ASYMPTOTIC THEORY FOR CURVE-CROSSING ANALYSIS

Zhibiao Zhao and Wei Biao Wu

Asymptotic theory for curve-crossing analysis

By Zhibiao Zhao and Wei Biao Wu

September 19, 2006

Department of Statistics, The University of Chicago

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.

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

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

where $\mathbf{1}_A = 1$ 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 μ include (i) $\mu_i = m_1(i/n)$ for some function m_1 on $[0,1]$ and (ii) $\mu_i = m_2(i)$ for some function m_2 on $\mathbb{R}$. These two formulations of μ 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 μ 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, 20, 21, 30, 34]. For example, if

$X_i = \sum_{j=1}^{k} [A_j \cos(i\omega_j) + B_j \sin(i\omega_j)] + e_i$ for independent normal random variables A_j, B_j and e_i, then one can estimate the frequencies ω_j by considering the zero-crossing counts of X_i and its differencing 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 Cabaña [1], Bulinskaya [2], Rice [31] and Ylvisaker [45] 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 [35] 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 [40, 41] considered linear processes.

This paper aims at establishing an asymptotic distributional theory for $C_n(\mu)$ for non-constant sequences μ. Such results can provide an inferential theory for statistical analysis using the curve-crossing method. For the number of curve-crossings by stationary continuous-time Gaussian processes, Slud [36] 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 can not 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

$$X_n = g(\varepsilon_n, \varepsilon_{n-1}, \ldots), \tag{2}$$

where g is a measurable function and ε_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 ε_i such that the distributional equality $(X_k)_{k \in \mathbb{Z}} =_\mathcal{D} (g(\varepsilon_k, \varepsilon_{k-1}, \ldots))_{k \in \mathbb{Z}}$ holds; see [17, 33, 39].

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 $\mathcal{F}'_n = (\varepsilon_n, \ldots, \varepsilon_1, \varepsilon'_0, \varepsilon_{-1}, \ldots)$ 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 $Z \in \mathcal{L}^q$, $q > 0$, 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 (μ_i). A general central limit theorem is given below.

Theorem 1. *Assume that, for each fixed integer k, the limit*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \text{cov}[\mathbf{1}_{(X_0 - \mu_{i-1})(X_1 - \mu_i) \leq 0}, \mathbf{1}_{(X_k - \mu_{i+k-1})(X_{k+1} - \mu_{i+k}) \leq 0}] = \gamma_k \quad (say) \tag{3}$$

exists. Further assume that $F(\cdot, \cdot)$ is continuous and that

$$\sum_{i=0}^{\infty} \delta_i < \infty, \quad where \quad \delta_i = \sup_{x,y \in \mathbb{R}} \|\mathcal{P}_0 \mathbf{1}_{(X_{i-1} - x)(X_i - y) \leq 0}\|_2. \tag{4}$$

Then $\sigma^2 := \sum_{k \in \mathbb{Z}} \gamma_k < \infty$, $\|C_n(\mu) - \mathbb{E}C_n(\mu)\|_2^2/n \to \sigma^2$ as $n \to \infty$ and

$$\frac{C_n(\mu) - \mathbb{E}C_n(\mu)}{\sqrt{n}} \Rightarrow N(0, \sigma^2). \tag{5}$$

3

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 μ; see Examples 1 and 2 below.

Example 1. For the level-crossing case with $\mu_i \equiv c$, we have $\gamma_k = \operatorname{cov}(V_0, V_k)$, where $V_i = \mathbf{1}_{(X_i-c)(X_{i+1}-c)\leq 0}$. 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

$$\gamma_k = \int_0^1 \operatorname{cov}\left(\mathbf{1}_{[X_0-m_1(t)][X_1-m_1(t)]\leq 0}, \ \mathbf{1}_{[X_k-m_1(t)][X_{k+1}-m_1(t)]\leq 0}\right)dt. \tag{6}$$

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

$$\gamma_k = \frac{1}{b}\sum_{j=1}^b \operatorname{cov}(V_{0,j}, V_{k,j}), \ \text{where } V_{i,j} = \mathbf{1}_{[X_i-\cos(2\pi\omega(i+j-1))][X_{i+1}-\cos(2\pi\omega(i+j))]\leq 0}. \tag{7}$$

If ω 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

$$\gamma_k = \int_0^1 \operatorname{cov}\left(\mathbf{1}_{Z_0(u)Z_1(u)\leq 0}, \ \mathbf{1}_{Z_k(u)Z_{k+1}(u)\leq 0}\right)du. \tag{8}$$

Remark 1. To apply Theorem 1, one needs to estimate the limiting variance σ^2. For stationary processes the latter problem is closely related to the long-run variance estimation. Unfortunately, due to the non-stationarity of $Z_i = \mathbf{1}_{(X_{i-1}-\mu_{i-1})(X_i-\mu_i)\leq 0} - G(\mu_{i-1}, \mu_i)$, the estimation of σ^2 is not straightforward. Let $\hat{G}$ be an estimate of G. For example, we can take $\hat{G}(x,y) = n^{-1}\sum_{i=1}^n \mathbf{1}_{(X_{i-1}-x)(X_i-y)\leq 0}$. As the proof of Theorem 1 (cf Section 3.1) suggests, one can propose the truncated estimate

$$\hat{\sigma}_k^2 = \frac{1}{n}\sum_{i,j=1}^n \hat{Z}_i\hat{Z}_j\mathbf{1}_{|i-j|\leq k}, \ \text{where } \hat{Z}_i = \mathbf{1}_{(X_{i-1}-\mu_{i-1})(X_i-\mu_i)\leq 0} - \hat{G}(\mu_{i-1}, \mu_i).$$

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. $\diamondsuit$

4

Remark 2. Recall that $(\varepsilon'_i)_{i \in \mathbb{Z}}$ 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 [43]. 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:

$$\tilde{\omega}_{i,q} = \|Y_i - Y_i^*\|_q, \quad \text{where} \quad Y_i^* = h(\varepsilon_i, \ldots, \varepsilon_1, \varepsilon'_0, \varepsilon'_{-1}, \ldots). \tag{9}$$

The difference between the two coupling versions Y'_i and Y_i^* of Y_i is that the former replaces ε_0 with ε'_0 while the latter replaces ε_i with ε'_i for all $i \leq 0$. By triangle inequality, $\tilde{\omega}_{i,q} \leq \sum_{j=i}^{\infty} \omega_{j,q}$. 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 ε_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 [39]: $X_n = a(X_{n-1} \vee 0) + b(X_{n-1} \wedge 0) + \varepsilon_n$, autoregressive conditional heteroscedasticity (ARCH) model [9]: $X_n = \varepsilon_n \sqrt{a^2 + b^2 X_{n-1}^2}$, random coefficient model [29]: $X_n = (a + b\varepsilon_n)X_{n-1} + \varepsilon_n$ and exponential autoregressive model [13]: $X_n = (a + be^{-cX_{n-1}^2})X_{n-1} + \varepsilon_n$ 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*

$$R_{\varepsilon_0}(x_0) \in \mathcal{L}^\alpha \quad and \quad \rho := \sup_{x \neq x'} \frac{\|R_{\varepsilon_0}(x) - R_{\varepsilon_0}(x')\|_\alpha}{|x - x'|} < 1 \tag{11}$$

Further assume that, for some $p > 2$,

$$\sup_{x \in \mathbb{R}} |F_X(x + t) - F_X(x)| = O[(\log |t|^{-1})^{-p}] \ as \ t \to 0. \tag{12}$$

Then (4) holds.

5

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 ε_i, $i \leq 0$, in $\mathcal{F}_n$ replaced by ε'_i. By Theorem 2 in [44], if (11) is satisfied, then X_n has a unique stationary solution of the form (2) with the function g satisfying

$$\|X_n - g(\mathcal{F}^*_n)\|_\alpha = O(\rho^n). \tag{13}$$

Let $X'_i = g(\mathcal{F}'_i)$, $X^*_i = g(\mathcal{F}^*_i)$ and $\kappa_i = \sup_x \|\mathbf{1}_{X_i \leq x} - \mathbf{1}_{X'_i \leq x}\|_2$. Since

$$\mathbb{E}[\mathbf{1}_{(X_{i-1}-x)(X_i-y)\leq 0}|\mathcal{F}_{-1}] = \mathbb{E}[\mathbf{1}_{(X'_{i-1}-x)(X'_i-y)\leq 0}|\mathcal{F}_{-1}] = \mathbb{E}[\mathbf{1}_{(X'_{i-1}-x)(X'_i-y)\leq 0}|\mathcal{F}_0],$$

by Jensen's inequality,

$$\|\mathcal{P}_0 \mathbf{1}_{(X_{i-1}-x)(X_i-y)\leq 0}\|_2 \leq \|\mathbf{1}_{(X_{i-1}-x)(X_i-y)\leq 0} - \mathbf{1}_{(X'_{i-1}-x)(X'_i-y)\leq 0}\|_2 \leq 2(\kappa_i + \kappa_{i-1}). \tag{14}$$

Note that $\|\mathbf{1}_{X_i \leq x} - \mathbf{1}_{X'_i \leq x}\|_2 \leq \|\mathbf{1}_{X_i \leq x} - \mathbf{1}_{X^*_i \leq x}\|_2 + \|\mathbf{1}_{X_{i+1} \leq x} - \mathbf{1}_{X^*_{i+1} \leq x}\|_2$. Let $\tau = \rho^{1/(2\alpha)}$. By (12) and (13),

$$\begin{aligned}
\mathbb{P}(X_i \leq x, X^*_i \geq x) &\leq \mathbb{P}(X_i \leq x, X^*_i \geq x, |X^*_i - X_i| \leq \tau^i) + \mathbb{P}(|X^*_i - X_i| > \tau^i) \\
&\leq \mathbb{P}(|X_i - x| \leq \tau^i) + \tau^{-\alpha i}\mathbb{E}(|X^*_i - X_i|^\alpha) = O(i^{-p} + \rho^{i/2}).
\end{aligned}$$

So $\sum_{i=0}^\infty \kappa_i < \infty$ and the proposition follows from (14). $\diamond$

Example 3. Consider the ARCH model $X_n = \varepsilon_n\sqrt{a^2 + b^2 X_{n-1}}$, where $\varepsilon_i, i \in \mathbb{Z}$, are iid innovations and a, b are real parameters. If $\varepsilon_0 \in \mathcal{L}^\alpha$ and $|b|\|\varepsilon_0\|_\alpha < 1$, then (11) hold.

2.2 Short-range dependent linear processes

Let ε_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 $\sum_{i=0}^\infty |a_i|^{q\wedge 2} < \infty$, 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 $\underline{X}_{n,m} = \sum_{j=n-m}^\infty a_j \varepsilon_{n-j}$. Note that $X_n = X_{n,m+1} + \underline{X}_{n,m}$ 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 $\varepsilon_0 \in \mathcal{L}^q, q > 0$. Assume that $F_\kappa \in \mathcal{C}^1$ for some $\kappa \in \mathbb{N}$. Then (4) holds provided that*

$$\sum_{i=0}^{\infty} |a_i|^{(q \wedge 2)/2} < \infty. \tag{16}$$

Proof. Let $L_i(x, y) = \mathbf{1}_{(X_{i-1}-x)(X_i-y) \leq 0} - G(x, y), G_j(x, y) = F_j(x, \infty) + F_j(y, \infty) - 2F_j(x, y)$, $\underline{X}'_{j,0} = \underline{X}_{j,0} + a_j(\varepsilon'_0 - \varepsilon_0)$ and $q_0 = q \wedge 2$. For $i \geq \kappa + 1$, by Schwarz's inequality,

$$
\begin{aligned}
\|\mathcal{P}_0 L_i(x, y)\|_2 &= \|\mathbb{E}[G_{i-1}(x - \underline{X}_{i-1,0}, y - \underline{X}_{i,0}) - G_{i-1}(x - \underline{X}'_{i-1,0}, y - \underline{X}'_{i,0})|\mathcal{F}_0]\|_2 \\
&\leq \|G_{i-1}(x - \underline{X}_{i-1,0}, y - \underline{X}_{i,0}) - G_{i-1}(x - \underline{X}'_{i-1,0}, y - \underline{X}'_{i,0})\|_{q_0}^{q_0/2} \\
&= O(|a_{i-1}|^{q_0/2} + |a_i|^{q_0/2})
\end{aligned}
$$

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

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 $\varepsilon_0 \in \mathcal{L}^q, q > 0$. Consider the ARMA$(k, p)$ process $X_n - \sum_{i=1}^{k} \varphi_i X_{n-i} = \varepsilon_n + \sum_{j=1}^{p} \psi_j \varepsilon_{n-j}$, where $\varphi_1, \ldots, \varphi_k, \psi_1, \ldots, \psi_p$ are real parameters. If all the roots of the equation $1 - \sum_{i=1}^{k} \varphi_i x^i = 0$ 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. $\diamond$

In Proposition 2 and Theorem 2 in Section 2.3, we assume that, for some $\kappa \in \mathbb{N}$, F_κ has bounded partial derivatives of certain order. Assume without loss of generality $a_0 \neq 0$. Clearly, if the distribution function of ε_0 has bounded derivatives up to order p, then so does F_1. An interesting example is the stable distribution with characteristic function

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

Here $\sqrt{-1}$ is the imaginary unit. By the inversion formula, ε_0 has bounded derivatives of all orders. Examples 5 provides other sufficient conditions under which $F_\kappa \in \mathcal{C}^p$.

Example 5. Let $\phi(t) = \mathbb{E}[\exp(\sqrt{-1}t\varepsilon_0)], t \in \mathbb{R}$, be the characteristic function of ε_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 $F_\kappa \in \mathcal{C}^p$. To this end, let

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

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

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

Let $r_k = \min_{i \leq k} \rho_i > 0$, $\kappa = n_{2k}$ with $k = \lfloor 4p/\lambda \rfloor + 4$. Then $|\phi_\kappa(t_1, t_2)| \leq c^{\kappa+1}[1 + r_k(|t_1| + |t_2|)^\lambda]^{-k}$. By the inversion formula, $F_\kappa \in \mathcal{C}^p$ 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

$$\Delta_r(g)(x, y) = \sum_{i=0}^{r} \frac{r!}{i!(r-i)!} g^{(r,i)}(x, y) \tag{17}$$

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

$$S_n(\mu; p) = \sum_{i=1}^{n} L_i(\mu_{i-1}, \mu_i; p) = C_n(\mu) + \sum_{i=1}^{n} \sum_{r=0}^{p} (-1)^{r+1} \Delta_r(G)(\mu_{i-1}, \mu_i) U_i(r), \tag{18}$$

where

$$L_i(x, y; p) = \mathbf{1}_{(X_{i-1}-x)(X_i-y)\leq 0} + \sum_{r=0}^{p}(-1)^{r+1}\Delta_r(G)(x, y)U_i(r), \qquad (19)$$

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

Note that $U_i(r)$ is a rth order polynomial in ε_j. Intuitively, $S_n(\mu; p)$ can be viewed as the pth order Taylor expansion of $C_n(\mu)$; see [14, 42] 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 $a > 1/2$, nonnegative integer $k < 1/(2a - 1)$ and a bounded function f, define the MWI

$$H_{k,a}(f) = \lambda(k, a)\int_{\mathbb{R}^k}\int_0^1 f(t)\prod_{i=1}^{k}[(t - u_i)^+]^{-a}dtdW(u_1)\ldots dW(u_k),$$

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 ε_i is in the domain of attraction of stable law with index $d \in (1, 2)$ (denoted by $\varepsilon_i \in \mathcal{D}(d)$) if there exists a slowly varying function L_ε 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 $\kappa \in \mathbb{N}$, $F_\kappa \in \mathcal{C}^{p+2}$. (i) Assume $\varepsilon_0 \in \mathcal{L}^q$, $q \geq 2$ and $2\beta \in (1, 1 + (p + 1)^{-1})$. Let $S_n(\mu; p)$ be as in (18). Then*

$$\frac{S_n(\mu; p)}{n^{1-(p+1)(\beta-1/2)}L^{p+1}(n)\|\varepsilon_0\|_2^{p+1}} \Rightarrow (-1)^{p+1}H_{p+1,\beta}[\Delta_{p+1}(G)(m_1(\cdot), m_1(\cdot))]. \qquad (20)$$

(ii) Let $\varepsilon_i \in \mathcal{D}(d)$, $d \in (1, 2)$. Assume that $1/d < \beta < 1$. Then there is a slowly varying function L_1 such that

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

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 $n^{3/2-\beta}L(n)\|\varepsilon_0\|_2$.

Example 6. Consider the FARIMA(k, θ, p) process X_n given by $\varphi(B)(1 - B)^\theta X_n = \psi(B)\varepsilon_n$. 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 $X_n = \sum_{i=0}^\infty a_i \varepsilon_{n-i}$, where

$$a_n = \frac{\Gamma(n + \theta)}{\Gamma(n + 1)\Gamma(\theta)} = n^{-(1-\theta)} \frac{\Gamma(n + \theta) n^{1-\theta}}{\Gamma(n + 1)\Gamma(\theta)} := n^{-(1-\theta)} L(n). \tag{21}$$

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, θ, 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_γ stand for generic positive constants which may vary among lines. Recall $\mathcal{F}_n = (\varepsilon_n, \varepsilon_{n-1}, \ldots)$, $\mathcal{F}'_n = (\varepsilon_n, \ldots, \varepsilon_1, \varepsilon'_0, \mathcal{F}_{-1})$, and the detail projection operator $\mathcal{P}_k \cdot = \mathbb{E}(\cdot | \mathcal{F}_k) - \mathbb{E}(\cdot | \mathcal{F}_{k-1}), k \in \mathbb{Z}$. For $m \leq n$ let $\mathcal{F}_n^m = (\varepsilon_n, \varepsilon_{n-1}, \ldots, \varepsilon_m)$.

3.1   Proof of Theorem 1.

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

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

$$= \sum_{j=-\infty}^i |\mathbb{E}[(\mathcal{P}_j Z_i)(\mathcal{P}_j Z_{i+r})]| \leq \sum_{j=-\infty}^i \delta_{i-j} \delta_{i+r-j} \leq \sum_{j=0}^\infty \delta_j \delta_{j+r} \tag{22}$$

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

$$\frac{1}{n} \|S_n\|_2^2 = \frac{1}{n} \sum_{|i-i'| \leq k} \text{cov}(Z_i, Z_{i'}) + \frac{1}{n} \sum_{|i-i'| > k} \text{cov}(Z_i, Z_{i'}) := I_{n,k} + J_{n,k}.$$

10

Let $s_k = \sum_{|\ell| \le k} \gamma_\ell$ and $t_k = \sum_{j=k}^\infty \delta_j$. By (3), $\lim_{n\to\infty} I_{n,k} = s_k$. So

$$
\begin{aligned}
\limsup_{n\to\infty} |\,\|S_n\|_2^2/n - \sigma^2| &\le \limsup_{n\to\infty} |I_{n,k} - s_k| + \limsup_{n\to\infty} |J_{n,k}| + |\sigma^2 - s_k| \\
&\le \sum_{r=k}^\infty \sum_{j=0}^\infty 2\delta_j \delta_{j+r} + |\sigma^2 - s_k| \le 2t_0 t_k + |\sigma^2 - s_k|.
\end{aligned}
$$

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

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

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

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

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

$$
\begin{aligned}
\mathcal{P}_0 \mathbb{E}(X|\mathcal{F}_i^{i-j}) &= \mathbb{E}[\mathbb{E}(X|\mathcal{F}_i^{i-j})|\mathcal{F}_0] - \mathbb{E}[\mathbb{E}(X|\mathcal{F}_i^{i-j})|\mathcal{F}_{-1}] \\
&= \mathbb{E}[\mathbb{E}(X|\mathcal{F}_i^{i-j})|\mathcal{F}_0^{i-j}] - \mathbb{E}[\mathbb{E}(X|\mathcal{F}_i^{i-j})|\mathcal{F}_{-1}^{i-j}] \\
&= \mathbb{E}(X|\mathcal{F}_0^{i-j}) - \mathbb{E}(X|\mathcal{F}_{-1}^{i-j}) = \mathbb{E}[(\mathcal{P}_0 X)|\mathcal{F}_0^{i-j}].
\end{aligned}
$$

Applying the preceding identity with $X = L_{i,m}(x,y)$, we have by Schwarz's inequality that $\sup_{x,y} \|\mathcal{P}_0 L_{i,m}(x,y)\|_2 \le \delta_i$ for all $m \ge 0$. So $\delta_{i,m} \le 2\delta_i$. On the other hand, $\delta_{i,m} \le \sup_{x,y\in\mathbb{R}} \|L_i(x,y) - L_{i,m}(x,y)\|_2 =: \eta_m$ (say). Therefore $\delta_{i,m} \le \min(2\delta_i, \eta_m)$. By Theorem 7.4.3 in [3], for any x,y, $\lim_{m\to\infty} \|\mathbf{1}_{X_0\le x, X_1\le y} - \mathbb{E}(\mathbf{1}_{X_0\le x, X_1\le y}|\mathcal{F}_1^{-m})\|_2 \to 0$. Hence by the continuity of F, we have $\sup_{x,y} \|\mathbf{1}_{X_0\le x, X_1\le y} - \mathbb{E}(\mathbf{1}_{X_0\le x, X_1\le y}|\mathcal{F}_1^{-m})\|_2 \to 0$ 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 $\|D_{i,j,m}\|_2 \le \delta_{i,m}$, $\|\sum_{i=1}^n D_{i,j,m}\|_2 \le \sqrt{n}\delta_{j,m}$. Let LIM denote $\limsup_{m\to\infty} \limsup_{n\to\infty}$. Then

$$
\begin{aligned}
\mathrm{LIM}\frac{\|S_n(\mu) - S_{n,m}\|_2}{\sqrt{n}} &= \mathrm{LIM}\frac{\|\sum_{i=1}^n \sum_{j=0}^\infty D_{i,j,m}\|_2}{\sqrt{n}} \\
&\le \mathrm{LIM}\frac{\sum_{j=0}^\infty \|\sum_{i=1}^n D_{i,j,m}\|_2}{\sqrt{n}} \le \lim_{m\to\infty} \Lambda_m = 0. \tag{24}
\end{aligned}
$$

Since $\sigma^2 > 0$, for sufficiently large m, $\liminf_{n\to\infty} \|S_{n,m}\|_2^2/n > \sigma^2/2 > 0$. 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 Rothensal's inequality (cf Theorem 1.5.11 in [6]). Details are omitted.

Lemma 1. *Let $\varepsilon_1 \in \mathcal{L}^\gamma$, $\gamma > 0$ 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_γ such that $\mathbb{E}|\sum_{i=1}^{n} b_i \varepsilon_i|^\gamma \leq c_\gamma (\sum_{i=1}^{n} |b_i|^{\gamma'})^{\gamma''/2}$ 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 $x, y, \delta_1, \delta_2 \in \mathbb{R}$,*

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

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 $f^{(1,0)}$,

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

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

Recall that $\underline{X}_{n,m} = \sum_{j=n-m}^{\infty} a_j \varepsilon_{n-j}$, $X_{n,m} = \sum_{j=0}^{n-m} a_j \varepsilon_{n-j}$ 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 $q' = q \wedge 4$, $A_n(k) = \sum_{i=n}^{\infty} |a_i|^k$, $B_n(k) = \sum_{i=n}^{\infty} |a_i - a_{i+1}|^k$,

$$\theta_{n,p,q} = |a_n - a_{n-1}| + |a_n|^{q'/2} + |a_n|[A_n^{p/2}(2) + A_n^{1/2}(q') + B_n^{1/2}(2)], \tag{25}$$

$\Theta_{n,p,q} = \sum_{k=1}^{n} \theta_{k,p,q}$ and $\Xi_{n,p,q} = n\Theta_{n,p,q}^2 + \sum_{i=1}^{\infty}(\Theta_{n+i,p,q} - \Theta_{i,p,q})^2$. Lemma 4 extends Lemma 10 in [42] to the bivariate case.

Lemma 3. *Assume that $F_\kappa \in \mathcal{C}^p$ 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 $0 \leq \gamma \leq r \leq p$,*

$$
\begin{aligned}
F_{n+m}^{(r,\gamma)}(x, y) &= \mathbb{E}F_m^{(r,\gamma)}\Big(x - \sum_{j=m}^{n+m-1} a_j \varepsilon'_{-j}, y - \sum_{j=m+1}^{n+m} a_j \varepsilon'_{1-j}\Big), \\
F^{(r,\gamma)}(x, y) &= \mathbb{E}F_m^{(r,\gamma)}\Big(x - \sum_{j=m}^{\infty} a_j \varepsilon'_{-j}, y - \sum_{j=m+1}^{\infty} a_j \varepsilon'_{1-j}\Big). \tag{26}
\end{aligned}
$$

12

Proof. For $r = 0$, a conditioning argument entails (26). For $r \geq 1$, (26) follows from the Lebesgue dominated convergence theorem. See Lemmas 6 and 7 in [42] for the details of the proof in the univariate case. $\diamondsuit$

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 $F_\kappa \in \mathcal{C}^{p+1}$ for some integer $\kappa > 0$. Then*

$$\sup_{x,y} \|\mathcal{P}_1 L_n(x, y; p)\|_2 = O(\theta_{n,p,q}).$$

Proof. For notational convenience we shall often drop the arguments x, y. For example, we write $F = F(x, y)$. Let $L_n^\circ(p) = \mathbf{1}_{X_{n-1} \leq x, X_n \leq y} + \sum_{r=0}^p (-1)^{r+1} \Delta_r(F)(x, y) U_n(r)$. It suffices to show that $\sup_{x,y} \|\mathcal{P}_1 L_n^\circ(p)\| = O(\theta_{n,p,q})$. We shall adopt the argument in the proof of Lemma 9 in [42]. As in [42], we may assume $n \geq \kappa + 1$. For $m \leq n$, let

$$u_m = u_{n,m} = (x - \underline{X}_{n,m}, y - \underline{X}_{n+1,m}), \quad u_m^* = u_{n,m}^* = (x - \underline{X}_{n,m}^*, y - \underline{X}_{n+1,m}^*),$$

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

$$\|F_n^{(r-1,\gamma)}(u_1) - F_n^{(r-1,\gamma)}(u_0) + F_n^{(r,\gamma+1)}(u_0)a_{n-1}\varepsilon_1 + F_n^{(r,\gamma)}(u_0)a_n\varepsilon_1\|_2^2$$
$$= O(|a_{n-1}|^{q'} + |a_n|^{q'}). \tag{27}$$

By Lemma 3,

$$F_n^{(r-1,\gamma)}(u_1) = \mathbb{E}[F_{n-1}^{(r-1,\gamma)}(u_1 - (a_{n-1}\varepsilon_1', a_n\varepsilon_1')) + F_{n-1}^{(r,\gamma+1)}(u_1)a_{n-1}\varepsilon_1' + F_{n-1}^{(r,\gamma)}(u_1)a_n\varepsilon_1'|\mathcal{F}_1].$$

Thus, by Schwarz's inequality and Lemma 2,

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

which together with (27) gives

$$\|F_{n-1}^{(r-1,\gamma)}(u_1) - F_n^{(r-1,\gamma)}(u_0) + F_n^{(r,\gamma+1)}(u_0)a_{n-1}\varepsilon_1 + F_n^{(r,\gamma)}(u_0)a_n\varepsilon_1\|_2^2$$
$$= O(|a_{n-1}|^{q'} + |a_n|^{q'}). \tag{28}$$

13

For $0 \leq \gamma + 1 \leq r \leq p$, define

$$M_n^{(r,\gamma)} = \Delta_1[F_n^{(r-1,\gamma)}(u_0)] + \sum_{i=r}^{p}(-1)^{i+r+1}\Delta_{i-r+1}(F^{(r-1,\gamma)})\mathbb{E}[U_{n+1}(i-r)|\mathcal{F}_0].$$

We shall use the induction argument to show that

$$\sup_{x,y} \|M_n^{(r,\gamma)}\|_2^2 = O[A_n^{p-r+1}(2) + A_n(q') + B_n(2)]. \tag{29}$$

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,

$$\begin{aligned}
\|F_n^{(r,\gamma)}(u_0) - F^{(r,\gamma)}(x,y)\|_2^2 &= \|\mathbb{E}[F_n^{(r,\gamma)}(u_0) - F_n^{(r,\gamma)}(u_0^*)|\mathcal{F}_0]\|_2^2 \\
&\leq \|F_n^{(r,\gamma)}(u_0) - F_n^{(r,\gamma)}(u_0^*)\|_2^2 = O[A_n(2)]. \tag{30}
\end{aligned}$$

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

$$\mathcal{P}_{-j}\Delta_1[F_n^{(r-2,\gamma)}(u_0)] = \Delta_1[F_{n+j}^{(r-2,\gamma)}(u_{-j}) - F_{n+j+1}^{(r-2,\gamma)}(u_{-j-1})]. \tag{31}$$

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

$$\begin{aligned}
I_n(j) &= \|\mathcal{P}_{-j}\Delta_1[F_n^{(r-2,\gamma)}(u_0)] + \Delta_1[F_{n+j+1}^{(r-1,\gamma+1)}(u_{-j-1})]a_{n+j}\varepsilon_{-j} \\
&\quad + \Delta_1[F_{n+j+1}^{(r-1,\gamma)}(u_{-j-1})]a_{n+j+1}\varepsilon_{-j}\|_2 = O(|a_{n+j}|^{q'/2} + |a_{n+j+1}|^{q'/2})
\end{aligned}$$

by (28) and (31) and

$$\begin{aligned}
J_n(j) &= \|\Delta_1[\Delta_1(F_{n+j+1}^{(r-2,\gamma)}(u_{-j-1}))]a_{n+j+1}\varepsilon_{-j} \\
&\quad + \sum_{i=r-1}^{p}(-1)^{i+r}\Delta_{i-r+2}[F^{(r-2,\gamma)}]\mathcal{P}_{-j}\mathbb{E}[U_{n+1}(i-r+1)|\mathcal{F}_0]\|_2.
\end{aligned}$$

Note that $\mathcal{P}_{-j}\mathbb{E}[U_{n+1}(i-r+1)|\mathcal{F}_0] = a_{n+j+1}\varepsilon_{-j}\mathbb{E}[U_{n+1}(i-r)|\mathcal{F}_{-j-1}]$ if $i \geq r$ and it vanishes if $i = r - 1$, and the decomposition $\Delta_{i-r+2}(F^{(r-2,\gamma)}) = \Delta_{i-r+1}[\Delta_1(F^{(r-2,\gamma)})]$.

14

Then $J_n(j) \leq c|a_{n+j+1}|[\|M_{n+j+1}^{(r,\gamma+1)}\|_2 + \|M_{n+j+1}^{(r,\gamma)}\|_2]$. Since $M_n^{(r-1,\gamma)} = \sum_{j=0}^{\infty} \mathcal{P}_{-j} M_n^{(r-1,\gamma)}$ and $\mathcal{P}_j, j \in \mathbb{Z}$, are orthogonal projections, the induction is completed by observing that

$$
\begin{aligned}
\|M_n^{(r-1,\gamma)}\|_2^2 &= \sum_{j=0}^{\infty} \|\mathcal{P}_{-j} M_n^{(r-1,\gamma)}\|_2^2 \\
&\leq 3 \sum_{j=0}^{\infty} [I_n(j)^2 + J_n(j)^2 + c|a_{n+j+1} - a_{n+j}|^2] \\
&= O[A_n^{p-r+2}(2) + A_n(q') + B_n(2)].
\end{aligned}
\tag{32}
$$

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

$$
\begin{aligned}
\|\mathcal{P}_1 L_{n+1}^{\circ}(p)\|_2 &= \|F_{n-1}(u_1) - F_n(u_0) + F_n^{(1,1)}(u_0)a_{n-1}\varepsilon_1 + F_n^{(1,0)}(u_0)a_n\varepsilon_1 \\
&\quad - a_n \varepsilon_1 M_n^{(1,0)} + F_n^{(1,1)}(u_0)(a_n - a_{n-1})\varepsilon_1\|_2.
\end{aligned}
\tag{33}
$$

$\diamondsuit$

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 $\varepsilon_0 \in \mathcal{L}^q$, $q > 0$. Assume that $F_\kappa \in \mathcal{C}^{p+1}$ 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 $(p+1)(2\beta-1) > 1$; $\sup_\mu \|S_n(\mu; p)\|_2 = O(\sqrt{n}) \sum_{i=0}^{n} |L^{p+1}(i)|/i$ 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 $\nu \in (1/\beta, q]$*

$$
\sup_\mu \|S_n(\mu; 1)\|_\nu = O[n^{1/\nu+(1-\beta q/\nu)^+} \sum_{i=1}^{n} |L^{q/\nu}(i)|/i + n^{1/\nu+(1+1/\nu-2\beta)^+} \sum_{i=1}^{n} |L^2(i)|/i].
$$

Proof. (i) Let $\lambda_n = \sup_{x,y} \|\mathcal{P}_1 L_n(x,y;p)\|_2$. Note that $S_n(\mu;p) = \sum_{j=-\infty}^{n} \mathcal{P}_j S_n(\mu;p)$. Since $\mathcal{P}_j, j \in \mathbb{Z}$, are orthogonal projections and $\mathcal{P}_j L_n(x,y;p) = 0$ for $j \geq n+1$, we have by Lemma 4 that

$$
\begin{aligned}
\|S_n(\mu;p)\|_2^2 &= \sum_{j=-\infty}^{n} \|\mathcal{P}_j S_n(\mu;p)\|_2^2 \leq \sum_{j=-\infty}^{n} \mathbb{E}\left\{ \sum_{i=1\vee j}^{n} \frac{[\mathcal{P}_j L_i(\mu_{i-1}, \mu_i; p)]^2}{\lambda_{i-j+1}} \sum_{i=1\vee j}^{n} \lambda_{i-j+1} \right\} \\
&\leq \sum_{j=-\infty}^{n} \left[\sum_{i=1\vee j}^{n} \lambda_{i-j+1} \right]^2 \leq \sum_{j=-\infty}^{n} \left[\sum_{i=1\vee j}^{n} \theta_{i-j+1,p,q} \right]^2 = O(\Xi_{n,p,q}).
\end{aligned}
$$

15

The argument in Lemma 5 in [42] shows that

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

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

$$\begin{aligned}
\sup_{x,y} \|\mathcal{P}_1 L_n(x,y;1)\|_\nu &= O[|a_n - a_{n-1}| + |a_n|^{q/\nu} + |a_n|A_n^{1/\nu}(\nu)] \\
&= O[n^{-\beta q/\nu}L^{q/\nu}(n) + n^{-2\beta+1/\nu}L^2(n)], \quad (34)
\end{aligned}$$

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

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

in view of Lemma 2; (28) now becomes

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

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

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

in view of Lemma 1. So (34) follows from (33). By the von Bahr-Esséen inequality,

$$\begin{aligned}
\|S_n(\mu;1)\|_\nu^\nu &\leq 2\sum_{j=-\infty}^{n} \|\mathcal{P}_j S_n(\mu;1)\|_\nu^\nu \leq 2\sum_{j=-\infty}^{n} \left[\sum_{i=1\vee j}^{n} \sup_{x,y} \|\mathcal{P}_1 L_{i-j+1}(x,y;1)\|_\nu\right]^\nu \\
&= O(n^{1+(\nu-\beta q)^+}[\sum_{i=1}^{n} |L^{q/\nu}(i)|/i]^\nu + n^{1+(1+\nu-2\beta\nu)^+}[\sum_{i=1}^{n} |L^2(i)|/i]^\nu).
\end{aligned}$$

$$\Diamond$$

Proof of Theorem 2. (i). By Proposition 3, if $(p+2)(2\beta-1) < 1$, $\|S_n(\mu;p+1)\|_2 = O[n^{1-(p+2)(\beta-1/2)}L^{p+2}(n)]$; if $(p+1)(2\beta-1) < 1$ and $(p+2)(2\beta-1) \geq 1$, $\|S_n(\mu;p+$

16

$1)\|_2 = O(\sqrt{n})[1 + \sum_{i=0}^{n} |L^{p+1}(i)|/i]$. In either case, $\|S_n(\mu; p+1)\|_2 = o(\sigma_{n,p+1})$, where $\sigma_{n,p+1} = n^{1-(p+1)(\beta-1/2)} L^{p+1}(n)$. Thus

$$
\begin{aligned}
S_n(\mu; p) &= S_n(\mu; p) - S_n(\mu; p+1) + o_p(\sigma_{n,p+1}) \\
&= (-1)^{p+1} \sum_{i=1}^{n} \Delta_{p+1}(G)(\mu_{i-1}, \mu_i) U_i(p+1) + o_p(\sigma_{n,p+1}).
\end{aligned}
$$

We complete the proof by noting that

$$
\frac{\sum_{i=1}^{n} \Delta_{p+1}(G)(\mu_{i-1}, \mu_i) U_i(p+1)}{\sigma_{n,p+1} \|\varepsilon_0\|_2^{p+1}} \Rightarrow H_{p+1,\beta}[\Delta_{p+1}(G)(m_1(\cdot), m_1(\cdot))],
$$

which follows from the argument of Theorem 5.1 in [38] (see also Theorem 2 in [37]).

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

$$
\frac{\sum_{i=1}^{n} \Delta_1(G)(\mu_{i-1}, \mu_i) X_i}{(1-\beta)\ell_n} \Rightarrow \int_{-\infty}^{1} \Big[\int_{(-u)^+}^{1-u} \Delta_1(G)(m_1(u+t), m_1(u+t)) t^{-\beta} dt \Big] dZ_d(u).
$$

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

$$
\eta = \max\Big\{ \frac{3\beta d - 1}{\beta(\beta d + 1)}, \frac{d}{\beta(1 + d - \beta d)} \Big\} \text{ and } \zeta = \max\Big\{ \frac{2d}{\beta d + 1}, \frac{d}{1 + d - \beta d} \Big\},
$$

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 $\nu \in (\zeta, (\beta q - 1)/(\beta - 1/d))$. Then $1/\nu + (1 - \beta q/\nu)^+ < 1 - \beta + 1/d$ and $1/\nu + (1 + 1/\nu - 2\beta)^+ < 1 - \beta + 1/d$. Since L is slowly varying, by the argument of Lemma 5 in [42], both $\sum_{i=1}^{n} |L^{q/\nu}(i)|/i$ and $\sum_{i=1}^{n} |L^2(i)|/i$ are slowly varying functions. By Proposition 3 with the above constructed q and ν, $\|S_n(\mu; 1)\|_\nu = o(\ell_n)$, which completes the proof since $S_n(\mu; 0) = S_n(\mu; 1) - \sum_{i=1}^{n} \Delta_1(G)(\mu_{i-1}, \mu_i) X_i$. $\Diamond$

Acknowledgments

References

[1] S. Benzaquen, E. M. Cabaña, The expected measure of the level sets of a regular stationary Gaussian process, Pacific J. Math. 103 (1982) 9-16.

[2] E.V. Bulinskaya, On the mean number of crossings of a level by a stationary Gaussian process, Theory Probab. Appl. 6 (1961) 435-438.

[3] Y.S. Chow, H. Teicher, Probability theory, Springer-Verlag, New York, 1998.

[4] J. Cuzick, A central limit theorem for the number of zeros of a stationary Gaussian process, Ann. Probab. 4 (1976) 547-556.

[5] J. Dedecker, C. Prieur, New dependence coefficients. Examples and applications to statistics, Probab. Th. Rel. Fields 132 (2005) 203-236.

[6] V. De La Peña, E. Giné, Decoupling: from dependence to independence: randomly stopped processes, U-statistics and processes, martingales and beyond, Springer-Verlag, New York, 1999.

[7] P. Diaconis, D. Freedman, Iterated random functions, SIAM Rev. 41 (1999) 41-76.

[8] M. Duflo, Random iterative models, Springer-Verlag, Heidelberg, Germany, 1997.

[9] R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation, Econometrica 50 (1982) 987-1007.

[10] J. Fan and Q. Yao, Nonlinear time series: nonparametric and parametric methods, Springer, New York, 2003.

[11] B. Gajic, K.K. Paliwal, Robust speech recognition using features based on zero crossings with peak amplitudes, Proc. IEEE Intern. Conf. on Acoustics, Speech and Signal Processing, vol. I, pp. 64-67, Hong Kong, April, 2003.

[12] L. Giraitis, H. L. Koul and D. Surgailis, Asymptotic normality of regression estimators with long memory errors, Statist. Probab. Lett. 29 (1996) 317-335.

[13] V. Haggan, T. Ozaki, Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model, Biometrika 68 (1981) 189-196.

[14] H.-C. Ho, T. Hsing, On the asymptotic expansion of the empirical process of long-memory moving averages, Ann. Statist. 24 (1996) 992-1024.

[15] H.-C. Ho, T. Hsing, A decomposition for generalized U-statistics of long-memory processes, in: P. Doukhan, G. Oppenheim, M.S. Taqqu, M.S. (Eds.), Theory and Applications of Long-range Dependence, Birkhäuser, Boston, 2003, pp. 143-155.

[16] W. Hoeffding, H. Robbins, The central limit theorem for dependent random variables, Duke Math. J. 15 (1948) 773-780.

[17] G. Kallianpur, Some ramifications of Wieners ideas on nonlinear prediction. In Norbert Wiener, Collected Works with Commentaries, MIT Press, MA, 1981, pp. 402-424.

[18] Y. Kasahara, M. Maejima, Weighted sums of i.i.d. random variables attracted to integrals of stable processes, Probab. Th. Rel. Fields 78 (1988) 75-96.

[19] B. Kedem, On the sinusoidal limit of stationary time series, Ann. Statist. 12 (1984) 665-674.

[20] B. Kedem, Higher-order crossings in time series model identification, Technometrics 29 (1987) 193-204.

[21] B. Kedem, Time series analysis by higher order crossings, IEEE press, New York, 1994.

[22] M.F. Kratz, J.R. León, Hermite polynomial expansion for non-smooth functionals of stationary Gaussian processes: crossings and extremes, Stoch. Proc. Appl. 66 (1997) 237-252.

[23] M.F. Kratz, J.R. León, Curve crossings and specular points, d'après Longuet-Higgins, preprint, 2005.

[24] L. Kuipers, H. Niederreiter, Uniform distribution of sequences, Wiley, New York, 1974.

[25] G. Letac, A contraction principle for certain Markov chains and its applications, in: Random matrices and their applications, Contemp. Math., 50, Amer. Math. Soc., Providence, RI, 1986, pp. 263–273.

[26] M.S. Longuet-Higgins, Reflection and refraction at a random moving surface, I, II, III, J. Opt. Soc. Am. 50 (1960) 838-856.

[27] P. Major, Multiple Wiener-Itô integrals: with applications to limit theorems, Springer, New York, 1981.

[28] A. Mokkadem, Propriétés de mélange des processus autorgréssifs polynomiaux, Ann. Inst. H. Poincaré Probab. Statist. 26 (1990) 219-260.

[29] D.F. Nicholls, B.G. Quinn, Random coefficient autoregressive models: an introduction, Springer-Verlag, New York, 1982.

[30] R.J. Niederjohn, P.F. Castelaz, Zero-crossing analysis methods for speech recognition, Proc. 1978 Conf. on Pattern Recognition and Image Processing, vol. CH1318 507-513 IEEE, New York, 1978.

[31] S.O. Rice, Mathematical analysis of random noise, Bell Syst. Tech. J. 24 (1945) 46-156.

[32] J.P. Romano, M. Wolf, A more general central limit theorem for m-dependent random variables with unbounded m, Statist. Probab. Lett. 47 (2000) 115-124.

[33] M. Rosenblatt, Stationary processes as shifts of functions of independent random variables, J. Math. Mech. 8 (1959) 665-681.

[34] J. Saunders, Real-time discrimination of broadcast speech/music, Proc. ICASSP'96, vol. II, Atlanta, GA. 993-996, 1996.

[35] K. Shimizu, M. Tanaka, Expected number of level-crossings for a strictly stationary ellipsoidal process, Statist. Probab. Lett. 64 (2003) 305-310.

[36] E.V. Slud, MWI representation of the number of curve-crossings by a differentiable Gaussian process, with applications, Ann. Probab. 22 (1994) 1355-1380.

[37] D. Surgailis, Zones of attraction of self-similar multiple integrals, Lithuanian Mathematical J. 22 (1982) 327-340.

[38] D. Surgailis, Non-CLTs: U-statistics, multinomial formula and approximations of multiple Itô-Wiener integrals, in: P. Doukhan, G. Oppenheim, M.S. Taqqu, M.S. (Eds.), Theory and Applications of Long-range Dependence, Birkhäuser, Boston, 2003, pp. 129-142.

[39] H. Tong, Nonlinear time series Analysis: a dynamic approach, Oxford University Press, Oxford, 1990.

[40] W.B. Wu, Central limit theorems for functionals of linear processes and their applications, Statist. Sinica 12 (2002) 635-649.

[41] W.B. Wu, Additive functionals of infinite-variance moving averages, Statist. Sinica 13 (2003) 1259-1267.

[42] W.B. Wu, Empirical processes of long-memory sequences, Bernoulli 9 (2003) 809-831.

[43] W.B. Wu, W. Min, On linear processes with dependent innovations, Stoch. Proc. Appl. 115 (2005) 939–958.

[44] W.B. Wu, X. Shao, Limit theorems for iterated random functions, J. Appl. Probab. 41 (2004) 425-436.

[45] N.D. Ylvisaker, The expected number of zeros of a stationary Gaussian process, Ann. Math. Statist. 36 (1965) 1043-1046.