

# **Asymptotic Theory for Extended Asymmetric Multivariate GARCH Processes\***

EI2016-35

**Manabu Asai**

Faculty of Economics  
Soka University, Japan

**Michael McAleer**

Department of Quantitative Finance  
National Tsing Hua University, Taiwan  
and  
Econometric Institute  
Erasmus School of Economics  
Erasmus University Rotterdam  
and  
Department of Quantitative Economics  
Complutense University of Madrid, Spain  
and  
Institute of Advanced Sciences  
Yokohama National University, Japan

September 2016

---

\*The authors are most grateful to Yoshi Baba and Chia-Lin Chang for very helpful comments and suggestions. The first author acknowledges the financial support of the Japan Ministry of Education, Culture, Sports, Science and Technology, Japan Society for the Promotion of Science, and Australian Academy of Science. The second author is most grateful for the financial support of the Australian Research Council, National Science Council, Ministry of Science and Technology (MOST), Taiwan, Japan Society for the Promotion of Science, and Institute of Advanced Sciences, Yokohama National University. Address for correspondence: Faculty of Economics, Soka University, 1-236 Tangi-machi, Hachioji, Tokyo 192-8577, Japan. Email address: m-asai@soka.ac.jp.

## Abstract

The paper considers various extended asymmetric multivariate conditional volatility models, and derives appropriate regularity conditions and associated asymptotic theory. This enables checking of internal consistency and allows valid statistical inferences to be drawn based on empirical estimation. For this purpose, we use an underlying vector random coefficient autoregressive process, for which we show the equivalent representation for the asymmetric multivariate conditional volatility model, to derive asymptotic theory for the quasi-maximum likelihood estimator. As an extension, we develop a new multivariate asymmetric long memory volatility model, and discuss the associated asymptotic properties.

**Keywords:** Multivariate conditional volatility, Vector random coefficient autoregressive process, Asymmetry, Long memory, Dynamic conditional correlations, Regularity conditions, Asymptotic properties.

**JEL Classification:** C13, C32, C58.

# 1 Introduction

Multivariate generalized autoregressive conditional heteroskedasticity (GARCH) models are frequently used in the analysis of dynamic covariance structure for multiple asset returns of financial time series (see the survey papers of, among others, Bauwens et al. (2006), McAleer (2005), and Silvennoinen and Teräsvirta (2009)). One of the most popular multivariate GARCH models is the BEKK model (see Baba, Engle, Kraft and Kroner (1985) and Engle and Kroner (1995)). The BEKK model has a positive definite covariance process, and it is easy to verify its stationary conditions. To reduce the number of parameters, and to show regularity conditions and asymptotic properties, the ‘diagonal BEKK’ and ‘scalar BEKK’ models are often used in empirical analysis. Comte and Lieberman (2003) show the consistency and asymptotic normality of the quasi-maximum likelihood (QML) estimator under conditions that are difficult to verify.

For accommodating the asymmetric effects in the multivariate framework, McAleer, Hoti and Chan (2009) consider the vector autoregressive and moving-average (VARMA) process with constant correlations and an asymmetric GARCH extension of the univariate asymmetric model of Glosten, Jagannathan, and Runkle (GJR) (1992). Taking account of dynamic correlations, Kroner and Ng (1998) develop the asymmetric BEKK (ABEKK) model. McAleer, Hoti and Chan (2009) show the consistency and asymptotic normality of the QML estimator of the asymmetric model with static correlations, but there are no asymptotic results for the ABEKK model.

In addition to asymmetric effects, another popular stylized fact is long-range dependence in volatility. In univariate conditional volatility models, Baillie, Bollerslev, and Mikkelsen (1996) developed the fractionally-integrated GARCH (FIGARCH) model, while Bollerslev and Mikkelsen (1996) suggested the fractionally-integrated exponential GARCH (FIEGARCH) model (see McAleer and Hafner (2014) and Martinet and McAleer (2016) for reservations regarding exponential GARCH).

Other studies have used the heterogeneous autoregressive (HAR) model of Corsi (2009), which is inspired by the heterogeneous ARCH model of Müller et al. (1997), to approximate the hyperbolic decay rates associated with long memory models.

The first purpose of the paper is to derive the consistency and asymptotic normality of the QML estimator for the VARMA-ABEKK model. For this purpose, we apply the approach of McAleer et al. (2008) based on the vector random coefficient autoregressive (RCA) process suggested by Nicholls and Quinn (1981) (see also Tsay (1987) for an application to conditional volatility models). The second purpose of the paper is to develop new extended asymmetric long memory BEKK (ALBEKK) and heterogeneous BEKK models, and to discuss the asymptotic properties of the associated QML estimators.

The remainder of the paper is organized as follows. Section 2 introduces the VARMA-ABEKK model, and shows a relationship between a vector RCA process and the conditional covariance model. Section 3 demonstrates the consistency and asymptotic normality of the QML estimator for the VARMA-ABEKK model. Section 4 presents the new ALBEKK and HABEKK models for long memory, and discusses the asymptotic properties of the associated QML estimators. Section 5 gives some concluding remarks. All proofs are given in the Appendix.

## 2 Asymmetric Multivariate GARCH Models

Let  $\mathbf{y}_t$  be an  $m \times 1$  vector, and consider the following asymmetric multivariate GARCH model:

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \quad (1)$$

$$\boldsymbol{\varepsilon}_t = \mathbf{H}_t^{1/2} \boldsymbol{\xi}_t, \quad \boldsymbol{\xi}_t \sim iid(0, I_m), \quad (2)$$

$$\mathbf{H}_t = \mathbf{W} + \sum_{i=1}^r [\mathbf{A}_i \boldsymbol{\varepsilon}_{t-i} \boldsymbol{\varepsilon}'_{t-i} \mathbf{A}'_i + \mathbf{C}_i \boldsymbol{\eta}_{t-i} \boldsymbol{\eta}'_{t-i} \mathbf{C}'_i] + \sum_{j=1}^s \mathbf{B}_s \mathbf{H}_{t-s} \mathbf{B}'_s, \quad (3)$$

where  $\mathbf{y}_t = (y_{1t}, \dots, y_{mt})'$ ,  $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \dots, \varepsilon_{mt})'$ ,  $\boldsymbol{\xi}_t = (\xi_{1t}, \dots, \xi_{mt})'$ ,  $\mathbf{A}_i$ ,  $\mathbf{B}_j$  and  $\mathbf{C}_i$  ( $i = 1, \dots, r$ ) ( $j = 1, \dots, s$ ) are  $m$ -dimensional square matrices,  $W$  is an  $m$ -dimensional positive definite matrix,  $\boldsymbol{\eta}_t = (n_{1t}\varepsilon_{1t}, \dots, n_{mt}\varepsilon_{mt})'$ , and  $n_{it} = \mathbf{1}(\varepsilon_{it} < 0)$ . For purposes of identification, the restrictions  $a_{11,i} \geq 0$ ,  $b_{11,j} \geq 0$  and  $c_{11,i} \geq 0$  are imposed. As the model encompasses the BEKK model of Engle and Kroner (1995), we will call this the ‘asymmetric BEKK’ (ABEKK) model. If  $r = s = 1$ , the ABEKK specification reduces to the model of Kroner and Ng (1998).

The vector form of the covariance matrix is given by:

$$\mathbf{h}_t = \mathbf{w} + \sum_{i=1}^r [(\mathbf{A}_i \otimes \mathbf{A}_i) + (\mathbf{C}_i \otimes \mathbf{C}_i)(\mathbf{N}_{t-i} \otimes \mathbf{N}_{t-i})] \tilde{\boldsymbol{\varepsilon}}_{t-i} + \sum_{j=1}^s (\mathbf{B}_j \otimes \mathbf{B}_j) \mathbf{h}_{t-j}, \quad (4)$$

where  $\mathbf{h}_t = \text{vec}(\mathbf{H}_t)$ ,  $\tilde{\boldsymbol{\varepsilon}}_t = \text{vec}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$ ,  $\mathbf{w} = \text{vec}(W)$ ,  $\mathbf{N}_t$  is a diagonal matrix with diagonal elements formed from the vector of indicator functions  $\mathbf{n}_t = (n_{1t}, \dots, n_{mt})'$ , and  $\otimes$  denotes the Kronecker product. As in Ling and McAleer (2003), we assume:

$$\boldsymbol{\mu}_t = \sum_{i=1}^p \boldsymbol{\Phi}_i L^i \mathbf{y}_t + \sum_{j=1}^q \boldsymbol{\Theta}_j L^j \boldsymbol{\varepsilon}_t, \quad (5)$$

where  $\boldsymbol{\Phi}_i$  and  $\boldsymbol{\Theta}_j$  are  $m \times m$  matrices, the roots of the characteristic polynomials  $|I_m - \sum_{i=1}^p \boldsymbol{\Phi}_i L^i|$  and  $|I_m - \sum_{j=1}^q \boldsymbol{\Theta}_j L^j|$  lie outside the unit circle, and  $L$  is the lag operator. Given the specification,  $\mathbf{y}_t$  follows the vector autoregressive moving-average (VARMA) process with the ABEKK structure, and we will call this the ‘VARMA-ABEKK’ model.

By extending the work of McAleer et al. (2008), we can derive the ABEKK model from a vector RCA process, as shown in the following proposition.

**Proposition 1.** (i) Consider the following vector RCA process:

$$\boldsymbol{\varepsilon}_t = \sum_{i=1}^r \left\{ \tilde{\mathbf{A}}_{it} + \tilde{\mathbf{C}}_{it} \right\} \boldsymbol{\varepsilon}_{t-i} + \boldsymbol{\zeta}_t, \quad \boldsymbol{\zeta}_t \sim iid(0, \Gamma), \quad (6)$$

where  $\zeta_t = (\zeta_{1t}, \dots, \zeta_{mt})'$ ,  $\Gamma$  is a positive definite covariance matrix, and the  $m \times m$  matrices of random coefficients  $\tilde{\mathbf{A}}_{it} = \{a_{j,l,it}\}$  and  $\tilde{\mathbf{C}}_{it} = \{c_{j,l,it}\}$  satisfy:

$$E_{\varepsilon,t-1}(\tilde{\mathbf{A}}_{it}) = O, \quad \forall i, t,$$

$$E_{\varepsilon,t-1}(\tilde{a}_{j_1,l_1,it}\tilde{a}_{l_2,j_2,it}) = a_{j_1,l_1}a_{l_2,j_2} \quad (j_1, j_2, l_1, l_2 = 1, \dots, k),$$

$$E_{\varepsilon,t-1}(\tilde{a}_{j_1,l_1,it}\tilde{a}_{l_2,j_2,js}) = 0 \text{ if } i \neq j \text{ and/or } t \neq s, (j_1, j_2, l_1, l_2 = 1, \dots, k),$$

$$E_{\varepsilon,t-1}(\tilde{\mathbf{C}}_{it}) = O, \quad \forall i, t,$$

$$E_{\varepsilon,t-1}(\tilde{c}_{j_1,l_1,it}\tilde{c}_{l_2,j_2,it}) = \begin{cases} c_{j_1,l_1}c_{l_2,j_2} & \text{if } \varepsilon_{l_1,t-1} < 0 \text{ and } \varepsilon_{l_2,t-1} < 0 \\ 0 & \text{otherwise} \end{cases} \\ (j_1, j_2, l_1, l_2 = 1, \dots, k),$$

$$E_{\varepsilon,t-1}(\tilde{c}_{j_1,l_1,it}\tilde{c}_{l_2,j_2,js}) = 0 \text{ if } i \neq j \text{ and/or } t \neq s, (j_1, j_2, l_1, l_2 = 1, \dots, k),$$

and  $\eta_t$ ,  $\tilde{\mathbf{A}}_{it}$  and  $\tilde{\mathbf{C}}_{it}$  are mutually independent for all  $i$  and  $t$ , but  $\tilde{\mathbf{C}}_{it}$  depends on  $\varepsilon_t$ . We denote  $E_{\varepsilon,t-1}$  as the expectation conditional on  $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$ , so that the conditional variance of  $\varepsilon_t$  is:

$$\mathbf{H}_t = E_{\varepsilon,t-1}(\varepsilon_t \varepsilon_t') = \sum_{i=1}^r [\mathbf{A}_i \varepsilon_{t-i} \varepsilon_{t-i}' \mathbf{A}_i' + \mathbf{C}_i \eta_{t-i} \eta_{t-i}' \mathbf{C}_i'] + \Gamma.$$

(ii) Consider the infinite-order vector RCA process:

$$\varepsilon_t = \sum_{i=1}^{\infty} \left\{ \tilde{\mathbf{A}}_{it}^* + \tilde{\mathbf{C}}_{it}^* \right\} \varepsilon_{t-i} + \zeta_t, \quad (7)$$

where  $\tilde{\mathbf{A}}_{it}^*$  and  $\tilde{\mathbf{C}}_{it}^*$  are defined similarly to  $\tilde{\mathbf{A}}_{it}$  and  $\tilde{\mathbf{C}}_{it}$ , respectively. Then the conditional variance is given by:

$$\mathbf{H}_t = \sum_{i=1}^{\infty} [\mathbf{A}_i^* \varepsilon_{t-i} \varepsilon_{t-i}' \mathbf{A}_i^{*'} + \mathbf{C}_i^* \eta_{t-i} \eta_{t-i}' \mathbf{C}_i^{*'}] + \Gamma, \quad (8)$$

which is also obtained by the ABEKK model (3), if the roots of the characteristic polynomials  $|I_{m^2} - \sum_{j=1}^s (\mathbf{B}_j \otimes \mathbf{B}_j) L^j|$  lie outside the unit circle. For the case  $r = s = 1$ , under the condition that the roots of  $|I_{m^2} - (\mathbf{B}_1 \otimes \mathbf{B}_1) L|$  lie outside the unit circle, the conditional covariance of  $\varepsilon_t$  in (7) is equivalent to (3) if and only if  $\mathbf{A}_i^* = \mathbf{B}^i \mathbf{A}$ ,  $\mathbf{C}_i^* = \mathbf{B}^i \mathbf{C}$ , and  $\text{vec}(\Gamma) = [\sum_{i=0}^{\infty} (\mathbf{B}^i \otimes \mathbf{B}^i)]^{-1} \text{vec}(W)$ .

For the equivalence of (2) and (7), we can derive the asymptotic theory of the VARMA-ABEKK model by applying the results in McAleer et al. (2008).

### 3 Structural and Statistical Properties

Denote the parameter vector  $\boldsymbol{\lambda} = (\boldsymbol{\theta}', \boldsymbol{\tau}')'$ ,  $\boldsymbol{\theta} = (\text{vec}(\boldsymbol{\Phi}_1)', \dots, \text{vec}(\boldsymbol{\Phi}_p)', \text{vec}(\boldsymbol{\Theta}_1)', \dots, \text{vec}(\boldsymbol{\Theta}_q)')$ ,  $\boldsymbol{\tau} = (\text{vech}(\boldsymbol{W})', \text{vec}(\boldsymbol{A}_1)', \dots, \text{vec}(\boldsymbol{A}_r)', \text{vec}(\boldsymbol{B}_1)', \dots, \text{vec}(\boldsymbol{B}_s)')$ , and the true parameter vector as  $\boldsymbol{\lambda}_0$ . We assume that the parameter space  $\Lambda$  is a compact subspace of Euclidean space, such that  $\boldsymbol{\lambda}_0$  is an interior point in  $\Lambda$ . We do not consider the situation in which the parameter is on the boundary of the parameter space.

For each  $\boldsymbol{\lambda} \in \Lambda$ , we make the following assumptions.

**Assumption 1.** *All the roots of:*

$$\left| I_{m^2} - \sum_{i=1}^r [(\boldsymbol{A}_i \otimes \boldsymbol{A}_i) + (\boldsymbol{C}_i \otimes \boldsymbol{C}_i)(\boldsymbol{N}_t \otimes \boldsymbol{N}_t)] L^i - \sum_{j=1}^s (\boldsymbol{B}_j \otimes \boldsymbol{B}_j) L^j \right| = 0$$

*are outside the unit circle. Moreover,  $I_{m^2} - \sum_{i=1}^r [(\boldsymbol{A}_i \otimes \boldsymbol{A}_i) + (\boldsymbol{C}_i \otimes \boldsymbol{C}_i)(\boldsymbol{N}_t \otimes \boldsymbol{N}_t)] L^i$  and  $\sum_{j=1}^s (\boldsymbol{B}_j \otimes \boldsymbol{B}_j) L^j$  are left coprime, and satisfy other identifiability conditions given in Ling and McAleer (2003).*

**Assumption 2.** *For the vector RCA process (7), the distribution of  $\boldsymbol{\zeta}_t$  is symmetric. For the vector of second moments,  $\tilde{\boldsymbol{\zeta}}_t = \text{vec}(\boldsymbol{\zeta}_t \boldsymbol{\zeta}_t')$ , we assume  $E(\tilde{\boldsymbol{\zeta}}_t) = \boldsymbol{\gamma} = \text{vec}(\boldsymbol{\Gamma})$  and  $\boldsymbol{\Gamma}_{\tilde{\boldsymbol{\zeta}} \tilde{\boldsymbol{\zeta}}'}$  is positive definite, where  $\boldsymbol{\Gamma}_{\tilde{\boldsymbol{\zeta}} \tilde{\boldsymbol{\zeta}}'} = E \left[ (\tilde{\boldsymbol{\zeta}}_t - \boldsymbol{\gamma}) (\tilde{\boldsymbol{\zeta}}_t - \boldsymbol{\gamma})' \right]$ . For the fourth moments of  $\tilde{\boldsymbol{A}}_{it}$  and  $\tilde{\boldsymbol{C}}_{it}$ , we assume:*

$$E|\tilde{a}_{j_1, l_1, it}^* \tilde{a}_{j_2, l_2, it}^* \tilde{a}_{j_3, l_3, it}^* \tilde{a}_{j_3, l_3, it}^*| < \infty,$$

$$E|\tilde{c}_{j_1, l_1, it}^* \tilde{c}_{j_2, l_2, it}^* \tilde{c}_{j_3, l_3, it}^* \tilde{c}_{j_3, l_3, it}^*| < \infty \quad (j_1, j_2, j_3, j_4, l_1, l_2, l_3, l_4 = 1, \dots, m),$$

respectively. Moreover, all the roots of:

$$\left| I_{m^4} - \sum_{i=1}^{\infty} E \left[ \left( \tilde{\mathbf{A}}_{it}^{*2} \otimes \tilde{\mathbf{A}}_{it}^{*2} \right) + \left( \tilde{\mathbf{C}}_{it}^{*2} \otimes \tilde{\mathbf{C}}_{it}^{*2} \right) \right] \right| = 0,$$

are outside the unit circle.

**Assumption 3.** The function  $\mathbf{h}_t$  is such that,  $\forall \lambda \in \Lambda$  and  $\forall \lambda_0 \in \Lambda$ ,  $\mathbf{h}_{t,\lambda} = \mathbf{h}_{t,\lambda_0}$  almost surely (a.s.), if and only if  $\lambda = \lambda_0$ .

Note that Assumption 3 is an identifiability condition, analogous to Assumption A4 of Jeantheau (1998). The structural properties of the model are developed and the analytical forms of the regularity conditions are derived in Proposition 2 and Theorem 1, respectively.

**Proposition 2.** Under Assumptions 1 and 2, the VARMA-ABEKK model based on the vector RCA process (7) possesses an  $\mathfrak{S}_{y,t}$ -measurable second-order stationary solution  $\{\mathbf{y}_t, \mathbf{\varepsilon}_t, \mathbf{h}_t\}$ , where  $\mathfrak{S}_{y,t}$  is a  $\sigma$ -field generated by  $\{\mathbf{y}_k : k \leq t\}$ . Define an  $m^2(s+r) \times 1$  vector as  $\mathbf{v}_t = (0, \dots, 0, \tilde{\mathbf{\varepsilon}}_t' - \mathbf{\omega}', 0, \dots, 0)'$ , with the sunbector consisting of the  $(m^2s+1)$ th to  $m^2(s+1)$ th columns as  $\tilde{\mathbf{\varepsilon}}_t - \mathbf{\omega}$ , where  $\mathbf{\omega} = \text{vec}(\mathbf{\Omega})$ . The solution  $\mathbf{h}_t$  has the following causal representation:

$$\mathbf{h}_t = \mathbf{\omega} + \underline{\mathbf{C}}' \sum_{j=1}^{\infty} \left( \prod_{i=1}^j \mathbf{\Psi}_{t+1-i} \right) \mathbf{v}_{t-i}, \text{ a.s.},$$

where  $\underline{\mathbf{C}} = [I_{m^2} \ O_{m \times m(s-1)}]'$ , which is an  $ms \times m$  matrix, and:

$$\mathbf{\Psi}_t = \begin{pmatrix} \mathbf{\Psi}_{11} & \mathbf{\Psi}_{12,t}^\dagger \\ O_{m^2r \times m^2s} & \mathbf{\Psi}_{22} \end{pmatrix}, \quad \mathbf{\Psi}_{11} = \begin{pmatrix} \mathbf{B}_1^\dagger & \cdots & \mathbf{B}_{s-1}^\dagger & \mathbf{B}_s^\dagger \\ I_{m^2(s-1)} & & & O_{m^2(s-1) \times m^2} \end{pmatrix},$$

$$\mathbf{\Psi}_{12,t} = \begin{pmatrix} \mathbf{A}_{1t}^\dagger & \cdots & \mathbf{A}_{rt}^\dagger \\ O_{m^2(s-1) \times m^2r} & & \end{pmatrix}, \quad \mathbf{\Psi}_{22} = \begin{pmatrix} O_{m^2 \times m^2r} \\ I_{m^2(r-1)} \end{pmatrix},$$

with  $\mathbf{B}_i^\dagger = (\mathbf{B}_i \otimes \mathbf{B}_i)$ ,  $\mathbf{A}_{it}^\dagger = (\mathbf{A}_i \otimes \mathbf{A}_i) + (\mathbf{C}_i \otimes \mathbf{C}_i)(\mathbf{N}_{t+1-i} \otimes \mathbf{N}_{t+1-i})$ , and  $\mathbf{N}_t$  is the  $m \times m$  diagonal matrix with the diagonal elements of  $(\mathbf{1}(\varepsilon_{1t} < 0), \dots, \mathbf{1}(\varepsilon_{mt} < 0))$ .

**Theorem 1.** (i) Under Assumptions 1 and 2 for the VARMA-ABEKK model without assuming the vector RCA structure, if  $\rho \left[ E \left( \Psi_t^{\otimes l} \right) \right] < 1$ , with  $l$  being a strictly positive integer, then the  $2l$ th moments of  $\{\mathbf{y}_t, \boldsymbol{\varepsilon}_t\}$  are finite, where  $\rho(\mathbf{A})$  denotes the largest modulus of the eigenvalues of a matrix  $\mathbf{A}$ ,  $\Psi_t$  is defined as in Proposition 2, and  $\mathbf{A}^{\otimes l}$  is the Kronecker product of the  $l$  matrices  $\mathbf{A}$ .

(ii) Under Assumptions 1 and 2 for the VARMA-ABEKK model based on the vector RCA process (7), if  $\rho \left[ E \left( \Psi_t^{\otimes l} \right) \right] < 1$ , with  $l$  being a strictly positive integer, and if  $2l$ th moments of  $\boldsymbol{\zeta}_t$  are finite, then the  $2l$ th moments of  $\{\mathbf{y}_t, \boldsymbol{\varepsilon}_t\}$  are finite.

Given these structural properties, the statistical properties of the model are established in Theorems 2–4, with sufficient multivariate log-moment conditions for consistency in Theorem 2, sufficient second-order moment conditions for consistency in Theorem 3, and sufficient conditions for asymptotic normality in Theorem 4.

The QMLE of the parameters in the model (1)–(3) are obtained by maximizing, conditional on the true  $(\mathbf{y}_t, \mathbf{h}_t)$ , the following log-likelihood function:

$$L_T(\boldsymbol{\lambda}) = \frac{1}{T} \sum_{t=1}^T l_t(\boldsymbol{\lambda}), \quad (9)$$

$$l_t(\boldsymbol{\lambda}) = -\frac{1}{2} \left( \log |\mathbf{H}_t| + \boldsymbol{\varepsilon} \mathbf{H}_t^{-1} \boldsymbol{\varepsilon}' \right),$$

where  $l_t(\boldsymbol{\lambda})$  takes the form of the Gaussian log-likelihood function, so that the QMLE is given as:

$$\hat{\boldsymbol{\lambda}} = \underset{\boldsymbol{\lambda} \in \Lambda}{\operatorname{argmax}} L_T(\boldsymbol{\lambda}).$$

Maximization of (9) leads to the following consistency result.

**Theorem 2.** Denote  $\hat{\boldsymbol{\lambda}}$  as the QMLE of  $\boldsymbol{\lambda}$ . Under Conditions C1–C6 in the Appendix,  $\hat{\boldsymbol{\lambda}} \rightarrow_p \boldsymbol{\lambda}$ .

An alternative proof of consistency of the QMLE based on second moments is to verify the sufficient conditions of Theorem 4.1.1 in Amemiya (1985), as demonstrated for the VARMA-GARCH model in Ling and McAleer (2003).

**Theorem 3.** *Denote  $\hat{\boldsymbol{\lambda}}$  as the QMLE of  $\boldsymbol{\lambda}_0$ . Under Conditions D1–D6 in the Appendix,  $\hat{\boldsymbol{\lambda}} \rightarrow_p \boldsymbol{\lambda}_0$ .*

Given the consistency of  $\hat{\boldsymbol{\lambda}}$ , the following theorem provides sufficient conditions for asymptotic normality.

**Theorem 4.** *Let  $\mathbf{y}_t$  be generated by VARMA-ABEKK model, based on the vector RCA process (7). Given the consistency of  $\hat{\boldsymbol{\lambda}}$  for  $\boldsymbol{\lambda}_0$ , under Conditions E1–E3 in the Appendix, it can be shown that:*

$$\sqrt{T} (\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0) \xrightarrow{d} N(0, \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\Omega}_\lambda \boldsymbol{\Sigma}_0^{-1}).$$

## 4 Multivariate Long Memory Asymmetric Conditional Volatility Models

In this section, we develop a new long memory ABEKK model as follows. Using the notation in Proposition 2, we can write equation (4) as:

$$\mathbf{h}_t = \mathbf{w} + \sum_{i=1}^r \mathbf{A}_i^\dagger \tilde{\boldsymbol{\varepsilon}}_{t-i} + \sum_{j=1}^s \mathbf{B}_j^\dagger \mathbf{h}_{t-j} = \mathbf{w} + \mathbf{A}^\dagger(L) \tilde{\boldsymbol{\varepsilon}}_t + \mathbf{B}^\dagger(L) \mathbf{h}_t.$$

For simplicity, we assume  $\mathbf{C}_i = \mathbf{O}$  so that  $\mathbf{A}_{it}^\dagger = \mathbf{A}_i^\dagger$ . Upon rearranging the terms, it follows that:

$$\left[ I_{m^2} - \mathbf{A}^\dagger(L) - \mathbf{B}^\dagger(L) \right] \tilde{\boldsymbol{\varepsilon}}_t = \mathbf{w} + [I_{m^2} - \mathbf{B}^\dagger(L)] \boldsymbol{\nu}_t,$$

where  $\boldsymbol{\nu}_t = \tilde{\boldsymbol{\varepsilon}}_t - \mathbf{h}_t$ , so that  $E_{\varepsilon,t-1}(\boldsymbol{\nu}_t) = 0$ . Following Bollerslev (1986) and Engle and Kroner (1995), we can interpret the volatility equation of the ABEKK model as a VARMA( $\max(r, s), r$ )

model for  $\tilde{\varepsilon}_t$ . As a multivariate extension of the integrated GARCH model of Engle and Bollerslev (1986), we can set  $I_{m^2} - \mathbf{A}^\dagger(L) - \mathbf{B}^\dagger(L) = (I_{m^2} - \mathbf{A}^\ddagger(L)) [(1 - L)I_{m^2}]$  to obtain:

$$(I_{m^2} - \mathbf{A}_t^\ddagger(L)) [(1 - L)I_{m^2}] \tilde{\varepsilon}_t = \mathbf{w} + [I_{m^2} - \mathbf{B}^\dagger(L)] \boldsymbol{\nu}_t.$$

By using the fractional differencing operator of a diagonal matrix, defined by:

$$\mathbf{D}(L) = \mathbf{D}_\varepsilon(L) \otimes \mathbf{D}_\varepsilon(L), \quad \mathbf{D}_\varepsilon(L) = \begin{pmatrix} (1 - L)^{d_1} & & O \\ & \ddots & \\ O & & (1 - L)^{d_m} \end{pmatrix},$$

where  $|d_j| < 1/4$  ( $j = 1, \dots, m$ ), we obtain a multivariate extension of the fractionally-integrated GARCH (FIGARCH) model of Baillie, Bollerslev, and Mikkelsen (1996) as:

$$(I_{m^2} - \mathbf{A}^\ddagger(L)) \mathbf{D}(L) \tilde{\varepsilon}_t = \mathbf{w} + [I_{m^2} - \mathbf{B}^\dagger(L)] \boldsymbol{\nu}_t,$$

which has an alternative form:

$$\mathbf{h}_t = \mathbf{w} + \left[ I_m - \left\{ I_{m^2} - \mathbf{A}^\ddagger(L) \right\} \mathbf{D}(L) \right] \tilde{\varepsilon}_t + \mathbf{B}^\dagger(L) \mathbf{h}_t,$$

to produce the long memory BEKK specification:

$$\mathbf{H}_t = \mathbf{W} + [\varepsilon_t \varepsilon_t' - \mathbf{D}_\varepsilon(L) \varepsilon_t \varepsilon_t' \mathbf{D}_\varepsilon(L)] + \sum_{i=1}^r \mathbf{A}_i \mathbf{D}_\varepsilon(L) \varepsilon_{t-i} \varepsilon_{t-i}' \mathbf{D}_\varepsilon(L) \mathbf{A}_i' + \sum_{j=1}^s \mathbf{B}_s \mathbf{H}_{t-s} \mathbf{B}_s'.$$

By extending the above result, we can develop the asymmetric long memory BEKK (ALBEKK) model (1), (2) and:

$$\begin{aligned} \mathbf{H}_t = \mathbf{W} + [\varepsilon_t \varepsilon_t' - \mathbf{D}_\varepsilon(L) \varepsilon_t \varepsilon_t' \mathbf{D}_\varepsilon(L)] + \sum_{i=1}^r \mathbf{A}_i \mathbf{D}_\varepsilon(L) \varepsilon_{t-i} \varepsilon_{t-i}' \mathbf{D}_\varepsilon(L) \mathbf{A}_i' \\ + \sum_{i=1}^r \mathbf{C}_i \mathbf{D}_\varepsilon(L) \boldsymbol{\eta}_{t-i} \boldsymbol{\eta}_{t-i}' \mathbf{D}_\varepsilon(L) \mathbf{C}_i' + \sum_{j=1}^s \mathbf{B}_s \mathbf{H}_{t-s} \mathbf{B}_s'. \end{aligned} \quad (10)$$

The following proposition shows the equivalence of the ALBEKK representation (10) and the infinite-order vector RCA process.

**Proposition 3.** Consider the infinite-order vector RCA process defined by (7) for  $\varepsilon_t$ . The conditional variance of  $\varepsilon_t$  given by (8) is also obtained from the ALBEKK model (10) if the roots of the characteristic polynomials,  $|I_{m^2} - \sum_{j=1}^s (\mathbf{B}_j \otimes \mathbf{B}_j) L^j|$ , lie outside the unit circle.

The proof is a straightforward extension of the proof of Proposition 1.

To prove consistency and asymptotic normality of the QML estimator for the ALBEKK model, we need to derive a causal representation, as in Proposition 2:

$$\mathbf{h}_t = \boldsymbol{\omega} + \underline{\mathbf{C}}' \sum_{j=1}^{\infty} \boldsymbol{\Psi}_{t+1-i}^{\dagger} \mathbf{v}_{t-i}, \text{ a.s.},$$

where  $\boldsymbol{\Psi}_{t+1-i}^{\dagger}$  are defined by  $\mathbf{D}_{\varepsilon}(L)$  in addition to the matrices in Proposition 2. Derivation of the exact conditions for consistency and asymptotic normality of ALBEKK will be considered in future work.

As an alternative approach for empirical analysis, we may extend the approximation of long-range dependence in volatility processes by using the heterogeneous autoregressive (HAR) model of Corsi (2000) and heterogeneous ARCH model of Müller et al. (1997). Assume  $t$  denotes time on a daily basis, and consider the mean of the residuals for the past  $h$  days as:

$$(\varepsilon_{t-1})_h = h^{-1}(\varepsilon_{t-1} + \cdots + \varepsilon_{t-h}).$$

Then we can obtain the weekly ( $h = 5$ ) and monthly ( $h = 22$ ) means of the past  $\varepsilon_t$  as  $(\varepsilon_{t-1})_5$  and  $(\varepsilon_{t-1})_{22}$ , so as to define  $(\boldsymbol{\eta}_{t-1})_5$  and  $(\boldsymbol{\eta}_{t-1})_{22}$ , to obtain the heterogeneous ABEKK (HABEKK) model as:

$$\begin{aligned} \mathbf{H}_t = & \mathbf{W} + \mathbf{A}_d \varepsilon_{t-1} \varepsilon'_{t-1} \mathbf{A}'_d + \mathbf{A}_w (\varepsilon_{t-1})_5 (\varepsilon_{t-1})'_5 \mathbf{A}'_w + \mathbf{A}_m (\varepsilon_{t-1})_{22} (\varepsilon_{t-1})'_{22} \mathbf{A}'_m \\ & + \mathbf{C}_d \boldsymbol{\eta}_{t-1} \boldsymbol{\eta}'_{t-1} \mathbf{C}'_d + \mathbf{C}_w (\boldsymbol{\eta}_{t-1})_5 (\boldsymbol{\eta}_{t-1})'_5 \mathbf{C}'_w + \mathbf{C}_m (\boldsymbol{\eta}_{t-1})_{22} (\boldsymbol{\eta}_{t-1})'_{22} \mathbf{C}'_m \\ & + \mathbf{B} \mathbf{H}_{t-1} \mathbf{B}'. \end{aligned}$$

Since the HABEKK model is a special case of ABEKK(22,1), we can apply Theorems 2–4 for the consistency and asymptotic normality of the associated QML estimator.

## 5 Concluding Remarks

This paper considered alternative versions of the vector ARMA and asymmetric BEKK GARCH, or VARMA-ABEKK, models as extensions of the widely-used univariate asymmetric (or threshold) GJR model of Glosten et al. (1992). We showed the equivalence of the ABEKK specification and the infinite-order random coefficient autoregressive process, and established the unique, strictly stationary and ergodic solution of the model, its causal expansion, and convenient sufficient conditions for the existence of moments. We derived sufficient conditions for consistency and asymptotic normality of the associated QML estimator. We also developed asymmetric long memory BEKK and heterogeneous BEKK models for capturing long-range dependence in the volatility matrix, and discussed the asymptotic properties of the QML estimators.

## References

Amemiya, T. (1985), *Advanced Econometrics*, Harvard University Press, Cambridge, MA, USA.

Baba, Y., R. Engle, D. Kraft and K. Kroner (1985), “Multivariate Simultaneous Generalized ARCH”, Unpublished Paper, University of California, San Diego.

Baillie R. T., T. Bollerslev, and H. O. Mikkelsen (1996), “Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity”, *Journal of Econometrics*, **74**, 3–30.

Bauwens, L., S. Laurent, and J.V.K. Rombouts (2006), “Multivariate GARCH: A Survey”, *Journal of Applied Econometrics*, **21**, 79–109

Bollerslev, T., and H. O. Mikkelsen (1996), “Modeling and Pricing Long-Memory in Stock Market Volatility”, *Journal of Econometrics*, **73**, 151–184.

Comte, F. and O. Lieberman (2003), “Asymptotic Theory for Multivariate GARCH Processes”, *Journal of Multivariate Analysis*, **84**, 61–4.

Engle, R.F. and K.F. Kroner (1995), “Multivariate Simultaneous Generalized ARCH”, *Econometric Theory*, **11**, 122–150.

Glosten, L., R. Jagannathan, and D. Runkle (1992), “On the Relation between the Expected Value and Volatility of Nominal Excess Returns on Stocks”, *Journal of Finance*, **46**, 1779–1801.

Jeantheau, T. (1998), “Strong Consistency of Estimators for Multivariate ARCH Models”, *Econometric Theory*, **14**, 70–86.

Kroner, K. and V. Ng (1998), “Modeling Asymmetric Comovements of Asset Returns”, *Review Financial Studies*, **11**, 817–844.

Ling, S., and M. McAleer (2002), “Necessary and Sufficient Moment Conditions for the GARCH(r,s) and Asymmetric Power GARCH(r,s) Models”, *Econometric Theory*, **18**, 722–729.

Ling, S. and M. McAleer (2003), “Asymptotic Theory for A Vector ARMA-GARCH Model”, *Econometric Theory*, **19**, 278–308.

Martinet, G. and M. McAleer (2016), “On The Invertibility of EGARCH( $p,q$ )”, to appear in *Econometric Reviews*.

McAleer, M. (2005), “Automated Inference and Learning in Modeling Financial Volatility”, *Econometric Theory*, **21**, 232–261.

McAleer, M., F. Chan, S. Hoti and O. Lieberman (2008), “Generalized Autoregressive Conditional Correlation”, *Econometric Theory*, **24**, 1554–1583.

McAleer, M. and C. Hafner (2014), “A One Line Derivation of EGARCH”, *Econometrics*, **2**, 92–97.

Silvennoinen, A., and T. Teräsvirta (2009), “Multivariate GARCH Models”, In T. G. Andersen, R.A. Davis, J.-P. Kreiss, and T. Mikosch (eds.), *Handbook of Financial Time Series*, 201–229, New York: Springer.

McAleer, M., S. Hoti and F. Chan (2009), “Structure and Asymptotic Theory for Multivariate Asymmetric Conditional Volatility”, *Econometric Reviews*, **28**, 422–440.

Müller, U., M. Dacorogna, R. Dav, R. Olsen, O. Pictet, and J. von Weizsäcker (1997), “Volatilities of Different Time Resolutions Analysing The Dynamics of Market Components”, *Journal of Empirical Finance*, **4**, 213–239.

Nicholls, D.F. and B.G. Quinn (1981), *Random Coefficient Autoregressive Models: An Introduction*, Lecture Notes in Statistics 11, New York: Springer.

Tsay, R. (1987), “Conditional Heteroscedastic Time Series Models”, *Journal of the American Statistical Association*, **82**, 590–604.

Tweedie, R. (1988), “Invariant Measure for Markov Chains with No Irreducibility Assumptions”, *Journal of Applied Probability*, **25A**, 275–285.

## Appendix

### Proof of Proposition 1

Under the assumptions of Proposition 1, the VRCA process (6) gives

$$\begin{aligned}
& \left\{ E_{\varepsilon, t-1} (\varepsilon_t \varepsilon_t') \right\}_{j_1, j_2} \\
&= \left\{ E_{\varepsilon, t-1} \left( \sum_{i=1}^r \sum_{n=1}^r \tilde{\mathbf{A}}_{it} \varepsilon_{t-i} \varepsilon_{t-n} \tilde{\mathbf{A}}_{nt}' \right) \right\} + \left\{ E_{\varepsilon, t-1} \left( \sum_{i=1}^r \sum_{n=1}^r \tilde{\mathbf{C}}_{it} \varepsilon_{t-i} \varepsilon_{t-n} \tilde{\mathbf{C}}_{nt}' \right) \right\} + \gamma_{j_1, j_2} \\
&= \sum_{i=1}^r \sum_{n=1}^r \sum_{l_1=1}^m \sum_{l_2=1}^m (\varepsilon_{t-i} \varepsilon_{t-m}')_{l_1, l_2} E_{\varepsilon, t-1} (\tilde{a}_{j_1, l_1, it} \tilde{a}_{l_2, j_2, mt}) \\
&\quad + \sum_{i=1}^r \sum_{n=1}^r \sum_{l_1=1}^m \sum_{l_2=1}^m (\varepsilon_{t-i} \varepsilon_{t-m}')_{l_1, l_2} E_{\varepsilon, t-1} (\tilde{c}_{j_1, l_1, it} \tilde{c}_{l_2, j_2, mt}) + \gamma_{j_1, j_2} \\
&= \sum_{i=1}^r \sum_{l_1=1}^m \sum_{l_2=1}^m \left[ (\varepsilon_{t-i} \varepsilon_{t-i}')_{l_1, l_2} a_{j_1, l_1, i} a_{l_2, j_2, i} + (\eta_{t-i} \eta_{t-i}')_{l_1, l_2} c_{j_1, l_1, i} c_{l_2, j_2, i} \right] + \gamma_{j_1, j_2},
\end{aligned}$$

which is equivalent to the matrix given in Proposition 1(i).

It is straightforward to derive equation (8) from the result of (i). From the vector representation of the variance equation of the ABEKK model (4), if the roots of  $\left| I_{m^2} - \sum_{j=1}^s (\mathbf{B}_j \otimes \mathbf{B}_j) L^j \right|$  lie outside the unit circle, we obtain

$$\begin{aligned}
\mathbf{h}_t &= \boldsymbol{\gamma} + \left[ I_{m^2} - \sum_{j=1}^s (\mathbf{B}_j \otimes \mathbf{B}_j) L^j \right]^{-1} \sum_{i=1}^r \left[ (\mathbf{A}_i \otimes \mathbf{A}_i) L^i + (\mathbf{C}_i \otimes \mathbf{C}_i) L_i (\mathbf{N}_t \otimes \mathbf{N}_t) \right] \tilde{\varepsilon}_t \\
&= \boldsymbol{\gamma} + \sum_{i=1}^{\infty} \left[ (\dot{\mathbf{A}}_i \otimes \dot{\mathbf{A}}_i) + (\dot{\mathbf{C}}_i \otimes \dot{\mathbf{C}}_i) (\mathbf{N}_{t-i} \otimes \mathbf{N}_{t-i}) \right] \dot{\varepsilon}_{t-i}
\end{aligned}$$

where  $\boldsymbol{\gamma} = \left[ I_{m^2} - \sum_{j=1}^s (\mathbf{B}_j \otimes \mathbf{B}_j) \right]^{-1} \mathbf{w}$ . Therefore, we establish the equivalence between (8) and the variance equation of ABEKK by setting  $\boldsymbol{\gamma} = \text{vec}(\Gamma)$ ,  $\dot{\mathbf{A}}_i = \mathbf{A}_i^*$ , and  $\dot{\mathbf{C}}_i = \mathbf{C}_i^*$ . For  $r = s = 1$ , we obtain the condition straightforwardly by substituting past  $\mathbf{H}_t$  recursively in equation (3).  $\square$

## Proof of Proposition 2

Let  $\mathbf{y}_t^\dagger = (\mathbf{y}'_t, \dots, \mathbf{y}'_{t-p+1})'$ . It is straightforward to show that:

$$\mathbf{y}_t^\dagger = \Phi^\dagger \mathbf{y}_{t-1}^\dagger + \Theta^\dagger \varepsilon_t^\dagger = \sum_{i=0}^{\infty} (\Phi^\dagger)^i \Theta^\dagger \varepsilon_{t-i}^\dagger,$$

where  $\varepsilon_t^\dagger = (\varepsilon'_t, \dots, \varepsilon'_{t-q})'$ , and

$$\Phi^\dagger = \begin{pmatrix} \Phi_1 & \cdots & \Phi_{p-1} & \Phi_p \\ & I_{m(p-1)} & & O_{m(p-1) \times m} \end{pmatrix}, \quad \Theta^\dagger = \begin{pmatrix} I & \Theta_1 & \cdots & \Theta_q \\ & O_{m(p-1) \times m(q+1)} & & \end{pmatrix}.$$

For the vector RCA process (7), which has the conditional covariance (3), we obtain:

$$E(\varepsilon_t) = \mathbf{0}, \quad V(\varepsilon_t) = \Omega, \quad Cov(\varepsilon_{t_1}, \varepsilon_{t_2}) = O \quad (t_1 \neq t_2),$$

where

$$\text{vec}(\Omega) = \left( I_{m^2} - \sum_{i=1}^r (\mathbf{A}_i \otimes \mathbf{A}_i) - \sum_{i=1}^r (\mathbf{C}_i \otimes \mathbf{C}_i) E(\mathbf{N}_t \otimes \mathbf{N}_t) - \sum_{j=1}^s (\mathbf{B}_j \otimes \mathbf{B}_j) \right)^{-1} \text{vec}(\mathbf{W}).$$

Note that the diagonal elements of the matrix  $E(\mathbf{N}_t \otimes \mathbf{N}_t)$  are  $E(\mathbf{1}(\varepsilon_{l_1,t} < 0))$  or  $E(\mathbf{1}(\varepsilon_{l_1,t} < 0)\mathbf{1}(\varepsilon_{l_2,t} < 0))$  ( $l_1, l_2 = 1, \dots, m$ ), with finite values. By Assumption 1,  $\Omega$  exists. Since  $\varepsilon_t$  satisfies the conditions of the white noise process,  $\mathbf{y}_t^\dagger$  is second-order stationary, as is  $\mathbf{y}_t$ .

Let  $\mathbf{x}_t = (\mathbf{h}'_t, \dots, \mathbf{h}'_{t-s+1}, \tilde{\varepsilon}'_t, \dots, \tilde{\varepsilon}'_{t-r+1})' - (\boldsymbol{\iota}_{s+r} \otimes \boldsymbol{\omega})$ , where  $\boldsymbol{\omega} = \text{vec}(\Omega)$ , and  $\boldsymbol{\iota}_l$  is  $l \times 1$  vector of ones. It is straightforward to show that:

$$\mathbf{x}_t = \Psi_t \mathbf{x}_{t-1} + \mathbf{v}_t = \mathbf{v}_t + \sum_{j=1}^{\infty} \left( \prod_{i=1}^j \Psi_{t+1-i} \right) \mathbf{v}_{t-i},$$

where  $\Psi_t$  and  $\mathbf{v}_t$  are defined in Proposition 2. Note that  $\mathbf{h}_t = \boldsymbol{\omega} + \underline{\mathbf{C}}' \mathbf{x}_t$ . Since  $\mathbf{v}_t$  consists of zero and  $(\tilde{\varepsilon}_t - \boldsymbol{\omega})$ , we consider the variance of  $\tilde{\varepsilon}_t$ . By Assumptions 1 and 2, and Proposition 1, we can

show that  $E(\tilde{\varepsilon}_t) = \omega$ , and the conditional covariance matrix of  $\tilde{\varepsilon}_t$  is given by:

$$\begin{aligned}
& E_{\varepsilon,t-1} [(\tilde{\varepsilon}_t - \omega)(\tilde{\varepsilon}_t - \omega)'] \\
&= \Gamma_{\tilde{\zeta}\tilde{\zeta}'} + \sum_{i=1}^{\infty} (\mathbf{A}_i^* \otimes \mathbf{A}_i^*) (\tilde{\varepsilon}_{t-i} - \omega)(\tilde{\varepsilon}_{t-i} - \omega)' (\mathbf{A}_i^* \otimes \mathbf{A}_i^*)' \\
&+ \sum_{i=1}^{\infty} [(\Gamma \otimes (\mathbf{A}_i^* \varepsilon_{t-i} \varepsilon'_{t-i} \mathbf{A}_i^{*'})) \\
&+ (I_m \otimes \mathbf{A}_i^*) E_{\varepsilon,t-1} [\text{vec}(\varepsilon_{t-i} \zeta_t') \text{vec}(\varepsilon_{t-i} \zeta_t')'] (\mathbf{A}_i^{*'} \otimes I_m) \\
&+ (\mathbf{A}_i^* \otimes I_m) E_{\varepsilon,t-1} [\text{vec}(\zeta_t \varepsilon'_{t-i}) \text{vec}(\zeta_t \varepsilon'_{t-i})'] (I_m \otimes \mathbf{A}_i^{*'}) + ((\mathbf{A}_i^* \varepsilon_{t-i} \varepsilon'_{t-i} \mathbf{A}_i^{*'}) \otimes \Gamma)] \\
&+ \sum_{i=1}^{\infty} (\mathbf{C}_i^* \mathbf{N}_{t-i} \otimes \mathbf{N}_{t-i} \mathbf{C}_i^*) (\tilde{\varepsilon}_{t-i} - \omega)(\tilde{\varepsilon}_{t-i} - \omega)' (\mathbf{C}_i^* \mathbf{N}_{t-i} \otimes \mathbf{C}_i^* \mathbf{N}_{t-i})' \\
&+ \sum_{i=1}^{\infty} [(\Gamma \otimes (\mathbf{C}_i^* \mathbf{N}_{t-i} \varepsilon_{t-i} \varepsilon'_{t-i} \mathbf{N}_{t-i} \mathbf{C}_i^{*'})) \\
&+ (I_m \otimes \mathbf{C}_i^* \mathbf{N}_{t-i}) E_{\varepsilon,t-1} [\text{vec}(\varepsilon_{t-i} \zeta_t') \text{vec}(\varepsilon_{t-i} \zeta_t')'] (\mathbf{N}_{t-i} \mathbf{C}_i^{*'} \otimes I_m) \\
&+ (\mathbf{C}_i^* \mathbf{N}_{t-i} \otimes I_m) E_{\varepsilon,t-1} [\text{vec}(\zeta_t \varepsilon'_{t-i}) \text{vec}(\zeta_t \varepsilon'_{t-i})'] (I_m \otimes \mathbf{N}_{t-i} \mathbf{C}_i^{*'}) \\
&+ ((\mathbf{C}_i^* \mathbf{N}_{t-i} \varepsilon_{t-i} \varepsilon'_{t-i} \mathbf{N}_{t-i} \mathbf{C}_i^{*'}) \otimes \Gamma)].
\end{aligned} \tag{11}$$

Note that  $E_{\varepsilon,t-1} (\text{vec}(\varepsilon_{t-i} \zeta_t') \text{vec}(\varepsilon_{t-i} \zeta_t')')$  and  $E_{\varepsilon,t-1} (\text{vec}(\zeta_t \varepsilon'_{t-i}) \text{vec}(\zeta_t \varepsilon'_{t-i})')$  consist of elements of  $(\Gamma \otimes \varepsilon_{t-i} \varepsilon'_{t-i})$ . By equation (11), the unconditional covariance matrix of the second moments of  $\varepsilon_t$  is given by:

$$\begin{aligned}
& \text{vec}(E[(\tilde{\varepsilon}_t - \omega)(\tilde{\varepsilon}_t - \omega)']) \\
&= \left( I_{m^4} - \sum_{i=1}^{\infty} E \left[ \left( \tilde{\mathbf{A}}_{it}^{*2} \otimes \tilde{\mathbf{A}}_{it}^{*2} \right) + \left( \tilde{\mathbf{C}}_{it}^{*2} \otimes \tilde{\mathbf{C}}_{it}^{*2} \right) \right] \right)^{-1} \\
&\times \text{vec} \left( \Gamma_{\tilde{\zeta}\tilde{\zeta}'} + \sum_{i=1}^{\infty} [(\Gamma \otimes (\mathbf{A}_i^* \Omega \mathbf{A}_i^{*'})) + ((\mathbf{A}_i^* \Omega \mathbf{A}_i^{*'}) \otimes \Gamma) \right. \\
&\quad \left. + (I_m \otimes \mathbf{A}_i^*) E [\text{vec}(\varepsilon_{t-i} \zeta_t') \text{vec}(\varepsilon_{t-i} \zeta_t')'] (\mathbf{A}_i^{*'} \otimes I_m) \right. \\
&\quad \left. + (\mathbf{A}_i^* \otimes I_m) E [\text{vec}(\zeta_t \varepsilon'_{t-i}) \text{vec}(\zeta_t \varepsilon'_{t-i})'] (I_m \otimes \mathbf{A}_i^{*'}) \right] \\
&+ \sum_{i=1}^{\infty} [(\Gamma \otimes (\mathbf{C}_i^* E(\mathbf{N}_t \Omega \mathbf{N}_t) \mathbf{C}_i^{*'})) + ((\mathbf{C}_i^* E(\mathbf{N}_t \Omega \mathbf{N}_t) \mathbf{C}_i^{*'}) \otimes \Gamma)]
\end{aligned} \tag{12}$$

$$\begin{aligned}
& + E \left( (I_m \otimes \mathbf{C}_i^* \mathbf{N}_t) E \left[ \text{vec}(\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\zeta}_t') \text{vec}(\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\zeta}_t')' \right] (\mathbf{N}_t \mathbf{C}_i^{*'} \otimes I_m) \right) \\
& + E \left( (\mathbf{C}_i^* \mathbf{N}_t \otimes I_m) E \left[ \text{vec}(\boldsymbol{\zeta}_t \boldsymbol{\varepsilon}_{t-i}') \text{vec}(\boldsymbol{\zeta}_t \boldsymbol{\varepsilon}_{t-i}')' \right] (I_m \otimes \mathbf{N}_t \mathbf{C}_i^{*'}) \right) .
\end{aligned}$$

By Assumption 2, the inverse on the right-hand side of (12) exists, and  $\Gamma_{\tilde{\zeta}\tilde{\zeta}'}$  is positive definite. By Assumption 1 and Proposition 1, we can show that the matrices comprising the second and third infinite sums in (12) are positive definite, and all elements take finite values. Note that,  $E(\text{vec}(\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\zeta}_t') \text{vec}(\boldsymbol{\varepsilon}_{t-i} \boldsymbol{\zeta}_t')')$  and  $E(\text{vec}(\boldsymbol{\zeta}_t \boldsymbol{\varepsilon}_{t-i}') \text{vec}(\boldsymbol{\zeta}_t \boldsymbol{\varepsilon}_{t-i}')')$  consist of elements of  $(\mathbf{\Gamma} \otimes \mathbf{\Omega})$ . By Assumptions 1 and 2, and by Proposition 1, we can show that all the elements of  $E[(\tilde{\boldsymbol{\varepsilon}}_t - \boldsymbol{\omega})(\tilde{\boldsymbol{\varepsilon}}_t - \boldsymbol{\omega})']$  are finite, and the matrix is positive definite.

Corresponding to the above causal representation, define:

$$\dot{\mathbf{x}}_t = \mathbf{v}_t + \sum_{j=1}^T \left( \prod_{i=1}^j \boldsymbol{\Psi}_{t+1-i} \right) \mathbf{v}_{t-i},$$

and let  $\mathbf{e}_l = (0, \dots, 0, 1, 0, \dots, 0)'$ , which is an  $m(r+s) \times 1$  vector, and 1 appears in the  $l$ th position. Denote the  $l$ th element of  $\left( \prod_{i=1}^j \boldsymbol{\Psi}_{t+1-i} \right) \mathbf{v}_{t-i}$  by  $s_t$ :

$$s_t = \mathbf{e}_l' \left( \prod_{i=1}^j \boldsymbol{\Psi}_{t+1-i} \right) \mathbf{v}_{t-i}.$$

By Assumption 1,  $E|s_t| < \infty$  if and only if  $E|\mathbf{e}_l' \mathbf{v}_t| < \infty$ , which we can show by applying Hölder's inequality:

$$E|\mathbf{e}_l' \mathbf{v}_t| \leq [\mathbf{e}_l' E(\mathbf{v}_t \mathbf{v}_t') \mathbf{e}_l]^{1/2},$$

which we can show by the above result that  $E(\tilde{\boldsymbol{\varepsilon}}_t \tilde{\boldsymbol{\varepsilon}}_t')$  is positive definite, corresponding to the fourth moment of  $\boldsymbol{\varepsilon}_t$ . By Assumption 1, we can show  $E|s_t| \rightarrow 0$  as  $T \rightarrow \infty$ . Therefore, each component of  $\dot{\mathbf{x}}_t$  converges almost surely (a.s.) as  $T \rightarrow \infty$ , as does  $\mathbf{h}_t$ . Hence, there exists an  $\mathfrak{S}_t$ -measurable second-order solution  $\boldsymbol{\varepsilon}_t$  to (4).

To show uniqueness, let  $\check{\varepsilon}_t$  be another  $\mathfrak{S}_t$ -measurable second-order stationary solution to (4). Propositions 1 and 2 suffice to apply Corollary 2.2.2 of Nicholls and Quinn (1982) to show the uniqueness of  $\varepsilon_t$ . Thus,  $\check{\mathbf{x}}_t = \Psi_t \check{\mathbf{x}}_{t-1} + \mathbf{v}_t$ , where  $\check{\mathbf{x}}_t = (\check{\mathbf{h}}'_t, \dots, \check{\mathbf{h}}'_{t-s+1}, \check{\varepsilon}'_t, \dots, \check{\varepsilon}'_{t-r+1})'$  and  $(\mathbf{h}'_t, \dots, \mathbf{h}'_{t-s+1}, \varepsilon'_t, \dots, \varepsilon'_{t-r+1})' = (\mathbf{h}_t, \dots, \mathbf{h}_{t-s+1}, \varepsilon_t, \dots, \varepsilon_{t-r+1})' - (\boldsymbol{\iota}_{s+r} \otimes \boldsymbol{\omega})$ . Let  $\mathbf{u}_t = \mathbf{x}_t - \check{\mathbf{x}}_t$  to obtain  $\mathbf{u}_t = \left( \prod_{i=1}^j \Psi_{t+1-i} \right) \mathbf{u}_{t-i}$ . By Assumption 1 and Hölder's inequality, we obtain:

$$E|e'_l \mathbf{u}_t| \leq [e'_l E(\mathbf{u}_t \mathbf{u}'_t) e_l]^{1/2} \rightarrow 0 \text{ as } T \rightarrow \infty,$$

since  $\text{vec}(E(\mathbf{u}_t \mathbf{u}'_t)) = E \left[ \left( \prod_{i=1}^j \Psi_{t+1-i} \right) \otimes \left( \prod_{i=1}^j \Psi_{t+1-i} \right) \right] \text{vec}(E(\mathbf{u}_{t-i} \mathbf{u}'_{t-i}))$ . Hence, the solution is unique. As  $\mathbf{h}_t = \boldsymbol{\omega} + \underline{\mathbf{C}}' \mathbf{x}_t$ , it follows the unique causal representation is given by:

$$\mathbf{h}_t = \boldsymbol{\omega} + \underline{\mathbf{C}}' \sum_{j=1}^{\infty} \left( \prod_{i=1}^j \Psi_{t+1-i} \right) \mathbf{v}_{t-i}, \text{ a.s. } \square$$

## Proof of Theorem 1

For the first part, using the results on finite moments in Tweedie (1988), Lemma A.3 in Ling and McAleer (2003), and Lemma 1 in McAleer et al. (2008), Hölder's inequality implies that  $E_{\pi_1} \|\varepsilon_t\|^2 < (E_{\pi_1} \|\varepsilon_t\|^{2l})^{1/l} < \infty$ , where  $\pi_1$  are the stationary distributions of  $\{\varepsilon_t\}$ . Furthermore,  $E_{\pi_2} \|\mathbf{y}_t\|^2 < \infty$  by the proof of Proposition 2. Thus,  $\{\mathbf{y}_t, \varepsilon_t\}$  is a secondary stationary solution of (4). Moreover, the solution  $\{\mathbf{y}_t, \varepsilon_t\}$  is unique and ergodic by Proposition 2. Therefore,  $\{\mathbf{y}_t, \varepsilon_t\}$  satisfying model (4) has finite  $2l$ th moment. For the second part, it is straightforward from the first part.  $\square$

## Proof of Theorem 2

It is sufficient to verify the following conditions for consistency in Jeantheau (1998).

**C1.**  $\Lambda$  is compact.

**C2.**  $\forall \boldsymbol{\lambda} \in \Lambda$ , the model admits a unique strictly stationary and ergodic solution  $\mathbf{y}_t$ .

**C3.** There exists a deterministic constant  $c > 0$  such that,  $\forall t$  and  $\forall \lambda \in \Lambda$ ,  $|\mathbf{H}_t| > c$ .

**C4.** Assumption 3.

**C5.**  $\mathbf{y}_t$  and  $\mathbf{H}_t$  are continuous functions of the parameter  $\lambda$ .

**C6.**  $E_{\lambda_0} |\log(\mathbf{H}_t)| < \infty$ ,  $\forall \lambda_0 \in \Lambda$ .

Under Proposition 2, (4) admits a unique strictly stationary and ergodic solution of  $\mathbf{y}_t$  (C2).

Furthermore, the model is identifiable under Assumption 3 (C4). Note that the determinant of the conditional covariance matrix is strictly positive, by the structure of the BEKK representation (3) for all  $t$ . Hence, there exists a constant  $c > 0$  such that  $|E_{\epsilon, t-1}(\epsilon_t \epsilon_t')| > c \forall t$  and  $\forall \lambda \in \Lambda$ , where  $\Lambda$  is a compact subspace of Euclidean space (C1 and C3). By the square integrability of  $\epsilon_t$ ,  $E_{\lambda_0}(\text{vech}(\mathbf{H}_{t,\lambda})) < \infty$ , which establishes C6 (for details, see Comte and Lieberman, 2003, p.67).

Under Assumption 1, C6, and the structure (4)–(5),  $\mathbf{y}_t$  and  $\mathbf{H}_t$  are continuous functions of the parameter  $\lambda$  (C5).  $\square$

### Proof of Theorem 3

It is sufficient to verify the following conditions in Theorem 4.1.1 in Amemiya (1985).

**D1.**  $\Lambda$  is compact.

**D2.**  $L_T(\lambda)$  is continuous in  $\lambda \in \Lambda$  for  $\mathbf{y}_t$  and is a measurable function of  $\mathbf{y}_t \forall \lambda \in \Lambda$ .

**D3.**  $T^{-1}L_T(\lambda)$  converges to a non-stochastic function  $L(\lambda)$  in probability uniformly in  $\lambda \in \Lambda$  as  $T \rightarrow \infty$ , and  $L(\lambda)$  attains a unique global maximum at  $\lambda_0$ .

Condition D1 is equivalent to C1 and D2 follows from C5, so D1 and D2 are satisfied under Theorem 2. To verify D3, it is convenient to introduce the unobserved process,  $\{\epsilon_t^*, \mathbf{H}_t^* : t = 0, \pm 1, \pm 2, \dots\}$ . Define the unobserved log-likelihood function conditional on the infinite past

observations:

$$L_T^*(\boldsymbol{\lambda}) = \frac{1}{T} \sum_{t=1}^T l_t^*(\boldsymbol{\lambda}),$$

$$l_t^*(\boldsymbol{\lambda}) = -\frac{1}{2} (\log |\mathbf{H}_t^*| + \boldsymbol{\varepsilon}_t^{*'} \mathbf{H}_t^{*-1} \boldsymbol{\varepsilon}_t^*).$$

Lemmas 4.2, 4.4 and 4.6 in Ling and McAleer (2003), and condition C3, imply that  $L(\boldsymbol{\lambda})$  exists for all  $\boldsymbol{\lambda} \in \Lambda$ ,  $\sup_{\lambda \in \Lambda} |L_T(\boldsymbol{\lambda}) - L(\boldsymbol{\lambda})| = o_p(1)$ ,  $L(\boldsymbol{\lambda})$  has a unique maximum at  $\boldsymbol{\lambda}_0$ , and  $|L_T(\boldsymbol{\lambda}) - L(\boldsymbol{\lambda})| = o_p(1)$ . Thus,

$$\sup_{\lambda \in \Lambda} |L_T(\boldsymbol{\lambda}) - L(\boldsymbol{\lambda})| \leq \sup_{\lambda \in \Lambda} |L_T^*(\boldsymbol{\lambda}) - L(\boldsymbol{\lambda})| + \sup_{\lambda \in \Lambda} |L_T^*(\boldsymbol{\lambda}) - L_T(\boldsymbol{\lambda})| = o_p(1).$$

Therefore,  $L_T(\boldsymbol{\lambda}) \rightarrow_p L(\boldsymbol{\lambda})$  uniformly in  $\Lambda$  (D3).  $\square$

## Proof of Theorem 4

Given the consistency of  $\hat{\boldsymbol{\lambda}}$  for  $\boldsymbol{\lambda}_0$  in Theorems 2 and 3, it is sufficient to verify the following conditions of Theorem 4.1.3 in Amemiya (1985):

- E1.**  $\partial^2 L_T / \partial \lambda \partial \lambda$  exists and is continuous in an open, convex neighborhood of  $\boldsymbol{\lambda}_0$ .
- E2.**  $T^{-1}(\partial^2 L_T / \partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}')|_{\lambda_T}$  converges to a finite nonsingular matrix  $\Sigma_0 = E [T^{-1}(\partial^2 L_T / \partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}')|_{\lambda_T}]$  in probability for any sequence  $\boldsymbol{\lambda}_T$ , such that  $\hat{\boldsymbol{\lambda}} \rightarrow_p \boldsymbol{\lambda}_0$ .
- E3.**  $T^{-1/2}(\partial L_T / \partial \boldsymbol{\lambda})|_{\lambda_0} \rightarrow^d N(0, \boldsymbol{\Omega}_\lambda)$ , where  $\boldsymbol{\Omega}_\lambda = \lim E [T^{-1}(\partial L_T / \partial \boldsymbol{\lambda})|_{\lambda_0} \times (\partial L_T / \partial \boldsymbol{\lambda}')|_{\lambda_0}]$ .

By Theorems 2 and 3,  $\hat{\boldsymbol{\lambda}}$  is consistent for  $\boldsymbol{\lambda}_0$ . It follows from the conditions in Theorem 2 that  $\partial^2 L_T / \partial \lambda \partial \lambda$  exists and is continuous in  $\Lambda$ . Lemma 5.4 in Ling and McAleer (2003) can be used to verify that conditions E1 and E2 hold. Under the existence of fourth moments of  $\boldsymbol{\zeta}_t$  in Assumption 2, using the central limit theorem of Stout (1974), and the Cramér-Wold device, it follows that

$$T^{-1/2} \sum_{t=1}^T \frac{\partial l_t}{\partial \lambda} \xrightarrow{d} N(0, \boldsymbol{\Omega}_\lambda),$$

where  $\boldsymbol{\Omega}_\lambda$  is positive definite (E3).  $\square$