Limit Theory for Cointegrated Systems with Moderately Integrated and Moderately Explosive Regressors[1]

Tassos Magdalinos
University of Nottingham, UK

Peter C. B. Phillips
Cowles Foundation for Research in Economics
Yale University
and
University of Auckland & University of York, UK

July 28, 2007

[1]A first draft of the paper was written in April, 2004. Phillips acknowledges partial research support from a Kelly Fellowship and the NSF under Grant Nos. SES-04-142254 and SES 06-47086. Magdalinos thanks the EPSRC and the Onassis Foundation for scholarship support.

Abstract

An asymptotic theory is developed for multivariate regression in cointegrated systems whose variables are moderately integrated or moderately explosive in the sense that they have autoregressive roots of the form $\rho_{ni} = 1 + c_i/n^{\alpha}$, involving moderate deviations from unity when $\alpha \in (0, 1)$ and $c_i \in \mathbb{R}$ are constant parameters. When the data are moderately integrated in the stationary direction (with $c_i < 0$), it is shown that least squares regression is consistent and asymptotically normal but suffers from significant bias, related to simultaneous equations bias. In the moderately explosive case (where $c_i > 0$) the limit theory is mixed normal with Cauchy-type tail behavior and the rate of convergence is explosive, as in the case of a moderately explosive scalar autoregression (Phillips and Magdalinos, 2007a). Moreover, the limit theory applies without any distributional assumptions and for weakly dependent errors under conventional moment conditions, so an invariance principle holds, unlike the well-known case of an explosive autoregression. This theory validates inference in cointegrating regression with mildly explosive regressors. The special case in which the regressors themselves have a common explosive component is also considered.

Keywords: Central limit theory, Cointegration, Diffusion, Explosive process, Invariance principle, Mixed normality, Moderate deviations, Unit root distribution, Weak dependence.

AMS 1991 subject classification: 62M10; JEL classification: C22

1. Introduction

The limit theory for cointegrated regressions has been fully worked out for integrated and near-integrated processes (Phillips, 1988a; Elliott, 1998). In the integrated case, optimal estimates can be obtained by a variety of system methods (Johansen, 1988; Phillips, 1991a and 1991b; Phillips and Hansen, 1990) and single equation (one-step) techniques (Phillips and Loretan, 1991; Saikkonen, 1991; Stock and Watson, 1993) and are known to have mixed normal limit distributions which depend on the long-run covariance matrix structure of the equation errors and first differences of the regressors. The limit distribution theory has also been worked out for cases of cointegrated regressors and instrumental variable methods (Phillips, 1995; Kitamura and Phillips, 1995 and 1997).

As is often emphasized in empirical work, economic time series seem to have autoregressive roots in the general neighborhood of unity and insistence that roots be at unity may well be too harsh a requirement. Accordingly, attention has been given to the case where the roots are local to unity in the sense that they have the form $\rho = 1 + c/n$, where n is the sample size (Phillips, 1987; Chan and Wei, 1987). Matrix cases, where the long-run autoregressive coefficient matrix has the form $R_n = I + C/n$ have been considered in Phillips (1988a, 1988b). This theory has been useful in developing power functions for testing problems in cointegrated regressions (Phillips, 1988a; Johansen, 1995) and in other local analyses (Elliot, 1998).

To characterize greater deviations from unity Phillips and Magdalinos (2007a & 2007b; hereafter PM_a and PM_b) recently investigated time series with an autoregressive root of the form $\rho_n = 1 + c/n^\alpha$, where the exponent α lies in the interval $(0, 1)$. Such roots represent moderate deviations from unity in the sense that they belong to larger neighborhoods of one than conventional local to unity roots. The parameter α measures the radial width of the local neighborhood, with smaller values of α being associated with larger neighborhoods. The boundary value as $\alpha \to 1$ includes the conventional local to unity case, whereas the boundary value as $\alpha \to 0$ includes the stationary or explosive AR(1) process, depending on the sign of c. In this paper, we call such time series *moderately integrated* or *mildly integrated* as distinct from near integrated (when $\alpha = 1$).

The present work shows how these ideas can be extended to cointegrated regression systems where the variables belong to a general class of moderately integrated time series, whose long run autoregressive coefficients deviate moderately from unity and are of the form $\rho_n = 1 + c/n^\alpha$ for some $\alpha \in (0, 1)$. We consider separately the moderately integrated, near stationary ($c < 0$) case and the moderately explosive ($c > 0$) case.

The main goal of the paper is to provide a framework of limit theory for such cointegrating regressions. Not all problems of inference in these systems are resolved, particularly in the case of mildly integrated regressors in stationary directions. However, some notable results are obtained, including a complete set of invariance principles in the moderately explosive case, including some important degenerate cases. These results provide an asymptotic validation of inference for least squares regression with moderately explosive and possibly cointegrated regressors and extend the theory for purely explosive cointegrated systems developed earlier in Phillips and Magdalinos (2007c).

The paper is organised as follows. Section 2 outlines a general modeling framework for a cointegrated system with moderately integrated regressors and weakly dependent innovation errors. Section 3 provides a limit theory in the near stationary case for $\alpha \in (0,1)$, where the rates of convergence for cointegrating regression estimates are now of order $O(n^{\frac{1+\alpha}{2}})$ instead of $O(n)$. The limit distribution theory is normal rather than mixed normal as in the case of integrated or near integrated regressors, and there are bias effects from serial dependence and endogeneity. Section 4 analyzes the limit theory in the near explosive case, showing that least squares regression is mixed normal even under non-Gaussian errors and regressors, so that an invariance principle applies. The case of cointegrated moderately explosive regressors is also treated and gives rise to a different mixed normal limiting distribution with asymptotic bias and slower rate of convergence. Section 5 provides some further discussion of the results and Section 6 is a notational glossary. Proofs and technical propositions are collected in Section 7.

2. Moderately Integrated Time Series and Cointegrated Systems

We consider the triangular system (cf. Phillips, 1991)

$$y_t = Ax_t + u_{0t} \tag{1}$$

$$x_t = R_n x_{t-1} + u_{xt} \tag{2}$$

$$R_n = I_K + \frac{C}{n^{\alpha}}, \ \alpha \in (0,1), \quad C = \mathrm{diag}(c_1, ..., c_K) \tag{3}$$

for $t = 1, ..., n$, where A is an $m \times K$ matrix of 'cointegrating' coefficients, x_t is a $K-$ vector of moderately integrated time series and the system is initialized at some $x_0 = o_p\left(n^{\alpha/2}\right)$. The vector $u_t = (u_{0t}', u_{xt}')'$ is a sequence of zero mean, weakly dependent errors that satisfy the following condition.

Assumption LP. *For each $t \in \mathbb{N}$, u_t has Wold representation*

$$u_t = F(L)\varepsilon_t = \sum_{j=0}^{\infty} F_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} j \|F_j\| < \infty, \tag{4}$$

where $F(z) = \sum_{j=0}^{\infty} F_j z^j$, $F_0 = I_{m+K}$, $F(1)$ has full rank,

$$\|M\| = \max_i \left\{ \lambda_i^{1/2} : \lambda_i \textit{ is an eigenvalue of } M'M \right\}$$

is the spectral norm of M, and $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a sequence of independent and identically distributed $(0,\Sigma_\varepsilon)$ random vectors satisfying $\Sigma_\varepsilon > 0$ and the moment condition $E\|\varepsilon_1\|^4 < \infty$ when $C < 0$ and $E\|\varepsilon_1\|^2 < \infty$ when $C > 0$.

Under **LP**, u_t has variance matrix $\Sigma = \sum_{j=0}^{\infty} F_j \Sigma_\varepsilon F_j'$, $E\|u_1\|^4 < \infty$ when $C < 0$, $E\|u_1\|^2 < \infty$ when $C > 0$ and partial sums that satisfy the central theorem (cf. Phillips and Solo, 1992)

$$n^{-1/2}\sum_{i=1}^{n} u_i \Rightarrow N(0,\Omega),$$

with variance matrix $\Omega = F(1)\Sigma_\varepsilon F(1)' > 0$. We partition Σ, $F(1)$ and Ω conformably with u_t as

$$\Sigma = \begin{bmatrix} \Sigma_{00} & \Sigma_{0x} \\ \Sigma_{x0} & \Sigma_{xx} \end{bmatrix}, \quad F(1) = \begin{bmatrix} F_0(1) \\ F_x(1) \end{bmatrix}, \quad \Omega = \begin{bmatrix} \Omega_{00} & \Omega_{0x} \\ \Omega_{x0} & \Omega_{xx} \end{bmatrix},$$

where $F_0(1)$ and $F_x(1)$ are $m\times(m+K)$ and $K\times(m+K)$ matrices respectively. Using the Beveridge Nelson (BN) decomposition, we obtain the following representation for u_t

$$u_t = F(1)\varepsilon_t - \Delta\tilde{\varepsilon}_t, \quad \text{for } \tilde{\varepsilon}_t = \sum_{j=0}^{\infty}\tilde{F}_j\varepsilon_{t-j},\ \tilde{F}_j = \sum_{k=j+1}^{\infty}F_k, \tag{5}$$

where $\sum_{j=0}^{\infty}\left\|\tilde{F}_j\right\| < \infty$ is assured by the summability condition in (4). The derivation of (5) as well as the summability of the sequence $\left(\tilde{F}_j\right)_{j\geq 0}$ follow as in Lemma 2.1 of Phillips and Solo (1992). Corresponding to the partition of u_t, we write

$$\begin{aligned} u_{0t} &= \sum_{j=0}^{\infty}F_{0j}\varepsilon_{t-j}, \quad u_{xt} = \sum_{j=0}^{\infty}F_{xj}\varepsilon_{t-j}, \\ \tilde{\varepsilon}_{0t} &= \sum_{j=0}^{\infty}\tilde{F}_{0j}\varepsilon_{t-j}, \quad \tilde{\varepsilon}_{xt} = \sum_{j=0}^{\infty}\tilde{F}_{xj}\varepsilon_{t-j}, \end{aligned}$$

where $F_{0j}, \tilde{F}_{0j} \in \mathbb{R}^{m\times(m+K)}$ and $F_{xj}, \tilde{F}_{xj} \in \mathbb{R}^{K\times(m+K)}$. Defining the one-sided long run covariance matrix

$$\Lambda := \sum_{h=1}^{\infty}E\left(u_t u_{t-h}'\right) = \begin{bmatrix} \Lambda_{00} & \Lambda_{0x} \\ \Lambda_{x0} & \Lambda_{xx} \end{bmatrix},$$

we have $\Omega = \Sigma + \Lambda + \Lambda'$ and, as usual,

$$\begin{aligned} \Lambda &= \sum_{h=1}^{\infty}\sum_{k=0}^{\infty}F_{k+h}\Sigma_\varepsilon F_k' = \sum_{k=0}^{\infty}\left(\sum_{h=1}^{\infty}F_{k+h}\right)\Sigma_\varepsilon F_k' \\ &= \sum_{k=0}^{\infty}\tilde{F}_k\Sigma_\varepsilon F_k' = E\left(\tilde{\varepsilon}_t u_t'\right). \end{aligned}$$

3. Limit Theory for Near Stationary Systems

We develop a limit theory for the centred least squares regression estimate

$$\hat{A}_n - A = \left(\sum_{t=1}^{n} u_{0t} x_t\right) \left(\sum_{t=1}^{n} x_t x_t'\right)^{-1}. \quad (6)$$

Our approach follows PM_a in the sense that we derive a law of large numbers and a martingale central limit theorem, respectively, for the denominator and numerator of the matrix quotient (6) and use this to extract the limit theory.

We start by considering the sample moment matrix $\sum_{t=1}^{n} x_t x_t'$. The following result allows an analysis of the asymptotic behavior of $\sum_{t=1}^{n} x_t x_t'$ analogous to the stationary case.

3.1 Lemma. *For model (2) with $c_i < 0$ for all i we have, as $n \to \infty$,*

(a) $x_n x_n' = O_p(n^{\alpha})$,

(b) $\tilde{\varepsilon}_n x_n' = O_p\left(n^{\alpha/2}\right)$,

(c) $\sum_{t=1}^{n} \tilde{\varepsilon}_t x_{t-1}' = O_p\left(n^{1+\frac{\alpha}{2}}\right)$

(d) $\frac{1}{n} \sum_{t=1}^{n} u_{xt} x_{t-1}' \to_p \Lambda_{xx}$.

Denote by $\mathcal{K}_s$ the $s^2 \times s^2$ commutation matrix (e.g., Abadir and Magnus, 2005). Expanding the expression for $x_t x_t'$ in (2), vectorising and summing over $t \in \{1, ..., n\}$ we obtain

$$\begin{aligned}
&[I_{K^2} - R_n \otimes R_n] \frac{1}{n} \sum_{t=1}^{n} vec\left(x_{t-1} x_{t-1}'\right) \\
= \; & \frac{1}{n} \left[vec\left(x_n x_n'\right) - R_n \otimes R_n vec\left(x_0 x_0'\right)\right] \\
& + \left(I_{K^2} + \mathcal{K}_K\right)\left(I_K \otimes R_n\right) \frac{1}{n} \sum_{t=1}^{n} vec\left(x_{t-1} u_{xt}'\right) + \frac{1}{n} \sum_{t=1}^{n} vec\left(u_{xt} u_{xt}'\right)
\end{aligned}$$

$$\begin{aligned}
= \; & \left(I_{K^2} + \mathcal{K}_K\right)\left(I_K \otimes R_n\right) \frac{1}{n} \sum_{t=1}^{n} vec\left(x_{t-1} u_{xt}'\right) + \frac{1}{n} \sum_{t=1}^{n} vec\left(u_{xt} u_{xt}'\right) + O_p\left(\frac{1}{n^{1-\alpha}}\right) \\
= \; & \left(I_{K^2} + \mathcal{K}_K\right) vec\left(\Lambda_{xx}\right) + vec\left(\Sigma_{xx}\right) + o_p(1) \\
= \; & vec\left(\Lambda_{xx} + \Lambda_{xx}' + \Sigma_{xx}\right) + o_p(1) \\
= \; & vec\left(\Omega_{xx}\right) + o_p(1),
\end{aligned}$$

as $n \to \infty$ by Lemma 3.1 and ergodicity of u_{xt}. Since

$$I_{K^2} - R_n \otimes R_n = -\frac{1}{n^\alpha}\left(I_K \otimes C + C \otimes I_K + \frac{C \otimes C}{n^\alpha}\right),$$

and $I_K \otimes C + C \otimes I_K$ is a negative definite matrix, the asymptotic behavior of the sample moment matrix is given by

$$\frac{1}{n^{1+\alpha}}\sum_{t=1}^{n} vec\left(x_{t-1}x_{t-1}'\right) = -\left(I_K \otimes C + C \otimes I_K\right)^{-1} vec\left(\Omega_{xx}\right) + o_p(1),$$

or

$$\frac{1}{n^{1+\alpha}}\sum_{t=1}^{n} x_{t-1}x_{t-1}' \to_p V_{xx} := \int_0^\infty e^{rC}\Omega_{xx}e^{rC}dr, \tag{7}$$

as $n \to \infty$. Note that V_{xx} is a $K \times K$ symmetric matrix with typical element

$$\left(V_{xx}\right)_{ij} = -\left(c_i + c_j\right)^{-1}\left(\Omega_{xx}\right)_{ij}, \tag{8}$$

so, in the scalar case, the limit expression in (7) reduces to $\frac{\omega^2}{-2c}$, the result obtained in PM$_b$.

The limit distribution of a suitably standardized version of the sample covariance $\sum_{t=1}^{n} u_{0t}x_t'$ is found by expanding this covariance in terms of components whose asymptotic behavior can be found directly, such as $\sum_{t=1}^{n}\varepsilon_t x_{t-1}'$ (see (9) below). The following results help to analyze these components and are proved in the Appendix.

3.2 Lemma. *For $\nu \in \{0, x\}$, define $\Gamma_{\nu x}^{\tilde{\varepsilon}}(k) = E\left(\tilde{\varepsilon}_{\nu t}\tilde{\varepsilon}_{xt-k}'\right)$ and*

$$M_{\nu n} = \Gamma_{\nu x}^{\tilde{\varepsilon}}(0) + \sum_{j=0}^{n-1}\tilde{F}_{\nu j+1}\Sigma_\varepsilon F_x(1)' R_n^j + \frac{1}{n^\alpha}\sum_{j=0}^{n-1}\Gamma_{\nu x}^{\tilde{\varepsilon}}(j+1) R_n^j C.$$

When $C < 0$ we obtain, for each $\alpha \in (0,1)$ and $\nu \in \{0, x\}$

(a) $n^{-\frac{1+\alpha}{2}}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_t u_t' - \Lambda\right) = O_p\left(n^{-\alpha/2}\right)$,

(b) $n^{-\frac{1+3\alpha}{2}}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_{\nu t}x_{t-1}' - M_{\nu n}\right) = O_p\left(n^{-\frac{\alpha\wedge(1-\alpha)}{2}}\right)$.

3.3 Lemma. *For each $\alpha \in (0,1)$ and $C < 0$*

$$\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\left(x_{t-1} \otimes \varepsilon_t\right) \Rightarrow N\left(0, V_{xx} \otimes \Sigma_\varepsilon\right) \quad \textit{as } n \to \infty.$$

Next, using the BN decomposition (5) and summation by parts, the sample covariance can be decomposed as follows:

$$\begin{aligned}
\sum_{t=1}^{n} u_{0t}x_t' &= \sum_{t=1}^{n} u_{0t}x_{t-1}'R_n + \sum_{t=1}^{n} u_{0t}u_{xt}' \\
&= F_0(1)\sum_{t=1}^{n} \varepsilon_t x_{t-1}'R_n - \sum_{t=1}^{n} \Delta\tilde{\varepsilon}_{0t}x_{t-1}'R_n + \sum_{t=1}^{n} u_{0t}u_{xt}' \\
&= F_0(1)\sum_{t=1}^{n} \varepsilon_t x_{t-1}'R_n + \sum_{t=1}^{n} \tilde{\varepsilon}_{0t}\Delta x_t' R_n + \sum_{t=1}^{n} u_{0t}u_{xt}' + O_p\left(n^{\alpha/2}\right) \\
&= F_0(1)\sum_{t=1}^{n} \varepsilon_t x_{t-1}'R_n + \left[\sum_{t=1}^{n} \tilde{\varepsilon}_{0t}u_{xt}' + \sum_{t=1}^{n} \tilde{\varepsilon}_{0t}x_{t-1}'\left(R_n - I_K\right)\right] R_n \\
&\quad + \sum_{t=1}^{n} u_{0t}u_{xt}' + O_p\left(n^{\alpha/2}\right) \\
&= F_0(1)\sum_{t=1}^{n} \varepsilon_t x_{t-1}'R_n + \sum_{t=1}^{n} \tilde{\varepsilon}_{0t}u_{xt}'R_n + \frac{1}{n^{\alpha}}\sum_{t=1}^{n} \tilde{\varepsilon}_{0t}x_{t-1}'CR_n \\
&\quad + \sum_{t=1}^{n} u_{0t}u_{xt}' + O_p\left(n^{\alpha/2}\right). \qquad (9)
\end{aligned}$$

The leading terms in the above expression for the sample covariance are the matrices $\sum_{t=1}^{n} \tilde{\varepsilon}_{0t}u_{xt}'$ and $\sum_{t=1}^{n} u_{0t}u_{xt}'$ with asymptotic order $O_{a.s.}(n)$ given by the ergodic theorem. Thus, if no correction is made to account for weak dependence, the sample covariance will converge to the constant probability limit of the leading term as follows

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n} u_{0t}x_t' &= \frac{1}{n}\sum_{t=1}^{n} \tilde{\varepsilon}_{0t}u_{xt}' + \frac{1}{n}\sum_{t=1}^{n} u_{0t}u_{xt}' + O_p\left(n^{-\frac{(1-\alpha)\wedge\alpha}{2}}\right) \\
&= \Lambda_{0x} + \Sigma_{0x} + o_p(1), \qquad (10)
\end{aligned}$$

by ergodicity of $u_{0t}u_{xt}'$ and $\tilde{\varepsilon}_{0t}u_{xt}'$. The above, together with (7), implies that for each $\alpha \in (0,1)$

$$\begin{aligned}
n^{\alpha}\left(\hat{A}_n - A\right) &= \left(\frac{1}{n}\sum_{t=1}^{n} u_{0t}x_t'\right)\left(\frac{1}{n^{1+\alpha}}\sum_{t=1}^{n} x_t x_t'\right)^{-1} \\
&\to_p \left(\Lambda_{0x} + \Sigma_{0x}\right) V_{xx}^{-1}. \qquad (11)
\end{aligned}$$

Obtaining a non degenerate weak limit for the sample covariance requires centering on the mean of the leading terms. Then, for each $\alpha \in (0,1)$ (9) gives, up to $o_p(1)$,

$$
\begin{aligned}
& \frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n} \left[u_{0t} x_t' - \Lambda_{0x} - \Sigma_{0x} - \frac{1}{n^{\alpha}} \left(\Lambda_{0x} C + M_{0n} C R_n \right) \right] \\
= \; & \frac{F_0(1)}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n} \varepsilon_t x_{t-1}' R_n + \frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n} \left(\tilde{\varepsilon}_{0t} u_{xt}' - \Lambda_{0x} \right) R_n \\
& + \frac{1}{n^{\frac{1+3\alpha}{2}}} \sum_{t=1}^{n} \left(\tilde{\varepsilon}_{0t} x_{t-1}' - M_{0n} \right) C R_n + \frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n} \left(u_{0t} u_{xt}' - \Sigma_{0x} \right) + O_p\left(n^{-1/2} \right),
\end{aligned}
$$

the last three terms being asymptotically negligible by Lemma 3.2 and the CLT for stationary ergodic processes. Thus, defining

$$
\begin{aligned}
U_{0x}^{(n)} &= \Lambda_{0x} + \Sigma_{0x} + \frac{1}{n^{\alpha}} \left(\Lambda_{0x} C + M_{0n} C R_n \right) \\
&= \Lambda_{0x} + \Sigma_{0x} + \frac{1}{n^{\alpha}} \left\{ \Lambda_{0x} + \Gamma_{0x}^{\tilde{\varepsilon}}(0) + \sum_{j=0}^{n-1} \tilde{F}_{0j+1} \Sigma_{\varepsilon} F_x(1)' R_n^{j+1} \right\} C \\
&\quad + \frac{1}{n^{2\alpha}} \sum_{j=0}^{n} \Gamma_{0x}^{\tilde{\varepsilon}}(j) R_n^j C^2,
\end{aligned}
$$

the sample covariance becomes

$$
\begin{aligned}
\frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n} vec\left[u_{0t} x_t' - U_{0x}^{(n)} \right] &= \left[I_n \otimes F_0(1) \right] \frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n} \left(x_{t-1} \otimes \varepsilon_t \right) + o_p(1) \\
&\Rightarrow N\left(0, V_{xx} \otimes \Omega_{00} \right). \qquad (12)
\end{aligned}
$$

As in PM$_b$, the weak dependence structure of the innovations induces an asymptotic bias for the least squares estimator $\hat{A}_n$. Explicit calculation of both the bias and the asymptotic distribution of $\hat{A}_n$ involves analysis of the limiting distribution of the denominator, $\sum_{t=1}^{n} x_{t-1} x_{t-1}'$ of $\hat{A}_n$ centered on its asymptotic mean. The latter can be obtained as an approximation of the centered sample covariance $\sum_{t=1}^{n} u_{xt} x_{t-1}'$ as we now show. Define the $K \times K$ symmetric matrices

$$
\Omega_{xx}^{(n)} := \Omega_{xx} + \frac{1}{n^{\alpha}} \left(\Lambda_{xx} C + C \Lambda_{xx}' + R_n C M_{xn}' + M_{xn} C R_n \right), \qquad (13)
$$

and $V_{xx}^{(n)}$ with typical element

$$
\left(V_{xx}^{(n)} \right)_{ij} = -\left(c_i + c_j \rho_{in} \right)^{-1} \left(\Omega_{xx}^{(n)} \right)_{ij}, \qquad (14)
$$

where $\rho_{in} = 1 + c_i / n^{\alpha}$ is the i-th diagonal element of R_n. Clearly, $\Omega_{xx}^{(n)} \to \Omega_{xx}$ and $V_{xx}^{(n)} \to V_{xx}$ as $n \to \infty$.

3.4 Lemma. *For each* $\alpha \in (0,1)$ *and* $C = \mathrm{diag}(c_i)$ with $c_i < 0$,

$$\begin{aligned}\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n} vec\left(u_{xt}x_{t-1}' - \Lambda_{xx} - \frac{1}{n^{\alpha}}M_{xn}C\right) &= \frac{[I_K \otimes F_x(1)]}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}(x_{t-1}\otimes\varepsilon_t) + o_p(1) \\ &\Rightarrow N(0, V_{xx}\otimes\Omega_{xx}).\end{aligned}$$

3.5 Lemma. *For each* $\alpha \in (0,1)$ *and* $C = \mathrm{diag}(c_i)$ *with* $c_i < 0$,

$$\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n} vec\left(\frac{x_{t-1}x_{t-1}'}{n^{\alpha}} - V_{xx}^{(n)}\right) = W_x\frac{[I_K \otimes F_x(1)]}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}(x_{t-1}\otimes\varepsilon_t) + o_p(1),$$

where $W_x := -(I_K \otimes C + C \otimes I_K)^{-1}(I_{K^2} + \mathcal{K}_K)$ *and* $V_{xx}^{(n)}$ *is defined in (14).*

The limiting distribution of $\hat{A}_n$ can now be derived by using (12) and Lemma 3.5. In particular, letting

$$\begin{aligned}\bar{A}_n &= A + \frac{1}{n^{\alpha}}U_{0x}^{(n)}\left[V_{xx}^{(n)}\right]^{-1} \\ &= A + \frac{1}{n^{\alpha}}(\Lambda_{0x} + \Sigma_{0x})\left[V_{xx}^{(n)}\right]^{-1} + \frac{1}{n^{2\alpha}}(\Lambda_{0x}C + M_{0n}CR_n)\left[V_{xx}^{(n)}\right]^{-1}, \quad (15)\end{aligned}$$

and centering the least squares estimator by $\bar{A}_n$ yields

$$\begin{aligned}\hat{A}_n - \bar{A}_n &= \left(\sum_{t=1}^{n}u_{0t}x_t'\right)\left(\sum_{t=1}^{n}x_tx_t'\right)^{-1} - \frac{1}{n^{\alpha}}U_{0x}^{(n)}\left[V_{xx}^{(n)}\right]^{-1} \\ &= \left\{\sum_{t=1}^{n}u_{0t}x_t' - U_{0x}^{(n)}\left[V_{xx}^{(n)}\right]^{-1}\left(\sum_{t=1}^{n}\frac{x_tx_t'}{n^{\alpha}}\right)\right\}\left(\sum_{t=1}^{n}x_tx_t'\right)^{-1} \\ &= \left\{\sum_{t=1}^{n}\left[u_{0t}x_t' - U_{0x}^{(n)}\right] - U_{0x}^{(n)}\left[V_{xx}^{(n)}\right]^{-1}\sum_{t=1}^{n}\left(\frac{x_tx_t'}{n^{\alpha}} - V_{xx}^{(n)}\right)\right\}\left(\sum_{t=1}^{n}x_tx_t'\right)^{-1}.\end{aligned}$$

Hence, since $n^{-(1+\alpha)}\sum_{t=1}^{n}x_tx_t' \to_p V_{xx}$ and $U_{0x}^{(n)} \to \Lambda_{0x} + \Sigma_{0x}$ we obtain

$$\begin{aligned}& n^{\frac{1+\alpha}{2}}vec\left(\hat{A}_n - \bar{A}_n\right) \\ =\ & \left(V_{xx}^{-1}\otimes I_m\right)\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}vec\left[u_{0t}x_t' - U_{0x}^{(n)}\right] \\ & - \left[V_{xx}^{-1}\otimes(\Lambda_{0x}+\Sigma_{0x})V_{xx}^{-1}\right]\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\left(\frac{x_tx_t'}{n^{\alpha}} - V_{xx}^{(n)}\right) + o_p(1) \\ =\ & \frac{[V_{xx}^{-1}\otimes F_0(1)] - [V_{xx}^{-1}\otimes(\Lambda_{0x}+\Sigma_{0x})V_{xx}^{-1}]W_x[I_K\otimes F_x(1)]}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}(x_{t-1}\otimes\varepsilon_t) + o_p(1)\end{aligned}$$

by (12) and Lemma 3.5. The limiting distribution of $n^{\frac{1+\alpha}{2}} vec\left(\hat{A}_n - \bar{A}_n\right)$, presented in the following theorem, is now an immediate consequence of Lemma 3.3.

3.6 Theorem. *For each $\alpha \in (0,1)$ and $C = \text{diag}(c_i)$ with $c_i < 0$,*

$$n^{\frac{1+\alpha}{2}} vec\left(\hat{A}_n - \bar{A}_n\right) \Rightarrow N\left(0, \mathbb{V}\right) \quad \text{as } n \to \infty,$$

where $\bar{A}_n$ is defined in (15),

$$\begin{aligned}
\mathbb{V} &= V_{xx}^{-1} \otimes \Omega_{00} - \left(I_K \otimes \Omega_{0x}\right) W_x \left[V_{xx}^{-1} \otimes V_{xx}^{-1} \left(\Lambda_{0x} + \Sigma_{0x}\right)'\right] \\
&\quad - \left[V_{xx}^{-1} \otimes \left(\Lambda_{0x} + \Sigma_{0x}\right) V_{xx}^{-1}\right] W_x \left(I_K \otimes \Omega_{x0}\right) \\
&\quad + \left[V_{xx}^{-1} \otimes \left(\Lambda_{0x} + \Sigma_{0x}\right) V_{xx}^{-1}\right] W_x \left(V_{xx}^{-1} \otimes \Omega_{xx}\right) W_x \left[V_{xx}^{-1} \otimes V_{xx}^{-1} \left(\Lambda_{0x} + \Sigma_{0x}\right)'\right],
\end{aligned}$$

$W_x = -\left(I_K \otimes C + C \otimes I_K\right)^{-1} \left(I_{K^2} + \mathcal{K}_K\right)$ and $\mathcal{K}_K$ denotes the $K^2 \times K^2$ commutation matrix.

3.4 Remarks.

(i) It is apparent from (15) that least squares estimation induces a bias that is of the same order as the moderate deviations present in the regressors, i.e., $O\left(n^{-\alpha}\right)$. As $\alpha \to 1$, this bias becomes smaller in relative terms and of course when $\alpha = 1$ we know that it becomes absorbed into the limit distribution (cf. Phillips and Durlauf, 1986; Park and Phillips, 1988) and takes the form of a second-order bias. On the other hand, as $\alpha \to 0$, the bias becomes larger in relative terms and, ultimately, becomes part of the conventional simultaneous equations bias. In particular, when the error process is iid we have $u_t = \varepsilon_t$, $F_0(1) = [I, 0]$, $\Lambda_{0x} = 0$, so that $F_0(1)\Sigma_{\varepsilon x} + \Lambda_{0x} = E\left(\varepsilon_{0t}\varepsilon_{xt}'\right)$, and then $\bar{A} = E\left(\varepsilon_{0t}\varepsilon_{xt}'\right) E\left(x_t x_t'\right)^{-1} = E\left(\varepsilon_{0t} x_t'\right) E\left(x_t x_t'\right)^{-1}$ is the familiar simultaneous equations bias in stationary systems.

(ii) FM-OLS, FM-IV and FM-GMM methods can also be used, along the lines suggested in Phillips and Hansen (1990), Phillips (1995), and Kitamura and Phillips (1997). FM-IV and FM-GMM were designed to deal with endogeneities when some of the variables are stationary while at the same time treating nonstationary endogeneities and weak dependence in the errors. All these methods are indeed successful in removing some of the endogeneity bias, but they continue to import bias through the moderate integration coefficients because the FM endogeneity corrections do not fully account for the effects of moderate integration in the regressors. The problem is analogous to that studied by Elliott (1998) in the case of cointegrating regressions with near integrated regressors. New procedures are needed for dealing with this complication in the near stationary case and these are being developed in ongoing work (Phillips and Magdalinos, 2007d).

4. Limit theory for Moderately Explosive Systems

We now turn to the limit behavior of $\hat{A}_n$ when $R_n = I_K + C/n^a$ and $c_i > 0$ for all i. In considering the limit theory for the least squares estimator $\hat{A}_n$, the precise nature of the moderately explosive behavior turns out to be important. In particular, both the rate of convergence and the asymptotic distribution of $\hat{A}_n$ are affected by the relationship between the regressors in (2), i.e. by the precise form of the matrix C. We distinguish between two cases:

(I) C has distinct diagonal elements so that $c_i \neq c_j$ for all $i \neq j$.

(II) C does not have distinct diagonal elements and $c_i = c_j$ for some $i \neq j$.

In case (II) some regressors exhibit common moderately explosive behavior. The limit theory concerning x_t and its sample moments applies in both cases. However, the regression theory differs in important ways between (I) and (II) because the presence of common moderately explosive behavior among the regressors leads to a singular limit for the normalized second moment matrix (see (20) below), affecting the order of magnitude of $\left(\sum_{t=1}^n x_t x_t'\right)^{-1}$ and hence that of $\hat{A}_n - A$.

The limit theory for $\hat{A}_n - A$ is obtained by approximating the sample moment matrix $\sum_{t=1}^n x_t x_t'$ and the sample covariance $\sum_{t=1}^n u_{0t} x_t'$ using a version of a sample splitting argument that dates back to Anderson (1959). First, $\sum_{t=1}^n x_t x_t'$ is approximated in terms of the standardized observation $n^{-\alpha/2} R_n^{-n} x_n$ which is in turn approximated using a weighted sum of the first κ_n shocks. Next, $\sum_{t=1}^n u_{0t} x_t'$ is approximated in terms of this component and a weighted sum of the remaining data points. Anderson (1959) used $\kappa_n = \lfloor n/2 \rfloor$ in the sample splitting but did not make use of weak convergence arguments because a central limit theory did not apply in the explosive case that he considered. Here, we employ weak convergence methods and require only that $(\kappa_n)_{n\in\mathbb{N}}$ be any sequence increasing to infinity such that

$$\|R_n\|^{-\kappa_n} \to 0, \quad n^\alpha \|R_n\|^{-(n-\kappa_n)} \to 0 \tag{16}$$

for case (I), and

$$\|R_n\|^{-\kappa_n} \to 0, \quad \kappa_n/n \to 0 \tag{17}$$

for case (II). For $(\kappa_n)_{n\in\mathbb{N}}$ satisfying either (16) or (17), define

$$Y_{Cn} := \frac{1}{n^{\alpha/2}} \sum_{j=1}^{\kappa_n} R_n^{-j} F_x(1)\, \varepsilon_j.$$

It turns out that the stochastic sequence Y_{Cn} plays an important role in determining the asymptotic behavior of moderately explosive systems. The following lemma gives two properties of Y_{Cn} used throughout this section.

4.1 Lemma. *For each sequence κ_n satisfying $\|R_n\|^{-\kappa_n} \to 0$ we have, as $n \to \infty$,*

(a) $n^{-\alpha/2} R_n^{-n} x_n = n^{-\alpha/2} \sum_{j=1}^{n} R_n^{-j} u_{xj} = Y_{Cn} + o_p(1)$,

(b) $Y_{Cn} \Rightarrow Y_C =_d N\left(0, \int_0^\infty e^{-pC} \Omega_{xx} e^{-pC} dp\right)$,

for all $\alpha \in (0,1)$ and $C > 0$.

The asymptotic behavior of the sample variance matrix of x_t can be determined in terms of the limiting random vector Y_C of Lemma 4.1 (b). Using the fact that

$$\frac{1}{n^\alpha}\left(R_n^{-n} \otimes R_n^{-n}\right) \sum_{t=1}^{n} vec\left(u_{xt} x_{t-1}'\right) = o_p(1) \quad \text{as } n \to \infty, \tag{18}$$

which is proved in the Appendix, we can use the recursive property (2) to obtain

$$\begin{aligned}
& \left(R_n \otimes R_n - I_{K^2}\right) \frac{1}{n^\alpha}\left(R_n^{-n} \otimes R_n^{-n}\right) \sum_{t=1}^{n} vec\left(x_t x_t'\right) \\
= \; & \frac{1}{n^\alpha}\left(R_n^{-n+1} \otimes R_n^{-n+1}\right) vec\left(x_n x_n'\right) + o_p(1) \\
= \; & \left[I_{K^2} + o_p(1)\right] vec\left(\frac{1}{n^{\alpha/2}} \sum_{j=1}^{n} R_n^{-j} u_{xj}\right)\left(\frac{1}{n^{\alpha/2}} \sum_{j=1}^{n} R_n^{-j} u_{xj}\right)' \\
= \; & vec Y_{Cn} Y_{Cn}' + o_p(1),
\end{aligned}$$

by part (a) of Lemma 4.1. Since $n^\alpha\left(R_n \otimes R_n - I_{K^2}\right) \to C \otimes I_K + I_K \otimes C$ as $n \to \infty$, we conclude that

$$\begin{aligned}
\frac{1}{n^{2\alpha}}\left(R_n^{-n} \otimes R_n^{-n}\right) \sum_{t=1}^{n} vec\left(x_t x_t'\right) &= \left(C \otimes I_K + I_K \otimes C\right)^{-1}\left(vec Y_{Cn} Y_{Cn}'\right) + o_p(1) \\
&= \int_0^\infty \left(e^{-pC} \otimes e^{-pC}\right) dp \left(vec Y_{Cn} Y_{Cn}'\right) + o_p(1) \\
&= vec\left(\int_0^\infty e^{-pC} Y_{Cn} Y_{Cn}' e^{-pC} dp\right) + o_p(1),
\end{aligned}$$

that is,

$$\frac{1}{n^{2\alpha}} R_n^{-n} \sum_{t=1}^{n} x_t x_t' R_n^{-n} = \int_0^\infty e^{-pC} Y_{Cn} Y_{Cn}' e^{-pC} dp + o_p(1) \quad \text{as } n \to \infty. \tag{19}$$

Thus, part (b) of Lemma 4.1 and the continuous mapping theorem yield the following limiting distribution for the sample variance matrix:

$$\frac{1}{n^{2\alpha}} \sum_{t=1}^{n} R_n^{-n} x_t x_t' R_n^{-n} \Rightarrow \int_0^\infty e^{-pC} Y_C Y_C' e^{-pC} dp \quad \text{as } n \to \infty. \tag{20}$$

Note that the weak limit (20) of the standardized sample moment matrix is not always non singular. For example, when all localising coefficients are the same, $C = cI_K$, $\int_0^\infty e^{-pC}Y_CY_C'e^{-pC}dp = \frac{1}{2c}Y_CY_C'$ has rank unity. For the general case, denote the i-th element of the random vector Y_C by $Y_C^{(i)}$ and define the matrices

$$M_C := \left[\frac{1}{c_i + c_j} : i,j \in \{1,...,K\}\right] \text{ and } \check{Y}_C := \text{diag}\left(Y_C^{(1)},...,Y_C^{(K)}\right).$$

Since Y_C is a Gaussian random vector, $Y_C^{(i)} \neq 0$ $a.s.$ for each i. Thus, the identity

$$\int_0^\infty e^{-pC}Y_CY_C'e^{-pC}dp = \check{Y}_C M_C \check{Y}_C \tag{21}$$

implies that $\int_0^\infty e^{-pC}Y_CY_C'e^{-pC}dp$ is nonsingular whenever the matrix M_C is nonsingular, i.e. if and only if $c_i \neq c_j$ for all $i \neq j$. On the other hand, the rank of M_C and $\int_0^\infty e^{-pC}Y_CY_C'e^{-pC}dp$ is reduced by one for every pair of equal localising coefficients.

We now turn our attention to the matrix of sample covariances. The following lemma provides the first step towards an approximation of the sample covariance matrix by a martingale array (cf. equation (23) below).

4.2 Lemma. *For each $\alpha \in (0,1)$, $C > 0$ and a sequence κ_n satisfying (16) we have, as $n \to \infty$,*

(a) $n^{-\alpha}E\left\|vec\sum_{t=1}^n u_{0t}\left(\sum_{j=t+1}^n R_n^{t-j}u_{xj}\right)'R_n^{-n}\right\| \to 0,$

(b) $n^{-\alpha}E\left\|vec\sum_{t=1}^{\kappa_n} u_{0t}\left(\sum_{j=1}^n R_n^{t-j}u_{xj}\right)'R_n^{-n}\right\| \to 0.$

Using Lemma 4.2 the sample covariance can be written as

$$\begin{aligned}
vec\frac{1}{n^\alpha}\sum_{t=1}^n u_{0t}x_t'R_n^{-n} &= \frac{1}{n^\alpha}vec\sum_{t=1}^n u_{0t}\left(\sum_{j=1}^t R_n^{t-j}u_{xj}\right)'R_n^{-n} + o_p(1) \\
&= \frac{1}{n^\alpha}vec\sum_{t=\kappa_n+1}^n u_{0t}\left(\sum_{j=1}^n R_n^{t-j}u_{xj}\right)'R_n^{-n} + o_p(1) \\
&= \frac{1}{n^{\alpha/2}}\sum_{t=\kappa_n+1}^n \left(R_n^{-(n-t)} \otimes u_{0t}\right)\left(\frac{1}{n^{\alpha/2}}\sum_{j=1}^n R_n^{-j}u_{xj}\right) + o_p(1) \\
&= [I_{mK} + o_p(1)]\frac{1}{n^{\alpha/2}}\sum_{t=\kappa_n+1}^n \left(R_n^{-(n-t)} \otimes u_{0t}\right)Y_{Cn} + o_p(1).
\end{aligned}$$

Applying the BN decomposition on the above expression and making use of the fact (proved in the Appendix) that

$$\frac{1}{n^{\alpha/2}}\sum_{t=\kappa_n+1}^{n}\left(R_n^{-(n-t)}\otimes\Delta\tilde{\varepsilon}_{0t}\right)=o_p(1)\quad\text{as } n\to\infty, \tag{22}$$

we obtain the following expression for the sample covariance as $n\to\infty$:

$$vec\frac{1}{n^{\alpha}}\sum_{t=1}^{n}u_{0t}x_t'R_n^{-n}=\left[I_{mK}+o_p(1)\right]\frac{1}{n^{\alpha/2}}\sum_{t=1}^{n-\kappa_n}\left[R_n^{-(n-\kappa_n-t)}Y_{Cn}\otimes F_0(1)\,\varepsilon_{t+\kappa_n}\right]. \tag{23}$$

Letting $\mathcal{F}_{n,i}:=\sigma(x_0,\varepsilon_i,\varepsilon_{i-1},...)$ it is clear that, since Y_{Cn} is $\mathcal{F}_{n,\kappa_n}$-measurable,

$$\xi_{n,t+\kappa_n}:=\frac{1}{n^{\alpha/2}}\left[R_n^{-(n-\kappa_n-t)}Y_{Cn}\otimes F_0(1)\,\varepsilon_{t+\kappa_n}\right] \tag{24}$$

is an $\mathbb{R}^{mK}$-valued martingale difference array with respect to $\mathcal{F}_{n,t+\kappa_n}$. Moreover, in the notation of Proposition A1, denote by $M_{n,k}:=\sum_{t=1}^{k}\xi_{n,t+\kappa_n}$ the martingale array corresponding to $\xi_{n,t+\kappa_n}$ and by $\langle M_n\rangle_k:=\sum_{t=1}^{k}E_{\mathcal{F}_{n,t+\kappa_n-1}}\left(\xi_{n,t+\kappa_n}\xi_{n,t+\kappa_n}'\right)$ the predictable quadratic variation of $M_{n,k}$. Since

$$\begin{aligned}\langle M_n\rangle_{n-\kappa_n} &= \left[\frac{1}{n^{\alpha}}\sum_{t=1}^{n-\kappa_n}R_n^{-(n-\kappa_n-t)}Y_{Cn}Y_{Cn}'R_n^{-(n-\kappa_n-t)}\right]\otimes\Omega_{00}\\ &= \left[I_{mK}+o(1)\right]\left(\int_0^{\infty}e^{-pC}Y_{Cn}Y_{Cn}'e^{-pC}dp\otimes\Omega_{00}\right)\\ &\Rightarrow \int_0^{\infty}e^{-pC}Y_CY_C'e^{-pC}dp\otimes\Omega_{00},\end{aligned} \tag{25}$$

and $\langle M_n\rangle_{n-\kappa_n}$ is $\mathcal{F}_{n,\kappa_n}$-measurable with $\mathcal{F}_{n,\kappa_n}\subseteq\mathcal{F}_{n,\kappa_n+1}$, i.e. the σ-algebra supporting $\langle M_n\rangle_{n-\kappa_n}$ is smaller than each of the elements of the filtration supporting $\xi_{n,t+\kappa_n}$, part (c) of Proposition A1 yields

$$\begin{aligned}vec\left(\frac{1}{n^{\alpha}}\sum_{t=1}^{n}u_{0t}x_t'R_n^{-n}\right) &= M_{n,n-\kappa_n}+o_p(1)\\ &\Rightarrow MN\left(0,\int_0^{\infty}e^{-pC}Y_CY_C'e^{-pC}dp\otimes\Omega_{00}\right)\end{aligned} \tag{26}$$

as $n\to\infty$. Verification of the Lindeberg and tightness conditions (see (36) and (37) in Proposition A1) needed for this result is provided in the Appendix.

The limit behavior of the regression coefficient in case (I) where the localising coefficients are distinct may now be deduced as follows. Using equations (19) and

(23) we obtain

$$
\begin{aligned}
& vec\left[n^{\alpha}\left(\hat{A}_n - A\right) R_n^n\right] \\
= & \left[\left(\frac{1}{n^{2\alpha}}\sum_{t=1}^{n} R_n^{-n} x_t x_t' R_n^{-n}\right)^{-1} \otimes I_m\right] vec\left(\frac{1}{n^{\alpha}}\sum_{t=1}^{n} u_{0t} x_t' R_n^{-n}\right) \\
= & \left[I_{mK} + o_p(1)\right]\left[\left(\int_0^{\infty} e^{-pC} Y_{Cn} Y_{Cn}' e^{-pC} dp\right)^{-1} \otimes I_m\right] \sum_{t=1}^{n-\kappa_n} \xi_{n,t} \\
= & \left[I_{mK} + o_p(1)\right]\left[\left(\int_0^{\infty} e^{-pC} Y_{Cn} Y_{Cn}' e^{-pC} dp\right)^{-1} \otimes \Omega_{00}^{-1}\right] (I_K \otimes \Omega_{00}) \sum_{t=1}^{n-\kappa_n} \xi_{n,t} \\
= & \left[I_{mK} + o_p(1)\right] \langle M_n \rangle_{n-\kappa_n}^{-1} (I_K \otimes \Omega_{00}) M_{n,n-\kappa_n}. \qquad (27)
\end{aligned}
$$

The limiting distribution of $M_{n,n-\kappa_n}$ is established in (26). We also show in the Appendix that $M_{n,n-\kappa_n}$ satisfies the requirements of Proposition A1 (c), so that joint convergence of $M_{n,n-\kappa_n}$ and $\langle M_n \rangle_{n-\kappa_n}$ applies. Thus, (25) and (26) imply that the regression coefficient matrix has a mixed normal limiting distribution provided in the following theorem.

4.3 Theorem. *For the model (1) - (2) in case (I) with $R_n = I_K + C/n^{\alpha}$, $c_i > 0$ for all i, $c_i \neq c_j$* for all $i \neq j$, $\alpha \in (0,1)$ *and weakly dependent errors satisfying Assumption **LP**, we have*

$$
n^{\alpha}\left(\hat{A}_n - A\right) R_n^n \Rightarrow MN\left(0, \left(\int_0^{\infty} e^{-pC} Y_C Y_C' e^{-pC} dp\right)^{-1} \otimes \Omega_{00}\right). \qquad (28)
$$

4.4 Remarks.

(i) In a 2-equation system, $K = 1$, so x_t, A, $C = c$ and $R_n = \rho_n = 1 + c/n^{\alpha}$ are scalar. Then Y_c is $N(0, \Omega_{xx}/2c)$. Thus, letting $Z =_d N(0,1)$, the limiting distribution of Theorem 4.3 reduces to

$$
\begin{aligned}
n^{\alpha}\left(\hat{A}_n - A\right) R_n^n \ \Rightarrow\ & MN\left(0, 2c\Omega_{00} Y_c^{-2}\right) =_d 2c\left(\frac{\Omega_{00}}{\Omega_{xx}}\right)^{1/2}\left[\left(\frac{2c}{\Omega_{xx}}\right)^{1/2} Y_c\right]^{-1} Z \\
=_d\ & 2c\left(\frac{\Omega_{00}}{\Omega_{xx}}\right)^{1/2} \mathcal{C},
\end{aligned}
$$

where $\mathcal{C}$ is a standard Cauchy variate. In the general case, the limit distribution (28) is mixed normal with Cauchy-type tails (cf. Phillips, 1994) and its exact form can be obtained by using a matrix quotient argument, as in Phillips (1985).

(ii) The limit theory (28) applies irrespective of the distribution of the errors u_t in (1) and (2). Thus, the result gives an invariance principle for regressions with moderately explosive processes. Note that the limit theory applies also in the case of weakly dependent errors and relies only on the long-run covariance structure of the errors. The results are also invariant to the initial condition x_0. Hence, (28) provides the basis for asymptotically valid testing and confidence interval construction. All that is needed is the estimation of the long run variance matrix Ω_{00} by conventional kernel methods from the regression residuals.

(iii) When $\alpha = 0$, the autoregressive roots of (2) no longer lie in a neighborhood of unity and we obtain a purely explosive cointegrated system. Setting $\Theta = I_K + C$ with $c_i \neq c_j$ for all $i \neq j$ and assuming $x_0 = 0$ and i.i.d. $N(0, \Sigma)$ innovations u_t with $\Sigma = \text{diag}(\Sigma_{00}, \Sigma_{xx})$, Phillips and Magdalinos (2007c) obtain the limiting distribution of the regression coefficient matrix to be

$$vec\left[\left(\hat{A}_n - A\right)\Theta^n\right] \Rightarrow MN\left(0, \left(\sum_{j=0}^{\infty}\Theta^{-j}X_\Theta X_\Theta'\Theta^{-j}\right)^{-1} \otimes \Sigma_{00}\right),$$

where $X_\Theta =_d N\left(0, \sum_{j=1}^{\infty}\Theta^{-j}\Sigma_{xx}\Theta^{-j}\right)$. The matrix $\sum_{j=0}^{\infty}\Theta^{-j}X_\Theta X_\Theta'\Theta^{-j}$ is positive definite in view of the assumption of distinct localising coefficients c_i. Note, however, that the limit theory in this case depends crucially on the assumptions concerning the initialization and the innovations. As in the AR(1) case of Anderson (1959), no central limit theory applies in general and the asymptotic distribution of the least squares estimator is characterized by the distributional assumptions imposed on the innovations.

We now consider case (II) where the matrix C does not have distinct diagonal elements. In this case two or more elements of x_t have comparable moderately explosive behavior governed by a common autoregressive root of the form $\rho_{nj} = 1 + c_j/n^\alpha$. We know by (21) that under such conditions the second moment matrix $\sum_{t=1}^{n} x_t x_t'$ in the regression (6) is asymptotically singular, which is explained by the fact that some elements of x_t have common moderately explosive behavior. These elements of x_t are then asymptotically multicollinear in much the same way as regressors that are cointegrated or have common deterministic trends (c.f., Park and Phillips, 1989; Phillips, 1995). To deal with this singularity in the regression, we can develop an asymptotic theory for the regression in a similar way by rotating the regression coordinates in the direction of the common explosive behavior and in an orthogonal direction. Here, however, the rotation is a random process determined by the regressor vector x_n. The randomness and the sample size dependence in the rotation present further complications in the development of the asymptotics because the limit theory segmentation then depends on weak convergence of the rotation matrix. In what follows we illustrate the process with a development for the special case where C is the scalar matrix $C = cI_K$ and $\rho_n = 1 + c/n^\alpha$.

Start by defining the orthogonal random matrix $H_n = [H_{cn}, H_{\perp n}]$, where

$$H_{cn} = \frac{x_n}{(x_n' x_n)^{1/2}}, \quad H_{\perp n}' H_{cn} = 0 \; a.s. \tag{29}$$

and the $K \times (K-1)$ random matrix $H_{\perp n}$ is an orthogonal complement to H_{cn} satisfying $H_{\perp n}' H_{\perp n} = I_{K-1}$ and $H_{\perp n} H_{\perp n}' = I_K - H_{cn} H_{cn}'$ almost surely. Thus, by Lemma 4.1, the asymptotic behavior of $H_{\perp n} H_{\perp n}'$ is given by

$$H_{\perp n} H_{\perp n}' = I_K - \frac{Y_{cn} Y_{cn}'}{Y_{cn}' Y_{cn}} + o_p(1) \Rightarrow I_K - \frac{Y_c Y_c'}{Y_c' Y_c} \quad \text{as } n \to \infty, \tag{30}$$

where Y_{cn} and Y_c are the random vectors Y_{Cn} and Y_C of Lemma 4.1 with $C = cI_K$. Applying this rotation to the moderately explosive regressor vector yields

$$z_t = H_n' x_t = \begin{bmatrix} H_{cn}' x_t \\ H_{\perp n}' x_t \end{bmatrix} =: \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix}$$

with z_{2t} satisfying the reverse autoregression $z_{2t} = \rho_n^{-1} z_{2t+1} - \rho_n^{-1} H_{\perp n}' u_{xt}$ which gives rise to

$$z_{2t} = -H_{\perp n}' \sum_{j=1}^{n-t} \rho_n^{-j} u_{xt+j},$$

since $z_{2n} = H_{\perp n}' x_n = 0$. Using orthogonality of H_n, we obtain the following expression for the least squares estimator after rotation of the regression space

$$\begin{aligned} n^{\frac{1+\alpha}{2}} \left(\hat{A}_n - A \right) &= \left(\frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^n u_{0t} z_t' \right) \left(\frac{1}{n^{1+\alpha}} \sum_{t=1}^n z_t z_t' \right)^{-1} H_n' \\ &= \left(\frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^n u_{0t} z_t' \right) \begin{bmatrix} \frac{Z_1' Z_1}{n^{1+\alpha}} & \frac{Z_1' Z_2}{n^{1+\alpha}} \\ \frac{Z_2' Z_1}{n^{1+\alpha}} & \frac{Z_2' Z_2}{n^{1+\alpha}} \end{bmatrix}^{-1} H_n', \end{aligned} \tag{31}$$

where $Z_1 = [z_{11}, z_{12}, ..., z_{1n}]' \in \mathbb{R}^n$, $Z_2 = [z_{21}', z_{22}', ..., z_{2n}']' \in \mathbb{R}^{n \times (K-1)}$,

$$\Pi_{1n} = (Z_1' Z_1)^{-1} Z_1' Z_2 \;\text{ and }\; Q_1 = I_n - Z_1 (Z_1' Z_1)^{-1} Z_1'.$$

Expressed in the above form, the inverse of the normalised second moment matrix can be obtained as

$$\begin{aligned} \left(\frac{\sum_{t=1}^n z_t z_t'}{n^{1+\alpha}} \right)^{-1} &= \begin{bmatrix} \left(\frac{Z_1' Z_1}{n^{1+\alpha}} \right)^{-1} + \Pi_{1n} \left(\frac{Z_2' Q_1 Z_2}{n^{1+\alpha}} \right)^{-1} \Pi_{1n}' & -\Pi_{1n} \left(\frac{Z_2' Q_1 Z_2}{n^{1+\alpha}} \right)^{-1} \\ -\left(\frac{Z_2' Q_1 Z_2}{n^{1+\alpha}} \right)^{-1} \Pi_{1n}' & \left(\frac{Z_2' Q_1 Z_2}{n^{1+\alpha}} \right)^{-1} \end{bmatrix} \\ &= \begin{bmatrix} O_p\left(n^{1-\alpha} \rho_n^{-2n} \right) & O_p\left(\rho_n^{-n} \right) \\ O_p\left(\rho_n^{-n} \right) & \left(\frac{1}{n^{1+\alpha}} \sum_{t=1}^n z_{2t} z_{2t}' \right)^{-1} + o_p(1) \end{bmatrix}. \end{aligned} \tag{32}$$

The asymptotic orders of magnitude in (32) are derived in the Appendix. The following result shows that the limiting behavior of the inverse of the second moment matrix is determined by the asymptotically non singular matrix $\left(\frac{Z_2'Q_1Z_2}{n^{1+\alpha}}\right)^{-1}$.

4.5 Lemma.

(a) *The matrix* $n^{-(1+\alpha)}\sum_{t=1}^{n} z_{2t}z_{2t}'$ *is asymptotically non singular a.s. with*

$$\left(\frac{1}{n^{1+\alpha}}\sum_{t=1}^{n} z_{2t}z_{2t}'\right)^{-1} = \left(\frac{1}{2c}H_{\perp n}'\Omega_{xx}H_{\perp n}\right)^{-1} + o_p(1).$$

(b) *The limiting distribution of the random matrix* $M_{H_{\perp n}} = H_{\perp n}\left(H_{\perp n}'\Omega_{xx}H_{\perp n}\right)^{-1}H_{\perp n}'$ *is given by*

$$M_{H_{\perp n}} \Rightarrow H_{\perp}\left(H_{\perp}'\Omega_{xx}H_{\perp}\right)^{-1}H_{\perp}' =: M_{H_{\perp}}$$

as $n\to\infty$*, where* $H_\perp$ *is a* $K\times(K-1)$ *random matrix (an orthogonal complement to* Y_c*) satisfying* $H_\perp H_\perp' = I_K - \left(Y_c'Y_c\right)^{-1}Y_cY_c'$*.*

Note that the random matrix $H_{\perp n}'\Omega_{xx}H_{\perp n}$ is positive definite *a.s.*, since $\Omega_{xx}>0$ and $H_{\perp n}$ has full column rank *a.s.*. In view of Lemma 4.5 and the fact that, by (23), $\sum_{t=1}^{n}u_{0t}z_{1t} = O_p\left(n^{\alpha}\rho_n^n\right)$, the least squares estimator becomes

$$n^{\frac{1+\alpha}{2}}\left(\hat{A}_n - A\right) = \left[I_{mK} + O_p\left(\frac{1}{n^{1/2}}\right)\right]\left(\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}u_{0t}z_{2t}'\right)\left(\frac{1}{n^{1+\alpha}}\sum_{t=1}^{n}z_{2t}z_{2t}'\right)^{-1}H_{\perp n}',$$

where by an identical argument to that used in the derivation of (10)

$$\frac{1}{n}\sum_{t=1}^{n}u_{0t}z_{2t}' = -\frac{1}{n}\sum_{t=1}^{n}u_{0t}\tilde{\varepsilon}_{xt}'H_{\perp n} + o_p(1) = -\Lambda_{0x}'H_{\perp n} + o_p(1).$$

Thus, the rates of convergence of $\sum_{t=1}^{n}u_{0t}z_{2t}'$ and $\sum_{t=1}^{n}z_{2t}z_{2t}'$ are the same as in the moderately stationary case.

As a result of cointegration within the vector of moderately explosive regressors, the regression signal is weaker than in the nonsingular case of Theorem 4.3 and so temporal dependence in the errors now affects the limit theory and induces an asymptotic bias in the least squares estimator. As in the moderately stationary case, the bias of $\hat{A}_n$ can be expressed in terms of the matrices

$$\bar{\Lambda}_{x\nu}^{(n)} = \rho_n^{-1}\Lambda_{x\nu} + \frac{\rho_n^{-1}c}{n^{\alpha}}\Gamma_{x\nu}^{\tilde{\varepsilon}}(1) - \frac{c}{n^{\alpha}}\sum_{j=1}^{n-1}\rho_n^{-j-1}\tilde{F}_{xj}\Sigma_{\varepsilon}F_{\nu}(1)' - \frac{c^2}{n^{2\alpha}}\sum_{j=1}^{n-1}\rho_n^{-j-1}\Gamma_{x\nu}^{\tilde{\varepsilon}}(j)$$

for $\nu \in \{0, x\}$ and $\bar{\Omega}_{xx}^{(n)} = \left(1 + \frac{c}{2n^\alpha}\right)^{-1} \left[\Sigma_{xx} + \bar{\Lambda}_{xx}^{(n)} + \bar{\Lambda}_{xx}^{(n)\prime}\right]$, whereas the asymptotic distribution of $\hat{A}_n$ will be obtained through a central limit theorem for the matrix martingale difference array

$$\Psi_n := n^{-\frac{1+\alpha}{2}} \sum_{t=1}^{n-1} \varepsilon_t \sum_{j=t+1}^{n} \rho_n^{-(j-t)} \varepsilon_j'.$$

4.6 Lemma. *Let* $\bar{\Psi}_n := n^{-\frac{1+\alpha}{2}} \sum_{t=\kappa_n+1}^{n-1} \varepsilon_t \sum_{j=t+1}^{n} \rho_n^{-(j-t)} \varepsilon_j'$ *with* $c > 0$ *and* $\alpha \in (0,1)$. *Then*

(a) $\Psi_n = \bar{\Psi}_n + o_p(1)$ *as* $n \to \infty$, *for any sequence* $(\kappa_n)_{n\in\mathbb{N}}$ *satisfying (17).*

(b) $\psi_n := vec\Psi_n \Rightarrow N\left(0, \frac{1}{2c}\Sigma_\varepsilon \otimes \Sigma_\varepsilon\right)$.

Moreover, under the condition $E\left\|\varepsilon_1\right\|^4 < \infty$, *we have, for* $\nu \in \{0, x\}$

(c) $n^{-\frac{1+\alpha}{2}} \sum_{t=1}^{n-1} \left[u_{\nu t} z_{2t}' + \bar{\Lambda}_{x\nu}^{(n)\prime} H_{\perp n}\right] = -F_\nu(1)\Psi_n F_x(1)' H_{\perp n} + o_p(1)$,

(d) $n^{-\frac{1+\alpha}{2}} \sum_{t=1}^{n-1} \left[\frac{z_{2t} z_{2t}'}{n^\alpha} - H_{\perp n}' \frac{\bar{\Omega}_{xx}^{(n)}}{2c} H_{\perp n}\right] = \frac{1}{2c} H_{\perp n}' F_x(1)\left(\Psi_n + \Psi_n'\right) F_x(1)' H_{\perp n} + o_p(1)$.

For the least squares estimator, combining (31) and (32) we obtain

$$n^{\frac{1+\alpha}{2}}\left(\hat{A}_n - A\right) = \left(\frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n} u_{0t} z_{2t}'\right)\left(\frac{1}{n^{1+\alpha}} \sum_{t=1}^{n} z_{2t} z_{2t}'\right)^{-1} H_{\perp n}' + o_p(1).$$

Thus, letting

$$B_n = 2cn^{\frac{1-\alpha}{2}} \bar{\Lambda}_{x0}^{(n)\prime} H_{\perp n}\left(H_{\perp n}' \bar{\Omega}_{xx}^{(n)} H_{\perp n}\right)^{-1} H_{\perp n}' \tag{33}$$

Lemma 4.5, Lemma 4.6 and the fact that $H_{\perp n} H_{\perp n}' M_{H_{\perp n}} = M_{H_{\perp n}}$ imply that

$$\begin{aligned}
& n^{\frac{1+\alpha}{2}}\left(\hat{A}_n - A\right) + B_n \\
= \; & \frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n-1}\left[u_{0t} z_{2t}' + \bar{\Lambda}_{x0}^{(n)\prime} H_{\perp n}\right]\left(\frac{1}{n^{1+\alpha}} \sum_{t=1}^{n} z_{2t} z_{2t}'\right)^{-1} H_{\perp n}' \\
& + \frac{2c\bar{\Lambda}_{x0}^{(n)\prime} H_{\perp n}\left(H_{\perp n}' \bar{\Omega}_{xx}^{(n)} H_{\perp n}\right)^{-1}}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n-1}\left[\frac{z_{2t} z_{2t}'}{n^\alpha} - \frac{H_{\perp n}' \bar{\Omega}_{xx}^{(n)} H_{\perp n}}{2c}\right]\left(\frac{\sum_{t=1}^{n} z_{2t} z_{2t}'}{n^{1+\alpha}}\right)^{-1} H_{\perp n}' \\
= \; & \left\{\Lambda_{x0}' M_{H_{\perp n}} F_x(1)\left(\Psi_n + \Psi_n'\right) F_x(1)' H_{\perp n} H_{\perp n}' - F_0(1)\Psi_n F_x(1)' H_{\perp n} H_{\perp n}'\right\} \\
& \times H_{\perp n}\left(\frac{1}{n^{1+\alpha}} \sum_{t=1}^{n} z_{2t} z_{2t}'\right)^{-1} H_{\perp n}' + o_p(1) \\
= \; & 2c\left\{\Lambda_{x0}' M_{H_{\perp n}} F_x(1)\left(\bar{\Psi}_n + \bar{\Psi}_n'\right) F_x(1)' M_{H_{\perp n}} - F_0(1)\bar{\Psi}_n F_x(1)' M_{H_{\perp n}}\right\} + o_p(1).
\end{aligned}$$

Denoting by $\mathcal{N}_{m+K} = \frac{1}{2}\left(I_{(m+K)^2} + \mathcal{K}_{m+K}\right)$ the $(m+K)^2 \times (m+K)^2$ symmetriser matrix, setting $\bar{\psi}_n = vec\bar{\Psi}_n$ and defining

$$\Delta_{H_{\perp n}} = 2\left[I_K \otimes \Lambda'_{x0} M_{H_{\perp n}}\right]\left[F_x(1) \otimes F_x(1)\right]\mathcal{N}_{m+K} - \left[F_x(1) \otimes F_0(1)\right],$$

we obtain

$$vec\left\{n^{\frac{1+\alpha}{2}}\left(\hat{A}_n - A\right) + B_n\right\} = 2c\left(M_{H_{\perp n}} \otimes I_m\right)\Delta_{H_{\perp n}}\bar{\psi}_n + o_p(1). \tag{34}$$

Since $\bar{\psi}_n$ is $\sigma(\varepsilon_{\kappa_n+1}, \varepsilon_{\kappa_n+2}, ...)$ measurable and, by (30), $H_{\perp n}H'_{\perp n}$ can be approximated by a $\sigma(\varepsilon_1, ..., \varepsilon_{\kappa_n})$ measurable process, $\bar{\psi}_n$ is asymptotically independent of $M_{H_{\perp n}}$ and $\Delta_{H_{\perp n}}$. Consequently, joint weak convergence applies in (34) and the limiting distribution of the least squares estimator follows immediately from Lemma 4.5 (b), Lemma 4.6 (b) and (30).

4.7 Theorem. *For the model (1) - (2) in case (II) with $R_n = (1 + c/n^\alpha) I_K$, $c > 0$, $\alpha \in (0,1)$ and weakly dependent errors satisfying Assumption LP with $E\|\varepsilon_1\|^4 < \infty$, the following mixed Gaussian limit theory applies*

$$vec\left\{n^{\frac{1+\alpha}{2}}\left(\hat{A}_n - A\right) + B_n\right\} \Rightarrow MN(0, 2c\mathbb{U}),$$

for the least squares estimator, with mixing matrix

$$\begin{aligned}\mathbb{U} &= 2M_{H_\perp} \otimes \Lambda'_{x0} M_{H_\perp}\Lambda_{x0} + M_{H_\perp} \otimes \Omega_{00} - M_{H_\perp} \otimes \Lambda'_{x0} M_{H_\perp}\Omega_{x0} - M_{H_\perp} \otimes \Omega_{0x} M_{H_\perp}\Lambda_{x0} \\ &\quad + \left(2M_{H_\perp}\Lambda_{x0} \otimes \Lambda'_{x0} M_{H_\perp} - M_{H_\perp}\Omega_{x0} \otimes \Lambda'_{x0} M_{H_\perp} - M_{H_\perp}\Lambda_{x0} \otimes \Omega_{0x} M_{H_\perp}\right)\mathcal{K}_{m+K},\end{aligned}$$

where $H_\perp$ is a $K \times (K-1)$ random matrix satisfying $H_\perp H'_\perp = I_K - (Y'_c Y_c)^{-1} Y_c Y'_c$ and $M_{H_\perp}$ is the random projection matrix $M_{H_\perp} = H_\perp (H'_\perp \Omega_{xx} H_\perp)^{-1} H'_\perp$.

4.8 Remarks.

(i) The limit distribution of $\hat{A}_n$ is mixed normal and the mixing matrix depends on Y_c, the limit variate of $n^{-\alpha/2}\rho_n^{-n} x_n$, through the identity $H_\perp H'_\perp = I_K - (Y'_c Y_c)^{-1} Y_c Y'_c$. The limit distribution of Theorem 4.7 is degenerate as the matrix $M_{H_\perp} = H_\perp (H'_\perp \Omega_{xx} H_\perp)^{-1} H'_\perp$ has rank $K-1$. In the direction of Z_c, the limit distribution is also mixed normal and of the Cauchy type, just as in case (I) above. In this direction the rate of convergence is faster, viz. $n^\alpha \|R_n\|^n$.

(ii) Since the limit distribution is mixed normal in both directions, we may proceed with inference about A in the absence of serial correlation ($\Lambda_{0x} = 0$). Standard errors are then obtained in the usual way and are based on combining the appropriate diagonal elements of $(X'X)^{-1}$ with those of an estimate of the long-run variance matrix $\Omega_{00.}$

(iii) The general case where some but not all localising coefficients are equal presents some further complications that relate to the position of the equal localising coefficients in the C matrix. Since the repeated localising coefficients that are the source of the degeneracy in Theorem 4.7 may be scattered along the diagonal of the C (or the R_n) matrix, rotation of the regression coordinates requires appropriate grouping of these repeated diagonal elements. Using the method of Phillips and Magdalinos (2007c), if there are p groups of repeated diagonal elements of R_n, the autoregressive matrix can be rearranged as $R_n = \Pi' \Phi_n \Pi$, where Π is a permutation matrix, Φ_n =diag(Φ_{n1}, Φ_{n2}) and

$$\begin{aligned} \Phi_{n1} &= \text{diag}\left(\rho_{n1} I_{r_1}, ..., \rho_{np} I_{r_p}\right), \\ \Phi_{n2} &= \text{diag}\left(\varphi_{n1}, ..., \varphi_{n,K-r}\right), \quad r = \textstyle\sum_{i=1}^{p} r_i, \end{aligned}$$

where all φ_{ni} and ρ_{ni} are diagonal elements of R_n with $\varphi_{ns} \neq \rho_{nl}$ for all s, l and $\varphi_{ni} \neq \varphi_{nj}$, $\rho_{ni} \neq \rho_{nj}$ for all $i \neq j$. This rearrangement transforms the system of equations in (2) into a system where the first r_1 equations contain the repeated root ρ_{n1}, the next r_2 equations contain the repeated root ρ_{n2} and so forth, whereas the last $K - r$ equations contain all distinct diagonal elements φ_{ni} of R_n. Letting Π_1 denote the matrix obtained from the first r rows of Π, we can construct an orthogonal complement to $\Pi_1 x_n =: \left(\psi_{n1}', ..., \psi_{np}'\right)'$, $\psi_{ni} \in \mathbb{R}^{r_i}$, as follows: for each $i \in \{1, ..., p\}$, consider an $r_i \times (r_i - 1)$ orthogonal complement $H_{\perp i}^{(n)}$ to each $\psi_{ni} / \left(\psi_{ni}' \psi_{ni}\right)^{1/2}$ satisfying $H_{\perp i}^{(n)\prime} \psi_{ni} = 0$ and $H_{\perp i}^{(n)\prime} H_{\perp i}^{(n)} = I_{r_i - 1}$ *a.s.*. Then $H_{n\perp}$:=diag$\left(H_{\perp 1}^{(n)}, ..., H_{\perp p}^{(n)}\right)$ is the required orthogonal complement satisfying $H_{n\perp}' \Pi_1 x_n = 0$ and $H_{n\perp}' H_{n\perp} = I_{r-p}$ almost surely. With this notation in place, we can rotate the regression coordinates in a direction orthogonal to $\Pi_1 x_n$ by using equation (13) of Phillips and Magdalinos (2007c). For simplicity, assume that the u_t are *iid* with $E u_t u_t' = \text{diag}(\Omega_{00}, \Omega_{xx})$. Then a similar analysis to Phillips and Magdalinos (2007c) and Lemma 4.5 yields the following asymptotic regression theory:

$$n^{\frac{1+\alpha}{2}} vec\left(\hat{A}_n - A\right) \Rightarrow MN\left(0, \Pi_1' \tilde{M}_{H_\perp} \Pi_1 \otimes \Omega_{00}\right), \tag{35}$$

where

$$\tilde{M}_{H_\perp} = H_\perp \left(H_\perp' \Pi_1 \int_0^\infty e^{-pC} \Omega_{xx} e^{-pC} dp \Pi_1' H_\perp\right)^{-1} H_\perp'.$$

Clearly, when $C = cI_K$, $r = K$, $\tilde{M}_{H_\perp} = 2cM_{H_\perp}$ and Theorem 4.7 follows from (35) by setting $\Lambda_{x0} = 0$ and $\Pi_1 = I_K$. Although the lack of correlation between u_{0t} and u_{xt} means that (35) applies to a less general class of models, (35) still provides useful insight into the (reduced) rank of the limiting distribution of $n^{\frac{1+\alpha}{2}}\left(\hat{A}_n - A\right)$. The rank of the limiting covariance matrix in (35) is given by

$$(r - p)\, m = \left(\sum_{i=1}^{p} r_i - p\right) m,$$

where p is the number of repeated roots of R_n and r_i is the number of times that the repeated root ρ_{ni} appears in R_n. Hence, the limiting covariance matrix assumes its maximum rank, $(K-1)\,m$, when all diagonal elements of R_n are equal. On the other hand, the minimum rank, m, occurs when $r = 2$ and $p = 1$, i.e. when R_n has exactly 2 equal diagonal elements. These findings are consistent with the theory for purely explosive cointegrated systems developed in Phillips and Magdalinos (2007c).

5. Discussion

This paper extends the study of cointegrated systems to models with moderately integrated and moderately explosive regressors and weakly dependent errors. In cointegrated systems with moderately integrated regressors, the limit distribution theory is normal, as distinct from mixed normal, and the convergence rate is $n^{(1+\alpha)/2}$ rather than n. As in the conventional case of cointegrated regression with integrated or near integrated regressors, least squares suffers from asymptotic bias. This bias is related to the simultaneous equations bias that manifests in least squares regression in stationary systems. In the mildly integrated case, the bias is not removed by conventional cointegrating regression approaches such as FM regression and in this respect the problem is analogous to the case of cointegration with near integrated regressors (Elliott, 1998). Several possibilities are possible for dealing with this further complication and these new approaches to estimation and inference are discussed in a companion work (Phillips and Magdalinos, 2007d).

In the case of moderately explosive regressors, however, the present paper shows that the situation is much more favorable to least squares regression. Here the limit theory is mixed normal and least squares regression is asymptotically median unbiased. These results apply for non-Gaussian errors and quite general weakly dependent errors. They are therefore suitable for statistical inference without further modification. The theory may be used to construct asymptotically valid confidence intervals in regressions with moderately explosive regressors.

Another notable result is the degeneracy that arises in the moderately explosive case when the regressors are themselves explosively cointegrated. This situation may arise in practice because in most empirical cases of extreme behavior, the behavior manifests from a single source, with contamination effects on other variables. The net result is a form of mildly explosive co-movement that produces degeneracy in the regressor moment matrix. The impact of this degeneracy on the limit theory is a mixed normal limit distribution with a slower rate of convergence ($n^{\frac{1+\alpha}{2}}$) in general, but a faster rate of convergence in the explosive direction. This corresponds to the effect of the degeneracy produced by purely explosive ($\alpha = 0$) co-movement where the rate of convergence is reduced to $n^{1/2}$ (Phillips and Magdalinos, 2007c). In all the above cases, central limit arguments apply and the limit distribution theory is mixed normal.

The interpretation of α as the degree of persistence of the regressors raises the issue of whether robust inference on the coefficient matrix A is possible. As long as the asymptotic bias of $\hat{A}_n$ is removed -see the comment on the first paragraph of this section- it is possible to employ inference procedures that do not rely on knowledge of α. Self-normalised statistics for hypothesis testing such as Wald statistics or "Studentised" statistics based on $\hat{A}_n$ are useful in this connection. Phillips and Magdalinos (2007d) discuss an inference procedure on A that is robust to all types of regressor persistence, including integrated and local-to-unity regressors. Cointegrated systems with different degree of persistence across equations is an interesting extension of the present work, but we do not persue it here.

6. Notation

$[\cdot]$	quadratic variation	$\lVert\cdot\rVert$	spectral norm
$\langle\cdot\rangle$	predictable quadratic variation	$\mathbf{1}\{\cdot\}$	indicator function
Ω	$:= F(1)\Sigma_\varepsilon F(1)'$	$\mathcal{F}_{nt}$	$:= \sigma(x_0, \varepsilon_t, \varepsilon_{t-1}, ...)$
Σ	$:= E(u_t u_t')$	$E_{\mathcal{F}}(X)$	$:= E(X\vert\mathcal{F})$
Λ	$:= \sum_{h=1}^{\infty} E\left(u_t u_{t-h}'\right)$	$:=$	definitional equality
$V_{xx}, \Omega_{xx}^{(n)}, V_{xx}^{(n)}$	see (8), (13), (14)	$\to_p$	convergence in probability
$\Gamma_{\cdot x}^{\tilde{\varepsilon}}(k), M_{\cdot n}$	see Lemma 3.2	$\to_{L_p}, \to_{a.s.}$	L_p, $a.s.$ convergence
W_x	see Lemma 3.5	$\Rightarrow$	weak convergence
$\bar{A}_n$	see (15)	$=_d$	distributional equivalence
$H_{cn}, H_{\perp n}$	see (29)	$MN(\cdot,\cdot)$	mixed normal distribution
$M_{H_{\perp n}}, M_{H_\perp}$	see Lemma 4.5	$o_p(1)$	tends to zero in probability
$\mathcal{K}_s$	$s^2 \times s^2$ commutation matrix	$a \wedge b, a \vee b$	$\min(a,b), \max(a,b)$

7. Technical appendix and proofs

We begin by establishing two useful technical propositions. Throughout this section, B denotes a bounding constant in $(0,\infty)$ that may assume different values.

Proposition A1. *Suppose that* $\left(\xi_{nj}, \mathcal{F}_{nj}\right)$, $1 \le j \le k_n$, $n \in \mathbb{N}$, *is a zero mean, square integrable martingale difference array in* $\mathbb{R}^d$ *satisfying the conditional Lindeberg condition*

$$\sum_{j=1}^{k_n} E_{\mathcal{F}_{nj-1}}\left(\left\lVert\xi_{nj}\right\rVert^2 \mathbf{1}\left\{\left\lVert\xi_{nj}\right\rVert > \delta\right\}\right) \to_p 0 \quad \delta > 0. \tag{36}$$

Denote by $M_{nk_n} := \sum_{j=1}^{k_n} \xi_{nj}$ *the corresponding martingale array with quadratic variation and predictable quadratic variation given respectively by*

$$[M_n]_{k_n} := \sum_{j=1}^{k_n} \xi_{nj}\xi_{nj}' \quad \text{and} \quad \langle M_n \rangle_{k_n} := \sum_{j=1}^{k_n} E_{\mathcal{F}_{nj-1}}\left(\xi_{nj}\xi_{nj}'\right).$$

(a) *If*

$$\sup_{n\in\mathbb{N}} \sum_{j=1}^{k_n} E\left\|\xi_{nj}\right\|^2 < \infty \tag{37}$$

holds, then $\left\|[M_n]_{k_n} - \langle M_n \rangle_{k_n}\right\| \to_p 0$ *as* $n \to \infty$.

(b) *Let* H *be a (possibly stochastic) non negative definite* $d \times d$ *matrix and* Z *a* $N(0, I_d)$ *random vector independent of* H. *If*

$$\langle M_n \rangle_{k_n} \to_p H, \tag{38}$$

then $M_{nk_n} \Rightarrow H^{1/2}Z =_d MN(0, H)$.

(c) *Convergence in probability in (38) can be replaced by convergence in distribution provided that* $\langle M_n \rangle_{k_n}$ *is* $\mathcal{G}_n$*-measurable with* $\mathcal{G}_n \subseteq \mathcal{F}_{n1}$*, and (37) holds. In this case joint convergence of* $\left(M_{nk_n}, \langle M_n \rangle_{k_n}\right)$ *applies.*

Proof. First note that (36) and (37) imply that $E\left(\max_{j\le k_n}\left\|\xi_{nj}\right\|^2\right) \to 0$. To see this, write for arbitrary $\delta > 0$,

$$\begin{aligned} E\left(\max_{j\le k_n}\left\|\xi_{nj}\right\|^2\right) &\le \delta^2 + E\left(\max_{j\le k_n}\left\|\xi_{nj}\right\|^2 \mathbf{1}\left\{\left\|\xi_{nj}\right\| > \delta\right\}\right) \\ &\le \delta^2 + E\sum_{j=1}^{k_n}\left\|\xi_{nj}\right\|^2 \mathbf{1}\left\{\left\|\xi_{nj}\right\| > \delta\right\} \\ &= \delta^2 + E\left[\sum_{j=1}^{k_n} E_{\mathcal{F}_{nj-1}}\left(\left\|\xi_{nj}\right\|^2 \mathbf{1}\left\{\left\|\xi_{nj}\right\| > \delta\right\}\right)\right] \to \delta^2, \end{aligned} \tag{39}$$

as $n \to \infty$ by (36) and the dominated convergence theorem, which applies in view of (37). Note also that (37) implies tightness of the sequence $\langle M_n \rangle_{k_n}$ since, using the Markov inequality,

$$\begin{aligned} \lim_{\nu\to\infty}\sup_{n\in\mathbb{N}} P\left\{\left\|\langle M_n \rangle_{k_n}\right\| > \nu\right\} &\le \lim_{\nu\to\infty}\frac{1}{\nu}\sup_{n\in\mathbb{N}} E\left\|\langle M_n \rangle_{k_n}\right\| \\ &\le \lim_{\nu\to\infty}\frac{1}{\nu}\left(\sup_{n\in\mathbb{N}}\sum_{j=1}^{k_n} E\left\|\xi_{nj}\right\|^2\right) = 0. \end{aligned} \tag{40}$$

Part (a) requires a generalization of Theorem 2.23 of Hall and Heyde (1980, hereafter H&H) for vector valued martingale difference arrays. Using a similar truncation argument to H&H, for arbitrary $\varepsilon > 0$, let

$$\zeta_{nj} := \xi_{nj}\mathbf{1}\left\{\left\|\xi_{nj}\right\| \leq \varepsilon\right\},$$

$C_{nk_n} := \sum_{j=1}^{k_n} \zeta_{nj}\zeta_{nj}'$ and $D_{nk_n} := \sum_{j=1}^{k_n} E_{\mathcal{F}_{nj-1}}\left(\zeta_{nj}\zeta_{nj}'\right)$. Then, for each $\delta > 0$,

$$\begin{aligned} P\left(\left\|[M_n]_{k_n} - \langle M_n\rangle_{k_n}\right\| > 3\delta\right) &\leq P\left(\left\|[M_n]_{k_n} - C_{nk_n}\right\| > \delta\right) \\ &\quad + P\left(\left\|\langle M_n\rangle_{k_n} - D_{nk_n}\right\| > \delta\right) \\ &\quad + P\left(\left\|C_{nk_n} - D_{nk_n}\right\| > \delta\right). \end{aligned} \tag{41}$$

For the first term of (41), the definition of ζ_{nj} yields

$$\begin{aligned} P\left(\left\|[M_n]_{k_n} - C_{nk_n}\right\| > \delta\right) &\leq P\left(\left\|\xi_{nj}\right\| > \varepsilon \text{ for some } j\right) \\ &\leq P\left(\max_{j\leq k_n}\left\|\xi_{nj}\right\| > \varepsilon\right) \to 0 \end{aligned} \tag{42}$$

as $n \to \infty$, by (39). For the second term of (41), we have

$$\langle M_n\rangle_{k_n} - D_{nk_n} = \sum_{j=1}^{k_n} E_{\mathcal{F}_{nj-1}}\left(\xi_{nj}\xi_{nj}'\mathbf{1}\left\{\left\|\xi_{nj}\right\| > \varepsilon\right\}\right)$$

so

$$P\left(\left\|\langle M_n\rangle_{k_n} - D_{nk_n}\right\| > \delta\right) \leq P\left[\sum_{j=1}^{k_n} E_{\mathcal{F}_{nj-1}}\left(\left\|\xi_{nj}\right\|^2\mathbf{1}\left\{\left\|\xi_{nj}\right\| > \varepsilon\right\}\right) > \delta\right] \to 0 \tag{43}$$

as $n \to \infty$, by (36). For the third term of (41), $\zeta_{nj}\zeta_{nj}' - E_{\mathcal{F}_{nj-1}}\left(\zeta_{nj}\zeta_{nj}'\right)$ is a matrix martingale difference array, so

$$\begin{aligned} P\left(\left\|C_{nk_n} - D_{nk_n}\right\| > \delta\right) &\leq \delta^{-2}E\left\|C_{nk_n} - D_{nk_n}\right\|^2 \\ &= \delta^{-2}\sum_{j=1}^{k_n} E\left\|\zeta_{nj}\zeta_{nj}' - E_{\mathcal{F}_{nj-1}}\left(\zeta_{nj}\zeta_{nj}'\right)\right\|^2 \\ &\leq 2\delta^{-2}\sum_{j=1}^{k_n} E\left[\left\|\zeta_{nj}\right\|^4 + \left(E_{\mathcal{F}_{nj-1}}\left\|\zeta_{nj}\right\|^2\right)^2\right] \\ &\leq 2\delta^{-2}\sum_{j=1}^{k_n} E\left[\left\|\zeta_{nj}\right\|^4 + E_{\mathcal{F}_{nj-1}}\left\|\zeta_{nj}\right\|^4\right] \\ &= 4\delta^{-2}\sum_{j=1}^{k_n} E\left\|\zeta_{nj}\right\|^4 \leq 4\varepsilon^2\delta^{-2}\sum_{j=1}^{k_n} E\left\|\xi_{nj}\right\|^2. \end{aligned}$$

Thus, recalling (41), (42) and (43) we obtain that for all $\varepsilon, \delta > 0$

$$\text{limsup}_{n\to\infty} P\left(\left\|[M_n]_{k_n} - \langle M_n\rangle_{k_n}\right\| > 3\delta\right) \leq 4\varepsilon^2\delta^{-2}\left(\sup_{n\in\mathbb{N}}\sum_{j=1}^{k_n} E\left\|\xi_{nj}\right\|^2\right)$$

and part (a) follows in view of (37) by letting $\varepsilon \to 0$.

Part (b) follows by applying the Cramér Wold device to Corollary 3.1 of H&H.

Part (c) is a consequence of Theorem 3.4 of H&H. Applying (39) to a univariate martingale difference array $(\xi_{nj}, \mathcal{F}_{nj})$, we obtain $E\left(\max_{j\leq k_n}\xi_{nj}^2\right) \to 0$, implying that $\max_{j\leq k_n}\left|\xi_{nj}\right| = o_p(1)$. Hence, the first two assumptions of Theorem 3.4 of H&H are satisfied. Moreover, part (a) implies that $[M_n]_{k_n} - \langle M_n\rangle_{k_n} \to_p 0$ as $n \to \infty$. Thus, we can choose $u_n = \langle M_n\rangle_{k_n}$ in condition (3.28) of H&H. Denoting by $\mathcal{G}_n \vee \mathcal{F}_{nj-1}$ the σ-algebra generated by $\mathcal{G}_n \cup \mathcal{F}_{nj-1}$, the measurability condition $\mathcal{G}_n \subseteq \mathcal{F}_{n1}$ for $\langle M_n\rangle_{k_n}$ implies that

$$E_{\mathcal{G}_n\vee\mathcal{F}_{nj-1}}\left(\xi_{nj}\right) = E_{\mathcal{F}_{nj-1}}\left(\xi_{nj}\right) = 0,$$

so assumption (3.29) of H&H is satisfied. Thus, for $d = 1$ part (b) follows from Theorem 3.4 of H&H.

The vector valued case is obtained by using the Cramér Wold device. If λ is an arbitrary non zero $\mathbb{R}^d$-vector, $\tilde{\xi}_{nj} := \lambda'\xi_{nj}$ is a univariate martingale difference array satisfying

$$\begin{aligned}\sum_{j=1}^{k_n} E_{\mathcal{F}_{nj-1}}\left(\tilde{\xi}_{nj}^2\mathbf{1}\left\{\left|\tilde{\xi}_{nj}\right| > \delta\right\}\right) &\leq \|\lambda\|^2\sum_{j=1}^{k_n} E_{\mathcal{F}_{nj-1}}\left(\left\|\xi_{nj}\right\|^2\mathbf{1}\left\{\left\|\xi_{nj}\right\| > \frac{\delta}{\|\lambda\|}\right\}\right) \\ \sup_{n\in\mathbb{N}}\sum_{j=1}^{k_n} E\tilde{\xi}_{nj}^2 &\leq \|\lambda\|^2\sup_{n\in\mathbb{N}}\sum_{j=1}^{k_n} E\left\|\xi_{nj}\right\|^2.\end{aligned}$$

Letting $\tilde{M}_{nk_n} := \lambda' M_{nk_n}$, $\left\langle\tilde{M}_n\right\rangle_{k_n} = \lambda'\langle M_n\rangle_{k_n}\lambda \Rightarrow \lambda' H\lambda$ and has the required measurability property. Thus, $\tilde{M}_{nk_n} \Rightarrow \lambda' H^{1/2}Z$ from the CLT derived for the $d = 1$ case. Since λ is arbitrary, $M_{nk_n} \Rightarrow H^{1/2}Z$ follows from the Cramér Wold theorem.

It remains to show joint convergence of $\left(M_{nk_n}, \langle M_n\rangle_{k_n}\right)$ which, in view of part (a), is equivalent to joint convergence of $\left(M_{nk_n}, [M_n]_{k_n}\right)$. Since $M_{nk_n} \Rightarrow H^{1/2}Z$, a sufficient condition for joint convergence of the martingale array M_{nk_n} and its quadratic variation is given by

$$\sup_{n\in\mathbb{N}} E\left(\max_{1\leq j\leq k_n}\left\|\xi_{nj}\right\|\right) < \infty \tag{44}$$

(cf. Jacod and Shiryaev, 1987, VI, Corollary 6.7). To establish (44), we use the

Lyapunov inequality to obtain

$$\begin{aligned}\sup_{n\in\mathbb{N}} E\left(\max_{1\le j\le k_n}\left\|\xi_{nj}\right\|\right) &\le \sup_{n\in\mathbb{N}}\left[E\left\{\left(\max_{1\le j\le k_n}\left\|\xi_{nj}\right\|\right)^2\right\}\right]^{1/2} \\ &= \sup_{n\in\mathbb{N}}\left[E\left(\max_{1\le j\le k_n}\left\|\xi_{nj}\right\|^2\right)\right]^{1/2} \le B\end{aligned}$$

for some $B \in (0,\infty)$, since $E\left(\max_{1\le j\le k_n}\left\|\xi_{nj}\right\|^2\right) \to 0$ by (39). Thus, joint convergence of $\left(M_{nk_n}, [M_n]_{k_n}\right)$ follows.

Proposition A2. *There exists a constant $B \in (0,\infty)$ such that*

$$\max_{1\le k\le n} E\left\|\frac{1}{n^{1/2}}\sum_{t=k}^{n} vec\left[\tilde{\varepsilon}_t\tilde{\varepsilon}'_{t-k} - \Gamma^{\tilde{\varepsilon}}(k)\right]\right\|^2 \le 2B.$$

Proof. For each k let $X_{t,k} = vec\left[\tilde{\varepsilon}_{0t}\tilde{\varepsilon}'_{xt-k} - \Gamma^{\tilde{\varepsilon}}_{0x}(k)\right]$. For each $h \ge 0$, $X_{t,k}$ has covariance matrix

$$\begin{aligned}\Gamma_X(h,k) &= E\left(X_{t,k}X'_{t-h,k}\right) \\ &= \sum_{j_1,\dots,j_4=0}^{\infty}\left(\tilde{F}_{j_2}\otimes\tilde{F}_{j_1}\right)E\left(\varepsilon_{t-k-j_2}\varepsilon'_{t-h-k-j_4}\otimes\varepsilon_{t-j_1}\varepsilon'_{t-h-j_3}\right)\left(\tilde{F}'_{j_4}\otimes\tilde{F}'_{j_3}\right) \\ &\quad - vec\Gamma^{\tilde{\varepsilon}}(k)\left[vec\Gamma^{\tilde{\varepsilon}}(k)\right]'.\end{aligned}$$

The expectation in the above expression will be non zero when the time indices of the innovations are equal or pairwise equal. The pair $j_1 = j_2 + k$, $j_3 = j_4 + k$ gives a value of $vec\Gamma^{\tilde{\varepsilon}}(k)\left[vec\Gamma^{\tilde{\varepsilon}}(k)\right]'$ for the infinite series which cancels out with the second term of $\Gamma_X(h,k)$. Thus,

$$\Gamma_X(h,k) = \Gamma_X^{(1)}(h,k) + \Gamma_X^{(2)}(h,k) + \Gamma_X^{(3)}(h,k),$$

where

$$\Gamma_X^{(1)}(h,k) = \sum_{j_3,j_4=0}^{\infty}\left(\tilde{F}_{j_4+h}\otimes\tilde{F}_{j_3+h}\right)\left(\Sigma_\varepsilon\otimes\Sigma_\varepsilon\right)\left(\tilde{F}'_{j_4}\otimes\tilde{F}'_{j_3}\right)$$

occurs from the pair $j_1 = j_3 + h$, $j_2 = j_4 + h$,

$$\Gamma_X^{(2)}(h,k) = \sum_{j_4=0}^{\infty}\sum_{j_3=(k-h)\vee 0}^{\infty}\left(\tilde{F}_{j_3+h-k}\otimes\tilde{F}_{j_4+h+k}\right)E\left(\varepsilon_1\varepsilon'_2\otimes\varepsilon_2\varepsilon'_1\right)\left(\tilde{F}'_{j_4}\otimes\tilde{F}'_{j_3}\right)$$

occurs from the pair $j_1 = j_4 + h + k$, $j_2 = j_3 + h - k$, and

$$\Gamma_X^{(3)}(h,k) = \sum_{j_4=0}^{\infty}\left(\tilde{F}_{j_4+h}\otimes\tilde{F}_{j_4+h+k}\right)E\left(\varepsilon_1\varepsilon'_1\otimes\varepsilon_1\varepsilon'_1\right)\left(\tilde{F}'_{j_4}\otimes\tilde{F}'_{j_4+k}\right)$$

occurs from the combination $j_1 = j_2 + k = j_3 + h = j_4 + h + k$. We will show that

$$\sum_{h=0}^{\infty} \|\Gamma_X(h,k)\| \le B < \infty, \quad \text{for some } B \text{ independent of } k \tag{45}$$

by proving the corresponding result for each $\Gamma_X^{(i)}(h,k)$. For $\Gamma_X^{(1)}(h,k)$,

$$\begin{aligned}
\sum_{h=0}^{\infty} \left\|\Gamma_X^{(1)}(h,k)\right\| &\le \|\Sigma_\varepsilon\|^2 \sum_{j_3,j_4=0}^{\infty} \left\|\tilde F_{j_3}\right\| \left\|\tilde F_{j_4}\right\| \sum_{h=0}^{\infty} \left\|\tilde F_{j_4+h}\right\| \left\|\tilde F_{j_3+h}\right\| \\
&\le \|\Sigma_\varepsilon\|^2 \sum_{j_3,j_4=0}^{\infty} \left\|\tilde F_{j_3}\right\| \left\|\tilde F_{j_4}\right\| \left(\sum_{h=0}^{\infty} \left\|\tilde F_{j_4+h}\right\|^2\right)^{1/2} \left(\sum_{h=0}^{\infty} \left\|\tilde F_{j_3+h}\right\|^2\right)^{1/2} \\
&\le \|\Sigma_\varepsilon\|^2 \left(\sum_{h=0}^{\infty} \left\|\tilde F_h\right\|^2\right) \left(\sum_{j_3=0}^{\infty} \left\|\tilde F_{j_3}\right\|\right)^2 =: B_1 < \infty.
\end{aligned}$$

For $\Gamma_X^{(2)}(h,k)$, the fact that $\|E(\varepsilon_1\varepsilon_2' \otimes \varepsilon_2\varepsilon_1')\| \le E\|\varepsilon_1\|^4 < \infty$ yields

$$\begin{aligned}
\sum_{h=0}^{\infty} \left\|\Gamma_X^{(2)}(h,k)\right\| &\le E\|\varepsilon_1\|^4 \sum_{h,j_4=0}^{\infty} \left\|\tilde F_{j_4}\right\| \left\|\tilde F_{j_4+h+k}\right\| \sum_{j_3=(k-h)\vee 0}^{\infty} \left\|\tilde F_{j_3+h-k}\right\| \left\|\tilde F_{j_3}\right\| \\
&\le E\|\varepsilon_1\|^4 \sum_{h,j_4=0}^{\infty} \left\|\tilde F_{j_4}\right\| \left\|\tilde F_{j_4+h+k}\right\| \left(\sum_{j_3=(k-h)\vee 0}^{\infty} \left\|\tilde F_{j_3+h-k}\right\|^2\right)^{1/2} \\
&\quad \times \left(\sum_{j_3=0}^{\infty} \left\|\tilde F_{j_3}\right\|^2\right)^{1/2} \\
&\le E\|\varepsilon_1\|^4 \left(\sum_{j_3=0}^{\infty} \left\|\tilde F_{j_3}\right\|^2\right) \sum_{j_4=0}^{\infty} \left\|\tilde F_{j_4}\right\| \sum_{h=0}^{\infty} \left\|\tilde F_{j_4+h+k}\right\| \\
&\le E\|\varepsilon_1\|^4 \left(\sum_{j_3=0}^{\infty} \left\|\tilde F_{j_3}\right\|^2\right) \left(\sum_{j_4=0}^{\infty} \left\|\tilde F_{j_4}\right\|\right)^2 =: B_2 < \infty.
\end{aligned}$$

Finally, for $\Gamma_X^{(3)}(h,k)$ we obtain, since $h,k \geq 0$,

$$\begin{aligned}
\sum_{h=0}^{\infty} \left\| \Gamma_X^{(3)}(h,k) \right\| &\leq E\left\|\varepsilon_1\right\|^4 \sum_{j_4=0}^{\infty} \left\|\tilde{F}_{j_4+k}\right\| \left\|\tilde{F}_{j_4}\right\| \sum_{h=0}^{\infty} \left\|\tilde{F}_{j_4+h}\right\| \left\|\tilde{F}_{j_4+h+k}\right\| \\
&\leq E\left\|\varepsilon_1\right\|^4 \sum_{j_4=0}^{\infty} \left\|\tilde{F}_{j_4+k}\right\| \left\|\tilde{F}_{j_4}\right\| \left(\sum_{h=0}^{\infty} \left\|\tilde{F}_{j_4+h}\right\|^2\right)^{\frac{1}{2}} \left(\sum_{h=0}^{\infty} \left\|\tilde{F}_{j_4+h+k}\right\|^2\right)^{\frac{1}{2}} \\
&\leq E\left\|\varepsilon_1\right\|^4 \left(\sum_{h=0}^{\infty} \left\|\tilde{F}_{h}\right\|^2\right) \left(\sum_{j_4=0}^{\infty} \left\|\tilde{F}_{j_4+k}\right\|^2\right)^{\frac{1}{2}} \left(\sum_{j_4=0}^{\infty} \left\|\tilde{F}_{j_4}\right\|^2\right)^{\frac{1}{2}} \\
&\leq E\left\|\varepsilon_1\right\|^4 \left(\sum_{h=0}^{\infty} \left\|\tilde{F}_{h}\right\|^2\right)^2 =: B_3 < \infty.
\end{aligned}$$

Thus, (45) follows with $B = B_1 + B_2 + B_3$. The proposition now follows by a standard argument for stationary processes. We can write

$$\begin{aligned}
\max_{1\leq k\leq n} E\left\| \frac{1}{n^{1/2}} \sum_{t=k}^{n} X_{t,k} \right\|^2 &= tr\left\{ \frac{1}{n} \max_{1\leq k\leq n} \sum_{t,s=k}^{n} E\left(X_{s,k} X_{t,k}'\right) \right\} \\
&= tr\left\{ \frac{1}{n} \max_{1\leq k\leq n} \sum_{s=k}^{n} \sum_{h=s-n}^{s-k} E\left(X_{s,k} X_{s-h,k}'\right) \right\} \\
&\leq \frac{1}{n} \max_{1\leq k\leq n} \sum_{s=k}^{n} \sum_{h=s-n}^{s-k} \left\|\Gamma_X(h,k)\right\| \\
&\leq \max_{1\leq k\leq n} \frac{n-k+1}{n} \sum_{h=-\infty}^{\infty} \left\|\Gamma_X(h,k)\right\| \\
&\leq \max_{1\leq k\leq n} \sum_{h=-\infty}^{\infty} \left\|\Gamma_X(h,k)\right\| \leq 2B,
\end{aligned}$$

by (45), since the bounding constant B does not depend on k.

Proof of Lemma 3.1. We begin by showing that

$$\max_{1\leq t\leq n} E\left\| \frac{1}{n^{\alpha/2}} \sum_{j=1}^{t} R_n^{t-j} u_{xj} \right\|^2 \leq B \tag{46}$$

for some $B \in (0,\infty)$. Denoting by $\Gamma_{u_x}(\cdot)$ the autocovariance matrix of u_{xj} we obtain

$$E\left\| \frac{1}{n^{\alpha/2}} \sum_{j=1}^{t} R_n^{t-j} u_{xj} \right\|^2 = tr \frac{1}{n^{\alpha}} \sum_{j,i=1}^{t} R_n^{t-j} \Gamma_{u_x}(j-i) R_n^{t-i} \leq B$$

since summability of $\|\Gamma_{u_x}(\cdot)\|$ gives, for all $t \in \{1, ..., n\}$,

$$\begin{aligned}
\frac{1}{n^\alpha}\sum_{j,i=1}^{t}\left\|R_n^{t-j}\Gamma_{u_x}(j-i)R_n^{t-i}\right\| &\leq \frac{1}{n^\alpha}\sum_{j=1}^{t}\|R_n\|^{t-j}\sum_{i=1}^{t}\|R_n\|^{t-i}\|\Gamma_{u_x}(j-i)\| \\
&\leq \frac{1}{n^\alpha}\sum_{j=1}^{t}\|R_n\|^{t-j}\sum_{i=1}^{t}\|\Gamma_{u_x}(j-i)\| \\
&\leq \frac{1}{n^\alpha}\sum_{j=1}^{t}\|R_n\|^{t-j}\left(\sum_{k=-\infty}^{\infty}\|\Gamma_{u_x}(k)\|\right) \\
&\leq \frac{1}{-\max_{1\leq i\leq K} c_i}\left(\sum_{k=-\infty}^{\infty}\|\Gamma_{u_x}(k)\|\right).
\end{aligned}$$

Now (46) implies that

$$\frac{1}{n^{\alpha/2}}x_n = \frac{1}{n^{\alpha/2}}\sum_{j=1}^{n}R_n^{n-j}u_{xj} + o_p(1) = O_p(1),$$

and parts (a) and (b) follow immediately. For part (c), write

$$\begin{aligned}
\left\|vec\left(\frac{1}{n^{1+\frac{\alpha}{2}}}\sum_{t=1}^{n}\tilde{\varepsilon}_t x_{t-1}'\right)\right\| &= \frac{1}{n^{1+\frac{\alpha}{2}}}\left\|\sum_{t=1}^{n}(x_{t-1}\otimes\tilde{\varepsilon}_t)\right\| \\
&= \frac{1}{n^{1+\frac{\alpha}{2}}}\left\|\sum_{t=1}^{n}\left[\left(\sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}\right)\otimes\tilde{\varepsilon}_t\right]\right\| + o_p(1) \\
&\leq \frac{1}{n^{1+\frac{\alpha}{2}}}\sum_{t=1}^{n}\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}\right\|\|\tilde{\varepsilon}_t\| + o_p(1).
\end{aligned}$$

The Cauchy-Schwarz inequality and (46) give

$$\begin{aligned}
E\left(\frac{1}{n^{1+\frac{\alpha}{2}}}\sum_{t=1}^{n}\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}\right\|\|\tilde{\varepsilon}_t\|\right) &\leq \frac{\left(E\|\tilde{\varepsilon}_1\|^2\right)^{\frac{1}{2}}}{n^{1+\frac{\alpha}{2}}}\sum_{t=1}^{n}\left(E\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}\right\|^2\right)^{\frac{1}{2}} \\
&\leq \left(E\|\tilde{\varepsilon}_1\|^2\right)^{\frac{1}{2}}\left(\max_{t\leq n}E\left\|\frac{\sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}}{n^{\alpha/2}}\right\|^2\right)^{\frac{1}{2}} \\
&\leq \left(E\|\tilde{\varepsilon}_1\|^2\right)^{\frac{1}{2}}B^{\frac{1}{2}}.
\end{aligned}$$

For part (d), note first that

$$\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t x_{t-1}' = \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t\left(R_n^{t-1}x_0 + \sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}\right)' = o_p(1). \tag{47}$$

To see this, write for the first term of (47)

$$\frac{1}{n}\sum_{t=1}^{n} vec\left(\varepsilon_t x_0' R_n^{t-1}\right) = \left[\frac{1}{n^{1-\frac{\alpha}{2}}}\sum_{t=1}^{n}\left(R_n^{t-1}\otimes\varepsilon_t\right)\right]\frac{x_0}{n^{\alpha/2}} = o_p\left(1\right)$$

since, by independence of ε_t, we obtain

$$E\left\|\frac{1}{n^{1-\frac{\alpha}{2}}}\sum_{t=1}^{n}\left(R_n^{t-1}\otimes\varepsilon_t\right)\right\|^2 = \frac{E\left\|\varepsilon_1\right\|^2}{n^{2-\alpha}}\sum_{t=1}^{n}\left\|R_n\right\|^{2(t-1)} = O\left(\frac{1}{n^{2(1-\alpha)}}\right).$$

The second term of (47) is a martingale array, so we can write

$$\begin{aligned}
E\left\|\frac{1}{n}\sum_{t=1}^{n} vec\left[\varepsilon_t\left(\sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}\right)'\right]\right\|^2 &= \frac{1}{n^2}E\left\|\sum_{t=1}^{n}\left[\left(\sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}\right)\otimes\varepsilon_t\right]\right\|^2 \\
&= \frac{1}{n^2}\sum_{t=1}^{n}E\left(\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}\right\|^2\right)E\left\|\varepsilon_t\right\|^2 \\
&\leq \frac{E\left\|\varepsilon_1\right\|^2}{n}\max_{2\leq t\leq n}E\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}\right\|^2 = O\left(\frac{1}{n^{1-\alpha}}\right)
\end{aligned}$$

by (46). Thus, the BN decomposition and summation by parts yields

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}u_{xt}x_{t-1}' &= F_x\left(1\right)\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t x_{t-1}' - \frac{1}{n}\sum_{t=1}^{n}\Delta\tilde{\varepsilon}_{xt}x_{t-1}' \\
&= -\frac{1}{n}\sum_{t=1}^{n}\Delta\tilde{\varepsilon}_{xt}x_{t-1}' + o_p\left(1\right) = \frac{1}{n}\sum_{t=1}^{n}\tilde{\varepsilon}_{xt}\Delta x_t' + o_p\left(1\right) \\
&= \frac{1}{n}\sum_{t=1}^{n}\tilde{\varepsilon}_{xt}x_t' - \frac{1}{n}\sum_{t=1}^{n}\tilde{\varepsilon}_{xt}x_{t-1}' = \frac{1}{n^{1+\alpha}}\sum_{t=1}^{n}\tilde{\varepsilon}_{xt}x_{t-1}'C + \frac{1}{n}\sum_{t=1}^{n}\tilde{\varepsilon}_{xt}u_{xt}' \\
&= \frac{1}{n}\sum_{t=1}^{n}\tilde{\varepsilon}_{xt}u_{xt}' + O_p\left(\frac{1}{n^{\alpha/2}}\right) = \Lambda_{xx} + o_p\left(1\right),
\end{aligned}$$

by parts (b) and (c) and ergodicity of $\tilde{\varepsilon}_{xt}u_{xt}'$.

Proof of Lemma 3.2 (a) Denoting by $\Gamma_{\tilde{\varepsilon}}\left(\cdot\right)$ the autocovariance matrix of $\tilde{\varepsilon}_t$, the BN decomposition on u_t yields

$$\begin{aligned}
\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_t u_t' - \Lambda\right) &= \frac{\tilde{F}_0}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\left(\varepsilon_t\varepsilon_t' - \Sigma_\varepsilon\right)F\left(1\right)' + \frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\left(\sum_{j=1}^{\infty}\tilde{F}_j\varepsilon_{t-j}\right)\varepsilon_t' F\left(1\right)' \\
&\quad - \frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\left[\tilde{\varepsilon}_t\tilde{\varepsilon}_t' - \Gamma_{\tilde{\varepsilon}}\left(0\right)\right] + \frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\left[\tilde{\varepsilon}_t\tilde{\varepsilon}_{t-1}' - \Gamma_{\tilde{\varepsilon}}\left(1\right)\right]. \quad (48)
\end{aligned}$$

The first term of (48) is $O_p\left(n^{-\alpha/2}\right)$ from the i.i.d. CLT since $E\left\|\varepsilon_1\right\|^4 < \infty$. The third and fourth terms of (48) have order $O_p\left(n^{-\alpha/2}\right)$ from a standard CLT for sample variances and covariances of linear processes (Hall and Heyde, 1980, Theorem 6.7) since $E\left\|\varepsilon_1\right\|^4 < \infty$ and, for some $\delta \in (0, 1/2)$,

$$
\begin{aligned}
\sum_{j=1}^{\infty} j^{\frac{1}{2}} \left\|\tilde{F}_j\right\|^2 &\leq \sum_{j=1}^{\infty} j^{-1-\delta} \left(j^{\frac{3}{4}+\frac{\delta}{2}} \sum_{k=j+1}^{\infty} \left\|F_k\right\| \right)^2 \leq \sum_{j=1}^{\infty} j^{-1-\delta} \left(\sum_{k=j+1}^{\infty} k^{\frac{3}{4}+\frac{\delta}{2}} \left\|F_k\right\| \right)^2 \\
&\leq \left(\sum_{k=2}^{\infty} k \left\|F_k\right\| \right)^2 \sum_{j=1}^{\infty} j^{-1-\delta} < \infty.
\end{aligned}
$$

The second term of (48) is a martingale array, so

$$
\begin{aligned}
E\left\| \frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n} \left[\varepsilon_t \otimes \left(\sum_{j=1}^{\infty} \tilde{F}_j \varepsilon_{t-j} \right) \right] \right\|^2 &= \frac{E\left\|\varepsilon_1\right\|^2}{n^{1+\alpha}} \sum_{t=1}^{n} E\left\| \sum_{j=1}^{\infty} \tilde{F}_j \varepsilon_{t-j} \right\|^2 \\
&= \frac{\left(E\left\|\varepsilon_1\right\|^2\right)^2}{n^{\alpha}} \sum_{j=1}^{\infty} \left\|\tilde{F}_j\right\|^2 = O\left(\frac{1}{n^{\alpha}}\right).
\end{aligned}
$$

This completes the proof of the lemma.

Proof of Lemma 3.2 (b) We show the result for $\nu = 0$. The argument for $\nu = x$ is identical. Since $x_t = \sum_{j=0}^{t-1} R_n^j u_{xt-j} + R_n^t x_0$ and $x_0 = o_p\left(n^{\alpha/2}\right)$

$$
\begin{aligned}
&\sum_{t=1}^{n} \tilde{\varepsilon}_{0t} x_{t-1}' \\
&= \sum_{t=1}^{n} \tilde{\varepsilon}_{0t} \sum_{j=0}^{t-1} u_{xt-1-j}' R_n^j + O_p\left(n^{\frac{3\alpha}{2}}\right) \\
&= \sum_{t=1}^{n} \tilde{\varepsilon}_{0t} \sum_{j=0}^{t-1} \varepsilon_{t-1-j}' F_x(1)' R_n^j + \sum_{t=1}^{n} \tilde{\varepsilon}_{0t} \tilde{\varepsilon}_{xt}' + \frac{1}{n^{\alpha}} \sum_{t=1}^{n} \tilde{\varepsilon}_{0t} \sum_{j=0}^{t-1} \tilde{\varepsilon}_{xt-1-j}' R_n^j C + O_p\left(n^{\frac{3\alpha}{2}}\right) \\
&= \sum_{j=0}^{n-1} \sum_{t=j+1}^{n} \tilde{\varepsilon}_{0t} \varepsilon_{t-1-j}' F_x(1)' R_n^j + \sum_{t=1}^{n} \tilde{\varepsilon}_{0t} \tilde{\varepsilon}_{xt}' + \frac{1}{n^{\alpha}} \sum_{j=0}^{n-1} \sum_{t=j+1}^{n} \tilde{\varepsilon}_{0t} \tilde{\varepsilon}_{xt-1-j}' R_n^j C + O_p\left(n^{\frac{3\alpha}{2}}\right)
\end{aligned}
$$

using the BN decomposition and summation by parts. Centering by M_{0n} and normalising, we obtain, up to $O_p\left(n^{-1/2}\right)$,

$$
\begin{aligned}
& \frac{1}{n^{\frac{1+3\alpha}{2}}}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_{0t}x_{t-1}'-M_{0n}\right) \\
= & \frac{1}{n^{\frac{1+3\alpha}{2}}}\sum_{j=0}^{n-1}\sum_{t=j+1}^{n}\left(\tilde{\varepsilon}_{0t}\varepsilon_{t-1-j}'-\tilde{F}_{0j+1}\Sigma_{\varepsilon}\right)F_x\left(1\right)'R_n^j \\
& +\frac{1}{n^{\frac{1+3\alpha}{2}}}\sum_{t=1}^{n}\left[\tilde{\varepsilon}_{0t}\tilde{\varepsilon}_{xt}'-\Gamma_{0x}^{\tilde{\varepsilon}}\left(0\right)\right]+\frac{1}{n^{\frac{1+5\alpha}{2}}}\sum_{j=0}^{n-1}\sum_{t=j+1}^{n}\left[\tilde{\varepsilon}_{0t}\tilde{\varepsilon}_{xt-1-j}'-\Gamma_{0x}^{\tilde{\varepsilon}}\left(j+1\right)\right]R_n^jC \\
& -\frac{1}{n^{\frac{1+3\alpha}{2}}}\sum_{j=0}^{n-1}j\tilde{F}_{0j+1}\Sigma_{\varepsilon}F_x\left(1\right)'R_n^j-\frac{1}{n^{\frac{1+5\alpha}{2}}}\sum_{j=0}^{n-1}j\Gamma_{0x}^{\tilde{\varepsilon}}\left(j+1\right)R_n^jC. \qquad (49)
\end{aligned}
$$

Since $\sum_{j=1}^{n}j\left\|R_n\right\|^j=O\left(n^{2\alpha}\right)$ and $\tilde{F}_{0j}$ and $\Gamma_{0x}^{\tilde{\varepsilon}}\left(j\right)$ are l_1-summable, the fourth and fifth terms in the above expression have order $O_p\left(n^{-\frac{1-\alpha}{2}}\right)$ and $O_p\left(n^{-\frac{1+\alpha}{2}}\right)$ respectively. The second term has been shown in the proof of Lemma 3.2 (a) to satisfy a standard CLT and is, therefore, of order $O_p\left(n^{-\frac{3\alpha}{2}}\right)$. For the third term, we have

$$
\begin{aligned}
& \frac{1}{n^{\frac{1+3\alpha}{2}}}E\left\|\sum_{j=0}^{n-1}\sum_{t=j+1}^{n}\left[\tilde{\varepsilon}_{0t}\tilde{\varepsilon}_{xt-1-j}'-\Gamma_{0x}^{\tilde{\varepsilon}}\left(j+1\right)\right]R_n^j\right\| \\
\leq & \frac{1}{n^{\frac{1+3\alpha}{2}}}\sum_{j=0}^{n-1}\left\|R_n\right\|^jE\left\|\sum_{t=j+1}^{n}\left[\tilde{\varepsilon}_{0t}\tilde{\varepsilon}_{xt-1-j}'-\Gamma_{0x}^{\tilde{\varepsilon}}\left(j+1\right)\right]\right\| \\
\leq & \frac{\left(\frac{1}{n^{\alpha}}\sum_{j=0}^{n-1}\left\|R_n\right\|^j\right)}{n^{\frac{1+\alpha}{2}}}\max_{0\leq j\leq n-1}E\left\|\sum_{t=j+1}^{n}\left[\tilde{\varepsilon}_{0t}\tilde{\varepsilon}_{xt-1-j}'-\Gamma_{0x}^{\tilde{\varepsilon}}\left(j+1\right)\right]\right\| \\
\leq & \frac{\left(2B\right)^{1/2}}{-\max_{i\leq K}c_i}\frac{1}{n^{\alpha/2}}, \qquad (50)
\end{aligned}
$$

by the Lyapunov inequality $E\left\|X\right\|\leq\left(E\left\|X\right\|^2\right)^{1/2}$ and Proposition A2. This shows that the third term of (49) has order $O_p\left(n^{-3\alpha/2}\right)$. In addition, the estimation in (50) shows that the first term of (49) has order $O_p\left(n^{-\alpha/2}\right)$ since, for each t and j, the sequence ε_{t-1-j} is a restriction of the linear process $\tilde{\varepsilon}_{xt-1-j}=\sum_{k=0}^{\infty}\tilde{F}_{xk}\varepsilon_{t-1-j-k}$ for $\tilde{F}_{x0}=I_K$ and $\tilde{F}_{xk}=0$ for all $k\geq 1$.

Proof of Lemma 3.3. The array $\xi_{nt} = n^{-\frac{1+\alpha}{2}} x_{t-1} \otimes \varepsilon_t$ is a martingale difference with respect to $\mathcal{F}_{nt} := \sigma(x_0, \varepsilon_t, \varepsilon_{t-1}, ...)$. The conditional variance of ξ_{nt} is given by

$$\begin{aligned}\sum_{t=1}^{n} E_{\mathcal{F}_{nt-1}}\left(\xi_{nt}\xi_{nt}'\right) &= \frac{1}{n^{1+\alpha}} \sum_{t=1}^{n} \left(x_{t-1}x_{t-1}' \otimes E_{\mathcal{F}_{nt-1}} \varepsilon_t \varepsilon_t'\right) \\ &= \left(\frac{1}{n^{1+\alpha}} \sum_{t=1}^{n} x_{t-1}x_{t-1}'\right) \otimes \Sigma_{\varepsilon} \to_p V_{xx} \otimes \Sigma_{\varepsilon},\end{aligned}$$

as $n \to \infty$ by (7). Thus, the lemma will follow from Proposition A1 (b) provided that the Lindeberg condition

$$\sum_{t=1}^{n} E_{\mathcal{F}_{nt-1}}\left(\|\xi_{nt}\|^2 \mathbf{1}\{\|\xi_{nt}\| > \delta\}\right) \to_p 0 \quad \delta > 0 \tag{51}$$

is satisfied. Since $n^{-(1+\alpha)} \sum_{t=1}^{n} \|x_{t-1}\|^2 \to_p trV_{xx}$ and the left side of (51) is given by

$$\frac{1}{n^{1+\alpha}} \sum_{t=1}^{n} \|x_{t-1}\|^2 E_{\mathcal{F}_{nt-1}}\left(\|\varepsilon_t\|^2 \mathbf{1}\left\{\|x_{t-1}\|\|\varepsilon_t\| > \delta n^{\frac{1+\alpha}{2}}\right\}\right),$$

it follows that

$$\max_{1\le t\le n} E_{\mathcal{F}_{nt-1}}\left(\|\varepsilon_t\|^2 \mathbf{1}\left\{\|x_{t-1}\|\|\varepsilon_t\| > \delta n^{\frac{1+\alpha}{2}}\right\}\right) \to_p 0 \quad \delta > 0 \tag{52}$$

is sufficient for (51). The inequality

$$\mathbf{1}\left\{\|x_{t-1}\|\|\varepsilon_t\| > \delta n^{\frac{1+\alpha}{2}}\right\} \le \mathbf{1}\left\{\|\varepsilon_t\| > \delta^{1/2} n^{\alpha/4}\right\} + \mathbf{1}\left\{\|x_{t-1}\| > \delta^{1/2} n^{\frac{1}{2}+\frac{\alpha}{4}}\right\}$$

implies that, for all $t \in \{1, ..., n\}$,

$$\begin{aligned}& E_{\mathcal{F}_{nt-1}}\left(\|\varepsilon_t\|^2 \mathbf{1}\left\{\|x_{t-1}\|\|\varepsilon_t\| > \delta n^{\frac{1+\alpha}{2}}\right\}\right) \\ \le\ & E_{\mathcal{F}_{nt-1}}\left(\|\varepsilon_t\|^2 \mathbf{1}\left\{\|\varepsilon_t\| > \delta^{1/2} n^{\alpha/4}\right\}\right) + E_{\mathcal{F}_{nt-1}}\left(\|\varepsilon_t\|^2 \mathbf{1}\left\{\|x_{t-1}\| > \delta^{1/2} n^{\frac{1}{2}+\frac{\alpha}{4}}\right\}\right) \\ =\ & E\left(\|\varepsilon_1\|^2 \mathbf{1}\left\{\|\varepsilon_1\| > \delta^{1/2} n^{\alpha/4}\right\}\right) + \mathbf{1}\left\{\|x_{t-1}\| > \delta^{1/2} n^{\frac{1}{2}+\frac{\alpha}{4}}\right\} E\|\varepsilon_1\|^2 \\ \le\ & E\left(\|\varepsilon_1\|^2 \mathbf{1}\left\{\|\varepsilon_1\| > \delta^{1/2} n^{\alpha/4}\right\}\right) + \mathbf{1}\left\{\max_{1\le t\le n}\|x_{t-1}\| > \delta^{1/2} n^{\frac{1}{2}+\frac{\alpha}{4}}\right\} E\|\varepsilon_1\|^2.\end{aligned}$$

Integrability of $\|\varepsilon_1\|^2$ implies that $E\left(\|\varepsilon_1\|^2 \mathbf{1}\left\{\|\varepsilon_1\| > \delta^{1/2} n^{\alpha/4}\right\}\right) \to 0$ as $n \to \infty$, so

$$\frac{1}{n^{\frac{1}{2}+\frac{\alpha}{4}}} \max_{1\le t\le n} \|x_{t-1}\| \to_p 0 \tag{53}$$

is sufficient for (52). Now

$$\begin{aligned}
\frac{1}{n^{\frac{1}{2}+\frac{\alpha}{4}}}\max_{1\le t\le n}\|x_{t-1}\| &\le \frac{\|x_0\|}{n^{\frac{1}{2}+\frac{\alpha}{4}}}\max_{1\le t\le n}\|R_n\|^{t-1}+\frac{1}{n^{\frac{1+\alpha-\nu}{2}}}\max_{1\le t\le n}\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}u_{xj}\right\| \\
&\le \frac{1}{n^{\frac{1}{2}+\frac{\alpha}{4}}}\max_{1\le t\le n}\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}F_x(1)\,\varepsilon_j\right\| \\
&\quad +\frac{1}{n^{\frac{1}{2}+\frac{\alpha}{4}}}\max_{1\le t\le n}\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}\Delta\tilde{\varepsilon}_{xj}\right\|+o_p\left(\frac{1}{n^{\frac{1}{2}-\frac{\alpha}{4}}}\right). \qquad (54)
\end{aligned}$$

For the first term on the right of (54) Kolmogorov's inequality gives

$$\begin{aligned}
P\left(\max_{1\le t\le n}\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}F_x(1)\,\varepsilon_j\right\|\ge \eta n^{\frac{1}{2}+\frac{\alpha}{4}}\right) &\le \frac{1}{\eta^2 n^{1+\frac{\alpha}{2}}}E\left\|\sum_{j=1}^{n-1}R_n^{n-1-j}F_x(1)\,\varepsilon_j\right\|^2 \\
&= \frac{1}{\eta^2 n^{1+\frac{\alpha}{2}}}tr\sum_{j=1}^{n-1}R_n^{n-1-j}\Omega_{xx}R_n^{n-1-j} \\
&= tr\frac{1}{\eta^2 n^{1+\frac{\alpha}{2}}}\sum_{j=0}^{n-2}R_n^{2j}\Omega_{xx}=O\left(\frac{1}{n^{1-\frac{\alpha}{2}}}\right).
\end{aligned}$$

For the second term on the right of (54), summation by parts gives

$$\max_{1\le t\le n}\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}\Delta\tilde{\varepsilon}_{xj}\right\|\le 2\max_{1\le t\le n}\|\tilde{\varepsilon}_{xt-1}\|+\frac{\|R_n\|^{-1}\|C\|}{n^{\alpha}}\max_{1\le t\le n}\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}\tilde{\varepsilon}_{xj}\right\|.$$

Now $n^{-\frac{1}{2}-\frac{\alpha}{4}}\max_{1\le t\le n}\|\tilde{\varepsilon}_{xt-1}\|=o_p(1)$, since, by the Chebyshev inequality,

$$\begin{aligned}
P\left(\frac{1}{n^{\frac{1}{2}+\frac{\alpha}{4}}}\max_{1\le t\le n}\|\tilde{\varepsilon}_{xt-1}\|\ge\eta\right) &\le \sum_{t=1}^{n}P\left(\|\tilde{\varepsilon}_{xt-1}\|\ge\frac{1}{n^{\frac{1}{2}+\frac{\alpha}{4}}}\eta\right) \\
&\le \frac{1}{\eta^2 n^{1+\frac{\alpha}{2}}}\sum_{t=1}^{n}E\|\tilde{\varepsilon}_{x1}\|^2=O\left(\frac{1}{n^{\alpha/2}}\right).
\end{aligned}$$

Finally, since $n^{-\alpha}\sum_{j=1}^{t-1}\|R_n\|^{2(t-1-j)} \leq B$ for all t,

$$
\begin{aligned}
\frac{1}{n^{\frac{1}{2}+\frac{5\alpha}{4}}}\max_{1\leq t\leq n}\left\|\sum_{j=1}^{t-1}R_n^{t-1-j}\tilde{\varepsilon}_{xj}\right\| &\leq \frac{1}{n^{\frac{1}{2}+\frac{5\alpha}{4}}}\max_{1\leq t\leq n}\sum_{j=1}^{t-1}\|R_n\|^{t-1-j}\|\tilde{\varepsilon}_{xj}\| \\
&\leq \frac{1}{n^{\frac{1}{2}+\frac{5\alpha}{4}}}\max_{1\leq t\leq n}\left(\sum_{j=1}^{t-1}\|R_n\|^{2(t-1-j)}\right)^{1/2}\left(\sum_{j=1}^{t-1}\|\tilde{\varepsilon}_{xj}\|^2\right)^{1/2} \\
&\leq B^{1/2}\frac{1}{n^{\frac{1}{2}+2\alpha}}\max_{1\leq t\leq n}\left(\sum_{j=1}^{t-1}\|\tilde{\varepsilon}_{xj}\|^2\right)^{1/2} \\
&\leq B^{1/2}\frac{1}{n^{2\alpha}}\left(\frac{1}{n}\sum_{j=1}^{n-1}\|\tilde{\varepsilon}_{xj}\|^2\right)^{1/2} = O_p\left(\frac{1}{n^{2\alpha}}\right)
\end{aligned}
$$

by the ergodic theorem. Hence, the second term on the right of (54) is $o_p(1)$, and (53) and the lemma follow.

Proof of Lemma 3.4. An identical argument to the derivation of (9) yields

$$
\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}u_{xt}x_{t-1}' = \frac{F_x(1)}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\varepsilon_t x_{t-1}' + \frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\tilde{\varepsilon}_{xt}u_{xt}' + \frac{1}{n^{\frac{1+3\alpha}{2}}}\sum_{t=1}^{n}\tilde{\varepsilon}_{xt}x_{t-1}'C + O_p\left(n^{-1/2}\right).
$$

Centering around the asymptotic means of $\tilde{\varepsilon}_{xt}u_{xt}'$ and $\tilde{\varepsilon}_{xt}x_{t-1}'$ we obtain

$$
\begin{aligned}
\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\left(u_{xt}x_{t-1}' - \Lambda_{xx} - \frac{1}{n^{\alpha}}M_{xn}C\right) &= \frac{F_x(1)}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\varepsilon_t x_{t-1}' + \frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_{xt}u_{xt}' - \Lambda_{xx}\right) \\
&\quad + \frac{1}{n^{\frac{1+3\alpha}{2}}}\sum_{t=1}^{n}\left(\tilde{\varepsilon}_{xt}x_{t-1}' - M_{xn}\right)C \\
&= \frac{F_x(1)}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}\varepsilon_t x_{t-1}' + o_p(1)
\end{aligned}
$$

by Lemma 3.2. By Lemma 3.3, $n^{-\frac{1+\alpha}{2}}\sum_{t=1}^{n}vec\left(x_{t-1}\otimes\varepsilon_t\right) \Rightarrow N\left(0, V_{xx}\otimes\Sigma_\varepsilon\right)$, so the lemma follows from the fact that $\Omega_{xx} = F_x(1)\Sigma_\varepsilon F_x(1)'$.

Proof of Lemma 3.5. From the calculation leading to (7) we have, up to $O_p\left(n^{-\frac{1-\alpha}{2}}\right)$,

$$
\begin{aligned}
&\left[n^{\alpha}\left(I_{K^2} - R_n\otimes R_n\right)\right]\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}vec\left(\frac{x_{t-1}x_{t-1}'}{n^{\alpha}}\right) \\
= &\frac{\left(\mathcal{K}_K + I_{K^2}\right)\left(R_n\otimes I_K\right)}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}vec\left(u_{xt}x_{t-1}'\right) + \frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}vec\left(u_{xt}u_{xt}'\right). \qquad (55)
\end{aligned}
$$

By Lemma 3.4, $n^{-\frac{1+\alpha}{2}}\sum_{t=1}^{n}\left(u_{xt}x_{t-1}'-\Lambda_{xx}-\frac{1}{n^{\alpha}}M_{xn}C\right)=O_p(1)$. Also a CLT for linear processes (Phillips and Solo, 1992) yields $n^{-\frac{1+\alpha}{2}}\sum_{t=1}^{n}\left[vec\left(u_{xt}u_{xt}'\right)-\Sigma_{xx}\right]=O_p\left(n^{-\alpha/2}\right)$. Hence, in order to achieve appropriate centering we need to subtract from both sides of (55) $n^{\frac{1-\alpha}{2}}$ times

$$\begin{aligned}
& \left[n^{\alpha}\left(I_{K^2}-R_n\otimes R_n\right)\right]^{-1}\left[\left(\mathcal{K}_K+I_{K^2}\right)\left(R_n\otimes I_K\right)vec\left(\Lambda_{xx}+\frac{1}{n^{\alpha}}M_{xn}C\right)+vec\Sigma_{xx}\right] \\
= & \left[n^{\alpha}\left(I_{K^2}-R_n\otimes R_n\right)\right]^{-1}vec\left[\Omega_{xx}+\frac{1}{n^{\alpha}}\left(\Lambda_{xx}C+C\Lambda_{xx}'+R_nCM_{xn}'+M_{xn}CR_n\right)\right] \\
= & -\left(I_K\otimes C+C\otimes R_n\right)^{-1}vec\Omega_{xx}^{(n)} \\
= & vecV_{xx}^{(n)},
\end{aligned}$$

in view of definitions (13) and (14). Thus, (55) yields, up to $o_p(1)$,

$$\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}vec\left(\frac{x_{t-1}x_{t-1}'}{n^{\alpha}}-V_{xx}^{(n)}\right)=W_x\frac{1}{n^{\frac{1+\alpha}{2}}}\sum_{t=1}^{n}vec\left(u_{xt}x_{t-1}'-\Lambda_{xx}-\frac{1}{n^{\alpha}}M_{xn}C\right)$$

and the result follows directly from Lemma 3.4.

Proof of Lemma 4.1. For part (a), we first show that

$$\frac{1}{n^{\alpha/2}}\sum_{j=1}^{n}R_n^{-j}u_{xj}=\frac{1}{n^{\alpha/2}}\sum_{j=1}^{n}R_n^{-j}F_x(1)\varepsilon_j+o_p(1). \tag{56}$$

In view of the BN decomposition, $n^{-\alpha/2}\sum_{j=1}^{n}R_n^{-j}\Delta\tilde{\varepsilon}_{xj}=o_p(1)$ is sufficient for (56). Summation by parts yields

$$\begin{aligned}
\frac{1}{n^{\alpha/2}}\sum_{j=1}^{n}R_n^{-j}\Delta\tilde{\varepsilon}_{xj} &= -\frac{1}{n^{\alpha/2}}R_n^{-1}\left(I_K-R_n\right)\sum_{j=1}^{n}R_n^{-j}\tilde{\varepsilon}_{xj}+O_p\left(\frac{1}{n^{\alpha/2}}\right) \\
&= \left[1+o_p(1)\right]C\frac{1}{n^{3\alpha/2}}\sum_{j=1}^{n}R_n^{-j}\tilde{\varepsilon}_{xj}=o_p(1),
\end{aligned}$$

since $E\left\|n^{-3\alpha/2}\sum_{j=1}^{n}R_n^{-j}\tilde{\varepsilon}_{xj}\right\|\leq E\left\|\tilde{\varepsilon}_{x1}\right\|O\left(n^{-\alpha/2}\right)$.

$$E\left\|\frac{1}{n^{3\alpha/2}}\sum_{j=1}^{n}R_n^{-j}\tilde{\varepsilon}_{xj}\right\|\leq E\left\|\tilde{\varepsilon}_{x1}\right\|\frac{1}{n^{3\alpha/2}}\sum_{j=1}^{n}\left\|R_n\right\|^{-j}=O\left(\frac{1}{n^{\alpha/2}}\right).$$

Part (a) follows since the last $n-\kappa_n$ terms of the sum on the right of (56) are asymptotically negligible, viz.

$$E\left\|\frac{1}{n^{\alpha/2}}\sum_{j=\kappa_n+1}^{n}R_n^{-j}F_x(1)\varepsilon_j\right\|^2\leq\frac{\left\|F_x(1)\right\|^2E\left\|\varepsilon_1\right\|^2}{n^{\alpha}}\sum_{j=\kappa_n+1}^{n}\left\|R_n\right\|^{-2j}=O\left(\left\|R_n\right\|^{-\kappa_n}\right).$$

For part (b), letting $\xi_{nj} = n^{-\alpha/2} R_n^{-j} F_x(1) \varepsilon_j$ we obtain

$$\sum_{t=1}^{\kappa_n} E\left(\xi_{nj}\xi_{nj}'\right) = \frac{1}{n^\alpha} \sum_{j=1}^{\kappa_n} R_n^{-j} \Omega_{xx} R_n^{-j} \to \int_0^\infty e^{-pC} \Omega_{xx} e^{-pC} dp.$$

Thus, by Proposition A1 (a), part (b) follows if (36) holds. For each $\delta > 0$, letting $\eta := \delta \|R_n\| \|F_x(1)\|$ we obtain

$$\begin{aligned}
\sum_{j=1}^{\kappa_n} E\left(\left\|\xi_{nj}\right\|^2 \mathbf{1}\left\{\left\|\xi_{nj}\right\| > \delta\right\}\right) &\leq \frac{\|F_x(1)\|^2}{n^\alpha} \sum_{j=1}^{\kappa_n} \|R_n\|^{-2j} E\left(\|\varepsilon_j\|^2 \mathbf{1}\left\{\|\varepsilon_j\| > n^{\alpha/2}\eta\right\}\right) \\
&= E\left(\|\varepsilon_1\|^2 \mathbf{1}\left\{\|\varepsilon_1\| > n^{\alpha/2}\eta\right\}\right) \frac{\|F_x(1)\|^2}{n^\alpha} \sum_{j=1}^{\kappa_n} \|R_n\|^{-2j} \\
&\leq BE\left(\|\varepsilon_1\|^2 \mathbf{1}\left\{\|\varepsilon_1\| > n^{\alpha/2}\eta\right\}\right) \to 0,
\end{aligned}$$

since the ε_j are identically distributed and $E\|\varepsilon_1\|^2 < \infty$.

Proof of (18). We can write

$$\begin{aligned}
&\left\|\frac{1}{n^\alpha}\left(R_n^{-n} \otimes R_n^{-n}\right) \sum_{t=1}^{n} vec\left(u_{xt} x_{t-1}'\right)\right\| \\
= \;& \frac{\|R_n\|^{-2n}}{n^\alpha} \left\| \sum_{t=1}^{n} vec\left(u_{xt} x_0' R_n^{t-1} + u_{xt} \sum_{j=1}^{t-1} u_{xj}' R_n^{t-1-j}\right)\right\| \\
\leq \;& \frac{\|R_n\|^{-2n}}{n^\alpha} \sum_{t=1}^{n} \sum_{j=1}^{t-1} \|R_n\|^{t-1-j} \|u_{xt}\| \|u_{xj}\| + O_p\left(n^{\alpha/2} \|R_n\|^{-n}\right).
\end{aligned}$$

The first term on the right tends to 0 in L_1, since

$$\begin{aligned}
&\frac{\|R_n\|^{-2n}}{n^\alpha} \sum_{t=1}^{n} \sum_{j=1}^{t-1} \|R_n\|^{t-1-j} E\|u_{xt}\| \|u_{xj}\| \\
\leq \;& \frac{E\|u_{x1}\|^2 \|R_n\|^{-2n}}{n^\alpha} \sum_{t=1}^{n} \|R_n\|^{t-1} \sum_{j=1}^{t-1} \|R_n\|^{-j} \\
\leq \;& B \|R_n\|^{-2n} \left(\sum_{t=1}^{n} \|R_n\|^{t-1} + n\right) = O\left(n^\alpha \|R_n\|^{-n}\right),
\end{aligned}$$

by the Cauchy Schwarz inequality and the fact that $\sum_{t=1}^{n} \|R_n\|^t = O\left(n^\alpha \|R_n\|^n\right)$.

Proof of Lemma 4.2. For part (a) we have

$$\begin{aligned}
&\frac{1}{n^{\alpha}}E\left\|vec\sum_{t=1}^{n}u_{0t}\left(\sum_{j=t+1}^{n}R_n^{t-j}u_{xj}\right)'R_n^{-n}\right\| \\
\leq\;& \frac{1}{n^{\alpha}}\sum_{t=1}^{n}\sum_{j=t+1}^{n}E\left\|vecu_{0t}u_{xj}'R_n^{-n+t-j}\right\| \\
\leq\;& \frac{1}{n^{\alpha}}\sum_{t=1}^{n}\sum_{j=t+1}^{n}\left\|R_n^{-n+t-j}\otimes I_m\right\|E\left\|u_{xj}\otimes u_{0t}\right\| \\
=\;& \frac{\|R_n\|^{-n}}{n^{\alpha}}\sum_{t=1}^{n}\|R_n\|^{t}\sum_{j=t+1}^{n}\|R_n\|^{-j}E\left(\|u_{xj}\|\,\|u_{0t}\|\right) \\
\leq\;& \left\{E\|u_{x1}\|^2E\|u_{01}\|^2\right\}^{1/2}\frac{\|R_n\|^{-n}}{n^{\alpha}}\sum_{t=1}^{n}\|R_n\|^{t}\sum_{j=t+1}^{n}\|R_n\|^{-j} \\
\leq\;& B\|R_n\|^{-n}n^{1-\alpha}\sum_{k=1}^{n}\|R_n\|^{-k}=O\left(n\|R_n\|^{-n}\right)=o(1)\,.
\end{aligned}$$

For part (b) a similar calculation gives

$$\begin{aligned}
&\frac{1}{n^{\alpha}}E\left\|vec\sum_{t=1}^{\kappa_n}u_{0t}\left(\sum_{j=1}^{n}R_n^{t-j}u_{xj}\right)'R_n^{-n}\right\| \\
\leq\;& \left\{E\|u_{x1}\|^2E\|u_{01}\|^2\right\}^{1/2}\frac{\|R_n\|^{-n}}{n^{\alpha}}\sum_{t=1}^{\kappa_n}\|R_n\|^{t}\sum_{j=1}^{n}\|R_n\|^{-j} \\
\leq\;& K\|R_n\|^{-n}\sum_{t=1}^{\kappa_n}\|R_n\|^{t}=O\left(n^{\alpha}\|R_n\|^{-(n-\kappa_n)}\right)=o(1)
\end{aligned}$$

by (16).

Proof of (22). The argument is identical to the proof of (56).

Proof of (26). We start with the conditional Lindeberg condition (36). For each $\delta > 0$, since Y_{Cn} is $\mathcal{F}_{n,\kappa_n}$-measurable we obtain

$$\begin{aligned}
& \sum_{t=1}^{n-\kappa_n} E_{\mathcal{F}_{n,\kappa_n+t-1}}\left(\left\|\xi_{n,t+\kappa_n}\right\|^2 \mathbf{1}\left\{\left\|\xi_{n,t+\kappa_n}\right\| > \delta\right\}\right) \\
\leq & \max_{t \leq n-\kappa_n} E_{\mathcal{F}_{n,\kappa_n+t-1}}\left(\left\|\varepsilon_{\kappa_n+t}\right\|^2 \mathbf{1}\left\{\left\|\xi_{n,t+\kappa_n}\right\| > \delta\right\}\right) \\
& \times \frac{\left\|F_0(1)\right\|^2 \left\|Y_{Cn}\right\|^2}{n^\alpha} \sum_{t=1}^{n-\kappa_n} \left\|R_n\right\|^{-2(n-\kappa_n-t)} \\
= & O_p(1) \max_{t \leq n-\kappa_n} E_{\mathcal{F}_{n,\kappa_n+t-1}}\left(\left\|\varepsilon_{\kappa_n+t}\right\|^2 \mathbf{1}\left\{\left\|\xi_{n,t+\kappa_n}\right\| > \delta\right\}\right), \qquad (57)
\end{aligned}$$

by Lemma 4.1 (b). Now letting $\eta = \delta \left\|F_0(1)\right\|^{-1}$

$$\begin{aligned}
\mathbf{1}\left\{\left\|\xi_{n,t+\kappa_n}\right\| > \delta\right\} & \leq \mathbf{1}\left\{\left\|\varepsilon_{\kappa_n+t}\right\| \left\|Y_{Cn}\right\| > \eta n^{\alpha/2}\right\} \\
& \leq \mathbf{1}\left\{\left\|\varepsilon_{\kappa_n+t}\right\| > \eta^{1/2} n^{\alpha/4}\right\} + \mathbf{1}\left\{\left\|Y_{Cn}\right\| > \eta^{1/2} n^{\alpha/4}\right\},
\end{aligned}$$

so the right side of (57) yields for each $\delta, \eta > 0$

$$\begin{aligned}
& \max_{t \leq n-\kappa_n} E_{\mathcal{F}_{n,\kappa_n+t-1}}\left(\left\|\varepsilon_{\kappa_n+t}\right\|^2 \mathbf{1}\left\{\left\|\xi_{n,t+\kappa_n}\right\| > \delta\right\}\right) \\
\leq & E\left(\left\|\varepsilon_1\right\|^2 \mathbf{1}\left\{\left\|\varepsilon_1\right\| > \eta^{1/2} n^{\alpha/4}\right\}\right) + \mathbf{1}\left\{\left\|Y_{Cn}\right\| > \eta^{1/2} n^{\alpha/4}\right\} E\left\|\varepsilon_1\right\|^2, \qquad (58)
\end{aligned}$$

since (ε_t) is an i.i.d. sequence and Y_{Cn} is $\mathcal{F}_{n,\kappa_n}$-measurable. The first term on the right of (58) tends to 0 as $n \to \infty$ by integrability of $\left\|\varepsilon_1\right\|^2$. The second term on the right of (58) tends to 0 in L_1 since $\left\|Y_{Cn}\right\| = O_p(1)$ as $n \to \infty$. This verifies (36).

For (37), independence of Y_{Cn} and $\varepsilon_{t+\kappa_n}$ yields

$$\begin{aligned}
\sup_{n \in \mathbb{N}} \sum_{t=1}^{n-\kappa_n} E\left\|\xi_{n,t+\kappa_n}\right\|^2 & \leq \sup_{n \in \mathbb{N}} \frac{\left\|F_0(1)\right\|^2}{n^\alpha} \sum_{t=1}^{n-\kappa_n} \left\|R_n\right\|^{-2(n-\kappa_n-t)} E\left(\left\|Y_{Cn}\right\|^2 \left\|\varepsilon_{t+\kappa_n}\right\|^2\right) \\
& = \left\|F_0(1)\right\|^2 E\left(\left\|\varepsilon_1\right\|^2\right) \sup_{n \in \mathbb{N}} E\left(\left\|Y_{Cn}\right\|^2\right) \frac{1}{n^\alpha} \sum_{t=1}^{n-\kappa_n} \left\|R_n\right\|^{-2(n-\kappa_n-t)} \\
& \leq B \sup_{n \in \mathbb{N}} E\left(\left\|Y_{Cn}\right\|^2\right) < \infty.
\end{aligned}$$

Now (26) follows from Proposition A1 (c).

Proof of Theorem 4.3. In the proof of (26) above we have established that $M_{n,n-\kappa_n}$ satisfies the conditions of Proposition A1 (c). Thus, joint convergence of $M_{n,n-\kappa_n}$ and $\langle M_n \rangle_{n-\kappa_n}$ applies. The theorem now follows directly by applying (25), (26) and the continuous mapping theorem to (27).

Proof of (32). We first show that $Z_1'Z_2 = O_p\left(n^{2\alpha}\rho_n^n\right)$. Towards this end, write

$$\begin{aligned}\left\|\frac{\rho_n^{-n}}{n^{2\alpha}}Z_1'Z_2\right\| &= \frac{\rho_n^{-n}}{n^{2\alpha}}\left\|\sum_{t=1}^{n} z_{1t}z_{2t}'\right\| \\ &\leq \|H_{cn}\|\,\|H_{\perp n}\|\,\frac{\rho_n^{-n}}{n^{2\alpha}}\sum_{t=1}^{n}\rho_n^t\left\|\sum_{k=1}^{t}\rho_n^{-k}u_{xk}\right\|\left\|\sum_{j=t+1}^{n}\rho_n^{-(j-t)}u_{xj}'\right\| + o_p(1),\end{aligned}$$

and

$$\begin{aligned}&\frac{\rho_n^{-n}}{n^{2\alpha}}\sum_{t=1}^{n}\rho_n^t E\left\|\sum_{k=1}^{t}\rho_n^{-k}u_{xk}\right\|\left\|\sum_{j=t+1}^{n}\rho_n^{-(j-t)}u_{xj}'\right\| \\ \leq\; &\frac{\rho_n^{-n}}{n^{\alpha}}\sum_{t=1}^{n}\rho_n^t\left\{E\left\|\frac{1}{n^{\alpha/2}}\sum_{k=1}^{t}\rho_n^{-k}u_{xk}\right\|^2\right\}^{1/2}\left\{E\left\|\frac{1}{n^{\alpha/2}}\sum_{j=1}^{n-t}\rho_n^{-j}u_{xt+j}'\right\|^2\right\}^{1/2} \\ \leq\; &\frac{\rho_n^{-n}}{n^{\alpha}}\sum_{t=1}^{n}\rho_n^t\left\{\frac{tr\left[\sum_{k=-\infty}^{\infty}\Gamma_{xx}^u(k)\right]}{n^{\alpha}\left(\rho_n^2-1\right)}\right\} \leq B.\end{aligned}$$

Thus, since $Z_1'Z_1 = O_p\left(n^{2\alpha}\rho_n^{2n}\right)$, $\Pi_{1n} = \left(Z_1'Z_1\right)^{-1}Z_1'Z_2 = O_p\left(\rho_n^{-n}\right)$ and

$$\frac{Z_2'Q_1Z_2}{n^{1+\alpha}} = \frac{Z_2'Z_2}{n^{1+\alpha}} - \frac{Z_2'Z_1\left(Z_1'Z_1\right)^{-1}Z_1'Z_2}{n^{1+\alpha}} = \frac{Z_2'Z_2}{n^{1+\alpha}} + O_p\left(\frac{1}{n^{1-\alpha}}\right),$$

and (32) follows.

Proof of Lemma 4.5. Since $z_{2n} = 0$ and $z_{2,1} = o_p\left(n^{\alpha/2}\right)$, the recursion $z_{2t} = \rho_n^{-1}z_{2t+1} - \rho_n^{-1}H_{\perp n}'u_{xt}$ gives

$$\begin{aligned}\frac{\left(\rho_n^2-1\right)}{n}\sum_{t=1}^{n-1}z_{2t}z_{2t}' &= H_{\perp n}'\frac{1}{n}\sum_{t=1}^{n-1}u_{xt+1}u_{xt+1}'H_{\perp n} - \frac{1}{n}\sum_{t=1}^{n-1}z_{2,t+1}u_{xt+1}'H_{\perp n} \\ &\quad - H_{\perp n}'\frac{1}{n}\sum_{t=1}^{n-1}u_{xt+1}z_{2,t+1}' + o_p\left(\frac{1}{n^{1-\alpha/2}}\right). \qquad (59)\end{aligned}$$

From (30) we know that $H_{\perp n} = O_p(1)$. Thus, the ergodic theorem yields

$$H_{\perp n}'\left(\frac{1}{n}\sum_{t=1}^{n-1}u_{xt+1}u_{xt+1}'\right)H_{\perp n}' = H_{\perp n}'\Sigma_{xx}^u H_{\perp n}' + o_p(1)$$

for the first term of (59). For the second term, we obtain

$$\frac{1}{n}\sum_{t=1}^{n-1}z_{2,t+1}u_{xt+1}'H_{\perp n} = -H_{\perp n}'\left[\frac{1}{n}\sum_{t=1}^{n-1}\left(\sum_{j=1}^{n-t-1}\rho_n^{-j}u_{xt+1+j}\right)u_{xt+1}'\right]H_{\perp n}.$$

Applying a BN decomposition on the inner sum, the part containing the i.i.d. innovations ε_{t+1+j} will be a martingale difference array that converges to 0 in L_2. Thus, summation by parts yields

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n-1} z_{2,t+1}u'_{xt+1}H_{\perp n} &= -H'_{\perp n}\left[\frac{1}{n}\sum_{t=1}^{n-1}\left(\sum_{j=1}^{n-t-1}\rho_n^{-j}\Delta\tilde{\varepsilon}_{xt+1+j}\right)u'_{xt+1}\right]H_{\perp n} \\
&= \rho_n^{-1}H'_{\perp n}\left[\frac{1}{n}\sum_{t=1}^{n-1}\tilde{\varepsilon}_{xt+1}u'_{xt+1}\right]H_{\perp n} \\
&\quad -H'_{\perp n}\left[\frac{c\rho_n^{-1}}{n^{1+\alpha}}\sum_{t=1}^{n-1}\sum_{j=1}^{n-t-1}\rho_n^{-j}\tilde{\varepsilon}_{xt+1+j}u'_{xt+1}\right]H_{\perp n} \\
&= H'_{\perp n}\left[\frac{1}{n}\sum_{t=1}^{n}\tilde{\varepsilon}_{xt+1}u'_{xt+1}\right]H'_{\perp n}+O_p\left(\frac{1}{n^{\alpha/2}}\right) \\
&= H'_{\perp n}\Lambda_{xx}H'_{\perp n}+o_p(1),
\end{aligned}$$

because, by an argument identical to derivation of (46),

$$\max_{1\le t\le n-1} E\left\|\sum_{j=1}^{n-t-1}\rho_n^{-j}\tilde{\varepsilon}_{xt+1+j}\right\| = O\left(n^{\alpha/2}\right).$$

Thus, recalling that $\Omega_{xx}=\Sigma_{xx}+\Lambda_{xx}+\Lambda'_{xx}$,

$$\frac{1}{n^{1+\alpha}}\sum_{t=1}^{n} z_{2t}z'_{2t} = \frac{1}{2c}H'_{\perp n}\Omega_{xx}H_{\perp n}+o_p(1),$$

showing part (a). For part (b), (30) and the Skorohod representation theorem imply that there exist random variables $\tilde{H}_{\perp n}$ and $\tilde{Y}_c$ defined on the same probability space for all $n\in\mathbb{N}$ such that $\tilde{H}_{\perp n}=_d H_{\perp n}$, $\tilde{Y}_c=_d Y_c$ and

$$H_{\perp n}H'_{\perp n}\to_{a.s.} I_K-\frac{\tilde{Y}_c\tilde{Y}'_c}{\tilde{Y}'_c\tilde{Y}_c}.$$

Denote by M^{+} the Moore-Penrose inverse of a matrix M. Since the rank of both $\tilde{H}_{\perp n}\tilde{H}'_{\perp n}\Omega_{xx}\tilde{H}_{\perp n}\tilde{H}'_{\perp n}$ and $\left(I_K-\frac{\tilde{Y}_c\tilde{Y}'_c}{\tilde{Y}'_c\tilde{Y}_c}\right)\Omega_{xx}\left(I_K-\frac{\tilde{Y}_c\tilde{Y}'_c}{\tilde{Y}'_c\tilde{Y}_c}\right)$ is $K-1$ *a.s.*, Theorem 2 of

Andrews (1987) yields

$$\begin{aligned}
H_{\perp n}\left(H_{\perp n}'\Omega_{xx}H_{\perp n}\right)^{-1}H_{\perp n}' &= \left(H_{\perp n}H_{\perp n}'\Omega_{xx}H_{\perp n}H_{\perp n}'\right)^{+} \\
&=_d \left(\tilde{H}_{\perp n}\tilde{H}_{\perp n}'\Omega_{xx}\tilde{H}_{\perp n}\tilde{H}_{\perp n}'\right)^{+} \\
&\to_{a.s.} \left[\left(I_K - \frac{\tilde{Y}_c\tilde{Y}_c'}{\tilde{Y}_c'\tilde{Y}_c}\right)\Omega_{xx}\left(I_K - \frac{\tilde{Y}_c\tilde{Y}_c'}{\tilde{Y}_c'\tilde{Y}_c}\right)\right]^{+} \\
&=_d \left[\left(I_K - \frac{Y_cY_c'}{Y_c'Y_c}\right)\Omega_{xx}\left(I_K - \frac{Y_cY_c'}{Y_c'Y_c}\right)\right]^{+} \\
&= \left[H_{\perp}H_{\perp}'\Omega_{xx}H_{\perp}H_{\perp}'\right]^{+} \\
&= H_{\perp}\left(H_{\perp}'\Omega_{xx}H_{\perp}\right)^{-1}H_{\perp}'.
\end{aligned}$$

Proof of Lemma 4.6. For part (a), we show that $\Psi_n - \bar{\Psi}_n \to 0$ in L_2. By independence of the sequence ε_t we obtain

$$\begin{aligned}
E\left\|\Psi_n - \bar{\Psi}_n\right\|^2 &= \frac{1}{n^{1+\alpha}}E\left\|\sum_{t=1}^{\kappa_n}\varepsilon_t\sum_{j=t+1}^{n}\rho_n^{-(j-t)}\varepsilon_j'\right\|^2 = \frac{E\left\|\varepsilon_1\right\|^2}{n^{1+\alpha}}\sum_{t=1}^{\kappa_n}E\left\|\sum_{j=t+1}^{n}\rho_n^{-(j-t)}\varepsilon_j'\right\|^2 \\
&= \frac{\left(E\left\|\varepsilon_1\right\|^2\right)^2}{n^{1+\alpha}}\sum_{t=1}^{\kappa_n}\sum_{j=t+1}^{n}\rho_n^{-2(j-t)} = O\left(\frac{\kappa_n}{n}\right) = o\left(1\right),
\end{aligned}$$

by (17). For part (b), letting $\xi_{nt} = n^{-\frac{1+\alpha}{2}}\left(\sum_{j=t+1}^{n}\rho_n^{-(j-t)}\varepsilon_j\right)\otimes\varepsilon_t$, we obtain

$$\begin{aligned}
\sum_{t=1}^{n}E_{\mathcal{F}_{nt}}\left(\xi_{nt}\xi_{nt}'\right) &= \frac{1}{n^{1+\alpha}}\sum_{t=1}^{n}\left(\sum_{j=t+1}^{n}\rho_n^{-2(j-t)}\Sigma_{\varepsilon}\otimes\varepsilon_t\varepsilon_t'\right) \\
&= \frac{1}{2c}\left(\Sigma_{\varepsilon}\otimes\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t\varepsilon_t'\right) + O_p\left(\frac{1}{n^{1-\alpha}}\right) \to_p \frac{1}{2c}\left(\Sigma_{\varepsilon}\otimes\Sigma_{\varepsilon}\right).
\end{aligned}$$

Thus, we can apply the CLT of Proposition A1 (b) to the martingale difference array $(\xi_{nt}, \mathcal{F}_{nt+1})$. The inequality

$$\mathbf{1}\left\{\left\|\xi_{nt}\right\| > \delta\right\} \leq \mathbf{1}\left\{\left\|\varepsilon_t\right\| > n^{\alpha/2}\delta\right\} + \mathbf{1}\left\{\left\|\sum_{j=t+1}^{n}\rho_n^{-(j-t)}\varepsilon_j\right\| > n^{1/2}\right\}$$

yields the following sufficient conditions for the Lindeberg condition (36) for $(\xi_{nt}, \mathcal{F}_{nt+1})$:

$$R_{1n} := \frac{1}{n^{1+\alpha}}\sum_{t=1}^{n}\left\|\varepsilon_t\right\|^2\mathbf{1}\left\{\left\|\varepsilon_t\right\| > n^{\alpha/2}\delta\right\}E\left\|\sum_{j=t+1}^{n}\rho_n^{-(j-t)}\varepsilon_j\right\|^2 = o_p\left(1\right)$$

$$R_{2n} := \frac{1}{n^{1+\alpha}} \sum_{t=1}^{n} \|\varepsilon_t\|^2 E\left(\left\| \sum_{j=t+1}^{n} \rho_n^{-(j-t)} \varepsilon_j \right\|^2 \mathbf{1}\left\{ \left\| \sum_{j=t+1}^{n} \rho_n^{-(j-t)} \varepsilon_j \right\| > n^{1/2} \right\} \right) = o_p(1).$$

Since $E\left\|\sum_{j=t+1}^{n} \rho_n^{-(j-t)} \varepsilon_j\right\|^2 \leq (\rho_n^2 - 1)^{-1} tr\Sigma_\varepsilon$, the result for R_{1n} follows from the fact that

$$R_{1n} \leq B \frac{1}{n} \sum_{t=1}^{n} \|\varepsilon_t\|^2 \mathbf{1}\left\{ \|\varepsilon_t\| > n^{\alpha/2}\delta \right\} \to_{L_1} 0,$$

by integrability of $\|\varepsilon_1\|^2$. For R_{2n}, applying the Cauchy Schwarz and Chebyshev inequalities, we obtain

$$\begin{aligned}
R_{2n} &\leq \frac{1}{n^{\frac{3}{2}+\alpha}} \sum_{t=1}^{n} \|\varepsilon_t\|^2 \left\{ E\left\| \sum_{j=t+1}^{n} \rho_n^{-(j-t)} \varepsilon_j \right\|^4 \right\}^{1/2} \left\{ E\left\| \sum_{j=t+1}^{n} \rho_n^{-(j-t)} \varepsilon_j \right\|^2 \right\}^{1/2} \\
&\leq \frac{B}{n^{\frac{3}{2}+\frac{\alpha}{2}}} \sum_{t=1}^{n} \|\varepsilon_t\|^2 \left\{ E\left\| \sum_{j=t+1}^{n} \rho_n^{-(j-t)} \varepsilon_j \right\|^4 \right\}^{1/2} \\
&\leq \frac{B\left(E\|\varepsilon_1\|^4\right)^{1/2}}{n^{\frac{3}{2}+\frac{\alpha}{2}}} \sum_{t=1}^{n} \|\varepsilon_t\|^2 \left\{ \sum_{j=1}^{n-t} \rho_n^{-4j} + 3\left(\sum_{j=1}^{n-t} \rho_n^{-2j} \right)^2 \right\}^{1/2} \\
&\leq \frac{B}{n^{\frac{3}{2}+\frac{\alpha}{2}}} \sum_{t=1}^{n} \|\varepsilon_t\|^2 \left\{ \frac{1}{\rho_n^4 - 1} + \frac{3}{(\rho_n^2 - 1)^2} \right\}^{1/2} = O_p\left(\frac{1}{n^{\frac{1-\alpha}{2}}} \right).
\end{aligned}$$

For parts (c) and (d), the proof follows closely the arguments of Lemmas 3.4 and 3.5, so we do not provide the full algebraic details. Recalling that $H_{\perp n} = O_p(1)$, repeated application of the BN decomposition and summation by parts yields

$$\begin{aligned}
\frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n-1} u_{0t} z_{2t}' &= -\frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n-1} u_{0t} \left(\sum_{j=t+1}^{n} \rho_n^{-(j-t)} u_{xj}' \right) H_{\perp n} \\
&= -F_0(1) \Psi_n F_x(1)' H_{\perp n} - \frac{\rho_n^{-1}}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n-1} u_{0t} \tilde{\varepsilon}_{xt}' H_{\perp n} \\
&\quad + \frac{\rho_n^{-1} c}{n^{\frac{1+3\alpha}{2}}} \sum_{t=1}^{n-1} u_{0t} \sum_{j=t+1}^{n} \rho_n^{-(j-t)} \tilde{\varepsilon}_{xj}' H_{\perp n} + O_p\left(\frac{1}{n^{\frac{(1-\alpha)\wedge\alpha}{2}}} \right). \quad (60)
\end{aligned}$$

The first term satisfies a martingale CLT, whereas the second term satisfies

$$n^{-\frac{1+\alpha}{2}} \sum_{t=1}^{n-1} \left(u_{0t} \tilde{\varepsilon}_{xt}' - \Lambda_{x0}' \right) = O_p\left(n^{-\alpha/2} \right),$$

by a linear process CLT. For the third term, a further BN decomposition on u_{0t} gives, up to $O_p\left(n^{\alpha/2}\right)$,

$$\begin{aligned}
&\sum_{t=1}^{n-1} u_{0t} \sum_{j=t+1}^{n} \rho_n^{-(j-t)} \tilde{\varepsilon}_{xj}' \\
&= F_0(1) \sum_{j=1}^{n-1} \rho_n^{-j} \sum_{t=1}^{n-j} \varepsilon_t \tilde{\varepsilon}_{xj+t}' + \frac{c}{n^\alpha} \sum_{j=1}^{n-1} \rho_n^{-j} \sum_{t=1}^{n-j} \tilde{\varepsilon}_{0t} \tilde{\varepsilon}_{xj+t}' - \sum_{t=1}^{n-1} \tilde{\varepsilon}_{0t} \tilde{\varepsilon}_{xt+1}'. \qquad (61)
\end{aligned}$$

It is easy to see that $\bar{\Lambda}_{x0}^{(n)\prime}$ is the asymptotic mean of the random matrices on the right of (60) and (61). Thus, centering by $\bar{\Lambda}_{x0}^{(n)\prime}$ yields the required result for $\sum_{t=1}^{n-1} u_{0t} z_{2t}'$, since $\sum_{j=1}^{n-1} \rho_n^{-j} \sum_{t=1}^{n-j} \left[\tilde{\varepsilon}_{0t} \tilde{\varepsilon}_{xj+t}' - \Gamma_{x0}^{\tilde{\varepsilon}}(j)'\right] = O_p\left(n^{\frac{1}{2}+\alpha}\right)$ by an identical argument to that used in Lemma 3.2 (b). The result for $\sum_{t=1}^{n-1} u_{xt} z_{2t}'$ follows similarly, with centering $\bar{\Lambda}_{xx}^{(n)\prime}$ instead of $\bar{\Lambda}_{x0}^{(n)\prime}$. For part (d), recalling that $z_{2n} = 0$, the usual recursion gives

$$\begin{aligned}
\frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=1}^{n-1} \left[\frac{z_{2t} z_{2t}'}{n^\alpha} - H_{\perp n}' \frac{\bar{\Omega}_{xx}^{(n)}}{2c} H_{\perp n}\right] &= -\frac{1}{2c} H_{\perp n}' \frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=2}^{n-1} \left[u_{xt} z_{2t}' + \bar{\Lambda}_{xx}^{(n)\prime} H_{\perp n}\right] \\
&\quad - \frac{1}{2c} \frac{1}{n^{\frac{1+\alpha}{2}}} \sum_{t=2}^{n-1} \left[z_{2t} u_{xt}' + H_{\perp n}' \bar{\Lambda}_{xx}^{(n)}\right] H_{\perp n} + o_p(1) \\
&= \frac{1}{2c} H_{\perp n}' F_x(1) \left(\Psi_n + \Psi_n'\right) F_x(1)' H_{\perp n} + o_p(1).
\end{aligned}$$

8. References

Abadir, K.M. and J. R. Magnus (2005). *Matrix Algebra.* Econometric Exercises, vol.1. Cambridge University Press.

Anrdrews, D.W.K. (1987). "Asymptotic results for generalized Wald tests," *Econometric Theory* 3, 348-358.

Anderson, T.W. (1959), "On Asymptotic Distributions of Estimates of Parameters of Stochastic Difference Equations", *Annals of Mathematical Statistics*, 30, 676-687.

Basawa, I.V. and P.J. Brockwell (1984). "Asymptotic conditional inference for regular nonergodic models with an application to autoregressive processes," *Annals of Statistics* 12, 161–171.

Chan, N. H. and C. Z. Wei (1987). "Asymptotic inference for nearly nonstationary AR(1) processes, *Annals of Statistics* 15, 1050–1063.

Elliott, G. (1998). "1998. On the Robustness of Cointegration Methods When Regressors Almost Have Unit Roots." *Econometrica* 66, 149-158.

Hall, P. and C.C. Heyde (1980). *Martingale Limit Theory and its Application.* Academic Press.

Jacod, J. and A.N. Shiryaev (1987). *Limit Theorems for Stochastic Processes.* Springer-Verlag.

Johansen, S. (1988). "Statistical analysis of cointegrating vectors," *Journal of Economic Dynamics and Control* 12, 231-254.

Johansen, S. (1995). "Identifying restrictions of linear equations with applications to simultaneous equations and cointegration," *Journal of Econometrics* 69(1), 111-132.

Kitamura, Y. and P. C. B. Phillips (1995). "Efficient IV Estimation in Nonstationary Regression: An Overview and Simulation Study". *Econometric Theory*, Vol. 12, No. 5, December 1995, pp. 1095-1130.

Kitamura, Y. and P. C. B. Phillips (1997). "Fully Modified IV, GIVE and GMM Estimation with Possibly Nonstationary Regressors and Instruments". *Journal of Econometrics*, Vol. 80, 1997, 85-123.

Park, J. Y. and P.C.B. Phillips (1989)."Statistical Inference in Regressions with Integrated Processes: Part 1," *Econometric Theory* 4, 468-497.

Phillips, P. C. B. (1985). "The distribution of matrix quotients," *Journal of Multivariate Analysis* 16:1, 157–161.

Phillips, P. C. B. (1987). "Towards a unified asymptotic theory for autoregression," *Biometrika* 74, 535–547.

Phillips, P. C. B. (1988a). "Regression Theory for Near-Integrated Time Series," *Econometrica*, 56, 1021-1044.

Phillips, P. C. B. (1988b). "Multiple regression with integrated processes." In N. U. Prabhu, (ed.), *Statistical Inference from Stochastic Processes, Contemporary Mathematics* 80, 79–106.

Phillips, P. C. B. (1991a). "Optimal Inference in Cointegrated Systems," *Econometrica* 59, 283-306

Phillips, P. C. B. (1991b). "Error Correction and Long-Run Equilibrium in Continuous Time," *Econometrica*, 59(4), 967-980

Phillips, P. C. B. (1994). "Some exact distribution theory for maximum likelihood estimators of cointegration coefficients in error correction models," *Econometrica* 62:1, 73–94.

Phillips, P. C. B. (1995). "Fully modified least squares and vector autoregression," *Econometrica,* 63, 1023-1078.

Phillips P. C. B. (2006). "Unit root log periodogram regression", *Journal of Econometrics* (forthcoming).

Phillips, P. C. B. and B. E. Hansen (1990). "Statistical inference in instrumental variables regression with I(1) processes," *Review of Economic Studies* 57, 99–125.

Phillips, P. C. B. and M. Loretan (1991). "Estimating Long-Run Economic Equilibria," *Review of Economic Studies*, 58, 407-436.

Phillips, P. C. B. and T. Magdalinos (2007a), "Limit theory for Moderate deviations from a unit root." *Journal of Econometrics*, 136, 115-130.

Phillips, P. C. B. and T. Magdalinos (2007b), "Limit theory for Moderate deviations from a unit root under weak dependence." in G. D. A. Phillips and E. Tzavalis (Eds.) *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis.* Cambridge University Press.

Phillips, P. C. B. and T. Magdalinos (2007c), "Limit theory for explosively cointegrated systems", *Econometric Theory* (forthcoming).

Phillips, P. C. B. and T. Magdalinos (2007d), "Inference in Cointegrated Systems with Mildly Integrated Regressors" Yale University (in preparation).

Phillips, P. C. B. and V. Solo (1992), "Asymptotics for Linear Processes," *Annals of Statistics* 20, 971–1001.

Saikkonen, P. (1991). "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory* 7, 1-21.

Stock, J.H. and M.W. Watson (1993). "A simple estimator of cointegrating vectors in higher order integrated systems", *Econometrica* 61, 783-821.

White, J. S. (1958). "The limiting distribution of the serial correlation coefficient in the explosive case," *Annals of Mathematical Statistics* 29, 1188–1197.