Asymptotic theory for curve-crossing analysis

Zhibiao Zhao*, Wei Biao Wu

Department of Statistics, The University of Chicago, United States

Received 12 April 2006; received in revised form 22 October 2006; accepted 22 October 2006
Available online 14 November 2006

This paper is dedicated to Professor Sándor Csörgö on the occasion of his 60th birthday

Abstract

We consider asymptotic properties of curve-crossing counts of linear processes and nonlinear time series by curves. Central limit theorems are obtained for curve-crossing counts of short-range dependent processes. For the long-range dependence case, the asymptotic distributions are shown to be either multiple Wiener–Itô integrals or integrals with respect to stable Lévy processes, depending on the heaviness of tails of the underlying processes.

© 2006 Elsevier B.V. All rights reserved.
Keywords: Central limit theorem; Curve-crossing; Linear processes; Multiple Wiener–Itô integral; Non-central limit theorem; Nonlinear time series

1. Introduction

Let \((X_i)_{i \in \mathbb{Z}}\) be a random process and \(\mu = (\mu_i)_{i \in \mathbb{Z}}\) be a sequence of real numbers. Define

\[
C_n(\mu) = \sum_{i=1}^{n} 1_{(X_{i-1} - \mu_{i-1})(X_i - \mu_i) \leq 0},
\]

where \(1_A = 1\) or \(0\) depending on whether event \(A\) occurs or not. The quantity \(C_n(\mu)\) is the number of \(i\), \(0 \leq i \leq n\), such that \(X_{i-1} \leq \mu_{i-1}\) and \(X_i \geq \mu_i\) or \(X_{i-1} > \mu_{i-1}\) and \(X_i \leq \mu_i\). In other words, \(C_n(\mu)\) is the number of times that the process \((X_i)_{i=0}^{n}\) crosses the sequence \((\mu_i)_{i=0}^{n}\). Special cases of \(\mu\) include (i) \(\mu_i = m_1(i/n)\) for some function \(m_1\) on \([0, 1]\) and (ii) \(\mu_i = m_2(i)\)

* Corresponding address: Department of Statistics, University of Chicago, 5734 S University Ave, 60637 Chicago, IL, United States. Tel.: +1 773 7028330; fax: +1 773 7029810.

E-mail address: zhao@galton.uchicago.edu (Z. Zhao).

0304-4149/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.spa.2006.10.010
for some function $m_2$ on $\mathbb{R}$. These two formulations of $\mu$ have different ranges of applicability. The former is a reasonable choice if one aims to capture the feature that the sequence being crossed is smoother than the underlying stochastic part. The latter can be used for curves such as $m_2(u) = \cos(u)$ that oscillate more heavily. See Fan and Yao [10, pp. 225–226] for further details on these two formulations.

When $\mu$ is a constant sequence, $C_n(\mu)$ is the level-crossing count. In particular, $C_n(0)$ is the zero-crossing count. Level-crossing analysis has been widely used in engineering, physics, speech recognition and other fields; see [11, 19–21, 30, 35]. For example, if $X_i = \sum_{j=1}^k [A_j \cos(i \omega_j) + B_j \sin(i \omega_j)] + \epsilon_i$ for independent normal random variables $A_j, B_j$ and $\epsilon_i$, then one can estimate the frequencies $\omega_j$ by considering the zero-crossing counts of $X_i$ and its differing sequences; see [21]. In the literature of level-crossing analysis, attention has been mainly focused on the mean number of crossings of processes at fixed levels. For example, Benzaquen and Cabana [1], Bulinskaya [2], Rice [31] and Ylvisaker [46] considered stationary Gaussian processes with continuous sample paths. In [21], asymptotic means and variances are obtained for level-crossing counts by Gaussian processes. More recently, Shimizu and Tanaka [36] obtained the expected number of two-level level-crossing counts by stationary ellipsoidal processes and Kratz and León [22] established an Hermite expansion for level-crossing counts by stationary Gaussian processes. For statistical inference including confidence intervals construction and hypothesis testing, however, it is desirable to have an asymptotic distributional theory for $C_n(\mu)$. In the context of level-crossing counts, Cuzick [4] and Kedem [21] obtained central limit theorems (CLT) for Gaussian processes and Wu [41, 42] considered linear processes.

This paper aims at establishing an asymptotic distributional theory for $C_n(\mu)$ for non-constant sequences $\mu$. Such results can provide an inferential theory for statistical analysis using the level-crossing method. For the number of curve-crossings by stationary continuous-time Gaussian processes, Slud [37] gave a multiple Wiener–Itô integral representation; Kratz and León [23] derived an Hermite expansion and applied their results to the specular points problem proposed in [26]. These results cannot be directly applied to discrete time non-Gaussian processes. Here we shall consider linear processes with finite or infinite variances and some popular nonlinear time series. For processes with infinite variances the classical spectral and time domain approaches seem inappropriate and the curve-crossing analysis provides a useful alternative.

In this paper we consider stationary processes of the form

$$X_n = g(\varepsilon_n, \varepsilon_{n-1}, \ldots),$$

(2)

where $g$ is a measurable function and $\varepsilon_i, i \in \mathbb{Z}$, are independent and identically distributed (iid) random variables. The framework (2) does not seem to be overly restrictive. The Wiener–Rosenblatt conjecture states that, for every stationary and ergodic process $(X_k)_{k \in \mathbb{Z}}$, there exists a measurable function $g$ and iid $\varepsilon_i$ such that the distributional equality $(X_k)_{k \in \mathbb{Z}} \Rightarrow (g(\varepsilon_k, \varepsilon_{k-1}, \ldots))_{k \in \mathbb{Z}}$ holds; see [17, 33, 40].

We now introduce some notation. Denote by $F_X$ and $F$ the distribution functions of $X_0$ and $(X_0, X_1)$, respectively, and let $G(x, y) = \mathbb{P}[(X_0 - x)(X_1 - y) \leq 0]$. Assume throughout the paper that $F$ is continuous. Then $G(x, y) = F_X(x) + F_X(y) - 2F(x, y)$. Let $\mathcal{F}_n = (\varepsilon_n, \varepsilon_{n-1}, \ldots)$ be the shift process and $\mathcal{F}'_n = (\varepsilon_n, \ldots, \varepsilon_1, \varepsilon'_1, \varepsilon_{-1}, \ldots)$ a coupled process of $\mathcal{F}_n$, where $(\varepsilon'_i)_{i \in \mathbb{Z}}$ is an iid copy of $(\varepsilon_i)_{i \in \mathbb{Z}}$. For a random variable $Z$ write $Z \in L^q$, $q > 0$, if $\|Z\|_q := [\mathbb{E}(|Z|^q)]^{1/q} < \infty$. Define the detail projection $\mathcal{P}_k Z = \mathbb{E}(Z|\mathcal{F}_k) - \mathbb{E}(Z|\mathcal{F}_{k-1}), k \in \mathbb{Z}$. For $a, b \in \mathbb{R}$ let $a \wedge b = \min(a, b), a \vee b = \max(a, b)$ and $a^+ = a \vee 0$. For two real
sequences \((a_n)\) and \((b_n)\), write \(a_n = O(b_n)\) if there exists a constant \(c\) such that \(|a_n| \leq c|b_n|\) for large \(n\) and \(a_n = o(b_n)\) if \(\lim_{n \to \infty} a_n/b_n = 0\). For a real function \(g(x, y)\) and integers \(r \geq s \geq 0\), denote by \(g^{(r,s)}\) the partial derivative \(\partial^r g(x, y)/\partial x^s \partial y^{r-s}\) if it exists. Denote by \(\mathcal{C}^p = \{ g : \sup_{x,y \in \mathbb{R}} |g^{(r,s)}(x, y)| < \infty, 0 \leq s \leq r \leq p \}\) the set of bivariate functions with bounded partial derivatives up to order \(p\).

The rest of the paper is structured as follows. We present main results in Section 2, where applications are made to nonlinear time series, long-memory linear processes and heavy-tailed processes which are widely used in practice. The limiting distributions are shown to be normal or non-normal, depending on the heaviness of the tails and the strength of dependence of the processes. Proofs are given in Section 3.

2. Main results

Asymptotic behavior of \(C_n(\mu)\) should certainly depend on the underlying process \((X_i)\) and the sequence \((\mu_i)\). A general central limit theorem is given below.

**Theorem 1.** Assume that, for each fixed integer \(k\), the limit
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \text{cov}[1_{(X_0 - \mu_{i-1})(X_1 - \mu_i) \leq 0}, 1_{(X_k - \mu_{i+k-1})(X_{k+1} - \mu_{i+k}) \leq 0}] = \gamma_k \quad \text{(say)}
\] (3)
exists. Further assume that \(F(\cdot, \cdot)\) is continuous and that
\[
\sum_{i=0}^{\infty} \delta_i < \infty, \quad \text{where} \quad \delta_i = \sup_{x,y \in \mathbb{R}} \| P_0 1_{(X_{i-1} - x)(X_{i} - y) \leq 0} \|_2.
\] (4)

Then \(\sigma^2 := \sum_{k \in \mathbb{Z}} \gamma_k < \infty, \|C_n(\mu) - EC_n(\mu)\|_2^2/n \to \sigma^2\) as \(n \to \infty\) and
\[
\frac{C_n(\mu) - EC_n(\mu)}{\sqrt{n}} \Rightarrow N(0, \sigma^2).
\] (5)

In Theorem 1, (4) is basically a short-range dependence condition. In Sections 2.1 and 2.2 we shall verify (4) for nonlinear time series and short-range dependent linear processes. For long-range dependent processes, (4) is violated and the limiting behavior of \(C_n(\mu)\) has more detailed structures; see Section 2.3. Condition (3) is imposed in such a way that it can allow various forms of \(\mu\); see Examples 1 and 2 below.

**Example 1.** For the level-crossing case with \(\mu_i = c\), we have \(\gamma_k = \text{cov}(V_0, V_k)\), where \(V_i = 1_{(X_{i} - c)(X_{i+1} - c) \leq 0}\). If \(\mu_i = m_1(i/n)\), where \(m_1\) is a piecewise continuous function on \([0, 1]\) with finitely many jumps, elementary calculations show that (3) holds with
\[
\gamma_k = \int_0^1 \text{cov}(1_{[X_0 - m_1(t)][X_1 - m_1(t)], 1_{[X_k - m_1(t)][X_{k+1} - m_1(t)] \leq 0]} dt.
\] (6)

**Example 2.** Let \(\mu_i = \cos(2\pi \omega i)\) for some \(\omega \in \mathbb{R}\). If \(\omega = a/b\) is a rational number with \(a, b \in \mathbb{N}\), then elementary calculations show that (3) holds with
\[
\gamma_k = \frac{1}{b} \sum_{j=1}^{b} \text{cov}(V_{0,j}, V_{k,j}), \quad \text{where} \quad V_{i,j} = 1_{[X_i - \cos(2\pi \omega(i+j-1))][X_{i+1} \cos(2\pi \omega(i+j)) \leq 0].
\] (7)
If \( \omega \) is irrational, then the sequence \((i \omega)_i \in \mathbb{N}\), is uniformly distributed modulo 1 [24]. Hence for \( Z_k(u) = X_k - \cos[2\pi(u + k\omega)] \), by Theorem 1.1 in [24] we have
\[
\gamma_k = \int_0^1 \text{cov}(1Z_0(u)Z_1(u) \leq 0, 1Z_k(u)Z_{k+1}(u) \leq 0) \, du. \tag{8}
\]

**Remark 1.** To apply Theorem 1, one needs to estimate the limiting variance \( \sigma^2 \). For stationary processes the latter problem is closely related to the long-run variance estimation. Unfortunately, due to the non-stationarity of \( Z_i = 1(X_{i-1} - \mu_{i-1})(X_i - \mu_i) \leq 0 - G(\mu_{i-1}, \mu_i) \), the estimation of \( \sigma^2 \) is not straightforward. Let \( \hat{G} \) be an estimate of \( G \). For example, we can take \( \hat{G}(x, y) = n^{-1} \sum_{i=1}^n 1(X_{i-1} - x)(X_i - y) \leq 0 \). As the proof of Theorem 1 (cf. Section 3.1) suggests, one can propose the truncated estimate
\[
\hat{\sigma}_k^2 = \frac{1}{n} \sum_{i,j=1}^n \hat{Z}_i \hat{Z}_j 1_{|i-j| \leq k}, \quad \text{where} \quad \hat{Z}_i = 1(X_{i-1} - \mu_{i-1})(X_i - \mu_i) \leq 0 - \hat{G}(\mu_{i-1}, \mu_i).
\]
Here \( k = k_n \) satisfies \( k \to \infty \) and \( k/n \to 0 \). We conjecture that, under (3) and (4), \( \hat{\sigma}_k^2 \to \sigma^2 \) in probability. It is also unclear how to choose optimal \( k \). \( \diamond \)

**Remark 2.** Recall that \((\varepsilon_i^1)_{i \in \mathbb{Z}}\) is an iid copy of \((\varepsilon_i)_{i \in \mathbb{Z}}\). Let \( h \) be a measurable function such that \( Y_n = h(\varepsilon_n, \varepsilon_{n-1}, \ldots) \in \mathcal{L}^q \) for some \( q \geq 1 \). Then (4) suggests a way of defining dependence in the sequence \((Y_n)_{n \in \mathbb{Z}}\): \( \delta_i(q) = \| \mathcal{P}_0 Y_i \|_q, i \geq 0 \); see [44]. Note that \( \mathbb{E}(Y_i|\mathcal{F}_{-1}) = \mathbb{E}(Y_i'|\mathcal{F}_0) \), where \( Y_i' = h(\mathcal{F}_i') \). By Jensen’s inequality, \( \delta_i(q) \leq \| Y_i - Y_i' \|_q =: \omega_i,q \) (say). Dedecker and Prieur [5] considered the following coupling coefficients:
\[
\omega_i,q = \| Y_i - Y_i'^* \|_q, \quad \text{where} \quad Y_i'^* = h(\varepsilon_i, \ldots, \varepsilon_1, \varepsilon_i, \varepsilon_i', \varepsilon_i', \ldots).
\tag{9}
\]
The difference between the two coupling versions \( Y_i' \) and \( Y_i'^* \) of \( Y_i \) is that the former replaces \( \varepsilon_0' \) with \( \varepsilon_i' \) while the latter replaces \( \varepsilon_i \) with \( \varepsilon_i' \) for all \( i \leq 0 \). By triangle inequality, \( \omega_i,q \leq \sum_{j=1}^\infty \omega_{j,q} \). See [5] for more details. It seems difficult to use their dependence coefficients to obtain results derived in the current paper, especially when the process exhibit long-range dependence. \( \diamond \)

### 2.1. Nonlinear time series

Let \( \varepsilon_i \) be iid random variables and define recursively
\[
X_n = R_{\varepsilon_n}(X_{n-1}), \tag{10}
\]
where \( R_{\varepsilon}(\cdot) = R(\cdot, \varepsilon) \) is a measurable random map. Many popular nonlinear time series models are of the form (10), for example, threshold autoregressive model [40]: \( X_n = a(X_{n-1} \lor 0) + b(X_{n-1} \land 0) + \varepsilon_n \), autoregressive conditional heteroscedasticity (ARCH) model [9]: \( X_n = \varepsilon_n \sqrt{a^2 + b^2 X_{n-1}^2} \), random coefficient model [29]: \( X_n = (a + b \varepsilon_n)X_{n-1} + \varepsilon_n \) among others. For properties of iterated random functions see [25,28,7,8].

**Proposition 1.** Assume that there exists \( x_0 \) and \( \alpha > 0 \) such that
\[
R_{\varepsilon_0}(x_0) \in \mathcal{L}^\alpha \quad \text{and} \quad \rho := \sup_{x \neq x'} \frac{\| R_{\varepsilon_0}(x) - R_{\varepsilon_0}(x') \|_\alpha}{|x - x'|} < 1. \tag{11}
\]
Further assume that, for some \( p > 2 \),
\[
\sup_{x \in \mathbb{R}} |F_X(x + t) - F_X(x)| = O((\log |t|)^{-p}) \quad \text{as} \quad t \to 0. \tag{12}
\]
Then (4) holds.

**Proof.** Recall \( \mathcal{F}_n = (\varepsilon_n, \varepsilon_{n-1}, \ldots) \) and \( \mathcal{F}_n' = (\varepsilon_n, \ldots, \varepsilon_1, \varepsilon_0', \varepsilon_{-1}, \ldots) \), where \( (\varepsilon_i')_{i \in \mathbb{Z}} \) is an iid copy of \( (\varepsilon_i)_{i \in \mathbb{Z}} \). Define the coupled process \( \mathcal{F}_n^* = (\varepsilon_n, \ldots, \varepsilon_1, \varepsilon_0', \varepsilon_{-1}, \ldots) \) with \( \varepsilon_i, i \leq 0 \), in \( \mathcal{F}_n \) replaced by \( \varepsilon_i' \). By Theorem 2 in [45], if (11) is satisfied, then (10) has a unique stationary solution of (2) with the function \( g \) satisfying \( X_n = g(\mathcal{F}_n) \in \mathcal{L}^\alpha \) and
\[
\|X_n - g(\mathcal{F}_n^*)\|_{\alpha} = O(\rho^n). \tag{13}
\]
Let \( X_i' = g(\mathcal{F}_i'), X_i^* = g(\mathcal{F}_i^*) \) and \( \kappa_i = \sup_x \|1_{X_i \leq x} - 1_{X_i'} \|_2 \). Since
\[
\mathbb{E}[1_{(X_{i-1} - x)(X_i - y) \leq 0}|\mathcal{F}_{i-1}] = \mathbb{E}[1_{(X_{i-1} - x)(X_i' - y) \leq 0}|\mathcal{F}_{i-1}] = \mathbb{E}[1_{(X_{i-1} - x)(X_i - y) \leq 0}|\mathcal{F}_0],
\]
by Jensen’s inequality,
\[
\|P_01_{(X_{i-1} - x)(X_i - y) \leq 0}\|_2 \leq \|1_{(X_{i-1} - x)(X_i - y) \leq 0} - 1_{(X_{i-1} - x)(X_i' - y) \leq 0}\|_2 \leq 2(\kappa_i + \kappa_{i-1}). \tag{14}
\]
Note that \( \|1_{X_i \leq x} - 1_{X_i^* \leq x}\|_2 \leq \|1_{X_i \leq x} - 1_{X_i^* \leq x}\|_2 + \|1_{X_i \leq x} - 1_{X_i^* \leq x}\|_2 \leq 2(\kappa_i + \kappa_{i-1}) \). Let \( \tau = \rho^{1/(2\alpha)} \). By (12) and (13),
\[
\mathbb{P}(X_i \leq x, X_i^* \geq x) \leq \mathbb{P}(X_i \leq x, X_i^* \geq x, |X_i^* - X_i| \leq \tau^i) + \mathbb{P}(|X_i^* - X_i| > \tau^i) \\
\leq \mathbb{P}(|X_i - x| \leq \tau^i) + \tau^{-\alpha i} \mathbb{E}|X_i^* - X_i|^\alpha = O(i^{-p} + \rho^{i/2}).
\]
So \( \sum_{i=0}^{\infty} \kappa_i < \infty \) and the proposition follows from (14). ∎

**Example 3.** Consider the ARCH(1) model \( X_n = \varepsilon_n \sqrt{a^2 + b^2 X_{n-1}^2} \), where \( \varepsilon_i, i \in \mathbb{Z} \), are iid innovations and \( a, b \) are real parameters. Then (11) holds under a very mild moment condition \( \varepsilon_0 \in \mathcal{L}^\alpha \) for some \( \alpha > 0 \) and \( |b||\varepsilon_0|_\alpha < 1 \).

### 2.2. Short-range dependent linear processes

Let \( \varepsilon_i \) be iid random variables with \( \varepsilon_0 \in \mathcal{L}^q \), \( q > 0 \). Assume \( \mathbb{E}(\varepsilon_0) = 0 \) if \( q \geq 1 \). For a real sequence \((a_i)_{i \geq 0}\) satisfying \( \sum_{i=0}^{\infty} |a_i|^q < \infty \), the linear process
\[
X_n = \sum_{i=0}^{\infty} a_i \varepsilon_{n-i} \tag{15}
\]
is well defined and is strictly stationary. Special cases of (15) include ARMA and fractional ARIMA (FARIMA) processes. For \( m \leq n \) define the truncated processes \( X_{n,m} = \sum_{j=0}^{n-m} a_j \varepsilon_{n-j} \) and \( X_{n,m} = \sum_{j=m-n}^{\infty} a_j \varepsilon_{n-j} \). Note that \( X_n = X_{n,m+1} + X_{n,m} \) and the two summands are independent. Denote by \( F_k \) the distribution function of \( (X_{0,1-k}, X_{1,1-k}) \).

**Proposition 2.** Let \( \varepsilon_0 \in \mathcal{L}^q, q > 0 \), and \( \mathbb{E}(\varepsilon_0) = 0 \) if \( q \geq 1 \). Assume that \( F_k \in \mathcal{C}_1 \) for some \( k \in \mathbb{N} \). Then (4) holds provided that
\[
\sum_{i=0}^{\infty} |a_i|^{(q \wedge 2)/2} < \infty. \tag{16}
\]
**Proof.** Let \( L_i(x, y) = 1_{(X_{i-1} - x)(X_i - y) \leq 0} - G(x, y), G_j(x, y) = F_j(x, \infty) + F_j(y, \infty) - 2F_j(x, y) \), \( X'_{j, 0} = X_{j, 0} + a_j(\varepsilon'_0 - \varepsilon_0) \) and \( q_0 = q \wedge 2 \). For \( i \geq k + 1 \), by Schwarz’s inequality,

\[
\|P_0 L_i(x, y)\|_2 = \|E[G_{i-1}(x - X_{i-1, 0}, y - X_{i, 0}) - G_{i-1}(x - X'_{i-1, 0}, y - X'_{i, 0})]\|_2 \\
\leq \|G_{i-1}(x - X_{i-1, 0}, y - X_{i, 0}) - G_{i-1}(x - X'_{i-1, 0}, y - X'_{i, 0})\|_{q_0/2} \leq O(|a_{i-1}|^{q_0/2} + |a_i|^{q_0/2})
\]

since by Lemma 3, \( G^{(1, 1)}_{i-1} \) and \( G^{(1, 0)}_{i-1} \) are bounded. So (16) entails (4).  

**Proposition 2** allows innovations with heavy tails, in which case the traditional covariance based spectral and time domain approaches are not directly applicable. If \( q = 2 \), then (16) becomes \( \sum_{i=1}^{\infty} |a_i| < \infty \), a classical condition for short-range dependence.

**Example 4.** Let \( \varepsilon_0 \in \mathcal{L}^q, q > 0 \), and \( E(\varepsilon_0) = 0 \) if \( q \geq 1 \). Consider the ARMA\((k, p)\) process \( X_n = \sum_{i=1}^{k} \psi_i \varepsilon_{n-i} + \varepsilon_n + \sum_{j=1}^{p} \psi_j \varepsilon_{n-j} \), where \( \psi_1, \ldots, \psi_k, \psi_1, \ldots, \psi_p \) are real parameters. If all the roots of the equation \( 1 - \sum_{i=1}^{k} \psi_i x^i = 0 \) lie outside of the unit circle, then \( X_n = \sum_{i=0}^{\infty} a_i \varepsilon_{n-i} \) with \( |a_i| = O(\lambda^i) \) for some \( \lambda \in (0, 1) \) and thus (16) holds.

In Proposition 2 and Theorem 2 in Section 2.3, we assume that, for some \( \kappa \in \mathbb{N} \), \( F_k \) has bounded partial derivatives of certain order. Assume without loss of generality \( a_0 \neq 0 \). Clearly, if the distribution function of \( \varepsilon_0 \) has bounded derivatives up to order \( p \), then so does \( F_1 \). An interesting example is the stable distribution with characteristic function

\[
\phi(t) = E[\exp(\sqrt{-1} t \varepsilon_0)] = \exp(\sqrt{-1} \delta t - \gamma |t|^{\alpha}), \quad \text{where} \, \delta \in \mathbb{R}, \, \gamma > 0, \, \alpha \in (0, 2].
\]

Here \( \sqrt{-1} \) is the imaginary unit. By the inversion formula, \( \varepsilon_0 \) has bounded derivatives of all orders. Example 5 provides other sufficient conditions under which \( F_k \in \mathcal{C}^p \).

**Example 5.** Let \( \phi(t) = E[\exp(\sqrt{-1} t \varepsilon_0)], \, t \in \mathbb{R} \), be the characteristic function of \( \varepsilon_0 \). Assume that for some \( \lambda > 0 \) and \( c < \infty \), \( |\phi(t)| \leq c/(1 + |t|^\lambda) \) (see [12]). Further assume that \( \mathcal{N} = \{a_n/a_{n-1} : a_{n-1} \neq 0, n \in \mathbb{N} \} \) contains infinitely many different elements. Then for any \( p \), there exists \( \kappa \in \mathbb{N} \) such that \( F_k \in \mathcal{C}^p \). To this end, let

\[
\phi_k(t_1, t_2) = E[e^{\sqrt{-1}(t_1 X_{0, n-1, k} + t_2 X_{1, 1-k})}], \quad t_1, t_2 \in \mathbb{R},
\]

be the characteristic function of the vector \((X_{0,1-k}, X_{1,1-k})\). By independence,

\[
|\phi_k(t_1, t_2)| = |\phi(a_0 t_2)| \prod_{i=1}^{k} |\phi(a_{i-1} t_1 + a_i t_2)| \leq \frac{c^{k+1}}{\prod_{i=1}^{k} [1 + |a_{i-1} t_1 + a_i t_2|^\lambda]}.
\]

Since \( \mathcal{N} \) has infinitely many elements, there exists a subsequence \( n_j, j = 1, 2, \ldots \), such that \( a_{n_j-1}, a_{n_j} \neq 0, b_j := a_{n_j}/a_{n_j-1} \neq 0 \) and \( b_j \neq b_{j'} \) if \( j \neq j' \). Simple calculations show that there exists \( \rho_j > 0 \) depending on \( a_{n_j-1}, a_{n_j}, a_{n_j+1-1} \) and \( a_{n_j+1} \) such that

\[
[1 + |a_{n_j-1} t_1 + a_n t_2|^\lambda][1 + |a_{n_j+1-1} t_1 + a_{n_j+1} t_2|^\lambda] \geq 1 + \rho_j(|t_1| + |t_2|)^\lambda.
\]

Let \( r_k = \min_{i \leq k} |a_i| \geq 0 \), \( \kappa = n_{2k} \) with \( k = [4p/\lambda] + 4 \). Then \( |\phi_k(t_1, t_2)| \leq c^{k+1}[1 + r_k(|t_1| + |t_2|)^\lambda]^{-k} \). By the inversion formula, \( F_k \in \mathcal{C}^p \) since \( k\lambda > 4p \).
2.3. Long-range dependent linear processes

In this section we shall study the linear process (15) with \( a_n = n^{-\beta} L(n) \), where \( \beta > 1/2 \) and \( L \) is a slowly varying function, i.e. for every \( \lambda > 0, L(\lambda n)/L(n) \to 1, n \to \infty \). Let \( \mathbb{E}(\varepsilon_0^2) < \infty \) and \( \beta < 1 \). By Karamata’s theorem, the covariances \( \mathbb{E}(X_0 X_n) \) are of order \( n^{1-2\beta} L^2(n) \) and not summable, suggesting long-range dependence. The case of \( \beta > 1 \) is covered by Proposition 2. In the long-range dependence case, the behavior of \( C_n(\mu) \) is much more complicated. Here we shall establish an asymptotic expansion of \( C_n(\mu) \).

For \( r \geq 0 \) and a differentiable bivariate function \( g \), define

\[
\Delta_r(g)(x, y) = \sum_{i=0}^{r} \frac{r!}{i!(r-i)!} g^{(r,i)}(x, y)
\]

if it exists. Recall that \( G(x, y) = F_X(x) + F_X(y) - 2F(x, y) \). Let

\[
S_n(\mu; p) = \sum_{i=1}^{n} L_i(\mu_{i-1}, \mu_i; p) = C_n(\mu) + \sum_{i=1}^{n} \sum_{r=0}^{p} (-1)^{r+1} \Delta_r(G)(\mu_{i-1}, \mu_i) U_i(r),
\]

where

\[
L_i(x, y; p) = 1(\xi_{i-1} = \xi_i = y) + \sum_{r=0}^{p} (-1)^{r+1} \Delta_r(G)(x, y) U_i(r),
\]

\[
U_i(r) = \sum_{0 \leq j_1 < \ldots < j_r \leq n} \prod_{s=1}^{r} a_{j_s} \varepsilon_{i-j_s} \quad \text{and} \quad U_i(0) = 1.
\]

Note that \( U_i(r) \) is a \( r \)th order polynomial in \( \varepsilon_j \). Intuitively, \( S_n(\mu; p) \) can be viewed as the \( p \)th order Taylor expansion of \( C_n(\mu) \); see [14,43] for the univariate case.

The long-range dependence case results in non-central limit theorems with limiting distributions being multiple Wiener–Itô integrals (MWI, [27]) or integrals with respect to stable Lévy processes. Let \( W(u) \) be a standard two-sided Brownian motion. For \( a > 1/2 \), non-negative integer \( k < 1/(2a - 1) \) and a bounded function \( f \), define the MWI

\[
H_{k,a}(f) = \lambda(k, a) \int_{\mathbb{R}}^1 f(t) \prod_{i=1}^{k} [(t - u_i)^+]^{-a} \mathrm{d}W(u_1) \cdots \mathrm{d}W(u_k),
\]

where \( \lambda(k, a) \) is a normalizing constant such that \( \|H_{k,a}(f_0)\|_2 = 1 \) with \( f_0 \equiv 1 \). Let \( Z_{d}(u), u \in \mathbb{R} \) be a two-sided stable Lévy process with index \( d \). We say that \( \varepsilon_i \) is in the domain of attraction of stable law with index \( d \in (1, 2) \) (denoted by \( \varepsilon_i \in D(d) \)) if there exists a slowly varying function \( L_\varepsilon \) such that \( n^{-1/d} L_\varepsilon^{-1} (n) \sum_{i=1}^{n} \varepsilon_i \) converges in distribution to a stable law with index \( d \).

**Theorem 2.** Let \( \mu_i = m_1(i/n) \), where \( m_1 \) is a piecewise continuous function on \([0, 1] \) with finitely many jumps, and \( a_n = n^{-\beta} L(n) \), where \( \beta > 1/2 \) and \( L \) is a slowly varying function satisfying \( L(n+1)/L(n) = 1 + O(1/n) \). Assume that for some \( \kappa \in \mathbb{N}, F_k \in C^{\kappa+2} \), and that \( \mathbb{E}(\varepsilon_0) = 0 \). (i) Assume \( \varepsilon_0 \in L^q, q \geq 2 \) and \( 2\beta \in (1, 1 + (p+1)^{-1}) \). Let \( S_n(\mu; p) \) be as in (18). Then

\[
\frac{S_n(\mu; p)}{n^{1-(p+1)(\beta-1/2)} L^{p+1}(n)} \Rightarrow (-1)^{p+1} H_{p+1,\beta}[\Delta_{p+1}(G)(m_1(\cdot), m_1(\cdot))].
\]
(ii) Let $\epsilon_i \in \mathcal{D}(d)$, $d \in (1, 2)$. Assume that $1/d < \beta < 1$. Then there is a slowly varying function $L_1$ such that

$$\frac{C_n(\mu) - \mathbb{E}C_n(\mu)}{n^{1-\beta}1/d L_1(n)} \Rightarrow (\beta - 1) \int_{-\infty}^{1} \int_{(-u)^+} \Delta_1(G)(m_1(u + t), m_1(u + t))t^{-\beta} dt dZ_d(u).$$

As a special case of (20), let $p = 0$. Then $C_n(\mu) - \mathbb{E}C_n(\mu) = S_n(\mu; 0)$ is asymptotically normal with a non-$\sqrt{n}$ norming sequence $n^{3/2-\beta}L(n)\|\epsilon_0\|_2$. It is worth pointing out that in [34] a central limit theorem is obtained for functionals of stationary Gaussian vector processes. The latter result can be applied to fixed level-crossing counts.

Example 6. Let $\epsilon_0 \in \mathcal{L}^q$, $q > 1$, and $\mathbb{E}(\epsilon_0) = 0$. Consider the FARIMA($k, \theta, p$) process $X_n$ defined by $\varphi(B)(1 - B)^\theta X_n = \psi(B)\epsilon_n$. Here $\varphi(x) = 1 - \sum_{i=1}^{k} \varphi_i x^i$ and $\psi(x) = 1 + \sum_{i=1}^{p} \psi_i x^i$ are two polynomials of degrees $k$ and $p$, respectively, $B$ is the backward shift operator defined by $B^j X_n = X_{n-j}$, $j \geq 0$, and $\theta \in (-1/2, 1/2)$. Let $\Gamma(x) = \int_{0}^{\infty} t^{x-1}e^{-t} dt$ be the gamma function. In the simple case of $p = k = 0$, we have $X_n = \sum_{i=0}^{\infty} a_i \epsilon_{n-i}$, where

$$a_n = \frac{\Gamma(n + \theta)}{\Gamma(n + 1)\Gamma(\theta)} = n^{-(1-\theta)}\frac{\Gamma(n + \theta)n^{1-\theta}}{\Gamma(n + 1)\Gamma(\theta)} = n^{-(1-\theta)}L(n).$$

By the Stirling formula $\Gamma(x) = \sqrt{2\pi}e^{-x}x^{x-1/2}[1 + O(1/x)], x \to \infty$, we can show that $L(n)$ is a slowly varying function satisfying $L(n + 1)/L(n) = 1 + O(1/n)$. If $\theta \in (0, 1/2)$ or equivalently $\beta := 1 - \theta \in (1/2, 1)$, then $X_n$ is long-range dependent. So Theorem 2 holds with FARIMA($0, \theta, 0$) process under the conditions specified therein. Furthermore, by a similar but more tedious argument, it can be shown that Theorem 2 also applies to general FARIMA($k, \theta, p$) processes provided that $\varphi(z) \neq 0$ for all complex $|z| \leq 1$ and $\psi(1) \neq 0$.

3. Proofs

In this section, $c$ and $c_\gamma$ stand for generic positive constants which may vary among lines. Recall $\mathcal{F}_n = (\epsilon_n, \epsilon_{n-1}, \ldots)$, $\mathcal{F}'_n = (\epsilon_n, \ldots, \epsilon_1, \epsilon_0, \mathcal{F}_0)$, and the detail projection operator $P_k \cdot = \mathbb{E}(\cdot | \mathcal{F}_k) - \mathbb{E}(\cdot | \mathcal{F}_{k-1})$, $k \in \mathbb{Z}$. For $m \leq n$ let $\mathcal{F}_n^m = (\epsilon_n, \epsilon_{n-1}, \ldots, \epsilon_m)$.

3.1. Proof of Theorem 1

Let $L_i(x, y) = 1_{(X_{i-1} - x)(X_i - y) \leq 0} - G(x, y)$, $S_n(\mu) = C_n(\mu) - \mathbb{E}C_n(\mu)$ and $Z_i \equiv L_i(\mu_{i-1}, \mu_i)$. For $r \geq 0$, since $P_j$, $j \in \mathbb{Z}$, are orthogonal projections, we have

$$|\text{cov}(Z_i, Z_{i+r})| = |\mathbb{E}(Z_i Z_{i+r})| = \left| \mathbb{E} \left[ \left( \sum_{j=-\infty}^{i} P_j Z_i \right) \left( \sum_{j'=-\infty}^{i+r} P_{j'} Z_{i+r} \right) \right] \right|$$

$$= \sum_{j=-\infty}^{i} |\mathbb{E}[P_j Z_i (P_j Z_{i+r})]| \leq \sum_{j=-\infty}^{i} \delta_{i-j} \delta_{i+r-j} \leq \sum_{j=0}^{\infty} \delta_j \delta_{j+r}$$

in view of Schwarz’s inequality. By (4), $\sigma^2 \leq \sum_{k \in \mathbb{Z}} |\gamma_k| < \infty$. For $k \in \mathbb{N}$ write

$$\frac{1}{n} \mathbb{E} S_n \mid_2^2 = \frac{1}{n} \sum_{|i-i'| \leq k} \text{cov}(Z_i, Z_{i'}) + \frac{1}{n} \sum_{|i-i'| > k} \text{cov}(Z_i, Z_{i'}) := I_{n,k} + J_{n,k}.$$
Let \( s_k = \sum_{|j| \leq k} \gamma_j \) and \( t_k = \sum_{j=k}^{\infty} \delta_j \). By (3), \( \lim_{n \to \infty} I_{n,k} = s_k \). So

\[
\lim \sup_{n \to \infty} \frac{\| S_n \|_2^2}{n - \sigma^2} \leq \lim \sup_{n \to \infty} |I_{n,k} - s_k| + \lim \sup_{n \to \infty} |J_{n,k}| + |\sigma^2 - s_k| \\
\leq \sum_{r=k}^{\infty} \sum_{j=0}^{\infty} 2\delta_j \delta_{j+r} + |\sigma^2 - s_k| \leq 2t_k \delta_{j} + |\sigma^2 - s_k|.
\]

Let \( k \to \infty \) we obtain \( \lim_{n \to \infty} \| S_n \|_2^2/n = \sigma^2 = \sum_{j \in \mathbb{Z}} \gamma_j \). If \( \sigma^2 = 0 \), then the central limit theorem (5) trivially holds. In the sequel we assume \( \sigma^2 > 0 \). For \( m \in \mathbb{N} \) let

\[
L_{i,m}(x, y) = \mathbb{E}[\mathbf{1}_{(X_{i-1} - x)(X_i - y) \leq 0} | \mathcal{F}_{i-m}^j] - G(x, y).
\]

Define \( S_{n,m} = \sum_{i=1}^{n} L_{i,m}(\mu_{i-1}, \mu_i), D_{i,j,m} = \mathcal{P}_{i-j}[L_i(\mu_{i-1}, \mu_i) - L_{i,m}(\mu_{i-1}, \mu_i)] \) and

\[
\Lambda_m = \sum_{i=0}^{\infty} \delta_{i,m}, \quad \text{where } \delta_{i,m} = \sup_{x,y \in \mathbb{R}} \| \mathcal{P}_0[L_i(x, y) - L_{i,m}(x, y)] \|_2. \tag{23}
\]

Next we show the identity \( \mathcal{P}_0 \mathbb{E}(X | \mathcal{F}_{i-j}^j) = \mathbb{E}[(\mathcal{P}_0 X) | \mathcal{F}_{i-j}^j], i, j \geq 0 \). If \( i \geq j + 1 \), then both sides are zero. If \( i \leq j \), since \( \mathcal{F}_{i-j}^j \) is independent of \( \mathcal{F}_{i-j-1}^j \), we have

\[
\mathcal{P}_0 \mathbb{E}(X | \mathcal{F}_{i-j}^j) = \mathbb{E}[\mathbb{E}(X | \mathcal{F}_{i-j}^j) | \mathcal{F}_0] - \mathbb{E}[\mathbb{E}(X | \mathcal{F}_{i-j}^j) | \mathcal{F}_{-1}]
\]

\[
= \mathbb{E}[\mathbb{E}(X | \mathcal{F}_{i-j}^j) | \mathcal{F}_0^j] - \mathbb{E}[\mathbb{E}(X | \mathcal{F}_{i-j}^j) | \mathcal{F}_{-1}^j]
\]

\[
= \mathbb{E}(X | \mathcal{F}_{0}^j) - \mathbb{E}(X | \mathcal{F}_{-1}^j) = \mathbb{E}(\mathcal{P}_0 X) | \mathcal{F}_{i-j}^j).
\]

Applying the preceding identity with \( X = L_{i,m}(x, y) \), we have by Schwarz’s inequality that \( \sup_{x,y} \| \mathcal{P}_0 L_{i,m}(x, y) \|_2 \leq \delta_i \) for all \( m \geq 0 \). So \( \delta_{i,m} \leq 2\delta_i \). On the other hand, \( \delta_{i,m} \leq \sup_{x,y \in \mathbb{R}} \| L_i(x, y) - L_{i,m}(x, y) \|_2 =: \eta_m \) (say). Therefore \( \delta_{i,m} \leq \min(2\delta_i, \eta_m) \). By Theorem 7.4.3 in [3], for any \( x, y, \lim_{m \to \infty} \| 1_{0 \leq x, x_1 \leq y} - \mathbb{E}(1_{0 \leq x, x_1 \leq y} | \mathcal{F}_{i-m}^j) \|_2 \to 0 \). Hence by the continuity of \( F \), we have \( \sup_{x,y} \| 1_{0 \leq x, x_1 \leq y} - \mathbb{E}(1_{0 \leq x, x_1 \leq y} | \mathcal{F}_{i-m}^j) \|_2 \to 0 \) as \( m \to \infty \). So \( \lim_{m \to \infty} \eta_m = 0 \), and by the Lebesgue dominated convergence theorem, \( \lim_{m \to \infty} \Lambda_m = 0 \).

Since \( \{D_{i,j,m}\}_{i=1}^n \) form martingale differences with respect to \( \mathcal{F}_{i-j} \) and \( \| D_{i,j,m} \|_2 \leq \delta_{i,m}, \sum_{i=1}^{n} D_{i,j,m} \|_2 \leq \sqrt{n} \delta_{j,m} \). Let LIM denote \( \lim \sup_{m \to \infty} \lim \sup_{n \to \infty} \). Then

\[
\text{LIM} \frac{\| S_n(\mu) - S_{n,m} \|_2}{\sqrt{n}} = \text{LIM} \frac{\| \sum_{i=1}^{n} \sum_{j=0}^{\infty} D_{i,j,m} \|_2}{\sqrt{n}}
\]

\[
\leq \text{LIM} \frac{\| \sum_{j=0}^{\infty} \sum_{i=1}^{n} D_{i,j,m} \|_2}{\sqrt{n}} \leq \lim_{m \to \infty} \Lambda_m = 0. \tag{24}
\]

Since \( \sigma^2 > 0 \), for sufficiently large \( m \), \( \lim_{n \to \infty} \| S_{n,m} \|_2^2/n > \sigma^2/2 > 0 \). By the central limit theorem for \( m \)-dependent random variables (cf. [16] or [32]), \( S_{n,m}/\| S_{n,m} \|_2 \Rightarrow N(0, 1) \). Hence \( S_n(\mu)/\sigma \Rightarrow N(0, 1) \). \( \diamond \)
3.2. Proof of Theorem 2

To prove Theorem 2, we need some lemmas. Lemma 1 is a simple consequence of Rosenthal’s inequality (cf. Theorem 1.5.11 in [6]). Details are omitted.

Lemma 1. Let \( \varepsilon_1 \in \mathcal{L}' \), \( \gamma > 0 \) and \( \mathbb{E}(\varepsilon_1) = 0 \) if \( \gamma \geq 1 \); let \( \gamma' = \gamma \wedge 2 \) and \( \gamma'' = \gamma \vee 2 \). Then there exists a \( c_\gamma \) such that \( \mathbb{E} \left| \sum_{i=1}^{n} b_i \varepsilon_i \right|^\gamma \leq c_\gamma \left( \sum_{i=1}^{n} |b_i|^{\gamma'} \right)^{\gamma''/2} \) holds for all \( b_i \in \mathbb{R} \).

Lemma 2. Let \( f \in \mathcal{C}^2 \). Then there exists a constant \( c < \infty \), depending only on \( f \), such that for all \( \gamma \in [1, 2] \) and \( x, y, \delta_1, \delta_2 \in \mathbb{R} \),

\[
|f(x + \delta_1, y + \delta_2) - f(x, y) - f^{(1, 1)}(x, y)\delta_1 - f^{(1, 0)}(x, y)\delta_2| \leq c \sum_{i=1}^{2} (|\delta_i|^\gamma + 1_{|\delta_i| \geq 1}).
\]

Proof. Let \( T = T(x, y, \delta_1, \delta_2) = f(x + \delta_1, y + \delta_2) - f(x, y) - f^{(1, 1)}(x, y)\delta_1 - f^{(1, 0)}(x, y)\delta_2 \). If \( |\delta_1| \leq 1, |\delta_2| \leq 1 \), by Taylor’s expansion, \( |T| \leq c(|\delta_1|^2 + |\delta_2|^2) \leq c(|\delta_1|^\gamma + |\delta_2|^\gamma) \). By the boundedness of \( f \), \( f^{(1, 1)} \) and \( f^{(1, 0)} \),

\[
|T|1_{|\delta_1| > 1, |\delta_2| \leq 1} \leq c(1 + |\delta_1|)1_{|\delta_1| > 1, |\delta_2| \leq 1} \leq c(|\delta_1|^\gamma + 1_{|\delta_1| > 1}).
\]

We can similarly deal with \( |T|1_{|\delta_1| \leq 1, |\delta_2| > 1} \) and \( |T|1_{|\delta_1| > 1, |\delta_2| > 1} \). So Lemma 2 follows.

Recall that \( X_{n,m} = \sum_{j=-m}^{\infty} a_j \varepsilon_n - j \), \( X_{n,m} = \sum_{j=-m}^{n-m} a_j \varepsilon_{n-j} \) and that \( F_X, F \) and \( F_k \) are the distribution functions of \( X_0, (X_0, X_1) \) and \((X_{0,1-k}, X_{1,1-k})\), respectively. For \( q \geq 2 \) let \( q' = q \wedge 4 \), \( A_n(k) = \sum_{i=k}^{\infty} |a_i| \), \( B_n(k) = \sum_{i=k}^{\infty} |a_i - a_{i+1}| \),

\[
\theta_{n,p,q} = |a_n - a_{n-1}| + |a_n|^{q'/2} + |a_n||A_n^{p/2}(2) + A_n^{1/2}(q') + B_n^{1/2}(2)|, 
\]

\[
\Theta_{n,p,q} = \sum_{k=1}^{n} \theta_{k,p,q} \quad \text{and} \quad \Xi_{n,p,q} = n \Theta_{n,p,q} + \sum_{i=1}^{\infty} (\Theta_{n+i,\cdot,p,q} - \Theta_{i,\cdot,p,q})^2. 
\]

Lemma 4 extends Lemma 10 in [43] to the bivariate case.

Lemma 3. Assume that \( F_k \in \mathcal{C}^{p} \) for some \( \kappa, p \in \mathbb{N} \). Then \( F_m(x, y), F(x, y) \in \mathcal{C}^{p} \) for all \( m \geq \kappa \). Furthermore, for all \( n \geq 0 \) and \( 0 \leq \gamma \leq r \leq p \),

\[
F^{(r,\gamma)}_{n+m}(x, y) = \mathbb{E} F^{(r,\gamma)}_m(x, y) \left( x - \sum_{j=m}^{n+m-1} a_j \varepsilon'_{-j}, y - \sum_{j=m+1}^{n+m} a_j \varepsilon'_{-j} \right). 
\]

\[
F^{(r,\gamma)}(x, y) = \mathbb{E} F^{(r,\gamma)}_m(x, y) \left( x - \sum_{j=m}^{\infty} a_j \varepsilon'_{-j}, y - \sum_{j=m+1}^{\infty} a_j \varepsilon'_{-j} \right). 
\]

Proof. For \( r = 0 \), a conditioning argument entails (25). For \( r \geq 1 \), (25) follows from the Lebesgue dominated convergence theorem. See Lemmas 6 and 7 in [43] for the details of the proof in the univariate case.

Lemma 4. Let \( L_n(x, y; p) \) be as in (19). Assume that \( \varepsilon_0 \in \mathcal{L}^q, q \geq 2 \). Let \( q' = q \wedge 4 \). Further assume that \( F_k \in \mathcal{C}^{p+1} \) for some integer \( \kappa > 0 \). Then

\[
\sup_{x,y} \| P_1 L_n(x, y; p) \|_2 = O(\theta_{n,p,q}).
\]
Proof. For notational convenience we shall often drop the arguments x, y. For example, we write \( F = F(x, y) \). Let \( L_n^p = 1_{x_{n-1} \leq x, x_n \leq y} + \sum_{r=0}^{p} (-1)^{r+1} \Delta_r(F)(x, y) U_n(r) \). It suffices to show that \( \sup_{x, y} \| P_i L_n^p(p) \| = O(\theta_{n, p, q}) \). We shall adopt the argument in the proof of Lemma 9 in [43]. As in [43], we may assume \( n \geq \kappa + 1 \). For \( m \leq n \), let

\[
u_m = u_{n,m} = (x - X_{m,n}, y - X_{m,n+1}), \quad u^*_m = u^*_{n,m} = (x - X^*_{m,n}, y - X^*_{m,n+1}),
\]

where \( X^*_{m,n} = \sum_{j=n-m}^{\infty} a_j \varepsilon''_{n-j} \) is a coupled process of \( X_{m,n} \). For \( 0 \leq \gamma \leq r - 1 \leq p - 1 \), since \( 1 \leq q'/2 \leq 2 \), by Lemmas 2 and 3,

\[
\| F_n^{(r-1, \gamma)}(u_1) - F_n^{(r-1, \gamma)}(u_0) + F_n^{(r, \gamma+1)}(u_0) a_{n-1} \varepsilon_1 + F_n^{(r, \gamma)}(u_0) a_{n} \varepsilon_1 \|_2^2 \\
= O(|a_{n-1}| q' + |a_n| q').
\]

By Lemma 3,

\[
F_n^{(r-1, \gamma)}(u_1) = \mathbb{E}[F_n^{(r-1, \gamma)}(u_1 - (a_{n-1} \varepsilon_1, a_n \varepsilon_1)) + F_n^{(r, \gamma+1)}(u_1) a_{n-1} \varepsilon_1 + F_n^{(r, \gamma)}(u_1) a_{n} \varepsilon_1 | F_1].
\]

Thus, by Schwarz’s inequality and Lemma 2,

\[
\| F_n^{(r-1, \gamma)}(u_1) - F_n^{(r-1, \gamma)}(u_1) \|_2^2 \leq \| F_n^{(r-1, \gamma)}(u_1) - F_n^{(r-1, \gamma)}(u_1 - (a_{n-1} \varepsilon_1', a_n \varepsilon_1')) \|_2^2 \\
- F_n^{(r, \gamma+1)}(u_1) a_{n-1} \varepsilon_1' - F_n^{(r, \gamma)}(u_1) a_{n} \varepsilon_1' \|_2^2 \\
= O(|a_{n-1}| q' + |a_n| q'),
\]

which together with (27) gives

\[
\| F_n^{(r-1, \gamma)}(u_1) - F_n^{(r-1, \gamma)}(u_0) + F_n^{(r, \gamma+1)}(u_0) a_{n-1} \varepsilon_1 + F_n^{(r, \gamma)}(u_0) a_{n} \varepsilon_1 \|_2^2 \\
= O(|a_{n-1}| q' + |a_n| q').
\]

For \( 0 \leq \gamma + 1 \leq r \leq p \), define

\[
M_n^{(r, \gamma)} = \Delta_1[F_n^{(r-1, \gamma)}(u_0)] + \sum_{i=r}^{p} (-1)^{i+r+1} \Delta_{i-r+1}(F^{(r-1, \gamma)}) \mathbb{E}[U_{n+1}(i - r) | F_0].
\]

We shall use the induction argument to show that

\[
\sup_{x, y} \| M_n^{(r, \gamma)} \|_2^2 = O(A_n^{p-r+1} + A_n(q') + B_n(2)).
\]

When \( r = p \),

\[
M_n^{(r, \gamma)} \|_2^2 \leq 2[\| F_n^{(r, \gamma+1)}(u_0) - F^{(r, \gamma+1)}(x, y) \|_2^2 + \| F_n^{(r, \gamma)}(u_0) - F^{(r, \gamma)}(x, y) \|_2^2].
\]

By Schwarz’s inequality and Lemma 3,

\[
\| F_n^{(r, \gamma)}(u_0) - F_n^{(r, \gamma)}(u_0) \|_2^2 = \| \mathbb{E}[F_n^{(r, \gamma)}(u_0) - F_n^{(r, \gamma)}(u_0^*) | F_0] \|_2^2 \\
\leq \| F_n^{(r, \gamma)}(u_0) - F_n^{(r, \gamma)}(u_0^*) \|_2^2 = O(A_n(2)).
\]
Similarly, \( \|F_n^{(r,\gamma+1)}(u_0) - F_n^{(r,\gamma+1)}(x, y)\|^2 = O[A_n(2)] \). So (29) holds for \( r = p \). Now suppose it also holds for some \( r \leq p \) and we consider \( r - 1 \). Note that for any \( j \geq 0 \) and \( \gamma \leq r - 2 \),

\[
\mathcal{P}_{-j}\Delta_1[F_n^{(r-2,\gamma)}(u_0)] = \Delta_1[F_n^{(r-2,\gamma)}(u_{-j}) - F_n^{(r-2,\gamma)}(u_{-j-1})].
\]  

(31)

It is easily seen that

\[
\|\mathcal{P}_{-j}M_n^{(r-1,\gamma)}\|_2 \leq I_n(j) + J_n(j) + c|a_{n+j+1} - a_{n+j}|,
\]

where

\[
I_n(j) = \|\mathcal{P}_{-j}\Delta_1[F_n^{(r-2,\gamma)}(u_0)] + \Delta_1[F_n^{(r-1,\gamma+1)}(u_{-j-1})]a_{n+j}\epsilon_j - \Delta_1[F_n^{(r-1,\gamma)}(u_{-j-1})]a_{n+j+1}\epsilon_j\|_2 = O(|a_{n+j}|^{q/2} + |a_{n+j+1}|^{q/2})
\]

by (28) and (31) and

\[
J_n(j) = \|\Delta_1[F_n^{(r-1,\gamma)}(u_{-j-1})]a_{n+j+1}\epsilon_j - \sum_{i=r-1}^{p} (-1)^{i+r}\Delta_{i-r+2}[F_n^{(r-2,\gamma)}]\mathcal{P}_{-j}\mathbb{E}[U_{n+1}(i - r + 1)|\mathcal{F}_0]\|_2.
\]

Note that \( \mathcal{P}_{-j}\mathbb{E}[U_{n+1}(i - r + 1)|\mathcal{F}_0] = a_{n+j+1}\epsilon_j \mathbb{E}[U_{n+1}(i - r)|\mathcal{F}_{-j-1}] \) if \( i \geq r \) and it vanishes if \( i = r - 1 \), and the decomposition \( \Delta_{i-r+2}(F_n^{(r-2,\gamma)}) = \Delta_{i-r+1}([\Delta_1(F_n^{(r-2,\gamma)})]) \). Then \( J_n(j) \leq c|a_{n+j+1}|(\|M_n^{(r,\gamma+1)}\|_2 + \|M_n^{(r,\gamma)}\|_2) \). Since \( M_n^{(r,\gamma)} = \sum_{j=0}^{\infty} \mathcal{P}_{-j}M_n^{(r-1,\gamma)} \) and \( \mathcal{P}_{-j}, j \in \mathbb{Z} \), are orthogonal projections, the induction is completed by observing that

\[
\|M_n^{(r-1,\gamma)}\|_2^2 = \sum_{j=0}^{\infty} \|\mathcal{P}_{-j}M_n^{(r-1,\gamma)}\|_2^2 \leq 3 \sum_{j=0}^{\infty} (I_n(j)^2 + J_n(j)^2 + c|a_{n+j+1} - a_{n+j}|^2)
\]

\[
= O[A_n^{p-r+2}(2) + A_n(q') + B_n(2)].
\]  

(32)

Lemma 4 now follows from (28) and (29) in view of

\[
\|\mathcal{P}_{-1}L_{n+1}^{\circ}(p)\|_2 = \|F_{n-1}(u_1) - F_n(u_0) + F_{n}^{(1,1)}(u_0)a_{n-1}\epsilon_1 + F_{n}^{(1,0)}(u_0)a_{n}\epsilon_1
\]

\[
- a_{n}\epsilon_1 M_{n}^{(1,0)} + F_{n}^{(1,1)}(u_0)(a_n - a_{n-1})\epsilon_1\|_2. \quad \diamond
\]  

(33)

Recall (18) for \( S_n(\mu; p) \). Proposition 3 below presents a reduction principle for \( S_n(\mu; p) \). Reduction principles are useful in proving asymptotic distributions for the long-range dependent case (cf. Theorem 2). See [15] for recent contributions of reduction principles. The results in the latter paper are not applicable here.

**Proposition 3.** Let \( \epsilon_0 \in \mathcal{L}^q, q > 0 \). Assume that \( F_k \in C^{p+1} \) for some integer \( \kappa > 0 \). Let \( a_n = n^{-\beta}L(n) \), where \( \beta > 1/2 \) and \( L \) is a slowly varying function satisfying \( L(n + 1)/L(n) = 1 + O(1/n) \). (i) Assume \( q \geq 2 \) and \( q > 2/\beta \). Then \( \sup_{\mu} \| S_n(\mu; p)\|_2 = O(\sqrt{n}) \) if \( (p + 1)(2\beta - 1) > 1 \); \( \sup_{\mu} \| S_n(\mu; p)\|_2 = O(\sqrt{n}) \) if \( (p + 1)(2\beta - 1) = 1 \) and \( \sup_{\mu} \| S_n(\mu; p)\|_2 = O([n^{1-(p+1)(\beta-1/2)}] L^{p+1}(n)) \) if \( (p + 1)(2\beta - 1) < 1 \). (ii) Assume \( q \in [1, 2) \)
and $\beta \in (1/q, 1)$. Then for any $v \in (1/\beta, q]$

$$\sup_{\mu} \| S_n(\mu; 1) \|_v = O\left[ n^{1/v+(1-\beta q/v)^+} \sum_{i=1}^{n} |L^{q/v}(i) / i \right. \]

$$ + n^{1/v+(1+1/v-2\beta)^+} \sum_{i=1}^{n} |L^2(i) / i \right] \right].$$

**Proof.** (i) Let $\lambda_n = \sup_{x,y} \| P_j L_n(x,y;p) \|_2$. Note that $S_n(\mu; p) = \sum_{j=-\infty}^{n} P_j S_n(\mu; p)$. Since $P_j$, $j \in \mathbb{Z}$, are orthogonal projections and $P_j L_n(x,y;p) = 0$ for $j \geq n + 1$, we have by Lemma 4 that

$$\| S_n(\mu; p) \|^2_2 = \sum_{j=-\infty}^{n} \| P_j S_n(\mu; p) \|^2_2$$

$$\leq \sum_{j=-\infty}^{n} \mathbb{E} \left\{ \sum_{i=1}^{n} \left[ \frac{|P_j L_i(\mu_{i-1}, \mu_i; p)|^2}{\lambda_{i-j+1}} \sum_{i=1}^{n} \lambda_{i-j+1} \right] \right\}$$

$$\leq \sum_{j=-\infty}^{n} \left[ \sum_{i=1}^{n} \lambda_{i-j+1} \right]^2 \leq \sum_{j=-\infty}^{n} \left[ \sum_{i=1}^{n} \theta_{i-j+1, p,q} \right]^2 = O(\Xi_{n,p,q}).$$

The argument in Lemma 5 in [43] shows that

$$\Xi_{n,p,q} = O(n), \quad O(n) \left[ \sum_{i=1}^{n} |L^{(p+1)}(i) / i \right] \right]^2 \text{ or } O[n^{2-(p+1)(2\beta-1)} L^{2(p+1)}(n)]$$

if $(p+1)(2\beta-1) > 1$, $(p+1)(2\beta-1) = 1$ or $(p+1)(2\beta-1) < 1$ holds respectively. Then we have (i).

(ii) The argument in the proof of Lemma 4 can be applied here to show that

$$\sup_{x,y} \| P_j L_n(x,y;1) \|_v = O[|a_n - a_{n-1}| + |a_n|^{q/v} + |a_n| A_{1/v}^{1/v}(v)]$$

$$= O[n^{-\beta q/v} L^{q/v}(n) + n^{-2\beta+1/v} L^2(n)], \quad (34)$$

where the second equality follows from Karamata’s theorem. For example, since $1 \leq q/v \leq 2$, (27) now becomes

$$\| F_n(u_1) - F_n(u_0) + F_n^{(1,1)}(u_0)a_{n-1}\varepsilon_1 + F_n^{(1,0)}(u_0)a_n\varepsilon_1 \|_v = O(|a_{n-1}|^q + |a_n|^q)$$

in view of Lemma 2; (28) now becomes

$$\| F_{n-1}(u_1) - F_n(u_0) + F_n^{(1,1)}(u_0)a_{n-1}\varepsilon_1 + F_n^{(1,0)}(u_0)a_n\varepsilon_1 \|_v = O(|a_{n-1}|^q + |a_n|^q);$$

when $p = 1$, by the argument of (30), we have

$$\| M_n^{(1,0)} \|_v = \| \Delta_1 [F_n(u_0) - F(x,y)] \|_v = O[A_n(v)]$$

in view of Lemma 1. So (34) follows from (33). By the von Bahr–Essén inequality,
Proposition 3. If \( p = 2 \), with the above constructed \( \frac{1}{\nu} \) and \( 1 < \frac{1}{\nu} \), we have \( \sum_{i=1}^{n} \left| L^q(v)/(i)/i \right| \nu < 1 \). The complete the proof by noting that \( \sigma_{n,p+1} = o(\sigma_{n,p+1}) \), where \( \sigma_{n,p+1} = 1 \). Thus

\[
S_n(\mu; p) = S_n(\mu; p) - S_n(\mu; p+1) + o_p(\sigma_{n,p+1})
\]

We complete the proof by noting that

\[
\frac{\sum_{i=1}^{n} \Delta_{p+1}(G)(\mu_{i-1}, \mu_i)U_i(p + 1)}{\sigma_{n,p+1}||\varepsilon_0||^{p+1}} \Rightarrow H_{p+1, \beta} [\Delta_{p+1}(G)(m_1(\cdot), m_1(\cdot))],
\]

which follows from the argument of Theorem 5.1 in [39] (see also Theorem 2 in [38]).

(ii) Let \( \ell_n = n^{1-p}/d L(n)L_{+}(n) \). Using the argument in the proof of Theorem 5.1 in [18], we can show that

\[
\frac{\sum_{i=1}^{n} \Delta_{1}(G)(\mu_{i-1}, \mu_i)X_i}{(1 - \beta)\ell_n} \Rightarrow \int_{-\infty}^{1} \left[ \int_{(-u)^+}^{1} \Delta_{1}(G)(m_1(u + t), m_1(u + t))t^{-\beta} dt \right] dZ_d(u).
\]

Since \( d \in (1, 2) \) and \( 1/d < \beta < 1 \), for

\[
\eta = \max \left\{ \frac{3\beta d - 1}{\beta(\beta d + 1)}, \frac{d}{\beta(1 + d - \beta d)} \right\} \quad \text{and} \quad \zeta = \max \left\{ \frac{2d}{\beta d + 1}, \frac{d}{1 + d - \beta d} \right\},
\]

we have \( 1/\beta < \eta < d \) and \( \zeta > 1/\beta \). Choose \( q \in (\eta, d) \). Then \( \zeta < (\beta q - 1)/(\beta - 1/d) < q \). Further choose \( v \in (\zeta, (\beta q - 1)/(\beta - 1/d)) \). Then \( 1/v + (1 - \beta q/v) + < 1 - \beta + 1/d \) and \( 1/v + (1 + 1/v - 2\beta)^+ < 1 - \beta + 1/d \). Since \( L \) is slowly varying, by the argument of Lemma 5 in [43], both \( \sum_{i=1}^{n} \left| L^q(v)/(i)/i \right| \) and \( \sum_{i=1}^{n} \left| L^2(i)/i \right| \) are slowly varying functions. By Proposition 3 with the above constructed \( q \) and \( v \), \( \|S_n(\mu; 1)\|_v = o(\ell_n) \), which completes the proof since \( S_n(\mu; 0) = S_n(\mu; 1) - \sum_{i=1}^{n} \Delta_{1}(G)(\mu_{i-1}, \mu_i)X_i \). ♦

Acknowledgments

The authors are grateful to the two referees and the associate editor for their many helpful comments. The work is supported in part by NSF grant DMS-0478704.

References