Fitting a \( p \)th order parametric generalized linear autoregressive multiplicative error model\(^1\)

N. Balakrishna  H. L. Koul  M. Ossiander  L. Sakhanenko\(^2\)

Cochin University of Science and Technology, Oregon State University and Michigan State University

Abstract

This paper is concerned with the problem of fitting a generalized linear model to the conditional mean function of multiplicative error time series models. These models are particularly suited to model nonnegative time series such as the duration between trades at a stock exchange and volume transactions. The proposed test, based on a marked residual empirical process whose marks are suitably defined residuals and which jumps at the estimated indices, is shown to be asymptotically distribution free.

1 Introduction

In the last two decades the family of multiplicative error models has attracted considerable attention for modeling nonnegative time series. Engle and Russell (1998) and Engle (2002) used these models for analyzing financial durations and trading volume of orders, respectively. For many other applications see Pacarur (2008), Hautsch (2012), and the references therein.

To proceed further, let \( Y_i, i \in Z := \{0, \pm1, \pm2, \ldots \} \), be a discrete time nonnegative stationary process with \( EY_0^2 < \infty \), and let \( \mathcal{H}_{i-1} \) denote the information available up to time \( i-1 \) for forecasting \( Y_i \). A multiplicative error model (MEM) takes the form

\[
Y_i = E[Y_i \mid \mathcal{H}_{i-1}] \varepsilon_i, \quad i \in Z,
\]

where \( \varepsilon_i, i \in Z \) are independent and identically distributed (i.i.d.) non-negative error random variables (r.v.’s) with \( E(\varepsilon_0) = 1, E(\varepsilon_0^2) < \infty \). Moreover, \( \varepsilon_i \) is assumed to be independent of the past information \( \mathcal{H}_{i-1} \), for all \( i \in Z \).

Pacarur (2008) and Hautsch (2012) discuss many parametric specifications for the conditional mean function \( E(Y_i \mid \mathcal{H}_{i-1}) \). The problem of fitting a given parametric MEM model is of practical importance since knowing the right parametric model can lead to optimal inference. Koul, Perera and Sillvaphule (2013) (KIS) proposed a lack of fit test for fitting a parametric model to this conditional mean function when the underlying time series is

\(^1\)Keywords & phrases: Martingale transform, AR conditional duration models.

\(^2\)This author’s research was supported in part by the NSF grant DMS–1612867
Markovian, i.e., when $E[Y_i \mid H_{i-1}] = \tau(Y_{i-1})$, $i \in \mathbb{Z}$, a.s., for some positive measurable function $\tau$ defined on $R^+ := [0, \infty)$. To describe the KIS test a bit more precisely, let $q \geq 1$ be a given positive integer, $\Omega \subset \mathbb{R}^q$, and $\{\psi(y, \omega), y \geq 0, \omega \in \Omega\}$ be a given family of positive functions. Consider the problem of testing

$$ H_0 : \tau(y) = \psi(y, \omega), \quad \text{for all } y \geq 0, \text{ and some } \omega \in \Omega, \text{ versus} $$

$$ H_1 : H_0 \text{ is not true.} $$

Let

$$ U_n(y, \omega) = n^{-1/2} \sum_{i=1}^{n} \left( \frac{Y_i}{\psi(Y_{i-1}, \omega)} - 1 \right) I(Y_{i-1} \leq y), \quad y \geq 0, \omega \in \Omega, $$

The test of $H_0$ proposed in KIS is based on the process $U_n(\cdot, \hat{\omega})$, where $\hat{\omega}$ is a $n^{1/2}$-consistent estimator of $\omega$ under $H_0$. Asymptotic null distribution of this process depends on the null model parameter $\omega_0$ under $H_0$ and its estimator in a complicated fashion. In KIS, an analog of the martingale transform of Stute, Thies and Zhu (1998) (STZ) of this process was shown to converge weakly, under $H_0$, to a time transform of Brownian motion. In particular the test based on the supremum of the absolute value of this transform is asymptotically distribution free, consistent against a large class of nonparametric alternatives, and has non-trivial asymptotic power against a large class of $n^{-1/2}$-local alternatives. In a finite sample simulation study this test was shown to have very desirable empirical level property and much larger empirical power, compared to several other model specification tests including the Ljung-Box $Q$ test and Lagrange multiplier type test of Meitz and Teräsvirta (2006).

Admittedly restricting the dependence of the conditional mean function to one lag variable has limited applications. Consider the situation where

$$ E[Y_i \mid H_{i-1}] = \gamma(Y_{i-1}, \cdots, Y_{i-p}), \quad i \in \mathbb{Z}, $$

(1.2)

for a known positive integer $p$ and a positive function $\gamma$ defined on $[0, \infty)^p$. In practice there are several time series where the conditional mean function $\gamma$ depends on the previous $p$ lags via a linear combination. Then the question of interest is which one of these models fit the given time series.

More precisely, let $\Theta_1 \subset \mathbb{R}^p, \Theta := [0, \infty) \times \Theta_1$, and $\varphi$ be a known positive link function defined on $\mathbb{R}$. Let $x'$ denote the transpose of an Euclidean vector $x$. The problem of interest is to test the hypotheses

$$ H_0 : \gamma(y) = \theta_0 + \varphi(\theta'y), \quad \text{for all } y \in [0, \infty)^p, \text{ and for some } (\theta_0, \theta')' \in \Theta, \quad \text{versus}$$

$$ H_1 : H_0 \text{ is not true.} $$
A version of the $U_n$-process suitable for this testing problem is

$$\mathcal{V}_n(y, \delta) = n^{-1/2} \sum_{i=1}^n \left( \frac{Y_i}{\vartheta_0 + \varphi(\vartheta'Y_{i-1})} - 1 \right) I\{\vartheta'Y_{i-1} \leq y\}, \quad y \geq 0,$$

where $\vartheta := (\vartheta_1, \ldots, \vartheta_p)'$, $\delta := (\vartheta_0, \vartheta')' \in \Theta$. Let $\delta_0$ denote the parameter vector for which $H_0$ holds, i.e., $\delta_0 := (\theta_0, \theta')'$. The proposed tests of $H_0$ will be based on the process $\mathcal{V}_n(\cdot, \hat{\delta})$, where $\hat{\delta}$ is a $n^{1/2}$-consistent estimator of $\delta_0$, under $H_0$. The linear combinations $\hat{\theta}'Y_{i-1}$ are known as estimated indices and $\mathcal{V}_n(\cdot, \hat{\delta})$ is a marked residual empirical process with the marks $\hat{\eta}_i := Y_i/[\hat{\theta}_0 + \varphi(\hat{\theta}'Y_{i-1})] - 1$ jumping at the indices $\hat{\theta}'Y_{i-1}$.

The asymptotic null distribution of this process under $H_0$ is discussed in the sub-section 2.1. An analog of the martingale transform of STZ test is given in the sub-section 2.2. Several proofs are relegated to section 3. In the sequel, $\| \cdot \|$ denotes the Euclidean norm, all limits are taken as $n \to \infty$, and $u_p(1)$ denotes a sequence of stochastic processes that converges to zero uniformly over its time domain, in probability.

## 2 Main Results

In this section we derive the asymptotic expansion of the process $\mathcal{V}_n(\cdot, \hat{\delta})$ under $H_0$ and describe a martingale type transformation that converges weakly to the standard Brownian motion on $[0, \infty)$.

### 2.1 Asymptotic linearity of $\mathcal{V}$ under $H_0$

Here we derive an asymptotic linearity result for the process $\mathcal{V}_n(x, \hat{\delta})$. To begin with we shall state the needed assumptions.

(F) The error variable $\varepsilon_0$ is positive, $E\varepsilon_0 = 1$, $E\varepsilon_0^2 < \infty$.

(G) The distribution function (d.f.) $G$ of $\theta'Y_0$ is continuous, where $\theta$ is as in $H_0$.

($) \ E\varphi^2(\theta'Y_0) < \infty$, and $\varphi$ is continuously differentiable with the derivative $\dot{\varphi}$ satisfying $E\|Y_0\dot{\varphi}(\theta'Y_0)\|^2 < \infty$, and for every $0 < c < \infty$,

$$\sup_{1 \leq i \leq n, \sqrt{n}\|\theta - \hat{\theta}\| \leq c} \sqrt{n}|\varphi(\theta'Y_{i-1}) - \varphi(\theta'Y_{i-1}) - (\theta - \hat{\theta})'Y_{i-1}\dot{\varphi}(\theta'Y_{i-1})| = o_p(1).$$

(C) There exist estimators $\hat{\theta}_0 > 0$ and $\hat{\theta} \in \mathbb{R}^p$ such that under $H_0$, $n^{1/2} \left( |\hat{\theta}_0 - \theta_0| + \|\hat{\theta} - \theta\| \right) = O_p(1)$.

Let $\delta := (\hat{\theta}_0, \hat{\theta}')'$. The derivation of the asymptotic linearity of the process $\mathcal{V}_n(\cdot, \hat{\delta})$ is facilitated by the following lemma of a general interest. Consider the following assumptions.
(E) The nonnegative process $Y_j, j \in \mathbb{Z}$ is stationary and ergodic. The r.v.’s $\eta_i, i \in \mathbb{Z}$ are i.i.d., and $\eta_i$ is independent of $Y_j, j \leq i - 1$, for all $i \in \mathbb{Z}$.

(g) $g(x, z)$ is a measurable function of $(x, z) \in [0, \infty)^p \times \mathbb{R}$ such that $Eg^2(Y_0, \eta_1) < \infty$.

Let $\mathcal{F}_i := \sigma\text{-field}\{Y_j, j \leq i\}, j \in \mathbb{Z}$, and define, for $s \in \mathbb{R}^p$ and $y \geq 0$,

$$V_n(y, s) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [g(Y_{i-1}, \eta_i) - E\{g(Y_{i-1}, \eta_i)|\mathcal{F}_{i-1}\}] I((\theta + n^{-1/2}s)'Y_{i-1} \leq y).$$

**Lemma 2.1** Assume that the processes $\eta_j, Y_j, j \in \mathbb{Z}$ satisfies assumption (E) with the d.f. $G$ of $\theta'Y_0$ satisfying condition (G), and that (g) holds. Then for every $c \in [0, \infty)$,

$$\sup_{y \geq 0, \|s\| \leq c} |V_n(y, s) - V_n(y, 0)| = o_P(1).$$

The proof of this lemma is given in the last section. We shall now use this to derive the limiting distribution of the process $V_n(y, \hat{\delta})$. To begin with consider the process

$$T_n(y, \vartheta) := n^{-1/2} \sum_{i=1}^{n} \left[ \frac{Y_i}{\theta_0 + \varphi(\theta'Y_{i-1})} - 1 \right] I(\vartheta'Y_{i-1} \leq y), \quad y \geq 0, \vartheta \in \mathbb{R}^p.$$

Note that $T_n(y, \theta) = V_n(y, \delta_0)$. We have the following corollary.

**Corollary 2.1** Assume that the model (1.1), (1.2), and $\mathcal{H}_0$ hold, and that the given time series $Y_j, j \in \mathbb{Z}$ is stationary and ergodic. In addition, suppose that the assumptions (F), (G) and (C) hold. Then

$$\sup_{y \geq 0} |T_n(y, \hat{\vartheta}) - V_n(y, \delta_0)| = o_P(1).$$

**Proof.** In view of the assumption (C), it suffices to prove that for every $0 < c < \infty$,

$$\sup_{y \geq 0, \|s\| \leq c} |T_n(y, \theta + n^{-1/2}s) - T_n(y, \theta)| = o_p(1).$$

But this follows from Lemma 2.1 upon taking

$$\eta_i \equiv \left[ Y_i / (\theta_0 + \varphi(\theta'Y_{i-1})) \right] - 1 \equiv \varepsilon_i - 1,$$

and $g(x, z) \equiv z$ in there, because, in this case $\eta_i$ is independent of $\mathcal{F}_{i-1}$ for every $i \in \mathbb{Z}$, and $E\{g(Y_{i-1}, \eta_i)|\mathcal{F}_{i-1}\} \equiv E(\eta_i) \equiv 0$, and $V_n(y, s) \equiv T_n(y, \theta + n^{-1/2}s)$.

Next, we state a weak convergence result for

$$\mathcal{V}_n(y, \delta_0) = n^{-1/2} \sum_{i=1}^{n} \eta_i I(\theta'Y_{i-1} \leq y),$$
where \( \eta_i \)'s are as in (2.2). By the classical CLT, \( \mathcal{V}_n(\infty, \delta_0) = n^{-1/2} \sum_{i=1}^{n} \eta_i = O_P(1) \).

Recall \( G \) is the d.f. of \( \theta' Y_0 \). Let \( \eta \) be a copy of \( \eta_i \) and \( \sigma^2 = \text{Var}(\eta) \). Under \( \mathcal{H}_0 \),

\[
\text{Cov}(\mathcal{V}_n(y, \delta_0), \mathcal{V}_n(z, \delta_0)) = \sigma^2 G(y \wedge z) = \text{Cov}(B(A(y)), B(A(z))), \quad y, z \geq 0,
\]

where \( B \) is the standard Brownian motion on \([0, \infty)\). The following lemma proves the weak convergence of the \( \mathcal{V}_n(\cdot, \delta_0) \) to the time transformed Brownian motion \( B(AG) \). Its proof appears in the last section.

**Lemma 2.2** Suppose (1.1), (1.2), \( \mathcal{H}_0, \ (G) \) hold. Then

\[
\text{V}_n(\cdot, \delta_0) \rightarrow_D \sigma B \circ A \quad \text{in} \ D[0, \infty] \quad \text{and uniform metric}.
\]

Next, consider \( \nu_n(\cdot, \delta) \). To simplify the exposition, we let

\[
h(\delta, y) := \vartheta_0 + \varphi(\vartheta' y), \quad \delta = (\vartheta_0, \vartheta)' \in \Theta, \ y \in \mathbb{R}^p.
\]

Note that \( \dot{h}(\delta, y) := \partial h(\delta, y)/\partial \delta = (1, y' \varphi'(\vartheta' y)') \). To state the main result about \( \nu_n(\cdot, \delta) \) we need to define

\[
\nu(y, \delta) := E\{\dot{h}(\delta, Y_0) I(\vartheta' Y_0 \leq y)\}, \quad \delta = (\vartheta_0, \vartheta)' \in \Theta, \ y \geq 0.
\]

Condition \( \varphi \) implies that \( E\|\dot{h}(\vartheta_0, Y_0)\|^2 < \infty \). This and the fact that \( \inf_y h(\vartheta_0, y) \geq \theta_0 > 0 \) imply that

\[
\sup_{y \geq 0} \|\nu(y, \delta_0)\|^2 \leq E\{\|\dot{h}(\vartheta_0, Y_0)\|/h(\vartheta_0, Y_0)\} < \infty.
\]

We have the following theorem.

**Theorem 2.1** Suppose (1.1), (1.2), \( \mathcal{H}_0, \ (F) , (G), \ (\varphi) \) and (C) hold. Then

\[
\mathcal{V}_n(y, \delta) = \mathcal{V}_n(y, \delta_0) - n^{1/2}(\delta - \delta_0)' \nu(y, \delta_0) + u_p(1).
\]

**Proof.** Recall that \( \eta_i \equiv [Y_i/h(\delta_0, Y_{i-1})] - 1 \). Consider the following decomposition.

\[
\mathcal{V}_n(y, \delta) = n^{-1/2} \sum_{i=1}^{n} \left[ \frac{Y_i}{h(\delta, Y_{i-1})} - 1 \right] I(\vartheta' Y_{i-1} \leq y)
\]

\[
= T_n(y, \delta) + n^{-1/2} \sum_{i=1}^{n} \eta_i h(\delta_0, Y_{i-1}) \left[ \frac{1}{h(\delta, Y_{i-1})} - \frac{1}{h(\delta_0, Y_{i-1})} \right] I(\vartheta' Y_{i-1} \leq y)
\]

\[
+n^{-1/2} \sum_{i=1}^{n} h(\delta_0, Y_{i-1}) \left[ \frac{1}{h(\delta, Y_{i-1})} - \frac{1}{h(\delta_0, Y_{i-1})} \right] I(\vartheta' Y_{i-1} \leq y)
\]

\[
= T_n(y, \hat{\delta}) + \hat{S}_n(y) + \hat{R}_n(y), \quad \text{say}.
\]

\[
= \mathcal{V}_n(y, \delta_0) + u_p(1) + \hat{S}_n(y) + \hat{R}_n(y), \quad \text{by (2.1)}.
\]
Upon combining (2.8), (2.9), (2.10) with (3.4) of Section 3 below, we obtain
\[
(2.11) \quad \hat{R}_n(y) = -n^{1/2}(\hat{\delta} - \delta_0)\nu(y, \delta_0) + \nu_p(1).
\]
Next, consider the term $\hat{S}_n$. Rewrite

\[
-\hat{S}_n(y) := n^{-1/2} \sum_{i=1}^{n} \frac{\eta_i}{h(\delta, Y_{i-1})} \xi_{ni} I(\vartheta' Y_{i-1} \leq y)
\]

\[+ \Delta' \frac{1}{n} \sum_{i=1}^{n} \left[ \frac{1}{h(\delta, Y_{i-1})} - \frac{1}{h(\delta_0, Y_{i-1})} \right] \eta_i \hat{h}(\delta_0, Y_{i-1}) I(\vartheta' Y_{i-1} \leq y)
\]

\[+ \Delta' \frac{1}{n} \sum_{i=1}^{n} \eta_i \hat{h}(\delta_0, Y_{i-1}) I(\vartheta' Y_{i-1} \leq y)
\]

\[= \hat{S}_{n1}(y) + \Delta' \hat{S}_{n2}(y) + \Delta' \hat{S}_{n3}(y), \text{ say.}
\]

Because $E|\eta_0| < \infty$, $n^{-1} \sum_{i=1}^{n} |\eta_i| = O_p(1)$, and by (2.7), on $A$,

\[
sup_{y \geq 0} |\hat{S}_{n1}(y)| \leq 2\theta_0^{-1} n^{1/2} \max_{1 \leq i \leq n} |\xi_{ni}| n^{-1} \sum_{i=1}^{n} |\eta_i| = o_p(1).
\]

An argument similar to the one used to prove (2.10) yields $\sup_{y \geq 0} |\hat{S}_{n2}(y)| = o_p(1)$. Upon combining these facts with (3.5) below we obtain $\sup_{y \geq 0} |\hat{S}_n(y)| = o_p(1)$. This fact together with (2.11), (2.6) and (2.1) completes the proof of the theorem.

### 2.2 Asymptotically distribution free tests

A consequence of Theorem 2.1 is that the asymptotic null distribution of the process $\mathcal{V}_n(\cdot, \hat{\delta})$ depends on the null model and estimator, and hence in general is unknown. We shall now describe a transformation of this process which will yield a process with known asymptotic null distribution. This transformation is an analog of the one in Stute, Thies and Zhu (1998), Koul and Stute (1999) and Stute and Zhu (2002).

Let $G_{\vartheta}$ denote the d.f. of $\vartheta' Y_0$. Note that $G_{\vartheta} = G$. Recall that $h(\delta_0, y) := \theta_0 + \varphi(\theta' y)$, \( \hat{h}(\delta_0, y) = (1, y' \hat{\varphi}(\theta' y))' \) and $\nu(y, \delta)$ given at (2.5). Let $\xi_{\vartheta}(x) = E(\vartheta' Y_0 | \vartheta' Y_0 = x)$, $x \geq 0$ and

\[
L(x, \delta) := E \left( \frac{\hat{h}(\delta, Y_0)}{\hat{h}(\delta_0, Y_0)} | \vartheta' Y_0 = x \right) = \frac{1}{\vartheta_0 + \varphi(x)} \left( \begin{array}{c} 1 \\ \xi_{\vartheta}(x) \varphi(x) \end{array} \right)
\]

Then $E\{Y_0 \varphi(\vartheta' Y_0) | \vartheta' Y_0 = x\} = \xi_{\vartheta}(x) \varphi(x)$, and

\[
\nu(y, \delta) = E \left( \frac{\hat{h}(\delta, Y_{i-1})}{\hat{h}(\delta_0, Y_{i-1})} \right) I(\vartheta' Y_{i-1} \leq y) = \int_{x \leq y} L(x, \delta) dG_{\vartheta}(x).
\]

Define

\[
A(x, \delta) := \int_x^{\infty} \nu(y, \delta) \nu(y, \delta)' dG_{\vartheta}(y).
\]
Assume

\[(2.12) \quad A(x, \delta_0) \text{ is positive definite for all } 0 \leq x < \infty.\]

Write \(A_\delta^{-1}(x)\) for \((A(x, \delta))^{-1}\). Define, for \(y \geq 0, \delta \in \Theta,\)

\[
J_n(y, \delta) := \int_0^y \frac{\nu(z, \delta)'A_\delta^{-1}(z)}{\int z} \nu(s, \delta) d\mathcal{V}_n(s, \delta) dG_\delta(z),
\]

\[
W_n(y, \delta) := \mathcal{V}_n(y, \delta) - J_n(y, \delta).
\]

We note here that by arguing as in Khmaladze and Koul (2009), one can verify that the transformation \(W_n(y, \delta_0)\) is well defined even when \((2.12)\) does not hold.

Next, suppose \(H_0\) holds. Then clearly, \(E W_n(y, \delta_0) \equiv 0\), and more importantly,

\[
(2.13) \quad \text{Cov} \left( W_n(y_1, \delta_0), W_n(y_2, \delta_0) \right) = \sigma^2 G(y_1 \wedge y_2), \quad y_1, y_2 \geq 0.
\]

The proof of this claim appears in the last section. In fact, arguing as in say Koul and Stute (1999) and using Lemma 2.4, one can show that \(W_n(y, \delta_0)\) converges weakly to \(\sigma^2 B \circ G\), in \(D[0, \infty)\) and uniform metric. However, this result is of little use from the practical point of view, since this process depends on the unknown parameters \(\theta_0, \theta\) and \(G\). We need the analog of the above process when all parameters are estimated. The needed estimated entities are defined below.

Analogous to the regression set up of Stute and Zhu (2002), an estimate of \(\xi_\theta\) here can be taken to be a kernel estimator. Because here observations are nonnegative, it is desirable to take kernels that are supported on \([0, 1)\). One such example is that of the inverse gamma kernel of Mnatsakanov and Sarkisian (2012) defined as

\[
K_\alpha(y, u) := \frac{1}{\Gamma(\alpha + 1)} \left( \frac{\alpha y}{u} \right)^{\alpha + 1} \exp \left\{ -\left( \frac{\alpha y}{u} \right) \right\}, \quad \alpha > 0, y > 0, u > 0.
\]

The corresponding estimator of \(\xi_\theta\) then is

\[
\tilde{\xi}_n(y) := \sum_{i=1}^n \frac{K_\alpha(y, \hat{\theta}'Y_{i-1})Y_i}{\sum_{i=1}^n K_\alpha(y, \hat{\theta}'Y_{i-1})}, \quad y > 0.
\]

Now we are ready to define estimates of the other entities needed in the transformation. Let

\[
(2.14) \quad G_n(y) := n^{-1} \sum_{i=1}^n I(\hat{\theta}'Y_{i-1} \leq y), \quad L_n(y) := \frac{1}{\theta_0 + \varphi(y)} \left( \frac{1}{\tilde{\xi}_n(y) \varphi(y)} \right),
\]

\[
\nu_n(y) := \int_{x \leq y} L_n(x) dG_n(x), \quad A_n(y) := \int_y^\infty \nu_n(z) \nu_n'(z) dG_n(z),
\]

\[
\tilde{J}_n(y) := \int_0^y \nu_n(z) A_n^{-1}(z) \int_z^\infty \nu_n(s) d\mathcal{V}_n(s, \hat{\theta}) dG_n(z),
\]

\[
\tilde{W}_n(y) := \mathcal{V}_n(y, \hat{\theta}) - \tilde{J}_n(y), \quad y > 0.
\]
Again, using the arguments similar to those used in Koul and Stute (1999) or Khmaladze and Koul (2004, 2009) and the tightness of the process $\mathcal{V}_n(\cdot, \theta)$ proved in Lemma 2.4 here, we can prove that $\hat{W}_n(y, \hat{\theta}_n) \rightarrow_D \sigma^2 B \circ G$, in $D[0, \infty)$ and uniform metric. Consequently, tests based on continuous functionals of the process $\hat{W}_n(y, \hat{\theta}_n)$, $y \geq 0$ will be asymptotically distribution free. In particular, the test that rejects $\mathcal{H}_0$ whenever

$$\hat{\sigma}_n^{-1} \sup_{0 \leq y < \infty} |\hat{W}_n(y, \hat{\theta})| > b_\alpha$$

is of the asymptotic size $\alpha$, where $b_\alpha$ be such that $P(\sup_{0 \leq t \leq 1} |B(t)| > b_\alpha) = \alpha$ and

$$\hat{\sigma}_n^2 := n^{-1} \sum_{i=1}^n \left( \frac{Y_i}{\theta_0 + \varphi(\theta Y_{i-1})} - 1 \right)^2.$$

A computational formula for the above process is as follows. Let $X_{i-1} = \hat{\theta}' Y_{i-1}$, $\hat{\eta}_i = \{Y_i/\hat{\theta}_0 + \varphi(X_{i-1})\} - 1$. Then

$$\tilde{J}_n(y) = n^{-3/2} \sum_{j=1}^n \hat{\eta}_j \nu_n(X_{j-1})' \sum_{i=1}^n A_n^{-1}(X_{i-1}) \nu_n(X_{i-1}) I(X_{j-1} \leq X_{i-1} \wedge y),$$

$$\tilde{W}_n(y) = n^{-1/2} \sum_{j=1}^n \hat{\eta}_j \left\{ I(X_{j-1} \leq y) - n^{-1} \sum_{i=1}^n \nu_n(X_{j-1})' A_n^{-1}(X_{i-1}) \nu_n(X_{i-1}) I(X_{j-1} \leq X_{i-1} \wedge y) \right\}.$$ 

As an example consider the problem of fitting an autoregressive conditional duration (ACD($p, 0$)) model given by

$$Y_i = \psi(Y_{i-1}, \delta_0) \varepsilon_i, \quad \psi(Y_{i-1}, \delta_0) = \theta_0 + \varphi(\theta' Y_{i-1}),$$

$$\varphi(\theta' Y_{i-1}) := \theta_1 Y_{i-1} + \theta_2 Y_{i-2} + \cdots + \theta_p Y_{i-p},$$

where $\{\varepsilon_i\}$ is a sequence of iid non-negative r.v.’s with unit mean and finite variance. The resulting sequence $\{Y_i\}$ defined by the above ACD($p, 0$) is a $p^{th}$ order Markov sequence, which is weakly stationary if $\theta_1 + \theta_2 + \cdots + \theta_p < 1$. The unconditional mean is given by $E(Y_i) = \theta_0/[1 - \theta_1 - \theta_2 - \cdots - \theta_p]$, which requires $\theta_0$ to be positive.

In general the rv $\varepsilon_i$ could follow any continuous distribution on the non-negative support with required moment conditions. If, for example, we assume that $\varepsilon_i$ follows a unit exponential distribution, then the conditional pdf of $Y_i$, given $Y_{i-1}$ is

$$f(y_i|Y_{i-1}) = \frac{1}{\psi(Y_{i-1}, \delta_0)} e^{-y_i/\psi(Y_{i-1}, \delta_0)}.$$
One can use likelihood based on this density to estimate $\delta$.

In this example $\varphi(x) \equiv x$ and the condition $(\varphi)$ is a priori satisfied with $\dot{\varphi} \equiv 1$. Thus the above testing procedure is applicable for fitting ACD($p, 0$) model to the given time series.

Similarly it is also applicable to fit an ARCH($p$) model given by

$$X_i = \sigma_i \zeta_i, \quad \sigma_i^2 = \theta_0 + \theta_1 X_{i-1}^2 + \theta_2 X_{i-2}^2 + \cdots + \theta_p X_{i-p}^2,$$

where $\{\zeta_i\}$ is a sequence of iid symmetric rvs with mean zero and finite variance. If we define $Y_i = X_i^2$ and $\zeta_i^2 = \varepsilon_i$ and $\sigma_i^2 = \psi_i$ then the ARCH($p$) model becomes an ACD($p, 0$) model. Hence the conditions for stationarity and other distributional properties will follow similarly. In particular if we assume standard normal distribution for $\zeta_i$ then $Y_i$ becomes an ACD model with Chi-square-1 innovations.

3 Proofs

This section contains the proof of Lemma 2.1, which is facilitated by the following preliminary result. Let $\gamma$ be a measurable function from $[0, \infty)^p \times \mathbb{R}$ to $\mathbb{R}$, $\eta_i := [\varepsilon_i/\varphi(\theta'Y_{i-1})] - 1$, $i \in \mathbb{Z}$, and define, for $y \in \mathbb{R}^+, s \in \mathbb{R}^p, b \in \mathbb{R}$,

$$\mathcal{H}_n(y, s, b) := n^{-1} \sum_{i=1}^n \gamma(Y_{i-1}, \eta_i) I((\theta + n^{-1/2}s)'Y_{i-1} \leq y + bn^{-1/2}\|Y_{i-1}\|),$$

$$H_n(y) := n^{-1} \sum_{i=1}^n \gamma(Y_{i-1}, \eta_i) I(\theta'Y_{i-1} \leq y), \quad H(y) := E\gamma(Y_0, \eta_1) I(\theta'Y_0 \leq y).$$

We are now ready to state the first preliminary result.

**Lemma 3.1** Let $\{Y_i, i \in \mathbb{Z}\}$ be a stationary and ergodic time series satisfying $E\|Y_0\|^2 < \infty$, (1.1), (1.2), and having continuous stationary distribution. Let $\gamma$ be a non-negative function on $[0, \infty)^p \times \mathbb{R}$ satisfying $E\gamma(Y_0, \eta_1) < \infty$. Then, for every $0 < c < \infty, b \in \mathbb{R}$,

$$\sup_{y \geq 0, \|s\| \leq c} \left| \mathcal{H}_n(y, s, b) - H(y) \right| = o_p(1).$$

**Proof.** By using a Glivenko-Cantelli type argument where $[0, \infty]$ is totally bounded by the nondecreasing bounded function $H(y)$, and by the Ergodic Theorem,

$$\sup_{y \geq 0} |H_n(y) - H(y)| = o(1), \quad \text{a.s.}$$

Stationarity and the assumption $E\|Y_0\|^2 < \infty$ imply

$$\max_{1 \leq i \leq n} n^{-1/2}\|Y_{i-1}\| = o_p(1).$$
Thus for every $\epsilon > 0$ there is an $N_{\epsilon}$ such that $P(A_n) > 1 - \epsilon$, for all $n > N_{\epsilon}$, where $A_n := \{\max_{1 \leq i \leq n} n^{-1/2}\|Y_i\| \leq \epsilon\}$. But, on $A_n$, for all $\|s\| \leq c$, $b \in \mathbb{R}$,

$$H_n(y - (c + b)\epsilon) \leq H_n(y, s, b) \leq H_n(y + (c + b)\epsilon), \quad \forall y \geq 0.$$  

The claim (3.1) follows from these inequalities, (3.2), and the continuity of the stationary distribution, thereby completing the proof of the lemma.

Next, consider, for $y \geq 0$,

$$\hat{R}_{n,3}(y) := n^{-1}\sum_{i=1}^{n}\frac{\hat{h}(\delta_0, Y_{i-1})}{\hat{h}(\delta_0, Y_{i-1})}I(\hat{\theta}Y_{i-1} \leq y), \quad \hat{S}_{n,3}(y) := \frac{1}{n}\sum_{i=1}^{n}\frac{\hat{h}(\delta_0, Y_{i-1})}{\hat{h}(\delta_0, Y_{i-1})}I(\hat{\theta}Y_{i-1} \leq y).$$

We have

**Corollary 3.1** Under the assumptions of Theorem 2.1,

$$\sup_{y \geq 0}\|\hat{R}_{n,3}(y) - \nu(y, \delta_0)\| = o_p(1). \quad (3.4)$$

$$\sup_{y \geq 0}\|\hat{S}_{n,3}(y)\| = o_p(1). \quad (3.5)$$

**Proof.** Proof of (3.4). Let

$$r_0(\delta_0, Y_{i-1}) := \frac{1}{\hat{h}(\delta_0, Y_{i-1})}, \quad r_j(\delta_0, Y_{i-1}) := Y_{i-j}\frac{\hat{\phi}(\theta'Y_{i-1})}{\hat{h}(\delta_0, Y_{i-1})}, \quad j = 1, \ldots, p,$$

$$T_j(y, s) := n^{-1}\sum_{i=1}^{n}r_j(\delta_0, Y_{i-1})I(\theta + n^{-1/2}s)'Y_{i-1} \leq y), \quad j = 0, 1, \ldots, p,$$

$$T(y, s) := (T_0(y, s), T_1(y, s), \ldots, T_p(y, s))', \quad y \geq 0, \quad s \in \Theta_1.$$ 

Then with $s = n^{1/2}(\hat{\theta} - \theta)$, $\hat{R}_{n,3}(y) = T(y, s)$. Thus, in view of the assumption (C), to prove (3.4), it suffices to show that for every $0 < c < \infty$,

$$\sup_{y \geq 0, \|s\| \leq c}\|T(y, s) - \nu(y, \delta_0)\| = o_p(1). \quad (3.6)$$

But note that the $j$th coordinates of $T(y, s)$ and $\nu(y, \delta_0)$ are like the $\mathcal{H}(y, s, 0)$ and $H(y)$ of Lemma 3.1 with $\gamma = r_j, \ j = 0, \ldots, p$, respectively. Thus (3.6) follows from Lemma 3.1 upon writing $r_j = r_j^+ - r_j^-$ and applying that lemma with $\gamma = r_j^+$ for each $j = 0, \ldots, p$.

**Proof of (3.5).** Follows similarly from Lemma 3.1 upon applying it with the gamma functions $\gamma_0(x, z) := z/h(\delta_0, x), \gamma_j(x, z) := (\hat{\phi}(\theta'x)/h(\delta_0, x))z x_j, \ j = 1, \ldots, p$. In this case the corresponding $H$ functions are as follows.

$$H_0(y) = E\gamma_0(Y_0, \eta_1)I(\theta'Y_0 \leq y) = E\left(\frac{1}{h(\delta_0, Y_0)}I(\theta'Y_0 \leq y)\right)E(\eta_1) = 0,$$

$$H_j(y) = E\frac{\hat{\phi}(\theta'Y_0)}{h(\delta_0, Y_0)}\eta_1 Y_0 I(\theta'Y_0 \leq y) \equiv 0.$$
because of the independence of $\eta_1$ and $Y_0$ and because $E\eta_1 = 0$. This completes the proof of the corollary.

**Proof of Lemma 2.1.** We will follow the structure of the proof of uniform Glivenko-Cantelli theorems in van der Vaart and Wellner (1997) (VW). Accordingly, let $\xi_i, i \in Z$ be an independent copy of the errors $\eta_i, i \in Z$ and $\zeta, \zeta_i, i \in Z$, be i.i.d. r.v.'s, independent of $Y_i, \eta_i, i \in Z$, with $P(\zeta = 1) = P(\zeta_i = -1) = 1/2$. For any r.v. $Z$, let $P_Z(E_Z)$ denote the conditional probability distribution (expectation) of $Z$, given all other r.v.'s. Then

$$E_{\eta} \sup_{y \geq 0, \|s\| \leq c} |V_{\eta}(y, s) - V_{\eta}(y, 0)|$$

$$= E_{\eta} \sup_{y \geq 0, \|s\| \leq c} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ g(Y_{i-1}, \eta_i) - E_{\xi} \{ g(Y_{i-1}, \xi_i) \} \right\} \right|$$

$$\times \left[ I((\theta + n^{-1/2}s)'Y_{i-1} \leq y) - I(\theta'Y_{i-1} \leq y) \right] \right|_{\mathcal{F}_0}$$

$$\leq E_{\xi} E_{\eta} \left\{ \sup_{y \geq 0, \|s\| \leq c} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \zeta_i [g(Y_{i-1}, \eta_i) - g(Y_{i-1}, \xi_i)] \right| \right\} \right|_{\mathcal{F}_0}$$

$$\leq \left| I((\theta + n^{-1/2}s)'Y_{i-1} \leq y) - I(\theta'Y_{i-1} \leq y) \right| \right|_{\mathcal{F}_0}$$

$$\leq 2E_{\xi} E_{\eta} \left\{ \sup_{y \geq 0, \|s\| \leq c} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \zeta_i g(Y_{i-1}, \eta_i) \right| + \left| I((\theta + n^{-1/2}s)'Y_{i-1} \leq y) - I(\theta'Y_{i-1} \leq y) \right| \right\}.$$ 

Now note that the process

$$W(y, s) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \zeta_i g(Y_{i-1}, \eta_i) \left[ I((\theta + n^{-1/2}s)'Y_{i-1} \leq y) - I(\theta'Y_{i-1} \leq y) \right]$$

is a Rademacher process $X_f$ indexed by the functions

$$f = \frac{1}{\sqrt{n}} g(Y_{i-1}, \eta_i) \left[ I((\theta + n^{-1/2}s)'Y_{i-1} \leq y) - I(\theta'Y_{i-1} \leq y) \right], \ i = 1, \ldots, n.$$

We shall follow the proof of 2.4.3 in VW.

Accordingly, let $P_n$ be the empirical measure based on $(Y_{i-1}, \eta_i), i = 1, \ldots, n$. For any $\epsilon > 0$, let $\mathcal{G}$ be an $\epsilon$-net in $L_1(P_n)$ over
\[ \mathcal{F} = \left\{ f : f(z, \eta) = g(z, \eta)[I((\theta + \frac{s}{\sqrt{n}})'z \leq y) - I(\theta'z \leq y)] : 
qquad z \in [0, \infty)^p, s \in \mathbb{R}^p, y \geq 0, \|s\| \leq c \right\}. \]

Then
\[ \sup_{y \geq 0, \|s\| \leq c} |W(y, s)| = \sup_{f \in \mathcal{F}} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \zeta_i f(Y_{i-1}, \eta_i) \right|. \]

Moreover,
\[ E_{\delta} \sup_{f \in \mathcal{F}} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \zeta_i f(Y_{i-1}, \eta_i) \right| \leq E_{\delta} \sup_{f \in \mathcal{G}} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \zeta_i f(Y_{i-1}, \eta_i) \right| + \epsilon. \]

Using a maximal inequality with the appropriate Orlicz norm for \( e^{x^2} - 1 \) taken over \( \zeta_1, \cdots, \zeta_n \) with the rest of variables fixed and using Hoeffding’s inequality for the Rademacher process as on pages 123–124 in VW, we bound the above expectation further by
\[ C(1 + \log N(\epsilon, \mathcal{F}, L_1(P_n)))^{1/2} \sup_{f \in \mathcal{F}} \left( \frac{1}{n} \sum_{i=1}^{n} f^2(Y_{i-1}, \eta_i) \right)^{1/2} + \epsilon, \]
where \( \{y \geq 0, \|s\| \leq c\} \) is a VC class, so the cardinality of \( \mathcal{G} \) is \( N(\epsilon, \mathcal{F}, L_1(P_n)) \leq C_V e^{-(V-1)} \) with a finite VC dimension \( V > 1 \). The last supremum above can be further bounded from the above by
\[
\sup_{f \in \mathcal{G}} \left( \frac{1}{n} \sum_{i=1}^{n} f^2(Y_{i-1}, \eta_i) \right) \\
= \sup_{y \geq 0, \|s\| \leq c} \left( \frac{1}{n} \sum_{i=1}^{n} g^2(Y_{i-1}, \eta_i) \left[ I((\theta + \frac{s}{\sqrt{n}})'Y_{i-1} \leq y) - I(\theta'Y_{i-1} \leq y) \right]^2 \right) \\
= \sup_{y \geq 0, \|s\| \leq c} \left( \frac{1}{n} \sum_{i=1}^{n} g^2(Y_{i-1}, \eta_i) \left| I((\theta + \frac{s}{\sqrt{n}})'Y_{i-1} \leq y) - I(\theta'Y_{i-1} \leq y) \right| \right) \\
\leq \sup_{y \geq 0, \|s\| \leq c} \left( \frac{1}{n} \sum_{i=1}^{n} g^2(Y_{i-1}, \eta_i) I(\theta'Y_{i-1} \text{ is between } y, y - \frac{s}{\sqrt{n}} Y_{i-1}) \right) \\
\leq \sup_{y \geq 0} \left( \frac{1}{n} \sum_{i=1}^{n} g^2(Y_{i-1}, \eta_i) I(\theta'Y_{i-1} \in [y - \frac{c\|Y_{i-1}\|}{\sqrt{n}}, y + \frac{c\|Y_{i-1}\|}{\sqrt{n}}]) \right) \\
= o_p(1). \]

The last claim follows from Lemma 3.1 because the process in the last but one bound equals \( \mathcal{H}_n(y, 0, c) - \mathcal{H}_n(y, 0, -c) \) of Lemma 3.1 with \( \gamma = g^2 \).

**Proof of Lemma 2.2.** Apply the CLT for martingales (Hall and Heyde: 1980, Corollary 3.1) to show that the fidiis tend to the right limit.
The proof of tightness is given in the following two lemmas, where $\sigma^2 = E\eta^2$. Since $\delta_0$ is fixed, we shall write $V_n(y)$ for $V_n(y, \delta_0)$ in this proof. Let $Z_i := \theta_i^t Y_{i-1}$, and $\eta'_i, i \in \mathbb{Z}$ be an independent copy of $\eta_i$. Let

$$V'_n(z) = n^{-1/2} \sum_{i=1}^n \eta'_i I(Z_{i-1} \leq z), \quad z \geq 0.$$ 

Recall the definition of $G_n$ from (2.14). Upon taking $\gamma \equiv 1, b = 0, s = 0$ in (3.1), we see that $H_n(y, 0, 0) \equiv G_n(y)$, $H(y) \equiv G(y)$. Hence from (3.1), we obtain

$$\lim_{n \to \infty} \sup_{z} |G_n(z) - G(z)| = 0, \quad \text{a.s.}$$

Let $\mathbb{R}^+ := [0, \infty]$, $\rho(y, z) = |G(y) - G(z)|, y, z \in \mathbb{R}^+$, and, for $f : \mathbb{R}^+ \to \mathbb{R}$,

$$\|f\|_{\delta} = \sup_{y, z \in \mathbb{R}^+: \rho(y, z) < \delta} |f(y) - f(z)|.$$

We are now ready to state an inequality useful in proving the tightness of $V_n$.

**Lemma 3.2 (Symmetrization Lemma):** For $\epsilon, \gamma, \delta > 0$ fixed with $\epsilon < \gamma$ and $\delta < \epsilon^2/3\sigma^2$, 

$$P(\|V_n\|_{\delta} > \gamma) \leq (1 - 3\delta \sigma^2/\epsilon^2)^{-1} P(\|V_n - V'_n\|_{\delta} > \gamma - \epsilon) + P(\sup_{k \geq n, z \in \mathbb{R}^+} |G_k(z) - G(z)| > \delta).$$

**Proof:** Let $P'_{\epsilon}$ and $E'_{\epsilon}$ denote the conditional probability measure and expectation, respectively, given $\eta_i, Z_i, i \in \mathbb{Z}$. Let $\rho_n(y, z) := |G_n(y) - G_n(z)|$, and

$$B_n(\gamma; y, z) := \left[|V_n(y) - V_n(z)| > \gamma\right], \quad B_n(\gamma, \delta) := \left[\|V_n\|_{\delta} > \gamma\right].$$

Note

$$B_n(\gamma, \delta) = \bigcup_{y, z \in \mathbb{R}^+: \rho(y, z) < \delta} B_n(\gamma; y, z).$$

For fixed $y, z \in \mathbb{R}^+$ with $\rho(y, z) < \delta$, 

$$B_n(\gamma; y, z) \cap \left[\|V_n - V'_n\|_{\delta} > \gamma - \epsilon\right] \supset B_n(\gamma; y, z) \cap \left[|V'_n(y) - V'_n(z)| < \epsilon\right].$$

Since $E'_{\epsilon}(V'_n(y) - V'_n(z)) = 0$ and $E'_{\epsilon}(V'_n(y) - V'_n(z))^2 = \sigma^2 \rho_n(y, z)$, Chebyshev's inequality gives $P'_{\epsilon}(|V'_n(y) - V'_n(z)| < \epsilon) \geq 1 - \sigma^2 \rho_n(y, z)/\epsilon^2$. Then

$$I(B_n(\gamma; y, z)) P'_{\epsilon} \left(\|V_n - V'_n\|_{\delta} > \gamma - \epsilon\right) \geq I(B_n(\gamma; y, z)) P'_{\epsilon} \left(|V'_n(y) - V'_n(z)| < \epsilon\right) \geq I(B_n(\gamma; y, z)) \left(1 - \sigma^2 \rho_n(y, z)/\epsilon^2\right).$$

Let

$$A_n(\delta) = \bigcup_{k \geq n} \left[\sup_{z \in \mathbb{R}^+} |G_k(z) - G(z)| > \delta\right].$$
Using the triangle inequality, on $A_n^c(\delta)$, $\rho_k(y, z) \leq 2\delta + \rho(y, z)$ for all $k \geq n$, $y, z \in \mathbb{R}^+$. Then for $y, z \in \mathbb{R}^+$ with $\rho(y, z) < \delta$,
\[
I(B_n(\gamma; y, z) \cap A_n^c(\delta)) P_\gamma(\|V_n - V'_n\|_\delta > \gamma - \epsilon) \\
\geq I(B_n(\gamma; y, z) \cap A_n^c(\delta))(1 - \sigma^2 \rho_n(y, z)/\epsilon^2) \\
\geq I(B_n(\gamma; y, z) \cap A_n^c(\delta))(1 - 3\sigma^2/\epsilon^2).
\]

This, in turn, gives
\[
I(B_n(\gamma, \delta) \cap A_n^c(\delta))(1 - 3\sigma^2/\epsilon^2) \leq I(B_n(\gamma, \delta) \cap A_n^c(\delta)) P_\gamma(\|V_n - V'_n\|_\delta > \gamma - \epsilon|\mathcal{F}) \\
\leq P_\gamma(\|V_n - V'_n\|_\delta > \gamma - \epsilon)
\]
and
\[
I(B_n(\gamma, \delta)) \leq I(B_n(\gamma, \delta) \cap A_n^c(\delta)) + I(A_n(\delta)) \\
\leq (1 - 3\sigma^2/\epsilon^2)^{-1} P_\gamma(\|V_n - V'_n\|_\delta > \gamma - \epsilon) + I(A_n(\delta)).
\]

Now take the expectation to obtain
\[
P(\|V_n\|_\delta > \gamma) \leq (1 - 3\delta\sigma^2/\epsilon^2)^{-1} P(\|V_n - V'_n\|_\delta > \gamma - \epsilon) + P(A_n(\delta)).
\]

From the above inequality it thus suffices to prove the tightness of the given process when $\eta_i$ are symmetric around zero, which is done in the next lemma.

**Lemma 3.3** Suppose $\eta$ is symmetrically distributed around 0 and $E\eta^2 < \infty$. Then for any fixed $\delta_0$ the process $V_n(\cdot, \delta_0)$ is sub-Gaussian with the semi-metric $d(y, z) = (\sigma^2 \rho(y, z))^{1/2}$, $y, z \in [0, \infty]$, that is, for any $y, z, x > 0$,
\[
\limsup_n P\left(\|V_n(z, \delta_0) - V_n(y, \delta_0)\| > x\right) \leq 2 \exp\left(-\frac{x^2}{2d^2(y, z)}\right).
\]

**Proof.** Without the loss of generality assume $y < z$. As before, let $Z_{i-1} := \theta'Y_{i-1}$. Recall $Z_{i-1}, i \geq 1$ is a stationary process and $G$ denotes the d.f. of $Z_0$.

Let $\zeta, \zeta_i, i = 1, \cdots, P_\zeta$ and $E_\zeta$ be as in the proof of Lemma 2.1. Then by the assumed symmetry of $\{\eta_i\}$, $\eta_i = \zeta_i|\eta_i|, i \geq 1$, in distribution. Then, by the Hoeffding (1964) inequality, $\forall x > 0, \forall n \geq 1$,
\[
P_\zeta\left(\|V_n(z, \delta_0) - V_n(y, \delta_0)\| > x\right) = P_\zeta\left(\sum_{i=1}^n \zeta_i|\eta_i| (I(y < Z_{i-1} \leq z)| > x\right) \\
\leq 2 \exp\left(-\frac{x^2}{2d^2(y, z)}\right),
\]
where
\[
d^2(y, z) = \frac{1}{n} \sum_{i=1}^n \eta_i^2 I(y < Z_{i-1} \leq z).
\]
Note that because of the independence of \( \eta_i \) and \( Y_{i-1} \), \( E d^2_n(y, z) = \sigma^2 \rho(y, z) \), and by (3.7), \( d^2_n(y, z) \to d^2(y, z) \), a.s., from which the lemma follows.

**Proof of Lemma 2.2 contd.** Use Corollary 2.2.8 of VW and its proof together with (3.8) to obtain that for every \( r > 0 \) and all sufficiently large \( n \),

\[
E \sup_{d(y, z) \leq r} |\mathcal{V}_n(z, \delta_0) - \mathcal{V}_n(y, \delta_0)| \leq K \sqrt{2\sigma^2} \int_0^r \sqrt{\ln D(\tau, d)} d\tau,
\]

where \( D(\tau, d) = 2\sigma^2/\tau^2 \) is the maximal packing number of \((\mathbb{R}^+, d)\), which is the maximal number of \( \tau \)-separated points; \( K \) is an absolute constant. Straightforward algebra shows that \( \sqrt{2\sigma^2} \int_0^r \sqrt{\ln D(\tau, d)} d\tau = \sigma^2 \int_{\ln(2\sigma^2/\tau^2)}^{\infty} \sqrt{v} e^{-v} dv \). Then, by the Markov inequality, for any \( \gamma > 0 \),

\[
P\{ \sup_{d(y, z) \leq r} |\mathcal{V}_n(z, \delta_0) - \mathcal{V}_n(y, \delta_0)| > \gamma \} \leq K \sigma^2 \gamma^{-1} \int_{\ln(2\sigma^2/\tau^2)}^{\infty} \sqrt{v} e^{-v} dv.
\]

Since the integral is bounded by \( \Gamma(3/2) \) then for any \( \gamma > 0 \) and \( \gamma_1 > 0 \) with \( \gamma \gamma_1 \geq K \sigma^2 \Gamma(3/2) \) and for every \( r > 0 \) and all sufficiently large \( n \),

\[
P\{ \sup_{d(y, z) \leq r} |\mathcal{V}_n(z, \delta_0) - \mathcal{V}_n(y, \delta_0)| > \gamma \} \leq \gamma_1,
\]

thereby completing the proof of the tightness of the process \( \mathcal{V}_n(\cdot, \delta_0) \) when \( \eta_i \) are symmetric around zero. This fact together with Lemma 3.2 in turn proves the tightness of this process for general \( \eta_i \)’s. This in turn completes the proof of Lemma 2.2.

**Proof of (2.13).** Recall \( \delta_0 = (\theta_0, \theta')' \) and \( G \) is the d.f. of \( \theta'Y_0 \). Fix \( 0 \leq y_1 < y_2 < \infty \). Note that \( EJ_n(y, \delta_0) \equiv 0 \) and by (2.3),

\[
\text{Cov} \left( W_n(y_1, \delta_0), W_n(y_2, \delta_0) \right) = \sigma^2 G(y_1 \wedge y_2) - EV_n(y_1, \delta_0)J_n(y_2, \delta_0) - EV_n(y_2, \delta_0)J_n(y_2, \delta_0) + EJ_n(y_1, \delta_0)J_n(y_2, \delta_0).
\]

Observe that

\[
J_n(y, \delta_0) = n^{-1/2} \sum_{i=1}^n \eta_i \int_0^y \nu(z, \delta_0)A_{\delta_0}^{-1}(z)\nu(\theta'Y_{i-1}, \delta_0)I(\theta'Y_{i-1} \geq z)dG(z).
\]

Fix \( 0 \leq y_1 \leq y_2 < \infty \). Then, by the Fubini Theorem,

\[
EJ_n(y_1, \delta_0)J_n(y_2, \delta_0) = \sigma^2 E \left\{ \int_0^{y_1} \nu(s, \delta_0)A_{\delta_0}^{-1}(s)\nu(\theta'Y_0, \delta_0)I(\theta'Y_0 \geq s)dG(s) \times \int_0^{y_2} \nu(\theta'Y_0, \delta_0)I(\theta'Y_0 \geq z)A_{\delta_0}^{-1}(z)\nu(z, \delta_0)dG(z) \right\}
\]

\[
= \sigma^2 \int_0^{y_1} \nu(s, \delta_0)A_{\delta_0}^{-1}(s) \int_0^{y_2} A(z \wedge s, \delta_0)A_{\delta_0}^{-1}(z)\nu(z, \delta_0)dG(z)dG(s).
\]
\[
\begin{align*}
&\quad = 2\sigma^2 \int_0^{y_1} \int_s^{y_1} \nu(z, \delta_0)' A_{\delta_0}^{-1}(s) \nu(s, \delta_0) dG(z) dG(s) \\
&\quad \quad + \sigma^2 \int_0^{y_2} \nu(s, \delta_0)' A_{\delta_0}^{-1}(s) \int_0^{y_2} \nu(z, \delta_0) dG(z) dG(s).
\end{align*}
\]

Similarly,

\[
\begin{align*}
E \mathcal{V}_n(y_1, \delta_0) J_n(y_2, \delta_0) \\
&= \sigma^2 \int_0^{y_2} E \nu(\theta' \mathbf{Y}_0 \leq y_1) \nu(\theta' \mathbf{Y}_0, \delta_0)' I(\theta' \mathbf{Y}_0 \geq s) A_{\delta_0}^{-1}(s) \nu(s, \delta_0) dG(s) \\
&= \sigma^2 \int_0^{y_1} \int_s^{y_1} \nu(z, \delta_0)' A_{\delta_0}^{-1}(s) \nu(s, \delta_0) dG(z) dG(s),
\end{align*}
\]

\[
\begin{align*}
E \mathcal{V}_n(y_2, \delta_0) J_n(y_1, \delta_0) \\
&= \sigma^2 \int_0^{y_1} E \nu(\theta' \mathbf{Y}_0, \delta_0)' I(\theta' \mathbf{Y}_0 \geq s, \theta' \mathbf{Y}_0 \leq y_2) A_{\delta_0}^{-1}(s) \nu(s, \delta_0) dG(s) \\
&= \sigma^2 \int_0^{y_1} \int_s^{y_1} \nu(z, \delta_0)' A_{\delta_0}^{-1}(s) \nu(s, \delta_0) dG(z) dG(s) \\
&\quad + \sigma^2 \int_0^{y_1} \int_y^{y_2} \nu(z, \delta_0)' A_{\delta_0}^{-1}(s) \nu(s, \delta_0) dG(z) dG(s).
\end{align*}
\]

From the above derivations one readily sees that

\[
E \mathcal{V}_n(y_1, \theta) J_n(y_2, \theta) + E \mathcal{V}_n(y_2, \theta) J_n(y_1, \theta) = E J_n(y_2, \theta) J_n(y_1, \theta),
\]

thereby completing the proof of (2.13).

References


