Quantile inference for near-integrated autoregressive time series under infinite variance and strong dependence

Ngai Hang Chan\textsuperscript{a,1}, Rong-Mao Zhang\textsuperscript{b,*}

\textsuperscript{a} Department of Statistics, Chinese University of Hong Kong, Shatin NT, Hong Kong
\textsuperscript{b} Department of Mathematics, Zhejiang University, Yuquan Campus, Hangzhou 310027, China

Received 25 August 2008; received in revised form 22 August 2009; accepted 12 September 2009
Available online 23 September 2009

Abstract

Consider a near-integrated time series driven by a heavy-tailed and long-memory noise $\epsilon_t = \sum_{j=0}^{\infty} c_j \eta_{t-j}$, where $\{\eta_j\}$ is a sequence of i.i.d random variables belonging to the domain of attraction of a stable law with index $\alpha$. The limit distribution of the quantile estimate and the semi-parametric estimate of the autoregressive parameters with long- and short-range dependent innovations are established in this paper. Under certain regularity conditions, it is shown that when the noise is short-memory, the quantile estimate converges weakly to a mixture of a Gaussian process and a stable Ornstein–Uhlenbeck (O–U) process while the semi-parametric estimate converges weakly to a normal distribution. But when the noise is long-memory, the limit distribution of the quantile estimate becomes substantially different. Depending on the range of the stable index $\alpha$, the limit distribution is shown to be either a functional of a fractional stable O–U process or a mixture of a stable process and a stable O–U process. These results indicate that although the quantile estimate tends to be more efficient for infinite variance time series, extreme caution should be exercised in the long-memory situation.

© 2009 Elsevier B.V. All rights reserved.

AMS 2000 Subject Classifications: primary 62M10; 62E20; secondary 60F17

Keywords: Heavy-tailed; Long-range dependent; Near-integrated time series and quantile regression

* Corresponding author. Tel.: +86 571 87952402.
E-mail addresses: nhchan@sta.cuhk.edu.hk (N.H. Chan), rmzhang@zju.edu.cn (R.-M. Zhang).
1 Tel.: +852 2609 8519.
1. Introduction

Consider a near-integrated first-order autoregressive (AR(1)) model

\[ Y_i = \gamma_n Y_{i-1} + \varepsilon_i, \quad (1.1) \]

where \( \gamma_n = 1 - \gamma/n \) and \( \gamma \) is a real number. The asymptotic theory of autoregressive time series with roots on or near the unit circle has been actively pursued by statisticians and econometricians alike. As of today, a relatively complete theory has been established under the finite variance situation. For a concise review on the recent developments of this topic, see Chan [5] and the references therein.

A large number of empirical studies ranging from signal processing, network traffic to insurance, however, indicates that time series with heavy tails offer a viable alternative. For background information on heavy-tailed time series and their applications, readers are referred to the Séminaire Européen de Statistique edited by Fiksenstädt and Rootzén [13], where exemplary theories and applications of extreme values in finance, insurance, the environment and telecommunications are surveyed. In financial econometrics, there has also been increasing interest in modeling financial phenomena by time series driven by heavy-tailed innovations. For example, Fama [12] and Mandelbrot [23,24] argued that distributions of commodity and stock returns are often heavy-tailed with possible infinite variance. Rachev and Mittnik [28] considered stable Paretian models in finance, [22] studied agent-based models in with heavy tails, and [2] studied financial market model where order flows follow heavy-tailed and long-memory durations.

Due to the intricacy of the asymptotic theory involved in the infinite variance model, much less is known when both long-range dependence and infinite variance structure are exhibited in the time series. Since the least squares procedure is known to be less robust and less effective when the time series is heavy-tailed, one of the main purposes of this paper is to establish a more robust estimate of \( \alpha_n \) for nearly nonstationary AR(1) models (1.1) driven by strongly dependent and infinite variance innovations. For more information and applications concerning strong dependent and infinite variance processes, we refer the readers to [11,30] and the references therein. It should also be pointed out that an alternate way to describe long-range dependence is by means of aggregating short-memory processes with random coefficients, see for example [9,3] and the references therein. In this paper, we follow the traditional method of describing long-range dependence through linear processes.

Specifically, Chan and Zhang [7] considered the least squares inference for a nearly nonstationary time series with errors defined by a heavy-tailed and long-memory noise

\[ \varepsilon_i = \sum_{j=0}^{\infty} c_j \eta_{i-j}, \]

where \( c_0 = 0 \) and \( c_j = j^{-\beta} l(j) \) when \( j \geq 1, \beta > 1/\alpha, l(\cdot) \) is a slowly varying function and \( \eta_i, \eta \in \mathbb{Z} \) are i.i.d variables and in the domain of attraction of a stable law. That is, there exists some sequence \( a_n = \inf\{ x : P( |\eta_0| > x ) \leq 1/n \} = n^{1/\alpha} L(n) \), \( L(x) \) which is a slowly varying function such that

\[ a_n^{-\alpha} \sum_{i=1}^{[ns]} \eta_i \Rightarrow J_1 Z_{\alpha}(s), \quad (1.2) \]
where $Z_n(t)$ is a stable random variable with index $\alpha \in (0, 2)$ and $\Rightarrow J_1$ denotes weak convergence in the $J_1$ topology, see [4].

It is well known that when $E\varepsilon_t^2 = \infty$, the least squares estimate is not efficient. An important method used to deal with this problem is the so-called quantile regression, which has been receiving considerable attention since the seminal work of Koenker and Bassett [20], see also [19] for more discussion on this topic. Knight [17,18] established the limit distribution for least absolute deviations estimate for $\gamma_n = 1$ when the infinite variance errors $\{\varepsilon_i\}$ are independent or weakly dependent. Chan et al. [8] considered quantile inference for a nearly non-stationary time series (i.e., $\gamma_n = 1 - \gamma/n$) when $\{\varepsilon_i\}$ are independent with infinite variance.

In this paper, we generalize the results to the case when $\{\varepsilon_t\}$ are long- and short-memory processes with heavy tails. Limit distributions of quantile regression estimate are established under different scenarios. As the limit distributions for long- and short-memory errors are substantially different, these results indicate that when applying quantile regression to infinite variance time series, extreme caution should be exercised. In particular, for short-memory noise ($\beta > 2/\alpha$), the process $n^{-1/2} \sum_{i=1}^{[nt]} \varphi_t(\varepsilon_i - \beta_0(\tau))$ converges weakly to a Gaussian process, where $\varphi_t(x) = \tau - I(x < 0)$. In this case, standard arguments together with the continuous mapping theorem can then be used to show that the limit distribution of the quantile estimate of $\gamma_n$ is a functional of a stable process and a Brownian motion. For the long-memory case ($\beta < 2/\alpha$), instead of converging weakly to a Gaussian process, the partial sum process $n^{-1/2} \sum_{i=1}^{[nt]} \varphi_t(\varepsilon_i - \beta_0(\tau))$ converges weakly to a stable process. The crux of the difficulty lies in establishing the limit of the process $\sum_{i=1}^{n} Y_i \varphi_t(\varepsilon_i - \beta_0(\tau))$. We show that when $1/\alpha < \beta < (\alpha + 2)/(3\alpha)$ and $\alpha > 1$, $\sum_{i=1}^{n} Y_i \varphi_t(\varepsilon_i - \beta_0(\tau))$ can be approximated by $\sum_{i=1}^{n} Y_i \varepsilon_i$ and as a result, the limit distribution of the quantile estimate of $\gamma_n$ can be deduced from [7] as a functional of a fractional Ornstein–Uhlenbeck (O–U) stable process. On the other hand, when $1 < \beta < 2/\alpha$ and $\alpha > 1$, applying a result of [21] shows that the partial sum $\sum_{i=1}^{n} Y_i \varphi_t(\varepsilon_i - \beta_0(\tau))$ converges weakly to a stable process and as a result, the limit distribution of the quantile estimate of $\gamma_n$ is a functional of two different stable processes.

The paper is organized as follows. Section 2 provides the asymptotic distribution of quantile regression estimate. As the limit process depends on unknown parameters of the density of $\varepsilon$ at the quantile and the variance in the weakly dependent case, estimation of these parameters and their corresponding limit distributions are given in Section 3. Proofs of the main results are given in Section 4 and technical lemmas are relegated to the Appendix.

2. Quantile regression

Given $\tau \in (0, 1)$, let $\gamma(\tau) = \gamma_n$ and denote the $\tau$-th quantile of $\varepsilon_t$ by $\beta(\tau)$. Define $\rho_t(\mu) = \mu(\tau - I(\mu < 0)), \theta(\tau) = (\beta(\tau), \gamma(\tau))^T$ and $X_t = (1, Y_{t-1})^T$. Let $Q_t(\tau | t - 1)$ be the $\tau$-th conditional quantile of $Y_t$, conditional on $Y_{t-1}$. Then $Q_t(\tau | t - 1) = X_t^T \theta(\tau)$. According to [20], the quantile regression estimate is defined as

$$\hat{\theta}(\tau) = \arg\min_{\theta(\tau)} \sum_{i=1}^{n} \rho_t(Y_t - X_t^T \theta(\tau)).$$  \hspace{1cm} (2.1)

We impose the following conditions throughout the entire paper.

H$_1$. Let $\{Y_t\}$ follow model (1.1) with $\{\eta_j\}$ satisfying (1.2).
H2. The density \( p(x) \) of \( \eta_1 \) satisfies \( |p'(x)| \leq C_1(1 + |x|)^{-(1+\delta)} \) for some \( \delta > \max\{0, \alpha - 1\} \) and for all \( x \in \mathbb{R} \) and \( |p'(x) - p'(y)| \leq C_2|x - y|(1 + |x|)^{-(1+\delta)} \) for all \( x, y \in \mathbb{R} \) with \( |x - y| < 1 \).

Let \( \lambda = \sum_{j=0}^{\infty} c_j \), \( f(x) \) be the density of \( \epsilon \) and \( \theta_0(\tau) = (\beta_0(\tau), \gamma_0)^T \) be the true value of \( \theta(\tau) \). Define \( A(x) = \int_0^1 (1, x(s))^T (1, x(s)) \, ds \). We have the following theorems.

**Theorem 2.1.** Assume conditions H1 and H2. If \( \beta > 2/\alpha \), then

\[
D_n(\hat{\theta}(\tau) - \theta_0(\tau)) \xrightarrow{d} \frac{\sigma}{f(\beta_0(\tau))} (A(S))^{-1} \left( W(\tau, 1), \int_0^1 S(s) \, dW(\tau, s) \right)^T. \tag{2.2}
\]

In particular,

\[
a_n \sqrt{n}(\hat{\alpha}(\tau) - \alpha_n) \xrightarrow{d} \frac{\sigma}{f(\beta_0(\tau))} \frac{\int_0^1 S(s) \, dW(\tau, s) - W(\tau, 1) \int_0^1 S(s) \, ds}{\int_0^1 S^2(s) \, ds - \left( \int_0^1 S(s) \, ds \right)^2}, \tag{2.3}
\]

and

\[
\left( \sum_{t=1}^{n} Y_{t-1}^2 - \left( \sum_{t=1}^{n} Y_t \right)^2 \right) \xrightarrow{d} \mathcal{N}(0, \frac{\sigma^2}{f^2(\beta_0(\tau))}), \tag{2.4}
\]

where \( D_n = \text{diag}(\sqrt{n}, a_n \sqrt{n}) \), \( \sigma^2 = \text{Var}(\varepsilon) + 2 \sum_{j=1}^{\infty} \text{E}[\varepsilon_j \phi(\varepsilon_0) \phi(\varepsilon_j)] \) and \( W(\tau, \cdot) \) is a standard Brownian motion independent of \( S(s) = \lambda(Z_\alpha(s) - \gamma \int_0^s e^{-\gamma(s-t)} \, dZ_\alpha(t)) \), with \( Z_\alpha(t) \) being defined in (1.2).

**Theorem 2.2.** Assume conditions H1 and H2. If \( c_j \sim b_0 j^{-\beta} \), \( 1/\alpha < \beta < (\alpha + 2)/(3\alpha) \) and \( \alpha > 1 \), then

\[
D_n(\hat{\theta}(\tau) - \theta_0(\tau)) \xrightarrow{d} (A(Z_{\alpha,\beta,\gamma}))^{-1} \left( Z_{\alpha,\beta,\gamma}(1), \gamma \int_0^1 Z_{\alpha,\beta,\gamma}(s) \, ds + \frac{1}{2} Z_{\alpha,\beta,\gamma}^2(1) \right)^T, \tag{2.5}
\]

where \( D_n = \text{diag}(a_n^{-1} n^{\beta}, n) \), \( Z_{\alpha,\beta}(t) = \int_{-\infty}^{t} f_0^{-\beta} \, dW_{\alpha}(s) \) and

\[
Z_{\alpha,\theta,\gamma}(t) = Z_{\alpha,\theta}(t) - \gamma \int_0^t e^{-\gamma(s-t)} Z_{\alpha,\theta}(s) \, ds, \quad Z_{\alpha,\theta,\gamma}(0) = 0.
\]

**Theorem 2.3.** Assume conditions H1 and H2. Suppose that \( c_j \sim b_0 j^{-\beta} \) and \( \lim_{t \to \infty} P(\eta_0 > x) / P(\eta_0 > x) = 1/2 \). Then for \( 1 + \sqrt{1-1/\alpha} < \beta < 2/\alpha \) and \( \alpha > 1 \),

\[
D_n(\hat{\theta}(\tau) - \theta_0(\tau)) \xrightarrow{d} \frac{1}{f(\beta_0(\tau))} (A(S))^{-1} \left( -L_{\alpha,\beta}(1), - \int_0^1 S(s) \, dL_{\alpha,\beta}(s) \right)^T, \tag{2.6}
\]

where \( D_n = \text{diag}(na_n^{-1/\beta}, na_n^{-1/\beta}) \), \( L_{\alpha,\beta}(s) \) is a stable process with index \( \alpha \beta \) defined by

\[
L_{\alpha,\beta} = c^+ Z_{\alpha,\beta}^+ + c^- Z_{\alpha,\beta}^-, \quad c^{\pm} = \alpha \int_0^\infty (F(\beta_0(t) \pm t) - \tau) t^{-1-1/\beta} \, dt,
\]

\[
\Lambda = \left( \frac{b_0^{\alpha}(\alpha\beta - 1)}{\Gamma(2 - \alpha\beta) \cos(\pi\alpha\beta/2) \beta^\alpha} \right)^{1/(\alpha\beta)}.
\]
and $Z_{\alpha\beta}(s)$ is an independent copy of $Z_{\alpha\beta}^+(s)$ with characteristic function

$$E e^{itZ_{\alpha\beta}^+(s)} = \exp\{-s|t|^\alpha \beta [1 - \text{isgn}(t) \tan(\pi \alpha \beta/2)]\}.$$  

Note that Theorems 2.1–2.3 point out the subtle differences and difficulties in quantile estimation of the near-integrated model (1.1). In the short-memory case ($\beta > 2/\alpha$), Theorem 2.1 shows that the limit distribution of the quantile estimate converges weakly to a functional of a mixture of a Brownian motion and a stable O–U process. On the other hand, in the long-memory case with $1/\alpha < \beta < (\alpha + 2)/(3\alpha)$, Theorem 2.2 shows that the limit distribution of the quantile estimate converges weakly to a functional of a different fractional stable O–U process. But for the long-memory case with $1 < \beta < 2/\alpha$, Theorem 2.3 gives a completely different characterization of the limit distribution of the quantile estimate of $\alpha_n$ as a functional of a mixture of a stable process and a stable O–U process. Consequently, one needs to be extremely cautious in applying the quantile regression procedure in the near-integrated model as there is an abrupt change in the behavior of the limit distributions.

3. Semi-parametric estimates of $\alpha_n$

Although Theorems 2.1–2.3 give the limit distributions of the quantile estimate of $\gamma_n$, they involve the unknown parameters $f(\beta_0(\tau))$ and $\sigma^2$. Likewise, $f(\beta_0(\tau))$ and $F(\beta_0(\tau) + t)$ in Theorem 2.3 are also unknown a priori. In this section, we propose a semi-parametric estimate $\tilde{\alpha}_n$ to tackle this problem. Note that according to Theorem 2.1, we have

$$\hat{t}_{\gamma_n} = \left( \frac{1}{n} \sum_{i=1}^{n} Y_{t-1}^2 - \left( \frac{1}{n} \sum_{i=1}^{n} Y_t \right)^2 \right)^{1/2} (\hat{\alpha}(\tau) - \alpha_n) \xrightarrow{d} N\left(0, \frac{\sigma^2}{f^2(\beta_0(\tau))}\right).$$

This implies that

$$\frac{f(\beta_0(\tau))}{\sigma} \hat{t}_{\gamma_n} \xrightarrow{d} N(0, 1).$$

Therefore, if we can construct consistent estimators $\hat{\sigma}$ and $\hat{f}(\beta_0(\tau))$ to estimate $\sigma$ and $f(\beta_0(\tau))$ respectively, then

$$\tilde{\alpha}_n = \frac{\hat{f}(\beta_0(\tau))}{\hat{\sigma}} \hat{t}_{\gamma_n} \xrightarrow{d} N(0, 1).$$

Similarly, to apply Theorem 2.3, we need to construct a consistent estimator $\hat{F}(\beta_0(\tau) + t)$ of $F(\beta_0(\tau) + t)$. Since

$$\sigma^2 = E \psi_\tau^2(\varepsilon_0 - \beta_0(\tau)) + 2 \sum_{k=1}^{\infty} E \psi_\tau(\varepsilon_0 - \beta_0(\tau)) \psi_\tau(\varepsilon_k - \beta_0(\tau)) = 2\pi f_{\varepsilon \varepsilon}(0),$$

where $f_{\varepsilon \varepsilon}(\cdot)$ is the spectral density of $\{\psi_\tau(\varepsilon_t - \beta_0(\tau))\}$, we can estimate $\sigma^2$ by

$$\hat{\sigma}^2 = 2\pi \hat{f}_{\varepsilon \varepsilon}(0) = \sum_{j=-M}^{M} (1 - j/M) \hat{r}(j),$$

where $\hat{r}(j) = \frac{1}{n} \sum_{t=1}^{n} \psi_\tau(\varepsilon_t - \hat{\beta}_0(\tau)) \psi_\tau(\varepsilon_{t+j} - \hat{\beta}_0(\tau))$ and $M = o(n^{1/2})$, $M \to \infty$. 

To estimate \( f(\beta_0(\tau)) \) in Theorems 2.1 and 2.3, we use a kernel density estimate method. Let \( \varepsilon_i = Y_i - \hat{\alpha}_n Y_{i-1} \) be the residuals and let \( K(\cdot) \) be a symmetric and monotone kernel function with a bounded derivative, a compact support, \([-1, 1]\), and \( \int_{-1}^{1} K(x) \, dx = 1 \). Since \( f(\cdot) \) is unknown, \( \beta_0(\tau) \) is also unknown. We estimate \( f(\beta_0(\tau)) \) by

\[
\hat{f}(\beta_0(\tau)) = \frac{1}{nh} \sum_{i=1}^{n} K\left( \frac{\varepsilon_i - \hat{\beta}_0(\tau)}{h} \right).
\]

To estimate the distribution \( F(\beta_0(\tau) + t) \) of \( \varepsilon \), we use an empirical process defined by

\[
\tilde{F}_n(\beta_0(\tau) + t) = \frac{1}{n} \sum_{i=1}^{n} I(\varepsilon_i - \hat{\beta}_0(\tau) \leq t).
\]

We have the following theorems.

**Theorem 3.1.** Under the conditions of Theorem 2.1,

\[
\hat{f}(\beta_0(\tau)) - f(\beta_0(\tau)) = O_p\left( (nh)^{-1/2} + h^2 \right) \tag{3.1}
\]

and

\[
\hat{\sigma} - \sigma = o_p(1). \tag{3.2}
\]

**Theorem 3.2.** Under the conditions of Theorem 2.3,

\[
\hat{f}(\beta_0(\tau)) - f(\beta_0(\tau)) = O_p\left( (nh)^{-1} a_n^{1/\beta} + h^2 \right) \tag{3.3}
\]

\[
\tilde{F}_n(\beta_0(\tau) + t) - F(\beta_0(\tau) + t) = O_p(n^{-1} a_n^{1/\beta}) \tag{3.4}
\]

and

\[
na_n^{-1/\beta} \tilde{F}_n(\beta_0(\tau) + t) - F(\beta_0(\tau) + t) \Rightarrow J_1 Z_{a \beta}^* (t), \tag{3.5}
\]

where \( a_n = n^{1/\alpha_1}(n) \) is defined in Section 1 and \( Z_{a \beta}^* (t) \) is a stable process with index \( a \beta \) defined in Appendix.

Theorems 3.1 and 3.2 show that the proposed semi-parametric estimates of \( f \), \( \sigma^2 \) and \( F \) are consistent and as a result, they can be used in conjunction with Theorems 2.1–2.3 to construct confidence intervals for the quantile estimates of the near-integrated process (1.1) in the long-memory and heavy-tailed situations.

### 4. Proofs

**Proof of Theorem 2.1.** Put \( v = (v_1, v_2)^T = \sqrt{n}(\beta(\tau) - \beta_0(\tau), a_n(\gamma(\tau) - \gamma_n))^T \) and

\[
Z_n(v) = \sum_{t=1}^{n} \rho_t(\varepsilon_t - \beta_0(\tau) - v^T D_n^{-1} X_t) - \rho_t(\varepsilon_t - \beta_0(\tau)).
\]

Then

\[
Z_n(v) = -\sum_{t=1}^{n} v^T D_n^{-1} X_t \varphi_\tau(\varepsilon_t - \beta_0(\tau))
\]
Furthermore, by Lemma A.3, it follows that

\[ H_1 \longrightarrow_d -v^T \sigma(W(\tau, 1), \int_0^1 S(t) dW(t, t))^T \]  

(4.2)

and for all \(|v| \leq C\) for some \(C > 0\),

\[ \max_{1 \leq t \leq n} |\sqrt{n} v^T D_n^{-1} X_t| \leq |v_1| + |v_2| \sup_{0 \leq t \leq 1} |Y_{int}/a_n| = O_p(1). \]

Let

\[ Z_{tn}(v) = (v^T D_n^{-1} X_t - \varepsilon_t + \beta_0(\tau)) I(v^T D_n^{-1} X_t > \varepsilon_t - \beta_0(\tau) > 0) \times I(0 < \sqrt{n} v^T D_n^{-1} X_t \leq \log n), \]

\[ F_t = \sigma(\varepsilon_s, s \leq t), \quad \mu_{tn} = E(Z_{tn}(v)|F_{t-1}), \]

\[ A_{tn}(v) = v^T D_n^{-1} X_t I(0 < \sqrt{n} v^T D_n^{-1} X_t \leq \log n). \]  

(4.3)

Then

\[ \sum_{t=1}^n \mu_{tn} = \sum_{t=1}^n \int_{\beta_0(\tau)}^{\beta_0(\tau)+A_{tn}(v)} (A_{tn}(v) + \beta_0(\tau) - x) f_{t-1}(x) dx = \sum_{t=1}^n \int_{\beta_0(\tau)}^{\beta_0(\tau)+A_{tn}(v)} \int_x^s f_{t-1}(x) ds f_{t-1}(x) dx = \sum_{t=1}^n \int_{\beta_0(\tau)}^{\beta_0(\tau)+A_{tn}(v)} f_{t-1}(x) dx ds = \sum_{t=1}^n \int_{\beta_0(\tau)}^{\beta_0(\tau)+A_{tn}(v)} (s - \beta_0(\tau)) f_{t-1}(\beta_0(\tau))(1 + o_p(1)) ds = \frac{1}{2} \sum_{t=1}^n f_{t-1}(\beta_0(\tau)) A_{tn}(v)^2 + o_p(1). \]  

(4.4)

Note that under \(H_2\), \(E[f_{t-1}(\beta_0(\tau))]^\vartheta < \infty\) for some \(\vartheta > 1\). Combining with the stationarity of \(\{f_{t-1}(\beta_0(\tau))\}\) yields for some \(\delta > 0\)

\[ \max_{1 \leq k \leq n} \frac{1}{n^{1-\delta}} \sum_{t=1}^k |f_{t-1}(\beta_0(\tau)) - f(\beta_0(\tau))| = o_p(1). \]  

(4.5)

This implies that

\[ \sum_{t=1}^n \mu_{tn} =^p f(\beta_0(\tau)) A_{tn}(v)^2. \]

Furthermore, by Lemma A.3, we have \(\max_{1 \leq t \leq n} A_{tn}(v) = o_p(1)\), which implies
\[ \sum_{t=1}^{n} E(Z_{tn}^2(v) | \mathcal{F}_{t-1}) \leq \left( \max_{1 \leq t \leq n} A_{tn}(v) \right) \sum_{t=1}^{n} \mu_{tn} \rightarrow^p 0. \] (4.6)

Therefore,

\[ II_3 \equiv^p \frac{1}{2} f(\beta_0(\tau)) \sum_{t=1}^{n} (v^T D_n^{-1} X_t)^2 I(v^T D_n^{-1} X_t > 0). \] (4.7)

Similarly,

\[ II_2 \equiv^p \frac{1}{2} f(\beta_0(\tau)) \sum_{t=1}^{n} (v^T D_n^{-1} X_t)^2 I(v^T D_n^{-1} X_t < 0). \] (4.8)

By Lemma A.3 and (4.2),

\[ Z_n(v) \rightarrow^d -v^T \sigma \left( W(\tau, 1), \int_0^1 S(t) dW(\tau, t) \right)^T + \frac{1}{2} f(\beta_0(\tau))v^T A(S)v. \]

Since \( Z_n(v) \) has a convex sample path, Theorem 2.1 follows from Lemma 2.2 of [10]. \( \square \)

**Proof of Theorem 2.2.** Let \( b_n = a_n n^{-1/2} \), \( v = (nb_n^{-1}(\beta(\tau) - \beta_0(\tau)), n(\gamma(\tau) - \gamma_n))^T \), \( D_n = \text{diag}(n/b_n, n) \) and \( Z_n(v) = \sum_{i=1}^{n} \rho_i (\epsilon_i - \beta_0(\tau)) - \rho_i (\epsilon_i - \beta_0(\tau)). \) Similar to (4.4) and (4.5), we have

\[ nb_n^{-2} Z_n(v) = -nb_n^{-2} \sum_{i=1}^{n} v^T D_n^{-1} X_i \psi(\epsilon_i - \beta_0(\tau)) + \frac{1}{2} nb_n^{-2} f(\beta_0(\tau)) \]

\[ \times \sum_{t=1}^{n} (v^T D_n^{-1} X_t)^2 I(nb_n^{-1} |v^T D_n^{-1} X_t| \leq \log n) + o_p(1). \] (4.9)

Under the condition that \( 1/\alpha < \beta < (\alpha + 2)/(3\alpha) \), it can be shown after tedious calculations that for any \( x \in R \),

\[ \sup_{0 \leq t \leq 1} \left| \sum_{i=1}^{[nt]} \psi(\epsilon_i - \beta_0(\tau)) - \sum_{i=1}^{[nt]} f(\beta_0(\tau)) \epsilon_i - \sum_{i=1}^{[nt]} \sum_{j=0}^{\infty} E[\psi(\epsilon_i - \beta_0(\tau))] \right| \]

\[ - f(\beta_0(\tau)) \epsilon_i |\eta_i - j| + \sum_{i=1}^{[nt]} \sum_{0 \leq j_1 < j_2} c_{j_1} c_{j_2} \eta_i - j_1 \eta_i - j_2 | = o_p(n^{-2/3} a_n^2). \] (4.10)

It follows from Theorem 3.3 of [31] that

\[ F_n(t) =: \frac{1}{n^{1-2\beta} a_n^2} \sum_{i=1}^{[nt]} \sum_{0 \leq j_1 < j_2} c_{j_1} c_{j_2} \eta_i - j_1 \eta_i - j_2 \]

\[ \Rightarrow J_1 C \int_{-\infty}^{t} \int_{u_2}^{t} \int_{u_1}^{1} (x - u_1)^{-\beta} (x - u_2)^{-\beta} dx d\mathbb{Z}_\alpha(u_1) d\mathbb{Z}_\alpha(u_2), \] (4.11)

for some constant \( C \). By (4.11) and the weak convergence of \( S_n(t) =: \sum_{i=1}^{[nt]} \epsilon_i/b_n \Rightarrow J_1 Z_{\alpha, \beta}(t) \), we have for \( \delta \) small enough (see [4]),

(a) \( \sup_{0 \leq t \leq 1} |S_n(t)| = O_p(1) \) and \( \sup_{0 \leq t \leq 1} |F_n(t)| = O_p(1) \);

(b) \( \sup_{|s-t| \leq \delta} |S_n(s) - S_n(t)| = o_p(1) \) and \( \sup_{|s-t| \leq \delta} |F_n(s) - F_n(t)| = o_p(1) \).
Partition \([0, 1]\) into sub-intervals each with length \(\delta\), say \(A_i = [(i - 1)\delta, i\delta], i = 1, 2, \ldots, [1/\delta]\) and \(A^* = [[1/\delta]\delta, 1]\). Then

\[
\frac{1}{b_n^2} \sum_{i=1}^{n} \sum_{l=1}^{\infty} \sum_{0 \leq j_1 < j_2} c_{j_1} c_{j_2} \eta_{i-j_1} \eta_{i-j_2} \varepsilon_i
\]

\[= \frac{1}{n} \sum_{i=1}^{n} \left( \sum_{l=1}^{\infty} \sum_{0 \leq j_1 < j_2} c_{j_1} c_{j_2} \eta_{i-j_1} \eta_{i-j_2} / (n^{1-2\beta}a_n^2) \right) \varepsilon_i
\]

\[= \frac{1}{n} \sum_{i=1}^{[1/\delta]} \left[ \sum_{l=[n(i-1)\delta]+1}^{[ni\delta]} \left( F_n \left( \frac{l}{n} \right) - F_n \left( \frac{[n(i-1)\delta]+1}{n} \right) \right) \varepsilon_l \right] + \frac{1}{n} \sum_{l=[n(1/\delta)]+1}^{n} F_n \left( \frac{l}{n} \right) \varepsilon_l
\]

\[\leq \sup_{|s-t| \leq \delta} |F_n(s) - F_n(t)| \left( \frac{1}{n} \sum_{i=1}^{n} |\varepsilon_i| \right)
\]

\[+ a_n n^{-\beta} [1/\delta] \sup_{|s-t| \leq \delta} |S_n(s) - S_n(t)| \left( \sup_{0 \leq t \leq 1} |F_n(t)| \right)
\]

\[+ \sup_{0 \leq t \leq 1} |F_n(t)| \left( \frac{1}{n} \sum_{l=[n(1/\delta)]+1}^{n} |\varepsilon_l| \right) = o_p(1), \quad (4.12)
\]

by letting \(n \to \infty\) and then \(\delta \to 0\).

Let \(H(\eta) = \sum_{j=0}^{\infty} \left( E[F_0(\tau) - c_j \eta_i] - E[F(\beta_0(\tau) - c_j \eta_i)] + f(\beta_0(\tau)c_j \eta_i) \right)\) and \(H'(\eta) = \sum_{j=0}^{\infty} E[\varphi(\varepsilon_i - \beta_0(\tau)) - f(\beta_0(\tau))\varepsilon_i | \eta_{i-j}]\). Then

\[
\frac{1}{b_n^2} \sum_{i=1}^{n} \sum_{j=0}^{\infty} E[\varphi(\varepsilon_i - \beta_0(\tau)) - f(\beta_0(\tau))\varepsilon_i | \eta_{i-j}]
\]

\[= \frac{Y_n-1}{b_n^2} \sum_{i=1}^{n} H'(\eta_i) - \frac{1}{b_n^2} \sum_{i=1}^{n-1} \sum_{l=1}^{i} \left( H'(\eta_l) + H(\eta_l) \right) (Y_i - Y_{i-1})
\]

\[+ \frac{1}{b_n^2} \sum_{i=1}^{n-1} \sum_{l=1}^{i} H(\eta_l) (Y_i - Y_{i-1})
\]

\[= \frac{Y_n-1}{b_n^2} \left[ 2 \sum_{i=1}^{n-1} H'(\eta_i) + H'(\eta_n) \right] - \frac{1}{b_n^2} \left[ \sum_{i=1}^{n-1} Y_{i-1} H(\eta_i) \right]
\]

\[- \frac{1}{b_n^2} \sum_{i=1}^{n-1} \sum_{l=1}^{i} \left( H'(\eta_l) + H(\eta_l) \right) \left( \varepsilon_i - \frac{\gamma Y_{i-1}}{n} \right) \]. \quad (4.13)

Similar to Lemma 3.2 of [33], we have \(a_n^{-1/\beta} \sum_{i=1}^{n} H(\eta_i) \longrightarrow^d Z_{\alpha\beta}'\), where \(Z_{\alpha\beta}'\) is a stable process with index \(\alpha\beta\). For any \(0 \leq t \leq 1\),

\[
\sum_{i=1}^{[nt]} (H'(\eta_i) + H(\eta_i)) = - \sum_{i=0}^{[nt]-i} \left[ F_j(\beta_0(\tau) - c_j \eta_i) - \tau + f(\beta_0(\tau)c_j \eta_i) \right]
\]
+ \sum_{i=1}^{[nt]} \sum_{j=[nt]−i+1}^{\infty} \{F(\beta_0(\tau) − c_j \eta_i) − E[F(\beta_0(\tau) − c_j \eta_i)] + f(\beta_0(\tau)c_j \eta_i)\} \\
− \sum_{i=1}^{[nt]} \sum_{j=1}^{[nt]−i} \{F_j(\beta_0(\tau) − c_j \eta_i) − F(\beta_0(\tau) − c_j \eta_i)\} − E[F_j(\beta_0(\tau) − c_j \eta_i) − F(\beta_0(\tau) − c_j \eta_i)]
\leq V_1(t) + V_2(t) + V_3(t), \quad (4.14)

with \(E|V_i(t)|^r \leq Cn^{1+r−αβ+κ}\) for any \(κ > 0, 1 < r < αβ, l = 1, 2\) and \(E|V_3(t)|^2 \leq Cn\). Therefore, the first term of the right-hand side in (4.13) is equal to zero in probability and by (3.9) and (3.10) of [33], we can show the third term is

\[-\frac{1}{b_n^2} \sum_{i=1}^{n-1} \sum_{l=1}^{i} (H'(η_l) + H(η_l))\varepsilon_i + o_p(1)\]
\[= -\frac{1}{b_n^2} \sum_{i=1}^{n-1} \sum_{l=1}^{i} (H'(η_l) + H(η_l))\varepsilon_i [I(|\varepsilon_i| > a_n \log n) + I(|\varepsilon_i| \leq a_n \log n)] + o_p(1)\]
\[= -\frac{1}{b_n^2} \sum_{i=1}^{n-1} [(V_1(i/n) + V_2(i/n))\varepsilon_i + V_3\varepsilon_i(i/n)I(|\varepsilon_i| \leq a_n \log n)] + o_p(1)\]
\[= -\frac{1}{b_n^2} \sum_{i=1}^{n-1} [V_3(i/n)\varepsilon_i I(|\varepsilon_i| \leq a_n \log n)] + o_p(1).\]

Let \(ε_{i,n} = ε_i I(|ε_i| \leq a_n \log n)\), by the Hölder inequality, we have

\[E \left| \frac{1}{b_n^2} \sum_{i=1}^{n-1} [V_3(i/n)\varepsilon_i I(|\varepsilon_i| \leq a_n \log n)] \right| \leq \frac{1}{b_n^2} \sum_{i=1}^{n-1} [E(|V_3(i/n)|^2)]^{1/2} [E(|\varepsilon_i|^2)]^{1/2}\]
\[= o(1).\]

By Theorem 2.7 of [21], we have the second term of the right hand side of (4.13) is equal to zero in probability. Thus, by (4.9), (4.10), (4.12) and (4.13),

\[nb_n^{-2} Z_n(v) \Rightarrow^p f(\beta_0(\tau))\left(−\sum_{t=1}^{n} v^T nb_n^{-2} D_n^{-1} X_t ε_t + \frac{1}{2} nb_n^{-2} \sum_{t=1}^{n} (v^T D_n^{-1} X_t)^2\right). \quad (4.15)\]

Since \(Z_n(v)\) is convex, it follows that

\[nb_n^{-1}(\hat{\beta}(\tau) − \beta_0(\tau), b_n(\hat{\alpha}(\tau) − γ_n)) \Rightarrow^p \Sigma_n^{-1}\left(\frac{1}{b_n^2} \sum_{i=1}^{n} ε_i, \frac{1}{b_n^2} \sum_{i=1}^{n} Y_{i−1} ε_i\right),\]

where
Theorem 2.2

Theorem 2.2

Theorem 2.3

and the convexity of

Theorem 2.1

and that

Lemma A.4

(3.1)

(3.2)

It follows from

Thus,

Proofs of Theorem 3.1.

By Theorem 2.3 of [7], we have Theorem 2.2. □

Proof of Theorem 2.3. Let \( v = (na_n^{-1/\beta}(\beta(\tau) - \beta_0(\tau)), na_n^{-1/\beta}(\gamma(\tau) - \gamma_n(\tau)))^T, D_n = \text{diag}(na_n^{-1/\beta}, na_n^{-1/\beta}) \) and \( Z_n(v) \) defined as above. Using a similar argument of Theorem 2.2, we have

\[
na_n^{-2/\beta}Z_n(v) = na_n^{-2/\beta} \sum_{i=1}^{\infty} v^T D_n^{-1} X_i \varphi_{\tau} (\varepsilon_i - \beta_0(\tau)) + \frac{1}{2} f(\beta_0(\tau)) na_n^{-2/\beta} \sum_{i=1}^{n} (v^T D_n^{-1} X_i)^2 + o_p(1).
\]

By Lemma A.4 and the convexity of \( Z_n(v) \), we have Theorem 2.3. □

Proofs of Theorem 3.1. To prove (3.1) and (3.2), it is enough to show that for some \( \sigma_1 < \infty \),

\[
(\sqrt{n}h) \left( \hat{f}(\beta_0(\tau)) - f(\beta_0(\tau)) - \frac{1}{2} f''(\beta_0(\tau)) h^2 \right) \longrightarrow^d N(0, \sigma_1^2)
\]

and that

\[
\hat{r}(j) - r(j) = O_p(n^{-1/2}).
\]

Note that

\[
\hat{e}_i - \hat{\beta}_0(\tau) = e_i - \beta_0(\tau) + (\hat{\gamma}_n - \gamma_0) Y_{i-1} + \hat{\beta}_0(\tau) - \beta_0(\tau) =: e_i - \beta_0(\tau) + \hat{\mu}_n.
\]

Put \( g(\beta_0(\tau), u) = \frac{1}{nh} \sum_{i=1}^{n} K \left( \frac{e_i - \beta_0(\tau) + u/\sqrt{n}}{h} \right) \). Then by \( H_2 \) and the monotone property of \( K(\cdot) \), for any \( 0 < C < \infty \),

\[
\mathbb{E} \sup_{0 \leq |u| \leq C} |g(\beta_0(\tau), u) - g(\hat{\beta}_0(\tau), 0)|
\]

\[
\leq \frac{1}{nh} \sum_{i=1}^{n} \mathbb{E} \sup_{0 \leq |u| \leq C} \left| K \left( \frac{e_i - \beta_0(\tau) + u/\sqrt{n}}{h} \right) - K \left( \frac{e_i - \beta_0(\tau)}{h} \right) \right|
\]

\[
\leq \frac{1}{nh} \sum_{i=1}^{n} \int_{-\infty}^{\infty} |K((y - \beta_0(\tau))/h)| \sup_{0 \leq |u| \leq C} |f(y - u/\sqrt{n}) - f(y)| dy
\]

\[
\leq 2(nh)^{-1} \sum_{i=1}^{n} h |f'(\beta_0(\tau))| Cn^{-1/2} = O(n^{-1/2}).
\]

Thus,

\[
\sup_{0 \leq |u| \leq C} |g(\beta_0(\tau), u) - g(\hat{\beta}_0(\tau), 0)| = O_p(n^{-1/2}).
\]

It follows from Theorem 2.1 that

\[
\hat{\mu}_n = (\hat{\gamma}_n - \gamma_0) Y_{i-1} + \hat{\beta}_0(\tau) - \beta_0(\tau) = O_p(n^{-1/2}).
\]
Combining (4.20) and (4.21) yields
\[
\frac{1}{nh} \sum_{i=1}^{n} K((\hat{\xi} - \hat{\beta}_0(\tau))/h) - \frac{1}{nh} \sum_{i=1}^{n} K((\varepsilon_i - \beta_0(\tau))/h)
= g(\beta_0(\tau), \sqrt{n}\hat{\mu}_n) - g(\beta_0(\tau), 0) = O_p(n^{-1/2}).
\] (4.22)

Similar to the proof of (A.4) (see Lemma A.1), we have
\[
\sqrt{nh} \left( g(\beta_0(\tau), 0) - f(\beta_0(\tau)) - f''(\beta_0(\tau))h^2/2 \right) \rightarrow^d N(0, \sigma_1^2).
\] (4.23)
where \( \sigma_1^2 = \text{E}[(nh)^{-1/2} \sum_{i=1}^{n} K((\varepsilon_i - \beta_0(\tau))/h)] < \infty \). Combining (4.22) and (4.23) gives (4.16).

For (4.17), let \( \xi_t = \varepsilon_t - \beta_0(\tau) \) and
\[
L_n(u) = \frac{1}{n} \sum_{t=1}^{n} \varphi_{\tau}(\xi_t + u/\sqrt{n})\varphi_{\tau}(\xi_{t+j} + u/\sqrt{n}).
\]

Then
\[
L_n(u) - EL_n(0) = L_n(u) - L_n(0) + L_n(0) - EL_n(0)
= \frac{1}{n} \sum_{t=1}^{n} \varphi_{\tau}(\xi_t + u/\sqrt{n}) \left( I(\xi_{t+j} < 0) - I(\xi_{t+j} + u/\sqrt{n} < 0) \right)
\times \frac{1}{n} \sum_{t=1}^{n} \varphi_{\tau}(\xi_{t+j}) \left( I(\xi_t < 0) - I(\xi_t + u/\sqrt{n} < 0) \right)
+ L_n(0) - EL_n(0)
=: L_n(u) + L_n(u) + L_n(0) - EL_n(0).
\] (4.24)

Observe that
\[
\sup_{|u| \leq C} |L_{n1}(u)| \leq \frac{\tau + 1}{n} \sum_{t=1}^{n} \left| I \left( -\frac{C}{\sqrt{n}} \leq \xi_{t+j} \leq \frac{C}{\sqrt{n}} \right) \right|.
\] (4.25)

Thus,
\[
\text{E} \sup_{|u| \leq C} |L_{n1}(u)| \leq \frac{\tau + 1}{n} \sum_{t=1}^{n} P(-C/\sqrt{n} \leq \varepsilon_{t+j} - \beta_0(\tau) \leq C/\sqrt{n})
= (\tau + 1) \int_{\beta_0(\tau) - C/\sqrt{n}}^{\beta_0(\tau) + C/\sqrt{n}} f(x)dx
\leq 3C(\tau + 1)f(\beta_0(\tau))/\sqrt{n}.
\] (4.26)

This implies that \( \sup_{|u| \leq C} |L_{n1}(u)| = O_p(n^{-1/2}) \). Similarly, we have \( \sup_{|u| \leq C} |L_{n2}(u)| = O_p(n^{-1/2}) \). In the following, we will apply the method of Woodroofe for the central limit theorem for functions of Markov chains to show that
\[
\sqrt{n}(L_n(0) - EL_n(0)) \rightarrow^d N(0, \sigma_2^2),
\] (4.27)
where $\sigma^2 = \text{Var}(\sqrt{n}L_n(0)) < \infty$. Put $g(\xi_t, \xi_{t+j}) = \varphi_t(\xi_t)\varphi_t(\xi_{t+j})$. Since $\sup_{x \in R} |f(x)| \leq C$, it follows that

$$
\sum_{i=1}^{\infty} \|E(g(\xi_t, \xi_{t+j})|\mathcal{F}_1) - E(g(\xi_t, \xi_{t+j})|\mathcal{F}_0)\| \\
= \sum_{i=1}^{\infty} \| - \tau E(I(\xi_t < 0 + I(\xi_{t+j} < 0)|\mathcal{F}_1) - E(I(\xi_t < 0 + I(\xi_{t+j} < 0)|\mathcal{F}_0) \| \\
+ E[I(\xi_t < 0)I(\xi_{t+j} < 0)|\mathcal{F}_1] - E[I(\xi_t < 0)I(\xi_{t+j} < 0)|\mathcal{F}_0]\| \\
\leq \tau \sum_{i=1}^{\infty} \|E(I(\xi_t < 0)|\mathcal{F}_1) - E(I(\xi_t < 0)|\mathcal{F}_0)\| \\
+ \tau \sum_{i=1}^{\infty} \|E(I(\xi_{t+j} < 0)|\mathcal{F}_1) - E(I(\xi_{t+j} < 0)|\mathcal{F}_0)\| \\
+ \sum_{i=1}^{\infty} \|E[I(\xi_t < 0)I(\xi_{t+j} < 0)|\mathcal{F}_1] - E[I(\xi_t < 0)I(\xi_{t+j} < 0)|\mathcal{F}_0]\| \\
=: \Gamma_1 + \Gamma_2 + \Gamma_3. \tag{4.28}
$$

Let $\xi_{t0}$, $\xi_{t2}$ be defined as that in Lemma A.1. By (A.2), we have

$$
\Gamma_1 = \tau \sum_{i=1}^{\infty} \|E(I(\xi_{t0} + c_{t-1}\eta_1 + \xi_{t2} < \beta_0(\tau)|\mathcal{F}_1) \\
- E(I(\xi_{t0} + c_{t-1}\eta_1 + \xi_{t2} < \beta_0(\tau)|\mathcal{F}_0)\| \\
= \tau \sum_{i=1}^{\infty} \|E(G_t(\beta_0(\tau) - \xi_{t0} - c_{t-1}\eta_1) - G_t(\beta_0(\tau) - \xi_{t0} - c_{t-1}\eta_1)|\mathcal{F}_1)\| \\
< \infty. \tag{4.29}
$$

Similarly, $\Gamma_2 < \infty$. To show that $\Gamma_3 < \infty$, let $F_{ij}$ be the distribution of $\sum_{i=0}^{j-1} c_i \eta_{t+j-i}$. Note that

$$
E(I(\xi_t < 0)I(\xi_{t+j} < 0)|\mathcal{F}_1) - E(I(\xi_t < 0)I(\xi_{t+j} < 0)|\mathcal{F}_0) \\
= E \left[ I(\xi_t < 0)F_{ij} \left( \beta_0(\tau) - \sum_{i=j}^{\infty} c_i \eta_{t+j-i} \right) |\mathcal{F}_1 \right] \\
- E \left[ I(\xi_t < 0)F_{ij} \left( \beta_0(\tau) - \sum_{i=j}^{\infty} c_i \eta_{t+j-i} \right) |\mathcal{F}_0 \right] \\
= E \left[ (\xi_{t0} + c_{t-1}\eta_1 + \xi_{t2} < \beta_0(\tau)) - I(\xi_{t0} + c_{t-1}\eta_1 + \xi_{t2} < \beta_0(\tau)) \right] \\
F_{ij}(\beta_0(\tau) - \xi_{t+j,0} - c_{t+j-1}\eta_1 - \sum_{i=2}^{t} c_{t+j-i}\eta_i) |\mathcal{F}_1 \right] \\
- E \left[ I(\xi_{t0} + c_{t-1}\eta_1 + \xi_{t2} < \beta_0(\tau))F_{ij} \left( \beta_0(\tau) \right) \right].
$$
\[-\xi_{t+j,0} - c_{t+j-1}\eta_1 - \sum_{i=2}^{t} c_{t+j-i}\eta_i\]

\[-F_{ij}\left(\beta_0(\tau) - \xi_{t+j,0} - c_{t+j-1}\eta_1 - \sum_{i=2}^{t} c_{t+j-i}\eta_i\right) |\mathcal{F}_1\]

\[=: \Gamma_{31r} + \Gamma_{32r}. \quad (4.30)\]

Let \(p_t(x), p_1^*(y)\) be the densities of \(\eta_t\) and \(\eta_1\). Then

\[
\Gamma_{31r} = E[E(\Gamma_{31r}|\mathcal{F}_{t-1})|\mathcal{F}_1]
\]

\[= E\left(\int_R \int_R [I(\xi_{t,t-1} + c_0x < \beta_0(\tau)) - I(\xi_{t,t-1} + c_0x + c_{t-1}(y - \eta_1) < \beta_0(\tau))] F_{ij}(\beta_0(\tau) - \xi_{t,t-1} - c_0x)p_t(x)p_1^*(y)dx dy |\mathcal{F}_1\right)
\]

\[= E\left[\int_R \min \left\{ \frac{1}{c_0} |c_{t-1}(y - \eta_1)|, 1 \right\} p_1^*(y)dy |\mathcal{F}_1\right]
\]

\[= CE(\min\{c_0^{-1}c_{t-1}(\eta_1 - \eta_1), 1\}|\mathcal{F}_1). \quad (4.31)\]

Thus,

\[
\sum_{i=1}^{\infty} \|\Gamma_{31r}\| \leq \sum_{i=1}^{\infty} C \left| E \left[\int_R \min \left\{ \frac{1}{c_0} |c_{t-1}(y - \eta_1)|, 1 \right\} p_1^*(y)dy |\mathcal{F}_1\right]\right|^2 < \infty.
\]

(4.32)

Furthermore, by (A.2), we have

\[
\sum_{t=1}^{\infty} \|\Gamma_{32r}\| \leq \sum_{t=1}^{\infty} \left| E \left[ F_{ij} \left(\beta_0(\tau) - \xi_{t+j,0} - c_{t+j-1}\eta_1 - \sum_{i=2}^{t} c_{t+j-i}\eta_i\right) - F_{ij}\left(\beta_0(\tau) - \xi_{t+j,0} - c_{t+j-1}\eta_1' - \sum_{i=2}^{t} c_{t+j-i}\eta_i\right) |\mathcal{F}_1\right] \right|\]

\[\leq \sum_{t=1}^{\infty} \left| \min\{1, c_{t+j-1}|\eta_1 - \eta_1'\} \right|\]

\[= O\left(\sum_{t=1}^{\infty} |c_{t+j-1}|^{\min\{\alpha'/2,1\}}\right) < \infty. \quad (4.33)\]

From (4.32) and (4.33), it follows that \(|\Gamma_3| < \infty\). Arguing along the same line as in the proof of Lemma A.1, we have (4.27). Thus, \(\sup_{|u| \leq C} |L_n(u) - EL_n(0)| = O_p(n^{-1/2})\). Combining this with (4.21) implies (4.17). The proof of Theorem 3.1 is complete. \(\square\)
Proof of Theorem 3.2. Let \( \hat{\mu}_n \) be defined in (4.18) and 
\[
\tilde{g}(\beta_0(\tau), \mu) = \frac{1}{nh} \sum_{i=1}^{n} K((\varepsilon_i - \beta_0(\tau)) + \mu a_n^{1/\beta})/h).
\]

Similar to the proof of (4.22), we have
\[
\frac{1}{nh} \sum_{i=1}^{n} K((\varepsilon_i - \hat{\beta}_0(\tau))/h) - \frac{1}{nh} \sum_{i=1}^{n} K((\varepsilon_i - \beta_0(\tau))/h)
\]
\[
= \tilde{g}(\beta_0(\tau), a_n^{1/\beta} \hat{\mu}_n) - \tilde{g}(\beta_0(\tau), 0) = O_p(a_n^{1/\beta}).
\]

Let \( \xi'(\eta, x) = \sum_{j=1}^{\infty} (F(x - c_j \eta_{i-j}) - EF(x - c_j \eta_{i-j})) \). Then by revising Lemma 5.2 of [32], we have
\[
\sup_{y \in [-1, 1]} \sum_{i=1}^{n} |\xi'(\eta, \beta_0(\tau) + yh) - \xi'(\eta, \beta_0(\tau))|
\]
\[
- [\xi(\eta, \beta_0(\tau) + yh) - \xi(\eta, \beta_0(\tau))] = O_p(a_n^{1/\beta}).
\]

Thus, by Lemmas 5.3 and 5.4 of [32], we have
\[
nh a_n^{-1/\beta} (\tilde{g}(\beta_0(\tau), 0) - \hat{E}\tilde{g}(\beta_0(\tau), 0))
\]
\[
= -a_n^{-1/\beta} \int \sum_{i=1}^{n} [I(\varepsilon_i < x) - F(x)]dK((x - \beta_0(\tau))/h)
\]
\[
= -a_n^{-1/\beta} \int I(\varepsilon_i < \beta_0(\tau) + yh) - F(\beta_0(\tau) + yh)
\]
\[
- I(\varepsilon_i < \beta_0(\tau)) - F(\beta_0(\tau)]dK(y)
\]
\[
= - \int a_n^{-1/\beta} \sum_{i=1}^{n} (\xi(\eta, \beta_0(\tau) + yh) - \xi(\eta, \beta_0(\tau)))dK(y) + O_p(1).
\]

By (3.9) of [32], there exists \( 1 < \gamma < \alpha \) such that for all \( y \in [-1, 1] \),
\[
|\xi(\eta, \beta_0(\tau) + yh) - \xi(\eta, \beta_0(\tau))| \leq C \max(|\eta|^{1/\beta} h^{1/\gamma^\beta}, 1).
\]

This yields that for any \( w > 0 \) and a large enough \( n \),
\[
P \left( \sup_{y \in [-1, 1]} \sum_{i=1}^{n} |\xi(\eta, \beta_0(\tau) + yh) - \xi(\eta, \beta_0(\tau))| \geq C w a_n^{1/\beta} \right)
\]
\[
\leq nP \{ \max(|\eta|^{1/\beta} h^{1/\gamma^\beta}, 1) > w a_n^{1/\beta} \} \leq C w^{-a^\beta},
\]
which combines with (4.36) implies
\[
nh a_n^{-1/\beta} (\tilde{g}(\beta_0(\tau), 0) - \hat{E}\tilde{g}(\beta_0(\tau), 0)) = O_p(1).
\]

Note that \( \hat{E}\tilde{g}(\beta_0(\tau), 0) - f(\beta_0(\tau)) = -\frac{1}{2} f''(\beta_0(\tau))h^2 \int_{-1}^{1} y^2 K(y)dy + o(h^2) \) and \( a_n^{-1/\beta} = o(n^{-1} n^{-1}) \). Thus (3.3) follows from (4.34) and (4.38).
Set \( \mathcal{F}_n(\beta_0(\tau) + t, \mu) = \sum_{i=1}^{n} I(\varepsilon_i - \beta_0(\tau) + \mu a_n^{-1/\beta} \leq t) \). It is easy to see that for any \( x, y \in \mathbb{R} \),

\[
E \left| \sup_{|\mu| \leq C} \left| \left[ \mathcal{F}_n(\beta_0(\tau) + x, \mu) - \mathcal{F}_n(\beta_0(\tau) + x, \mu) \right] - \left[ \mathcal{F}_n(\beta_0(\tau) + y, \mu) - \mathcal{F}_n(\beta_0(\tau) + y) \right] \right| \leq 2 \int_{0}^{C/a_n^{1/\beta}} \int_{0}^{[x-y]} f'(\beta_0(\tau) + a + b) \, db \, da \leq C|x-y|/a_n^{1/\beta}.
\]

This implies that

\[
\sup_{t \in \mathbb{R}} |\tilde{\mathcal{F}}_n(\beta_0(\tau) + t) - \mathcal{F}_n(\beta_0(\tau) + t)| = O_p(a_n^{-1/\beta}). \tag{4.39}
\]

From Theorem 2.1 of [32], it follows that

\[
n a_n^{-1/\beta} (\mathcal{F}_n(\beta_0(\tau) + t) - F(\beta_0(\tau) + t)) \Rightarrow \Lambda Z_{a\beta}^+ \int_{0}^{\infty} [F(\beta_0(\tau) + t - s) - F(\beta_0(\tau) + t)] s^{-1-1/\beta} \, ds
\]

\[+ \Lambda Z_{a\beta}^- \int_{0}^{\infty} [F(\beta_0(\tau) + t + s) - F(\beta_0(\tau) + t)] s^{-1-1/\beta} \, ds =: Z_{a\beta}^*(t). \tag{4.40}
\]

By (4.39) and (4.40), we have (3.4) and (3.5). This completes the proof of Theorem 3.2. \( \square \)

**Acknowledgements**

We would like to thank the Editor, an associate editor and two anonymous referees for critical comments and helpful suggestions, which lead to a substantial improvement of the paper. This research was supported in part by grants from HKSAR-RGC-GRF 400306 and 400408, the National Science Foundation of China (10801118) and the Ph.D. Programs Foundations of the Ministry of Education of China (200803351094).

**Appendix**

Let

\[
S_n(s) = \frac{1}{a_n} \sum_{t=1}^{[ns]} \eta_t, \quad T_n(s) = \frac{1}{a_n} Y_{[ns]}, \quad W_n(\tau, s) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[ns]} \varphi_\tau(\varepsilon_t - \beta_0(\tau))
\]

where \( \varphi_\tau(x) = \tau - I(x < 0) \). To prove Theorems 2.1–2.3 and Theorems 3.1 and 3.2, we need the following lemmas.

**Lemma A.1.** Under conditions \( H_1 \) and \( H_2 \), for \( \beta > 2/\alpha \), we have

\[
\begin{pmatrix} S_n(s_1) \\ W_n(\tau, s_2) \end{pmatrix} \Rightarrow J_1 \begin{pmatrix} Z_{a\beta}(s_1) \\ \sigma W(\tau, s_2) \end{pmatrix} \text{ on } D(0, 1) \times D(0, 1). \tag{A.1}
\]
Proof. Let \( \|X\| \) be the norm \((E|X|^2)^{1/2}\), \(e_0 = \sum^\ast_{j=\infty} c_{t-j} \eta_j \), \( \varepsilon_t = \sum^t_{j=\infty} c_{t-j} \eta_j \) and \( \mathcal{F}_t = \sigma\{e_s, s \leq t\} \). Let \( \{\eta'_j\} \) be an independent copy of \( \{\eta_j\} \) and \( G_t \) be the distribution of \( \varepsilon_t \). Since \( \sup_{x \in R} |f(x)| < C < \infty \), it follows that \( g_t(x) = G'_t(x) \) is also bounded by \( C \). This gives

\[
\sum_{t=1}^{\infty} \|E(\varphi_t(\varepsilon_t - \beta_0(\tau))|\mathcal{F}_t) - E(\varphi_t(\varepsilon_t - \beta_0(\tau))|\mathcal{F}_0)\|
\]

\[
= \sum_{t=1}^{\infty} \|E(G_t(\beta_0(\tau) - \varepsilon_t) - G_t(\beta_0(\tau) - \varepsilon_t)|\mathcal{F}_t)\|
\]

\[
\leq \sum_{t=1}^{\infty} \|G_t(\beta_0(\tau) - \varepsilon_t) - G_t(\beta_0(\tau) - \varepsilon_t)|\mathcal{F}_t)\|
\]

\[
\leq \sum_{t=1}^{\infty} \min\{C|c_{t-1}(\eta_1 - \eta'_1)|, 1\}
\]

\[
\leq \sum_{t=1}^{\infty} \left[ E(C|c_{t-1}(\eta_1 - \eta'_1)) \right]^{\min(\alpha', 2)} \]^{1/2}

\[
= O\left( \sum_{t=0}^{\infty} |c_{t}\alpha'/2 \right) < \infty \quad (A.2)
\]

for some \( \alpha' < \alpha \), where we have used the fact that \( [\min(1, |a|)^2 \leq |a|^{\min(2, \alpha')} \). It follows from \( (A.2) \) that

\[
\|E(W_n(\tau, 1)|\mathcal{F}_0)\|^2 = \left\| \sum_{k=\infty}^0 E(W_n(\tau, 1)|\mathcal{F}_k) - E(W_n(\tau, 1)|\mathcal{F}_{k-1}) \right\|^2
\]

\[
\leq \sum_{k=\infty}^0 \|E(W_n(\tau, 1)|\mathcal{F}_k) - E(W_n(\tau, 1)|\mathcal{F}_{k-1})\|^2
\]

\[
\leq \frac{1}{n} \sum_{k=\infty}^0 \left( \sum_{i=1}^n \|E(\varphi_t(\varepsilon_t - \beta_0(\tau))|\mathcal{F}_k) - E(\varphi_t(\varepsilon_t - \beta_0(\tau))|\mathcal{F}_{k-1})\| \right)^2
\]

\[
= \left( \frac{1}{n} \right) O\left( \sum_{k=\infty}^0 \sum_{i=1}^n \|E(\varphi_t(\varepsilon_{t+k+1} - \beta_0(\tau))|\mathcal{F}_k) - E(\varphi_t(\varepsilon_{t+k+1} - \beta_0(\tau))|\mathcal{F}_{k-1})\| \right)
\]

\[
= o(1). \quad (A.3)
\]

Since \( \{\varphi_t(\varepsilon_t - \beta_0(\tau))\} \) is a stationary process with \( E\varphi_t^2(\varepsilon_t - \beta_0(\tau)) < \infty \), from \( (A.2) \) and \( (A.3) \) and Theorem 1 of [34] (see also [35]), it follows that

\[
W_n(\tau, 1) \longrightarrow_d N(0, \sigma^2). \quad (A.4)
\]

By \( (A.4) \) and a standard argument, we have

\[
W_n(\tau, s) \longrightarrow_d \sigma W(\tau, s) \quad \text{in } D(0, 1).
\]

Thus, the marginal distributions of \( S_n(\cdot) \) and \( W_n(\cdot) \) are \( Z_\alpha(\cdot) \) and \( \sigma W(\tau, \cdot) \) respectively. Following the argument of [29], we have the conclusion as desired. \( \square \)
Lemma A.2. If $1 < \beta < 2/\alpha$ and $1 < \alpha < 2$, there exists a $\nu > 0$ such that for any $\mu > 0$ and for any $0 \leq t_1 < t_2 \leq 1$,

$$
P \left\{ \frac{1}{n^{1/(\alpha \beta)}} \left| \sum_{i=[nt_1]}^{[nt_2]} v_i \left( \varphi_i (\xi_i - \beta_0(\tau)) + \zeta_i(\eta_i, \beta_0(\tau)) \right) \right| \geq \mu \right\} \leq C_{12} (t_2 - t_1)n^{-\nu}, \quad (A.5)
$$

and for any $0 \leq t \leq 1$,

$$
P \left\{ a_n^{-1/\beta} \left| \sum_{i=1}^{[nt]} \left( \zeta_i^0(\eta_i, \beta_0(\tau)) - \zeta(\eta_i, \beta_0(\tau)) \right) \right| \geq \mu \right\} \leq C_{12} \mu^{-\alpha} n^{-(\alpha(\beta-1)+1-1/\beta)}, \quad (A.6)
$$

where $F(x)$ is the distribution of $\xi_0$, \( \{v_i, 1 \leq i \leq n\} \) is a non-random real-valued sequence with $\max_{1 \leq i \leq n} |v_i| = O(\log n)$ and

$$
\zeta(\eta_i, \beta_0(\tau)) = \sum_{j=1}^{\infty} (F(\beta_0(\tau) - c_j \xi_i) - EF(\beta_0(\tau) - c_j \xi_i)),
$$

$$
\zeta^0(\eta_i, \beta_0(\tau)) = \sum_{j=1}^{\infty} (F(\beta_0(\tau) - c_j \xi_{i-j}) - EF(\beta_0(\tau) - c_j \xi_{i-j})).
$$

Proof. Let $X_{i,j} = \sum_{l=0}^{j} c_l \xi_{i-l}, \overline{X}_{i,j} = \sum_{l=j+1}^{\infty} c_l \xi_{i-l}, F_j(\beta_0(\tau)) = P(X_{i,j} \leq \beta_0(\tau))$ and

$$
U_{i,j}(\beta_0(\tau)) = F_{j-1}(\beta_0(\tau) - \overline{X}_{i,j-1}) - F_j(\beta_0(\tau) - \overline{X}_{i,j}) - F(\beta_0(\tau) - c_j \xi_{i-j}) + EF(\beta_0(\tau) - c_j \xi_{i-j}).
$$

Then

$$
\sum_{i=[nt_1]}^{[nt_2]} v_i \left( I(\xi_i \leq \beta_0(\tau)) - \tau - \eta(\xi_i, \beta_0(\tau)) \right) = \sum_{i=[nt_1]}^{[nt_2]} \sum_{j=1}^{\infty} v_i U_{i,j}(\beta_0(\tau)) = \sum_{j=-\infty}^{[nt_2]-1} \sum_{i=[nt_1] \vee (j+1)}^{[nt_2]} v_i U_{i,j}(\beta_0(\tau)).
$$

Let $M_{[nt_2],j} = \{ \sum_{i=[nt_1] \vee (j+1)}^{[nt_2]} U_{i,j}(\beta_0(\tau)) \}$. Then $\{M_{[nt_2],j}\}$ is a martingale difference. By Bahr–Essen’s inequality for martingales, we have that for any $1 \leq v \leq 2$,

$$
E \left( \left| \sum_{j=-\infty}^{[nt_2]-1} M_{[nt_2],j} \right|^v \right) \leq 2 \sum_{j=-\infty}^{[nt_2]-1} E |M_{[nt_2],j}|^v \leq 2 \sum_{j=-\infty}^{[nt_2]-1} \left( \sum_{i=[nt_1] \vee (j+1)}^{[nt_2]} E^{1/v} (U_{i,j} - |U_{i,j}|)^v \right). \quad (A.7)
$$

Using (A.7) and a similar argument to that of Lemma 5.3 in [32] (see also Lemma 7 of [6]), we have (A.5) as desired. (A.6) can be shown by a similar argument of Lemma 5.2 in [32]. \hfill \Box

Lemma A.3. Assume conditions $H_1$ and $H_2$ hold. Then,

$$
\int_0^1 T_n(s)ds \to^d \int_0^1 S(s)ds, \quad \int_0^1 T_n^2(s)ds \to^d \int_0^1 S^2(s)ds, \quad (A.8)
$$
\[
\int_{0}^{1} T_n(s-\epsilon)dW_n(\tau, s) \longrightarrow^d \int_{0}^{1} S(s)dW(\tau, s), \quad \text{(A.9)}
\]

where

\[
S(s) = \lambda \int_{0}^{s} e^{-\gamma(s-t)}dZ_\alpha(t) = \lambda( Z_\alpha(s) - \gamma \int_{0}^{s} e^{-\gamma(s-t)} Z_\alpha(t)dt)
\]

and the convergence in (A.8) and (A.9) holds jointly.

**Proof.** The proof of \( \int_{0}^{1} T_n(s)ds \longrightarrow^d \int_{0}^{1} S(s)ds \) is similar to that of \( \int_{0}^{1} T_n^2(s)ds \longrightarrow^d \int_{0}^{1} S^2(s)ds \), so we only give the proof of the latter. Note that

\[
T_n(s) = Y_{[ns]}(s)/a_n = \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j} \varepsilon_j/a_n,
\]

and

\[
\varepsilon_i = \lambda \eta + (\varepsilon_i - \lambda \eta_i).
\]

In view of (A.10) and (A.11), we have

\[
T_n(s) = \frac{\lambda}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j} \eta_j + \frac{1}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j}(\varepsilon_j - \lambda \eta_j).
\]

This yields

\[
\int_{0}^{1} T_n^2(s)ds = \int_{0}^{1} \left( \frac{\lambda}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j} \eta_j + \frac{1}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j}(\varepsilon_j - \lambda \eta_j) \right)^2 ds
\]

\[
= \lambda^2 \int_{0}^{1} \left( \frac{1}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j} \eta_j \right)^2 ds + \int_{0}^{1} \left( \frac{1}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j}(\varepsilon_j - \lambda \eta_j) \right)^2 ds
\]

\[
+ 2\lambda \int_{0}^{1} \left( \frac{1}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j} \eta_j \right) \left( \frac{1}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j}(\varepsilon_j - \lambda \eta_j) \right) ds
\]

\[
=: I_1 + I_2 + I_3.
\]

By Lemma 2 of [8], we have

\[
\frac{1}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j} \eta_j \Rightarrow J_t Z_\alpha(s) - \gamma \int_{0}^{s} e^{-\gamma(s-t)} Z_\alpha(t)dt \quad \text{in } D(0, 1).
\]

By means of the continuous mapping theorem, \( I_1 \longrightarrow^d \int_{0}^{1} S^2(t)dt \). Note that \( I_3 \leq 2(I_1 I_2)^{1/2} \). It is enough to show \( I_2 \longrightarrow^p 0 \). Observe that

\[
\frac{1}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j}(\varepsilon_j - \lambda \eta_j) = \frac{1}{a_n} \sum_{j=1}^{[ns]} (\varepsilon_j - \lambda \eta_j) - \frac{\gamma}{na_n} \sum_{j=1}^{[ns]} \sum_{k=1}^{[ns]} \gamma_n^{[ns]-j}(\varepsilon_k - \lambda \eta_k)
\]

\[
= \frac{1}{a_n} \left( \sum_{j=1}^{[ns]} \sum_{i=i-j+1}^{\infty} c_i \eta_j - \sum_{j=-\infty}^{0} \sum_{i=1-j}^{[ns]} c_i \eta_j \right) - \frac{\gamma}{na_n}
\]
When $\alpha > 1$, since
\[
\sup_{1 \leq m \leq n} \mathbb{E} \left\{ \frac{1}{a_n} \sum_{j=1}^{m} \sum_{i=m-j+1}^{\infty} c_i \eta_j - \sum_{j=-\infty}^{m-j} \sum_{i=1-j}^{\infty} c_i \eta_j \right\} \to 0,
\]
(A.16)

it follows that
\[
\frac{1}{n} \sum_{i=1}^{n} \left| \frac{1}{a_n} \sum_{j=1}^{i} \gamma_n^{i-j} (\varepsilon_j - \lambda \eta_j) \right| \to P 0.
\]
(A.17)

On the other hand, by the continuous mapping and Theorem 2 of [1],
\[
\sup_{0 \leq i \leq n} \left| \frac{1}{a_n} \sum_{j=1}^{i} \gamma_n^{i-j} (\varepsilon_j - \lambda \eta_j) \right| \\
\leq \sup_{1 \leq i \leq n} \left| \frac{1}{a_n} \sum_{j=1}^{i} \gamma_n^{i-j} \varepsilon_j \right| + \sup_{1 \leq i \leq n} \left| \frac{1}{a_n} \sum_{j=1}^{i} \gamma_n^{i-j} \lambda \eta_j \right| = O_P(1).
\]

Thus,
\[
I_2 = \int_0^1 \left( \frac{1}{a_n} \sum_{j=1}^{[ns]} \gamma_n^{[ns]-j} (\varepsilon_j - \lambda \eta_j) \right)^2 \, ds \\
= \frac{1}{n} \sum_{t=1}^{n} \left( \frac{1}{a_n} \sum_{j=1}^{l} \gamma_n^{l-j} (\varepsilon_j - \lambda \eta_j) \right)^2 \to P 0.
\]
(A.18)

When $\alpha \leq 1$, by the ‘Beveridge–Nelson’ decomposition of $\{\varepsilon_t\}$, we have
\[
\varepsilon_t - \lambda \eta_t = \varepsilon_{t-1} - \varepsilon_t
\]
(A.19)

where $\varepsilon_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}$ and $\bar{c}_j = \sum_{j=i}^{\infty} c_j$. Eq. (A.8) can be shown as in Theorem 2.1 of [27]. The proof of (A.8) is complete.

Next, we adopt an idea of [18] to show (A.9). Let $\{\xi_t\}$ be stationary ergodic martingale differences with respect to $\sigma$-fields generated by $\{\eta_k, k \leq t\}$ for $t = 1, 2, \ldots, n$ with $\mathbb{E} \xi_t^2 < \infty$ and $Z_t$ be a stationary process with $\mathbb{E} Z_t^2 < \infty$ such that
\[
\varphi_t(\eta_t) = \xi_t + Z_t - Z_{t+1}, \quad t = 1, 2, \ldots, n.
\]

Let $W^*_n(\tau, s) = \frac{1}{\sqrt{n}} \sum_{j=1}^{[ns]} \xi_t$. Then
\[
\left| \int_0^1 T_n(s-\tau) dW_n(\tau, s) - \int_0^1 T_n(s-\tau) dW^*_n(\tau, s) \right| \\
= |T_n(1)(W^*_n(\tau, 1) - W_n(\tau, 1)) - \int_0^1 (W^*_n(\tau, s) - W_n(\tau, s)) dT_n(s)| \\
\leq |T_n(1)| \sup_{0 \leq s \leq 1} |W^*_n(\tau, s) - W_n(\tau, s)| + \left| \frac{1}{a_n} \sum_{i=1}^{n} (Z_{t+1} - Z_t) \varepsilon_t \right|.
and bigger than $C$. Since $\varepsilon < \alpha$ cases:

Next, we use Karamata’s theorem to show that $I$ and $0 = 52 \leq I_{51} \leq I_{52} \leq I_{53} = 51 = 52$

Thus $I_4 = o_p(1)$. For any $\zeta > 0$, define

$$I_{51} = P \left\{ \sum_{i=1}^{n} \left( \frac{1}{\sqrt{n}} (Z_{t+1} - Z_1) I(|Z_{t+1} - Z_1| > \sqrt{n}\delta) \right) \frac{\varepsilon_i}{a_n} > \zeta \right\},$$

$$I_{52} = P \left\{ \sum_{i=1}^{n} \left( \frac{1}{\sqrt{n}} (Z_{t+1} - Z_1) I(|Z_{t+1} - Z_1| \leq \sqrt{n}\delta) \right) \frac{\varepsilon_i}{a_n} I(|\varepsilon_i| \geq Ma_n) > \zeta \right\}$$

and

$$I_{53} = P \left\{ \sum_{i=1}^{n} \left( \frac{1}{\sqrt{n}} (Z_{t+1} - Z_1) I(|Z_{t+1} - Z_1| \leq \sqrt{n}\delta) \right) \frac{\varepsilon_i}{a_n} I(|\varepsilon_i| < Ma_n) > \zeta \right\}.$$

Since $E Z_i^2 < \infty$, it follows that

$$I_{51} \leq P \left( \max_{1 \leq t \leq n} |Z_{t+1} - Z_1| > \sqrt{n}\delta \right) \to 0.$$

Next, we use Karamata’s theorem to show that $I_{52} \to 0$. For this quantity, we split $\alpha$ into two cases: $\alpha > 1$ and $0 < \alpha \leq 1$. For $\alpha > 1$, by Karamata’s theorem, we have

$$I_{52} \leq P \left\{ \sum_{i=1}^{n} \delta |\varepsilon_i| I(|\varepsilon_i| \geq Ma_n) > \zeta a_n \right\}$$

$$\leq \frac{\delta}{\zeta a_n} \sum_{i=1}^{n} \mathbb{E} \left[ |\varepsilon_i| I(|\varepsilon_i| \geq Ma_n) \right]$$

$$\leq C \delta \left( \frac{\alpha}{\alpha - 1} \right). \quad (A.22)$$

For $0 < \alpha \leq 1$ and $\nu < \alpha$, we have

$$I_{52} \leq \frac{\delta^\nu}{(\zeta a_n)^\nu} \sum_{i=1}^{n} \mathbb{E}[|\varepsilon_i|^\nu I(|\varepsilon_i| \geq (Ma_n)^\nu)]. \quad (A.23)$$

Since $\varepsilon_i^\nu$ has index $\alpha/\nu$, similar to the argument of (A.22), the right-hand side of (A.23) is no bigger than $C \delta^\nu \alpha / (\alpha - \nu)$. Thus, by taking $\delta \to 0$ small enough in (A.22) and (A.23), we have $I_{52} \to 0$. Finally, using Karamata’s theorem again, we have

$$I_{53} \leq \frac{1}{\zeta} \mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=1}^{n} (Z_{t+1} - Z_1)^2 \right)^{1/2} \left( \frac{1}{\alpha^2} \sum_{i=1}^{n} \varepsilon_i^2 I(|\varepsilon_i| < Ma_n) \right)^{1/2} \right]$$

$$\leq \frac{1}{\zeta} \left[ \mathbb{E} \left( \frac{1}{n} \sum_{i=1}^{n} (Z_{t+1} - Z_1)^2 \right)^{1/2} \left( \mathbb{E} \left( \frac{1}{\alpha^2} \sum_{i=1}^{n} \varepsilon_i^2 I(|\varepsilon_i| < Ma_n) \right) \right)^{1/2} \right]$$

$$\to 0.$$
by first letting \( n \to \infty \) and then \( M \to 0 \). Thus \( I_5 \to^p 0 \). Note that since \( 1 - \gamma_n = \gamma/n \), similar to \( I_5, I_6 \) can be shown converging to zero in probability. Combining these with \( I_4 \to^p 0 \) and \((A.20)\) yields \( \int_0^1 T_n d(W_n - W_n^\#) \to^p 0 \). We can therefore work with \( \int T_n dW_n^\# \) instead of \( \int T_n dW_n \). Since \( \sup_{1 \leq i \leq n} |\varphi_n (\varepsilon_i - \beta_0(\tau))| \leq 2 \), it follows that \( W_n^\# \) is a martingale with bounded jumps. By \( T_n(s) = \lambda \sum_{j=1}^{[ns]} -j \eta_j/a_n + o_p(1) \), we have

\[
(T_n(s), W_n^\#(s)) \overset{f.d.d.}{\longrightarrow} (S(s), W(\tau, s)).
\]

Since \( T_n(s) \Rightarrow M_1 S(s) \) (see Theorem 2 of [1]) and \( W_n^\#(s) \Rightarrow J_1 \sigma_W(\tau, s) \) on \([0, 1]\) by Lemma A.1 and \((A.21)\). It follows from Theorem 3 of [15] that

\[
\int_0^1 T_n(s-) dW_n^\#(\tau, s) \overset{d}{\longrightarrow} \sigma \int_0^1 S(s-) dW(\tau, s).
\]

This gives \((A.9)\). The joint convergence follows from the joint weak convergence of \( \{S_n(\cdot)\} \) and \( \{W_n(\tau, \cdot)\} \). The proof of Lemma A.4 is complete. \( \square \)

**Lemma A.4.** Let \( L_n(s) = a_n^{-1/\beta} \sum_{i=1}^{[ns]} \xi(\eta_i, \beta_0(\tau)), \quad S_n(s) = \sum_{i=1}^{[ns]} \eta_i/a_n \) and \( W_n(\tau, s) = \frac{1}{\alpha \beta} \sum_{i=1}^{[ns]} \varphi_n(\varepsilon_i - \beta_0(\tau)) \). When \( 1 + \sqrt{1 - 1/\alpha} < \beta < \alpha/2 \), then under condition \( H_1 \) and \( H_2 \), we have

\[
(S_n(s), L_n(s)) \Rightarrow J_1 (Z_\alpha(s), L_{\alpha \beta}(s)) \quad \text{in } D[0, 1] \times D[0, 1] \quad \text{(A.24)}
\]

and

\[
\left( \int_0^1 T_n(s) ds, \int_0^1 T_n^2(s) ds, \int_0^1 T_n(s-) dW_n' (s) \right) \overset{d}{\longrightarrow} \left( \int_0^1 S(s) ds, \int_0^1 S^2(s) ds, \int_0^1 S(s) dL_{\alpha \beta}(s) \right). \quad \text{(A.25)}
\]

**Proof.** For the proof of \((A.24)\), let \( \kappa_i = (a_n^{-1} \eta_i, a_n^{-1/\beta} \xi(\eta_i, \beta_0(\tau))), \quad 1 \leq i \leq n \). Then \( \{\kappa_i\} \) is a sequence of i.i.d random vectors. Since \( \eta_i \) belongs to the domain of attraction of a stable law with index \( \alpha \), it follows that \( \xi(\eta_i, \beta_0(\tau)) \) belongs to the domains of attraction of a stable law with index \( \alpha \beta \) (see [32]). By a similar argument of Theorem 1 in [25], it can be shown that \( \kappa_i \) belongs to a generalized domain of an operator stable law on \( R^2 \). That is, there exists a stable vector process \( \kappa(t) = (\kappa_1(t), \kappa_2(t)) \) with \( \kappa_1(t) = d Z_\alpha(t) \), and \( \kappa_2(t) = d L_{\alpha \beta}(t) \) such that

\[
Z_n(t) \Rightarrow J_1 \kappa(t),
\]

where \( Z_n(t) = \sum_{i=1}^{[nt]} \kappa_i \). This gives \((A.24)\).

For \((A.24)\), put \( T_n'(s) = \lambda \sum_{i=1}^{[ns]} -j \eta_i/a_n \). By a similar argument of [26], we have \( T_n'(s) \) is a semi-martingale satisfying the so-called UT conditions defined in Kurtz and Protter [21] (see also [16]). Further, by \((A.24)\) and the continuous mapping theorem (see Lemma 2 of [8]), \( T_n'(s) \Rightarrow J_1 S(s) \) on \([0, 1]\). Thus, by Theorem 2.7 of [21] (see also [14]), it follows that

\[
(T_n'(s), L_n(1)T_n'(1) - \int_0^1 L_n(s-) dT_n'(s)) \Rightarrow J_1 (S(s), \int_0^1 S(s) dL_{\alpha \beta}(s)).
\]

Therefore, for the proof of \((A.24)\), it is enough to show

\[
\left( \int_0^1 T_n(s) ds, \int_0^1 T_n^2(s) ds, \int_0^1 T_n(s-) dW_n'(s) \right)
\]
Therefore, by (A.27), it follows that

\[ \sum_{t=1}^{n} Y_{t-1} \varphi_{\tau}(\varepsilon_{t} - \beta_{0}(\tau)) = \frac{\lambda}{a_{n} n^{1/(a \beta)}} \sum_{t=1}^{n} \left( \sum_{j=1}^{t-1} \gamma_{n}^{t-1-j} \eta_{j} \right) \varphi_{\tau}(\varepsilon_{t} - \beta_{0}(\tau)) + o_{p}(1) \]

\[ = \frac{\lambda}{a_{n} n^{1/(a \beta)}} \left( \sum_{t=1}^{n} \gamma_{n}^{n-1-t} \eta_{t} \right) \left( \sum_{t=1}^{n} \varphi_{\tau}(\varepsilon_{t} - \beta_{0}(\tau)) \right) \]

\[ - \frac{\lambda}{a_{n} n^{1/(a \beta)}} \sum_{t=2}^{n} \left( \sum_{j=1}^{t-1} \varphi_{\tau}(\varepsilon_{j} - \beta_{0}(\tau)) \right) \eta_{t} + o_{p}(1). \]  

By Lemma A.2, we have \( \sum_{t=1}^{n} \varphi_{\tau}(\varepsilon_{t} - \beta_{0}(\tau)) = L_{n}(1) + o_{p}(1) \) and when \( 1 + \sqrt{1-1/\alpha} < \beta < 2/\alpha \),

\[ E \left| \frac{1}{a_{n} n^{1/(a \beta)}} \sum_{i=1}^{n} \sum_{j=1}^{t} \left[ \zeta(\eta_{j}, \beta_{0}(\tau)) - \zeta'(\eta_{j}, \beta_{0}(\tau)) \right] \eta_{i} \right| \]

\[ \leq \frac{\lambda}{a_{n} n^{1/(a \beta)}} \sum_{i=1}^{n} \sum_{j=1}^{t} \left| \zeta(\eta_{j}, \beta_{0}(\tau)) \right| \eta_{i} = o(1). \]
Further,
\[
\frac{1}{n^{1/(\alpha\beta)}} \sum_{i=1}^{[nt]} \left[ \phi \left( \varepsilon_j - \beta_0(\tau) \right) - \zeta'(\eta_j, \beta_0(\tau)) \right] \Rightarrow J_1 0.
\]

Invoking Theorem 2.7 of [21] yields
\[
\lambda a_n n^{1/(\alpha\beta)} \sum_{t=1}^{n-1} \sum_{j=1}^{t-1} \left[ \phi \left( \varepsilon_j - \beta_0(\tau) \right) - \zeta'(\eta_j, \beta_0(\tau)) \right] \eta_t \longrightarrow d 0.
\]

Thus, by (A.28), we have (A.26) and the proof of Lemma 5.5 is completed. □

References


