Statistics and Its Interface Volume 4 (2011) 227–233

# A note on asymptotic inference for FIGARCH$(p, d, q)$ models

Ngai Hang Chan* and Chi Tim Ng

Parameters estimation for a FIGARCH$(p, d, q)$ model is studied in this paper. By constructing a compact parameter space $\Theta$ satisfying the non-negativity constraints for the FIGARCH model, it is shown that the results of Robinson and Zaffaroni (2006) can be applied to establish the strong consistency and asymptotic normality of the quasi-maximum likelihood (QML) estimator of the FIGARCH model.

AMS 2000 subject classifications: Primary 62F12, 62E20; secondary 91B84.
Keywords and phrases: Asymptotic normality, Consistency, Fractionally-integrated GARCH model, Non-negativity, Quasi-maximum likelihood estimator.

## 1. INTRODUCTION

The fractionally-integrated (FI) GARCH model discussed in Baillie et al. (1996) and Bollerslev and Mikkelson (1996) has attracted a considerable amount of attention among economists and practitioners. In the empirical literature, the parameters of the FIGARCH model are commonly estimated using the quasi-maximum likelihood estimator (QMLE). Baillie et al. (1996) claimed that strong consistency and asymptotic normality of the QMLE can be established following similar arguments in Lee and Hansen (1994) for the GARCH $(1, 1)$ model. However, their claim was queried by Mikosch and Stărică (2002). On the other hand, Robinson and Zaffaroni (2006) develop a general theory of the QMLE of an ARCH$(\infty)$ model under a general framework. As an illustration, they construct a fractional (F) GARCH model that resembles the FIGARCH model in the sense that both models incorporate the Taylor coefficients of the function $\pi(z) = (1 - z)^d$. They show that their results are applicable to the FIGARCH model. However, their proof entails the following assumption.

NN. The searching region for the maximization of the quasi-log likelihood function contains only parameters that give no negative coefficient in the ARCH$(\infty)$ representation.

The validity of NN is non-trivial for the FIGARCH models with orders $(p, q) \neq (0, 0)$, see Conrad and Haag (2006). To apply the results of Robinson and Zaffaroni (2006), a compact searching region $\Theta$ satisfying the assumption NN has to be constructed. In this note, we show under certain conditions that the assumptions A-H of Robinson and Zaffaroni (2006) are fulfilled for $\theta \in \Theta$. In this way, the asymptotic behavior of the QMLE of the FIGARCH models within $\Theta$ can then be directly established. It should be noted that due to the difficulties of explicitly expressing $\Theta$, in practice, we have to search for the stationary points of the quasi-log likelihood function globally. The link between these stationary points and the QMLE in $\Theta$ is furnished in Proposition 2.

Throughout this paper, the assumptions A to H of Robinson and Zaffaroni (2006) are denoted by RZ-A to RZ-H. Consider the FIGARCH model,

$$X_t^2 = \sigma_t^2 \epsilon_t^2,$$
$$\sigma_t^2 = \omega + \sum_{j=1}^{\infty} \psi_j X_{t-j}^2,$$

where

$$\sum_{j=1}^{\infty} \psi_j z^j = 1 - \frac{\phi(z)}{1 - \beta(z)} (1 - z)^d,$$

$\phi(z)$ is a polynomial of order $q$ with constant term 1, $\beta(z)$ is a polynomial of order $p$ with a zero constant term, $d \in (0, 1)$ and $\omega > 0$ are non-negative real numbers.

Suppose that the data generating process is obtained from the FIGARCH model with $\theta = \theta_0$. We need assumptions A1–A3 for the data generating process.

A1. Douc et al. (2008): The coefficients $\psi_j$ of the true model satisfy

$$\sum_{j=1}^{\infty} \psi_j \log \psi_j + \mathrm{E}(\epsilon_0^2 \log(\epsilon_0^2)) \in (0, +\infty];$$

A2. For all $j = 1, 2, \ldots, \psi_j > 0$;
A3. All roots of $\phi(z)$ and $1 - \beta(z)$ lie outside the unit disc and $\phi(z)$ and $1 - \beta(z)$ are co-prime.

**Remark 1.** According to Theorem 1 and Corollary 2 in Douc et al. (2008), under A1, the true model admits a strictly stationary solution $\{X_t\}$ and its moments $\mathrm{E}|X_t|^{2\rho}$ are finite for all $\rho \in (0, 1)$. Therefore, RZ-E holds. The special case FIGARCH$(0, d, 0)$ model has been studied in Corol-

*Corresponding author.

lary 3 of Douc et al. (2008), where it was shown that there exists $0 < d^* < 1$ such that for all $d \in (d^*, 1)$, A1 holds.

**Remark 2.** A2 guarantees that $\sigma_t^2$ generated by the FIGARCH equation in the ARCH($\infty$) expression are all positive, see Conrad and Haag (2006). However, it should be noted these conditions do not necessarily require the non-negativity of the coefficients of $\beta(z)$ and $1 - \beta(z) - \phi(z)(1 - z)^d$, see Conrad and Haag (2006). For the special case FIGARCH$(1, d, 0)$, the necessary and sufficient condition for A2 is that any one of the following is satisfied.

1. $0 < \beta_1 < 1$, $d - \beta_1 \geq 0$; or
2. $-1 < \beta_1 < 0$, $2d - \sqrt{4 - 2d} \leq 2\beta_1 \geq 0$.

Suppose that the parameters to be estimated are $\theta = (\omega, d, \phi, \beta)$ and $\{X_t^2 : 1 < t \leq n\}$ are the observed values. We are interested in the asymptotic properties of the estimator obtained by maximizing the quasi log-likelihood function locally over a searching region $\Theta$ as defined in Section 2. The quasi log-likelihood function is constructed as follows. Define

$$\hat{q}_t(\theta) = \frac{X_t^2}{h_t(\theta)} + \log h_t(\theta)$$

$$q_t(\theta) = \frac{X_t^2}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta).$$

Here, $\sigma_t^2(\theta)$ is the stationary stochastic process

$$\sigma_t^2(\theta) = \omega + \sum_{j=1}^{\infty} \psi_j(\theta) X_{t-j}^2,$$

while $h_t(\theta; \{X_t^2\}_{-\infty<t\leq n})$ is a predictable stochastic process chosen to approximate the unobservable random variables $\sigma_t^2$. The process $h_t(\theta)$ is constructed as follows,

$$h_t(\theta) = \omega + \sum_{j=1}^{t-1} \psi_j(\theta) X_{t-j}^2.$$

The quasi log-likelihood function has the form

$$Q_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \hat{q}_t(\theta).$$

We show that under certain conditions, the consistency and asymptotic normality of the QMLE estimator constructed above can be established by applying Theorems 1 and 2 of Robinson and Zaffaroni (2006). To achieve that, it remains to prove the validity of the assumptions RZ-F(3) and RZ-G. One should note that RZ-A to RZ-D follow immediately from the FIGARCH$(p, d, q)$ model and RZ-H is satisfied when $d > 1/2$.

This note is organized as follows. In Section 2, a compact searching space $\Theta$ that contains the true parameter $\theta_0$ as an interior point is constructed. Two main theorems for the asymptotic behavior of the QMLE over $\Theta$ are presented. In Section 3, we show that the assumptions RZ-F(3) and RZ-G are satisfied for $\theta \in \Theta$ and thereby proofs of the theorems in Section 2 are complete. Some preliminary results used in the proofs are given in the Appendix.

## 2. MAIN RESULTS

Two theorems on consistency and asymptotic normality of the QMLE are presented in this section. Suppose that the data generating process is obtained from the FIGARCH model with $\theta = \theta_0$, satisfying assumptions A1–A3. The parameter space $\Theta$ is constructed in Proposition 1.

Introduce the following notations. Define

$$B(\beta) = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \cdots & \beta_p \\ 1 & 0 & \cdots & & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Let $B_{ij}^n$ be the $(i, j)$-th entry of the matrix $B^n$. Let $\lambda(\beta)$ be the eigenvalue of $B(\beta)$ with the largest modulus and let $p_\beta(\Theta)$ be the coordinate-mapping on the parameter space $\Theta$.

**Proposition 1.** *There exists a parameter space* $\Theta \subset R^{p+q+2}$ *satisfying conditions B1 to B6 as follows.*

*B1.* $\Theta$ *is compact and contains* $\theta_0$ *as an interior point;*
*B2.* $0 < \omega^L < \omega < \omega^U$ *and* $1/2 < d^L < d < d^U$;
*B3. There exist constants* $0 < \lambda_1^L < \lambda_1^U < 1$ *such that for all* $\beta \in p_\beta(\Theta)$, $\lambda_1^L \leq |\lambda(\beta)| \leq \lambda_1^U$;
*B4. There exist constants* $0 < \lambda_2^L < \lambda_2^U < 1$ *such that for all* $\phi \in p_\phi(\Theta)$, $\lambda_2^L \leq |\lambda(\phi)| \leq \lambda_2^U$;
*B5. For all* $\theta \in \Theta$,

$$K_L j^{-d-1} \leq \psi_j(\theta) \leq K_U j^{-d-1}$$

*for some constants* $0 < K_L < K_U$;
*B6. Within* $\Theta$, *the polynomials* $\phi(z)$ *and* $1 - \beta(z)$ *do not have common zeros.*

The results of Robinson and Zaffaroni (2006) can be applied to establish the asymptotic behavior of

$$\theta_n = \arg\min_{\theta \in \Theta} Q_n(\theta).$$

In practice, it would not be possible to verify if a given set $\Theta$ satisfies B1–B6 because these conditions involve the unknown parameter $\theta_0$. Instead, we have to search for the stationary points of the quasi-log likelihood function $Q_n(\theta)$ globally (in a space $\Gamma$ that is large enough to include $\Theta$). Denote the set of such stationary points by $T_n$. What remains is to furnish the link between $T_n$ and $\theta_n$. Consider the event $E = \{\exists N \text{ such that } \forall n > N,\ \theta_n \in T_n\}$. It can be shown that $P(E) = 1$ (see Proposition 2), which allows us to establish the asymptotic properties of $T_n$ from the results of $\theta_n$.

**Proposition 2.** $P(E) = P\{\exists N$ such that $\forall n > N,\ \theta_n \in T_n\} = 1$.

*Proof.* Let $\delta_0$ be a positive constant such that the ball $S_{\delta_0}(\theta_0)$ lies inside $\Theta$ (existence of $\delta_0$ is guaranteed by Proposition 1). Recall the following two facts.

1. By Theorem 1, $\theta_n$ is strongly consistent, i.e., $P\{\theta_n \to \theta_0\} = 1$.
2. $\{\theta_n \in S_{\delta_0}(\theta_0)\}$ is a subset of $\{\theta_n \in T_n\}$. To see this, note that when $\theta_n \in S_{\delta_0}(\theta_0)$, since $S_{\delta_0}(\theta_0)$ is an open subset of the compact parameter space $\Theta$, $\theta_n$ is not maximized on the boundary of $\Theta$. Therefore, $\theta_n \in T_n$.

Combining these two facts, the required result can be deduced as follows.

$$\begin{aligned} P(E) &= P\{\exists N \text{ such that } \forall n > N, \theta_n \in T_n\} \\ &\geq P\{\exists N \text{ such that } \forall n > N, \text{ we have } \theta_n \in S_{\delta_0}(\theta^0)\} \\ &\geq P\{\forall \delta > 0, \exists N \text{ such that } \forall n > N, \\ &\qquad \text{we have } \theta_n \in S_{\delta}(\theta^0)\} \\ &= P\{\theta_n \to \theta_0\} \\ &= 1. \end{aligned}$$ □

With RZ-F(3) and RZ-G being satisfied, the following results follow immediately from Robinson and Zaffaroni (2006).

**Corollary 1.** *If A1–A3 and RZ-A($\alpha$) are satisfied by some $\alpha > 2$, then for $\Theta$ prescribed in Proposition 1, then the QMLE $\theta_n = \arg\min_\Theta Q_n(\theta)$ is strongly consistent, i.e., with probability one, $\theta_n \to \theta_0$.*

Let $\nabla$ and $\nabla^2$ be the gradient operator and the Hessian matrix respectively. Define

$$G_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}[\nabla \hat{q}_t(\theta)][\nabla \hat{q}_t(\theta)]^T$$

and

$$H_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\nabla^2 \hat{q}_t(\theta).$$

**Corollary 2.** *Suppose that $d > 1/2$. If A1–A3 and RZ-A(4) hold, then for $\Theta$ prescribed in Proposition 1, there exist positive definite matrices $\Omega_1$ and $\Omega_2$, such that*

$$\Omega_1 = \mathrm{E}[\nabla q_t(\theta)][\nabla q_t(\theta)]^T\ ,\ \Omega_2 = \mathrm{E}\nabla^2 q_t(\theta),$$

*and*

$$\sqrt{n}(\theta_n - \theta_0) \to^d \mathrm{N}(0, \Omega_2^{-1}\Omega_1\Omega_2^{-1}).$$

*Here, the matrix $\Omega_2^{-1}\Omega_1\Omega_2^{-1}$ can be approximated as*

$$H_n^{-1}(\theta_n)G_n(\theta_n)H_n^{-1}(\theta_n) \to^{a.s.} \Omega_2^{-1}\Omega_1\Omega_2^{-1}.$$

# 3. PROOFS

In this section, Proposition 1, RZ-F(3) and RZ-G are established. For convenience, RZ-F(3) and RZ-G are restated here.

RZ-F(3): Let $k \leq 3$, and $1 \leq i_1, \ldots, i_k \leq p+q+1$. Suppose that $\theta \in \Theta$. Consider the derivatives of $\psi_j(\theta)$ with respect to the parameters $\theta^{i_1}, \ldots, \theta^{i_k}$, with the parameter $d$ appearing $m$-times, where $m \geq 0$. For each $\eta > 0$, a constant $K > 0$ can be found such that the derivatives satisfy

$$\left|\frac{\partial^k \psi_j(\theta)}{\partial\theta_{i_1}\cdots\partial\theta_{i_k}}\right| \leq K\psi_j^{1-\eta}(\theta).$$

RZ-G: Let $r = p+q+1$. For each $\theta \in \Theta$, there exist integers

$$1 \leq j_1(\theta) < \cdots < j_r(\theta) < \infty,$$

such that

$$\mathrm{rank}\{\nabla\psi_{j_1}(\theta), \ldots, \nabla\psi_{j_r}(\theta)\} = r,$$

where $\nabla$ is the gradient operator.

*Proof of Proposition 1.* Let $\mathcal{D}_\beta = \{\beta | \lambda_1^L \leq \lambda(\beta) \leq \lambda_1^U\}$. Define

$$\Theta_0 = \mathcal{D}_\beta \times \mathcal{D}_\phi \times [d^L, d^U] \times [\omega^L, \omega^U].$$

By Theorem 2.1 of Hosking (1981),

$$\lim_{j\to\infty} j^{d+1}\psi_j(\theta) = \frac{-\phi(1)}{\Gamma(d)} \cdot \frac{1}{1-\beta(1)} = K(\theta).$$

Moreover, from Lemma 1 in Appendix A, the convergence is uniform over $\Theta_0$. When $0 < \delta < \inf_{\theta\in\Theta_0} K(\theta)$, an integer $N$ can always be found so that for $j > N$ and $\theta \in \Theta_0$,

$$[K(\theta)+\delta]j^{-d-1} > \psi_j(\theta) > [K(\theta)-\delta]j^{-d-1} > 0.$$

For $j \leq N$, since $\psi_j(\theta_0) > 0$ (by assumption A2), we can find $\delta_j$ such that $0 < \delta_j < \psi_j(\theta_0)$. Let

$$\Theta_j = \{\theta | \psi_j(\theta_0) - \delta_j < \psi_j(\theta) < \psi_j(\theta_0) + \delta_j\}.$$

By the continuity of the functions $\psi_j(\theta)$, the sets $\Theta_j$ are open. From Lemma 1 in Appendix A, the set $\Theta_1 = \overline{\cap_{j=0}^{N}\Theta_j}$ is compact and contains an open set in which $\theta_0$ is an interior point. In addition, $\Theta_1$ satisfies B1–B5. With assumption A3, it can be checked from continuity arguments that there exists a neighbourhood of $\theta_0$ such that B6 holds. Let $N(\theta_0)$ be such a neighbourhood. Then the conditions B1–B6 are fulfilled for $\Theta = \overline{\Theta_1 \cap N(\theta_0)}$. The constants $K_L$ and $K_U$ prescribed in B5 are defined as follows.

$$K_L = \min_{\theta\in\Theta}\left\{\frac{-\phi(1)}{\Gamma(-d)[1-\beta(1)]} - \delta, \frac{\psi_1(\theta)}{1^{-d-1}}, \ldots, \frac{\psi_N(\theta)}{N^{-d-1}}\right\},$$

$$K_U = \max_{\theta\in\Theta}\left\{\frac{-\phi(1)}{\Gamma(-d)[1-\beta(1)]}+\delta, \frac{\psi_1(\theta)}{1^{-d-1}}, \dots, \frac{\psi_N(\theta)}{N^{-d-1}}\right\}.$$

□

*Proof of RZ-F(3).* From condition B5, it is sufficient to show that a constant $K > 0$ can be found such that for $j > 1$,

$$\left|\frac{\partial^k \psi_j(\theta)}{\partial\theta_{i_1}\cdots\partial\theta_{i_k}}\right| \le K j^{-d-1}(\log j)^m.$$

By Lemma 2 in Appendix A, there exists a continuous function $K(\theta)$ such that for any given $\delta < \inf_{\theta\in\Theta} K(\theta)$, an integer $N$ can be found so that for all $j > N$ and $\theta \in \Theta$,

$$\left|\frac{\partial^k \psi(\theta)}{\partial\theta_{i_1}\cdots\partial\theta_{i_k}}\right| \le \left\{\max_{\theta\in\Theta}|K(\theta)+\delta|\right\} j^{-d-1}\log^m j.$$

The constant $K$ chosen below fulfills the goal.

$$K = \max\left\{\delta + \sup_{\theta\in\Theta} K(\theta), \frac{\psi_2(\theta)}{2^{-d-1}\log^m 2}, \dots, \frac{\psi_N(\theta)}{N^{-d-1}\log^m N}\right\}.$$

□

*Proof of RZ-G.* First, we show that a sufficient condition for RZ-G is that $\lambda^T\nabla_\theta\psi(z;\theta) = 0$ implies $\lambda = 0$, where $\lambda$ is an $r$-dimensional vector.

To check this, assume by contradiction that RZ-G does not hold. To simplify notation, $\psi(\theta)$ is written as $\psi$ for the rest of the proof. Then, for all $\theta \in \Theta$ and $1 \le j_1 < \cdots < j_r < \infty$, we have

$$\operatorname{rank}\{\nabla\psi_{j_1}, \dots, \nabla\psi_{j_r}\} < r.$$

Choose $1 \le j_1 < \cdots < j_r < \infty$ such that

$$\operatorname{rank}\{\nabla\psi_{j_1}, \dots, \nabla\psi_{j_r}\}$$

is maximum. Denote the maximum rank by $r^*$. Select a maximal linear independent subset from $\{j_1, \dots, j_r\}$ and denote it by $\{j_1^*, \dots, j_{r^*}^*\}$. Then, for any $j$ not belonging to such a subset, $\nabla\psi_j$ can be expressed as a linear combination of $\nabla\psi_{j_1^*}, \dots, \nabla\psi_{j_{r^*}^*}$. On the other hand, there exist non-trivial solutions to the system of linear equations

$$\lambda^T\nabla\psi_{j_1^*} = \cdots = \lambda^T\nabla\psi_{j_{r^*}^*} = 0,$$

since $r^* < r$. Therefore, $\lambda^T\nabla\psi_j = 0$ for all $j = 1, 2, \dots,$ and consequently, $\lambda^T\nabla_\theta\psi(z;\theta) = 0$. Let $\lambda = (\lambda^\phi, \lambda^\beta, \lambda^d)$, $\lambda^\phi(z)$, and $\lambda^\beta(z)$ be two polynomials with zero constant term. Suppose that $\lambda^T\nabla_\theta\psi(z;\theta) = 0$. Simple algebraic manipulations yield

$$\psi(z) = 1 - \frac{[\phi(z)+\lambda^\phi(z)+\lambda^d\phi(z)\log(1-z)](1-z)^d}{1-\beta(z)-\lambda^\beta(z)}.$$

Since the Taylor coefficients $\psi_j = O(j^{-d-1})$, $\lambda^d$ must be zero, $\psi_j = O(j^{-d-1}\log j)$. What remains is to show that

$$\frac{\phi(z)}{1-\beta(z)} = \frac{\phi(z)+\lambda^\phi(z)}{1-\beta(z)-\lambda^\beta(z)}$$

implies $\lambda^\phi(z) = \lambda^\beta(z) = 0$. By condition B6, there exists a polynomial $p(z)$ such that $\phi(z)\lambda^\phi(z) = \phi(z)p(z)$ and $1-\beta(z)-\lambda^\beta(z) = [1-\beta(z)]p(z)$. Since the orders and constant terms of $\phi$ and $\beta$ are the same as those of $\phi$ and $\beta$, we must have $p(z) = 1$. Therefore, $\lambda^\phi(z) = \lambda^\beta(z) = 0$ and we have the desired results. □

## APPENDIX A. TECHNICAL LEMMAS

The purpose of this appendix is to establish the uniform convergence of $\psi_j(\theta)$ and its derivatives over $\Theta_0$ defined in the proof of Proposition 1. Let

$$\psi(z;\beta,\phi,d) = 1 - \frac{\phi(z)(1-z)^d}{1-\beta(z)},$$
$$\pi(d) = 1-(1-z)^d.$$

Notations like $\psi_j(\theta)$ are used to represent the Taylor coefficients of the functions, e.g. $\psi(z;\theta) = \sum_{j=0}^\infty \psi_j(\theta)z^j$. Let $f : R^r \to R$ be a real-valued function of $(\theta_1, \dots, \theta_r)$. For any integers $k > 0$ and $1 \le i_1, i_2, \dots, i_k \le r$, define

$$\partial^{i_1\cdots i_k} f = \frac{\partial^k f}{\partial\theta_{i_1}\cdots\partial\theta_{i_k}}.$$

**Lemma 1.** *(1) The set $\{\beta|\lambda_1^L < \lambda(\beta) < \lambda_1^U\}$ is open.*
*(2) The set $\mathcal{D}_\beta = \{\beta|\lambda_1^L \le \lambda(\beta) \le \lambda_1^U\}$ is compact.*
*(3) Within $\mathcal{D}_\beta$, $1-\beta(1)$ is bounded above and below by some positive constants.*
*(4) Within $\mathcal{D}_\beta$, for all $\delta > 0$, there exists a constant $K > 0$ which does not depend on the choice of $\beta$, such that for $j = 1, 2, \dots,$ we have*

$$|B_{11}^j(\beta)| \le K(|\lambda^U| + \delta)^j.$$

*Proof.* (1) This can be seen from the continuity of spectral norm $\lambda(\beta)$.
(2) By the relationship between roots and coefficients, $|\beta_i|$ is bounded by $C_i^p|\lambda_1^U|^i$.
(3) To see this, consider the characteristic equation of $B(\beta)$, which is

$$f(\lambda) = \lambda^p - \beta_1\lambda^{p-1} - \cdots - \beta_p = 0.$$

Let $\lambda_1, \dots, \lambda_p$ be the roots of the above equation, then

$$1-\beta(z) = f(1) = (1-\lambda_1 z)\cdots(1-\lambda_p z).$$

Simple calculations yield that

$$0 < (1-|\lambda_1^U|\cdot|z|)^p < 1-\beta(1) < (1+|\lambda_1^U|\cdot|z|)^p.$$

(4) Let $R(\beta) = (|\lambda_1| + \delta)^{-1}$. Applying Cauchy's estimation (see Theorem 10.26 in Rudin, 1987), we have an upper bound,

$$B_{11}^j(\beta) \le \frac{1}{R^j(\beta)} \cdot \frac{1}{(1 - R(\beta)|\lambda_1(\beta)|)^p} = \frac{(|\lambda_1(\beta)| + \delta)^{p+j}}{\delta^p}.$$

Choosing $K = (\frac{\lambda^U + \delta}{\delta})^p$ fulfills the need. □

**Lemma 2.** *Consider the derivatives of $\psi_j(\theta)$ with respect to the parameters $\theta^{i_1}, \ldots, \theta^{i_k}$, with $d$ appearing $m$-times, where $m \ge 0$. Then the derivatives satisfy*

$$\lim_{j\to\infty} j^{d+1} \log^{-m} j \partial^{i_1\cdots i_k} \psi_j(\theta) = K^{i_1\cdots i_k}(\theta)$$

*and these convergences are uniform in $\Theta$.*

*Proof for the special case $k = 0$, $\phi = 0$, $\beta = 0$.* From Theorem 2.1 of Hosking (1981), we have

$$\lim_{j\to\infty} j^{d+1} \pi_j(d) = \frac{-1}{\Gamma(-d)}.$$

To establish the uniform convergence in $[d^L, d^U]$, we give bounds for the following expression,

$$\begin{aligned}&\log \pi_j + (d+1)\log j \\ &\quad = \log d + \sum_{k=2}^{j} \left\{ \log \frac{k-1-d}{k} + (d+1) \log \frac{k}{k-1} \right\}.\end{aligned}$$

Below, we show that for all $j = 1, 2, \ldots$, the term in the brace bracket is monotonic decreasing. The derivative of these terms with respect to $d$ are

$$\frac{-1}{j-1-d} + \log \frac{j}{j-1} < \frac{-1}{j-1-d} + \frac{1}{j-1} < 0.$$

Here, the inequality $\log x < x - 1$ is used. Then,

$$\begin{aligned}&\log \pi_j + (d+1)\log j \\ &\quad \le \log d^U + \sum_{k=2}^{j} \left\{ \log \frac{k-1-d^L}{k} + (d^L+1) \log \frac{k}{k-1} \right\} \\ &\quad \to \log \left[ \frac{d^U}{d^L} \frac{1}{\Gamma(-d^L)} \right],\end{aligned}$$

and

$$\begin{aligned}&\log \pi_j + (d+1)\log j \\ &\quad \ge \log d^L + \sum_{k=2}^{j} \left\{ \log \frac{k-1-d^U}{k} + (d^U+1) \log \frac{k}{k-1} \right\} \\ &\quad \to \log \left[ \frac{d^L}{d^U} \frac{1}{\Gamma(-d^U)} \right].\end{aligned}$$

These yield the required results. □

*Proof for the special case $k \ge 1$, $\phi = 0$, $\beta = 0$.* Here, only the case that $k = 1$ is considered as the result for $k > 1$ can be shown inductively in a similar manner. Consider the recursive relationship

$$\pi_1 = d \text{ and } \pi_j = \frac{j-1-d}{j} \pi_{j-1} \text{ for } j = 2, 3, \ldots.$$

It can be shown by induction that

$$\frac{\partial \pi_j(d)}{\partial d} = \left( \frac{1}{d} - \frac{1}{1-d} - \cdots - \frac{1}{j-1-d} \right) \pi_j.$$

It suffices to establish the uniform convergence of

$$\frac{1}{\log j} \left( \frac{1}{d} - \frac{1}{1-d} - \cdots - \frac{1}{j-1-d} \right)$$

for $d \in [d^L, d^U]$ as $j \to \infty$. Below, an upper bound and a lower bound for the quantity

$$\frac{1}{\log j} \left( \frac{1}{1-d} + \cdots + \frac{1}{j-1-d} \right)$$

are given. By the inequality

$$\frac{1}{j} < \log \frac{j}{j-1} < \frac{1}{j-1},$$

we have,

$$\begin{aligned} 1 &= \frac{1}{\log j} \left\{ \log \frac{j}{j-1} + \cdots + \log \frac{2}{1} \right\} \\ &< \frac{1}{\log j} \left\{ \frac{1}{j-1} + \cdots + 1 \right\} \\ &< \frac{1}{\log j} \left\{ \frac{1}{j-1-d} + \cdots + \frac{1}{1-d} \right\} \\ &< \frac{1}{\log j} \left\{ \frac{1}{j-2} + \cdots + \frac{1}{1} + \frac{1}{1-d^U} \right\} \\ &< \frac{1}{\log j} \left\{ \log \frac{j-2}{j-3} + \cdots + \log \frac{2}{1} + 1 + \frac{1}{1-d^U} \right\} \\ &= \frac{1}{\log j} \left\{ \log(j-2) + 1 + \frac{1}{1-d^U} \right\}. \end{aligned}$$

Consequently,

$$\begin{aligned} &\frac{1}{d^U \log j} - \frac{1}{\log j} \left\{ \log(j-2) + 1 + \frac{1}{1-d^U} \right\} \\ &\quad < \frac{1}{\log j} \left\{ \frac{1}{d} - \frac{1}{1-d} - \cdots - \frac{1}{j-1-d} \right\} \\ &\quad < -1 + \frac{1}{d^L \log j}. \end{aligned}$$

Here, both the upper and lower bounds converge to $-1$. □

*Proof for the general cases.* Below, we show that if the convergence of

$$\lim_{j\to\infty} j^{d+1}(\log j)^m \cdot \varsigma_j(\theta) = K(\theta) > 0$$

is uniform over a region $\Theta$ and $K(\theta)$ is bounded in $\Theta$, then

$$\lim_{j\to\infty} j^{d+1}(\log j)^m \sum_{i=1}^{j} B^{i-1}(\theta)\varsigma_{j-i}(\theta) = \frac{K(\theta)}{1-\beta(1)}$$

uniformly over $\Theta$.

For any integers $M$, $N$ and $j > M+N$, consider

$$\begin{aligned}\varsigma_j(\theta) &+ B_{11}\varsigma_{j-1}(\theta) + \cdots + B_{11}^{j-1}\varsigma_1(\theta)\\ &= \sum_{i=0}^{M} + \sum_{i=M+1}^{j-N-1} + \sum_{i=j-N}^{j-1} B_{11}^{j-i}\varsigma_i(\theta)\\ &= S_1(\theta) + S_2(\theta) + S_3(\theta).\end{aligned}$$

By Lemma 1, the first sum

$$\begin{aligned}j^{d+1}&(\log j)^m S_1(\theta)\\ &= j^{d+1}(\log j)^m \sum_{i=1}^{M} B_{11}^{j-i}(\theta)\varsigma_i(\theta)\\ &\le K j^{d^U+1}(\log j)^m \sum_{i=1}^{M}(|\lambda^U| + \delta)^j \max_{\theta\in\Theta}\varsigma_i(\theta)\\ &\to 0.\end{aligned}$$

Let $M$ be chosen so that for $j > M$,

$$(K(\theta)-\delta)j^{-d-1} < \varsigma_i(\theta) < (K(\theta)+\delta)j^{-d-1}.$$

By Lemma 1 and the fact that $\frac{j}{j-i} < i+1$ and $\frac{(\log(j-i))^m}{(\log j)^m} < 1$ when $j > i+1$, we have for sufficiently large $N$ and $j > M+N$,

$$\begin{aligned}|j^{d+1}&(\log j)^m S_2(\theta)|\\ &\le (K(\theta)+\delta)\sum_{i=N+1}^{j-M-1}\left|B_{11}^i(\theta)\right|\left(\frac{j}{j-i}\right)^{d+1}\cdot\frac{(\log(j-i))^m}{(\log j)^m}\\ &\le (K(\theta)+\delta)\sum_{i=N+1}^{j-M-1}\left|B_{11}^i(\theta)\right|(i+1)^{d+1}\\ &\le (K(\theta)+\delta)\sum_{i=N+1}^{\infty}\left|B_{11}^i(\theta)\right|(i+1)^{d+1}\\ &\le K(K(\theta)+\delta)\sum_{i=N+1}^{\infty}(\lambda^U+\delta)^i(i+1)^{d^U+1},\end{aligned}$$

which is arbitrarily small. For the third term,

$$\begin{aligned}&\left|j^{d+1}(\log j)^m S_3(\theta) - K(\theta)\sum_{i=1}^{\infty}B_{11}^i(\theta)\right|\\ &\quad\le |j^{d+1}(\log j)^m S_3(\theta) - K(\theta)(1 + B_{11}(\theta) + \cdots + B_{11}^N(\theta))|\\ &\qquad + K(\theta)\sum_{i=1}^{\infty}\left|B_{11}^i(\theta)\right|.\end{aligned}$$

The last term can be bounded by

$$K(\theta)\sum_{i=1}^{\infty}\left|B_{11}^i(\theta)\right| \le K\sum_{i=1}^{\infty}\left|\lambda^U+\delta\right|^i,$$

which is arbitrarily small. An upper bound for the first term is

$$\begin{aligned}&\sum_{i=0}^{N}|B_{11}^i(\theta)|\cdot\left|\frac{j^{d+1}(\log(j-i))^m}{(j-i)^{d+1}(\log j)^m}(j-i)^{d+1}\right.\\ &\qquad\left.\times(\log(j-i))^m\varsigma_{j-i} - K(\theta)\right|\\ &\quad\le K\sum_{i=0}^{N}|\lambda^U+\delta|^i\cdot\left|\frac{j^{d+1}(\log(j-i))^m}{(j-i)^{d+1}(\log j)^m}(j-i)^{d+1}\right.\\ &\qquad\left.\times(\log(j-i))^m\varsigma_{j-i} - K(\theta)\right|,\end{aligned}$$

which converges to zero uniformly in $\Theta$ as $j\to\infty$. As a result,

$$\lim_{j\to\infty} j^{d+1}\sum_{i=1}^{j}B^{i-1}(\theta)\varsigma_{j-i}(\theta) = K(\theta)\sum_{j=0}^{\infty}B^j(\theta) = \frac{K(\theta)}{1-\beta(1)},$$

and the convergence is uniform. □

## ACKNOWLEDGEMENTS

We would like to thank the Editors, an Associate Editor and two anonymous referees for helpful comments and suggestions, which lead to an improved version of this paper. Research supported in part by the General Research Fund of the Research Grants Council of Hong Kong under grants 400408 and 400410 and the Collaborative Research Fund of the Research Grants Council of Hong Kong under grant CityU8/CRF/09.

*Received 24 July 2010*

## REFERENCES

Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **74**, 3–30. MR1409033

Bollerslev, T. and Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics* **73**, 151–184.

Conrad, C. and Haag, B. R. (2006). Inequality constraints in the fractionally integrated GARCH model, *Journal of Financial Econometrics* **4**, 413–449.

Douc, R., Roueff, F. and Soulier, P. (2008). On the existence of some ARCH($\infty$) processes. *Stochastic Processes and their Applications* **118**, 755–761. MR2411519

Hosking, J. M. R. (1981). Fractional differencing. *Biometrika* **68**, 165–176. MR0614953

Lee, S. W. and Hansen, B. E. (1994). Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory* **10**, 29–52. MR1279689

Mikosch, T. and Stărică, C. (2002). Long-range dependence effects and ARCH modeling. In Doukhan, P., Oppenheim, G. and Taqqu, M. S. (eds.), *Theory and Applications of Long-Range Dependence*, 439–459. Birkhauser, Boston. MR1957503

Robinson, P. M. and Zaffaroni, P. (2006). Pseudo-maximum likelihood estimation of ARCH($\infty$) models. *Annals of Statistics* **34,** 1049–1074. MR2278351

Rudin, W. (1987). *Real and Complex Analysis*, 3rd ed. McGraw-Hill, New York. MR0924157

Ngai Hang Chan
Department of Statistics
The Chinese University of Hong Kong
Shatin, NT, Hong Kong
E-mail address: nhchan@sta.cuhk.edu.hk

Chi Tim Ng
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hung Hum, Kowloon, Hong Kong
E-mail address: machitim@inet.polyu.edu.hk