COINTEGRATION RANK TESTS IN INFINITE ORDER VECTOR AUTOREGRESSIVE PROCESSES

MITSUHIRO ODAKI

Abstract

This paper discusses asymptotic properties of hypothesis tests to determine the cointegration ranks for vector time series that are allowable to be represented as infinite lag order vector autoregressive processes. First, it is shown that the test proposed by Saikkonen and Luukkonen (1997) (the SL test) based on long but finite lag order vector autoregression approximation as a simple extension of the well-known Johansen's likelihood ratio (LR) approach, is not on the same level of asymptotic validity as that for the Johansen's LR test in the conventional sense. Second, we propose a new test as some correction of the previous one, constructed by manipulatively altering first-step regression in the SL test, and it is established that this test tends to be more powerful performances than the SL test and possesses essentially the same asymptotic properties as the SL test under the nulls. Through some Monte Carlo experiments, we finds some examples that finite sample performances of the two tests do not sufficiently reflect their asymptotics.

1 Introduction

The number of independent cointegrating relations, called *cointegration rank*, is essential and indispensable for the model formulation, parametrization and inferences under the occurrence of cointegration, particularly in the time series systems which consist of more than 3 series and consequently the cointegration rank might be greater than 1. Conventionally, the detection of the cointegration rank has been done by Johansen's (1988, 1992, 1996) likelihood ratio (LR) tests and the methodology based on those (Johansen methodology) with a limitation that the data generating processes (DGPs) represented as finite lag order vector autoregressions (VARs) in levels need to be supposed.

Several semiparametric and nonparametric approaches to rank determination have been proposed to compensate the limitation of the Johansen's LR test. Among these, Saikkonen and Luukkonen (1997) studied applicability of the Johansen's LR test (and its methodology) to the DGPs represented as infinite lag order VARs by approximating VAR models whose lag orders are finite but “long”, following the results in Saikkonen (1992) that originated inferences based on such a finite lag order VAR approximation. For the test established by this article (hereafter referred as SL test), the lag order of the VAR considered is supposed to be not a fixed integer in the conventional sense but the one that goes to infinity at a “slower” rate as the sample size increases, and the limiting distributions for the SL test statistics under the nulls are the same as the counterparts for the conventional Johansen's LR test. In addition, Luukkonen and Saikkonen (2000) adapted the SL test to the DGPs with deterministic trends on the basis of the above methodology, and Luukkonen and Saikkonen (1999) and Qu and Perron (2007) have discussed the determination of an optimal lag order based on information criteria such as AIC or HQ in such approximations, pointing out that the SL test should be executed in the VAR attached to such an optimal lag order. However, the asymptotics of the above tests under the alternatives are not explicitly provided in past articles.

The present paper theoretically shows that when the time series system considered is cointegrated, the SL test statistics diverge at slower rate than in the conventional LR tests under the alternatives, resulting in poor asymptotic powers. As it turns out in later sections, this are derived mainly from the matter that the presence of the error correction term in the vector error correction model (VECM) form does not significantly contribute to improve the innovation in terms of the magnitude of the covariance matrix of the error vector, resulting in that the effect of the error correction term vanishes as the sample size goes to infinity. To explicitly evaluate the matter, we derive a VAR expression such that there are no error correction term and that the lag order is finite but goes to infinity at a slower rate as the sample size increases. The paper also focuses its interest upon the asymptotics of the tests under the nulls by paying attention to certain quantities that are “nuisance parameter free”, in the sense that both their well-known limiting distributions and their finite sample distributions as the data series are distributed as Gaussian are so, and asymptotic differences between the test statistics and the quantities. The convergence of the differences to zero in the presence of certain non degenerate random variables is evaluated, pointing out that the convergence rate is slower than that of the conventional one but not so severe.

Another purpose of the present paper is to propose a new test for the cointegration rank in infinite lag order VARs, so that the divergence under the alternative is achieved at faster rate and the limiting distributions and the convergence rate to nuisance parameter free quantities leading to those under the nulls are the same as those of the SL test (or the Johansen's LR test). The test proposed utilizes the structure of the SL test as much as possible in construction and partly corrects the regressant and regressors in its first-stage regression, so that the term corresponding to the first-lagged one of the differences is excluded from the list of the regressors

and as a result, the error correction term can take effect. It should be strongly noted that the divergence rate of the test under the alternatives is either equal to that of the conventional Johansen's LR test, suggesting that in large samples, the tests might be more powerful than the SL test.

Monte Carlo experiments are also executed in several particular DGPs taking VARMA(2,2) forms in order to investigate finite sample performances of the tests discussed, particularly on how the detection of the true rank value is achieved as the sample size is 150 or 250 in connection with reflection of the asymptotics. The experimental results reveal that the performances vary according to the DGP, in the sense that those are satisfactory in some DGPs but far from desirability in other DGPs. and that not only the roots of the reverse characteristic polynomial of the VAR/VMA part but also the values of the coefficient matrices of the VAR/VMA part might seriously affect the asymptotics established theoretically, partly because of the forms of the DGPs adopted that are not the simplest ones, relating this to phenomenon similar to near-integration. The test newly proposed roughly seems to perform better than the SL test for many cases when a suitable VAR lag order is chosen, but its superiority was occasionally insignificant and its performances are rather worse than that of the SL test in several DGPs, unlike the matter expected from the asymptotics, and the SL seems to be more stable through the whole DGPs. We might not assert that each of tests sufficiently reflects the asymptotics even when the sample size is 250 owing to some unsatisfactory results for several DGPs.

The paper is organized as follows. Section 2 formulates the DGP and discusses the derivation of a finite lag order VAR expression with no error correction term. The tests proposed are presented in Section 3, and asymptotics for the tests are established in Section 4. Section 5 deals with Monte Carlo experiments. The remaining issues including some concluding remarks are discussed in Section 6. The proofs of lemmas and theorems in the text are provided in the Appendix.

2 The DGP and Finite Lag Order VAR Expression

Consider an observed p-dimensional vector time series y_t expressed as

$$y_t = u_t + t\mu + \mu_0, \qquad \forall t \geq 1, \tag{1}$$

where μ and μ_0 are p-dimensional constant vectors and $\{u_t\}$ takes an infinite lag order VAR form as

$$A(L)u_t = \epsilon_t, \qquad \forall \ integer \ t,$$

where $\{\epsilon_t\}$ is iid $(0, \Lambda)$ with a $p \times p$ positive definite matrix Λ and finite fourth order cumulants, L stands for the lag operator, $A(z)$ is defined as

$$A(z) = -\alpha\beta' z + (1-z)\left(I_p - \sum_{i=1}^{\infty} H_i z^i\right),$$

with $I_{\bar{m}}$ denoting the $\bar{m} \times \bar{m}$ identity matrix, α and β of $p \times r$ column full rank matrices, where r is an integer satisfying $1 \leq r \leq p-1$, and H_i of $p \times p$ matrices. We make the following assumptions: det $A(z) \neq 0$ for $\forall \, |z| \leq 1$ except $z = 1$,

$$\sum_{i=1}^{\infty} i \parallel H_i \parallel < \infty,$$

where $\|\cdot\|$ denotes the Euclidean norm, i.e., $\|H_i\| = \{tr(H_i' H_i)\}^{1/2}$, implying that $H_i = O(i^{-1-\bar{a}})$ for sufficiently large i, where $\bar{a} > 1$, μ is spanned by columns of γ, where γ is a $p \times (p-r)$ column full rank matrix such that $\beta'\gamma = 0$, $y_0 = O_p(1)$ or $O(1)$ as an initial condition for y_t and $rank \ \delta' H(1)\gamma = p - r$, where δ is a $p \times (p-r)$ column full rank matrix such that $\delta'\alpha = 0$ and $\delta'\Lambda\delta = I_{p-r}$, and $H(z) = I_p - \sum_{i=1}^{\infty} H_i z^i$, noting that these assumptions except the last one are generally made in the literature dealing with $I(1)$ and $I(0)$ series or cointegration (e.g., Saikkonen (1992), Phillips (1995) or Saikkonen and Luukkonen (1997)) and that the last one is required to exclude the occurrence of higher order cointegration in the definition of cointegration by Engle and Granger (1987) and is exactly the same as Assumption $A3$ of Banerjee et al. (1993, p. 147) or the condition stated by Equation (4.5) of Johansen (1996, p. 49). [1]

Now, define

$$C(z) = [\beta, \ \gamma] \begin{bmatrix} (1-z)I_r & 0 \\ 0 & I_{p-r} \end{bmatrix}$$

$$\begin{bmatrix} \alpha' A(z)\beta & \alpha' H(z)\gamma \\ (1-z)\delta' H(z)\beta & \delta' H(z)\gamma \end{bmatrix}^{-1} \begin{bmatrix} \alpha' \\ \delta' \end{bmatrix}$$

$$= \sum_{i=0}^{\infty} C_i z^i,$$

$$C^{(1)}(z) = \sum_{i=0}^{\infty} C_i^{(1)} z^i,$$

$$with \ C_i^{(1)} = -\sum_{h=i+1}^{\infty} C_h, \quad \forall i \geq 1,$$

$$\mu = C(1)\left(-\alpha\beta' + H(1)\gamma\delta'\right)\bar{\mu}_0.$$

It is easy to see that $C_0 = I_p$, that det $C(z) \neq 0$ for $\forall \, |z| \leq 1$ except $z = 1$, that $C(1) = \gamma\delta'$, where γ and δ are some matrices satisfying the requirements given above for the ones expressed by the notations, and that $C_i = O(i^{-1-\bar{a}})$ and $C_i^{(1)} = O(i^{-\bar{a}})$ for sufficiently large i, and based on these results or essentially owing to the Granger representation theorem by Engle and Granger (1987) (e.g., Johansen (1996, p. 49), Theorem 4.2), we can convert (1) into an infinite order VMA expression for Δy_t:

$$\Delta y_t = C(L)\epsilon_t + \mu$$

$$= C(1)\epsilon_t + (1-L)C^{(1)}(L)\epsilon_t + \mu. \tag{2}$$

We now note that y_t is cointegrated with $d = b = 1$ in the definition of cointegration by Engle and Granger (1987) and with the cointegration rank r. For discussion later, for the case $\mu \neq 0$ and $r < p-1$, partition γ into a p-dimensional vector γ_1 and a $p \times (p-1-r)$ matrix γ_2, as $\gamma = [\gamma_1, \ \gamma_2]$, and without losing generality, suppose that $\gamma_1'\mu \neq 0$ and $\gamma_2'\mu = 0$.

Now, let $\{K_T\}$ and $\{L_T\}$ be sequences of positive integers such that $\lim_{T\to\infty} K_T = \infty$ and $\lim_{T\to\infty} K_T/T^{1/3} = 0$ and $K_T/m \geq L_T > K_T/m - 1$, where m is an integer either greater than or equal to 2, and further for $h = 0$, 1, L_T and $s = 0, 1, \ldots, L_T$, define

$$\zeta_{t-s-K_T-h-1;K_T-h}$$

$$= \sum_{i=K_T-h+1}^{t-s-2} H_i(\Delta y_{t-s-i} - \mu) + \sum_{i=0}^{\infty} H_{t-s-1+i}\breve{u}_{-i+1},$$

$$if \ t \geq K_T - h + 3 + s,$$

$$=\sum_{i=0}^{\infty} H_{t-s-1+i}\breve{u}_{-i+1}, \quad if\ t=K_T-h+2+s,$$

where $\breve{u}_{-i+1}=\sum_{h=0}^{\infty} C_h\,\varepsilon_{-i+1-h}$. It is well-known that (2) can be converted into an infinite order VAR or VECM representation (e.g., Johansen (1996, p. 55), Theorem 4.5), and as done in Saikkonen (1992) or Saikkonen and Luukkonen (1997), the representation can be further approximated by a K_T-th order VAR: for $t\geq K_T+2$,

$$\Delta y_t - \mu = \alpha(\beta' y_{t-1} - E\beta' y_{t-1}) + \sum_{i=1}^{K_T} H_i(\Delta y_{t-i} - \mu) + \epsilon_t + \zeta_{t-K_T-1;K_T}, \qquad (3)$$

where $\zeta_{t-K_T-1;K_T}$ is of $O_p\,(K_T^{-1/2-\bar{a}})$. We next set up that there exists a VAR scheme comprising K_T pieces of lagged differences (i.e., Δy_{t-i}, $i=1, \ldots, K_T$) only, such that the variances of the error terms except for ε_t converge to zero as T tends to ∞, similarly to those for (3).

Lemma 1 *For y_t generated by* (1)*, we have*

$$\Delta y_t - \mu = \sum_{i=1}^{K_T} G_{i;K_T;L_T}(\Delta y_{t-i} - \mu) + \epsilon_t - L_T^{-1}(\sum_{s=1}^{L_T}\epsilon_{t-s}) + \bar{\zeta}_{t-K_T+L_T-1;K_T-L_T}, \quad \forall t\geq K_T+2, \quad (4)$$

where $G_{i;K_T;L_T}$ are $p\times p$ matrices satisfying

$$I_p - \sum_{i=1}^{K_T} G_{i;K_T;L_T} z^i = \left(\sum_{s=0}^{L_T-1}\frac{L_T-s}{L_T}z^s\right)\{-\alpha\beta' z + (1-z) \bullet \left(I_p - \sum_{i=1}^{K_T-L_T} H_i z^i\right)\}$$

and

$$\bar{\zeta}_{t-K_T+L_T-1;K_T-L_T} = \zeta_{t-K_T+L_T-1;K_T-L_T} - L_T^{-1}\sum_{s=1}^{L_T}\zeta_{t-s-K_T+L_T-1;K_T-L_T},$$

with the property that $\bar{\zeta}_{t-K_T+L_T-1;K_T-L_T}=O_p(K_T^{-1/2-\bar{a}})$.

Now, define $v_t=\sum_{i=0}^{\infty} C_i^{(1)}\varepsilon_{t-i}$ and $\bar{\alpha}=\alpha(\alpha'\alpha)^{-1}$, noting that $\beta' y_{t-1}=\beta' v_{t-1}+\beta'(y_0-v_0)$, therefore, $E\beta' y_{t-1}=E\beta' y_0$. Then, by comparing (3) and (4) after multiplication of the both sides of those equations by $\bar{\alpha}'$, we obtain

$$\beta' v_{t-1} = \sum_{i=1}^{K_T}\bar{\alpha}'(G_{i;K_T;L_T} - H_i)(\Delta y_{t-i} - \mu) - L_T^{-1}\bar{\alpha}'(\sum_{s=1}^{L_T}\epsilon_{t-s}) - \beta'(y_0 - v_0 - Ey_0) + \bar{\alpha}'(\bar{\zeta}_{t-K_T+L_K-1;K_T-L_T} - \zeta_{t-K_T-1;K_T}). \quad (5)$$

(5) naturally suggests that residuals obtained by regressing $\beta' y_{t-1}$ on $\Delta y_{t-1}, \ldots, \Delta y_{t-K_T}$, and 1 are dominated by $L_T^{-1}\,\bar{\alpha}'(\sum_{s=1}^{L_T}\varepsilon_{t-s})$ of $O_p(K_T^{-1/2})$, though it will be formally established in discussion to prove the theorems stated later, and consequently this might lead to deterioration of the asymptotic power of the SL test. It will be also seen from (4) or (5) that an innovation squares error or a linear least squares prediction error in an infinite order VAR/VECM is not improved by the presence of its error correction term.

3 Test Statistics

First, define

$$\Delta z_{t-1} = \left(\Delta y'_{t-1}, \cdots, \Delta y'_{t-K_T}\right)',$$
$$\Delta z_{t-2;-} = \left(\Delta y'_{t-2}, \cdots, \Delta y'_{t-K_T}\right)',$$
$$z_{t-2;-} = \left(\Delta y'_{t-2}, \cdots, \Delta y'_{t-K_T},\ y'_{t-1}\right)',$$
$$\bar{y}_{-1-i} = T^{-1}\sum_{t=K_T+2}^{T} y_{t-1-i},$$
$$\overline{\Delta y}_{-i} = T^{-1}\sum_{t=K_T+2}^{T}\Delta y_{t-i}, \quad i=0,1,$$
$$\overline{\Delta z} = T^{-1}\sum_{t=K_T+2}^{T}\Delta z_{t-1},$$
$$\overline{\Delta z}_{-1} = T^{-1}\sum_{t=K_T+2}^{T}\Delta z_{t-2;-},$$
$$\bar{z}_{-1} = T^{-1}\sum_{t=K_T+2}^{T} z_{t-2;-},$$

in order to construct some statistics. Second, based these, demeaned sample matrix products are constructed as follows:

$$\hat{M}_{11} = T^{-1} \sum_{t=K_T+2}^{T} y_{t-1} y'_{t-1} - (T-K_T-1)^{-1} T\, \overline{y}_{-1}\, \overline{y}'_{-1},$$

$$\hat{M}_{01} = T^{-1} \sum_{t=K_T+2}^{T} \Delta y_t y'_{t-1} - (T-K_T-1)^{-1} T\, \overline{\Delta y}_0\, \overline{y}'_{-1},$$

$$\hat{M}_{00} = T^{-1} \sum_{t=K_T+2}^{T} \Delta y_t \Delta y'_t - (T-K_T-1)^{-1} T\, \overline{\Delta y}_0\, \overline{\Delta y}'_0,$$

$$\hat{M}_{22} = T^{-1} \sum_{t=K_T+2}^{T} \Delta z_{t-1} \Delta z'_{t-1} - (T-K_T-1)^{-1} T\, \overline{\Delta z}\, \overline{\Delta z}',$$

$$\hat{M}_{02} = T^{-1} \sum_{t=K_T+2}^{T} \Delta y_t \Delta z'_{t-1} - (T-K_T-1)^{-1} T\, \overline{\Delta y}_0\, \overline{\Delta z},$$

$$\hat{M}_{12} = T^{-1} \sum_{t=K_T+2}^{T} y_{t-1} \Delta z'_{t-1} - (T-K_T-1)^{-1} T\, \overline{y}_{-1}\, \overline{\Delta z}',$$

with $\hat{M}_{10} = \hat{M}'_{01}$, $\hat{M}_{20} = \hat{M}'_{02}$ and $\hat{M}_{21} = \hat{M}'_{12}$,

$$\hat{M}_{00;ij} = T^{-1} \sum_{t=K_T+2}^{T} \Delta y_{t-i} \Delta y'_{t-j} - (T-K_T-1)^{-1} T\, \overline{\Delta y}_{-i}\, \overline{\Delta y}'_{-j}, \quad i,j = 0,1,$$

$$\hat{M}_{01;1} = T^{-1} \sum_{t=K_T+2}^{T} \Delta y_{t-1} y'_{t-1} - (T-K_T-1)^{-1} T\, \overline{\Delta y}_{-1}\, \overline{y}'_{-1},$$

$$\hat{M}_{33} = T^{-1} \sum_{t=K_T+2}^{T} \Delta z_{t-2;-} \Delta z'_{t-2;-} - (T-K_T-1)^{-1} T\, \overline{\Delta z}_{-1}\, \overline{\Delta z}'_{-1},$$

$$\hat{M}_{13} = T^{-1} \sum_{t=K_T+2}^{T} y_{t-1} \Delta z'_{t-2;-} - (T-K_T-1)^{-1} T\, \overline{y}_{-1}\, \overline{\Delta z}'_{-1},$$

$$\hat{M}_{03;i} = T^{-1} \sum_{t=K_T+2}^{T} \Delta y_{t-i} \Delta z'_{t-2;-} - (T-K_T-1)^{-1} T\, \overline{\Delta y}_{-i}\, \overline{\Delta z}'_{-1}, \quad i = 0,1,$$

$$\hat{M}_{44} = T^{-1} \sum_{t=K_T+2}^{T} z_{t-2;-} z'_{t-2;-} - (T-K_T-1)^{-1} T\, \overline{z}_{-1}\, \overline{z}'_{-1},$$

$$\hat{M}_{04;i} = T^{-1} \sum_{t=K_T+2}^{T} \Delta y_{t-i} z'_{t-2;-} - (T-K_T-1)^{-1} T\, \overline{\Delta y}_{-i}\, \overline{z}'_{-1}, \; i = 0,1,$$

with $\hat{M}_{00;ji} = \hat{M}'_{00;ij}$, $\hat{M}_{10;1} = \hat{M}'_{01;1}$, $\hat{M}_{31} = \hat{M}'_{13}$, $\hat{M}_{30;i} = \hat{M}'_{03;i}$ and $\hat{M}_{40;i} = \hat{M}'_{04;i}$, and noting that $\hat{M}_{00;00} = \hat{M}_{00}$. Third, define

$$\begin{aligned}
\tilde{S}_{11} &= \hat{M}_{11} - \hat{M}_{12}\hat{M}_{22}^{-1}\hat{M}_{21}, \\
\tilde{S}_{01} &= \hat{M}_{01} - \hat{M}_{02}\hat{M}_{22}^{-1}\hat{M}_{21}, \\
\tilde{S}_{00} &= \hat{M}_{00} - \hat{M}_{02}\hat{M}_{22}^{-1}\hat{M}_{20}, \\
\hat{S}_{11} &= \hat{M}_{11} - \hat{M}_{13}\hat{M}_{33}^{-1}\hat{M}_{31}, \\
\hat{S}_{00;d} &= \hat{M}_{00;01} - \hat{M}_{04;0}\hat{M}_{44}^{-1}\hat{M}_{40;1}, \\
\hat{S}_{00;e} &= \hat{M}_{00;11} - \hat{M}_{04;1}\hat{M}_{44}^{-1}\hat{M}_{40;1}, \\
\hat{H}_1 &= \hat{S}_{00;d}\hat{S}_{00;e}^{-1},
\end{aligned}$$

with $\tilde{S}_{10} = \tilde{S}'_{01}$, and based on $\hat{H}_1$, further define

$$\begin{aligned}
\hat{M}_{01;H} &= \hat{M}_{01} - \hat{H}_1\hat{M}_{01;1}, \\
\hat{M}_{03;H} &= \hat{M}_{03;0} - \hat{H}_1\hat{M}_{03;1}, \\
\hat{M}_{30;H} &= \hat{M}'_{03;H}, \\
\hat{M}_{00;H} &= \hat{M}_{00} - \hat{H}_1\hat{M}_{00;10} \\
&\quad - \hat{M}_{00;01}\hat{H}'_1 + \hat{H}_1\hat{M}_{00;11}\hat{H}'_1, \\
\hat{S}_{01} &= \hat{M}_{01;H} - \hat{M}_{03;H}\hat{M}_{33}^{-1}\hat{M}_{31;0}, \\
\hat{S}_{00} &= \hat{M}_{00;H} - \hat{M}_{03;H}\hat{M}_{33}^{-1}\hat{M}_{30;H},
\end{aligned}$$

with $\hat{S}_{10} = \hat{S}'_{01}$. Fourthly, let $\tilde{\lambda}_1 \geq ... \geq \tilde{\lambda}_p$ be the ordered eigenvalues of $\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}\tilde{S}_{01}$, and similarly $\hat{\lambda}_1 \geq ... \geq \hat{\lambda}_p$ be the ordered eigenvalues of $\hat{S}_{11}^{-1}\hat{S}_{10}\hat{S}_{00}^{-1}\hat{S}_{01}$. Now, for $j = 0, 1, ..., p-1$, two statistics to test the null $r = j$ and the alternative $r \geq j+1$, are given as

$$\hat{S}_j = -\sum_{h=j+1}^{p} T\,log(1 - \tilde{\lambda}_h),$$

$$\hat{Q}_j = -\sum_{h=j+1}^{p} T\,log(1 - \hat{\lambda}_h).$$

Note that $\hat{S}_j$ corresponds to the SL test, and another is newly proposed by the present paper and hereafter is referred as "CSL test", implying "a corrected SL test".

The CSL test is motivated by the matter that a consistent estimator of H_1 in (1)/(3) is utilized for some correction of the first-stage regression as in the SL test, and for this purpose, we first obtain $\hat{H}_1$ as an estimated coefficient matrix of Δy_{t-1} in a regression equation of Δy_t on $\Delta y_{t-1}, ..., \Delta y_{t-K_T}, y_{t-1}$, and 1, and then $\hat{S}_{ij}$ are obtained based on residuals obtained by regressing of $\Delta y_t - \hat{H}_1\Delta y_{t-1}$ or y_{t-1} on $\Delta y_{t-2}, ..., \Delta y_{t-K_T}$, and 1. It will be shown in the proof of Theorem 2 in the following section that the residuals obtained by regressing $\beta' y_{t-1}$ on $\Delta y_{t-2}, \ldots, \Delta y_{t-K_T}$, and 1 are of $O_p(1)$ without being degenerate, resulting in asymptotic powers that are as good as that of the Johansen's LR test. We also see later that the convergence rate of $\hat{H}_1$ to H_1 plays a decisive role for the asymptotics of the test.

4 Asymptotics

To formulate the distributions of the test statistics, define the $(p-r)$-dimensional vector/scalar ξ_{t-1} as $\sum_{h=1}^{t-1} \delta' \varepsilon_h$ for the case $\mu = 0$, as $((t-1), \sum_{h=1}^{t-1} \varepsilon_h' \bar{\delta})'$ for the case $\mu \neq 0$ and $r < p-1$, where $\bar{\delta} = \delta \gamma' \gamma_2 (\gamma_2' \gamma \gamma' \gamma_2)^{-1/2}$, and as $t-1$ for the case $\mu \neq 0$ and $r = p-1$, let $\bar{\xi}_{-1} = T^{-1} \sum_{t=K_T+2}^{T} \xi_{t-1}$ and based on ξ_{t-1} and $\bar{\xi}_{-1}$, further define

$$
\begin{aligned}
\bar{N}_{11} &= T^{-2} \sum_{t=K_T+2}^{T} \xi_{t-1}\xi_{t-1}' \\
&\quad - (T-K_T-1)^{-1}\bar{\xi}_{-1}\bar{\xi}_{-1}', \\
\bar{N}_{01} &= T^{-1} \sum_{t=K_T+2}^{T} \epsilon_t \xi_{t-1}' \\
&\quad - (T-K_T-1)^{-1} T (T^{-1} \sum_{t=K_T+2}^{T} \epsilon_t) \bar{\xi}_{-1}'.
\end{aligned}
$$

Note that the quantity $tr\, \bar{N}_{11}^{-1} \bar{N}_{01}' \delta\delta' \bar{N}_{01}$ is nuisance parameter free in its limiting distribution, as seen from the theorem stated below, and in its finite sample one as ε_t are distributed as Gaussian. In addition to the above notations, let the conventional symbols $\Rightarrow$ and $W_m(u)$ stand for weak convergence of probability measures on the unit interval [0, 1] and an m-dimensional standard Brownian motion on [0, 1] respectively. Also, define $\bar{W}_m(u) = W_m(u) - \int_0^1 W_m(u)\, du$, $\breve{W}_{p-r}(u)$ as $\bar{W}_{p-r}(u)$ for the case $\mu = 0$, as $(u - 1/2, \bar{W}_{p-r-1}')'$ for the case $\mu \neq 0$ and $r < p-1$ and as $u - 1/2$ for the case $\mu \neq 0$ and $r = p-1$, and

$$
\begin{aligned}
&\bar{W}(p-r) \\
= &\left(\int_0^1 dW_{p-r}(u) \breve{W}_{p-r}'(u) \right) \\
&\bullet \left(\int_0^1 \breve{W}_{p-r}(u) \breve{W}_{p-r}'(u) du \right)^{-1} \\
&\bullet \left(\int_0^1 \breve{W}_{p-r}(u) dW_{p-r}(u)' \right).
\end{aligned}
$$

Furthermore, define the $(p-r) \times (p-r)$ matrices/scalars D_T^{-1} and $\bar{D}_T^{-1}$, as I_{p-r} and $(\gamma'\gamma)^{-1}$ for the case $\mu = 0$, as

$$
\begin{bmatrix} T^{-1/2} & 0 \\ 0 & I_{p-r-1} \end{bmatrix} \quad and \quad \begin{bmatrix} T^{-1/2}(\mu'\gamma_1)^{-1} & 0 \\ 0 & (\gamma_2'\gamma\gamma'\gamma_2)^{-1/2} \end{bmatrix}
$$

for the case $\mu \neq 0$ and $r < p-1$ and as $T^{-1/2}$ and $T^{-1/2} (\mu'\gamma)^{-1}$ for the case $\mu \neq 0$ and $r = p-1$, and define

$$
\begin{aligned}
\hat{d} &= K_T^{-1/2} T^{-1/2} \sum_{t=K_T+2}^{T} \left(\sum_{s=1}^{K_T} \epsilon_{t-s} \right) \epsilon_t', \\
\breve{B} &= (\alpha - \delta\delta'\Lambda\alpha)(\alpha'\Lambda\alpha)^{-1}\alpha'\hat{d}\delta\delta', \\
&\qquad \bar{S}_{11;*} = D^{-1}\bar{N}_{01}' + \bar{D}_T^{-1\prime}\gamma' C(1)\Lambda, \\
\tilde{\theta} &= tr\left(D_T^{-1}\bar{N}_{11}D_T^{-1} \right)^{-1} \\
&\quad \bullet \{ D_T^{-1}\bar{N}_{01}' \left(\breve{B} + \breve{B}' \right) \bar{N}_{01} D_T^{-1} \\
&\quad - D_T^{-1}\bar{N}_{01}' \left(\breve{B} + \breve{B}' \right) \alpha\bar{\alpha}'\bar{S}_{11;*}' \\
&\quad - \bar{S}_{11;*}\bar{\alpha}\alpha' \left(\breve{B} + \breve{B}' \right) \bar{N}_{01} D_T^{-1} \}, \\
\bar{S}_{11;**} &= \bar{S}_{11;*} - \bar{D}_T^{-1\prime}\gamma'\{I_p - C(1)\}\Lambda\beta\alpha', \\
\bar{S}_{**} &= \bar{N}_{01}D_T^{-1} + (I_p - \alpha\beta') \\
&\quad \bullet \Lambda\{I_p - C(1)'\}\gamma\bar{D}_T^{-1} - \alpha\bar{\alpha}'\bar{S}_{11;**}', \\
\tilde{C} &= \hat{d}'\bar{\alpha}(\bar{\alpha}'\Lambda\bar{\alpha})^{-1}\beta', \qquad \hat{C} = -\beta'\Lambda\tilde{C}, \\
\hat{F} &= \tilde{C}\bar{S}_{**}, \\
\hat{D} &= -\tilde{C}\hat{M}_{00;10} - \hat{M}_{00;01}\tilde{C}' + \tilde{C}\hat{M}_{00;11}H_1' \\
&\quad + H_1\hat{M}_{00;11}\tilde{C}' + \tilde{C}\hat{M}_{03;1}\hat{M}_{33}^{-1}(\hat{M}_{30;0} \\
&\quad - \hat{M}_{30;1}H_1') + (\hat{M}_{03;0} - H_1\hat{M}_{03;1}) \\
&\quad \bullet \hat{M}_{33}^{-1}\hat{M}_{30;1}\tilde{C}', \\
\breve{G} &= (\alpha - \delta\delta'\Lambda\alpha)(\alpha'\alpha)^{-1}(\beta'\Lambda\beta)^{-1}\hat{C}\delta\delta', \\
\hat{\theta} &= tr\left(D_T^{-1}\bar{N}_{11}D_T^{-1} \right)^{-1} \{ D_T^{-1}\bar{N}_{01}' \left(\breve{G} + \breve{G}' \right) \\
&\quad \bar{N}_{01}D_T^{-1} - D_T^{-1}\bar{N}_{01}' \left(\breve{G} + \breve{G}' \right) \alpha\bar{\alpha}'\bar{S}_{11;**}' \\
&\quad - \bar{S}_{11;**}\bar{\alpha}\alpha' \left(\breve{G} + \breve{G}' \right) \bar{N}_{01} D_T^{-1} \\
&\quad - D_T^{-1}\bar{N}_{01}'\delta\delta'\hat{D}\delta\delta'\bar{N}_{01}D_T^{-1} \\
&\quad + D_T^{-1}\bar{N}_{01}'\delta\delta'\hat{F}D_T^{-1} \\
&\quad + D_T^{-1}\hat{F}'\delta\delta'\bar{N}_{01}D_T^{-1} \}.
\end{aligned}
$$

We shall now state the asymptotics of the test statistics in the previous section.

Theorem 1 *Suppose that y_t is generated by* (1). *Then, we have*

$$T^{-1}K_T\hat{S}_j = \sum_{h=j+1}^{r} \nu_h + o_p(1), \quad for\ \forall j < r, \quad (6)$$

where $\nu_1 \geq ... \geq \nu_r$ *are the ordered eigenvalues of* $\alpha' \Lambda^{-1} \alpha\, \bar{\alpha}' \Lambda \bar{\alpha}$,

$$\hat{S}_r = tr\bar{N}_{11}^{-1}\bar{N}_{01}'\delta\delta'\bar{N}_{01} + T^{-1/2}K_T^{1/2}\tilde{\theta} + o_p(T^{-1/2}K_T^{1/2}) + O_p(K_T^{-\bar{a}}), \quad (7)$$

with the properties that $\tilde{\theta} = O_p(1)$ *and* $\tilde{\theta}^{-1} = O_p(1)$, *and*

$$tr\bar{N}_{11}^{-1}\bar{N}_{01}'\delta\delta'\bar{N}_{01} \Rightarrow tr\bar{W}(p-r). \quad (8)$$

We see from (6) of the above theorem that $\hat{S}_j = O_p(TK_T^{-1})$ under the alternative $r > j$, implying that the divergence rate TK_T^{-1} is slower than the conventional T, and therefore the SL test might be less powerful. On the other hand, (7) formulates an asymptotic difference between the SL test statistic under the null $j = r$ and the nuisance parameter free quantity, characterized by the well-known limiting distribution stated by (8), which correspond to both (6.21) and (6.32) of Johansen (1996, pp. 94 and 98). As indicated by (7), the difference is greatly affected by the size of K_T characterizing the VAR approximation. It will be stated in the proof of Theorem 1 that the term expressed as $O_p(K_T^{-\bar{a}})$ is related to a quantity constructed based on $\zeta_{t-K_T-1;K_T}$ or $\bar{\zeta}_{t-K_T+L_T-1;K_T-L_T}$. It should be noted that if $\bar{a}$ is so large that $\lim_{T\to\infty} T^{1/2}/K_T^{\bar{a}+1/2} = 0$ holds, this term can be included in the term expressed as $o_p(T^{-1/2}K_T^{1/2})$, emphasizing that this is valid for the case in which H_i decay exponentially, such as $A(z)$ taking a rational or VARMA form.

Theorem 2 *Suppose that* y_t *is generated by* (1). *Then, we have*

$$T^{-1}\hat{Q}_j = \sum_{h=j+1}^{r} \bar{\nu}_h + o_p(1), \quad for\ \forall j < r, \quad (9)$$

where $\bar{\nu}_1 \geq ... \geq \bar{\nu}_r$ *are the ordered eigenvalues of* $\alpha' (\Lambda + \alpha\beta'\Lambda\beta\alpha')^{-1} \alpha\beta'\Lambda\beta$,

$$\hat{Q}_r = tr\bar{N}_{11}^{-1}\bar{N}_{01}'\delta\delta'\bar{N}_{01} + T^{-1/2}K_T^{1/2}\hat{\theta} + o_p(T^{-1/2}K_T^{1/2}) + O_p(K_T^{-\bar{a}}), \quad (10)$$

with the properties that $\hat{\theta} = O_p(1)$ and $\hat{\theta}^{-1} = O_p(1)$.

We see from this theorem that under the null, the limiting distribution of the CSL test statistic is equal to that of the SL test and the difference between the CSL test statistic and the nuisance parameter free quantity is of $O_p(T^{-1/2}K_T^{1/2})$, which is also the same as that of the SL test statistic. In addition, $\hat{Q}_j = O_p(T)$ under the alternative, in other words, the CSL test is of the same probability order as the Johansen's LR, suggesting that the CSL test might be expected to be more powerful than the SL test as T becomes large. The term expressed as $O_p(K_T^{-\bar{a}})$ might be included in the term expressed as $o_p(T^{-1/2}K_T^{1/2})$ in many cases, based on the same reason as the one for the SL test stated previously. However, considering how the asymptotics established theoretically above are preserved for the tests under finite samples, there might exist DGPs that this term greatly affects the distance from the nuisance parameter free quantity even if T is not so small and H_i decay exponentially. This might be caused by some roots of det $C(z) = 0$, and some examples adopted in Section 5 might be related to such cases.

To examine the size of $\hat{\theta}$ in comparison to that of $\tilde{\theta}$ for the SL might be difficult because both are subject to complicated dependence on various parameters.

5 Additional Matters and Remarks

If stronger initial conditions such as $y_t = 0$ and $\varepsilon_{-i} = 0$ for $\forall i \geq 0$ is supposed, reconstructing $\hat{Q}_j$ based on the data series that are not demeaned is allowed for the case $\mu = 0$, just as done for $\hat{S}_j$, and to establish results as in Theorems 1 and 2 for those in this case might be rather simpler. We note that the results for the alternatives are the same as in the above theorems in other words, each of the results of the right-hand sides in (6) and (9) still holds for each of the test statistics replaced, whereas the results under the nulls require to replace $\bar{N}_{11}$, $\bar{N}_{01}$, D_T and $\bar{W}(p-r)$ in (7), (8) and (10) with suitable ones corresponding to the data series that are not demeaned (Johansen (1996, p. 94) e.g.).

It should be strongly noted that the DGP (1) does not exclude the possibility of a finite lag order VAR, i.e., it is allowable to suppose that $H_i = 0$ for any $i \geq k+1$ with a nonnegative integer k. In such a situation, (6) reveals that the SL test is disadvantageous to desirable rank detection, emphasizing that it is reasonable to adopt the Johansen's LR test if k is known. It can also be shown within the well-known asymptotic theory of the $I(0)$ and $I(1)$ series (e.g., Hamilton (1994, p. 192 and p. 548)) that if K_T is replaced with a fixed and short one, $\hat{S}_j$ is of $O_p(T^{-1})$ for $j < r$ even if $A(z)$ in (1) is not finite. The conventional divergence rate ensures desirable rejection of the nulls, though the test adopting such a value of K_T should not be considered as the "SL" test and is not suitable for the purpose of the study stated the introduction. On the other hand, the CSL test always requires a large value of K_T to preserve its asymptotics, since those strongly depend upon the consistency property of $\hat{H}_1$.

We can easily include the case for which r defined as $1 - rank\ C(1)$ is either p or 0 in our discussion above, noting that y_t is weakly stationary except y_0 or $\sum_{i=0}^{\infty} C_i^{(1)} \varepsilon_{-i}$ as $r = p$ and it is not cointegrated as $r = 0$. The formulation of the DGP under such cases is rather simpler, in the sense that (1) can be adapted by replacing α and β' forming the error correction term with a $p \times p$ matrix $\breve{A}$ having full rank and I_p respectively for the case $r = p$ and by letting the term $\alpha(\beta' y_{t-1} - E\beta' y_{t-1})$ vanish for the case $r = 0$, and then it is easy to derive such an expression as (2) or (3) by handling such alteration. However, it is meaningless for the case $r = 0$ to derive such an expression as (4) of Lemma 1 or (5) since (3) can be already formed as a pure VAR with no error correction terms, whereas it is possible for the case $r = p$ to derive such an equation as (5) from (3) and (4) altered. Also, for the case $r = p$, $C(1) = 0$ and δ, γ, $\bar{N}_{11}$ and $\bar{N}_{11}$ are not defined, and for the case $r = 0$, $C(1)$ has full rank, $\gamma = C(1)\Lambda^{1/2}$, $\delta = \Lambda^{-1/2}$, and the definitions of $\bar{N}_{11}$ and $\bar{N}_{01}$ are still valid. We shall now present the asymptotic results of the tests for these cases.

Theorem 3 (i) *Suppose that y_t is generated by (1) as $r = p$. Then, for the test statistics and quantities defined above, instead of (6) and (9) in Theorems 1 and 2, we have*

$$T^{-1} K_T \hat{S}_j = (p - j) + o_p(1), \quad for\ \forall j < p, \quad (6)'$$

noting that $\breve{A}'\Lambda^{-1}\breve{A}\breve{A}^{-1}\Lambda\breve{A}^{-1\prime} = I_p$, and

$$T^{-1}\hat{Q}_j = \sum_{h=j+1}^{p} \bar{\nu}_h + o_p(1), \qquad for\ \forall j < p, \quad (9)'$$

where $\bar{\nu}_1 \geq \ldots \geq \bar{\nu}_p$ are the ordered eigenvalues of

$$\breve{A}'\left(\Lambda + \breve{A}\Lambda\breve{A}'\right)^{-1}\breve{A}\Lambda.$$

(ii) *Suppose that y_t is generated by (1) as $r = 0$. Then, for the test statistics and quantities defined above, (8) still holds and instead of (7) and (10) in Theorems 1 and 2, we have*

$$\hat{S}_0 = tr\bar{N}_{11}^{-1}\bar{N}_{01}'\delta\delta'\bar{N}_{01} + o_p(T^{-1/2}K_T^{1/2}) + O_p(K_T^{-\bar{a}}), \quad (7)'$$

and

$$\hat{Q}_0 = tr\bar{N}_{11}^{-1}\bar{N}_{01}'\delta\delta'\bar{N}_{01} + o_p(T^{-1/2}K_T^{1/2}) + O_p(K_T^{-\bar{a}}). \quad (10)'$$

For the case $r = p$, we note that the SL test exhibits desirable performances even if K_T is replaced by a fixed and short one, since the test statistic is of $O_p(T)$ as stated above and all the nulls must be rejected in order to detect the true rank value. Also, the differences between the test statistics and the nuisance parameter free quantities become smaller under the null $r = 0$ for the case $r = 0$, as stated by (7)′ and (10)′.

The examination of the tests under DGPs with near-integrated series or for local alternatives, discussed in such literatures as Johansen (1996, pp. 201-210), Elliott and Pesavento (2009), Hjalmarsson and Österholm (2010) and Kurita (2011), might force us to adopt a simple approach on the formulation of (1), supposing that $H_i = 0$ for any $i \geq 1$. However, such simplification or an extension to a VAR whose lag order is not one but finite stated in Exercises of Johansen (1996, pp. 209-210) is not appropriate for the present paper based on an infinite order VAR approximation at all, and it will be not easy to find some suitable formulation to handle the near-integration or local alternatives since adaptation over the whole H_i might be needed, as anticipated from the Exercises, and then the asymptotic results are expected to be more complicated as well. In addition, this issue might be frequently discussed within the framework of a nested test, for which transitions between a null and alternative hypotheses are implicitly hypothesized, and some DGPs within such a framework might not conform to the requirements for the roots of $\det A(z) = 0$.

Needless to say, the theoretical results established above hold for any K_T satisfying the requirements in Section 2. However, the SL and CSL tests practically need to determine a value from the set of such values of K_T in advance, and there are no preceding articles that propose an approach to determining the optimal choice from such values of K_T and nor any concrete values were provided in Qu and Perron (2007) mentioned in Introduction.

6 Monte Carlo Experiments

In this section, we execute Monte Carlo experiments for the cointegration rank tests dealt with in this paper, trying to adopt DGP examples to cover the parameter space of (1) as well as possible. The purpose of the experiments is to observe how the SL and CSL tests behave as certain rank determination methods in finite samples. Taking this and some circumstances of our results of the empirical sizes and powers explained later account on or following Bierens (1997), we put our attention on the entire procedure to determine the rank value than the empirical size/power of the individual hypothesis test, though the latter might more directly reflect the asymptotics established theoretically in the previous section and numerous articles including Saikkonen and Luukkonen (1997) focus their attention on the latter.

The DGPs adopted are of 4-variates systems ($p = 4$), paying attention to that the system dimensions are not so small and supposing that $\mu_0 = 0$, $y_{-j} = 0$ for any $j \geq 0$ and ε_t are distributed as Gaussian with mean zero and covariance matrix I_4 ($\Lambda = I_4$), and possess VARMA(2,2) representations, emphasizing that the error structures of the DGPs that the present paper aims to adopt are a little more complicated than those adopted by the previous researches: $A(z)$ is expressed as

$$A(z) = \left(I_p + D_1 z + D_2 z^2\right)^{-1} \bullet \{-\check{\alpha}\beta' z + (1-z)\left(I_p - \check{H}_1 z\right)\},$$

where D_1, D_2 and $\check{H}_1$ are matrices of 4×4 and $\check{\alpha} = (I_p + D_1 + D_2)\alpha$ such that $\det\{-\check{\alpha}\beta' z + (1-z)(I_p - \check{H}_1 z)\} = 0$ for $\forall |z| \leq 1$ except $z = 1$, $\det(I_p + D_1 z + D_2 z^2) \neq 0$ for $\forall |z| \leq 1$, and $\det(I_p - \check{H}_1 z) = 0$ and $\det(I_p + D_1 z + D_2 z^2) = 0$ have no common roots. Now, let $\check{\mu} = (I_p - \check{H}_1)\mu$ and define $\check{C}(z)$ as the one satisfying

$$\check{C}(z)\{-\check{\alpha}\beta' z + (1-z)\left(I_p - \check{H}_1 z\right)\} = (1-z) I_p.$$

Then, multiplying both sides of (1) by $\{-\check{\alpha}\beta' L + (1-L)(I_p - \check{H}_1 L)\}$ yields

$$\Delta y_t = \check{\alpha}\beta' y_{t-1} + \bar{H}_1 \Delta y_{t-1} + \check{\mu} + \epsilon_t + D_1 \epsilon_{t-1} + D_2 \epsilon_{t-2}, \quad \forall t \geq 3, \tag{11}$$

and as done from (1) into (2), (11) can be converted into a VMA representation:

$$\Delta y_t = \check{C}(L)\left(I_p + D_1 L + D_2 L^2\right)\epsilon_t + \check{C}(1)\check{\mu}, \tag{12}$$

It is easy to see from (2) and (11) that $C(z) = \check{C}(z)(I_p + D_1 z + D_2 z^2)$ and that $\mu = \check{C}(1)\check{\mu}$, implying that $\mu \neq 0$ if and only if $\check{\mu}$ is not spanned by the columns of $\check{\alpha}$.

We classify our DGP examples in two groups, with each group containing 24 DGPs, so that the DGPs within each group cover the values of r ranging from 0 to 3, and specify $\check{H}_1$, $\check{\alpha}$, β, D_i and $\check{\mu}$, so that they are as similar to each other as possible.[2].

First, for the formulation of the VAR parts in the first group, a single $\check{H}_1$ is designed as

$$\check{H}_1 = \begin{bmatrix} 0.05 & 0.16 & 0.25 & 0.16 \\ 1.02 & -0.02 & 0.22 & -0.82 \\ 0.05 & 0.5 & 0.05 & 0.5 \\ -0.29 & 0.19 & 0.22 & 0.48 \end{bmatrix},$$

and $\check{\alpha}$ and β for the cases $r \geq 1$ are specified as

$$r = 1: \quad \check{\alpha} = [-0.2, -0.2, -0.5, -0.2]', \quad \beta = [1, 1, 1, 1]',$$

$$r=2:\quad \breve{\alpha}=-\begin{bmatrix}-0.2 & -0.2 & -0.5 & -0.2\\ 0 & -1 & 0 & 0.5\end{bmatrix}',$$

$$\beta=\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 0 & 0 & -1\end{bmatrix}',$$

$$r=3:\quad \breve{\alpha}=-\begin{bmatrix}-0.2 & -0.2 & -0.5 & -0.2\\ 0 & -1 & 0 & 0.5\\ -0.045 & -0.015 & -0.045 & -0.015\end{bmatrix}',$$

$$\beta=\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 0 & 0 & -1\\ 1 & -1 & 1 & -1\end{bmatrix}',$$

noting that these form nested models.

In the second group, the VAR formation is given as putting

$$r=0:\quad \breve{H}_1=\begin{bmatrix}-0.5 & -0.5 & 0.7 & 0\\ 0.5 & -0.5 & 0.6 & -0.2\\ 0 & 0 & -0.6 & 0.6\\ 0 & 0 & -0.2 & -0.7\end{bmatrix},$$

$$r=1:\quad \breve{H}_1=\begin{bmatrix}0 & 0 & 0.7 & 0\\ 0 & 0 & -0.6 & -0.2\\ 0 & 0 & -0.6 & 0.6\\ 0 & 0 & -0.2 & -0.3\end{bmatrix},$$

$$\breve{\alpha}=[-0.5,\ -0.5,\ 0,\ 0]',\quad \beta=[1,\ 1,\ 0,\ 0]',$$

$$r=2:\quad \breve{H}_1=\begin{bmatrix}0 & 0 & 0.7 & -0.7\\ 0 & 0 & -0.6 & -0.2\\ 0.25 & 0.25 & 0 & 0\\ 0 & -2 & -0.2 & -0.3\end{bmatrix},$$

$$\breve{\alpha}=-\begin{bmatrix}-0.5 & -0.5 & 0 & 0\\ 0 & 0 & -0.1 & 0\end{bmatrix}',$$

$$\beta=\begin{bmatrix}1 & 1 & 0 & 0\\ 2.5 & 2.5 & 6 & -6\end{bmatrix}',$$

$$r=3:\quad \breve{H}_1=\begin{bmatrix}0 & 0 & 1.2 & 1.5\\ 0 & 0 & -0.1 & -0.1\\ 0.05 & 0.25 & -0.2 & -0.6\\ 0 & 0 & 0 & -0.1\end{bmatrix},$$

$$\breve{\alpha}=-\begin{bmatrix}-0.5 & -0.5 & 0 & 0\\ 0 & 0 & -0.1 & 0\\ -0.5 & 0 & 0.2 & -0.2\end{bmatrix}',$$

$$\beta=\begin{bmatrix}1 & 1 & 0 & 0\\ 2.5 & 2.5 & 6 & -6\\ 0 & 0 & 1 & 3\end{bmatrix}',$$

noting that the four marices of $\breve{H}_1+I_p+\breve{\alpha}\beta'$ within the group are similar.

Next, D_i are designed as

$$D_1=\begin{bmatrix}d_1 & 0 & 0 & 0\\ 0 & d_2 & 0 & d_4\\ 0 & 0 & d_3 & 0\\ 0 & 0 & 0 & d_2\end{bmatrix},$$

$$D_2=\begin{bmatrix}d_5 & 0 & 0 & 0\\ 0 & d_6 & 0 & 0\\ 0 & 0 & d_7 & 0\\ 0 & 0 & 0 & d_6\end{bmatrix}.$$

with d_i taking the below values of (i), (ii) or (iii)

(*i*): $d_1=d_2=d_3=0.4$, $d_4=0$ *and* $d_5=d_6=d_7=-0.32$,
(*ii*): $d_1=1.5$, $d_2=0.9$, $d_3=0.4$, $d_4=0.5$, $d_5=0.56$, $d_6=0.2$ *and* $d_7=0.04$,
(*iii*): $d_1=d_2=d_3=-0.1$, $d_4=0.5$ *and* $d_5=d_6=d_7=-0.2$.

Finally, $\breve{\mu}$ is set to be either zero vector or (1, 0, 0.5, 1)′.

We ran 10,000 simulations using pseudo normal random variables for ε_t as $T=150$, 250, in order to obtain the relative frequencies that the rank determination procedure based on the SL/CSL test for one value of K_T ranging from 3 to 6 correctly detects r over those 10,000 trials. The tests are executed consecutively for several nulls at the 5 % level as long as those are rejected. The critical values are taken from MacKinnon *et al.* (1999), noting that those are not of the empirical distributions of the nuisance parameter free quantities at each finite sample but of their limiting distributions, and all calculations were made in Gauss. In the tables below, those relative frequency results are presented.

Now, let us survey finite sample performances of the rank determination procedures based on the SL and CSL tests through tabular comparison. As observed, there exist more DGP examples for which the CSL test seems to be superior to the SL test in terms of correct rank detection, though the superiority of the CSL test might be not observed for all the values of K_T, the difference between both tests is not so great in many cases and the CSL test becomes comparatively worse in its performances in several examples. Those might naturally reflect results of the empirical sizes and powers of the tests that are not reported in the present paper: in general, the CSL test tends to be more powerful than its worse upward size distortions can be covered, resulting in more correct rank detection by the CSL, though there exist cases such that the stances of the CSL test and the SL test stated above are reversed. It is also observed that for the SL test, the power tends to be greater as K_T decreases, emphasizing that it might be supported by the theoretical results stated in Theorem 1 or its related statement. However, in many cases, those are accompanied by serious upward size distortion, and therefore, it is difficult to find certain clues to decide whether each test is desirable or not from only the results of the sizes and powers.[3]

The cases for which d_i take the values of (iii) in the first group, particularly as $r=0$ or $r=1$, reveal conspicuous inferiority of the CSL test, suggesting that there might exist some problems in the convergence property of $\hat{H}_1$ under such a case. In really, experimental results for the CSL statistics constructed by substituting H_1 itself for $\hat{H}_1$ certainly support this suggestion, though those are not provided in the present paper as well, and from the viewpoint of the compatibility of these experimental results with (A.51) in the Appendix, which provides an elaborate formulation of the convergence of $\hat{H}_1$, we might consider the term as $o_p(T^{-1/2}K_T^{1/2})$ or $O_p(K_T^{-\bar{a}})$ in this equation not to be insignificant in finite samples, particularly in the cases stated above. The negative values of d_i in (iii) might form situations in which the stochastic trend $\check{C}(1)(I_p+D_1+D_2)(\sum_{h=1}^{t}\varepsilon_h)$ is not easily distinguished from the weakly stationary part of y_t and consequently correct rank detection is not smoothly done, though the results for the SL test are desirable under those cases and the performances of the CSL test in the second group are rather considerably better than those of the SL test as an appropriate K_T is chosen. We note that the negative values of d_i does not necessarily imply that det $(I_p+D_1z+D_2z^2)=0$ has roots close to the unit root, unlike VARMA(1,1) or VARMA(2,1), emphasizing that observing the performances of the tests in such cases is one of the aims of the experiments. There also exist several cases in which almost perfect rank detection is led to by the CSL test: particularly in the cases of $r=1$ in the second group, extremely small sizes associated with adequate powers result in such rank detection, though direct results of the sizes and powers are not provided in the present paper. In these cases, there exist relatively numerous roots of det $\{-\check{\alpha}\beta'z+(1-z)(I_4-\check{H}_1z)\}=0$ that are not 1 but relatively close to 1 in absolute value, which might cause "favorable" distortions of the sizes and powers for the CSL statistics, resulting in the desirable rank detection discussed above. We next point out that the performances of the CSL test tend to vary according to the value of K_T chosen: they are noticeably desirable at several values of K_T, whereas the remaining values of K_T lead to severe results, pointing out that this appears to follow from the convergence property of $\hat{H}_1$ associated with the value of K_T adopted. On the other hand, the SL test exhibits relatively stable performances regardless of the value of K_T or the DGPs.

The results for $T=250$ are more desirable than those for $T=150$ in both tests, reflecting certain dependence of the asymptotics stated in the previous section upon the sample size.

In general, the finite sample performances of the two methods are not so stable in the sense that those are desirable in several DGP examples but not so in others, though the SL test seems to be more robust. We also note that values of K_T other than those adopted above might not satisfy the requirements in Section 2 and tend to fail in successfully detecting the rank value.

7 Concluding remarks

We have been dealt with two cointegration rank tests, the one discussed by Saikkonen and Luukkonen (1997), abbreviated as the SL test, and the version corrected by the present paper, named the CSL, via the finite lag order VAR approximation and discussed their performances through the asymptotic results established theoretically and finite sample performances by some Monte Carlo experiments. As a result, it is pointed out that with the exception of the

CSL under the alternatives, the tests discussed here are not free of some slower convergence/divergence properties, as seen in other cointegration rank tests for the DGPs with serially correlated errors, such as the ones in Bierens (1997) or Shintani (2001). Under the alternatives, the divergence of the SL test statistics is slower than the conventional one, whereas the CSL is considered to overcome such weakness in the sense that its divergence rate is the same as that for the conventional Johansen's LR test, and similarly under the nulls, the speed of the convergence of the test statistics to the nuisance parameter free quantities leading to the well-known limiting distributions is certainly reduced, resulting in deterioration of the asymptotic sizes and powers of the tests. We also emphasize that the conventional divergence of the CSL test under the alternative was led to from the matter that the CSL test successfully manage the first step regression used in the SL test by altering one vector series included in the list of regressors.

In addition to the theoretical outcomes on the asymptotic properties of the SL and CSL tests, the experiments in the previous section reveal finite sample performances of the tests in several specific examples. We might find several DGP examples that the asymptotic properties do not work sufficiently, particularly for the CSL test, though the number of such cases does not outnumber that of the cases which led to satisfactory results, and it might be stated that the CSL test is generally superior to the SL test. On the other hand, the SL test generally exhibits satisfactory performances in the examples for which the CSL performances were not so. This seems to reflect the matter that the consistency property of the estimator of H_1 is severely deteriorated in some DGPs and for some values of K_T, implying that the CSL test is not robust at all. For the cases of $r=3$, the tests must be rejected three times to attain to the true rank value, and if K_T adopted is a comparatively small value, the tests, particularly the SL test, might tend to lead to desirable rank detection probably because of the divergence properties of the test statistics under the alternatives stated in Sections 4 and 5. We might recognize from the experimental results that an appropriate value of K_T is needed to achieve satisfactory rank detection, particularly for the CSL test.

Discussion on the asymptotics/performances of the tests when the DGP considered corresponds to a local alternative has not been dealt with in the present paper, emphasizing that the simple DGP formulations adopted by the previous researches are not appropriate in our situation and that there exist some DGPs that are not suitable for local alternatives. Also, the case in which one root in det $C(z)=0$ is close to -1, mentioned as the "near" non-invertible case, has not been paid attention to, noting that this issue has been discussed similarly in unit root testing for time series with serially correlated errors (e.g., Perron and Ng (1996)) Furthermore, the issue of how one value of K_T as an optimal lag order should be determined has not been discussed as well. It might be not easy to establish a method giving one concrete value with some validity beyond the requirements given above, as predicted from articles such as Qu and Perron (2007). On the other hand, the experiments in this paper might be worthier if the critical value of the empirical distribution of the nuisance parameter free quantity was adopted instead of the counterpart of the limiting one for each test, though the present paper's work was done under the supposition that the differences are inconsiderable. In addition, in order to derive firmer views on the finite sample performances of the tests, we might need to compare the tests discussed in this paper with other cointegration rank tests such as those mentioned above. We leave these issues to future research.

FOOTNOTES

1 Note that (1) is essentially the same as the DGP formulation by Saikkonen (1992) or Saikkonen and Luukkonen (1997).

2 For the change of the values, it is not easy for the roots of the equation det$\{-\check{\alpha}\beta' z+(1-z)(I_p-\check{H}_1 z)\}=0$ to keep satisfying the requirements, and thus we need to chose suitable values with great care.

[3] The report on the empirical sizes and powers and the technical report for detailed proofs of the theorems and lemma are available from the author upon request.

Appendix

Proof of Lemma 1 Noting that (3) for which K_T is replaced by $K_T - 1$ or $K_T - L_T$ is still valid, we replace K_T and t in this equation with $K_T - L_T$ and $t - s$, respectively, for $s = 0, 1, \ldots, L_T$, which yields

$$\begin{aligned} &\Delta y_{t-s} - \mu - \alpha(\beta' y_{t-s-1} - E\beta' y_{t-s-1}) \\ &- \sum_{i=1}^{K_T - L_T} H_i(\Delta y_{t-s-i} - \mu) \\ = \; &\epsilon_{t-s} + \zeta_{t-s-K_T+L_T-1;K_T-L_T}. \\ &\qquad \forall\, t \geq K_T - L_T + 2 + s. \quad (A.1) \end{aligned}$$

The lemma can be proved essentially by multiplying both sides of (A.1) as $s = 0$ by the polynomial

$$\left(\sum_{s=0}^{L_T-1} \frac{L_T - s}{L_T} L^s\right)(1 - L) = 1 - L_T^{-1}\left(\sum_{s=0}^{L_T-1} L^s\right)L,$$

taking into account that

$$\begin{aligned} &L^s \zeta_{t-K_T+L_T-1;K_T-L_T} \\ = \; &L^s\{\Delta y_t - \mu - \alpha(\beta' y_{t-1} - E\beta' y_{t-1}) \\ &- \sum_{i=1}^{K_T-L_T} H_i(\Delta y_{t-i} - \mu) - \epsilon_t\} \\ = \; &\zeta_{t-s-K_T+L_T-1;K_T-L_T}, \\ &\qquad if\ t \geq K_T - L_T + 2 + s, \qquad s = 1, \cdots, L_T. \end{aligned}$$

Proof of Theorem 1 We first note that the quantities for which $\zeta_{t-K_T-1;K_T}$ or $\bar{\zeta}_{t-K_T+L_T-1;K_T-L_T}$ are used in construction are at most of $O_p(K_T^{-\bar{a}})$ in the asymptotics desired for the theorem, though to show it is omitted in this paper. The well-known asymptotics for weakly stationary and ergodic, $I(0)$ and $I(1)$ series (e.g., Hannan (1970, pp. 209-212), Hamilton (1994, p. 192 and p. 548), Park and Phillips (1888, 1989) or Phillips (1995)) play an active part for the proof, and it should also be noted that the proof needs some arguments on sums of K_T-pieces of quantities being of various orders of probability, such as $O_p(1)$ or $O_p((K_T - s + 1))$ for $s = 1, \ldots, K_T$. The arguments/results in the foregoing articles such as Saikkonen (1992) or Saikkonen and Luukkonen (1997) are not applied to the present proof, since those are not provided for the convenience of the present paper and in addition, quantities constructed based on $\sum_{s=1}^{L_T} \varepsilon_{t-s}$ have been not dealt with in these studies. The evaluation of these quantities is more specific and elaborate in the present paper, though doing so makes the proof tedious and long. The role of Equation (4) or (5) and asymptotics on $\sum_{s=1}^{L_T} \varepsilon_{t-s}$ should be particularly emphasized. Taking these points into account, we proceed with tasks needed to prove the theorem below.

First, define a series of pK_T-dimensional vectors and $pK_T \times pK_T$ matrices as follows:

$$\eta_{t-iK_T-1} = \left(\epsilon'_{t-iK_T-1}, \cdots, \epsilon'_{t-iK_T-K_T}\right)', \quad i = 0, 1, \quad \mu_{K_T} = (\mu', \cdots, \mu)',$$

$$\check{C}_{K_T} = \begin{bmatrix} I_p & C_1 & C_2 & \cdots & C_{K_T-2} & C_{K_T-1} \\ 0 & I_p & C_1 & \cdots & C_{K_T-3} & C_{K_T-2} \\ & & I_p & \ddots & \vdots & \vdots \\ & & & \ddots & C_1 & \vdots \\ \mathbf{0} & & & & I_p & C_1 \\ & & & & & I_p \end{bmatrix},$$

$$\check{F}_{K_T} = \begin{bmatrix} C_{K_T} & C_{K_T+1} & \cdots & C_{2K_T-1} \\ C_{K_T-1} & C_{K_T} & \cdots & C_{2K_T-2} \\ \vdots & \vdots & \ddots & \vdots \\ C_2 & C_3 & \cdots & C_{K_T+1} \\ C_1 & C_2 & \cdots & C_{K_T} \end{bmatrix},$$

$$\bar{F}_{k_T+1+m} = \begin{bmatrix} C_{2K_T+m} \\ C_{2K_T-1+m} \\ \vdots \\ C_{K_T+2+m} \\ C_{K_T+1+m} \end{bmatrix}, \quad \forall m \geq 0,$$

$$\overline{\eta} = T^{-1} \sum_{t=K_T+2}^{T} \eta_{t-1},$$

$$\overline{\phi} = T^{-1} \sum_{t=K_T+2}^{T} (\Delta z_{t-1} - \mu_{K_T}),$$

$$\begin{aligned} \tilde{N}_{02} = T^{-1} &\sum_{t=K_T+2}^{T} \eta_{t-1}(\Delta z_{t-1} - \mu_{K_T})' \\ &- (T - K_T - 1)^{-1} T\, \overline{\eta}\, \overline{\phi}', \end{aligned}$$

with $\tilde{N}_{20} = \tilde{N}'_{02}$. Second, for any sequence of positive integers $\{N_T\}$ such that $\lim_{T\to\infty} N_T = \infty$ and $\lim_{T\to\infty} N_T/K_T \leq 1$, let $\bar{I}'_{N_T}$ and $\check{I}'_{N_T}$ represent the $p \times pN_T$ matrices, gievn as

$$\bar{I}'_{N_T} = [I_p, \cdots, I_p,\ I_p], \qquad \breve{I}'_{N_T} = [I_p, \cdots, I_p,\ 0],$$

noting that the $p \times p$ submatrix comprising the last p columns of $\breve{I}'_{N_T}$ is set to zero matrix. Third, let $\bar{I}_{\bar{N}_T;N_T}$ represent the $p \times pN_T$ matrix such that

$$\bar{I}'_{\bar{N}_T;N_T} = \left[\bar{I}'_{\bar{N}_T},\ 0\right] = \left[\breve{I}'_{\bar{N}_T+1},\ 0\right],$$

where $\{\bar{N}_T\}$ is a sequence of positive integers such that $\lim_{T\to\infty} N_T = \infty$ and $\bar{N}_T < N_T$.

Using the above notations, we obtain an expression for Δz_{t-1}:

$$\breve{C}_{K_T}^{-1}(\Delta z_{t-1} - \mu_{K_T}) = \eta_{t-1} + \breve{C}_{K_T}^{-1}\breve{F}_{K_T}\eta_{t-K_T-1} + \nu_{t-2K_T-1}, \qquad (A.2)$$

where $\nu_{t-2K_T-1} = \sum_{m=0}^{\infty}\breve{C}_{K_T}^{-1}\bar{F}_{K_T+1+m}\,\varepsilon_{t-2K_T-1-m}$ Next, let $\{K_{2;T}\}$ be a sequence of positive integers such that $\lim_{T\to\infty} K_{2;T}/K_T^{1/2} = 0$ and $\lim_{T\to\infty} K_T^{1/(1+\bar{a})}/K_{2;T} = 0$, let $K_{1;T} = K - K_{2;T}$, and define

$$\psi_{t-K_T-1} = \alpha\beta'\{\epsilon_{t-K_T-1} + \sum_{j=1}^{\infty}(I_p + \sum_{h=1}^{j} C_h)\epsilon_{t-K_T-1-j}\},$$

$$\omega_{t-K_T-1;1} = \left(0, \cdots, 0,\ \omega'_{t-K_T-1;1;K_{2;T}}, \cdots, \omega'_{t-K_T-1;1;1}\right)',$$

$$\omega_{t-K_T-1;2} = \left(\omega'_{t-K_T-1;2;K_T+1}, \cdots, \omega'_{t-K_T-1;2;K_{2;T}+1},\ 0, \cdots, 0\right)',$$

$$\omega_{t-2K_T-1;*} = \sum_{m=0}^{\infty}\begin{bmatrix} H_{2K_T+m} + \sum_{h=1}^{K_T+m} H_{2K_T+m-h}C_h \\ H_{2K_T-1+m} + \sum_{h=1}^{K_T+m} H_{2K_T+m-1-h}C_h \\ \vdots \\ H_{K_T+2+m} + \sum_{h=1}^{K_T+m} H_{K_T+2+m-h}C_h \\ C_{K_T+1+m} \end{bmatrix} \bullet \epsilon_{t-2K_T-1-m},$$

where

$$\omega_{t-K_T-1;1;K_{2;T}-j+1} = H_{K_{2;T}-j+1}\epsilon_{t-K_T-1} + \sum_{i=1}^{K_T-1}\left(H_{K_{2;T}-j+1+i} + \sum_{h=1}^{i} H_{K_{2;T}-j+1+i-h}C_h\right)\epsilon_{t-K_T-1-i}, \quad j = 1, \cdots, K_{2;T}-1,$$

$$\omega_{t-K_T-1;1;1} = \sum_{i=1}^{K_T} C_i\epsilon_{t-K_T-i},$$

$$\omega_{t-K_T-1;2;K_T-j'+1} = H_{K_T-j'+1}\epsilon_{t-K_T-1} + \sum_{i=1}^{K_T-1}\left(H_{K_T-j'+1+i} + \sum_{h=1}^{i} H_{K_T-j'+1+i-h}C_h\right)\epsilon_{t-K_T-1-i}, \quad j' = 1, \cdots, K_{1;T}.$$

Then, based on these notations and several relationships such as

$$C(z)\{-\alpha\beta' z + (1-z)(I_p - \sum_{i=1}^{\infty} H_i z^i)\} = (1-z)I_p,$$

(A.2) can be converted into

$$\breve{C}_{K_T}^{-1}(\Delta z_{t-1} - \mu_{K_T}) = \eta_{t-1} + \breve{I}_{K_T}\psi_{t-K_T-1} + \sum_{i=1}^{2}\omega_{t-K_T-1;i} + \omega_{t-2K_T-1;*}. \qquad (A.3)$$

We shall now turn to the derivation of an asymptotic result for

$$L_T^{-1}\bar{I}'_{L_T;K_T}\tilde{N}_{02}\hat{M}_{22}^{-1}\tilde{N}_{20}\bar{I}_{L_T;K_T},$$

which will be completed in (A.20) below through long and tedious arguments. First, define

$$\hat{N}_{10;h;t} = \sum_{s=1}^{L_T}\epsilon_{t-s})\epsilon'_{t-h}, \qquad h = 0, 1, \cdots, K_T,$$

with $\hat{N}_{01;h;t} = \hat{N}'_{10;h;t}$. Then it is not difficult to see that

$$T^{-1}\sum_{t=K_T+2}^{T}\hat{N}_{10;h;t} = \Lambda + O_p(T^{-1/2}K_T^{1/2}), \quad h = 1, \cdots, L_T, \qquad (A.4)$$

and that

$$T^{-1}\sum_{t=K_T+2}^{T}\hat{N}_{10;h;t} = O_p(T^{-1/2}K_T^{1/2}), \quad h = 0 \ or\ h = L_T+1, \cdots, t-1, \qquad (A.5)$$

which is in turn followed by

$$\hat{N}_{11;+} = \Lambda + O_p(T^{-1/2}K_T^{1/2}), \qquad (A.6)$$

where

$$\hat{N}_{11;+} = L_T^{-1}T^{-1}\sum_{t=K_T+2}^{T}(\sum_{s=1}^{L_T}\epsilon_{t-s})(\sum_{s=1}^{L_T}\epsilon_{t-s})'.$$

We next state some elementary results for the sums of $\bar{M}_T$-pieces of statistics. Let $\{\bar{M}_T\}$ and $\{\bar{N}_T\}$ be sequences of positive integers such that $\lim_{T\to\infty}\bar{M}_T = \infty$ and $\lim_{T\to\infty}\bar{N}_T = \infty$. Then, for any $\bar{M}_Tp$-dimensional vector $\bar{b}$, any $p\bar{N}_T$-dimensional vector $\bar{f}$

such that $\bar{b}'\bar{b} = O_p(1)$ and $\bar{f}'\bar{f} = O_p(1)$, and any $p\bar{M}_T \times p\bar{N}_T$ matrix $\check{L}$ whose elements are of $O_p(1)$, it is easy to see that

$$\bar{M}_T^{-1/2}\bar{b}'\check{L} = O_p(1), \quad \bar{N}_T^{-1/2}\check{L}\bar{f} = O_p(1),$$
$$\bar{M}_T^{-1/2}\bar{N}_T^{-1/2}\bar{b}'\check{L}\bar{f} = O_p(1). \qquad (A.7)$$

In the discussion on the sums of $\bar{M}_T$-pieces of statistics below, (A.7) is used without particularly mentioning it.

Now, return to the derivation of the asymptotic targeted and for this purpose, define

$$\hat{M}_\eta = T^{-1} \sum_{t=K_T+2}^{T} \eta_{t-1}\eta'_{t-1},$$
$$\hat{M}_{\omega 1} = T^{-1} \sum_{t=K_T+2}^{T} \omega_{t-K_T-1;1}\omega'_{t-K_T-1;1}, \quad i = 1, 2,$$
$$\hat{M}_\psi = T^{-1} \sum_{t=K_T+2}^{T} \psi_{t-K_T-1}\psi'_{t-K_T-1},$$
$$\hat{M}_{\omega 1:\psi} = T^{-1} \sum_{t=K_T+2}^{T} \omega_{t-K_T-1;i}\psi'_{t-K_T-1}.$$

In view of (A.2) and (A.3), it can be seen that $\tilde{N}_{02}\check{C}'^{-1}_{K_T}$ and $\check{C}^{-1}_{K_T}\hat{M}_{22}\check{C}'^{-1}_{K_T}$ are described by the quantities defined above and undefined ones that are at most of $O_p(T^{-1/2})$, which will be shown through the standard asymptotic theory for weakly stationary and ergodic series, and from this, it is obvious that

$$\hat{M}_\eta = I_{K_T} \otimes \Lambda + O_p(T^{-1/2}). \qquad (A.8)$$

Next, partition $\hat{M}_\eta$ into four block matrices as

$$\hat{M}_\eta = \begin{bmatrix} \hat{M}_{\eta;11} & \hat{M}_{\eta;12} \\ \hat{M}_{\eta;21} & \hat{M}_{\eta;22} \end{bmatrix},$$

where $\hat{M}_{\eta;11}$, $\hat{M}_{\eta;12}$, $\hat{M}_{\eta;21}$ and $\hat{M}_{\eta;22}$ are matrices of $pK_{1;T} \times pK_{1;T}$, $pK_{1;T} \times pK_{2;T}$, $pK_{2;T} \times pK_{1;T}$ and $pK_{2;T} \times pK_{2;T}$ and define

$$\bar{M}_{\Lambda;ii} = I_{K_{i;T}} \otimes \Lambda, \quad \hat{P}_{ii} = \hat{M}_{\eta;ii} - \bar{M}_{\Lambda;ii}, \quad i = 1, 2.$$

Then, it is trivial again by the standard asymptotic theory for weakly stationary and ergodic series to show that

$$\hat{M}^{-1}_{\eta;ii} = \bar{M}^{-1}_{\Lambda;ii} + \hat{P}_{ii;*}, \quad \hat{P}_{ii;*} = O_p(T^{-1/2}),$$
$$i = 1, 2. \quad (A.9)$$

Also, define

$$\hat{N}_{11;*} = L_T^{-1}T^{-1} \sum_{t=K_T+2}^{T} y_{t-1}(\sum_{s=1}^{L_T}\epsilon_{t-s})'$$
$$- L_T^{-1}(T - K_T - 1)^{-1}T\,\bar{y}_{-1}\,\bar{\eta}'\bar{I}_{L_T;K_T}.$$

Then, by (A.4) to (A.6) and by showing

$$T^{-1} \sum_{t=K_T+2}^{T} \left(T^{1/2},\ (t-1)/T^{1/2},\ (\sum_{h=1}^{t-L_T-1}\epsilon_h)'\right)'$$
$$\bullet\, (\epsilon_t - \epsilon_{t-s})'$$
$$= O_p(T^{-1/2}s^{1/2}), \qquad s = 1, \cdots, L_T,$$

we can obtain

$$\bar{D}_T^{-1\prime}\gamma'\hat{N}_{11;*} = \bar{S}_{11;*} + O_p(T^{-1/2}K_T^{1/2}). \qquad (A.10)$$

Next, let $\hat{M}_{\omega_1;-}$ and $\hat{M}_{\omega_1:\psi;-}$ represent the matrices of $pK_{2;T} \times pK_{2;T}$ and $pK_{2;T} \times p$ respectively such that

$$\hat{M}_{\omega 1} = \begin{bmatrix} 0 & 0 \\ 0 & \hat{M}_{\omega 1;-} \end{bmatrix}, \quad \hat{M}_{\omega 1:\psi} = \begin{bmatrix} 0 \\ \hat{M}_{\omega 1:\psi;-} \end{bmatrix}'.$$

Then, it can be shown based on the properties of ψ_{t-K_T-1} or $\omega_{t-K_T-1;1}$ that the submatrix comprising columns from the $(j-1)p+1$-th to the jp-th of $\bar{M}_{\omega_1:\psi;-}$ is of $O((K_{2;T}-j+1)^{-1-\bar{a}})$ for $j = 1, \ldots, K_{2;T}$. It is also led to by the standard asymptotic theory that

$$\hat{M}_{\omega 1;-} = \bar{M}_{\omega 1;-} + O_p(T^{-1/2}),$$
$$\hat{M}_{\omega 1:\psi:-} = \bar{M}_{\omega 1:\psi;-} + O_p(T^{-1/2}), \qquad (A.11)$$

where $\bar{M}_{\omega_1;-}$ and $\bar{M}_{\omega_1:\psi;-}$ are matrices of $pK_{2;T} \times pK_{2;T}$ and $pK_{2;T} \times p$ respectively, such that

$$\bar{M}_{\omega 1} = \begin{bmatrix} 0 & 0 \\ 0 & \bar{M}_{\omega 1;-} \end{bmatrix}, \quad \bar{M}_{\omega 1:\psi} = \begin{bmatrix} 0 \\ \bar{M}_{\omega 1:\psi;-} \end{bmatrix}'.$$

Similarly, defining $\bar{M}_\psi = E\,\psi_{t-K_T-1}\,\psi'_{t-K_T-1}$, it follows from the definition of ψ_{t-K_T-1} and the standard asymptotic theory that $\bar{M}_\psi = O(1)$ and $\hat{M}_\psi = \bar{M}_\psi + O_p(T^{-1/2})$. Also, in view of the form of $\bar{M}_{\omega_1:\psi;-}$, we can find a $pK_{2;T} \times r$ column full rank matrix $\bar{M}_{\omega_1:\psi;+}$ such that its columns correspond to some of those of $\bar{M}_{\omega_1:\psi;-}$,

$$\bar{M}'_{\omega 1:\psi;+}\bar{M}_{\omega 1:\psi;+} = O(1) \ \mathit{and} \ \left(\bar{M}'_{\omega 1:\psi;+}\bar{M}_{\omega 1:\psi;+}\right)^{-1}$$
$$= O(1).$$

Also, defining

$$\tilde{M}_{\omega 1:\psi} = \bar{M}_{\omega 1:\psi:+}\left(\bar{M}'_{\omega 1:\psi:+}\bar{M}_{\omega 1:\psi:+}\right)^{-1/2},$$

$$\bar{Q}_{\omega 1:\psi} = \bar{M}'_{\omega 1:\psi;-}\tilde{M}_{\omega 1:\psi},$$

$$\bar{M}_{\eta} = I_{K_T}\otimes\Lambda + \bar{M}_{\omega 1} + \bar{M}_{\omega 1:\psi}\breve{I}'_{K_T} + \breve{I}_{K_T}\bar{M}'_{\omega 1:\psi},$$

$$\bar{M}_{\eta;*} = \bar{M}_{\eta} + \breve{I}_{K_T}\bar{M}_{\psi}\breve{I}'_{K_T},$$

$$\hat{P}_1 = L_T^{-1/2}\bar{I}'_{L_T;K_T}\tilde{N}_{02}\breve{C}'^{-1}_{K_T} - L_T^{-1/2}\Lambda\bar{I}'_{L_T;K_T},$$

$$\tilde{P}_2 = \breve{C}^{-1}_{K_T}\hat{M}_{22}\breve{C}'^{-1}_{K_T} - \bar{M}_{\eta;*},$$

$$\hat{P}_2 = \left(\bar{M}_{\eta;*} + \tilde{P}_2\right)^{-1} - \bar{M}^{-1}_{\eta;*}.$$

it is trivial from the form of $\bar{M}_{\omega_1:\psi}$ and the definitions of $\hat{P}_1$, $\tilde{P}_2$ and $\hat{P}_2$ to see that

$$rank\,\bar{Q}_{\omega 1:\psi} = rank\,\bar{M}_{\omega 1:\psi;+} = rank\,\bar{M}_{\omega 1:\psi;-} = rank\,\bar{M}_{\omega 1:\psi} = r,$$

and that

$$L_T^{-1/2}\bar{I}'_{L_T;K_T}\tilde{N}_{02}\breve{C}'^{-1}_{K_T} = L_T^{-1/2}\Lambda\bar{I}'_{L_T;K_T}+\hat{P}_1,$$

$$\left(\breve{C}^{-1}_{K_T}\hat{M}_{22}\breve{C}'^{-1}_{K_T}\right)^{-1} = \bar{M}^{-1}_{\eta;*} + \hat{P}_2. \qquad (A.12)$$

On the other hand, in view of the forms of $\tilde{M}_{\omega_1:\psi}$ and $\bar{Q}_{\omega_1:\psi}$ and based on that the submatrix comprising columns from the $(j-1)p+1$-th to the jp-th of $\bar{M}_{\omega_1:\psi;-}$ is of $O((K_{2;T}-j+1)^{-1-\bar{\alpha}})$, $j=1, \ldots, K_{2;T}$, and (A.8), (A.9) and (A.11) with some manipulation and via tedious arguments, it can be shown that for any pK_T-dimensional vectors b and f such that $b'b=O_p(1)$ and $f'f=O_p(1)$,

$$b'\bar{M}^{-1}_{\eta}f = O_p(1). \qquad (A.13)$$

Also, defining

$$\breve{\Lambda}_{\omega 1}=\left(\bar{Q}'_{\omega 1:\psi}\Lambda^{-1}\bar{Q}_{\omega 1:\psi}\right)^{-1},\quad \bar{\Lambda}_{\omega 1:Q} = \bar{Q}_{\omega 1:\psi}\breve{\Lambda}_{\omega 1}\bar{Q}'_{\omega 1:\psi}$$

and by using arguments similar to those used to derive (A.13), we obtain

$$L_T^{-1}\Lambda\bar{I}'_{L_T;K_T}\bar{M}^{-1}_{\eta}\bar{I}_{L_T;K_T}\Lambda = \Lambda - (1/m)\bar{\Lambda}_{\omega 1:Q} + O(K_T^{-1}K_{2;T}), \qquad (A.14)$$

and

$$L_T^{-1}\Lambda\bar{I}'_{L_T;K_T}\bar{M}^{-1}_{\eta}\breve{I}_{K_T} = I_p - \bar{\Lambda}_{\omega 1:Q}\Lambda^{-1} + O(K_T^{-1}K_{2;T}). \qquad (A.15)$$

Furthermore, taking the form of $\bar{\Lambda}_{\omega_1:Q}$ into account, it can be shown that

$$rank\left(\Lambda - \bar{\Lambda}_{\omega 1:Q}\right) = p-r,$$

and this in turn ensures the presence of a $p\times p$ orthogonal matrix $\tilde{L}$ such that $\tilde{L}=[\tilde{L}_1, \tilde{L}_2]$ with $\tilde{L}_1$ of $p\times(p-r)$ and $\tilde{L}_2$ of $p\times r$,

$$rank\,\tilde{L}'_1\left(\Lambda - \bar{\Lambda}_{\omega 1:Q}\right)\tilde{L}_1 = p-r,\ and\ \tilde{L}'_2\left(\Lambda - \bar{\Lambda}_{\omega 1:Q}\right) = 0.$$

Consequently,

$$\left(L_T^{-1}\breve{I}'_{K_T}\bar{M}^{-1}_{\eta}\breve{I}_{K_T} + L_T^{-1}\bar{M}_{\psi}\right)^{-1} = (1/m)\Lambda\tilde{L}_T \bullet\{\begin{bmatrix}\{\tilde{L}'_1\left(\Lambda-\bar{\Lambda}_{\omega 1:Q}\right)\tilde{L}_1\}^{-1} & 0\\ 0 & \left(\tilde{L}'_2\breve{R}\tilde{L}_2\right)^{-1}\end{bmatrix} + O(K_T^{-1/2}K_{2;T}^{1/2})\}\tilde{L}'_T\Lambda, \qquad (A.16)$$

where $\breve{R}$ is a $p\times p$ constant matrix such that $\breve{R}=O(1)$ and $(\tilde{L}'_2\breve{R}\tilde{L}_2)^{-1}=O(1)$ and $\tilde{L}_T=[\tilde{L}_1, K_{2;T}^{-1/2}K_T^{1/2}\tilde{L}_2]$.

Now, putting (A.14) to (A.16) together yields

$$L_T^{-1}\Lambda\bar{I}'_{L_T;K_T}\bar{M}^{-1}_{\eta;*}\bar{I}_{L_T;K_T}\Lambda = \{(m-1)/m\}\Lambda + o(1), \qquad (A.17)$$

It is not so difficult to see from (A.13) and (A.16) that for any pK_T-dimensional vectors b and f such that $b'b=O_p(1)$ and $f'f=O_p(1)$,

$$b'\bar{M}^{-1}_{\eta;*}f = O_p(1). \qquad (A.18)$$

It can also be shown by (A.3) and some results including (A.8) or (A.11) that $\hat{P}_1=O_p(T^{-1/2}K_T^{1/2})$ and $\tilde{P}_2=O_p(K_{T;*}^{-1})$, where $K_{T;*}=min\{T^{1/2}, K_{2;T}^{\bar{\alpha}+1}\}$. Furthermore, defining

$$\hat{P}_{2;*} = \sum_{i=1}^{\infty}(-1)^i(\bar{M}^{-1}_{\eta;*}\tilde{P}_2)^i,$$

it can be shown based on the above result of $\tilde{P}_2$ and via some arguments that for any pK_T-dimensional vector b such that $b'b=O_p(1)$,

$$\hat{P}_2 = -\bar{M}^{-1}_{\eta;*}\left(\tilde{P}_2 + \tilde{P}_2\hat{P}_{2;*}\right)\bar{M}^{-1}_{\eta;*},$$

$$b'\hat{P}_{2;*} = O_p(K^{-1}_{T;*}K_T^{1/2}),$$

$$\tilde{P}_2\hat{P}_{2;*}b = O_p(K^{-2}_{T;*}K_T^{3/2}),$$

$$b'\tilde{P}_2\hat{P}_{2;*} = O_p(K^{-2}_{T;*}K_T^{3/2}),$$

$$b'\bar{M}^{-1}_{\eta;*}\tilde{P}_2 = O_p(K^{-1}_{T;*}K_T^{1/2}),$$

$$b'\bar{M}^{-1}_{\eta;*}\tilde{P}_2\hat{P}_{2;*} = O_p(K^{-2}_{T;*}K_T^{3/2}), \qquad (A.19)$$

Hence, by applying (A.17) and (A.19) as well as the above result of $\hat{P}_1$ to (A.12), we obtain

$$L_T^{-1}\bar{I}'_{L_T;K_T}\tilde{N}_{02}\hat{M}_{22}^{-1}\tilde{N}_{20}\bar{I}_{L_T;K_T}$$
$$= \{(m-1)/m\}\Lambda + o_p(1). \qquad (A.20)$$

Next, define

$$\tilde{N}_{12} = T^{-1}\sum_{t=K_T+2}^{T}\xi_{t-1}(\Delta z_{t-1} - \mu_{K_T})' - (T-K_T-1)^{-1}T\,\bar{\xi}_{-1}\,\bar{\phi}',$$
$$\tilde{N}_* = T^{-1}D_T^{-1}\sum_{t=K_T+2}^{T}\xi_{t-1}(\epsilon_{t-1} + \psi_{t-K_T-1})' - \{(T-K_T-1)^{-1}T^{1/2}D_T^{-1}\bar{\xi}_{-1}\} \bullet \{T^{-1/2}\sum_{t=K_T+2}^{T}(\epsilon_{t-1} + \psi_{t-K_T-1})'\},$$
$$\tilde{\Omega}_* = D_T^{-1}T^{-1}\sum_{t=K_T+2}^{T}\xi_{t-1}\omega'_{t-K_T-1;1} - \{(T-K_T-1)^{-1}T^{1/2}D_T^{-1}\bar{\xi}_{-1}\} \bullet \left(T^{-1/2}\sum_{t=K_T+2}^{T}\omega'_{t-K_T-1;1}\right),$$

with $\bar{\xi}_{-1} = T^{-1}\sum_{t=K_T+2}^{T}\xi_{t-1}$ and $\tilde{N}_{21} = \tilde{N}'_{12}$. Using (A.2) and (A.3) with tedious arguments, it can be led to that

$$D_T^{-1}\tilde{N}_{12}\check{C}'^{-1}_{K_T}$$
$$= \tilde{N}_*\check{I}'_{K_T} + \tilde{\Omega}_* + \tilde{Q}_* + O_p(T^{-1/2}K_T^{1/2}), \quad (A.21)$$

where

$$\tilde{Q}_* = \left[O_p(K_T^{-\bar{a}}),\ O_p((K_T-1)^{-\bar{a}}),\cdots,\ O_p((K_{2;T}+2)^{-\bar{a}}),\ O_p((K_{2;T}+1)^{-\bar{a}}),\ 0\right] + \left[O_p((2K_T)^{-\bar{a}}),\ O_p((2K_T-1)^{-\bar{a}}),\cdots,\ O_p((K_T+2)^{-\bar{a}}),\ O_p((K_T+1)^{-\bar{a}})\right],$$

noting that the terms expressed by the orders of probability are all matrices of $p \times p$ and 0 in

$$\left[O_p(K_T^{-\bar{a}}),\ O_p((K_T-1)^{-\bar{a}}),\cdots,\ O_p((K_{2;T}+2)^{-\bar{a}}),\ O_p((K_{2;T}+1)^{-\bar{a}}),\ 0\right]$$

is zero matrix of $p \times pK_{2;T}$. We note that (A.21) is associated with (A.8) of Saikkonen and Luukkonen (1997, p. 114). It can also be shown by the well-known asymptotics for $I(0)$ and $I(1)$ or (A.16) and the properties of the elements of ω_{t-K_T-1} in the statement subsequent to (A.3) that $\tilde{N}_* = O_p(1)$ and that $\tilde{\mathbf{\Omega}}_*\tilde{\mathbf{\Omega}}'_* = O_p(1)$. Furthermore, it can be derived from arguments similar to those used to derive (A.13) and (A.16) that

$$\check{I}'_{K_T}\bar{M}_{\eta;*}^{-1} = \left[O(K_T^{-1}),\ O(K_{2;T}^{-1/2})\right],$$

where the term expressed as $O(K_T^{-1})$ is $p \times pK_{1;T}$ and the one as $O(K_{2;T}^{-1/2})$ is $p \times pK_{2;T}$, and based on this result as well as (A.21), the following can be derived:

$$D_T^{-1}\tilde{N}_{12}\check{C}'^{-1}_{K_T}\bar{M}_{\eta;*}^{-1}$$
$$= \{\tilde{\Omega}_* + \tilde{Q}_* + O_p(T^{-1/2}K_T^{1/2})\}\bar{M}_{\eta;*}^{-1} + \tilde{N}_*\left[O(K_T^{-1}),\ O(K_{2;T}^{-1/2})\right]. \qquad (A.22)$$

Consequently, by using the second equation of (A.12), (A.13), (A.18), (A.19), (A.21) and (A.22) with the asymptotics of $\tilde{P}_2$, $\tilde{N}_*$ and $\tilde{\mathbf{\Omega}}_*$ above, it can be shown that

$$K_T^{-1/2}D_T^{-1}\tilde{N}_{12}\hat{M}_{22}^{-1}\tilde{N}_{21}D_T^{-1} = O_p(1). \qquad (A.23)$$

Also, from the results and the arguments used to derive (A.23), we obtain

$$K_T^{-1/2}D_T^{-1}\tilde{N}_{12}\hat{M}_{22}^{-1}\tilde{N}_{20}\bar{I}_{L_T;K_T} = O_p(1). \qquad (A.24)$$

For the derivation of other fundamental asymptotics, define

$$\bar{\epsilon}_0 = T^{-1}\sum_{t=K_T+2}^{T}\epsilon_t,$$
$$\bar{v}_{-1-i} = T^{-1}\sum_{t=K_T+2}^{T}v_{t-1-i}, \qquad i = 0, 1,$$
$$\bar{N}_{02} = T^{-1}\sum_{t=K_T+2}^{T}\epsilon_t(\Delta z_{t-1} - \mu_{K_T})' - (T-K_T-1)^{-1}T\,\bar{\epsilon}_0\,\bar{\phi}',$$
$$\hat{M}_{12;+} = T^{-1}\sum_{t=K_T+2}^{T}v_{t-1}(\Delta z_{t-1} - \mu_{K_T})' - (T-K_T-1)^{-1}T\,\bar{v}_{-1}\,\bar{\phi}',$$

where $\bar{N}_{20} = \bar{N}'_{02}$ and $\hat{M}_{21;+} = \hat{M}'_{12;+}$, partition $\hat{M}_{12;+}\check{C}'^{-1}_{K_T}$ into blocks as

$$\hat{M}_{12;+}\check{C}_{K_T}^{-1} = \left[\hat{M}_{12;+;1},\cdots\beta'\hat{M}_{12;+;K_T}\right],$$

where $\hat{M}_{12;+;i}$ of $p \times p$, $i = 1, \ldots, K_T$ and also put $\hat{M}_{21;+;i} = \hat{M}'_{12;+;i}$. It is easy to see from (A.2) and (A.3) as well as from the standard asymptotic theory for weakly stationary and ergodic series that

$$\bar{N}_{02}\check{C}'^{-1}_{K_T} = O_p(T^{-1/2}),$$
$$\hat{M}_{12;+;i} = C_{i-1}^{(1)}\Lambda + O(K_T^{-\bar{a}}) + O_p(T^{-1/2}), \quad i = 1,\cdots K_T. \qquad (A.25)$$

By suitably using the first result of (A.25) in addition to the second equation of (A.12), (A.13), (A.18) and the asymptotics of $\tilde{P}$ and based on arguments similar to those used to derive (A.24), it can be shown that

$$\bar{N}_{02}\hat{M}_{22}^{-1}\bar{N}_{20} = O_p(T^{-1}K_T). \qquad (A.26)$$

Also, based on (A.9) as $i=1$, the second equation of (A.12), (A.13), (A.18), (A.19) and (A.21) with the asymptotics of $\tilde{N}_*$ and $\tilde{Q}_*$ in addition to the first result of (A.25),

$$D_T^{-1}\tilde{N}_{12}\hat{M}_{22}^{-1}\bar{N}_{20} = o_p(T^{-1/2}K_T^{1/2}). \qquad (A.27)$$

Furthermore, it can be established by using (A.9) as $i=1$, (A.19), the property of $\tilde{\boldsymbol{\Omega}}_*$ and the second result of (A.25) that

$$D_T^{-1}\tilde{N}_{12}\hat{M}_{22}^{-1}\hat{M}_{21;+} = o_p(1). \qquad (A.28)$$

Similarly, based on arguments similar to those used to derive (A.26), the result of $\tilde{P}_2$, (A.19) and (A.25), it can be shown that

$$\begin{aligned}&\hat{M}_{12;+}\hat{M}_{22}^{-1}\hat{M}_{21;+}\\ &= \sum_{i=0}^{\infty} C_i^{(1)}\Lambda C_i^{(1)\prime} + O_p(K_T^{-1}),\end{aligned} \qquad (A.29)$$

and that

$$\bar{N}_{02}\hat{M}_{22}^{-1}\hat{M}_{21;+} = O_p(T^{-1/2}). \qquad (A.30)$$

On the other hand, defining

$$\tilde{d} = K_T^{-1/2}T^{-1/2}\sum_{t=K_T+2}^{T}\left(\sum_{s=1}^{L_T}\epsilon_{t-s}\right)\epsilon_t',$$

it can be shown based on the results obtained previously and via long and tedious arguments that

$$\begin{aligned}&\bar{I}_{L_T;K_T}'\tilde{N}_{02}\hat{M}_{22}^{-1}\bar{N}_{20}\\ &= T^{-1/2}K_T^{1/2}\tilde{d} - T^{-1/2}K_T^{1/2}(1/m)\hat{d}\\ &\quad + o_p(T^{-1/2}K_T^{1/2}),\\ &\quad \tilde{d} = O_p(1), \quad \hat{d} = O_p(1),\\ &\quad \tilde{d} = O_p(1), \quad \hat{d}^{-1} = O_p(1).\end{aligned} \qquad (A.31)$$

It is now remarked that (A.31) plays a decisive role in evaluating the difference between the test statistic and the nuisance parameter free quantity stated by (7) of Theorem 1. Also, from (A.20) and (A.29) we obtain

$$\hat{M}_{12;+}\hat{M}_{22}^{-1}\tilde{N}_{20}\bar{I}_{L_T;K_T} = O_p(K_T^{1/2}). \qquad (A.32)$$

Now, we turn to the derivation of asymptotics of $\tilde{S}_{ij}$ and their related quantities. For this purpose, define

$$\begin{aligned}\bar{M}_{10} &= T^{-1}\sum_{t=K_T+2}^{T} y_{t-1}\epsilon_t'\\ &\quad - (T-K_T-1)^{-1}T\overline{y}_{-1}\overline{\epsilon}_0',\\ \bar{S}_{10} &= \bar{M}_{10} - \hat{M}_{12}\hat{M}_{22}^{-1}\bar{N}_{20}, \quad \bar{S}_{01} = \bar{S}_{10}'.\end{aligned}$$

We emphasize the role of (5) to establish the asymptotics of $\beta'\bar{S}_{10}$, $\beta'\tilde{S}_{11}\beta$, $\beta'\tilde{S}_{10}$ and $\beta'\tilde{S}_{11}\gamma\bar{D}_T^{-1}$ below, and note that they are constructed based on residuals obtained by regressing $\beta' v_{t-1}$, ε_t, $\Delta y_t - \mu$ and $\bar{D}_T^{-1}\gamma' y_{t-1}$ on $\Delta z_{t-1}' - \mu_{K_T}'$ and 1 and that $\beta'\hat{M}_{12} = \beta'\hat{M}_{12;+}$. Substituting the right-hand side of (5) for $\beta' v_{t-1}$ in the expression of $\beta'\bar{S}_{10}$ and applying (A.5), $\bar{I}_{L_T;K_T}\overline{\eta} = O_p(T^{-1/2}K_T)$ shown easily and (A.31) to the expression obtained by the substitution, it can be led to that

$$\begin{aligned}&K_T\beta'\bar{S}_{10}\\ &= -T^{-1/2}K_T^{1/2}\bar{\alpha}'\hat{d} + o_p(T^{-1/2}K_T^{1/2}).\end{aligned} \qquad (A.33)$$

Similarly, it is led to by substitution similar to that used to derive (A.33) as well as (A.6) and (A.20) that

$$K_T\beta'\tilde{S}_{11}\beta = \bar{\alpha}'\Lambda\bar{\alpha} + o_p(1). \qquad (A.34)$$

Note that $N^{-1}\underline{C}_{11}$ of Saikkonen and Luukkonen (1997, pp. 114-115) corresponds to our $\beta'\tilde{S}_{11}\beta$, though their statistics are not demeaned versions, and that their (A.12) should be restated as $N^{-1}\underline{C}_{11} = O_p(K^{-1})$ in view of (A.34), emphasizing that $Eu_{1,t-1}u_{1,t-1}'$ is equal to $\sum_{1q}H'(H\sum_{qq}H')^{-1}H\sum_{1q}$, which also follows from our (A.29) and (A.44). Also, by substituting the right-hand side of (3) for $(\Delta y_t - \mu)'$ in the expression of $K_T\beta'\tilde{S}_{10}$, we obtain

$$\begin{aligned}&K_T\beta'\tilde{S}_{10}\\ &= K_T\beta'\tilde{S}_{11}\beta\alpha' + K_T\beta'\bar{S}_{10} + O_p(K_T^{-\bar{a}}).\end{aligned} \qquad (A.35)$$

Furthermore, based on substitution of the right－hand side of (5) for $v_{t-1}'\beta$ in the expression of $\bar{D}_T^{-1\prime}\gamma'\tilde{S}_{11}\beta$ and from (A.10), (A.24) and (A.32) with arguments similar those used to derive (A.33), (A.34) or (A.35), we obtain

$$\bar{D}_T^{-1\prime}\gamma'\tilde{S}_{11}\beta = -\bar{S}_{11;*}\bar{\alpha} + o_p(1), \qquad (A.36)$$

implying that (A.14) of Saikkonen and Luukkonen (1997, p. 115) should be replaced by a more suitable one. Similarly, substitution of the right-hand side of (3) for $\Delta y_t - \mu$ and $\Delta y_t' - \mu'$ in the expression of $\tilde{S}_{00}$, together with

$$T^{-1} \sum_{t=K_T+2}^{T} \epsilon_t \epsilon_t' - (T-K_T-1)^{-1} T \bar{\epsilon}_0 \bar{\epsilon}_0$$
$$= \Lambda + O_p(T^{-1/2})$$

as the standard result, (A.26) and (A.33) as well as arguments similar to those used to derive (A.34), yields

$$\tilde{S}_{00} = \Lambda + \alpha\beta'\tilde{S}_{11}\beta\alpha' + O_p(T^{-1/2}) + O_p(K_T^{-1-\bar{a}}). \quad (A.37)$$

On the other hand, the theory for standard Brownian motion formulates the limiting distributions of $D_T^{-1}\bar{N}_{11}D_T$, $\bar{N}_{01}D_T^{-1}$ and $\delta'\bar{N}_{01}D_T^{-1}$, resulting in (8). It is also led to by (2), (A.23), (A.28) and (A.29) that

$$\bar{D}_T^{-1\prime}\gamma'\tilde{S}_{11}\gamma\bar{D}_T^{-1}$$
$$= D_T^{-1}\bar{N}_{11}D_T^{-1} + O_p(T^{-1}K_T^{1/2}), \quad (A.38)$$

noting that (A.38) is a detailed version of (A.15) of Saikkonen and Luukkonen (1997, p. 115). Similarly, based on substitution of the right-hand side of (3) for $\Delta y_t - \mu$ in the expression of $\tilde{S}_{01}\gamma\bar{D}_T^{-1}$ and from (A.27), (A.30) and the definitions of $\tilde{S}_{11}$ and $\tilde{S}_{01}$, we obtain

$$\tilde{S}_{01}\gamma\bar{D}_T^{-1} = \bar{N}_{01}D_T^{-1} + \alpha\beta'\tilde{S}_{11}\gamma\bar{D}_T^{-1}$$
$$+ o_p(T^{-1/2}K_T^{1/2}) + O_p(K_T^{-\bar{a}}). \quad (A.39)$$

The remaining parts of the proof can be proved following manners similar to those used to show the counterparts for the conventional LR test (e.g., Johansen (1988) or Johansen (1996), ch. 11) based on (A.33) to (A.39).

Proof of Theorem 2 This theorem can be proved essentially based on a part of arguments used in the proof of Theorem 1. First, define the $p(K_{T-1})$-dimensional vectors

$$\eta_{t-2;-} = \left(\epsilon_{t-2}', \cdots, \epsilon_{t-K_T}'\right)', \quad \mu_{K_T-1} = (\mu', \cdots, \mu)',$$

and based on those, define

$$\bar{\eta}_{-1} = T^{-1} \sum_{t=K_T+2}^{T} \eta_{t-2;-},$$

$$\bar{\phi}_{-1} = T^{-1} \sum_{t=K_T+2}^{T} (\Delta z_{t-2;-} - \mu_{K_T-1}),$$

$$\tilde{N}_{30} = T^{-1} \sum_{t=K_T+2}^{T} (\Delta z_{t-2;-} - \mu_{K_T-1})\eta_{t-2;-}'$$
$$- (T-K_T-1)^{-1} T\bar{\phi}_{-1}\bar{\eta}_{-1}',$$

$$\tilde{N}_{13} = T^{-1} \sum_{t=K_T+2}^{T} \xi_{t-1}(\Delta z_{t-2;-} - \mu_{K_T-1})'$$
$$- (T-K_T-1)^{-1} T\bar{\xi}_{-2}\bar{\phi}_{-1}',$$

$$\bar{N}_{30} = T^{-1} \sum_{t=K_T+2}^{T} (\Delta z_{t-2;-} - \mu_{k_T-1})\epsilon_t'$$
$$- (T-K_T-1)^{-1} T\bar{\phi}_{-1}'\bar{\epsilon}_0,$$

$$\hat{M}_{31;+;i} = T^{-1} \sum_{t=K_T+2}^{T} (\Delta z_{t-2;-} - \mu_{K_T-1})v_{t-1-i}'$$
$$-(T-K_T-1)^{-1} T\bar{\phi}_{-1}\bar{v}_{-1-i}', \quad i = 0, 1.$$

Based on arguments similar to those to derive (A.24), it can be shown that

$$K_T^{-1/2}D_T^{-1}\tilde{N}_{13;0}\hat{M}_{33}^{-1}\bar{N}_{30}\bar{I}_{L_T;K_T-1} = O_p(1), \quad (A.40)$$

arguments similar to those to show (A.29) yields

$$\beta'\hat{M}_{13}\hat{M}_{33}^{-1}\hat{M}_{31}\beta$$
$$= \sum_{i=1}^{\infty} \beta' C_i^{(1)}\Lambda C_i^{(1)\prime}\beta + O_p(K_T^{-1}), \quad (A.41)$$

and by argumants similar to those to show (A.28),

$$D_T^{-1}\tilde{N}_{13}\hat{M}_{33}^{-1}\hat{M}_{31;+;i} = o_p(1). \quad (A.42)$$

Also, following the arguments used in (A.31) with minor alterations, it is easily led to that

$$\bar{I}_{L_T;K_T-1}'\tilde{N}_{03}\hat{M}_{33}^{-1}\bar{N}_{30}$$
$$= T^{-1/2}K_T^{1/2}\tilde{d} - T^{-1/2}K_T^{1/2}(1/m)\hat{d}$$
$$+ o_p(T^{-1/2}K_T^{1/2}). \quad (A.43)$$

Furthermore, it is trivial from the standard theory for weakly stationary and ergodic series to show that

$$\beta'\hat{M}_{11}\beta = \sum_{i=0}^{\infty} \beta' C_i^{(1)}\Lambda C_i^{(1)\prime}\beta + O_p(T^{-1/2}). \quad (A.44)$$

Then, it follows immediately from (A.42) and (A.44) that

$$\beta'\hat{S}_{11}\beta = \beta'\Lambda\beta + O_p(K_T^{-1}). \quad (A.45)$$

On the other hand, defining

$$\hat{N}_{11;**} = L_T^{-1}T^{-1}\sum_{t=K_T+2}^{T} y_{t-1}(\sum_{s=1}^{L_T}\epsilon_{t-1-s})' - L_T^{-1}(T-K_T-1)^{-1}T\,\overline{y}_{-1}\,\overline{\eta}'_{-1}\bar{I}_{L_T;K_T-1},$$

it follows from essentially the same arguments as those to show (A.10) that

$$\bar{D}_T^{-1\prime}\gamma'\hat{N}_{11;**} = \bar{S}_{11;*} + O_p(K_T^{-1}). \tag{A.46}$$

Using (A.42) and (A.46) etc., it can be led to that

$$\beta'\hat{S}_{11}\gamma\bar{D}_T^{-1} = \bar{S}_{11;**} + o_p(1), \tag{A.47}$$

and similarly to (A.38),

$$T^{-1}\bar{D}_T^{-1\prime}\gamma'\hat{S}_{11}\gamma\bar{D}_T^{-1} = D_T^{-1}\bar{N}_{11}D_T^{-1} + O_p(T^{-1}K_T^{1/2}). \tag{A.48}$$

Also, defining

$$\hat{S}_{01;1} = \hat{M}_{01;1} - \hat{M}_{03;1}\hat{M}_{33}^{-1}\hat{M}_{13},$$

based on substitution of the right-hand side of (3) and arguments similar to those used for (A.43), the asymptotics of $\hat{S}_{01;1}$, $\hat{S}_{00;i}$, $\hat{S}_{00;d}$ and $\hat{S}_{00;e}$ can be derived: for example, from (A.47) etc. we obtain

$$\hat{S}_{01;1}\gamma\bar{D}_T^{-1} = \bar{S}_{**} + o_p(1). \tag{A.49}$$

Consequently, it is established by using the results obtained above that

$$\hat{H}_1 - H_1 = T^{-1/2}K_T^{1/2}\tilde{C} + o_p(T^{-1/2}K_T^{1/2}) + O_p(K_T^{-\bar{a}}). \tag{A.50}$$

Based on (A.50) and the substitution of the right-hand side of (3) for $\Delta y_t' - \mu' - (\Delta y_{t-1}' - \mu')H_1'$, it is now established that

$$\beta'\hat{S}_{10} = \beta'\hat{S}_{11}\beta\alpha' + T^{-1/2}K_T^{1/2}\hat{C} + o_p(T^{-1/2}K_T) + O_p(K_T^{-\bar{a}}). \tag{A.51}$$

Similarly, by arguments similar to those used to derive (A.37) and (A.39), it is led to that

$$\hat{S}_{00} = \Lambda + \alpha\beta'\hat{S}_{11}\beta\alpha' + T^{-1/2}K_T^{1/2}\hat{D} + o_p(T^{-1/2}K_T^{1/2}) + O_p(K_T^{-\bar{a}}). \tag{A.52}$$

and that

$$\hat{S}_{01}\gamma\bar{D}_T^{-1} = \bar{N}_{01}D_T^{-1} + \alpha\beta'\hat{S}_{11}\gamma\bar{D}_T^{-1} + T^{-1/2}K_T^{1/2}\hat{F} + o_p(T^{-1/2}K_T^{1/2}) + O_p(K_T^{-\bar{a}}). \tag{A.53}$$

The rest of the proof can be proved as has been done for the counterpart in the proof of Theorem 1 based on (A.45), (A.47), (A.48) and (A.51) to (A.53).

Proof of Theorem 3 The results required for the theorem can be led to by suitably arranging the arguments used to prove Theorems 1 and 2, noting that the quantities accompanied by γ or δ vanish for the case $r = p$ and that the ones accompanied by β or $\bar{\alpha}$ or based on (4) or (5) are so for the case $r = 0$.

References

Banerjee, A., J.J. Dolado, J.W. Galbraith, and D.F. Hendry, 1993, Co-Integration, Error-Correction and the Econometric Analysis of Non-stationary Data, Oxford: Oxford University Press.

Bierens, H.J., 1997, Nonparametric cointegration analysis, Journal of Econome-trics 77, 379-404.

Elliott, G. and E. Pesavento, 2009, Testing the null of no cointegration when covariates are known to have a unit root, Econometric Theory 25, 1829-1850.

Engle, R.F. and C.W.J. Granger, 1987, Co-integration and error-correction: Representation, Estimation and Testing, Econometrica 55, 251-276.

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

Hamilton, J. D., 1994, Time Series Analysis, New Jersey, Princeton University Press.

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

Hjalmarsson, H. and P. Österholm, 2010, Testing for cointegration using the Johansen methodology when variables are near-integrated: size distortions and partial remedies, Empirical Economics 39, 51-76.

Johansen, S., 1988, Statistical analysis of cointegration vectors, Journal of Economic Dynamics and Control 12, 231-254.

Johansen, S., 1992, Determination of cointegrating rank in the presence of a linear trend, Oxford Bulletin of Economics and Statistics 54, 383-397.

Johansen, S., 1996, Likelihood-Based Inference in Cointegrated Vector Autoregre-ssive Models, (2nd

printing), Oxford, Oxford University Press.
Kurita, T. 2011, Local power of likelihood-based tests for cointegrating rank: comparative analysis of full and partial systems, Journal of Time Series Analysis 32, 672-679.
Lütkepohl, H. and P. Saikkonen, 1999, Order selection in testing for the cointegrating rank of a VAR process, in: R.F. Engle and H. White, (Eds.), Co-integration, Causality, and Forecasting, Festschrift in Honour of Clive Granger, Oxford, Oxford University Press, 200-211.
Lütkepohl, H. and P. Saikkonen, 2000, Testing for the cointegrating rank of a VAR process with a time trend, Journal of Econometrics 95, 177-198.
MacKinnon, J.G., Haug A. A. and L. Michelis, 1999, Numerical distribution functions of likelihood ratio tests for cointegration, Journal of Applied Econometrics 14, 563-577.
Park, J.Y. and P.C.B. Phillips, 1988, Statistical inference in regressions with integrated processes: Part I, Econometric Theory 4, 468-497.
Park, J.Y. and P.C.B. Phillips, 1989, Statistical inference in regressions with integrated processes: Part II, Econometric Theory 5, 95-131.
Perron, P. and S. Ng, 1996, Useful modifications to some unit root tests with dependent errors and their local asymptotic properties, Review of Economic Studies 63, pp 435-463.
Phillips, P. C. B., 1995, Fully modified least squares and vector autoregression, Econometrica 63, 1023-1078.
Qu, Z. and P. Perron, 2007, A modified information criterion for cointegration tests based on a VAR approximation, Econometric Theory 23, 638-685.
Saikkonen, P., 1992, Estimation and testing of cointegrated systems by an autoregressive approximation, Econometric Theory 9, 1-27.
Saikkonen, P. and R. Luukkonen, 1997, Testing cointegration in infinite order vector autoregressive processes, Journal of Econometrics 81, 93-126.
Shintani M., 2001, A simple cointegrating rank test without vector autoreg-ression, Journal of Econometrics 105, 337-362.

TABLE 1 *Relative frequencies of determining correctly*: The first group

DGP/Test	SL				CSL			
K_T	3	4	5	6	3	4	5	6
(i): $\breve{\mu}=0$								
				$T=150$				
$r=0$	0.891	0.869	0.822	0.783	0.645	0.617	0.903	0.875
$r=1$	0.931	0.83	0.727	0.592	0.811	0.813	0.867	0.811
$r=2$	0.907	0.81	0.782	0.546	0.905	0.917	0.813	0.791
$r=3$	0.823	0.545	0.527	0.295	0.88	0.75	0.754	0.6
				$T=250$				
$r=0$	0.938	0.932	0.918	0.906	0.81	0.815	0.982	0.978
$r=1$	0.965	0.954	0.946	0.901	0.897	0.896	0.986	0.962
$r=2$	0.938	0.946	0.931	0.901	0.937	0.945	0.934	0.937
$r=3$	0.952	0.928	0.914	0.818	0.941	0.954	0.948	0.929
(ii): $\breve{\mu}=0$								
				$T=150$				
$r=0$	0.824	0.83	0.731	0.712	0.589	0.618	0.984	0.979
$r=1$	0.829	0.648	0.587	0.491	0.879	0.898	0.339	0.308
$r=2$	0.606	0.548	0.359	0.355	0.875	0.88	0.103	0.1
$r=3$	0.317	0.153	0.139	0.088	0.655	0.372	0.468	0.295
				$T=250$				
$r=0$	0.898	0.913	0.877	0.885	0.761	0.805	0.997	0.998
$r=1$	0.935	0.913	0.865	0.785	0.923	0.942	0.533	0.454
$r=2$	0.828	0.84	0.67	0.693	0.928	0.943	0.173	0.161
$r=3$	0.661	0.578	0.464	0.391	0.909	0.819	0.83	0.699
(iii): $\breve{\mu}=0$								
				$T=150$				
$r=0$	0.902	0.892	0.859	0.817	0.428	0.385	0.548	0.473
$r=1$	0.945	0.912	0.836	0.723	0.635	0.639	0.443	0.437
$r=2$	0.933	0.924	0.888	0.8	0.916	0.909	0.811	0.767
$r=3$	0.877	0.81	0.717	0.568	0.874	0.841	0.798	0.733
				$T=250$				
$r=0$	0.938	0.943	0.933	0.922	0.633	0.623	0.827	0.802
$r=1$	0.954	0.962	0.957	0.941	0.759	0.772	0.438	0.445
$r=2$	0.947	0.947	0.944	0.94	0.944	0.944	0.907	0.884
$r=3$	0.959	0.96	0.959	0.947	0.943	0.949	0.948	0.947

TABLE 1 (continued)

DGP/Test	SL				CSL			
K_T	3	4	5	6	3	4	5	6
(i): $\breve{\mu} \neq 0$								
				$T = 150$				
$r = 0$	0.776	0.716	0.659	0.564	0.432	0.392	0.908	0.864
$r = 1$	0.903	0.829	0.779	0.67	0.695	0.672	0.89	0.83
$r = 2$	0.625	0.373	0.379	0.27	0.786	0.809	0.724	0.681
$r = 3$	0.888	0.708	0.686	0.489	0.916	0.841	0.517	0.37
				$T = 250$				
$r = 0$	0.863	0.844	0.825	0.778	0.588	0.583	0.979	0.971
$r = 1$	0.931	0.904	0.908	0.873	0.77	0.775	0.986	0.963
$r = 2$	0.865	0.713	0.689	0.522	0.856	0.877	0.89	0.879
$r = 3$	0.945	0.912	0.909	0.856	0.944	0.928	0.731	0.655
(ii): $\breve{\mu} \neq 0$								
				$T = 150$				
$r = 0$	0.675	0.675	0.556	0.528	0.394	0.42	0.972	0.965
$r = 1$	0.809	0.713	0.652	0.574	0.737	0.723	0.406	0.372
$r = 2$	0.436	0.45	0.317	0.335	0.771	0.8	0.166	0.155
$r = 3$	0.422	0.276	0.233	0.168	0.763	0.528	0.003	0.002
				$T = 250$				
$r = 0$	0.797	0.823	0.763	0.765	0.575	0.624	0.995	0.995
$r = 1$	0.891	0.875	0.855	0.81	0.802	0.799	0.604	0.516
$r = 2$	0.654	0.685	0.499	0.543	0.84	0.872	0.26	0.245
$r = 3$	0.704	0.674	0.55	0.515	0.897	0.87	0.007	0.004
(iii): $\breve{\mu} \neq 0$								
				$T = 150$				
$r = 0$	0.806	0.792	0.733	0.672	0.27	0.225	0.437	0.361
$r = 1$	0.901	0.879	0.825	0.757	0.719	0.697	0.454	0.444
$r = 2$	0.709	0.582	0.453	0.38	0.783	0.768	0.678	0.646
$r = 3$	0.919	0.882	0.836	0.751	0.92	0.89	0.672	0.558
				$T = 250$				
$r = 0$	0.855	0.865	0.844	0.82	0.415	0.39	0.717	0.69
$r = 1$	0.923	0.924	0.913	0.899	0.821	0.815	0.449	0.45
$r = 2$	0.887	0.868	0.799	0.718	0.85	0.846	0.83	0.815
$r = 3$	0.946	0.936	0.928	0.919	0.943	0.933	0.868	0.819

TABLE 2 *Relative frequencies of determining r correctly*: The second group

DGP/Test	SL				CSL			
K_T	3	4	5	6	3	4	5	6
(*i*): $\check{\mu}=0$								
				$T=150$				
$r=0$	0.943	0.901	0.893	0.839	0.57	0.415	0.971	0.939
$r=1$	0.919	0.659	0.691	0.497	0.806	0.768	0.963	0.946
$r=2$	0.61	0.66	0.404	0.379	0.847	0.84	0.886	0.881
$r=3$	0.946	0.938	0.839	0.627	0.947	0.965	0.958	0.904
				$T=250$				
$r=0$	0.961	0.942	0.946	0.927	0.783	0.677	0.998	0.995
$r=1$	0.961	0.946	0.938	0.832	0.894	0.873	0.99	0.988
$r=2$	0.943	0.945	0.83	0.823	0.898	0.896	0.99	0.99
$r=3$	0.951	0.97	0.958	0.964	0.952	0.972	0.999	0.999
(*ii*): $\check{\mu}=0$								
				$T=150$				
$r=0$	0.908	0.846	0.846	0.769	0.855	0.737	0.993	0.986
$r=1$	0.786	0.483	0.555	0.424	0.896	0.81	0.758	0.657
$r=2$	0.433	0.665	0.29	0.39	0.801	0.877	0.538	0.656
$r=3$	0.928	0.593	0.628	0.268	0.923	0.867	0.386	0.178
				$T=250$				
$r=0$	0.944	0.912	0.927	0.9	0.923	0.869	1.0	1.0
$r=1$	0.955	0.843	0.862	0.689	0.948	0.94	0.959	0.896
$r=2$	0.861	0.935	0.696	0.817	0.941	0.931	0.884	0.934
$r=3$	0.954	0.964	0.955	0.887	0.955	0.965	0.884	0.653
(*iii*): $\check{\mu}=0$								
				$T=150$				
$r=0$	0.914	0.915	0.901	0.875	0.155	0.122	0.915	0.888
$r=1$	0.936	0.867	0.757	0.645	0.579	0.545	0.884	0.859
$r=2$	0.676	0.643	0.506	0.422	0.706	0.727	0.966	0.961
$r=3$	0.96	0.959	0.939	0.843	0.953	0.954	0.987	0.975
				$T=250$				
$r=0$	0.932	0.945	0.944	0.939	0.348	0.339	0.991	0.987
$r=1$	0.949	0.957	0.953	0.931	0.731	0.725	0.964	0.959
$r=2$	0.942	0.939	0.893	0.844	0.794	0.819	0.999	0.999
$r=3$	0.959	0.961	0.962	0.961	0.955	0.959	0.998	0.996

TABLE 2 (continued)

DGP/Test	SL				CSL			
K_T	3	4	5	6	3	4	5	6
(*i*): $\check{\mu} \neq 0$								
				$T = 150$				
$r = 0$	0.88	0.792	0.801	0.701	0.563	0.404	0.953	0.91
$r = 1$	0.912	0.739	0.76	0.602	0.653	0.575	0.951	0.908
$r = 2$	0.716	0.741	0.556	0.53	0.739	0.687	0.91	0.906
$r = 3$	0.924	0.903	0.882	0.775	0.926	0.889	0.969	0.934
$T = 250$								
$r = 0$	0.918	0.862	0.886	0.834	0.718	0.598	0.996	0.99
$r = 1$	0.936	0.882	0.913	0.842	0.729	0.68	0.988	0.977
$r = 2$	0.923	0.885	0.872	0.843	0.8	0.765	0.994	0.994
$r = 3$	0.934	0.913	0.925	0.908	0.933	0.915	0.992	0.988
(*ii*): $\check{\mu} \neq 0$								
				$T = 150$				
$r = 0$	0.833	0.699	0.735	0.594	0.604	0.467	0.987	0.975
$r = 1$	0.83	0.609	0.653	0.528	0.722	0.673	0.818	0.717
$r = 2$	0.57	0.719	0.416	0.53	0.801	0.758	0.634	0.729
$r = 3$	0.862	0.74	0.748	0.454	0.876	0.863	0.566	0.322
				$T = 250$				
$r = 0$	0.894	0.804	0.856	0.782	0.715	0.623	0.999	0.998
$r = 1$	0.926	0.847	0.881	0.772	0.778	0.764	0.972	0.927
$r = 2$	0.902	0.851	0.798	0.827	0.891	0.816	0.932	0.957
$r = 3$	0.907	0.902	0.902	0.885	0.902	0.897	0.907	0.788
(*iii*): $\check{\mu} \neq 0$								
				$T = 150$				
$r = 0$	0.86	0.846	0.819	0.771	0.236	0.187	0.918	0.884
$r = 1$	0.909	0.872	0.808	0.726	0.622	0.562	0.856	0.818
$r = 2$	0.79	0.741	0.641	0.571	0.696	0.689	0.975	0.972
$r = 3$	0.922	0.914	0.907	0.876	0.901	0.519	0.961	0.949
				$T = 250$				
$r = 0$	0.88	0.895	0.885	0.868	0.369	0.352	0.992	0.989
$r = 1$	0.925	0.92	0.917	0.899	0.74	0.708	0.961	0.953
$r = 2$	0.932	0.913	0.898	0.87	0.752	0.765	1.0	0.999
$r = 3$	0.941	0.936	0.934	0.925	0.952	0.862	0.984	0.98