Estimation and Testing for Partially Nonstationary Vector Autoregressive Models with GARCH: WLS versus QMLE *

Chor-yiu SIN
Department of Economics
Hong Kong Baptist University, Hong Kong

Preliminary version. Comments are welcome.

Abstract

Macroeconomic or financial data are often modelled with cointegration and GARCH. Noticeable examples include those studies of price discovery, in which stock prices of the same underlying asset are cointegrated and they exhibit multivariate GARCH. Modifying the asymptotic theories developed in Li, Ling and Wong (2001) and Sin and Ling (2004), this paper proposes a WLS(weighted least squares) for the parameters of an ECM(error-correction model). Apart from its computational simplicity, by construction, the consistency of WLS is insensitive to possible misspecification in conditional variance. Further, asymmetrically distributed deflated error is allowed, at the expense of more involved asymptotic distributions of the statistics. Efficiency loss relative to QMLE(quasi-maximum likelihood estimator) is discussed within the class of LABF(locally asymptotically Brownian functional) models. The insensitivity and efficiency of WLS in finite samples are examined through Monte Carlo experiments. We also apply the WLS to an empirical example of HSI(Hang Seng Index), HSIF(Hang Seng Index Futures) and TraHK(Hong Kong Tracker Fund).

Key Words: Asymmetric distribution; Cointegration; LABF models; Multivariate GARCH; Price discovery; WLS.

JEL Codes: C32, C51, G14

* *Acknowledgments:* I thank helpful comments and/or assistance from Chi-shing CHAN and Shiqing LING. This research is partially supported by the Hong Kong Research Grant Council competitive earmarked grant HKBU2014/02H. The usual disclaimers apply.

1 Introduction

Throughout this paper, we consider an $m-$dimensional autoregressive (AR) process $\{Y_t\}$, which is generated by

$$Y_t = \Phi_1 Y_{t-1} + \cdots + \Phi_s Y_{t-s} + \varepsilon_t, \tag{1.1}$$

$$E(\varepsilon_t \mid \Sigma_{t-1}) = E((\varepsilon_{1t}, \ldots, \varepsilon_{mt})' \mid \Sigma_{t-1}) = 0, \tag{1.2}$$

where Φ_j's are constant matrices, and Σ_{t-1} is an increasing σ-algebra.

Assuming the ε_t's are i.i.d., under further conditions on Φ_j's (See Assumptions 2.1, 2.2 and 2.3 below), Johansen (1988) (see also Ahn and Reinsel, 1990) shows that, although some component series of $\{Y_t\}$ exhibit nonstationary behaviour, there are r linear combinations of $\{Y_t\}$ that are stationary. This phenomenon, which is called cointegration in the literature of economics, was first investigated in Granger (1983) (see also Engle and Granger, 1987). The partially nonstationary multivariate AR model or cointegrating time series models without GARCH have been extensively discussed over the past twenty years. Other noticeable examples include Phillips and Durlauf (1986), Stock and Watson (1993), Johansen (1996), and Rahbek and Mosconi (1999).

Economic time series related to financial markets often exhibit time-varying variances. As far as we know, Li, Ling and Wong (2001) (henceforth LLW (2001)) first investigate multivariate time series that exhibit both cointegration and time-varying variances. In LLW (2001), the heteroskedasticity part is the random coefficient AR model [see e.g. Tsay (1987)] and thus the scope of applications is relatively limited. Sin and Ling (2004) modify LLW (2001)'s model a bit and consider a multivariate GARCH model first suggested by Bollerslev (1990) and widely used in many papers in the literature. More precisely, the conditional variance-covariance matrix, denoted as V_t is *modelled* as $D_t \Gamma D_t$, where $D_t = diag(\sqrt{h_{1t}}, \ldots, \sqrt{h_{mt}})$ and:

$$h_{it} \quad = \quad a_{i0} + \sum_{j=1}^{q} a_{ij} \varepsilon_{it-j}^2 + \sum_{k=1}^{p} b_{ik} h_{it-k}, \tag{1.3}$$

$$\Gamma \equiv (\sigma_{ij})_{m\times m}, \text{ a symmetric positive definite matrix with } \sigma_{ii} = 1. \quad (1.4)$$

Following Sin and Ling (2004), this paper assumes the existence of some *pseudo* true parameters of this multivariate GARCH process, which satisfies Assumptions 2.4-2.5 below. However, in view of the possible misspecification in variance (see, for instance, the GJR model first suggested in Glosten, Jagannathan and Runkle, 1993 and the time-varying correlation model first suggested in Tse and Tsui, 2002), instead of a QMLE(quasi-maximum likelihood estimator), we consider a WLS(weighted least squares), which is computationally simpler. Unlike Sin and Ling (2004), asymmetrically distributed deflated error is allowed, at the expense of a more involved distribution. Efficiency loss relative to QMLE is discussed within the class of LABF(locally asymptotically Brownian functional) models.

In this paper, we first investigate the full rank and the reduced rank WLS. Using these two estimators, we construct a Wald-type test for reduced rank. We show that the asymptotic distribution of this test is a functional of a standard Brownian motion and a standard normal vector with d unknown nuisance parameters, where $d \equiv m - r$. The critical value can thus be simulated via Monte Carlo method. It is expected that the test based on the WLS of process (1.1)-(1.4) is more powerful than Johansen's test or Reinsel-Ahn's test which ignores GARCH.

This paper proceeds as follows. Section 2 discusses the structure of model (1.1)-(1.4). Section 3 and section 4 derive the asymptotic distribution of the full rank estimators and the reduced rank estimators, respectively. Section 5 devises a test for reduced rank. The extension to asymmetric distribution, the efficiency loss, the Monte Carlo experiments and an illustrative empirical example are discussed in the subsequent sections. We conclude in the last section.

2 Basic Properties of Models

Denote L as the lag operator. Refer to (1.1)-(1.2) and define $\Phi(L) = I_m - \sum_{j=1}^{s} \Phi_j L^j$. We first make the following assumption:

Assumption 2.1. *$\mid \Phi(z) \mid = 0$ implies that either $\mid z \mid > 1$ or $z = 1$.* □

Define $W_t = Y_t - Y_{t-1}$, $\Phi_j^* = -\sum_{k=j+1}^{s} \Phi_k$ and $C = -\Phi(1) = -(I_m - \sum_{j=1}^{s} \Phi_j)$. By a Taylor's formula, $\Phi(L)$ can be decomposed as:

$$\Phi(z) = (1-z)I_m - Cz - \sum_{j=1}^{s-1} \Phi_j^*(1-z)z^j. \tag{2.1}$$

Thus, we can reparameterize process (1.1) as:

$$W_t = CY_{t-1} + \sum_{j=1}^{s-1} \Phi_j^* W_{t-j} + \varepsilon_t. \tag{2.2}$$

Following Johansen (1988,1996) and Reinsel and Ahn (1990), we can decompose $C = AB$, where A and B are respectively $m \times r$ and $r \times m$ matrices of rank r. Define $d = m - r$. Denote $B_\perp$ as a $d \times m$ matrix of full rank such that $BB'_\perp = 0_{r\times d}$, $\bar{B} = (BB')^{-1}B$ and $\bar{B}_\perp = (B_\perp B'_\perp)^{-1}B_\perp$, and $A_\perp$ as an $m \times d$ matrix of full rank such that $A'A_\perp = 0_{r\times d}$, $\bar{A} = A(A'A)^{-1}$ and $\bar{A}_\perp = A_\perp(A'_\perp A_\perp)^{-1}$. We impose the following condition:

Assumption 2.2. *$\mid A'_\perp(I_m - \sum_{j=1}^{s-1} \Phi_j^*)B'_\perp \mid \neq 0$.* □

Assumption 2.3. *$E(\varepsilon_t\varepsilon_t') < \infty$ and $E(vec[\varepsilon_t\varepsilon_t']vec[\varepsilon_t\varepsilon_t']') < \infty$.* □

By the proof of Theorem (4.2) in Johansen (1996),

$$\tilde{\Phi}(L) \begin{bmatrix} (1-L)B_\perp Y_t \\ BY_t \end{bmatrix} = (\bar{A}_\perp, \bar{A})'\varepsilon_t, \tag{2.3}$$

where $\tilde{\Phi}(z) = (\bar{A}_\perp, \bar{A})'\Phi(z)(\bar{B}'_\perp, \bar{B}'(1-z)^{-1})$ is invertible for $\mid z \mid < 1 + \rho$ for some $\rho > 0$. Denote $Q' = [Q_1, Q_2]$, where $Q'_1 = B_\perp$ and $Q'_2 = B$. Let $P = Q^{-1} = [P_1, P_2]$, where $P_1 = \bar{B}'_\perp$ and $P_2 = \bar{B}'$. Thus,

$$P_1Q'_1 + P_2Q'_2 = I_m,\ Q'_1P_1 = I_d,\ Q'_1P_2 = 0_{d\times r},\ Q'_2P_1 = 0_{r\times d} \text{ and } Q'_2P_2 = I_r.$$

Define $Z_t = QY_t \equiv (Z_{1t}, Z_{2t})'$. As in Johansen (1988, 1996) and Ahn and Reinsel (1990), we have the following decomposition:

$$Z_{1t} = Q_1' Y_t = Z_{1t-1} + u_{1t}, \text{ and } Z_{2t} = Q_2' Y_t = u_{2t}, \tag{2.4}$$

where $u_t = (u_{1t}', u_{2t}')' = \psi(L) a_t$, $\psi(L) \equiv \tilde{\Phi}^{-1}(L)$ and $a_t \equiv (\bar{A}_\perp, \bar{A})' \varepsilon_t$. By Assumption 2.3, ε_t is an $I(0)$ process. Thus, Z_{1t} is $I(1)$ while Z_{2t} is $I(0)$.

We close this section with the following assumptions on (1.3)-(1.4).

Assumption 2.4. *For $i = 1, \ldots, m$, $a_{i0} > 0$, $a_{i1}, \ldots, a_{iq}, b_{i1}, \ldots, b_{ip} \geq 0$, and $\sum_{j=1}^q a_{ij} + \sum_{k=1}^p b_{ik} < 1$.* □

Assumption 2.5. *For $i = 1, \ldots, m$, define $\eta_{it} \equiv \varepsilon_t / \sqrt{h_{it}}$. All eigenvalues of $E(A_{it} \otimes A_{it})$ lie inside the unit circle, where $\otimes$ denotes the Kronecker product and*

$$A_{it} = \begin{pmatrix} a_{i1}\eta_{it}^2 & \ldots & a_{iq}\eta_{it}^2 & b_{i1}\eta_{it}^2 & \ldots & b_{ip}\eta_{it}^2 \\ & I_{q-1} & 0_{(q-1)\times 1} & & 0_{(q-1)\times p} & \\ a_{i1} & \ldots & a_{iq} & b_{i1} & \ldots & b_{ip} \\ & 0_{(p-1)\times q} & & & I_{p-1} & 0_{(p-1)\times 1} \end{pmatrix}. \quad \square$$

Assumption 2.6. *$\eta_t \equiv (\eta_{1t}, \ldots, \eta_{mt})'$ is symmetrically distributed.* □

3 Full Rank Estimation

We first let $X_{t-1} \equiv [Y_{t-1}', W_{t-1}', \ldots, W_{t-s+1}']'$, $\varphi \equiv vec[C, \Phi_1^*, \ldots, \Phi_{s-1}^*]$ and $\delta \equiv [\delta_1', \delta_2']'$, where $\delta_1 \equiv [a_0', a_1', \ldots, a_q', b_1', \ldots, b_p']'$, $a_j \equiv [a_{1j}, ..., a_{mj}]'$, $b_k \equiv [b_{1k}, ..., b_{mk}]'$, $j = 0, 1, ..., q$, $k = 1, ..., p$, and $\delta_2 \equiv \tilde{\nu}(\Gamma)$, which is obtained from $vec(\Gamma)$ by eliminating the supradiagonal and the diagonal elements of Γ [see Magnus (1988, p.27)].

Given $\{Y_t : t = 1, \cdots, n\}$, conditional on the initial values $Y_s = 0$ for $s \leq 0$, the *normal* log-likelihood function (LF) (with a constant ignored) can be written as

$$l(\tilde{\varphi}, \tilde{\delta}) = \sum_{t=1}^n \tilde{l}_t \text{ and } \tilde{l}_t = -\frac{1}{2} \tilde{\varepsilon}_t' \tilde{V}_t^{-1} \tilde{\varepsilon}_t - \frac{1}{2} \ln |\tilde{V}_t|, \tag{3.1}$$

where $\tilde{V}_t = \tilde{D}_t \tilde{\Gamma} \tilde{D}_t$. In (3.1), $\tilde{\varepsilon}_t$ and $\tilde{V}_t$ are functions of the *generic* parameter $(\tilde{\varphi}, \tilde{\delta})$. Further denote $\tilde{h}_t = (\tilde{h}_{1t}, \ldots, \tilde{h}_{mt})'$ and $\vec{\tilde{h}}_t = (\tilde{h}_{1t}^{-1}, \ldots, \tilde{h}_{mt}^{-1})'$. Using the Hadamard product $\odot$ [see Magnus and Neudecker (1988, p.27)], the score function, with respect to $\tilde{\varphi}$, can be written as

$$\nabla_{\varphi}\tilde{l}_t = -\frac{1}{2}\nabla_{\varphi}\tilde{h}_t(\iota - w(\tilde{\varepsilon}_t\tilde{\varepsilon}_t'\tilde{V}_t^{-1})) \odot \vec{\tilde{h}}_t + (X_{t-1}\otimes I_m)\tilde{V}_t^{-1}\tilde{\varepsilon}_t, \qquad (3.2)$$

where $\nabla_x f$ denotes $\partial f/\partial x$, $\iota = (1, 1, \ldots, 1)'_{m\times 1}$ and $w(A)$ is a vector containing the diagonal elements of the square matrix A. In Sin and Ling (2004), the score function (3.2) is used. As one can see in that paper, the algorithm for the one-step estimator is quite involved. More importantly, if the multivariate GARCH is misspecified-specified and for all $(\tilde{\varphi}, \tilde{\delta})$, $Prob\{E[\nabla_{\varphi}\tilde{h}_t(\iota - w(\tilde{\varepsilon}_t\tilde{\varepsilon}_t'\tilde{V}_t^{-1})) \odot \vec{\tilde{h}}_t \mid \Sigma_{t-1}] = 0\} < 1$, it is unclear what the asymptotic properties of the one-step estimator carries. In view of that, for our *WLS*, we only consider the second part of the score function:

$$\tilde{f}_t \equiv (X_{t-1}\otimes I_m)\tilde{V}_t^{-1}\tilde{\varepsilon}_t. \qquad (3.3)$$

Denote $\bar{Q}^* = \text{diag}(Q\otimes I_m, I_{(s-1)m^2})$ and $\bar{D}^* = \text{diag}(nI_{dm}, \sqrt{n}I_{rm+(s-1)m^2})$. For any fixed positive constant K, let $\Theta_n \equiv \{(\tilde{\varphi}, \tilde{\delta}) : \|\bar{D}^*\bar{Q}^{*'-1}(\tilde{\varphi}-\varphi)\| \le K \ \textit{and}\ \|\sqrt{n}(\tilde{\delta}-\delta)\| \le K\}$, where (φ, δ) is the *true* parameter. Using Assumptions 2.1-2.5 and a similar method as in Ling and Li (1998), the derivative of $\tilde{f}_t$ on Θ_n can be *simplified* as follows:

$$\bar{D}^{*-1}\bar{Q}^*(\sum_{t=1}^{n}\nabla_{\varphi'}\tilde{f}_t)\bar{Q}^{*'}\bar{D}^{*-1} = \sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^*\tilde{F}_t\bar{Q}^{*'}\bar{D}^{*-1} + o_p(1), \qquad (3.4)$$

where $o_p(1)$ denotes convergence to zero in probability, and $\tilde{F}_t \equiv -(X_{t-1}X'_{t-1}\otimes\tilde{V}_t^{-1})$.

Similar to the arguments in Ling et al. (2003) and Ling and Li (2003), we can show that the following results hold uniformly in Θ_n:

$$\sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^*(\tilde{F}_t - F_t)\bar{Q}^{*'}\bar{D}^{*-1} = o_p(1), \qquad (3.5)$$

$$\sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^*(\tilde{f}_t - f_t) = \sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^*F_t(\tilde{\varphi}-\varphi) + o_p(1), \qquad (3.6)$$

In practice, we first find an initial estimator $(\tilde{\varphi}, \tilde{\delta})$ such that $\bar{D}^*\bar{Q}^{*'-1}(\tilde{\varphi}-\varphi) = O_p(1)$ and $\sqrt{n}(\tilde{\delta}-\delta) = O_p(1)$. For instance, it can be obtained following the procedure in LLW (2001) and Ling et al. (2003). Using this initial estimator and a one-step iteration as in Ling and Li (2003), we obtain a new estimator $(\dot{\varphi}, \dot{\delta})$ such that:

$$\bar{D}^*\bar{Q}^{*'-1}(\dot{\varphi}-\varphi) = -(\sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^* F_t \bar{Q}^{*'}\bar{D}^{*-1})^{-1}(\sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^* f_t) + o_p(1). \quad (3.7)$$

Let $(W_m'(u),\, W_m^{*\prime}(u))'$ be a $2m-$dimensional Brownian motion (BM) with the covariance matrix:

$$u\Omega \equiv u\begin{pmatrix} V_* & I_m \\ I_m & \Omega_1^* \end{pmatrix},$$

where $V_* = E\varepsilon_t\varepsilon_t'$, and $\Omega_1^* = E(V_t^{-1}\varepsilon_t\varepsilon_t' V_t^{-1})$. Let $B_d(u) = \Omega_{a_1}^{-1/2}[I_d, 0]\,\Omega_a^{1/2}V_*^{-1/2}W_m(u)$, where $\Omega_a = E(a_t a_t')$ and $\Omega_{a_1} = [I_d, 0]\Omega_a[I_d, 0]'$. We first give the following basic lemma, which resembles Lemma 3.1 in Sin and Ling (2004).

Lemma 3.1. *Suppose Assumptions 2.1-2.6 hold. Then*

$$(a) \quad \sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^*\nabla_{\varphi} l_t \longrightarrow_{\mathcal{L}} \Big\{vec[(\int_0^1 B_d(u)dW_m^*(u)')'\Omega_{a_1}^{1/2}\psi_{11}']', [N(0,\Omega_2^*)]'\Big\}',$$

$$(b) \quad -\sum_{t=1}^{n}\bar{D}^{*-1}\bar{Q}^* F_t\bar{Q}^*\bar{D}^{*-1} \longrightarrow_{\mathcal{L}} diag\Big\{\psi_{11}\Omega_{a_1}^{1/2}\int_0^1 B_d(u)B_d(u)'\Omega_{a_1}^{1/2}\psi_{11}' \otimes \Omega_1, \Omega_2\Big\},$$

where $\longrightarrow_{\mathcal{L}}$ *denotes convergence in distribution,* $\psi_{11} \equiv [I_d, 0](\sum_{k=1}^{\infty}\psi_k)[I_d, 0]'$, $\Omega_1 \equiv E(V_t^{-1})$, $\Omega_2 \equiv E(U_{t-1}U_{t-1}' \otimes V_t^{-1})$, $\Omega_2^* \equiv E(U_{t-1}U_{t-1}' \otimes V_t^{-1}\varepsilon_t\varepsilon_t' V_t^{-1})$, *and* $U_t = [(BY_t)', W_t', \cdots, W_{t-s+2}']'$. □

The following theorem comes from Lemma 3.1.

Theorem 3.1. *Under the assumptions in Lemma 3.1,*

$$(a) \quad n(\dot{C}-C)P_1 \longrightarrow_{\mathcal{L}} \Omega_1^{-1}M^*,$$

$$(b) \quad \sqrt{n}vec[(\dot{C}-C)P_2, (\dot{\Phi}_1^*-\Phi_1^*), \ldots, (\dot{\Phi}_{s-1}^*-\Phi_{s-1}^*)] \longrightarrow_{\mathcal{L}} N(0, \Omega_2^{-1}\Omega_2^*\Omega_2^{-1}),$$

where $M^* = (\int_0^1 B_d(u)dW_m^*(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}\Omega_{a_1}^{-1/2}\psi_{11}^{-1}$. □

When $E(\varepsilon_t\varepsilon_t' \mid \Sigma_{t-1}) = V_t$, $\Omega_1^* = \Omega_1$ and $\Omega_2^* = \Omega_2$. On the other hand, when the h_{it}'s are not constant, $\dot{C}$ is more efficient than the LSE of C in Ahn and Reinsel (1990), in the sense discussed in Ling and McAleer (2003b). Moreover, the simplicity of the distributions in Theorem 3.1(a)-(b) relies on the symmetry assumption (Assumption 2.6). Detailed discussions on these and the related issues can be found in subsequent sections below.

4 Reduced Rank Estimation

We first rewrite (2.2) in a reduced rank form:

$$W_t = ABY_{t-1} + \sum_{j=1}^{s-1} \Phi_j^* W_{t-j} + \varepsilon_t, \tag{4.1}$$

where A and B are defined as in section 2. Denote $\alpha = [\alpha_1', \alpha_2']'$ with $\alpha_1 \equiv vec[B]$ and $\alpha_2 \equiv vec[A, \Phi_1^*, \dots, \Phi_{s-1}^*]$. In the next sub-section, we first show the asymptotic properties of Johansen's estimator, which is used as an initial estimator for the reduced rank estimation that incorporates GARCH.

4.1 Initial Estimator for Parameters in AR Part

Johansen's estimator is essentially the QMLE which ignores the possible GARCH, i.e., the maximizer of the LF in (3.1) with $V_t(\tilde{\varphi}, \tilde{\delta})$ replaced by a constant matrix $\tilde{V}_*$. Denote this estimator as $\hat{\alpha} = [\hat{\alpha}_1', \hat{\alpha}_2']'$ with $\hat{\alpha}_1 = vec[\hat{B}]$ and $\hat{\alpha}_2 = vec[\hat{A}, \hat{\Phi}_1^*, \dots, \hat{\Phi}_{s-1}^*]$. Similar to Lemma 13.2 in Johansen (1996), we obtain the asymptotic distributions of the *normalized* estimators for α_1 and α_2 as follows. The details are omitted.

Theorem 4.1. *Suppose Assumptions 2.1-2.5 hold. Then*

(a) $\quad n((\hat{B}\bar{B}')^{-1}\hat{B} - B)P_1 \longrightarrow_{\mathcal{L}} (A'V_*^{-1}A)^{-1}A'V_*^{-1}(A_\perp, A)M,$

(b) $\quad \sqrt{n} vec[(\hat{A}(\hat{B}\bar{B}') - A), (\hat{\Phi}_1^* - \Phi_1^*), \dots, (\hat{\Phi}_{s-1}^* - \Phi_{s-1}^*)] \longrightarrow_{\mathcal{L}} N(0, \Sigma_2^{-1}\Sigma_2^*\Sigma_2^{-1}),$

where $M = \Omega_a^{1/2}(\int_0^1 B_d(u)dB_m(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}\Omega_{a1}^{-1/2}\psi_{11}^{-1}$, $B_m(u) = V_*^{-1/2}W_m(u)$, $\Sigma_2 = E(U_{t-1}U_{t-1}' \otimes I_m)$, $\Sigma_2^* = E(U_{t-1}U_{t-1}' \otimes \varepsilon_t\varepsilon_t')$, *and the remaining variables are defined as in Lemma 3.1.* □

It should be emphasized that the results above does not rely on the symmetry assumption (Assumption 2.6). From Theorem 4.1(b), one can see that in case of conditional heteroskedasticity, $E(\varepsilon_t\varepsilon_t' \mid \Sigma_{t-1}) \neq V_*$, a constant matrix, the asymptotic distribution of the normalized estimator for α_2 is different from that in Johansen (1988,1996). In fact, the distribution here is also different from that in

Theorem 4.1(b) of Sin and Ling (2004), who assume correct specification in variance. On the other hand, one can see from Theorem 4.1(a) that the asymptotic distribution of $(\hat{B}\bar{B}')^{-1}\hat{B}P_1$ is the same as that in Johansen (1988,1996), regardless of the presence of GARCH. As in Ahn and Reinsel (1990), if the components of Y_t can be arranged so that the last d components are non-cointegrated, then we can impose the structure $B = [I_r, B_0]$. Decompose $\hat{B} = [\hat{B}_1, \hat{B}_2]$, where $\hat{B}_1$ is rxr and $\hat{B}_2$ is rxd. Provided that $\hat{B}_1$ is invertible, it is easy to show that

$$n(\hat{B}_1^{-1}\hat{B}_2 - B_0) \longrightarrow_{\mathcal{L}} (A'V_*^{-1}A)^{-1}A'V_*^{-1}(A_\perp, A)MP_{21}^{-1}, \tag{4.2}$$

$$\sqrt{n}vec[(\hat{A}\hat{B}_1 - A), (\hat{\Phi}_1^* - \Phi_1^*), \ldots, (\hat{\Phi}_{s-1}^* - \Phi_{s-1}^*)] \longrightarrow_{\mathcal{L}} N(0, \Sigma_2^{-1}\Sigma_2^*\Sigma_2^{-1}), \tag{4.3}$$

where P_{21} is a $d \times d$ matrix such that $[0_{d\times r}, I_d]P = [P_{21}, P_{22}]$. The distribution in (4.2) is exactly the same as that in Ahn and Reinsel (1990), if their Jordan canonical form applies and $A = P_2$ up to an rxr invertible matrix.

4.2 Reduced Rank Estimation that Incorporates GARCH

This sub-section uses Johansen's estimator $\hat{\alpha}$ and some estimator $\tilde{\tilde{\delta}}$ to obtain a new reduced rank estimation that incorporates GARCH. The LF based on the error-correction form (4.1) is the same as that in (3.1), but now it is a function of the *generic* parameter $\tilde{\alpha}$ and $\tilde{\delta}$. Denote $U_t^* \equiv [(Y_t \otimes A')', (U_t \otimes I_m)']'$. Similar to (3.2),

$$\nabla_\alpha \tilde{l}_t = \nabla_\alpha l_t(\tilde{\alpha}, \tilde{\delta}) = -\frac{1}{2}(\nabla_\alpha \tilde{h}_t)(\iota - w(\tilde{\varepsilon}_t\tilde{\varepsilon}_t'\tilde{V}_t^{-1})) \odot \vec{\tilde{h}}_t + \tilde{U}_{t-1}^*\tilde{V}_t^{-1}\tilde{\varepsilon}_t. \tag{4.4}$$

For the same reasons discussed in Section 3, our WLS only considers the second term in (4.4), that is:

$$\tilde{r}_t \equiv (\tilde{r}_{1t}', \tilde{r}_{2t}')' \equiv \tilde{U}_{t-1}^*\tilde{V}_t^{-1}\tilde{\varepsilon}_t. \tag{4.5}$$

Denote $\bar{D}^{**} \equiv diag(nI_{rd}, \sqrt{n}I_{rm+(s-1)m^2})$ and $\bar{Q}^{**} \equiv diag((Q_1' \otimes I_r), I_{rm+(s-1)m^2})$. For any fixed positive constant K, let $\Xi_n \equiv \{(\tilde{\alpha}, \tilde{\delta}) : \|\bar{D}^{**}\bar{Q}^{**\prime -1}(\tilde{\alpha} - \alpha)\| \leq K \ and \ \|\sqrt{n}(\tilde{\delta} - \delta)\| \leq K\}$. Similar to (3.4), on Ξ_n, the derivative of $\tilde{r}_t$ can be *simplified* as follows:

$$\bar{D}^{**-1}\bar{Q}^{**}\sum_{t=1}^{n}\nabla_{\alpha'}\tilde{r}_t\bar{Q}^{**\prime}\bar{D}^{**-1} \quad = \quad \bar{D}^{**-1}\bar{Q}^{**}\sum_{t=1}^{n}\tilde{R}_t\bar{Q}^{**\prime}\bar{D}^{**-1} + o_p(1), \qquad (4.6)$$

where $\tilde{R}_t = diag\{\tilde{R}_{1t}, \tilde{R}_{2t}\}$, $\tilde{R}_{1t} = -(Y_{t-1}Y'_{t-1} \otimes \tilde{A}'\tilde{V}_t^{-1}\tilde{A})$, $\tilde{R}_{2t} = -(\tilde{U}_{t-1}\tilde{U}'_{t-1} \otimes \tilde{V}_t^{-1})$.

Similar to (3.5)-(3.6), the following results hold uniformly in Ξ_n:

$$\bar{D}^{**-1}\bar{Q}^{**}\sum_{t=1}^{n}(\tilde{R}_t - R_t)\bar{Q}^{**\prime}\bar{D}^{**-1} = o_p(1), \qquad (4.7)$$

$$\bar{D}^{**-1}\bar{Q}^{**}\sum_{t=1}^{n}(\tilde{r}_t - r_t) = \bar{D}^{**-1}\bar{Q}^{**}\sum_{t=1}^{n}R_t(\tilde{\alpha} - \alpha) + o_p(1), \qquad (4.8)$$

where R_t and r_t are $\tilde{R}_t$ and $\tilde{r}_t$ evaluated at the *true* parameters α and δ. Consequently, with the initial estimators $\hat{\alpha}$ and $\tilde{\delta}$, we perform a one-step iteration:

$$\dot{\alpha}_1 \quad = \quad \hat{\alpha}_1 - (\sum_{t=1}^{n}R_{1t}|_{\hat{\alpha},\tilde{\delta}})^{-1}(\sum_{t=1}^{n}r_{1t}|_{\hat{\alpha},\tilde{\delta}}), \qquad (4.9)$$

$$\dot{\alpha}_2 \quad = \quad \hat{\alpha}_2 - (\sum_{t=1}^{n}R_{2t}|_{\hat{\alpha},\tilde{\delta}})^{-1}(\sum_{t=1}^{n}r_{2t}|_{\hat{\alpha},\tilde{\delta}}). \qquad (4.10)$$

The asymptotic distributions of the *normalized* estimators for α are given as follows.

Theorem 4.2. *Suppose the assumptions in Lemma 3.1 hold. Then*

(a) $$n((\dot{B}\bar{B}')^{-1}\dot{B} - B)P_1 \longrightarrow_{\mathcal{L}} (A'\Omega_1 A)^{-1}A'M^*,$$

(b) $$\sqrt{n}vec[(\dot{A}(\dot{B}\bar{B}') - A), (\dot{\Phi}_1^* - \Phi_1^*), \ldots, (\dot{\Phi}_{s-1}^* - \Phi_{s-1}^*)] \longrightarrow_{\mathcal{L}} N(0, \Omega_2^{-1}\Omega_2^*\Omega_2^{-1}),$$

where M^* is defined as in Theorem 3.1, and the remaining variables are defined as in Lemma 3.1. □

As one can see in a section below, in fact the result in Theorem 4.2(b) does not rely on the symmetry assumption (Assumption 2.6). Decompose $\dot{B} = [\dot{B}_1, \dot{B}_2]$, where $\dot{B}_1$ is rxr and $\dot{B}_2$ is rxd. If the components of Y_t can be arranged as in Ahn and Reinsel (1990) such that the last d components are non-cointegrated, and $\dot{B}_1$ is invertible, it is easy to show that

$$n(\dot{B}_1^{-1}\dot{B}_2 - B_0) \longrightarrow_{\mathcal{L}} (A'\Omega_1 A)^{-1}A'M^*P_{21}^{-1}, \qquad (4.11)$$

$$\sqrt{n}vec[(\dot{A}\dot{B}_1 - A), (\dot{\Phi}_1^* - \Phi_1^*), \ldots, (\dot{\Phi}_{s-1}^* - \Phi_{s-1}^*)] \longrightarrow_{\mathcal{L}} N(0, \Omega_2^{-1}\Omega_2^*\Omega_2^{-1}), \qquad (4.12)$$

where P_{21} is defined around (4.2). The distribution in (4.11) is essentially the same as that in LLW (2001), with slightly different definitions of Ω_1 and $W_m^*(u)$ because of the different ARCH-type errors and we do not assume correct specification in variance.

5 Testing for Reduced Rank

This section applies the asymptotic distributions in Theorems 3.1 and 4.2 to construct tests for reduced rank. The null and the alternative hypotheses are:

$$H_0 : rank(C) = r < m \text{ vs } H_a : rank(C) = m. \tag{5.1}$$

We first consider the Wald-type test statistic:

$$W_G \equiv vec(\dot{C} - \dot{A}\dot{B})'(-\sum_{t=1}^{n} \tilde{F}_t) vec(\dot{C} - \dot{A}\dot{B}), \tag{5.2}$$

Recall that $\dot{C}$ is the full rank estimator defined in Section 3, $\dot{A}$ and $\dot{B}$ are the reduced rank estimators defined in Sub-section 4.2, while $\tilde{F}_t = -(X_{t-1}X'_{t-1} \otimes \tilde{V}_t^{-1})$, where $\tilde{V}_t$ is evaluated at some estimator on Θ_n or Ξ_n. See Sections 3 and 4. The following lemma gives the asymptotic distribution of W_G.

Lemma 5.1. *Suppose the assumptions in Lemma 3.1 hold. Then under the null H_0, the Wald-type test for rank,*

$$W_G \longrightarrow_{\mathcal{L}} tr[(\int_0^1 B_d(u) dV_d^*(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}(\int_0^1 B_d(u)dV_d^*(u)')],$$

where $V_d^*(u) = \Upsilon B_d(u) + [(A'_\perp \Omega_1^{-1} A_\perp)^{-1/2} A'_\perp \Omega_1^{-1} \Omega_1^* \Omega_1^{-1} A_\perp (A'_\perp \Omega_1^{-1} A_\perp)^{-1/2} - \Upsilon\Upsilon']^{1/2} V_d(u)$, $\Upsilon = (A'_\perp \Omega_1^{-1} A_\perp)^{1/2} (A'_\perp V_* A_\perp)^{-1/2}$, *and* $(B'_d(u), V'_d(u))'$ *is a* $2d-$*dimensional standard Brownian motion.* □

When $\Omega_1^* = \Omega_1$, the distribution of W_G can be simplified as follows.

Theorem 5.1. *If the assumptions in Lemma 5.1 hold and* $\Omega_1^* = \Omega_1$, *then*

$$W_G \longrightarrow_{\mathcal{L}} tr\{[\zeta(I_d - \Lambda_d)^{1/2} + \Phi\Lambda_d^{1/2}]'[\zeta(I_d - \Lambda_d)^{1/2} + \Phi\Lambda_d^{1/2}]\}, \tag{5.3}$$

where Λ_d *is a diagonal matrix containing the* d *eigenvalues of* $(I_d - \Upsilon\Upsilon')$*,* $\Phi \sim N(0, I_d)$ *and independent of* $\zeta = [\int_0^1 B_d(u)B_d(u)'du]^{-1/2}\int_0^1 B_d(u)dB_d(u)'$*.* □

Some of the critical values are tabulated in Appendix A. When the ε_t's are conditional homoskedastic, $\Omega_1^* = \Omega_1 = V_*^{-1}$ and hence $\Lambda_d = 0_{d\times d}$. The distribution of W_G is exactly the same as that in Johansen (1988,1996) and Reinsel and Ahn (1992). On the other hand, when $\Omega_1^* \neq \Omega_1$, we may define a modified Wald-type test statistic:

$$W_G^* \equiv vec(\dot{C}^* - \dot{A}\dot{B}^*)'(-\sum_{t=1}^{n}\tilde{F}_t^*)vec(\dot{C}^* - \dot{A}\dot{B}^*), \tag{5.4}$$

where $vec(\dot{C}^*) = (\sum_{t=1}^n \tilde{F}_t^*)^{-1}(\sum_{t=1}^n \tilde{F}_t)vec(\dot{C})$, $\dot{B}^* = (\dot{A}'\dot{\Omega}_1^*\dot{A})^{-1}(\dot{A}'\dot{\Omega}_1\dot{A})\dot{B}$. $\tilde{F}_t^* = -(X_{t-1}X_{t-1}' \otimes \tilde{V}_t^{-1}\dot{\varepsilon}_t\dot{\varepsilon}_t'\tilde{V}_t^{-1})$. The following corollary gives the asymptotic distribution of W_G^*.

Corollary 5.1. *Suppose the assumptions in Lemma 5.1 hold.*

$$W_G^* \longrightarrow_{\mathcal{L}} tr\{[\zeta(I_d - \Lambda_d^*)^{1/2} + \Phi\Lambda_d^{*1/2}]'[\zeta(I_d - \Lambda_d^*)^{1/2} + \Phi\Lambda_d^{*1/2}]\}, \tag{5.5}$$

where Λ_d^* *is a diagonal matrix containing the* d *eigenvalues of* $(I_d - (A_\perp'\Omega_1^{*-1}A_\perp)^{1/2} \cdot (A_\perp' V_* \ A_\perp)^{-1}(A_\perp'\Omega_1^{*-1}A_\perp)^{1/2})$*.* □

The critical values of the distribution in (5.3) can be simulated via Monte Carlo method. Using 100,000 replications and sample size, $n = 2,000$ of i.i.d. normal processes, we simulate the critical values when $d = 1$ and $d = 2$ and $(\lambda_1,\ \lambda_2)$ range from 0.0 to 0.9. $(\lambda_1,\ \lambda_2)$ are the diagonal elements of Λ_2 (see Theorem 5.1) or those of Λ_2^* (see Corollary 5.1). The critical values are given in Appendix A. For intermediate values of (λ_1, λ_2), the critical values could be obtained by interpolation.

Refer to Theorem 5.1 and Corollary 5.1. In actual empirical applications, one needs to estimate the d eigenvalues of $I_d-(A_\perp'\Omega_1^{-1}A_\perp)^{1/2}(A_\perp'V_*A_\perp)^{-1}(A_\perp'\Omega_1^{-1}A_\perp)^{1/2}$, or those of $I_d - (A_\perp'\Omega_1^{*-1}A_\perp)^{1/2}(A_\perp'V_*A_\perp)^{-1}(A_\perp'\Omega_1^{*-1}A_\perp)^{1/2}$. By the definition of V_* (see around (3.10) above), it can be consistently estimated by $n^{-1}\sum_{t=1}^n \dot{\varepsilon}_t\dot{\varepsilon}_t{}'$, where $\dot{\varepsilon}_t$ is the residual in Sub-section 4.2. Similarly, by the definition of $A_\perp$ (see around (2.2) above), it can be consistently estimated by $(I_m - c(\dot{A}'c)^{-1}\dot{A}')c_\perp$, where $c = (I_r, 0_{rxd})'$ and $c_\perp = (0_{dxr}, I_d)'$. See p.48 of Johansen (1996) for details. Lastly, refer to the definitions of Ω_1 and Ω_1^* (see Lemma 3.1 and around (3.10) respectively), they can respectively be consistently estimated by $\frac{1}{n}\sum_{t=1}^n \tilde{V}_t^{-1}$ and $\frac{1}{n}\sum_{t=1}^n \tilde{V}_t^{-1}\dot{\varepsilon}_t\dot{\varepsilon}_t{}'\tilde{V}_t^{-1}$.

6 Conclusions

Macroeconomic or financial data are often modelled with cointegration and GARCH. Noticeable examples include those studies of price discovery, in which stock prices of the same underlying asset are cointegrated and they exhibit multivariate GARCH. Modifying the asymptotic theories developed in Li, Ling and Wong (2001) and Sin and Ling (2004), this paper proposes a WLS(weighted least squares) for the parameters of an ECM(error-correction model). Apart from its computational simplicity, by construction, the consistency of WLS is insensitive to possible misspecification in conditional variance. Further, asymmetrically distributed deflated error is allowed, at the expense of more involved asymptotic distributions of the statistics. Efficiency loss relative to QMLE(quasi-maximum likelihood estimator) is discussed within the class of LABF(locally asymptotically Brownian functional) models. The insensitivity and efficiency of WLS in finite samples are examined through Monte Carlo experiments. We also apply the WLS to an empirical example of HSI(Hang Seng Index), HSIF(Hang Seng Index Futures) and TraHK(Hong Kong Tracker Fund).

A Appendix: Critical Values

TABLE A.1

Quantiles of the Limiting Distribution (5.3) or (5.5)

$d = 1$, no Constant Term

	$\alpha-$th simulated quantiles							
λ_1	.500	.750	.800	.850	.900	.950	.975	.990
0.0	0.602	1.550	1.891	2.343	2.995	4.153	5.357	7.018
0.1	0.575	1.539	1.869	2.315	2.978	4.140	5.365	6.941
0.2	0.553	1.511	1.850	2.308	2.964	4.138	5.362	6.939
0.3	0.533	1.489	1.824	2.282	2.941	4.108	5.305	6.921
0.4	0.515	1.462	1.800	2.254	2.914	4.083	5.286	6.929
0.5	0.499	1.441	1.770	2.223	2.883	4.043	5.242	6.895
0.6	0.490	1.414	1.743	2.197	2.845	4.013	5.225	6.824
0.7	0.481	1.385	1.718	2.171	2.811	3.963	5.174	6.839
0.8	0.470	1.364	1.693	2.139	2.782	3.920	5.097	6.774
0.9	0.461	1.354	1.674	2.105	2.746	3.867	5.047	6.718
1.0	0.455	1.326	1.649	2.078	2.711	3.827	5.068	6.633

The table values were computed from $100,000$ simulations with $n = 2,000$.
λ_1 is the eigenvalue of Λ_1 in (5.3) or Λ_1^* in (5.5).

TABLE A.2

Quantiles of the Limiting Distribution (5.3) or (5.5)

$d = 2$, no Constant Term

		$\alpha-$th simulated quantiles							
λ_1	λ_2	.500	.750	.800	.850	.900	.950	.975	.990
0.0	0.0	5.508	7.844	8.522	9.365	10.479	12.286	14.065	16.278
0.0	0.1	5.405	7.739	8.413	9.267	10.386	12.237	13.971	16.144
0.0	0.2	5.298	7.645	8.313	9.159	10.312	12.158	13.886	16.041
0.0	0.3	5.189	7.541	8.210	9.062	10.234	12.073	13.793	15.986
0.0	0.4	5.068	7.440	8.112	8.959	10.119	11.987	13.722	15.895
0.0	0.5	4.952	7.330	8.008	8.865	10.003	11.887	13.659	15.802
0.0	0.6	4.839	7.216	7.909	8.744	9.906	11.789	13.542	15.716
0.0	0.7	4.726	7.112	7.783	8.647	9.796	11.676	13.440	15.623
0.0	0.8	4.619	6.981	7.668	8.525	9.680	11.559	13.354	15.530
0.0	0.9	4.504	6.867	7.542	8.410	9.551	11.446	13.230	15.435
0.0	1.0	4.393	6.745	7.417	8.268	9.443	11.306	13.172	15.450
0.1	0.1	5.287	7.635	8.325	9.172	10.295	12.140	13.885	16.105
0.1	0.2	5.178	7.534	8.229	9.079	10.217	12.071	13.817	15.991
0.1	0.3	5.058	7.440	8.123	8.979	10.125	11.987	13.736	15.920
0.1	0.4	4.945	7.341	8.023	8.865	10.018	11.902	13.612	15.806
0.1	0.5	4.832	7.224	7.920	8.750	9.919	11.818	13.539	15.643
0.1	0.6	4.718	7.108	7.791	8.643	9.808	11.692	13.422	15.552
0.1	0.7	4.605	6.987	7.677	8.533	9.679	11.578	13.296	15.482
0.1	0.8	4.498	6.856	7.559	8.413	9.561	11.434	13.179	15.337
0.1	0.9	4.382	6.749	7.430	8.290	9.455	11.284	13.064	15.247
0.1	1.0	4.278	6.627	7.307	8.157	9.307	11.147	12.950	15.229
0.2	0.2	5.070	7.445	8.137	8.987	10.116	11.973	13.707	15.898
0.2	0.3	4.945	7.336	8.037	8.881	10.028	11.879	13.601	15.812
0.2	0.4	4.828	7.225	7.916	8.761	9.916	11.791	13.501	15.647
0.2	0.5	4.711	7.111	7.807	8.658	9.819	11.691	13.383	15.556
0.2	0.6	4.596	6.998	7.682	8.532	9.691	11.566	13.298	15.405
0.2	0.7	4.488	6.881	7.560	8.415	9.579	11.433	13.191	15.319
0.2	0.8	4.383	6.753	7.435	8.288	9.453	11.293	13.027	15.191
0.2	0.9	4.266	6.621	7.309	8.165	9.322	11.141	12.902	15.023
0.2	1.0	4.160	6.502	7.190	8.031	9.182	10.985	12.768	15.020
0.3	0.3	4.830	7.232	7.929	8.781	9.931	11.752	13.491	15.702
0.3	0.4	4.717	7.118	7.809	8.657	9.816	11.669	13.411	15.609
0.3	0.5	4.598	7.001	7.688	8.540	9.693	11.570	13.285	15.471

TABLE A.2 (Continued)

		α−th simulated quantiles							
λ_1	λ_2	.500	.750	.800	.850	.900	.950	.975	.990
0.3	0.6	4.489	6.877	7.570	8.415	9.565	11.432	13.179	15.318
0.3	0.7	4.369	6.758	7.442	8.281	9.442	11.296	13.051	15.202
0.3	0.8	4.263	6.636	7.302	8.160	9.310	11.158	12.897	15.021
0.3	0.9	4.152	6.505	7.187	8.042	9.163	11.010	12.743	14.870
0.3	1.0	4.052	6.374	7.045	7.882	9.046	10.819	12.592	14.853
0.4	0.4	4.600	7.006	7.695	8.549	9.707	11.557	13.290	15.510
0.4	0.5	4.486	6.877	7.577	8.420	9.576	11.438	13.180	15.374
0.4	0.6	4.373	6.760	7.444	8.287	9.440	11.310	13.061	15.231
0.4	0.7	4.255	6.631	7.318	8.148	9.313	11.171	12.881	15.087
0.4	0.8	4.150	6.506	7.179	8.012	9.176	11.024	12.733	14.928
0.4	0.9	4.040	6.378	7.050	7.883	9.018	10.847	12.567	14.747
0.4	1.0	3.941	6.233	6.911	7.735	8.875	10.678	12.395	14.651
0.5	0.5	4.376	6.751	7.437	8.298	9.444	11.322	13.053	15.298
0.5	0.6	4.261	6.625	7.299	8.171	9.310	11.176	12.919	15.115
0.5	0.7	4.151	6.497	7.178	8.016	9.177	11.049	12.759	14.954
0.5	0.8	4.036	6.362	7.039	7.870	9.030	10.854	12.567	14.820
0.5	0.9	3.937	6.235	6.907	7.727	8.866	10.693	12.398	14.612
0.5	1.0	3.836	6.098	6.758	7.588	8.685	10.541	12.202	14.486
0.6	0.6	4.152	6.495	7.161	8.015	9.153	11.035	12.781	14.993
0.6	0.7	4.045	6.356	7.027	7.874	9.015	10.894	12.580	14.809
0.6	0.8	3.930	6.214	6.890	7.719	8.857	10.713	12.401	14.622
0.6	0.9	3.828	6.086	6.749	7.577	8.698	10.529	12.218	14.480
0.6	1.0	3.733	5.959	6.612	7.428	8.512	10.358	12.002	14.298
0.7	0.7	3.936	6.213	6.885	7.721	8.847	10.719	12.432	14.668
0.7	0.8	3.827	6.082	6.738	7.564	8.688	10.555	12.247	14.435
0.7	0.9	3.724	5.933	6.598	7.413	8.520	10.353	12.036	14.259
0.7	1.0	3.630	5.811	6.464	7.251	8.347	10.151	11.794	14.091
0.8	0.8	3.728	5.934	6.586	7.400	8.526	10.342	12.053	14.255
0.8	0.9	3.626	5.791	6.434	7.240	8.345	10.144	11.857	14.064
0.8	1.0	3.528	5.666	6.303	7.084	8.154	9.952	11.588	13.825
0.9	0.9	3.531	5.655	6.286	7.071	8.166	9.932	11.656	13.770
0.9	1.0	3.446	5.521	6.142	6.913	7.972	9.703	11.390	13.553
1.0	1.0	3.359	5.378	5.977	6.734	7.777	9.471	11.120	13.264

The table values were computed from 100,000 simulations with $n = 2,000$.

$\lambda_1 \leq \lambda_2$ are the eigenvalues of Λ_2 in (5.3) or Λ_2^* in (5.5).

B Appendix: Technical Proofs

Lemma B.1. *Under the assumptions in Theorem 4.2, it follows that*

(*a*) $(\hat{B}\bar{B}')^{-1}(\dot{B}-\hat{B}) = O_p(n^{-1/2})$,

(*b*) $\hat{A}(\dot{B}\bar{B}') = \hat{A}(\hat{B}\bar{B}') + O_p(n^{-1/2}) = A + O_p(n^{-1/2})$,

(*c*) $(\dot{B}\bar{B}')^{-1}\hat{B}P_1 = (\hat{B}\bar{B}')^{-1}\hat{B}P_1 + O_p(n^{-3/2}) = BP_1 + O_p(n^{-1})$,

(*d*) $(\dot{B}\bar{B}')^{-1}\hat{B}P_2 = (\hat{B}\bar{B}')^{-1}\hat{B}P_2 + O_p(n^{-1/2}) = BP_2 + O_p(n^{-1/2})$. □

Proof. (a). We first denote $D_{\alpha_1} = diag(nI_{rd}, \sqrt{n}I_{r^2})$ and $\hat{Q}^{**} = \mathcal{Q}(I_m \otimes (\hat{B}\bar{B}')')$, with $\mathcal{Q} = (Q \otimes I_r)$. Also denote $\hat{\alpha}_1 = vec(\hat{B})$, $\breve{\alpha}_1 = vec((\hat{B}\bar{B}')^{-1}\hat{B})$ and $\dot{\alpha}_1 = vec(\dot{B})$. $\hat{\alpha}_2$, $\breve{\alpha}_2$ and $\dot{\alpha}_2$ are defined accordingly. $\hat{\alpha}$, $\breve{\alpha}$ and $\dot{\alpha}$ are also defined accordingly. Since $\hat{Q}^{**\prime -1} = (P' \otimes I_r)(I_m \otimes (\hat{B}\bar{B}')^{-1})$, we have

$$\begin{aligned}(I_m \otimes (\hat{B}\bar{B}')^{-1})(\dot{\alpha}_1 - \hat{\alpha}_1) &= \mathcal{Q}'D_{\alpha_1}^{-1}D_{\alpha_1}(P' \otimes I_r)(I_m \otimes (\hat{B}\bar{B}')^{-1})(\dot{\alpha}_1 - \hat{\alpha}_1) \\ &= \mathcal{Q}'D_{\alpha_1}^{-1}[D_{\alpha_1}\hat{Q}^{**\prime -1}(\dot{\alpha}_1 - \hat{\alpha}_1)].\end{aligned}$$

As $\mathcal{Q}'D_{\alpha_1}^{-1} = O(n^{-1/2})$, it suffices to show $D_{\alpha_1}\hat{Q}^{**\prime -1}(\dot{\alpha}_1 - \hat{\alpha}_1) = O_p(1)$. By (4.9),

$$\begin{aligned}&D_{\alpha_1}\hat{Q}^{**\prime -1}(\dot{\alpha}_1 - \hat{\alpha}_1) \\ &\quad = -[\sum_{t=1}^n D_{\alpha_1}^{-1}\hat{Q}^{**}(R_{1t}|_{\hat{\alpha},\dot{\delta}})\hat{Q}^{**\prime}D_{\alpha_1}^{-1}]^{-1}[\sum_{t=1}^n D_{\alpha_1}^{-1}\hat{Q}^{**}(r_{1t}|_{\hat{\alpha},\dot{\delta}})] \\ &\quad = -[\sum_{t=1}^n D_{\alpha_1}^{-1}\mathcal{Q}(R_{1t}|_{\breve{\alpha},\dot{\delta}})\mathcal{Q}'D_{\alpha_1}^{-1}]^{-1}[\sum_{t=1}^n D_{\alpha_1}^{-1}\mathcal{Q}(r_{1t}|_{\breve{\alpha},\dot{\delta}})].\end{aligned}$$

By Theorem 4.1 and Theorem 3.1(c), $n(\breve{\alpha}_1 - \alpha_1) = O_p(1)$, $\sqrt{n}(\breve{\alpha}_2 - \alpha_2) = O_p(1)$, and $\sqrt{n}(\dot{\delta} - \delta) = O_p(1)$. Similar to the arguments for (4.7), it follows that:

$$\sum_{t=1}^n D_{\alpha_1}^{-1}\mathcal{Q}(R_{1t}|_{\breve{\alpha},\dot{\delta}})\mathcal{Q}'D_{\alpha_1}^{-1} = \sum_{t=1}^n D_{\alpha_1}^{-1}\mathcal{Q}R_{1t}\mathcal{Q}'D_{\alpha_1}^{-1} + o_p(1). \qquad \text{(B. 1)}$$

On the other hand, by a Taylor's expansion and (B.1), with R_{1t}^* and r_{1t}^* being evaluated at a mid-point of $(\breve{\alpha}, \dot{\delta})$ and (α, δ),

$$\sum_{t=1}^n D_{\alpha_1}^{-1}\mathcal{Q}(r_{1t}|_{\breve{\alpha},\dot{\delta}})$$

$$
\begin{aligned}
&= \sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q} r_{1t} + \sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}(R_{1t}^*)(\check{\alpha}_1 - \alpha_1) + \sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}(\nabla_{\alpha_2'} r_{1t}^*)(\check{\alpha}_2 - \alpha_2) \\
&= \sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}_1 r_{1t} + [\sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q} R_{1t}\mathcal{Q}' D_{\alpha_1}^{-1} + o_p(1)]\frac{1}{n} D_{\alpha_1}(P' \otimes I_r)[n(\check{\alpha}_1 - \alpha_1)] \\
&\quad + [\frac{1}{\sqrt{n}} \sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}(\nabla_{\alpha_2'} r_{1t}^*)]\sqrt{n}(\check{\alpha}_2 - \alpha_2). \qquad \text{(B. 2)}
\end{aligned}
$$

It is not difficult to show that $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} D_{\alpha_1}^{-1}\mathcal{Q}(\nabla_{\alpha_2'} r_{1t}^*)$ is $O_p(1)$. So is the RHS of (B.2). By Lemmas 3.1(a)-(b), (B.1) and (B.2), (a) holds.

(b). By the $\sqrt{n}$-consistency of $\hat{A}(\hat{B}\bar{B}')$ for A, and (a) of this lemma,

$$
\hat{A}(\dot{B}\bar{B}') = \hat{A}(\hat{B}\bar{B}') + \hat{A}(\hat{B}\bar{B}')(\hat{B}\bar{B}')^{-1}(\dot{B} - \hat{B})\bar{B}' = \hat{A}(\hat{B}\bar{B}') + O_p(1)O_p(n^{-1/2}).
$$

Thus, (b) holds.

(c) and (d). Denote $\check{B} = (\hat{B}\bar{B}')^{-1}\hat{B}$.

$$
(\dot{B}\bar{B}')^{-1}\hat{B} = [(\hat{B}\bar{B}')^{-1}\dot{B}\bar{B}']^{-1}(\hat{B}\bar{B}')^{-1}\hat{B} = [(\hat{B}\bar{B}')^{-1}\dot{B}\bar{B}']^{-1}\check{B}. \qquad \text{(B. 3)}
$$

Using the formula $dF^{-1} = -F^{-1}(dF)F^{-1}$ for the $r \times r$ matrix F with $F(x) = [x\bar{B}]^{-1}$, and applying a Taylor's expansion to $[(\hat{B}\bar{B}')^{-1}\dot{B}\bar{B}']^{-1}$ around $\check{B}\bar{B}'$, we have

$$
[(\hat{B}\bar{B}')^{-1}\dot{B}\bar{B}']^{-1} = [\check{B}\bar{B}']^{-1} - [B^*\bar{B}']^{-1}[(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B}]\bar{B}'[B^*\bar{B}']^{-1},
$$

where B^* lies between $(\hat{B}\bar{B}')^{-1}\dot{B}$ and $\check{B}$. Therefore, the RHS of (B.3) equals:

$$
\begin{aligned}
&[(\hat{B}\bar{B}')^{-1}\hat{B}\bar{B}']^{-1}(\hat{B}\bar{B}')^{-1}\hat{B} - [B^*\bar{B}']^{-1}[(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B}]\bar{B}'[B^*\bar{B}']^{-1}\check{B} \\
&\quad = (\hat{B}\bar{B}')^{-1}\hat{B} - [B^*\bar{B}']^{-1}[(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B}]\bar{B}'[B^*\bar{B}']^{-1}\check{B}. \qquad \text{(B. 4)}
\end{aligned}
$$

By (a) of this lemma, $(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B} = O_p(n^{-1/2})$. From this, we can show that $[B^*\bar{B}']^{-1} = O_p(1)$. $\bar{B}$ and $\check{B}$ are also $O_P(1)$. By (B.4), (d) holds. By Theorem 4.1, $\check{B}P_1 = O_p(n^{-1})$ because $BP_1 = 0$. By (B.4),

$$
\begin{aligned}
&[(\hat{B}\bar{B}')^{-1}\hat{B}\bar{B}']^{-1}(\hat{B}\bar{B}')^{-1}\hat{B}P_1 - [B^*\bar{B}']^{-1}[(\hat{B}\bar{B}')^{-1}\dot{B} - \check{B}]\bar{B}'[B^*\bar{B}']^{-1}\check{B}P_1 \\
&\quad = (\hat{B}\bar{B}')^{-1}\hat{B}P_1 + O_p(n^{-3/2}).
\end{aligned}
$$

Thus, (c) holds. This completes the proof. □

Proof of Theorem 4.2. Denote $\dot{Q}_1^{**} = (Q_1' \otimes I_r)(I_m \otimes (\dot{B}\bar{B}')')$, $\dot{Q}_2^{**} = diag((\dot{B}\bar{B}')^{-1} \otimes I_m, I_{(s-1)m^2})$, $\grave{\alpha}_1 = vec((\dot{B}\bar{B}')^{-1}\hat{B})$, $\grave{\alpha}_2 = vec[\hat{A}(\dot{B}\bar{B}'), \hat{\Phi}_1^*, \ldots, \hat{\Phi}_{s-1}^*]$, and $\grave{\alpha} = [\grave{\alpha}_1', \grave{\alpha}_2']'$. By Lemmas B.1(b)-(c), $(\grave{\alpha}, \dot{\delta}) \in \Xi_n$. Thus by (4.7),

$$
\begin{aligned}
n^{-2}\sum_{t=1}^{n} \dot{Q}_1^{**}(R_{1t}|_{\hat{\alpha},\dot{\delta}})\dot{Q}_1^{**\prime} &= n^{-2}\sum_{t=1}^{n}(Q_1' \otimes I_r)(R_{1t}|_{\grave{\alpha},\dot{\delta}})(Q_1 \otimes I_r) \\
&= n^{-2}\sum_{t=1}^{n}(Q_1' \otimes I_r)R_{1t}(Q_1 \otimes I_r) + o_p(1), \qquad \text{(B. 5)} \\
n^{-1}\sum_{t=1}^{n} \dot{Q}_2^{**}(R_{2t}|_{\hat{\alpha},\dot{\delta}})\dot{Q}_2^{**\prime} &= n^{-1}\sum_{t=1}^{n}(R_{2t}|_{\grave{\alpha},\dot{\delta}}) = n^{-1}\sum_{t=1}^{n} R_{2t} + o_p(1). \qquad \text{(B. 6)}
\end{aligned}
$$

Refer to (4.6). Due to the block-diagonality of $\tilde{R}_t$, by (4.8),

$$
\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n} \dot{Q}_1^{**}(r_{1t}|_{\hat{\alpha},\dot{\delta}}) = \frac{1}{n}\sum_{t=1}^{n}(Q_1' \otimes I_r)(r_{1t}|_{\grave{\alpha},\dot{\delta}}) \\
= \ &\frac{1}{n}\sum_{t=1}^{n}(Q_1' \otimes I_r)r_{1t} + (\frac{1}{n}\sum_{t=1}^{n}(Q_1' \otimes I_r)R_{1t}(Q_1 \otimes I_r))(P_1' \otimes I_r)(\grave{\alpha}_1 - \alpha_1) + o_p(1), \quad \text{(B. 7)} \\
&\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \dot{Q}_2^{**}(r_{2t}|_{\hat{\alpha},\dot{\delta}}) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(r_{2t}|_{\grave{\alpha},\dot{\delta}}) \\
= \ &\frac{1}{\sqrt{n}}\sum_{t=1}^{n} r_{2t} + (\frac{1}{\sqrt{n}}\sum_{t=1}^{n} R_{2t})(\grave{\alpha}_2 - \alpha_2) + o_p(1). \qquad \text{(B. 8)}
\end{aligned}
$$

(a). Recall that $\dot{Q}_1^{**\prime -1}\hat{\alpha}_1 = (P_1' \otimes I_r)\grave{\alpha}_1$. By (4.9), (B.5) and (B.7),

$$
\begin{aligned}
n\dot{Q}_1^{**\prime -1}\dot{\alpha}_1 &= n\dot{Q}_1^{**\prime -1}\hat{\alpha}_1 - [n^{-2}\sum_{t=1}^{n} \dot{Q}_1^{**}(R_{1t}|_{\hat{\alpha},\dot{\delta}})\dot{Q}_1^{**\prime}]^{-1}[n^{-1}\sum_{t=1}^{n} \dot{Q}_1^{**}(r_{1t}|_{\hat{\alpha},\dot{\delta}})] \\
&= n(P_1' \otimes I_r)\grave{\alpha}_1 - [n^{-2}\sum_{t=1}^{n}(Q_1' \otimes I_r)R_{1t}(Q_1 \otimes I_r)]^{-1}[n^{-1}\sum_{t=1}^{n}(Q_1' \otimes I_r)r_{1t}] \\
&\quad -n(P_1' \otimes I_r)(\grave{\alpha}_1 - \alpha_1) + o_p(1) \\
&= n(P_1' \otimes I_r)\alpha_1 - [\frac{1}{n^2}\sum_{t=1}^{n}(Q_1' \otimes I_r)R_{1t}(Q_1 \otimes I_r)]^{-1}[\frac{1}{n}\sum_{t=1}^{n}(Q_1' \otimes I_r)r_{1t}] \\
&\quad +o_p(1). \qquad \text{(B. 9)}
\end{aligned}
$$

Note that $\dot{Q}_1^{**\prime -1}\dot{\alpha}_1 - (P_1' \otimes I_r)\alpha_1 = vec[((\dot{B}\bar{B}')^{-1}\dot{B} - B)P_1]$. By (B.9) and Lemma 3.1(a)-(b), (a) holds.

(b). By (4.10), (B.6) and (B.8),

$$
\sqrt{n}\dot{Q}_2^{**\prime -1}\dot{\alpha}_2 = \sqrt{n}\dot{Q}_2^{**\prime -1}\hat{\alpha}_2 - [n^{-1}\sum_{t=1}^{n} \dot{Q}_2^{**}(R_{2t}|_{\hat{\alpha},\dot{\delta}})\dot{Q}_2^{**\prime}]^{-1}[n^{-1/2}\sum_{t=1}^{n} \dot{Q}_2^{**}(r_{2t}|_{\hat{\alpha},\dot{\delta}})]
$$

$$
\begin{aligned}
&= \sqrt{n}\dot{\alpha}_2 - [n^{-1}\sum_{t=1}^{n} R_{2t}]^{-1}[n^{-1/2}\sum_{t=1}^{n} r_{2t}] - \sqrt{n}(\dot{\alpha}_2 - \alpha_2) + o_p(1) \\
&= \sqrt{n}\alpha_2 - [n^{-1}\sum_{t=1}^{n} R_{2t}]^{-1}[n^{-1/2}\sum_{t=1}^{n} r_{2t}] + o_p(1). \qquad \text{(B. 10)}
\end{aligned}
$$

By (B.10) and Lemma 3.1(a)-(b), (b) holds. This completes the proof. $\square$

Proof of Lemma 5.1. Let $\dot{\varphi}^* = vec[CP_1, \dot{C}P_2, \dot{\Phi}_1^*, \cdots, \dot{\Phi}_{s-1}^*]$, and $l^*(\dot{\varphi}^*, \dot{\delta})$ be $l(\dot{\varphi}, \dot{\delta})$ with $\dot{C}P_1 Z_{1t-1}$ replaced by $CP_1 Z_{1t-1}$. By Lemma 3.1, Theorem 3.1 and a Taylor's expansion, we can show that

$$
2[l(\dot{\varphi}, \dot{\delta}) - l^*(\dot{\varphi}^*, \dot{\delta})] = vec[n(\dot{C} - C)P_1]'[\frac{1}{n^2}\sum_{t=1}^{n} L_{1t}]vec[n(\dot{C} - C)P_1] + o_p(1), \qquad \text{(B. 11)}
$$

where $L_{1t} = (Z_{1t-1}Z_{1t-1}' \otimes V_t^{-1}) + \sum_{l=1}^{t-1}[Z_{1t-l-1}Z_{1t-l-1}' \otimes ((\Gamma^{-1} \odot \Gamma + I_m) \odot \nu_l \nu_l' \odot \Pi_{lt})]$.

Denote $\ddot{A} = \dot{A}(\dot{B}\bar{B}')$ and $\ddot{B} = (\dot{B}\bar{B}')^{-1}\dot{B}$. Note $\dot{A}\dot{B} = \ddot{A}\ddot{B}$. Moreover,

$$
\ddot{A}\ddot{B} - AB = (\ddot{A} - A)B + A(\ddot{B} - B) + (\ddot{A} - A)(\ddot{B} - B).
$$

Recall that $BP_1 = 0$. By Theorem 4.2, $(\ddot{B} - B)P_1 = O_p(n^{-1})$ and $(\ddot{A} - A) = O_p(n^{-1/2})$ under H_0. Hence,

$$
\begin{aligned}
n(\ddot{A}\ddot{B} - AB)P_1 &= n(\ddot{A} - A)BP_1 + nA(\ddot{B} - B)P_1 + (\ddot{A} - A)n(\ddot{B} - B)P_1 \\
&= nA(\ddot{B} - B)P_1 + O_p(n^{-1/2}). \qquad \text{(B. 12)}
\end{aligned}
$$

Let $\dot{\alpha}^* = vec[ABP_1, \dot{A}\dot{B}P_2, \dot{\Phi}_1^*, \cdots, \dot{\Phi}_{s-1}^*]$, and $l^*(\dot{\alpha}^*, \dot{\delta})$ be $l(\dot{\alpha}, \dot{\delta})$ with $\dot{A}\dot{B}P_1 Z_{1t-1}$ replaced by $ABP_1 Z_{1t-1} = CP_1 Z_{1t-1}$. By Lemma 3.1, Theorem 4.2, a Taylor's expansion and (A.12), we can show that:

$$
\begin{aligned}
&2[l(\dot{\alpha}, \dot{\delta}) - l^*(\dot{\alpha}^*, \dot{\delta})] \\
&= vec[n(\ddot{A}\ddot{B} - AB)P_1]'[n^{-2}\sum_{t=1}^{n} L_{1t}]vec[n(\ddot{A}\ddot{B} - AB)P_1] + o_p(1) \\
&= vec[nA(\ddot{B} - B)P_1]'[n^{-2}\sum_{t=1}^{n} L_{1t}]vec[nA(\ddot{B} - B)P_1] + o_p(1). \qquad \text{(B. 13)}
\end{aligned}
$$

It is straightforward to show that $l^*(\dot{\varphi}^*, \dot{\delta}) - l^*(\dot{\alpha}^*, \dot{\delta}) = o_p(1)$. Furthermore, by (A.11), (A.13) and Lemma 3.1, it follows that

$$
LR_G \longrightarrow_{\mathcal{L}} vec[\Omega_1^{-1}M^*]'[Z \otimes \Omega_1]vec[\Omega_1^{-1}M^*] - vec[DM^*]'[Z \otimes \Omega_1]vec[DM^*]
$$

$$= \quad vec[\Omega_1^{-1}M^*]'vec[\Omega_1\Omega_1^{-1}M^*Z] - vec[DM^*]'vec[\Omega_1 DM^*Z]$$

$$= \quad tr[M^{*\prime}\Omega_1^{-1}M^*Z] - tr[M^{*\prime}D\Omega_1 DM^*Z]$$

$$= \quad tr[(\Omega_1^{-1} - A(A'\Omega_1 A)^{-1}A')M^*ZM^{*\prime}]. \qquad \text{(B. 14)}$$

where $D \equiv A(A'\Omega_1 A)^{-1}A'$, $Z \equiv \psi_{11}\Omega_{a1}^{1/2}\int_0^1 B_d(u)B_d(u)'\Omega_{a1}^{1/2}\psi_{11}'$ and M^* is defined as in Theorem 3.1. Following the lines on p.359 of Reinsel and Ahn (1992), we can rewrite $\Omega_1^{-1} - A(A'\Omega_1 A)^{-1}A'$ as:

$$\Omega_1^{-1}(\Omega_1 - \Omega_1 A(A'\Omega_1 A)^{-1}A'\Omega_1)\Omega_1^{-1} = \Omega_1^{-1}A_\perp(A_\perp'\Omega_1^{-1}A_\perp)^{-1}A_\perp'\Omega_1^{-1}.$$

Therefore, we can rewrite the asymptotic distribution in (A.13) as:

$$tr[(\int_0^1 B_d(u)dV_d^*(u)')'(\int_0^1 B_d(u)B_d(u)'du)^{-1}(\int_0^1 B_d(u)dV_d^*(u)')],$$

where $V_d^*(u) \equiv (A_\perp'\Omega_1^{-1}A_\perp)^{-1/2}A_\perp'\Omega_1^{-1}W_m^*(u)$. Note $E[B_d(u)V_d^*(u)'] = u\Omega_{a1}^{-1/2}(A_\perp'\Omega_1^{-1}A_\perp)^{1/2} = u\Upsilon'$. Thus, we can rewrite $V_d^*(u)$ as a linear combination of two independent $d-$dimensional standard BMs:

$$\Upsilon B_d(u) + [(A_\perp'\Omega_1^{-1}A_\perp)^{-1/2}A_\perp'\Omega_1^{-1}\Omega_1^*\Omega_1^{-1}A_\perp(A_\perp'\Omega_1^{-1}A_\perp)^{-1/2} - \Upsilon\Upsilon']^{1/2}V_d(u). \text{(B. 15)}$$

The proof is complete. □

Proof of Theorem 5.1. When $\Omega_1^* = \Omega_1$, (A.15) in the proof of Lemma 5.1 can be simplified as $\Upsilon B_d(u) + [I_d - \Upsilon\Upsilon']^{1/2}V_d(u)$. Thus, the asymptotic distribution can be simplified as:

$$tr\{[\int_0^1 \Upsilon B_d(u)dB_d(u)'\Upsilon' + \int_0^1 \Upsilon B_d(u)dV_d(u)'(I_d - \Upsilon\Upsilon')^{1/2}]'$$
$$\cdot[\int_0^1 \Upsilon B_d(u)B_d(u)'\Upsilon' du]^{-1}[\int_0^1 \Upsilon B_d(u)dB_d(u)'\Upsilon' + \int_0^1 \Upsilon B_d(u)dV_d(u)'(I_d - \Upsilon\Upsilon')^{1/2}]\}.$$

However, $\Upsilon B_d(u) \sim N(0, \Upsilon\Upsilon')$. Abusing the notation, we write $\Upsilon B_d(u)$ as $(\Upsilon\Upsilon')^{1/2}$ $B_d(u)$, where $B_d(u)$ is (another) $d-$dimensional standard BM independent of $V_d(u)$.

Therefore, cancelling some of the $(\Upsilon\Upsilon')^{1/2}$ terms, the asymptotic distribution can be expressed as:

$$tr\{[\int_0^1 B_d(u)dB_d(u)'(\Upsilon\Upsilon')^{1/2} + \int_0^1 B_d(u)dV_d(u)'(I_d - \Upsilon\Upsilon')^{1/2}]'$$
$$[\int_0^1 B_d(u)B_d(u)'du]^{-1}[\int_0^1 B_d(u)dB_d(u)'(\Upsilon\Upsilon')^{1/2} + \int_0^1 B_d(u)dV_d(u)'(I_d - \Upsilon\Upsilon')^{1/2}]\}.$$

Since $(I_d - \Upsilon\Upsilon')$ is a real symmetric matrix, we can decompose it as $\Theta\Lambda_d\Theta'$, where Θ is an orthogonal matrix such that $\Theta'\Theta = I_d$. In view of $(\Upsilon\Upsilon')^{1/2} = \Theta(I_d - \Lambda_d)^{1/2}\Theta'$ and $(I_d - \Upsilon\Upsilon')^{1/2} = \Theta\Lambda_d^{1/2}\Theta'$ and due to the orthogonality of Θ, we can write the asymptotic distribution as:

$$tr\{[\int_0^1 \Theta' B_d(u) dB_d(u)'\Theta(I_d - \Lambda_d)^{1/2}\Theta' + \int_0^1 \Theta' B_d(u) dV_d(u)'\Theta\Lambda_d^{1/2}\Theta']'$$
$$\cdot[\int_0^1 \Theta' B_d(u) B_d(u)' du\Theta]^{-1}$$
$$\cdot[\int_0^1 \Theta' B_d(u) dB_d(u)'\Theta(I_d - \Lambda_d)^{1/2}\Theta' + \int_0^1 \Theta' B_d(u) dV_d(u)'\Theta\Lambda_d^{1/2}\Theta']\}.$$

Since $\Theta' B_d(u) \sim N(0, \Theta'\Theta) = N(0, I_d)$, similar to the previous arguments, and abusing the notation, we can write $\Theta' B_d(u)$ and $\Theta' V_d(u)$ as two independent standard BMs $B_d(u)$ and $V_d(u)$ respectively. Cancelling the orthogonal Θ, we have:

$$tr\{[\int_0^1 B_d(u) dB_d(u)'(I_d - \Lambda_d)^{1/2} + \int_0^1 B_d(u) dV_d(u)'\Lambda_d^{1/2}]'$$
$$\cdot[\int_0^1 B_d(u) B_d(u)' du]^{-1}[\int_0^1 B_d(u) dB_d(u)'(I_d - \Lambda_d)^{1/2} + \int_0^1 B_d(u) dV_d(u)'\Lambda_d^{1/2}]\}$$
$$= tr\{[\zeta(I_d - \Lambda_d)^{1/2} + \Phi\Lambda_d^{1/2}]'[\zeta(I_d - \Lambda_d)^{1/2} + \Phi\Lambda_d^{1/2}]\}.$$

This completes the proof. □

REFERENCES

Ahn, S. K. and Reinsel, G.C. (1990), Estimation for Partially Nonstationary Multivariate models, *Journal of American Statistical Association*, 85, 813-823.

Bollerslev, T. (1990), Modeling the Coherence in the Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach, *Review of Economics and Statistics*, 72, 498-505.

Engle, R.F. and Granger, C.W.J. (1987), Cointegration and Error Correction: Representation, Estimation and Testing, *Econometrica*, 55, 251-276.

Glosten, L.R., Jagannathan, R. and Runkle, D.E. (1993), On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48, 1779-1802.

Granger, C.W.J. (1983), Cointegrated Variables and Error Correction models, Discussion Paper, Department of Economics, University of California at San Diego.

Johansen, S. (1988), Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. (1996), *Likelihood-Based Inference in Cointegrated Vector Autoregressive models*. Oxford: Oxford University Press.

Journal of Econometrics, 73, 401-410.

Li, W.K., Ling, S. and Wong, H. (2001), Estimation for Partially Nonstationary Multivariate Autoregressive models with Conditional Heteroskedasticity, *Biometrika*, 88, 1135-1152.

Ling, S. and Li, W.K. (1997), Diagnostic Checking of Nonlinear Multivariate Time Series with Multivariate ARCH Errors, *Journal of Time Series Analysis*, 18, 447-464.

Ling, S. and Li, W.K. (1998), Limiting Distributions of Maximum Likelihood Estimators for Unstable ARMA models with GARCH Errors, *Annals of Statistics*, 26, 84-125.

Ling, S. and Li, W.K. (2003), Asymptotic Inference for Unit Root with GARCH(1,1) Errors, *Econometric Theory*, 19, 541-564.

Ling, S., Li, W.K. and McAleer, M. (2003), Estimation and Testing for Unit Root Processes with GARCH(1,1) Errors: Theory and Monte Carlo Study, *Econometric Reviews*, 22, 179-202.

Ling, S. and McAleer, M. (2003a), Asymptotic Theory for a Vector ARMA-GARCH model, *Econometric Theory*, 19, 280-310.

Ling, S. and McAleer, M. (2003b), On Adaptive Estimation in Nonstationary ARMA models with GARCH Errors, *Annals of Statistics*, 31, 642-674.

Magnus, J.R. (1988), *Linear Structures*. New York: Oxford University Press.

Magnus, J.R. and Neudecker, H. (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Chichester: John Wiley and Sons.

Phillips, P.C.B. and Durlauf, S.N. (1986), Multiple Time Series Regression with Integrated Processes, *Review of Economic Studies*, 53, 473-495.

Rahbek, A. and Mosconi, R. (1999), Cointegration Rank Inference with Stationary Regressions in VAR models, *Econometrics Journal*, 2, 76-91.

Reinsel, G.C. and Ahn, S.K. (1992), Vector AR models with Unit Root and Reduced Rank Structure: Estimation, Likelihood Ratio Test, and Forecasting, *Journal of Time Series Analysis*, 13, 133-145.

Sin, C.-y. and Ling, S. (2004), Estimation and Testing for Partially Nonstationary Vector Autoregressive models with GARCH, Discussion Paper, Hong Kong Baptist University, and Hong Kong University of Science and Technology.

Stock, J.H. and Watson, M.W. (1993), A Simple Estimator of Cointegrating Vectors in Higher Order Integrated System, *Econometrica*, 61, 783-820.

Tsay, R.S. (1987), Conditional Heteroscedastic Time Series models, *Journal of the American Statistical Association*, 82, 590-604.

Tse, Y.K. and Tsui, A.K.C. (2002), A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations, *Journal of Business and Economic Statistics*, 20, 351-362.