0,1]$. Convergence in law of finite dimensional distributions of a sequence of random processes will be signified by "convergence (fdd) in law." The term L always stands for the back-shift operator; $\mathcal{I}$ is reserved for the "Fisher information"; but I_r stands for the unit matrix of order r. The largest integer less than or equal to s will be noted by $[s]$.

2. INFERENCE PROCEDURES FOR THE MODEL (2)

2.1. Construction of Inference Procedures

In this section we first obtain, when γ is in a suitable vicinity of the interval $[-1,1]$, the quadratic approximation to the log-likelihood ratios of the observables $(X_1,Y_1),\ldots,(X_n,Y_n)$ and then describe the construction of inference procedures based on the resulting approximating quadratic structure. It is important to note that the specification of the joint density of the observables will involve the distributions of certain unobservables also. We obtain the indicated approximation without assuming any conditions on the density of the unobservables because they in general are not verifiable. In this section we let $\theta = (\beta,\gamma,\alpha)$.

We shall assume that the common density of i.i.d. variable $\mathcal{E}_i$ has an absolutely continuous Lebesgue density $p(u,v)$. In addition we shall assume that the matrix

$$\mathcal{I} = \begin{bmatrix} \mathcal{I}_{11} & \mathcal{I}_{12} \\ \mathcal{I}_{21} & \mathcal{I}_{22} \end{bmatrix} \tag{6}$$

is finite and nonsingular, where $\mathcal{I}_{ij} = \int \psi_i(u,v)\psi_j(u,v)p(u,v)\,du\,dv$, $i,j = 1,2$, with

$$\psi_1(u,v) = -\frac{\partial \log p(u,v)}{\partial u} \quad \text{and} \quad \psi_2(u,v) = -\frac{\partial \log p(u,v)}{\partial v}.$$

We next introduce the "score functions" that will be involved in the quadratic approximations. For this purpose let

$$U_j(\beta,\gamma) = (Y_j - \beta X_{j-1}, X_j - \gamma X_{j-1})'.$$

The idea here, in constructing the approximation to the log-likelihood ratios, is to treat the variables

$$\mathcal{E}_{ni}(\theta) = \sum_{l=0}^{i-2} d_l(\alpha)U_{i-l}(\beta,\gamma), \tag{7}$$

which are obtained by truncating (5) and which involve only the observables, as i.i.d. with the common density $p(u,v)$ of $\mathcal{E}_1$. We then take the resulting joint den-

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

sity of these variables to be the joint density of the observables $(X_1,Y_1),\ldots,(X_n,Y_n)$. The derivative of the logarithm of such a density with respect to the parameter θ can be taken approximately to be of the 3-vector

$$W_n(\theta) = \sum_{i=k_n+1}^{n} (\widetilde{X}_{i-1}(\alpha)\psi_1(\mathcal{E}_{ni}(\theta)), \widetilde{X}_{i-1}(\alpha)\psi_2(\mathcal{E}_{ni}(\theta)), H_i'(\theta)\psi(\mathcal{E}_{ni}(\theta)))', \quad \textbf{(8)}$$

where $\psi = (\psi_1,\psi_2)'$,

$$\widetilde{X}_{i-1}(\alpha) = \sum_{j=0}^{i-2} d_j(\alpha)X_{i-j-1} \quad \text{and} \quad H_i(\theta) = \sum_{j=1}^{j-2} d_j^{(1)}(\alpha)U_{i-j}(\beta,\gamma)$$

with $d_j^{(1)}(\alpha)$ the derivative, with respect to the scalar α, of $d_j(\alpha)$. Here, k_n is any conveniently chosen sequence of positive integers such that $k_n \uparrow \infty$ and $(k_n/n) \to 0$. It will be seen that having the sum $\sum_{i=k_n+1}^{n}$ in (8), ignoring the first k_n terms, is convenient in many ways.

Next the 3×3 symmetric matrix that will be involved in the approximating quadratic is as follows:

$$M_n(\theta) = \begin{bmatrix} \left(\sum_{i=k_n+1}^{n} \widetilde{X}_{i-1}^2(\alpha)\right)\mathcal{I} & \sum_{i=k_n+1}^{n} \widetilde{X}_{i-1}(\alpha)\mathcal{I}H_i(\theta) \\ & \sum_{i=k_n+1}^{n} H_i'(\theta)\mathcal{I}H_i(\theta) \end{bmatrix}.$$

Next define

$$\delta_{*n} = \operatorname{diag}\{\delta_n,\delta_n,n^{-1/2}\}$$

with

$$\delta_n = \begin{cases} n^{-1/2} & \text{if } |\gamma| < 1 \\ n^{-(1+\alpha)} & \text{if } \gamma = 1 \quad \text{and} \quad 0 \le \alpha < \frac{1}{2} \\ n^{-1} & \text{otherwise.} \end{cases}$$

Let $P_{\theta,n}$ be the joint distribution of the observables $((X_i,Y_i), i = 1,\ldots,n)$, where recall that $\theta = (\beta,\gamma,\alpha)'$, and define the log-likelihood ratios:

$$\Lambda_n(\theta^*,\theta) = \log \frac{dP_{\theta^*,n}}{dP_{\theta,n}}.$$

We are now in a position to state the quadratic approximation result.

THEOREM 1. *Assume that the matrix* (6) *is finite and nonsingular. Then for any* θ *in* $\Omega = \{\theta = (\beta,\gamma,\alpha)' : \beta$ *real*, $-\frac{1}{2} < \alpha < \frac{1}{2}$, $|\gamma| \le 1\}$, *and for any bounded* 3*-vectors* h_n, *one has the following conclusions*:

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

(a)

$$\Lambda_n(\theta + \delta_{*n}h_n, \theta) = h_n'\delta_{*n}W_n(\theta) - \frac{1}{2}h_n'\delta_{*n}M_n(\theta)\delta_{*n}h_n + o_p(1),$$

$$\delta_{*n}W_n(\theta + \delta_{*n}h_n) = \delta_{*n}W_n(\theta) - \delta_{*n}M_n(\theta)\delta_{*n}h_n + o_p(1),$$

and

$$\delta_{*n}M_n(\theta + \delta_{*n}h_n)\delta_{*n} = \delta_{*n}M_n(\theta)\delta_{*n} + o_p(1)$$

in $P_{\theta,n}$.

(b) *The sequences* $\{P_{\theta,n}\}$ *and* $\{P_{\theta+\delta_{*n}h_n,n}\}$ *are contiguous.*

(c) *When* $\gamma = \pm 1$, *the matrix* $M_n(\theta)$ *involved in the preceding approximation can be taken to be* $\mathrm{diag}\{(\sum_{i=k_n+1}^{n}\widetilde{X}_{i-1}^2(\alpha))\mathcal{I}, \sum_{i=k_n+1}^{n}H_i'(\theta)\mathcal{I}H_i(\theta)\}$.

To use the approximating quadratic structure of the preceding theorem, one needs auxiliary estimators of the parameters. One possible approach, which is known as the conditional sum of squares method, to obtaining them can be described as follows. First, fix the parameter α and construct $\tilde{\beta}_n(\alpha)$ and $\tilde{\gamma}_n(\alpha)$ by

$$(\tilde{\beta}_n(\alpha), \tilde{\gamma}_n(\alpha))' = \left(\sum_{i=2}^{n}\widetilde{X}_{i-1}^2(\alpha)\right)^{-1}\left(\sum_{i=2}^{n}\widetilde{X}_{i-1}(\alpha)\widetilde{Y}_i(\alpha), \sum_{i=2}^{n}\widetilde{X}_{i-1}(\alpha)\widetilde{X}_i(\alpha)\right)',$$

where $\widetilde{X}_i(\alpha)$ is as defined earlier and $\widetilde{Y}_i(\alpha) = \sum_{j=0}^{i-1}Y_{i-j}d_j(\alpha)$. These are just the solutions of the equation

$$0 = \sum_{i=2}^{n}\widetilde{X}_{i-1}(\alpha)\left(\sum_{j=0}^{i-2}d_j(\alpha)U_{i-j}(\beta,\gamma)\right).$$

Note that this is essentially the Appx.L based estimators when α is known and when the error variables $\mathcal{E}_i$ are assumed to be i.i.d. Gaussian.

Now, a possible auxiliary estimator of α can be obtained by acting as if $\mathcal{E}_{ni}(\tilde{\theta}_n(\alpha))$ are i.i.d. Gaussian, where $\tilde{\theta}_n(\alpha) = (\tilde{\beta}_n(\alpha), \tilde{\gamma}_n(\alpha), \alpha)$ and $\mathcal{E}_{ni}(\theta)$ is as in (7). Under this scheme, if Appx.L approach is followed, the resulting $\tilde{\alpha}_n$ will satisfy

$$n^{-1/2}\sum_{i=k_n+1}^{n}H_i'(\tilde{\theta}_n(\tilde{\alpha}_n))\psi(\mathcal{E}_{ni}(\tilde{\theta}_n(\tilde{\alpha}_n))) = o_p(1)$$

with ψ corresponding to the Gaussian case. Unfortunately, the explicit construction of such $\tilde{\alpha}_n$ is not simple in view of the fact that α is involved in a complicated nonlinear way in the preceding scheme, though it can be taken to be an implicitly defined maximizing solution of $\sum[\mathcal{E}_{ni}(\tilde{\theta}_n(\alpha))]^2$, or the analogous Whittle-type estimators (e.g., similar to that in Dahlhaus, 1989; for a review, see further Robinson, 1994a). As far as we are aware, a step by step explicit construction of estimators of α is not readily available in the literature in nonlinear situations such as the present one. We shall return to this problem in a more general context elsewhere.

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

It can be shown that $\tilde{\theta}_n = \tilde{\theta}_n(\tilde{\alpha}_n)$ constructed as previously will have the property that

$$\delta_{*n}^{-1}(\tilde{\theta}_n - \theta) = O_p(1) \tag{9}$$

in $P_{\theta,n}$ for all θ in Ω. Now discretize $\tilde{\theta}_n$ in the sense that, letting $\tilde{\theta}_n = (\tilde{\beta}_n, \tilde{\gamma}_n, \tilde{\alpha}_n)$, compute $\tilde{\beta}_n$ and $\tilde{\gamma}_n$ up to an approximation of order $(\sum_{j=k_n}^{n} \tilde{X}_{i-1}^2(\tilde{\alpha}_n))^{-1/2}$, obtaining β_n^* and γ_n^*, and compute $\tilde{\alpha}_n$ up to an approximation of order $n^{-1/2}$, obtaining α_n^*. The resulting $\theta_n^* = (\beta_n^*, \gamma_n^*, \alpha_n^*)$ will still satisfy the property (9) of $\tilde{\theta}_n$. For the details of the advantages of using θ_n^* as the preliminary estimator instead of $\tilde{\theta}_n$ itself, see Jeganathan (1997, Sect. 2.2) and the references contained therein.

Next construct

$$\hat{\theta}_n = (\hat{\beta}_n, \hat{\gamma}_n, \hat{\alpha}_n) = \theta_n^* + M_n^{-1}(\theta_n^*) W_n(\theta_n^*), \tag{10}$$

which is called the centering estimator of θ in the sense that when θ is in the probable range of $\theta_n^*, \Lambda_n(\theta, \theta_n^*)$ will be approximately of the form

$$-\frac{1}{2} |M_n^{1/2}(\theta_n^*)(\theta - \hat{\theta}_n)|^2 + \frac{1}{2} |M_n^{-1/2}(\theta_n^*) W_n(\theta_n^*)|^2$$

in view of Theorem 1 and the discretized nature of θ_n^*. Thus a Wald–type procedure for testing $\gamma = \gamma_{0n}$, when β and α are treated as nuisance parameters, is given by

$$(M_n^{22}(\theta_n^*))^{-1}(\hat{\gamma}_n - \gamma_{0n})^2 \tag{11}$$

with $M_n^{ij}(\theta)$ the (i,j)th element of $M_n^{-1}(\theta)$. When $\gamma_{0n} = \pm 1 + \delta_n h_{2n}$ with $\{h_{2n}\}$ bounded, (11) tests the unit root hypothesis. An analogous test statistic for testing the hypothesis $\beta = \beta_{0n}$ is given by

$$(M_n^{11}(\theta_n^*))^{-1}(\hat{\beta}_n - \beta_{0n})^2. \tag{12}$$

Using the approximations in statement (a) of Theorem 1, the centering estimator (10) has the representation

$$\delta_{*n}^{-1}(\hat{\theta}_n - \theta) = (\delta_{*n} M_n(\theta) \delta_{*n})^{-1} \delta_{*n} W_n(\theta) + o_p(1). \tag{13}$$

If the hypothesis $\gamma = \pm 1 + \delta_n h_{2n}$ is accepted, it can be assumed that the range of γ is a suitable unknown contiguous neighborhood of unity, so that by statement (c) of Theorem 1, $M_n(\theta)$ in (10) can be replaced by $\mathrm{diag}\{(\sum_{i=k_n+1}^{n} \tilde{X}_{i-1}^{q2}(\alpha))\mathcal{I}, \sum_{i=k_n+1}^{n} H_i'(\theta)\mathcal{I}H_i(\theta)\}$, in which case $(\hat{\beta}_{*n}, \hat{\gamma}_n) = (\hat{\beta}_{*n}, \hat{\gamma}_{*n})$ with

$$(\hat{\beta}_{*n}, \hat{\gamma}_{*n}) = (\beta_n^*, \gamma_n^*)' + \left(\sum_{i=k_n}^{n} \tilde{X}_{i-1}^2(\alpha_n^*) \right)^{-1} \mathcal{I}^{-1} W_n^*(\theta_n^*),$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

where $W_n^*(\theta)$ is the subvector of $W_n(\theta)$ formed by its first two components. In particular (12) can be replaced by

$$\left(\sum_{i=k_n}^{n} \tilde{X}_{i-1}^2(\alpha_n^*)\right)\lambda_{11}^{-1}(\hat{\beta}_{*n}-\beta_{0n})^2, \tag{12a}$$

where λ_{11} is the (1,1) element of $\mathcal{I}^{-1}$.

Remark 1. Very often in the literature, when the hypothesis $\gamma=\gamma_{0n}=\pm 1+\delta_n h_{2n}$ is accepted, it is imposed in the model structure when obtaining the test procedures for β, though such procedures may lead to certain difficulties (see Jeganathan, 1997, Sect. 1). In the present context, one such procedure analogous to the preceding ones is given by

$$\left(\sum_{i=k_n+1}^{n} \tilde{X}_{i-1}^2(\alpha_n^*)\right)\mathcal{I}_{11}(\hat{\beta}_n^{(0)}-\beta_0)^2 \tag{14}$$

with

$$\hat{\beta}_n^{(0)}=\beta_n^*+\left(\sum_{i=k_n}^{n} \tilde{X}_{i-1}^2(\alpha_n^*)\right)^{-1}\mathcal{I}_{11}^{-1}W_n^{(1)}(\beta_n^*,\gamma_{0n},\alpha_n^*),$$

having the representation

$$\delta_n^{-1}(\hat{\beta}_n^{(0)}-\beta)=\left(\mathcal{I}_{11}\delta_n^2\sum_{i=k_n}^{n} \tilde{X}_{i-1}^2(\alpha)\right)^{-1}\delta_n W_n^{(1)}(\beta,\gamma_{0n},\alpha)+o_p(1),$$

where $W_n^{(1)}(\theta)$ is the first component of $W_n(\theta)$. It will be seen that the limiting distribution of (14) under the null hypothesis is central chi-squared, one of the main reasons for preferring (14) (see Phillips, 1991), whereas that of (11) or (12) will be different from chi-squared, as will be seen in the next section. However, as noted in Jeganathan (1997), the fact that a test procedure has a limiting chi-squared distribution alone cannot be taken as a distinct advantage over the ones that adapt themselves to the underlying model structure in an efficient way such as (11) and (12).

When the underlying common density $p(u,v)$ is non-Gaussian, another way of obtaining the procedures that will have the limiting chi-squared is to use the approximating quadratic structure of the likelihood ratios of the conditional distribution of $(Y_1,\ldots,Y_n)$ given $(X_1,\ldots,X_n)$, instead of the unconditional likelihoods as was done previously. (Details of this procedure can be obtained in analogy with Remark 4 of Jeganathan, 1997, Sect. 2.4.) Unfortunately, such a procedure cannot be employed in the Gaussian case because the parameters β and γ will be involved in the conditional likelihoods through a linear combination of them. In addition, the reservation expressed at the end of the previous paragraph applies here also.

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

2.2. Limiting Behaviors

In this section we describe the limiting distributional behaviors of the random quantities involved in the quadratic approximation of Theorem 1, which will then lead to the description of the analogous behaviors of the statistics (11)–(14), under the null and under contiguous alternatives.

For this purpose, let $B_H(t)$, $0 \le t \le 1$, $H = \alpha + \frac{1}{2}$ be the fractional Brownian motion, that is, a Gaussian process with

$$E[|B_H(t) - B_H(s)|^2] = C_H|t-s|^{2H}, \tag{15}$$

where

$$C_H = \frac{\mathrm{Var}(\varepsilon_{12})}{2\pi} \int_R \left| \frac{e^{ix}-1}{ix} \right|^2 |x|^{1-2H}\, dx.$$

Here $\mathcal{E}_{12}$ is the second component of $\mathcal{E}_1$. When $H = \frac{1}{2}$, $B_H(t)$ reduces to a Brownian motion. Note that $B_H(t)$ can be written as a functional of $\{B_{1/2}(s) : -\infty < s \le 1\}$ (see Mandelbrot and van Ness, 1968). In what follows, define

$$Z_n = \sum_{i=k_n+1}^{n} H_i' \psi(\mathcal{E}_i).$$

PROPOSITION 2.

(a) *The process*

$$\left(n^{-H} \sum_{i=1}^{[nt]} v_i, n^{-1/2} \sum_{i=1}^{[nt]} \psi_1(\mathcal{E}_i), n^{-1/2} \sum_{i=1}^{[nt]} \psi_2(\mathcal{E}_i), n^{-1/2} Z_n \right)$$

converges (fdd) *in law to* $(B_H(t), W_1(t), W_2(t), Z)$ *where* $W_1(t)$ *is independent of* $B_H(t)$,

$$E(B_H(t)W_2(s)) = \frac{1}{2}\,\mathrm{Cov}(\mathcal{E}_{12}, \psi_2(\mathcal{E}_1))[|t|^{H+1/2} + |s|^{H+1/2} - |t-s|^{H+1/2}],$$

$(W_1(t), W_2(t))$ *is a two-dimensional Brownian motion with covariance matrix* $\mathcal{I}$, *and* Z *is Gaussian, independent of* $(B_H(t), W_1(t), W_2(t))$, *with zero mean and variance as specified subsequently.*

(b) *The process*

$$\left(n^{-1/2} \sum_{i=1}^{[nt]} (-1)^{[nt]-i} v_{[nt]-i}, n^{-1/2} \sum_{i=1}^{[nt]} (-1)^{i-1} \psi_1(\mathcal{E}_i), n^{-1/2} \right.$$
$$\left. \times \sum_{i=1}^{[nt]} (-1)^{i-1} \psi_2(\mathcal{E}_i), n^{-1/2} Z_n \right)$$

converges (fdd) *in law to* $(2^{-\alpha} B_{1/2}(t), W_1(t), W_2(t), Z)$ *whose variance–covariance function is as in* (a) *with* $H = \frac{1}{2}$.

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

The variance of the preceding Z is given by

$$\tau^2 = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} d_j^{(1)} d_k^{(1)} E(U_j' \mathcal{I} U_k). \tag{16}$$

The following process will be involved in the next result.

$$F(t) = \begin{cases} \left(\sum_{j=0}^{\infty} d_j(\alpha)\right) B_H(t) & \text{if } \gamma = 1, 0 \le \alpha < \frac{1}{2} \\ \dfrac{1}{(-\alpha-1)!}\displaystyle\int_0^t \frac{1}{(t-u)^{1+\alpha}} B_H(u)\, du & \text{if } \gamma = 1, -\frac{1}{2} < \alpha < 0 \\ B_{1/2}(t) & \text{if } \gamma = -1. \end{cases}$$

PROPOSITION 3.

(a) *When* $\gamma = +1$, *the process*

$$\left(\sqrt{n}\delta_n \widetilde{X}_{[nt]}(\alpha),\, n^{-1/2}\sum_{i=1}^{[nt]} \psi_1(\mathcal{E}_i),\, n^{-1/2}\sum_{i=1}^{[nt]} \psi_2(\mathcal{E}_i), n^{-1/2} Z_n\right)$$

converges (fdd) *in law to* $(F(t), W_1(t), W_2(t), Z)$ *where* $F(t)$ *is as earlier for* $\gamma = +1$ *and the remaining components are as in* (a) *of Proposition* 2.

(b) *When* $\gamma = -1$, *the process*

$$\left((-1)^{[nt]} n^{-1/2} \widetilde{X}_{[nt]}(\alpha),\, n^{-1/2}\sum_{i=1}^{[nt]} (-1)^{i-1}\psi_1(\mathcal{E}_i),\, n^{-1/2} \times \sum_{i=1}^{[nt]} (-1)^{i-1}\psi_2(\mathcal{E}_i), n^{-1/2} Z_n\right)$$

converges (fdd) *in law to* $(B_{1/2}(t), W_1(t), W_2(t), Z)$ *with the limit as in* (b) *of Proposition* 2. (*Note that* $B_{1/2}(t) = F(t)$ *when* $\gamma = -1$.)

The preceding result leads to the first part of the following theorem.

THEOREM 4.

(a) *When* $\gamma = \pm 1$, *the random quantity* $(\delta_{*n} W_n(\theta), \delta_{*n} M_n(\theta) \delta_{*n})$ *involved in Theorem* 1 *converges in law to*

$$\left(\begin{bmatrix} \int F(t)\, dW_1(t) \\ \int F(t)\, dW_2(t) \\ Z \end{bmatrix}, \begin{bmatrix} \left(\int F^2(t)\, dt\right)\mathcal{I} & 0 \\ 0 & \tau^2 \end{bmatrix}\right).$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

(b) *When* $|\gamma| < 1$, *the vector* $n^{-1/2}W_n(\theta)$ *converges in law to the Gaussian vector* $(G', Z)'$ *with zero mean vector and covariance matrix as specified subsequently. In addition,* $n^{-1}M_n(\theta)$ *converges in probability to the covariance matrix of* $(G', Z)'$.

The covariance matrix of the preceding G is $E[\mathcal{E}_{12}^2](1-\gamma^2)^{-1}\mathcal{I}$, and the covariance between G and Z is given by

$$\sum_{j=0}^{\infty}\sum_{k=1}^{\infty} d_j d_k^{(1)} E(X_0^* \mathcal{I} U_{j+1-k}), \tag{17}$$

where X_0^* is the second component of $(1-\gamma L)^{-1}(1-L)^{-\alpha}\mathcal{E}_0$. It is important to note that the quantities (16) and (17) have played no role in the construction of the procedures of Section 2.1 whose limiting distributions, to be given subsequently, also will not involve them.

The next result gives statement (a) of the preceding theorem under contiguous alternatives.

THEOREM 5. *Under the alternatives* $(\beta + \delta_n h_{1n}, \mp 1 + \delta_n h_{2n}, \alpha + n^{-1/2}h_{3n})'$, $(h_{1n}, h_{2n}, h_{3n})' \to (h_1, h_2, h_3)'$, *the random quantity* $(\delta_{*n}W_n(\theta), \delta_{*n}M_n(\theta)\delta_{*n})$ *converges in law to*

$$\left(\begin{bmatrix} \int F_{h_2}(t)\, dW_1(t) \\ \int F_{h_2}(t)\, dW_2(t) \\ Z \end{bmatrix} + \begin{bmatrix} \left(\int F_{h_2}^2(t)\, dt\right)\mathcal{I} & 0 \\ 0 & \tau^2 \end{bmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}, \begin{bmatrix} \left(\int F_{h_2}^2(t)\, dt\right) & 0 \\ 0 & \tau^2 \end{bmatrix} \right),$$

where

$$F_{h_2}(t) = F(t) + h_2 \int_0^t \exp[(t-u)h_2]F(u)\, du.$$

Theorems 4 and 5, together with the representation (13), give the limiting distributions of the test statistics (11)–(14) under both the null and the contiguous alternatives. For instance, when $\gamma_0 = \mp 1$, the limiting distribution of (11) is given by

$$\left(\int F^2(t)\, dt\right)^{-1} \lambda_{22}^{-1} \left(\lambda_{21}\int F(t)\, dW_1(t) + \lambda_{22}\int F(t)\, dW_2(t)\right)^2, \tag{18}$$

where λ_{ij} is the (i,j)th element of $\mathcal{I}^{-1}$ and that of (12) or (12a), when γ is assumed to be in an unknown neighborhood of unity, is given by

$$\left(\int F^2(t)\, dt\right)^{-1} \lambda_{11}^{-1} \left(\lambda_{11}\int F(t)\, dW_1(t) + \lambda_{12}\int F(t)\, dW_2(t)\right)^2. \tag{19}$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

Now, assume, in addition, that $\mathcal{E}_1$ is Gaussian with mean zero and covariance matrix $\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$, so that $\lambda_{ij} = \sigma_{ij}$. Let $(B^*(t), B_{1/2}(t))'$ be the limit of the process $n^{-1/2} \sum_{i=1}^{[nt]} \mathcal{E}_i$. Let $\{B(s) : \infty < s < \infty\}$ be the standard Brownian motion. Let $F_*(t)$ be the functional of $B(s)$, constructed in the same way as $F(t)$ is constructed from $\{B_{1/2}(s) : \infty < s \leq 1\}$. Let $\mathcal{N}$ be a standard normal independent of $B(t)$.

Then (18) reduces to

$$\left(\int F^2(t)\, dt \right)^{-1} \left(\sigma_{22}^{-1/2} \int F(t)\, dB_{1/2}(t) \right)^2 = \left(\int F_*^2(t)\, dt \right)^{-1} \left(\int F_*(t)\, dB(t) \right)^2$$

and (19) reduces to

$$\left(\int F^2(t)\, dt \right)^{-1} \left(\sigma_{11}^{-1/2} \int F(t)\, dB^*(t) \right)^2 = \left(\rho \frac{\int F_*(t)\, dB(t)}{\sqrt{\int F_*^2(t)\, dt}} + \sqrt{1-\rho^2}\mathcal{N} \right)^2,$$

where $\rho = \lambda_{12}/\sqrt{\lambda_{11}\lambda_{22}}$. See further Remark 3 in Jeganathan (1997) for more details in this connection.

Note that the limiting distribution of the statistic (14) is given by

$$\mathcal{I}_{11}^{-1} \left(\int F_{h_2}(t)\, dt \right)^{-1} \left(\int F_{h_2}(t)\, dW_1(t) \right)^2,$$

which has the central chi-squared distribution with one degree of freedom, because the processes $F_{h_2}(t)$ (which is constructed from $B_H(t)$) and $W_1(t)$ are independent (see Proposition 2).

2.3. Some Extensions

First, as noted earlier, when γ is in the contiguous neighborhood of ± 1, the inference procedures of the present system can be transformed to those of (1) through the transformation $(\beta, \gamma) \mapsto (\beta\gamma, \gamma)$ (for more details see Jeganathan, 1997, Remark 1, Sect. 1).

Remark 2. We now indicate the modification needed for the situation in which the first component u_i of U_i has the fractional parameter α_1 with the associated expansions (4) and (5) with the coefficients $d_{1j}(\alpha_1)$ and $c_{1j}(\alpha_1)$, and similarly for the second component v_i having the parameter α_2 with the corresponding coefficients $d_{2j}(\alpha_2)$ and $c_{2j}(\alpha_2)$. Then, the score function $W_n(\theta)$ will involve the following modifications: First, it will now have four components corresponding to the four parameters. The term $\mathcal{E}_{ni}(\theta)$ will now involve both $d_{1j}(\alpha_1)$'s and $d_{2j}(\alpha_2)$'s. The $\widetilde{X}_i$ involved in the first component (associated with β) of $W_n(\theta)$ will now

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

involve $d_{1j}(\alpha_1)$'s in place of the earlier $d_j(\alpha)$'s, and similarly that corresponding to the second component will be based on $d_{2j}(\alpha_2)$'s. A similar modification will be involved for the last two components of $W_n(\theta)$ associated, respectively, with the parameters α_1 and α_2 and for the matrix $M_n(\theta)$, which will now be 4×4.

To indicate the corresponding modification for the normalizing matrix δ_{*n}, first define

$$\delta_{1n} = \begin{cases} n^{-1/2} & \text{if } |\gamma| < 1 \\ n^{-(1+\alpha_2)} & \text{if } \gamma = 1 \quad \text{and} \quad 0 \le \alpha_1 < \frac{1}{2} \\ n^{-(1+\alpha_2-\alpha_1)} & \text{otherwise.} \end{cases}$$

Let δ_{2n} be the δ_n defined earlier with $\alpha = \alpha_2$. Then the appropriate normalizing matrix is $\delta_{*n} = \text{diag}\{\delta_{1n}, \delta_{2n}, n^{-1/2}, n^{-1/2}\}$. With these modifications, Theorem 1 and the rest of the arguments of Section 2.1 hold.

Regarding the limiting distributions, the first two components of the previously modified $\delta_{*n} W_n(\theta)$ will converge in law to $\begin{bmatrix} \int F_1(t)\, dW_1(t) \\ \int F_2(t)\, dW_2(t) \end{bmatrix}$ where $F_2(t)$ is the earlier $F(t)$ involving $B_H(t)$ now with $\alpha = \alpha_2$, and

$$F_1(t) = \begin{cases} \left(\sum_{j=0}^{\infty} d_{1j}(\alpha) \right) B_H(t) & \text{if } \gamma = 1, 0 \le \alpha_1 < \frac{1}{2} \\ \dfrac{1}{(-\alpha_1 - 1)!} \displaystyle\int_0^t \frac{1}{(t-s)^{1+\alpha_1}} B_H(s)\, ds & \text{if } \gamma = 1, -\frac{1}{2} < \alpha_1 < 0 \\ 2^{\alpha_1 - \alpha_2} B_{1/2}(t) & \text{if } \gamma = -1, \end{cases}$$

where $B_H(t)$ as before but again with $\alpha = \alpha_2$. Similarly, the leading 2×2 submatrix of $\delta_{*n} M_n(\theta) \delta_{*n}$ associated with the first two components of $W_n(\theta)$ will converge in law to $[\mathcal{I}_{ij} \int F_i(t) F_j(t)\, dt]_{2\times 2}$.

We omit the straightforward details of the limiting distributions for the remaining components of $W_n(\theta)$ and $M_n(\theta)$.

Remark 3. Suppose that one has the system (2) but with u_i and v_i, respectively, replaced by the linear combinations $au_i + bv_i$ and $cu_i + dv_i$, for some constants a, b, c, d. Assume that u_i and v_i are as in Remark 2 with the respective fractional parameters α_1 and α_2. If a and b (respectively, c and d) are nonzero, then the role of α_2 (respectively, α_1) in the limiting distribution specified in Remark 2 will be played by $\max(\alpha_1, \alpha_2)$.

Remark 4. We now indicate that $\mathcal{E}_i$ themselves can be assumed to be a linear process with expansions similar to (4) and (5) in terms of i.i.d. vectors e_j, which are assumed to satisfy the same earlier conditions imposed for $\mathcal{E}_j$. Specifically assume that

$$\mathcal{E}_i = \sum_{j=0}^{\infty} \Delta_j e_{i-j} \quad \text{and} \quad e_i = \sum_{j=0}^{\infty} \Pi_j \mathcal{E}_{i-j}.$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

The coefficients may possibly involve unknown parameters. Further, the coefficients are assumed to be absolutely summable, in addition to the further remaining conditions of Section 3 of Jeganathan (1997). This will be the case when for instance $\mathcal{E}_i$ form a stationary autoregressive moving average (ARMA) model, and this special case may be kept in mind as an aid here.

To indicate the required modifications, take $e_{ni}(\theta) = \sum_{j=0}^{k_n^*} \Pi_j \mathcal{E}_{n,i-j}(\theta)$ with $\mathcal{E}_{ni}(\theta)$ as defined in (7). Here the vector θ now comprises all the parameters involved. Further, k_n^* is similar to the earlier k_n, but in situations such as stationary ARMA, it may be chosen to increase extremely slowly, for instance, at the rate of log n (see Jeganathan, 1997, Sect. 3), so that for the actual given sample size the sum involved will contain only a few terms. As before, the score function $W_n(\theta)$ is taken to be the appropriate derivative but now with $e_{ni}(\theta)$ involved in place of $\mathcal{E}_{ni}(\theta)$. This will induce the corresponding modification in the matrix $M_n(\theta)$. Interested readers can easily supply the details based on the details of the present Section 2.1 and those of Section 3 of Jeganathan (1997).

Regarding the normalizing matrix, it can be taken to be the same earlier δ_{*n}, in which case there will however be modifications in the limiting distributions: that of the first two components of the present $W_n(\theta)$ will be $(\sum_{j=0}^{\infty} \Pi_j)(\int F(t)\, dW_1(t), \int F(t)\, dW_2(t))'$ with $F(t), W_1(t), W_2(t)$ as in Section 2.2, and the limit of the associated 2×2 submatrix of $M_n(\theta)$ will be $(\int F^2(t))(\sum_{j=0}^{\infty} \Pi_j)\mathcal{I}(\sum_{j=0}^{\infty} \Pi_j)'$.

It is also important to note that in the special case when Π_j is of the diagonal form $\mathrm{diag}\{\pi_{1,j}, \pi_{2,j}\}$ for constants $\pi_{1,j}$ and $\pi_{2,j}$, there will be no change in the limiting distributions of Wald procedures of Section 2.1 as can be easily checked (in this connection see Jeganathan, 1997, Remark 8).

3. INFERENCE PROCEDURES IN A GENERAL MODEL

3.1. Construction of Inference Procedures

The primary purpose of this section is to treat the general model (3), in which, to recall, the q-variate observations $Z_1, \ldots, Z_n$ are such that

$$\left.\begin{aligned} Z_i - AZ_{i-1} &= U_i \\ (1-L)^{\alpha} U_i &= \mathcal{E}_i \end{aligned}\right\} \tag{20}$$

with $\mathcal{E}_i$ i.i.d. with zero mean and finite nonsingular covariance matrix that will be assumed to be known. The first step is to obtain the quadratic approximation to the log-likelihood ratios in the vicinity of suitable auxiliary estimators of the parameters A and α. In the next step we shall use this approximation to make a rough identification of the approximate unit eigenvalues and their configuration, and then represent the model in a triangular form expressing the cointegrating relationships and the nonstationary trends, which will then lead to more precise inference procedures regarding the cointegrating and unit roots parameters.

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

Here $\mathcal{E}_1$ is assumed to have a Lebesgue density $p(z)$ with $z = (z_1, \ldots, z_q)'$. Introduce the q-vector

$$\psi(z) = (\psi_1(z), \ldots, \psi_q(z))',$$

where $\psi_i(z) = -\partial \log p(z)/\partial z_i$. As in Section 2, we shall require the following assumption.

Assumption. The variance–covariance matrix of $\psi(U_1)$, to be denoted by $\mathcal{I}$, is finite and nonsingular.

We shall let θ be the $q^2 + 1$-vector composed of the subvector $\mathrm{vec}(A)$ and α, with α the first element of θ. Similar to (7), define

$$\mathcal{E}_{ni}(\theta) = \sum_{l=0}^{i-2} d_l(\alpha) U_{i-l}(A)$$

with $U_i(A) = Z_i - AZ_{i-1}$. Then the role of (8) is now played by

$$W_n(\theta) = \begin{bmatrix} \sum_{i=k_n+1}^{n} H_i'(\theta)\psi(\mathcal{E}_{ni}(\theta)) \\ \sum_{i=k_n+1}^{n} \psi(\mathcal{E}_{ni}(\theta)) \otimes \tilde{Z}_{i-1}(\alpha) \end{bmatrix},$$

where

$$\tilde{Z}_{i-1}(\alpha) = \sum_{j=0}^{i-2} d_j(\alpha) Z_{i-j-1} \quad \text{and} \quad H_i(\theta) = \sum_{j=1}^{i-2} d_j^{(1)}(\alpha) U_{i-j}(A).$$

Also, define the symmetric matrix

$$M_n(\theta) = \begin{bmatrix} \sum_{i=k_n+1}^{n} H_i'(\theta)\mathcal{I}H_i(\theta) & \\ \sum_{i=k_n+1}^{n} \mathcal{I}H_i(\theta) \otimes \tilde{Z}_{i-1}(\alpha) & \mathcal{I} \otimes \sum_{i=k_n+1}^{n} \tilde{Z}_{i-1}(\alpha)\tilde{Z}_{i-1}'(\alpha) \end{bmatrix}.$$

We shall let Ω be the set of all θ for which all the eigenvalues of the corresponding A are bounded in absolute values by unity. *It is important to note here that by considering such θ in Ω, the intent is actually to consider all the alternatives contiguous to θ.* To proceed further we need to consider the Jordan decomposition of the matrix A. Recall that for any A corresponding to fixed θ there is a nonsingular matrix Q such that QAQ^{-1} has a quasi-diagonal form called the Jordan matrix. When θ is in Ω, we can choose Q such that the Jordan matrix will be of the form $\mathrm{diag}\{\Upsilon, \Gamma\}$, with the square matrix Υ, of order r, say, (respectively, Γ, of order $d = q - r$) made up of the Jordan blocks corresponding to the eigenvalues with absolute values strictly less than unity (respectively, equal to unity).

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

For convenience in the next subsection, we shall employ the real Jordan form (see, e.g., Shilov, 1977, p. 155), which will in particular entail that the elements Q, Υ, and Γ are real.

Now, if we write Q in the form $\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}$, with Q_1 and Q_2 of respective orders $r \times q$ and $d \times q$, then

$$QZ_i = \operatorname{diag}\{\Upsilon, \Gamma\} QZ_{i-1} + QU_i \tag{21}$$

so that

$$V_{i1} = Q_1 Z_i \quad \text{and} \quad V_{i2} = Q_2 Z_i$$

are, respectively, under θ, purely stationary and nonstationary processes.

To proceed further, let δ_n be the $d \times d$ normalizing matrix as defined in (36), which follows (whose exact nature is not required for the present). Define, with

$$\delta_{*n} = \operatorname{diag}\{\delta_n, \ldots, \delta_n\}_{qd \times qd}$$

and $\mathbf{Q} = \operatorname{diag}\{Q, \ldots, Q\}_{q^2 \times q^2}$, the $(q^2+1) \times (q^2+1)$ matrix

$$\eta_{*n} = \operatorname{diag}\{n^{-1/2}, \operatorname{diag}\{n^{-1/2} I_{qr}, \delta_{*n}\}\mathbf{Q}\}.$$

Then one has

$$\eta_{*n} W_n(\theta) = \begin{bmatrix} n^{-1/2} \sum_{i=k_n+1}^{n} H_i'(\theta) \psi(\mathcal{E}_{ni}(\theta)) \\ n^{-1/2} \sum_{i=k_n+1}^{n} \psi(\mathcal{E}_{ni}(\theta)) \otimes \widetilde{V}_{*,i-1}(\alpha) \\ \delta_{*n} \sum_{i=k_n+1}^{n} \psi(\mathcal{E}_{ni}(\theta)) \otimes \widetilde{V}_{**,i-1}(\alpha) \end{bmatrix}, \tag{22}$$

where

$$\widetilde{V}_{*,i-1}(\alpha) = \sum_{j=0}^{i-2} d_j(\alpha) V_{i-j-1,1} \quad \text{and} \quad \widetilde{V}_{**,i-1}(\alpha) = \sum_{j=0}^{i-2} d_j(\alpha) V_{i-j-1,2}.$$

It will be shown in the next section (see Propositions 10 and 12) that

$$\begin{aligned} &\delta_{*n} \sum_{i=k_n+1}^{n} \psi(\mathcal{E}_{ni}(\theta)) \otimes \widetilde{V}_{**,i-1}(\alpha) \quad \text{and} \\ &\delta_{*n} \left(\mathcal{I} \otimes \sum_{i=k_n+1}^{n} \widetilde{V}_{**,i-1}(\alpha) \widetilde{V}_{**,i-1}'(\alpha) \right) \delta_{*n}' \end{aligned} \tag{23}$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

converge in law. In addition it will be shown (see Proposition 13) that the vector

$$\begin{bmatrix} n^{-1/2}\sum_{i=k_n+1}^{n} H_i'(\theta)\psi(\mathcal{E}_{ni}(\theta)) \\ n^{-1/2}\sum_{i=k_n+1}^{n} \psi(\mathcal{E}_{ni}(\theta)) \otimes \widetilde{V}_{*,i-1}(\alpha) \end{bmatrix} \tag{24}$$

converges in law to the Gaussian random vector, independent of the preceding limits, with zero mean vector and covariance matrix C, say. Further, it will follow that

$$\eta_{*n} M_n(\theta) \eta_{*n}' = \operatorname{diag}\left\{ C, \delta_{*n}\left(\mathcal{I} \otimes \sum_{i=k_n+1}^{n} \widetilde{V}_{**,i-1}(\alpha)\widetilde{V}_{**,i-1}'(\alpha) \right)\delta_{*n}' \right\} + o_p(1). \tag{25}$$

These facts will give the joint distribution of (22) and (25) (see Theorem 14 of the next subsection).

We are now in a position to state the following generalization of Theorem 1. Here the likelihood ratios $\Lambda_n(\theta^*, \theta)$ are as defined in Section 2.1.

THEOREM 6. *Assume that the parameter θ is in Ω. Then for the preceding normalizing matrices η_{*n} and for any bounded q^2+1-vectors h_n, one has*

(a)

$$\Lambda_n(\theta + \eta_{*n}h_n, \theta) = h_n \eta_{*n} W_n(\theta) - \frac{1}{2} h_n' \eta_{*n} M_n(\theta) n_{*n} h_n + o_p(1),$$

$$\eta_{*n}(\theta + \eta_{*n}h_n)\eta_{*n} = \eta_{*n} M_n(\theta)\eta_{*n} + o_p(1)$$

and

$$\eta_{*n} W_n(\theta + \eta_{*n}h_n) = \eta_{*n} W_n(\theta) - \eta_{*n} M_n(\theta)\eta_{*n} h_n + o_p(1)$$

in $P_{\theta,n}$.

(b) *The sequences $\{P_{\theta,n}\}$ and $\{P_{\theta+\eta_{*n}h_n,n}\}$ are contiguous.*

(c) *The limiting distributions of the random quantities involved in* (a) *are as in Theorem* 14, *which follows.*

We shall now transform the quadratic approximation around θ in this result into that around the vicinity of a suitable auxiliary estimator to make a rough identification of the possible approximate unit eigenvalues and their configuration. Thus, let $\tilde{\theta}_n$ be the preliminary estimator constructed in analogy with the construction of Section 2.1. Then, discretize it, that is, divide the space R^{q^2+1} into cubes of side length unity and construct a discretized θ_n^* by taking $M_n^{1/2}(\tilde{\theta}_n)\theta_n^*$ to be any convenient point of the unit cube in which $M_n^{1/2}(\tilde{\theta}_n)\tilde{\theta}_n$ falls. Then it can be

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

seen that $\Lambda(\theta_n^* + \eta_{*n} h_n, \theta_n^*)$, the log-likelihood ratio around θ_n^*, has the quadratic approximation with the center

$$\hat{\theta}_n = \theta_n^* + M_n^{-1}(\theta_n^*) W_n(\theta_n^*), \tag{26}$$

from which one obtains a Wald-type quadratic for the inference purposes of the matrix A:

$$R(A, \hat{A}_n) = (\mathrm{vec}(A) - \mathrm{vec}(\hat{A}_n))' D_n(\hat{\theta}_n)(\mathrm{vec}(A) - \mathrm{vec}(\hat{A}_n)), \tag{27}$$

where the inverse $D_n^{-1}(\theta)$ is the submatrix of $M_n^{-1}(\theta)$ formed by its last q^2 rows and the last q^2 columns and $\mathrm{vec}(\hat{A}_n)$ is the subvector of $\hat{\theta}_n$ formed by its last q^2 elements. From the fact

$$(\eta'_{*n})^{-1}(\hat{\theta}_n - \theta) = (\eta_{*n} M_n(\theta) \eta'_{*n})^{-1} \eta_{*n} W_n(\theta) + o_p(1)$$

the limiting distribution of (27) can also be obtained (from Theorem 14), but its precise form is not very essential for the present.

Now, using the distance (27), the identification of the approximate unit eigenvalues and the multiplicities of their occurrences is done in exactly the same way as in Jeganathan (1997, Sect. 4.1). Briefly, first find $\hat{Q}_n$ such that $\hat{Q}_n \hat{A}_n \hat{Q}_n^{-1} = \hat{J}_n$, the Jordan matrix. Then, through the consideration of the minors and the corresponding "elementary divisors" of $\hat{A}_n$, the eigenvalues involved in $\hat{J}_n$ that are close to unity in absolute values are adjusted to unity, which will simultaneously result in the adjustment of their multiplicities also, so that the resulting adjusted $\bar{J}$ will be such that, with $\bar{A} = \hat{Q}_n \bar{J} \hat{Q}_n^{-1}$, the distance $R(\bar{A}, \hat{A}_n)$ will become "reasonable." Both the "true" parameters θ and $\hat{\theta}_n$ will then be in the probable range of $\bar{\theta} = \mathrm{vec}(\bar{A})$.

The next step is to confine the parameter space to the probable range of $\bar{\theta}$. Here the range is probable in the sense that any parameter point in that range will be suitably contiguous to $\bar{\theta}$, so that for the purpose of the study of the asymptotic behaviors of the inference procedures, $\bar{\theta}$ *itself will be taken as the true parameter* (independent of n). Note that although this procedure will be approximately reasonable, it is not entirely without difficulties (of logical nature) when the entire procedure is viewed in a single stage (for the indications, see Jeganathan, 1997, Sect. 4.1). (Note however that no such difficulties arise when the procedure is viewed as a two-stage process, with the first stage corresponding to the approximate identification of the unit roots.)

The next step is to obtain a reduced form in the range indicated previously, identifying the cointegrating relationships, as was done in Jeganathan (1997). To indicate this, assume again that $\bar{J}$ is of the form $\mathrm{diag}\{\bar{\Upsilon}, \bar{\Gamma}\}$ with $\bar{\Upsilon}$, of order r, say, consisting of all the Jordan blocks corresponding to the eigenvalues less than unity in absolute values. For any A in the probable range $\bar{A}$, consider its Jordan decomposition $A = Q^{-1} J Q$ in the form conformable to that of $\bar{A}$. Partition J and Q in the forms

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

$$J = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$$

where J_{11} and Q_{11} (respectively, J_{22} and Q_{22}) are square matrices of order r (respectively, of order $d = q - r$). Further, let $Q_1 = [Q_{11}, Q_{12}]$ and $Q_2 = [Q_{21}, Q_{22}]$. Define $V_{i1} = Q_1 Z_i$ and $V_{i2} = Q_2 Z_i$. Note that when $A = \bar{A}$, V_{i1} and V_{i2} are, respectively, purely stationary and purely nonstationary processes, and the same will be true approximately in a suitable sense when A is in the probable range of $\bar{A}$. Assuming without loss of generality that Q_{11} is nonsingular, one has $Q_{11}^{-1} V_{i1} = Q_{11}^{-1} Q_1 Z_i = Z_{i1} + Q_{11}^{-1} Q_{21} Z_{i2}$ for disjoint subvectors z_{i1} and Z_{i2} of Z_i such that $Z_i = (Z_{i1}', Z_{i2}')'$, of respective orders r and d. Then following the arguments in Section 4.1 of Jeganathan (1997), it can be shown that the model (20) can be approximately represented, with L the back-shift operator, in the form

$$Z_{i1} - B_* Z_{i2} = (I - L Q_{11}^{-1} J_{11} Q_{11})^{-1} Q_{11}^{-1} Q_1 U_i \tag{28}$$

and

$$\left.\begin{aligned} Z_{i2} + T^{-1} Q_{21} (Z_{i1} + Q_{11}^{-1} Q_{12} Z_{i2}) &= T^{-1} V_{i2} \\ T^{-1} V_{i2} - \Gamma_* T^{-1} V_{i-1,2} &= T^{-1} J_{21} (I - L J_{11})^{-1} Q_1 U_{i-1} + T^{-1} Q_2 U_i \end{aligned}\right\}, \tag{29}$$

where the quantities involved are defined as follows: $T = Q_{22} - Q_{21} Q_{11}^{-1} Q_{12}$, $\Gamma_* = T^{-1} J_{22} T$ and $B_* = -Q_{11}^{-1} (Q_{12} + (I - L J_{11})^{-1} J_{12} T)$. In addition, the component $T^{-1} Q_{21} (Z_{i1} + Q_{11}^{-1} Q_{12} Z_{i2})$ in the left-hand side (l.h.s.) of (29) can be approximately replaced by $\bar{T}^{-1} \bar{Q}_{21} (Z_{i1} + \bar{Q}_{11}^{-1} \bar{Q}_{21} Z_{i2})$ where the parameters involved are at $\bar{A}$. Thus (28) is the cointegrating relationship, and Z_{i2} represents the nonstationary trend.

Further note that $Z_{i1} - B_* Z_{i2}$ can be approximately written in the form $Z_{i1} - B_{**} Z_{i-1,2}$ with $B_{**} = B_* \Gamma_*$.

In view of the preceding discussion, one possible approach is to treat the parameters B_* and Γ_*, together with those involved in the error components of (28) and (29), as freely varying and treat the situation as a suitable generalization of the situation of Section 2. Now using essentially the arguments given in Jeganathan (1997) (see further Remark 4 in the current paper) it can be shown that the quadratic approximation to the marginal log-likelihood ratios with respect to the parameters B_* and Γ_* exists, which will in addition recover the nonstationary component in the approximation of Theorem 6, having the matrix of the quadratic $\delta_{*n} (\mathcal{I} \otimes \sum_{i=k_n+1}^{n} \tilde{V}_{**,i-1}(\alpha) \tilde{V}_{**,i-1}'(\alpha)) \delta_{*n}'$ (see equation (25)).

This will lead, as in Jeganathan (1997, Sect. 4.1), to the fact, letting

$$\mu(\theta) = \begin{bmatrix} \mathrm{vec}(B_{**}) \\ \mathrm{vec}(\Gamma_*) \end{bmatrix},$$

that for a suitable subvector $\theta_{(2)}$ of θ, the vector $(\mu'(\theta), \theta_{(2)}')'$ will be in one–one correspondence with θ in the contiguous vicinity of $\bar{\theta}$, and, in addition, the esti-

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

mator $\mu(\hat{\theta}_n)$ can be used to fit the indicated marginal quadratic approximation. Thus for the inference purposes of the parameters in (28) and (29), one can use the estimator $\hat{\theta}_n$ obtained under the unrestricted level (20).

To describe this more precisely, let $\mathcal{I}^*$ be the Fisher information of the density of $\begin{bmatrix} Q_{11}^{-1}Q_1U_1 \\ T^{-1}Q_2U_1 \end{bmatrix}$. Now, in computing the matrix that will be involved in the quadratic approximation, (29) can be taken to be $Z_{i,2} = \Gamma_* Z_{i-1,2} + T^{-1}Q_2U_i$. Then the indicated matrix will take the form

$$C_n(\theta) = \begin{bmatrix} \mathcal{I}_{11}^* \otimes \sum_{i=k_n+1}^{n} \tilde{Z}_{i-1,2}^*(\alpha)\tilde{Z}_{i-1,2}^{*\prime}(\alpha) & \mathcal{I}_{12}^* \otimes \sum_{i=k_n+1}^{n} \tilde{Z}_{i-1,2}^*(\alpha)\tilde{Z}_{i-1,2}^{\prime}(\alpha) \\ \mathcal{I}_{21}^* \otimes \sum_{i=k_n+1}^{n} \tilde{Z}_{i-1,2}(\alpha)\tilde{Z}_{i-1,2}^{*\prime}(\alpha) & \mathcal{I}_{22}^* \otimes \sum_{i=k_n+1}^{n} \tilde{Z}_{i-1,2}(\alpha)\tilde{Z}_{i-1,2}^{\prime}(\alpha) \end{bmatrix},$$

where

$$\begin{bmatrix} \mathcal{I}_{11}^* & \mathcal{I}_{12}^* \\ \mathcal{I}_{21}^* & \mathcal{I}_{22}^* \end{bmatrix} = \mathcal{I}^*$$

with $\mathcal{I}_{11}^*$ of order $r \times r$ and

$$\tilde{Z}_{i-1,2}(\alpha) = \sum_{j=0}^{i-2} d_j(\alpha) Z_{i-j-1,2} \quad \text{and} \quad \tilde{Z}_{i-1,2}^*(\alpha) = (I - Q_{11}^{-1}J_{11}Q_{11})^{-1}\tilde{Z}_{i-1,2}(\alpha).$$

Further define $\eta_n = \operatorname{diag}\{\eta_{1n}, \eta_{2n}\}$ with $\eta_{1n} = \operatorname{diag}\{\delta_n T(I - Q_{11}^{-1}J_{11}Q_{11}), \ldots, \delta_n T(I - Q_{11}^{-1}J_{11}Q_{11})\}_{d^2 \times d^2}$ and $\eta_{2n} = \operatorname{diag}\{\delta_n T, \ldots, \delta_n T\}_{rd \times rd}$ where δ_n (defined in (36)) is as involved earlier. Denote the distribution of the observations $Z_i, i = 1, \ldots, n$ under the reparameterization $(\mu', \theta_{(2)}')'$ by $P_{\mu,\theta_{(2)},n}$. We shall let $\mu(\bar{\theta})$ by $\bar{\mu}$. Further μ_n will be said to be in the vicinity of $\bar{\mu}$ whenever $\mu_n = \bar{\mu} + \eta_n' u_n$ for bounded vectors u_n.

THEOREM 7. *Assume that the information matrix $\mathcal{I}$ is nonsingular and finite. Then for any μ_n in the vicinity, in the preceding sense, of $\bar{\mu}$, and for bounded vectors h_n and t_n of appropriate orders, one has the following conclusions*:

(a)

$$\log \frac{dP_{\mu_n+\eta_n'h_n,\bar{\theta}_{(2)}+n^{-1/2}t_n,n}}{dP_{\mu_n,\bar{\theta}_{(2)}+n^{-1/2}t_n,n}} = h_n'\eta_n C_n(\hat{\theta}_n)(\hat{\mu}_n - \mu_n) - \frac{1}{2}h_n'\eta_n C_n(\hat{\theta}_n)\eta_n' h_n + o_p(1)$$

in $P_{\bar{\mu},\bar{\theta}_{(2)},n}$.

(b) *The sequences* $\{P_{\bar{\mu}+\eta_n'h_n,\bar{\theta}_{(2)}+n^{-1/2}t_n,n}\}$ *and* $\{P_{\bar{\mu},\bar{\theta}_{(2)},n}\}$ *are contiguous.*

Note that $\eta_n C_n(\theta)\eta_n'$ involved in the preceding quadratic reduces to that in (23) except that $\mathcal{I}^*$ is involved instead of $\mathcal{I}$. Thus, in view of the relationship between

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

$\mathcal{I}^*$ and $\mathcal{I}$, the preceding quadratic approximation recovers the nonstationary component of that of Theorem 6.

In view of this result a Wald-type statistic for the parameter μ and for the marginal parameters $\mu_{(1)} = \text{vec}(B_{**})$ and $\mu_{(2)} = \text{vec}(\Gamma_*)$ can be constructed in the usual manner, and their limiting behaviors will follow from the limiting behaviors of the quantities involved in the approximating quadratic part of the statement of (a) of the preceding result, in exactly the same manner as in Jeganathan (1997, Sect. 4).

3.2. Limiting Behaviors

In this section we obtain the limiting distributional structures of the random quantities involved in the preceding section. We shall first consider the quantities in (23). For this purpose, assume that the Jordan decomposition $QAQ^{-1} = \text{diag}\{\Upsilon, \Gamma\}$ is in the real form. Recall that under θ, the eigenvalues of Γ are assumed to lie on the unit circle. We then have the form

$$\Gamma = \text{diag}\{\Gamma_1, \ldots, \Gamma_s\}, \tag{30}$$

where each block Γ_l, of order r_l, say, corresponds to an eigenvalue, (such that $r_1 + \cdots + r_s = d$). (It may be recalled that the eigenvalues corresponding to distinct blocks need not be distinct.) The forms of the blocks corresponding, respectively, to the eigenvalues $+1$ and -1 are

$$\begin{bmatrix} 1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 1 & \ldots & 0 & 0 \\ . & . & . & \ldots & . & . \\ 0 & 0 & 0 & \ldots & 1 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ . & . & . & \ldots & . & . \\ 0 & 0 & 0 & \ldots & -1 & 1 \\ 0 & 0 & 0 & \ldots & 0 & -1 \end{bmatrix}.$$

The blocks corresponding to the complete conjugate roots are of the form

$$\begin{bmatrix} \Delta' & I & 0 & \ldots & 0 & 0 \\ 0 & \Delta' & I & \ldots & 0 & 0 \\ . & . & . & \ldots & . & . \\ 0 & 0 & 0 & \ldots & \Delta' & I \\ 0 & 0 & 0 & \ldots & 0 & \Delta' \end{bmatrix},$$

where I is the unit matrix of order 2 and Δ is the rotation matrix

$$\Delta = \begin{bmatrix} \cos\phi & \sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}, \qquad 0 < \phi < \pi. \tag{31}$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

Convention. For notational convenience, we shall denote $V_{i2} = Q_2 Z_i$ of Section 3.1 by V_i.

Then, with $Q_2 U_i = \xi_i$, one has by (21),

$$V_i = \Gamma V_{i-1} + \xi_i$$

so that, by (30),

$$V_i = (V_{1i}', \ldots, V_{li}')'$$

with

$$V_{li} = \Gamma_l V_{l,i-1} + \xi_{li}, \tag{32}$$

for each $1 \le l \le s$ and $1 \le i \le n$, where V_{li}'s and ξ_{li}'s are, respectively, the conformally defined subvectors of V_i and ξ_i.

The idea involved in introducing the subvectors V_{li}'s is to express the quantities in (23) in terms of V_{li}'s, which are easier and convenient to handle both individually and in combinations. Then, ignoring the normalizing δ_{*n}, the qd-vector (the first quantity) in (23) will be composed of the subvectors

$$\sum_{i=k_n+1}^{n} \psi(\mathcal{E}_{ni}(\theta)) \otimes \widetilde{V}_{l,i-1}(\alpha), \qquad l = 1, \ldots, s \tag{33}$$

with

$$\widetilde{V}_{l,i-1}(\alpha) = \sum_{j=0}^{i-2} d_j(\alpha) V_{l,i-j-1}. \tag{34}$$

Similarly, the matrix in (23) will be composed of the submatrices

$$\mathcal{I} \otimes \sum_{i=k_n}^{n} \widetilde{V}_{l,i-1}(\alpha) \widetilde{V}_{m,i-1}'(\alpha), \qquad l, m = 1, \ldots, s. \tag{35}$$

These quantities involve the vectors V_{li}'s whose limiting distributions will be handled directly as follows.

Convention. In what follows it would be convenient to identify the eigenvalues of Γ_l, when the eigenvalues involved are complex conjugates $e^{i\phi_l}$ and $e^{-i\phi_l}$, by ϕ_l, $0 < \phi_l < \pi$. When the eigenvalue of Γ_l is $+1 = e^{i0}$ or $-1 = e^{i\pi}$ identify it by the corresponding $\phi_l = 0$ or π. Thus $0 \le \phi_l \le \pi$. *Throughout what follows* $\phi_l, l = 1, \ldots, s$ *will be fixed.*

We shall first obtain the limiting distributions of the matrices in (35). We shall show presently that the limiting behavior of V_{li} will involve effectively only the following processes:

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

$$\rho_{n,l}(t) = \begin{cases} n^{-H} \sum_{j=0}^{[nt]} \xi^*_{l,[nt]-j} & \text{if } \phi_l = 0 \\ n^{-1/2} \sum_{j=0}^{[nt]} (-1)^{[nt]-j} \xi^*_{l,[nt]-j} & \text{if } \phi_l = \pi \\ n^{-1/2} \sum_{j=0}^{[nt]} \Delta^{[nt]-j} \xi^*_{l,[nt]-j} & \text{if } 0 < \phi_l < \pi \end{cases}$$

involving the following components of $\xi_{l,i}$ of (32):

$$\xi^*_{l,i} = \begin{cases} \text{The last element of the vector } \xi_{l,i} \text{ when } \phi_l = 0 \text{ or } \pi \\ \text{Vector of order 2 formed by the last two elements of } \xi_{l,i} \text{ when } 0 < \phi_l < \pi. \end{cases}$$

To indicate the reason for this, consider for instance the case Γ_l corresponding to the eigenvalue $+1$. Then, solving (32), $V_{li} = \sum_{j=0}^{i} \Gamma_l^j \xi_{l,i-j}$, and because the (a,b)th element of Γ_l^j is $j(j-1)\ldots(j-b+a+1)/(1)(2)\ldots(b-a)$, it turns out that the last columns of Γ_l^j's dominate the asymptotic behavior of the sum so that only the last element $\xi^*_{l,i}$ of $\xi_{l,i}$ will effectively be involved.

Following is then a generalization of Proposition 2.

PROPOSITION 8. *The process* $(\rho_{n,1}(t),\ldots,\rho_{n,s}(t))$ *converges in law to the Gaussian process* $(C_1\rho_1(t),\ldots,C_s\rho_s(t))$, *whose distributional behavior is described subsequently*, *where* C_l *are some constants*.

In this proposition, the exact nature of the constants C_l will not be required because they will disappear in the next proposition. To describe the limiting distribution of Proposition 8, we need to specify only the covariance function of the limiting process. For this purpose define

$$e_{l,i} = (1-L)^{\alpha} \xi^*_{l,i},$$

where $\xi^*_{l,i}$ are as earlier. Note that $e_{l,i}$ are i.i.d.

We now specify the covariances of $\rho_l(t)$ and $\rho_m(t)$. First, if ϕ_l and ϕ_m are such that $\phi_l \neq \phi_m$, then these processes are independent.

If $\phi_l = \phi_m = \pi$, then $\rho_l(t)$ and $\rho_m(t)$ form a two-dimensional Brownian motion with the covariance matrix given by that of the vector $(e_{l,1}, e_{m,1})'$.

When $0 < \phi_l = \phi_m < \pi$, note that both $\rho_l(t)$ and $\rho_m(t)$ are two-dimensional; they form a four-dimensional Brownian motion with the covariance matrix specified by $\mathrm{Var}(\rho_l(t)) = \mathrm{Tr}\ \mathrm{Var}(e_{l,1})$, $\mathrm{Var}(\rho_m(t)) = \mathrm{Tr}\ \mathrm{Var}(e_{m,1})$, $\mathrm{Cov}(\rho_l(t), \rho_m(t)) = \mathrm{Tr}\,\mathrm{Cov}(e_{l,1}, e_{m,1})$, where Tr stands for the trace of the matrix.

When $\phi_l = \phi_m = 0$, the joint distribution of $\rho_l(t)$ and $\rho_m(t)$ is given by that of the two-dimensional fractional Brownian motion with covariance matrix function $E[\rho_l(t)\rho'_m(s)]$ given by $\frac{1}{2}C_H[|t|^{2H} + |s|^{2H} - |t-s|^{2H}]\mathrm{Var}[(\xi_{l,1}, \xi_{m,1})']$ with C_H as in (15). Now define

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

$$\tilde{\rho}_{n,l}(t) = \begin{cases} n^{-H^*} \sum_{j=0}^{[nt]} d_j(\alpha)\rho_{l,[nt]-j} & \text{if } \phi_l = 0 \\ n^{-1/2} \sum_{j=0}^{[nt]} d_j(\alpha)(-1)^j \rho_{l,[nt]-j} & \text{if } \phi_l = \pi \\ n^{-1/2} \sum_{j=0}^{[nt]} d_j(\alpha)\Delta^j \rho_{l,[nt]-j} & \text{if } 0 < \phi_l < \pi \end{cases}$$

with Δ as defined in (31) and $H^* = \frac{1}{2} + \max\{0,\alpha\}$. Then one has

PROPOSITION 9. *The process* $(\tilde{\rho}_{n,1}(t),\ldots,\tilde{\rho}_{n,s}(t))$ *converges in law to the Gaussian process* $(\tilde{\rho}_1(t),\ldots,\tilde{\rho}_s(t))$ *where*

$$\tilde{\rho}_l(t) = \begin{cases} \rho_l(t) & \text{if } 0 < \phi_l \le \pi \\ \left(\sum_{j=0}^{\infty} d_j(\alpha)\right)\rho_l(t) & \text{if } \phi_l = 0, \quad 0 \le \alpha < \frac{1}{2} \\ \dfrac{1}{(-\alpha-1)!}\displaystyle\int_0^t \frac{1}{(t-u)^{1+\alpha}}\,\rho_l(u)\,du & \text{if } \phi_l = 0, \quad -\frac{1}{2} < \alpha < 0. \end{cases}$$

Now, to deal with the matrices in (35), define normalizing matrices

$$N_l = \begin{cases} \operatorname{diag}\left\{\dfrac{1}{n^{r_l-1+H^*}},\ldots,\dfrac{1}{n^{1+H^*}},\dfrac{1}{n^{H^*}}\right\} & \text{if } \phi_l = 0 \\ \operatorname{diag}\left\{\dfrac{(-1)^{r_l-1}}{n^{r_l-1+1/2}},\ldots,\dfrac{(-1)^1}{n^{1+1/2}},\dfrac{1}{n^{1/2}}\right\} & \text{if } \phi_l = \pi \\ \operatorname{diag}\left\{\dfrac{\Delta^{r_l/2-1}}{n^{r_l/2-1+1/2}},\ldots,\dfrac{\Delta^1}{n^{1+1/2}},\dfrac{\Delta^0}{n^{1/2}}\right\} & \text{if } 0 < \phi_l < \pi. \end{cases}$$

It is also convenient to introduce the processes

$$\hat{V}_{n,l}(t) = \begin{cases} N_l \widetilde{V}_{l,[nt]}(\alpha) & \text{if } \phi_l = 0 \\ N_l(-1)^{[nt]}\widetilde{V}_{l,[nt]}(\alpha) & \text{if } \phi_l = \pi, \\ N_l \mathbf{\Delta}^{[nt]}\widetilde{V}_{l,[nt]}(\alpha) & \text{if } 0 < \phi_l < \pi \end{cases}$$

where $\widetilde{V}_{l,i}$ is as defined in (34) and $\mathbf{\Delta} = \operatorname{diag}\{\Delta,\ldots,\Delta\}_{r_l \times r_l}$.

We now describe the limits of these processes. Define the r_l-vector process (r_l is the order of the Jordan block Γ_l)

$$S_l(t) = (S_{l,1}(t),\ldots,S_{l,r_l}(t))',$$

the right-hand side (r.h.s.) of which is defined as follows.

Case $\phi_l = 0$ or π.

$$S_{l,r_l}(t) = \tilde{\rho}_l(t)$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

and, for $a = 1, \ldots, r_l - 1$,

$$S_{l,a}(t) = \int_0^t S_{l,a+1}(u)\, du.$$

Case $0 < \phi_l < \pi$. (Note that r_l is even in this case)

$$(S_{l,r_l-1}(t), S_{l,r_l}(t))' = \tilde{\rho}_l(t)$$

and, for $a = 1, \ldots, (r_l/2) - 1$,

$$(S_{l,2a-1}(t), S_{l,2a}(t))' = \int_0^t (S_{l,2a+1}(u), S_{l,2a+2}(u))'\, du.$$

We are now in a position to describe the limiting distributions of the random matrices in (35) in the following result where

$$\delta_n = n^{-1/2}\mathrm{diag}\{N_1, \ldots, N_s\}. \tag{36}$$

PROPOSITION 10. *The process* $(\hat{V}'_{n,1}(t), \ldots, \hat{V}'_{n,s}(t))'$ *converges in law to* $S(t) = (S'_1(t), \ldots, S'_s(t))'$. *In particular* $\mathcal{I} \otimes (\sum_{i=k_n+1}^n \delta_n \widetilde{V}_{l,i-1}(\alpha)\widetilde{V}'_{m,i-1}(\alpha)\delta'_n) = \mathcal{I} \otimes (n^{-1}\sum_{i=k_n+1}^n \hat{V}_{n,l}(i-1/n)\hat{V}'_{n,m}(i-1/n))$ *converges in law to* $\mathcal{I} \otimes (\int S_l(t)S'_m(t)\, dt)$, $l, m = 1, \ldots, s$. *Hence the matrix* $\delta_{*n}(\mathcal{I} \otimes \sum_{i=k_n}^n \times \widetilde{V}_{**,i-1}(\alpha)\widetilde{V}'_{**,i-1}(\alpha))\delta'_{*n}$ *in* (23) *converges in law to* $\mathcal{I} \otimes (\int S(t)S'(t)dt)$.

We next obtain the limiting distributions of the vectors in (33). Note that $n^{-1/2}\mathbf{N}_1 \sum_{i=k_n+1}^n \psi(\mathcal{E}_{ni}(\theta)) \otimes \widetilde{V}_{l,i-1}(\alpha)$

$$= \begin{cases} \displaystyle\sum_{i=k_n+1}^n n^{-1/2}\psi(\mathcal{E}_{ni}(\theta)) \otimes N_l \widetilde{V}_{l,i+1}(\alpha) & \text{if } \phi = 0 \\ \displaystyle\sum_{i=k_n+1}^n n^{-1/2}(-1)^{i-1}\psi(\mathcal{E}_{ni}(\theta)) \otimes (-1)^{i-1} N_l \widetilde{V}_{l,i-1}(\alpha) & \text{if } \phi = \pi, \\ \displaystyle\sum_{i=k_n+1}^n n^{-1/2}\psi(\mathcal{E}_{ni}(\theta)) \otimes (\boldsymbol{\Delta}')^{i-1} N_l \boldsymbol{\Delta}^{i-1} \widetilde{V}_{l,i-1}(\alpha) & \text{if } 0 < \phi < \pi \end{cases}$$

(37)

where $\boldsymbol{\Delta}^i = \mathrm{diag}\{\Delta^i, \ldots, \Delta^i\}_{r_l \times r_l}$ and $\mathbf{N}_1 = \mathrm{diag}\{N_l, \ldots, N_l\}_{qr_l \times qr_l}$. To obtain the limiting distribution of this define

$$\mathcal{X}_{n,l}(t) = \begin{cases} n^{-1/2}\displaystyle\sum_{i=2}^{[nt]} \psi(\mathcal{E}_{ni}(\theta)) & \text{if } \phi_l = 0 \\ n^{-1/2}\displaystyle\sum_{i=2}^{[nt]} (-1)^{i-1}\psi(\mathcal{E}_{ni}(\theta)) & \text{if } \phi_l = \pi. \\ n^{-1/2}\displaystyle\sum_{i=2}^{[nt]} ([\cos(i-1)\phi_l]\psi'(\mathcal{E}_{ni}(\theta)), [\sin(i-1)\phi_l]\psi'(\mathcal{E}_{ni}(\theta)))' & \\ \quad = (\mathcal{X}'_{n,l1}(t), \mathcal{X}'_{n,l2}(t))', \text{ say}, & \text{if } 0 < \phi_l < \pi \end{cases}$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

Note that when $0 < \phi_l < \pi$, $\mathcal{X}_{n,l}(t)$ is a $2q$-vector. Further note that when $\phi_l = \phi_m$, $\mathcal{X}_{n,l}(t) = \mathcal{X}_{n,m}(t)$, so that it is enough to find the joint distribution of the preceding processes for distinct ϕ_l's. When ϕ_l's are distinct, the processes $\mathcal{X}_{n,l}(t)$'s converge in law jointly to independent limits $\mathcal{X}_l(t)$'s, say. When $\phi_l = 0$ or π, the limit $\mathcal{X}_l(t)$ is Gaussian with zero mean vector and covariance matrix $\mathcal{I}$. When $0 < \phi_l < \pi$, $\mathcal{X}_l(t) = (\mathcal{X}'_{l1}(t), \mathcal{X}'_{l2}(t))'$ with $\mathcal{X}_{l1}(t)$ and $\mathcal{X}_{l2}(t)$ i.i.d. Gaussian with zero mean vector and covariance matrix $\frac{1}{2}\mathcal{I}$.

The $\mathcal{X}_{n,l}(t)$'s also converge in law jointly with the random vectors of Proposition 8, such that the joint distribution of $\mathcal{X}_l(t)$'s and $\rho_l(t)$'s is Gaussian. When $\phi_l \neq \phi_m$, $\rho_l(t)$ and $\mathcal{X}_m(t)$ are independent. When $\phi_l = \phi_m = \pi$ (respectively, $0 < \phi_l = \phi_m < \pi$), $\rho_l(t)$ and $\mathcal{X}_m(t)$ (respectively, $\rho_l(t)$ and $\mathcal{X}_{mi}(t)$ $(i = 1,2)$) form a two-dimensional Brownian motion. When $\phi_l = \phi_m = 0$, $\rho_l(t)$ and $\mathcal{X}_m(t)$ form a two-dimensional Gaussian process with $\mathrm{Cov}(\rho_l(u), \mathcal{X}_m(v)) = \frac{1}{2}[|u|^{H+1/2} + |v|^{H+1/2} - |u - v|^{H+1/2}]\mathrm{Cov}(e_{l,1}, \psi(\mathcal{E}_1))$.

PROPOSITION 11. *The processes $\mathcal{X}_{n,l}(t)$'s converge jointly in law with those of Proposition* 8 *to Gaussian processes, whose joint distribution is as described earlier.*

From this and Proposition 10, it follows easily that when $\phi_l = 0$ or π, (37) converges in law to $\int (d\mathcal{X}_l(t)) \otimes S_l(t)$. The corresponding description when $0 < \phi_l < \pi$ requires further care. In this case, following exactly the arguments given in Jeganathan (1997, Sect. 4.2, step III), it follows that (37) converges in law to

$$\mathrm{vec}([\zeta_{kl}])$$

with

$$\zeta_{k,2j-1} = \int S_{l,2j-1}(t)\, d\mathcal{X}_{l1k}(t) + \int S_{l,2j}(t)\, d\mathcal{X}_{l2k}(t),$$

$$\zeta_{k,2j} = \int S_{l,2j}(t)\, d\mathcal{X}_{l1k}(t) + \int S_{l,2j-1}(t)\, d\mathcal{X}_{l2k}(t),$$

where $\mathcal{X}_{l1k}(t)$ and $\mathcal{X}_{l2k}(t)$ are, respectively, the kth components of $\mathcal{X}_{l1}(t)$ and $\mathcal{X}_{l2}(t)$.

Now note that if we denote the vector (37) by W_{nl}, then for any qd-vector h, $h'\delta_{*n}\sum_{i=k_n}^{n}\psi(\mathcal{E}_{ni}(\theta)) \otimes \widetilde{V}_{**,i-1}(\alpha) = \sum_{l=1}^{s} h'_{(l)} W_{nl}$ for suitable subvectors $h_{(l)}$ of h. Thus if we denote the limit, as obtained previously, of W_{nl} by W_l, then these limits form a qd-vector W such that $\sum_{l=1}^{s} h'_{(l)} W_l = h'W$.

PROPOSITION 12. *The quantity $\delta_{*n}\sum_{i=k_n}^{n}\psi(\mathcal{E}_{ni}(\theta)) \otimes \widetilde{V}_{**,i-1}(\alpha)$ in* (23) *converges in law, jointly with the convergence in Proposition* 10, *to W defined previously.*

Next let us state the limiting distribution of the vector (24).

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

PROPOSITION 13. *The vector* (24) *converges in law, jointly with those of* (23), *to a Gaussian vector Z, say, with zero mean vector and a suitable covariance matrix C.*

The indicated covariance matrix C can be obtained in analogy with the covariance matrix of (G', Z) of statement (b) of Theorem 4, but it will not be involved in the construction, or the limiting distributions, of the Wald procedures.

We are now in a position to state the limiting distribution of the quantities involved in the approximation of Theorem 6.

THEOREM 14. *The collection* $\{\eta_{*n} W_n(\theta), \eta_{*n} M_n(\theta) \eta_{*n}\}$ *involved in Theorem* 6 *converges in law to*

$$\left\{ \begin{bmatrix} Z \\ W \end{bmatrix}, \begin{bmatrix} C & 0 \\ 0 & \mathcal{I} \otimes \int S(t) S'(t)\, dt \end{bmatrix} \right\}.$$

In addition Z and W are independent. The variables involved in the limit are as in Propositions 10, 12, *and* 13.

As noted earlier the weak limits of the random quantities involved in the quadratic approximation of Theorem 7 will also be the previous same W (which depends on $\mathcal{I}$) and $\mathcal{I} \otimes \int S(t) S'(t)\, dt$ but with $\mathcal{I}^*$ of that theorem involved in place of $\mathcal{I}$.

4. THE REMAINING PROOFS

Throughout this section i stands for the complex number $\sqrt{-1}$. Further we let $f(u) = |1 - \exp(-iu)|^{1-2H} = |1 - 2\cos u|^{1-2H}$. Note that $(E[\mathcal{E}_{12}^2]/2\pi) f(u)$ is the spectral density of the process $\{v_l\}$ involved in (2), where recall that $\mathcal{E}_{12}$ is the second component of $\mathcal{E}_1$. In addition, for notational convenience, the appearance of the parameters will be suppressed. Thus, for instance, $\psi(\mathcal{E}_i)$ stands for $\psi(\mathcal{E}_i(\theta))$.

Proof of Statement (a) of Proposition 2. We shall first give the proof for $t = 1$. First recall that $H_k = \sum_{j=1}^{k-2} d_j^{(1)} U_{k-j}$. Also note that the summands $H_k' \psi(\mathcal{E}_k)$ are martingale differences, with respect to the σ-fields $\sigma(\mathcal{E}_j; j \le k)$. Further $n^{-H} \sum_{k=1}^{n} v_k$, the first component of the vector involved, can be written in the form $\sum_{j=-\infty}^{n} n^{-H} a_{nj} \mathcal{E}_{j2}$ for suitable constants a_{nj}.

Thus for each n, the vector involved can be viewed as the sum of martingale differences. Hence we need to verify the Lindeberg condition and the convergence of conditional variances and covariances that, with respect to the first three elements of the vector involved, will be the same as the corresponding unconditional ones.

To start with consider the first three components of the vector. First note that the Lindeberg condition for the first component follows from the fact that

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

$\sup_j |n^{-H} a_{nj}| \to 0$. Next recall that, because $(E[\mathcal{E}_{12}^2]/2\pi) f(u)$ is the spectral density of the process $\{v_l\}$,

$$\operatorname{Var}\left(n^{-H}\sum_{k=1}^{n} v_k\right) = \frac{E[\mathcal{E}_{12}^2]}{2\pi} n^{-2H} \int_{-\pi}^{\pi} \left|\sum_{k=1}^{n} \exp(ik\lambda)\right|^2 f(\lambda)\, d\lambda$$

$$= \frac{E[\mathcal{E}_{12}^2]}{2\pi} \int_{-n\pi}^{n\pi} \left|n^{-1}\sum_{k=1}^{n} \exp\left(i\frac{k}{n}u\right)\right|^2 \left(n^{1-2H} f\left(\frac{u}{n}\right)\right) du, \tag{38}$$

which converges to the limit C_H as defined in connection with (15). Similarly,

$$\operatorname{Cov}\left(\frac{1}{n^H}\sum v_k, \frac{1}{n^{1/2}}\sum \psi_2(\mathcal{E}_k)\right)$$

$$= \frac{E[\mathcal{E}_{12}\psi_2(\mathcal{E}_1)]}{2\pi} \int_{-n\pi}^{n\pi} \left|\frac{1}{n}\sum \exp\left(i\frac{k}{n}u\right)\right|^2 \left(n^{1/2-H} a^*\left(\frac{u}{n}\right)\right) du,$$

which converges to $E[\mathcal{E}_{12}\psi_2(\mathcal{E}_1)]$, where $a^*(u)$ is the real part of $(1 - \exp(-iu))^{(1/2)-H}$. The same holds when $\psi_2(\mathcal{E}_k)$ is replaced by $\psi_1(\mathcal{E}_k)$, but the limit will be zero because $\operatorname{Cov}(\mathcal{E}_{12}, \psi_1(\mathcal{E}_1)) = 0$ (see Jeganathan, 1997, Proposition 2). The covariance matrix of the middle two components of the vector converges to $\mathcal{I}$.

Now consider the last component $n^{-1/2}\sum_{k=k_n+1}^{n} H_k' \psi(\mathcal{E}_k)$. First note that it is directly verified that the covariances between this and the remaining components of the vector are zero. We shall verify the convergence in probability of the sum $n^{-1}\sum_{k=k_n+1}^{n} H_k' \mathcal{I} H_k$ of the conditional variances to the limit (16), and the conditional Lindeberg condition, in the proof that follows of statement (b) of Theorem 4.

Thus the required convergence in law at $t = 1$ follows. Similarly the convergence of fdd can be obtained. To identify the limiting covariances note, for instance, that, for $s < t$, in view of the stationarity of v_i's, $\operatorname{Var}(n^{-H}\sum_{k=1}^{[nt]} v_k - n^{-H}\sum_{k=1}^{[ns]} v_k) = \operatorname{Var}(n^{-H}\sum_{k=1}^{[nt]-[ns]} v_k)$, which converges, by (38), to $C_H|t-s|^{2H}$. The identifications of the remaining limits are the same. ■

We remark that under the Gaussian restriction, the preceding convergence of fdd can be strengthened to convergence in the Skorokhod space $D[0,1]$ (see, e.g., Davydov, 1970; Gorodetskii, 1977; Taqqu, 1975), a fact that will be used subsequently for convenience though not essential. In Chung (1995) it is shown that this strengthening (for vectors with components having the forms $n^{-H}\sum_{k=1}^{n} v_k$ with possibly different H's) holds without the Gaussian restriction; earlier results in this direction in Davydov (1970) and Gorodetskii (1977) involved further moment conditions. This situation corresponds to the root $+1$, but unfortunately the same strengthening is not available presently for the non-Gaussian situations that

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

will be dealt with subsequently involving other possible unit roots. However, for our purpose the convergence in law of fdd is sufficient.

Proof of Statement (a) of Proposition 3. First consider the case $\gamma = +1$ and $0 < \alpha < \frac{1}{2}$. In this case, because $X_{[nt]} = \sum_{k=1}^{[nt]} v_k$ and in view of statement (a) of Proposition 2, the result will follow from the fact that

$$n^{-H}\widetilde{X}_{[nt]}(\alpha) - \left(\sum_{j=0}^{\infty} d_j\right) n^{-H} X_{[nt]} = o_p(1) \tag{39}$$

for every $0 \le t \le 1$, which can be verified as follows. First note that it is enough to show that the expected value of the square of the preceding difference converges to zero. In particular one can assume Gaussianity without loss of generality, in which case the convergence in this sense is same as the convergence in law. But, under Gaussianity, $n^{-H}X_{[nt]}$ converges in the Skorokhod space as noted earlier. Hence (39) follows easily because d_j's are absolutely summable when $0 < \alpha < \frac{1}{2}$.

Next consider the case $\gamma = +1$ and $-\frac{1}{2} < \alpha < 0$. (Note that now d_j are not absolutely summable.) In this case the difference

$$n^{-1/2}\widetilde{X}_{[nt]}(\alpha) - \sum_{l=1}^{k} n^{-H} X_{[nt]-[((l-1)/k)tn]-1}\left(n^{-((1/2)-H)} \sum_{j=[((l-1)/k)tn]+1}^{[(1/k)tn]} d_j\right) = o_p(1) \tag{40}$$

for every $0 \le t \le 1$, by first letting $n \to \infty$ and then letting $k \to \infty$. Assuming that this is true, the required result will follow because $d_j \sim 1/(-\alpha - 1)!\, j^{1+\alpha}$, entailing, for each fixed k,

$$n^{-((1/2)-H)} \sum_{j=[((l-1)/k)tn]+1}^{[(1/k)tn]} d_j \to \frac{1}{(-\alpha-1)!} \int_{t(l-1)/k}^{tl/k} \frac{1}{x^{1+\alpha}}\, dx, \quad \text{as } n \to \infty,$$

so that the r.h.s. of the difference in (40) can be approximated, for each fixed k, in view of statement (a) of Proposition 2, by

$$\frac{1}{(-\alpha-1)!} \sum_{l=1}^{k} B_H\left(t - t\left(\frac{l-1}{k}\right)\right) \int_{t(l-1)/k}^{tl/k} \frac{1}{x^{1+\alpha}}\, dx,$$

which converges, as $k \to \infty$, to the required limit. The proof of (40) uses the same ideas of the proof of (39) in the sense of restricting to the Gaussian situation. ■

Proof of Statement (b) of Proposition 2. For the latter purpose, we take the first component of the vector involved to be the more general sum $n^{-1/2} \sum_{k=0}^{[nt]} v_{[nt]-k} \exp(i(k - [nt])\phi)$. Then note that when $\phi = \pi$, this reduces to the first component of the vector involved.

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

The required proof parallels that of statement (a) given earlier. To indicate, first assume that $0 < \phi < \pi$. Then note that

$$\begin{aligned}\operatorname{Var}&\left(\frac{1}{n^{1/2}}\sum_{k=0}^{n} v_{[nt]-k}\exp(i(k-[nt])\phi)\right)\\ &= \frac{E[\mathcal{E}_{12}^2]}{2\pi}\, n^{-1}\int_{-\pi}^{\pi}\left|\sum_{k=1}^{n}\exp(ik(\lambda+\phi))\right|^2 f(\lambda)\,d\lambda\\ &= \frac{E[\mathcal{E}_{12}^2]}{2\pi}\int_{-n(\phi-\pi)}^{n(\phi+\pi)}\left|\frac{1}{n}\sum_{k=1}^{n}\exp\left(i\,\frac{k}{n}\,u\right)\right|^2 f\left(\frac{u}{n}-\phi\right)du, \qquad \textbf{(41)}\end{aligned}$$

which converges to

$$f(\phi)\,\frac{E[\mathcal{E}_{12}^2]}{2\pi}\int_{-\infty}^{\infty}\left|\frac{e^{itu}-1}{iu}\right|^2 du = f(\phi)E[\mathcal{E}_{12}^2]$$

when $0 < \phi < \pi$. It can be seen that the same holds when $\phi = \pi$ (because the integrand of the first integral in the preceding discussion is periodic with period 2π so that the range of integration can be replaced by any range of length 2π that will include π). Similarly the remaining covariance functions can be dealt with. ■

Proof of statement (b) of Proposition 3. Again for the latter purpose, we shall consider a more general case. First define $n^{-1/2}X^*_{[nt]} = n^{-1/2}\sum_{k=0}^{[nt]} v_{[nt]-k}\times \exp(ik\phi)$. Then note that $\exp(-[nt]i\phi)\sum_{j=0}^{[nt]} X^*_{[nt]-j}d_j$, which is the first component of the vector involved when $\phi = \pi$, reduces to $\sum_{k=0}^{[nt]} v_{[nt]-k}\exp((k-[nt])i\phi)(\sum_{j=0}^{k}\exp(-ij\phi)d_j)$. We shall show that

$$\exp(-[nt]i\phi)n^{-1/2}\sum_{j=0}^{[nt]} X^*_{[nt]-j}d_j - a(\phi)\sum_{k=0}^{[nt]} v_{[nt]-k}\exp((k-[nt])i\phi) = o_p(1), \qquad \textbf{(42)}$$

where $a(\phi) = \sum_{j=0}^{\infty} e^{-ij\phi}d_j = (1-e^{-i\phi})^{\alpha}$. Note that $f(\lambda) = |a(\lambda)|^{-2}$. In view of the preceding proof, the required result will follow from (42). To prove (42), note that the variance of the difference in (42) is

$$n^{-1}\int_{-\pi}^{\pi}\left|\sum_{k=0}^{[nt]}\exp((k-[nt])i(\phi+\lambda))\left(\sum_{j=0}^{k} e^{-ij\phi}d_j - a(\phi)\right)\right|^2 f(\lambda)\,d\lambda. \qquad \textbf{(43)}$$

Now note that $d_j = \pi^{-1}\int_{-\pi}^{\pi} a(u)\cos ju\,du = -i\pi^{-1}\int_{-\pi}^{\pi} a(u)\sin ju\,du$. Hence

$$\begin{aligned}\sum_{j=0}^{k} e^{-ij\phi}d_j &= \frac{1}{2\pi}\int_{-\pi}^{\pi} a(u)\,du + \frac{1}{\pi}\\ &\quad\times\sum_{j=1}^{k}\left[\cos j\phi\int_{-\pi}^{\pi} a(u)\cos ju\,du + \sin j\phi\int_{-\pi}^{\pi} a(u)\sin ju\,du\right]\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin(k+\frac{1}{2})u}{\sin\frac{1}{2}u}\,a(\phi+u)\,du.\end{aligned}$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

It can be shown, using the arguments that lead to Theorem 10.8 in Zygmund (1968), together with some further elementary arguments, that the absolute value of the difference

$$\sum_{j=0}^{k} e^{-ij\phi} d_j - a(\phi) = \pi^{-1} \int_{-\pi}^{\pi} \frac{\sin(k+\frac{1}{2})u}{\sin\frac{1}{2}u} [a(\phi+u) - a(\phi)]\, du$$

is of order $O(1/k^{1-\alpha^*})$ with $\alpha^* = \max\{0,\alpha\}$. Thus it follows that (43) is bounded in absolute value by

$$n^{2\alpha^*-1} \int_{-\pi}^{\pi} f(\lambda)\, d\lambda \to 0$$

because $-\frac{1}{2} < \alpha < \frac{1}{2}$. This proves the result. ■

Proof of Statement (a) of Theorem 4. We first consider $\delta_{*n} W_n(\theta)$. First assume that $\gamma = 1$. Note that the first component of $\delta_{*n} W_n(\theta)$ is given by $\sum_{j=k_n+1}^{n} n^{1/2}\delta_n \widetilde{X}_{j-1} n^{-1/2}\psi_1(\mathcal{E}_{nj})$. It will be shown later that $\mathcal{E}_{nj}$ involved in this sum can be replaced by $\mathcal{E}_j$ without disturbing the limiting behavior. Now fix k and define $\hat{X}_{j,nk} = n^{1/2}\delta_n \widetilde{X}[((l-1)n)/k]$ for $[((l-1)n)/k] \le j < [ln/k]$, $l = 1,\ldots,k$. Then the variance of the difference $\sum_{j=k_n+1}^{n}(n^{1/2}\delta_n \widetilde{X}_{j-1} - \hat{X}_{j-1,nk}) n^{-1/2}\psi_1(\mathcal{E}_j)$ (which is a sum of martingale differences) is given by $1/n \sum_{j=k_n+1}^{n} E[(n^{1/2}\delta_n \widetilde{X}_{j-1} - \hat{X}_{j-1,nk})^2]$. Using arguments similar to the ones involved in the preceding proofs, this last quantity can be computed explicitly and can be shown to converge to zero by first letting $n \to \infty$ and then letting $k \to \infty$. Similar approximation holds when $\psi_2(\mathcal{E}_j)$ is involved instead of $\psi_1(\mathcal{E}_j)$. Thus it is enough to consider the sums $\sum_{j=k_n+1}^{n} \hat{X}_{j-1,nk} n^{-1/2}\psi_l(\mathcal{E}_j)$, $l = 1,2$. It is easy to see that these sums together with the last component of $\delta_{*n} W_n(\theta)$ converge in law to the desired limit for $\delta_{*n} W_n(\theta)$, by first letting $n \to \infty$ and then letting $k \to \infty$, in view of statement (a) of Proposition 2. (Note that only the convergence of fdd is involved in the last step.) The same type of arguments can be used for the convergence of the quadratic $\delta_{*n} M_n \delta_{*n}$.

Similarly, when $\gamma = -1$, the first component of $\delta_{*n} W_n(\theta)$ can be approximated by $\sum_{j=k_n+1}^{n} \hat{X}^*_{j-1,nk} n^{-1/2}(-1)^{j-1}\psi_1(\mathcal{E}_j)$ where $\hat{X}^*_{j,nk} = n^{-1/2}(-1)[((l-1)n)/k]\widetilde{X}_{[((l-1)n)/k]}$ for $[((l-1)n)/k] \le j < [ln/k]$, $l = 1,\ldots,k$, and the same holds for the second component. Hence the required convergence follows from statement (b) of Proposition 2. ■

Proof of Statement (b) of Theorem 4. First consider $n^{-1/2}\sum_{k=k_n+1}^{n} \hat{X}_{k-1} \times n^{-1/2}\psi(\mathcal{E}_{nk})$, the vector consisting of the first two components of $\delta_{*n} W_n(\theta)$. As in the previous proof $\mathcal{E}_{nk}$ can be replaced by $\mathcal{E}_k$. Then we are dealing with a sum of martingale differences, so that we need to obtain

$$n^{-1}\sum \widetilde{X}^2_{k-1} = E[\mathcal{E}^2_{12}](1-\gamma^2)^{-1} + o_p(1) \quad \text{and} \quad \sup_{1<k\le n} |n^{-1/2}\widetilde{X}_{k-1}| = o_p(1). \quad \textbf{(44)}$$

For this purpose, define $X^*_k = X_k + R_k$ with $R_k = \sum_{l=0}^{\infty} \gamma^{k+l} U_l = \gamma^k R_0$. Assume for simplicity that $X_0 = 0$. Then note that $X^*_k = \sum_{l=0}^{\infty} \gamma^l U_{k-l}$. Now recall that

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

$\widetilde{X}_k = \sum_{j=0}^{k-2} d_j X_{k-j}$. Note that $\sum_{j=0}^{k-2} d_j R_{k-j} = R_0(\sum_{j=0}^{k-2} d_j \gamma^{k-j})$. Because $\sum_{j=0}^{k-2} d_j \gamma^{k-j} \to 0$ as $k \to \infty$, it follows that

$$n^{-1} \sum_{k=k_n+1}^{n} E\left[\left(\sum_{j=0}^{k-2} d_j R_{k-j}\right)^2\right] \to 0. \tag{45}$$

We shall next show that

$$n^{-1} \sum_{k=k_n+1}^{n} E\left[\left(\sum_{j=k-1}^{\infty} d_j X^*_{k-j}\right)^2\right] \to 0. \tag{46}$$

Note that (45) and (46) show that it is enough to prove (44) with $\widetilde{X}_{k-1}$ replaced by $\sum_{j=0}^{\infty} d_j X^*_{k-j} = \sum_{l=0}^{\infty} \gamma^l \mathcal{E}_{k-l,2}$, in which case (44) follows. To see that (46) holds, note that the spectral density of the process $\{X^*_k\}$ is given by $g(u) = |1 - \gamma \exp(iu)|^{-2} f(u)$. Hence $e[(\sum_{j=k-1}^{\infty} d_j X^*_{k-j})^2]$ is a constant multiple of $\int_{-\pi}^{\pi} |\sum_{j=k-1}^{\infty} d_j \exp(-iju)|^2 g(u)\, du$ that, in view of the arguments of the earlier proof of statement (b) of Proposition 3, is of order $O(k^{-2(1-\alpha)})$, from which (46) follows.

Now consider the last component $n^{-1/2} \sum_{k=k_n+1}^{n} H'_k \psi(\mathcal{E}_k)$. Now, for each fixed r, define $U_{k-j,r} = \sum_{s=0}^{r} c_k \mathcal{E}_{k-j-s}$. It can be easily shown, using arguments similar to those in connection with (46), together with the absolute summability of $d_j^{(1)}$, that $n^{-1} \sum_{k=k_n+1}^{n} E[|H_k - \sum_{j=1}^{r} d_j^{(1)} U_{k-j,r}|^2] \to 0$ by first letting $n \to \infty$ and then $r \to \infty$. From this, the fact (44) with $\widetilde{X}_{k-1}$ replaced by H_k, and the convergence in probability of the sum $n^{-1} \sum_{k=k_n+1}^{n} \widetilde{X}_{k-1} H_k$ of the conditional covariances to the desired limit, follow easily using arguments similar to those given previously. This completes the proof. ■

The proof of Theorem 5 is essentially the same as that of Jeganathan (1997, Statement (b), Theorem 3), and hence we omit it.

We next turn to the proof of Theorem 1. Given that certain conditions are verified, it will be seen that the proof, by applying the appropriate results in Jeganathan (1995), is essentially the same as that of Theorem 4 in Jeganathan (1997). The next lemma will contain most of the required conditions. To state it, recall that $\mathcal{E}_k = \sum_{j=0}^{\infty} d_j U_{k-j}$ stands for $\sum_{j=0}^{k-1} d_j(\alpha) U_{k-j}(\beta,\gamma) + \sum_{j=k}^{\infty} d_j(\alpha) U_{k-j}$. (Note that for $j \le 0$, U_j does not depend on the parameters.) Further, define the 2×3 matrix $\chi_{nk} = [\mathrm{diag}\{\delta_n \widetilde{X}_{k-1}, \delta_n \widetilde{X}_{k-1}\}, n^{-1/2} H_k]$. The corresponding quantities under the parameter $\theta + \delta_{*n} h_n$ will be denoted, respectively, by $\mathcal{E}_k(h_n)$ and $\chi_{nk}(h_n)$. Similarly d_j under $\theta + \delta_{*n} h_n$ will be of the form $d_j(\alpha + n^{-1/2} h_{3n})$, which will be denoted by $d_{jh_{3n}}$, with h_{3n} the last component of h_n.

LEMMA. *Under θ, one has*

$$\sum_{k=k_n+1}^{n} |\mathcal{E}_k(h_n) - \mathcal{E}_k - \chi_{nk} h_n|^2 = o_p(1), \tag{47}$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

$$\sup_{k_n+1\le k\le n} |\chi_{nk}| = o_p(1), \tag{48}$$

$$\sum_{k=k_n+1}^{n} |\chi_{nk}|^2 = O_p(1), \tag{49}$$

$$\sum_{k=k_n+1}^{n} |\chi_{nk}(h_n) - \chi_{nk}|^2 = o_p(1). \tag{50}$$

Proof. First note that the difference $\mathcal{E}_k(h_n) - \mathcal{E}_k - \chi_{nk}h_n$ can be written in the form

$$\delta_n \widetilde{X}_{k-1}(h_n) - \delta_n \widetilde{X}_{k-1} + \sum_{j=0}^{\infty} (d_{jh_{3n}} - d_j - n^{-1/2}h_{3n}d_j^{(1)})U_{k-j} + n^{-1/2}h_{3n}\sum_{j=k}^{\infty} d_j^{(1)}U_{k-j}.$$

Using the absolute summability of $d_j^{(1)}$, it follows that $n^{-1}\sum_{k=k_n+1}^{n} E \times [|\sum_{j=k}^{\infty} d_j^{(1)}U_{k-j}|^2]$ converges to zero. It is also easy to see that $\sum_{k=k_n+1}^{n} |\delta_n \widetilde{X}_{k-1}(h_n) - \delta_n \widetilde{X}_{k-1}|^2$ is of order $o_p(1)$. Further, because U_k are stationary, $\sum_{k=k_n+1}^{n} E[|\sum_{j=0}^{\infty}(d_{jh_{3n}} - d_j - n^{-1/2}h_{3n}d_j^{(1)})U_{k-j}|^2]$ is bounded by $nE[|\sum_{j=0}^{\infty}(d_{jh_{3n}} - d_j - n^{-1/2}h_{3n}d_j^{(1)})U_{-j}|^2]$, which converges to zero in view of the fact that $d_{jh_{3n}} = d_j(\alpha + n^{-1/2}h_{3n})$. Thus (47) follows. Further (48) and (49) follow from the arguments of the proof of statement (a) of Theorem 4. By similar reasoning, (50) also follows, completing the proof. ■

Proof of Theorem 1. For the moment assume that $\psi(\mathcal{E}_k(\theta))$ are involved instead of $\psi(\mathcal{E}_{nk}(\theta))$ in $W_n(\theta)$ and $M_n(\theta)$. Then using exactly the same arguments as that of Theorem 4 in Jeganathan (1997), it is seen that the result follows, in view of the preceding lemma and the weak convergence statement (a) of Theorem 4 (proved earlier with $\psi_k(\mathcal{E}_{nk})$ replaced by $\psi_k(\mathcal{E}_k)$), now using Corollary 15 of Jeganathan (1995). (Note that the proof in Jeganathan, 1997, consisted of verifying the conditions of the preceding lemma under both the null and the alternatives and then using Theorem 13 in Jeganathan, 1995, instead of the present use of the indicated weak convergence results.)

Thus it remains to show that $\psi_k(\mathcal{E}_k)$ involved previously can be replaced by $\psi_k(\mathcal{E}_{nk})$ in the quantities used in the approximation. This will follow in exactly the same way as in Jeganathan (1997, Lemma d, Sect. 5) provided, in the present context, the requirement

$$\sum_{k=k_n+1}^{n} E\left[\left|\sum_{j=k-1}^{\infty} d_j U_{k-j}\right|^2\right] \to 0$$

is satisfied. (Note that this is the condition used in Jeganathan, 1997, except that the sum $\sum_{j=k_n}^{\infty}$ was involved inside the square instead of $\sum_{j=k-1}^{\infty}$ in view of the fact that $\mathcal{E}_{nk}$ involved the existence of $\sum_{j=0}^{k_n} d_j U_{k-j}$.) This requirement is true because, as was seen in the preceding proof of statement (b) of Theorem 4, $E[|\sum_{j=k-1}^{\infty} d_j \times$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

$U_{k-j}|^2]$ is bounded by a constant multiple of $\int_{-\pi}^{\pi}(\sum_{j=k-1}^{\infty} d_j \exp(-iju))^2 f(u)\,du$ that is of order $O(k^{-2(1-\alpha)})$ with $\alpha < \frac{1}{2}$. This completes the proof. ■

Proof of Proposition 8. It only remains to find the limits of the variance and covariance functions, in view of the arguments of the proof of Proposition 2, from which the convergences of the variances also follow; note that when $0 < \phi_l \leq \pi$, $\rho_{n,l}(t)$ can be expressed in terms of the processes dealt with in the proof of statement (b) of Proposition 2. When $\phi_l = \phi_m$, the same arguments give the convergence of the covariance between $\rho_l(t)$ and $\rho_m(t)$. Therefore, it remains to check that when $\phi_l \neq \phi_m$, the limit of the covariance between $\rho_l(t)$ and $\rho_m(t)$ is zero. First consider the case $0 < \phi_l \neq \phi_m \leq \pi$. Then, writing $f(\lambda)$ in the form $|b(\lambda)|^2$, it is enough to check that

$$n^{-1}\int_{-\pi}^{\pi}\left(\sum_{k=1}^{n}\exp(ik(\lambda+\phi_l))\right)\left(\sum_{k=1}^{n}\exp(-ik(\lambda+\phi_m))\right)|b(\lambda)|^2\,d\lambda \to 0. \quad \textbf{(51)}$$

The step involved in (41) shows that

$$n^{-1}\int_{-\pi}^{\pi}\left|\sum_{k=1}^{n}\exp(ik(\lambda+\phi_l))\right|^2 (b(\lambda)-b(\phi_l))^2\,d\lambda \to 0, \quad \textbf{(52)}$$

and the same holds, involving the complex conjugate $\bar{b}(\lambda)$, when ϕ_l is replaced by ϕ_m. Hence an application of Cauchy–Schwarz inequality shows that the integral in (51) is approximately a constant multiple of, with $\phi^* = \phi_l + \phi_m$,

$$\begin{aligned} &n^{-1}\int_{-\pi}^{\pi}\left(\sum_{k=1}^{n}\exp(ik(\lambda+\phi_l))\right)\left(\sum_{k=1}^{n}\exp(-ik(\lambda+\phi_m))\right)d\lambda \\ &\quad = n^{-1}\sum_{k=1}^{n}\exp(ik\phi^*) \\ &\quad = \frac{\exp(i\phi^*)(\exp(in\phi^*)-1)}{n(\exp(i\phi^*)-1)} \\ &\quad \to 0, \end{aligned}$$

because $\exp(i\phi^*) \neq 1$. Next consider the case $\phi_m = 0$ and $0 < \phi_l \leq \pi$. Similar to the preceding arguments and in view of (52), it is enough to show that

$$\frac{1}{n^{1/2+H}}\int_{-\pi}^{\pi}\left(\sum_{k=1}^{n}\exp(ik(\lambda+\phi_l))\right)\left(\sum_{k=1}^{n}\exp(-ik\lambda)\right)b^*(\lambda)\,d\lambda \to 0, \quad \textbf{(53)}$$

where $b^*(\lambda)$ is the real part of $b(\lambda) = (1-\exp(i\lambda))^{-\alpha}$. But the integral involved reduces to

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

$$\begin{aligned}
&\frac{1}{n^{1/2+H}}\sum_{j=1}^{n}\exp(ij\phi_l)\sum_{k=1}^{j}c_k \\
&\quad=\frac{1}{n^{1/2+H}}\sum_{k=1}^{n}c_k\sum_{j=k}^{n}\exp(ij\phi_l) \\
&\quad=\frac{1}{n^{1/2+H}}\sum_{k=1}^{n}c_k\frac{[\exp(i(n+1)\phi_l)-\exp(ik\phi_l)]}{(\exp(i\phi_l)-1)} \\
&\quad=\frac{1}{n^{1/2+H}(\exp(i\phi_l)-1)}\left[\exp(i(n+1)\phi_l)\sum_{k=1}^{n}c_k-\sum_{k=1}^{n}c_k\exp(ik\phi_l)\right] \\
&\quad\to 0
\end{aligned}$$

because $\exp(i\phi_l)\neq 1$. This completes the proof. ■

Proof of Proposition 9. The required arguments are contained in the proof of Proposition 3. Note that when $0<\phi_l\leq\pi$, the $\tilde{\rho}_{n,l}$'s can be expressed in terms of the sums considered in the proof of statement (b) of Proposition 3. ■

Proof of Proposition 10. In view of the joint convergence in Propositions 8 and 9, it would be sufficient for the most part to handle $\hat{V}_{n,l}(t)$'s individually. In view of (32) one has $\widetilde{V}_{l,k}=\Gamma_l\widetilde{V}_{l,k-1}+\tilde{\xi}_{l,k}$ with $\tilde{\xi}_{l,k}=\sum_{j=0}^{k-1}d_j\xi_{l,k-j}$. Solving this one has

$$\widetilde{V}_{l,k}=\sum_{j=0}^{k}\Gamma_l^j\tilde{\xi}_{l,k-j}. \tag{54}$$

The arguments involved in handling this parallel those of the corresponding situation (the case $\alpha=0$) treated in Jeganathan (1997, Sect. 4.2, step III).

First consider Case $\phi_l=0$. In this case the ath row and the b-column element of Γ_l^j are given by

$$\mathcal{C}_{b-a}^{j}=\frac{j(j-1)\ldots(j-b+a+1)}{(1)(2)\ldots(b-a)},$$

$a,b=1,\ldots,r_l$, where r_l is the order of Γ_l. First note that, with $\tilde{\xi}_{l,k}^{*}$ the last element of $\tilde{\xi}_{l,k}$ and with $\tilde{\rho}_{n,l}(t)$ as in Proposition 9, $n^{-1/2}\sum_{j=0}^{[nt]}\tilde{\xi}_{l,[nt]-j}^{*}=\tilde{\rho}_{n,l}(t)$. Hence

$$\begin{aligned}
&N_l\sum_{0}^{[nt]}(\mathcal{C}_{r_l-1}^{j},\ldots,\mathcal{C}_{0}^{j})'\tilde{\xi}_{[nt]-j,l}^{*} \\
&\quad=\operatorname{diag}\left\{\frac{1}{n^{r_l-1}},\ldots,\frac{1}{n^{1}},1\right\}\sum_{0}^{[nt]}(\mathcal{C}_{r_l-1}^{j},\ldots,\mathcal{C}_{0}^{j})' \\
&\qquad\times\left(\tilde{\rho}_{n,l}\left(\frac{j}{n}\right)-\tilde{\rho}_{n,l}\left(\frac{j-1}{n}\right)\right),
\end{aligned} \tag{55}$$

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

which is directly shown to converge in law to the process $S_l(t)$. Now note that this sum is, apart from the normalizing N_l, just that part of (54) formed by the last columns of Γ_l^j's, which dominates the part corresponding to the remaining columns in the sense that the difference between $\hat{V}_{n,l}(t) = N_l \widetilde{V}_{l,[nt]}$ and (55) converges to zero in probability.

The remaining cases corresponding to $0 < \phi_l \leq \pi$ are handled similarly in analogy with the treatment in Jeganathan (1997). ■

Proof of Proposition 11. Only the convergences of the covariances between the processes $\mathcal{X}_{n,l}(t)$ and $\rho_{n,l}(t)$ to the indicated limits need to be verified, but this is contained in the proof of Proposition 8. ■

Proof of Proposition 13. The marginal convergence in law of (24) is just a multivariate extension of statement (b) of Theorem 4, the proof of the corresponding statement (a) of which further essentially contains the arguments for the required joint convergence and the independence requirement of the limits. ■

REFERENCES

Baillie, R.T. (1996) Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5–59.

Chan, N.H. & N. Terrin (1995) Inference for unstable long-memory processes with applications to fractional unit root autoregressions. *Annals of Statistics* 23, 1662–1683.

Chan, N.H. & C.Z. Wei (1988) Limiting distributions of least squares estimates of unstable autoregressive process. *Annals of Statistics* 16, 367–401.

Cheung, Y.-W. & K.S. Lai (1993) A fractional integration analysis of purchasing power parity. *Journal of Business and Economic Statistics* 11, 103–112.

Chung, C.-F. (1995) Sample Variances, Sample Covariances, and Linear Regression of Stationary Multivariate Long Memory Processes. Technical report, Michigan State University.

Dahlhaus, R. (1989) Efficient parameter estimation for self similar processes. *Annals of Statistics* 17, 1749–1766.

Davydov, Y.A. (1970) The invariance principle for stationary processes. *Theory of Probability and Its Applications* 15, 487–498.

Dolado, J.J. & F. Marmol (1996) Efficient Estimation of Cointegrating Relationships among Higher Order and Fractionally Integrated Process. Documento de Trabafo 9617, Banco de Espana—Servicio de Estudios.

Gorodetskii, V.V. (1977) On convergence to semi-stable Gaussian processes. *Theory of Probability and Its Applications* 22, 498–508.

Granger, C.W.J. & R. Joyeux (1980) An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis* 1, 15–39.

Hosking, J.R.M. (1981) Fractional differencing. *Biometrika* 68, 165–176.

Jeganathan, P. (1991) On the asymptotic behavior of least square estimators in AR time series with roots near the unit circle. *Econometric Theory* 7, 269–306.

Jeganathan, P. (1995) Some aspects of asymptotic theory with applications to time series models. *Econometric Theory* 11, 818–887.

Jeganathan, P. (1997) On asymptotic inference in linear cointegrated time series systems. *Econometric Theory* 13, 692–745.

Johansen, S. (1995) *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

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

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press

LeCam, L. & G.L. Yang (1990) *Asymptotics in Statistics: Some Basic Concepts*. New York: Springer-Verlag.
Lee, H. (1996) *Analysis of Economic Time Series with Long Memory*. Ph.D Dissertation, Michigan State University.
Mandelbrot, B.B. & J.W. van Ness (1968) Fractional Brownian motions, fractional Brownian noises and applications. *SIAM Review* 10, 422–437.
Phillips, P.C.B. (1991) Optimal inference in cointegrated systems. *Econometrica* 59, 283–306.
Pötscher, B.M. (1991) Noninvertibility and pseudo-maximum likelihood estimation in misspecified ARMA models. *Econometric Theory* 7, 435–449.
Reinsel, G.C. (1993) *Elements of Multivariate Time Series Analysis*. New York: Springer-Verlag.
Robinson, P.M. (1994a) Time series with strong dependence. In C.A. Sims (ed.), *Advances in Econometrics: Sixth World Congress*, vol. I. New York: Cambridge University Press.
Robinson, P.M. (1994b) Time Series Regression with Long Range Dependence. Technical report, London School of Economics.
Samorodnitsky, G. & M.S. Taqqu (1994) *Stable Non-Gaussian Random Processes*. New York: Chapman and Hall.
Shilov, E.G. (1977) *Linear Algebra*. New York: Dover. (Translated from the Russian).
Sowell, F.B. (1990) The fractional unit root distribution. *Econometrica* 58, 495–505.
Stock, J.H. (1994) Unit roots, structural breaks and trends. In R.F. Engle & D.L. McFadden (eds.), *Handbook of Econometrics*, vol. IV. Amsterdam: Elsevier Science B.V.
Tanaka, K. (1996) *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*. New York: Wiley.
Taqqu, M.S. (1975) Weak convergence to fractional Brownian motion and to Rosenblatt process. *Zeitschrift für wahrscheinlichkeitstheorie und Verwandte Gebiete* 31, 287–302.
Toda, H.Y. & P.C.B. Phillips (1993) Vector autoregressions and causality. *Econometrica* 61, 1367–1393.
Tsay, R.S. & G.C. Tiao (1991) Asymptotic properties of multivariate nonstationary processes with applications to autoregressions. *Annals of Statistics* 18, 220–250.
Watson, M.W. (1994) Vector autoregression and cointegration. In R.F. Engle & D.L. McFadden (eds.), *Handbook of Econometrics*, vol. IV. Amsterdam: Elsevier Science B.V.
Zygmund, A. (1968) *Trigonometric Series*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S0266466699154057 Published online by Cambridge University Press