The Annals of Probability
1990, Vol. 18, No. 3, 1361–1387

WEAK CONVERGENCE OF SERIAL RANK STATISTICS UNDER DEPENDENCE WITH APPLICATIONS IN TIME SERIES AND MARKOV PROCESSES

BY MICHEL HAREL AND MADAN L. PURI[1]

Institut Universitaire de Technologie de Limoges and U.A. C.N.R.S., Rouen, and Indiana University

The asymptotic normality of linear serial rank statistics introduced by Hallin, Ingenbleek and Puri (1985) for the problem of testing white noise against ARMA alternatives is established for φ-mixing as well as strong mixing sequences of random variables. Applications in Markov processes and ARMA processes in time series are provided.

1. Introduction. Let $\{X_{n,i},\ 1 \le i \le n,\ n \ge 1\}$ be real-valued random variables with continuous distribution functions $F_n(x) = P(X_{n,i} \le x)$, $1 \le i \le n$, $n \ge 1$. Consider the statistics

$$(1.1)\qquad \mathscr{S}_n = (n-k)^{-1} \sum_{i=k+1}^{n} c_{n,i} a_n(R_{n,i-k}, \dots, R_{n,i}),$$

where the $c_{n,i}$ are known constants, $a_n(\cdot, \dots, \cdot)$ are the scores, $R_{n,i}$ denotes the rank of $X_{n,i}$ among $(X_{n,1}, \dots, X_{n,n})$ and $k \ge 1$ is a fixed integer $(< n)$. Our aim is to study the asymptotic behavior of $\mathscr{S}_n$ when the sequence $\{X_{n,i}\}$ is φ-mixing with rates

$$(1.2)\qquad \varphi(m) = O(m^{-1-\varepsilon}) \quad \text{for some } \varepsilon > 0,\ m \ge 1$$

or

$$(1.3)\qquad \sum_{m=1}^{\infty} m^{-1} \varphi^{1/2(3+k)}(m) < \infty,$$

or is strong-mixing with rates

$$(1.4)\qquad \sum_{m=1}^{\infty} m^{2(k+2)} \alpha^{\varepsilon}(m) < \infty \quad \text{for some } \varepsilon \in (0, 1/2(3+k)).$$

Recall that the array $\{X_{n,i},\ 1 \le i \le n,\ n \ge 1\}$ is φ-mixing if

$$\sup_{m \le n} \sup_{1 \le j \le n-m} \{|P(A|B) - P(A)| \colon B \in \sigma(X_{n,i}, 1 \le i \le j),$$
$$A \in \sigma(X_{n,i}, i \ge j+m)\} = \varphi(m) \downarrow 0$$

Received February 1988; revised July 1989.
[1]Research supported by Office of Naval Research Contract N00014-85-K-0648.
AMS 1980 *subject classifications*. 60F05, 60J05, 62M10.
Key words and phrases. Serial rank statistics, φ-mixing, strong mixing, graduate empirical process, graduate rank process, Skorohod topology, weak convergence, Markov process, ARMA process.

as $m \uparrow \infty$ for positive integers j and m, and it is strong-mixing if

$$\sup_{m \le n} \sup_{1 \le j \le n-m} \{|P(A \cap B) - P(A)P(B)|, A \in \sigma(X_{n,i}, 1 \le i \le j), B \in \sigma(X_{n,i}, i \ge j+m)\} = \alpha(m) \downarrow 0 \quad \text{as } m \uparrow \infty,$$

for positive integers j and m. Here $\sigma(X_{n,i}, i \le j)$ and $\sigma(X_{n,i}, i \ge j+m)$ are the σ-fields generated by $(X_{n,1}, \dots, X_{n,j})$ and $(X_{n,j+m}, X_{n,j+m+1}, \dots)$, respectively. The asymptotic behavior of the statistic $\mathscr{S}_n$ under strong-mixing conditions leads to interesting applications in ARMA processes in time series as well as in Markov processes (Section 6). In passing we may mention that Hallin, Ingenbleek and Puri (1985) established the asymptotic normality of linear *serial* rank statistics $\mathscr{S}_n$ defined in (1.1) for an ARMA process contiguous to white noise. We show (in Section 2) that contiguity is not necessary for the derivation of the asymptotic distribution theory derived in Hallin, Ingenbleek and Puri (1985) and our results also lead to applications in some Markov processes which are either geometrically ergodic or Doeblin recurrent, and to some ARMA processes. For a related problem dealing with the applications of U-statistics [see Harel and Puri (1989a, 1990)] to some Markov processes and ARMA models, the reader is referred to Harel and Puri (1989b).

2. Asymptotic normality. We start with a few preliminaries.

Denote by $\hat{F}_n(x)$, the right continuous empirical distribution function of $X_{n,i}$, $i = 1, \dots, n$; i.e., let $\hat{F}_n(x) = n^{-1}\sum_{i=1}^n I_{\{X_{n,i} \le x\}}$ where $I_{\{\cdot\}}$ denotes the indicator function. Denote by G_n the distribution function of the $k+1$ of the successive random variables $X_{n,1}, \dots, X_{n,n}$. Let H_n (for each $n \ge k+1$) be a sequence of continuous distribution functions on $(0,1)^{k+1}$, defined by

$$H_n(\mathbf{t}) = G_n(F_n^{-1}(t_1), \dots, F_n^{-1}(t_{k+1})) \quad \text{for all } \mathbf{t} = (t_1, \dots, t_{k+1}) \in (0,1)^{k+1}, \tag{2.1}$$

where $F_n^{-1}(u) = \inf\{t\colon F_n(t) \ge u\}$, $0 < u < 1$. Since H_n is continuous, it is actually well defined on $[0,1]^{k+1}$. Though G_n, H_n and $\mathbf{t}$ depend on k, we have suppressed this fact for notational convenience.

Denote by C_{k+2}, the space of all continuous maps $f\colon [0,1]^{k+2} \to \mathbb{R}$, and by $C_{k+1}(j)$, $1 \le j \le k+1$, the space of all continuous and bounded maps $f\colon A(j) \to \mathbb{R}$, where $A(j) = [0,1]^{j-1} \times (0,1) \times [0,1]^{k+1-j}$.

Definition. We say that the sequence $\{H_n\}$ satisfies the differentiability condition if (a) $\partial H_n/\partial t_j$ exists on $A(j)$ and belongs to $C_{k+1}(j)$, $1 \le j \le k+1$, and (b) $\partial H_n/\partial t_j \to l_j$ in the uniform topology on any compact subset of $A(j)$ as $n \to \infty$, and l_j belongs to $C_{k+1}(j)$.

We define the *graduate empirical process* [also called the *copula process*, see, e.g., Gaenssler and Stute (1987), Chapter V] W_n as

$$(2.2)\qquad W_n(t) = (n-k)^{-1/2} \sum_{i=k+1}^{[nt_0]} \left\{ \prod_{j=1}^{k+1} I_{\{F_n(X_{n,i+j-k-1}) \le t_j\}} - H_n(\mathbf{t}) \right\}$$

for all $t = (t_0, \mathbf{t}) = (t_0, t_1, \dots, t_{k+1}) \in (0,1)^{k+2}$, where $[nt_0]$ denotes the integral part of the real number nt_0.

We also consider the rank process L_n (called the *graduate rank process*) defined as

$$(2.3)\qquad L_n(t) = (n-k)^{-1/2} \sum_{i=k+1}^{[nt_0]} \left\{ \prod_{j=1}^{k+1} I_{\{\hat{F}_n(X_{n,i+j-k-1}) \le t_j\}} - H_n(\mathbf{t}) \right\}.$$

For any $n \ge 1$, we define a signed measure λ_n concentrated on $\{1/n, \dots, (n-1)/n, 1\}^{k+2}$ by setting

$$\lambda_n\left(\prod_{j=0}^{k+1} \left[\frac{i_j}{n}, 1 \right] \right) = c_{n,i_0} a_n(i_1, \dots, i_{k+1}),$$

for all $(i_0, \dots, i_{k+1}) \in \{1, \dots, n\}^{k+2}$. (By convention, $c_{n,i_0} = 0$ if $i_0 < k+1$.)

We also define a centering coefficient b_n by

$$(2.4)\qquad b_n = \int_{[0,1]^{k+2}} \hat{H}_n(t) \lambda_n(dt),$$

where $\hat{H}_n$ is the function $[0,1]^{k+2} \to \mathbb{R}^+$ such that $\hat{H}_n(t) = ([nt_0] - k)H_n(\mathbf{t})$.

We now state the following theorem, the proof of which is given in Section 5.

THEOREM 2.1. *Assume that there exists a Radon measure λ_0 on $[0,1]^{k+2}$ such that*

$$(2.5)\qquad \lim_{n\to\infty} \int f\, d\lambda_n = \int f\, d\lambda_0 \quad \textit{for all } f \in C_{k+2}$$

and

$$(2.6)\qquad \sup_{n\in\mathbb{N}} \int f\, d|\lambda_n| < \infty; \qquad \mathbb{N} = \{0, 1, 2, \dots\},$$

where $|\lambda_n|$ denotes the measure of total variation.

Assume that the sequence $\{X_{n,i}\}$ is (a) *φ-mixing with rates* (1.2) *or* (b) *strong-mixing with rates* (1.4). *Furthermore, assume that* (c) *the covariance functions $\{K_n,\ n \ge 1\}$ of the empirical processes $\{W_n,\ n \ge 1\}$ defined in* (2.2) *converge to a function $K(\cdot,\cdot)$ in pointwise topology as $n \to \infty$ and* (d) *$\{H_n\}$ satisfies the differentiability conditions. Then L_n converges weakly in uniform topology to a Gaussian process L_∞ with trajectories a.s. in C_{k+2}, and*

$(n-k)^{1/2}(\mathscr{S}_n - b_n)$ converges in law to the normal distribution with mean 0 *and variance σ^2, where*

$$(2.7)\qquad \sigma^2 = \int_{[0,1]^{k+2}} \cdots \int_{[0,1]^{k+2}} E[L_\infty(t), L_\infty(t')]\, d\lambda_0(t)\, d\lambda_0(t') < \infty.$$

REMARK 2.1. Theorem 2.1 is proved under the assumption that the sequence $\{X_{n,i}\}$ is nonstationary and either φ-mixing with rates (1.2) or strong-mixing with rates (1.4). The theorem does not hold with the φ-mixing rates (1.3) unless one assumes stationarity (which implies that the distribution functions F_n, G_n and H_n are equal to unique distribution functions F, G and H, respectively) and the special case when $c_{n,i} = 1$ for all i.

Let $\tilde{\mathscr{S}}_n$ denote the statistics $\mathscr{S}_n$ when $c_{n,i} = 1$ for all i, i.e., let

$$(2.8)\qquad \tilde{\mathscr{S}}_n = \sum_{i=k+1}^{n} a_n(R_{n,i-k}, \dots, R_{n,i}),$$

and let $\tilde{b}_n$ denote the corresponding centering constant, i.e.,

$$(2.9)\qquad \tilde{b}_n = \int_{[0,1]^{k+1}} H_n(\mathbf{t}) \tilde{\lambda}_n(d\mathbf{t}),$$

where $\tilde{\lambda}_n$ is a measure concentrated on $\{1/n, \dots, (n-1)/n, 1\}^{k+1}$ and

$$\tilde{\lambda}_n\left(\prod_{j=1}^{k+1}\left[\frac{i_j}{n}, 1\right]\right) = a_n(i_1, \dots, i_{k+1}).$$

Then, we have the following theorem.

THEOREM 2.2. *Assume there exists a Radon measure $\tilde{\lambda}_0$ on* $[0,1]^{k+1}$ *such that*

$$(2.10)\qquad \lim_{n\to\infty} \int \tilde{f}\, d\tilde{\lambda}_n = \int \tilde{f}\, d\tilde{\lambda}_0$$

and

$$(2.11)\qquad \sup_{n\in\mathbb{N}} \int \tilde{f}\, d|\tilde{\lambda}_n| < \infty,$$

where $|\tilde{\lambda}_n|$ denotes the measure of total variation.

Assume that the sequence $\{X_{n,i}\}$ is (a′) *φ-mixing with rates* (1.3) *and* (b′) *H satisfies the differentiability condition. Then $\bar{L}_n(1, \mathbf{t})\cdot$ converges weakly in uniform topology to a Gaussian process $\tilde{L}_\infty$ with trajectories a.s. in C_{k+1}, and $(n-k)^{1/2}(\tilde{\mathscr{S}}_n - \tilde{b}_n)$ converges in law to the normal distribution with mean* 0 *and variance $\tilde{\sigma}^2$, where*

$$(2.12)\qquad \tilde{\sigma}^2 = \int_{[0,1]^{k+1}} \cdots \int_{[0,1]^{k+1}} E[\tilde{L}_\infty(\mathbf{t}), \tilde{L}_\infty(\mathbf{t}')]\, d\tilde{\lambda}_0(\mathbf{t})\, d\tilde{\lambda}_0(\mathbf{t}') \qquad (< \infty).$$

The proof follows from Theorem 2.1 by putting $t_0 = 1$ for the processes W_n and L_n, and showing that the finite projections of W_n converge to a normal law (the proof of which is given in Proposition 3.4).

The following corollary gives sufficient conditions under which the conditions (2.5) and (2.6) are satisfied.

COROLLARY 2.2. *Let J be a function on $[0,1]^{k+2}$ such that*

$$J(i_0/n, \dots, i_{k+1}/n) = c_{n,i_0} a_n(i_1, \dots, i_{k+1})$$

for all $(i_0, \dots, i_{k+1}) \in \{1, \dots, n\}^{k+2}$, $J = J_d + J_c$, where J_d is a step function taking only a finite number of jumps, and where for any $I \subset \{0, \dots, k+1\}$, J_c has a continuous derivative $\partial^I J_c / (\partial t_j)_{j \in I}$, then the conditions (2.5) *and* (2.6) *are satisfied.*

PROOF. It suffices to prove the above corollary in the case when J_d has only one jump, say at $a = (a_0, \dots, a_{k+1}) \in [0,1]^{k+2}$. Let λ'_n and λ''_n be measures on $[0,1]^{k+2}$ defined by

$$\lambda'_n\left(\prod_{j=0}^{k+1}\left[\frac{i_j}{n}, 1\right]\right) = J_c\left(\frac{i_0}{n}, \dots, \frac{i_{k+1}}{n}\right)$$

and

$$\lambda''_n\left(\prod_{j=0}^{k+1}\left[\frac{i_j}{n}, 1\right]\right) = J_d\left(\frac{i_0}{n}, \dots, \frac{i_{k+1}}{n}\right),$$

for all $(i_0, \dots, i_{k+1}) \in \{1/n, \dots, (n-1)/n, 1\}^{k+2}$.

It is easy to check that

$$\lim_{n \to \infty} \int_{[0,1]^{k+2}} f \, d\lambda'_n = \sum_{I \subset \{0, \dots, k+1\}} \int_{[0,1]^{k+2}} f \frac{\partial^I J_c}{(\partial t_j)_{j \in I}} \left((t_j)_{j \in I}, (1)^{k+2-i}\right),$$

for all $f \in C_{k+2}$ whose $i = \operatorname{card} I$.

Thus, we obtain a measure λ'_0 satisfying

$$\lim_{n \to \infty} \int_{[0,1]^{k+2}} f \, d\lambda'_n(t) = \int_{[0,1]^{k+2}} f \, d\lambda'_0(t).$$

Analogously, we obtain

$$\lim_{n \to \infty} \int_{[0,1]^{k+2}} f \, d\lambda''_n(t) = f(a) \sum_{I \subset \{0, \dots, k+1\}} (-1)^i J_d\Big(\underset{i \in I}{(a_{i^-})}, \underset{i \notin I}{(a_{i^+})}\Big),$$

for all $f \in C_{k+2}$, where $i = \operatorname{card} I$. □

3. Weak convergence of the graduate empirical process and the graduate rank process. We start with preliminaries.

3.1A. *The spaces D_{k+2} and C_{k+2}.* Let f: $[0,1]^{k+2} \to \mathbb{R}$. For $\rho \in \{0,1\}^{k+2}$, define

$$f_\rho(t) = \lim_{\substack{s_i \uparrow t_i,\ \rho(i)=1 \\ s_i \downarrow t_i,\ \rho(i)=0}} f(s); (s,t) \in \left([0,1]^{k+2}\right)^2, \qquad i = 0,1,\dots,k+1,$$

if it exists, in which case, call $f_\rho(t)$ the ρ-limit of f at t. Denote by D_{k+2}, the space of all maps f: $[0,1]^{k+2} \to \mathbb{R}$ such that for all $\rho \in \{0,1\}^{k+2}$, f_ρ exists and $f_\rho = f$ for $\rho = (0,\dots,0)$.

We say that we have *special Skorohod topology* on D_{k+2} if we have the uniform topology for the first coordinate and the J_1-Skorohod topology for the other coordinates. [For definition of Skorohod topology, cf. Skorohod (1956) and Billingsley (1968).]

We define a modulus of continuity for any bounded function f: $[0,1]^{k+2} \to \mathbb{R}^+$ to be denoted by $\omega(f,\delta)$, $(\delta > 0)$, by setting

$$\omega(f,\delta) = \sup_{(t,t') \in ([0,1]^{k+2})^2} |f(t) - f(t')|, \qquad \|t - t'\| < \delta, \tag{3.1}$$

where $\|t\| = \sup\{|t_j|, 0 \le j \le k+1\}$. Note that f belongs to C_{k+2} if and only if $\lim_{\delta \to 0} \omega(f,\delta) = 0$.

The following proposition, which is a variant of Theorem 1.2 of Dudley (1978), will be used to prove Proposition 3.4.

PROPOSITION 3.1. *Let Y_n be a process with values in D_{k+2} and measurable with respect to $\mathscr{U}_{k+2}$, the σ-field generated by the uniform topology (on D_{k+2}). Let P_n denote the law of Y_n. Then, there exists a probability measure P, with $P(C_{k+2}) = 1$, for which P_n converges weakly with respect to the uniform topology if and only if*

(a) *for all finite subsets U of $[0,1]^{k+2}$, $\phi_U(P_n)$ converges weakly to $\phi_U(P)$ (ϕ_U is the projection of D_{k+2} on $\mathbb{R}^U$),*

(b) $\forall\ \varepsilon > 0$, $\lim_{\delta \to 0} \limsup_{n \to \infty} P_n[\{f;\ \omega(f,\delta) \ge \varepsilon\}] = 0$.

The proof is given in the Appendix.

3.1B. *Grid accompanying a sequence of probability measures.* We call a *grid* T of $[0,1]^{k+2}$ a subset of $[0,1]^{k+2}$ such that $T = \prod_{j=0}^{k+1} T^{(j)}$, where $T^{(j)}$ is a finite subset of $[0,1]$ which includes 0 and 1.

We call a *pace* τ of a grid $T = \prod_{j=0}^{k+1} T^{(j)}$ the number $\tau = \max_{0 \le j \le k+1} \tau_j$, where $\tau_j = \max\{|t_j' - t_j|,\ t_j'$ and t_j are successive elements in $T^{(j)}\}$.

We denote the *lower boundary* of T by $\underleftarrow{T}$ where

$$\underleftarrow{T} = \bigcup_{j=0}^{k+1} \left(\prod_{l=0}^{j-1} T^{(l)} \times \{0\} \times \prod_{l=j+1}^{k+1} T^{(l)} \right).$$

We call *block* B of T any part of T in the form

$$B = \prod_{j=0}^{k+1} \{(t_j, t_j'] \cap T^{(j)}, \text{ where } t_j \text{ and } t_j' \text{ belong to } T^{(j)} \text{ and } t_j < t_j'\}.$$

We call *evaluation* $e_T^{(B)}$ of B into T, the operator $e_T^{(B)}$: $D_{k+2} \to \mathbb{R}^+$ such that

$$e_T^{(B)}(f) = \sum_{(\varepsilon_0, \dots, \varepsilon_{k+1}) \in \{0,1\}^{k+2}} (-1)^{\sum_{i=0}^{k+1} \varepsilon_i} f[(1-\varepsilon_0)t_0 + \varepsilon_0 t_0', \dots, (1-\varepsilon_{k+1})t_{k+1} + \varepsilon_{k+1} t_{k+1}']. \tag{3.2}$$

Let ν be a finite measure on $[0,1]^{k+2}$ and let T be a grid of $[0,1]^{k+2}$. We call *reduction* $\tilde{\nu}$ of ν on T the measure on T defined by

$$\tilde{\nu}(\{t\}) = \begin{cases} 0 & \text{if } t \in \underset{\leftarrow}{T}, \\ \nu\left(\prod_{j=0}^{k+1} (t_j', t_j]\right) & \text{if } t \notin \underset{\leftarrow}{T}, \end{cases}$$

where

$$t_j' = \max\{x;\, x \in T^{(j)};\, x < t_j,\, t_j \in T^{(j)}\}.$$

For any $\delta > 0$, we set

$$\omega_T(f, \delta) = \sup\{|f(t) - f(t')|;\, (t, t') \in T^2,\, \|t - t'\| \le \delta\}.$$

We say that a sequence $\{T_n\}_{n \in \mathbb{N}^*}$ of grids is asymptotically dense in $[0,1]^{k+2}$ if the pace τ_n of T_n satisfies $\lim_{n\to\infty} \tau_n = 0$, $\mathbb{N}^* = \mathbb{N} - \{0\}$, $\mathbb{N} = 0, 1, 2, \dots$.

Let P_n, $n \in \mathbb{N}^*$, be a sequence of probability measures on $(D_{k+2}, \mathscr{D}_{k+2})$, where $\mathscr{D}_{k+2}$ is the σ-field generated by the Skorohod topology (on D_{k+2}). We say that the sequence $\{T_n\}$ of grids accompanies the sequence $\{P_n\}$ if and only if $\forall\, \varepsilon > 0$, $\exists\, \varepsilon' > 0$ and $\forall\, \delta \in [0, \frac{1}{2})$, $\exists\, N_0 \in N^*$, we have

$$P_n\left[\{f \in D_{k+2},\, \omega(f, \delta) \ge \varepsilon \text{ and } \omega_{T_n}(f, 2\delta) < \varepsilon'\}\right] = 0 \qquad \forall\, n \ge N_0.$$

The following propositions [Propositions (3.2) and (3.3)] are variants of a result of Neuhaus (1971) [see, e.g., Theorems 2 and 4 in Balacheff and Dupont (1980)] and will be used in Section 4.

PROPOSITION 3.2. *Let P_n, $n \in \mathbb{N}$, be probability measures on $(D_{k+2}, \mathscr{D}_{k+2})$ such that the following conditions are satisfied*:

(3.3) *$\phi_U(P_n)$ converges weakly to some probability measure P_U on $\mathbb{R}^U$ for any finite subset U of $[0,1]^{k+2}$*

and

$$\forall\, \varepsilon > 0, \qquad \lim_{\delta \to 0} \limsup_{n\to\infty} P_n[f \in D_{k+2};\, \omega(f, \delta) \ge \varepsilon] = 0. \tag{3.4}$$

Then P_n converges weakly with respect to the Skorohod topology to some probability measure P and $P(C_{k+2}) = 1$.

PROPOSITION 3.3. *Let ν be a positive finite measure on $[0,1]^{k+2}$ with continuous marginals. Let P_n be a sequence of probability measures on $(D_{k+2}, \mathscr{D}_{k+2})$ such that $\forall\, n \in \mathbb{N}$, $P_n[f \in D_{k+2};\ f|[0,1]^{k+2} = 0] = 1$. Let T_n be a sequence of grids asymptotically dense in $[0,1]^{k+2}$ and accompanying P_n. Furthermore, suppose that for any block B_n of T_n,*

$$P_n\left[f \in D_{2+k};\ |e_{T_n}^{(B_n)}(f)| \geq \lambda\right] \leq \lambda^{-\gamma}(\tilde{\nu}_n(B_n))^{\beta}, \tag{3.5}$$

where $\tilde{\nu}_n$ is the reduction of ν on T_n, and $\beta > 1$ and $\gamma > 0$. Then, we have $\forall\, \varepsilon > 0$, $\exists\, \delta \in (0,1)$ and $N_0 \in \mathbb{N}$, such that

$$P_n\left[f \in D_{k+2};\ \omega(f, \delta) \geq \varepsilon\right] \leq \varepsilon \qquad \forall\, n \geq N_0. \tag{3.6}$$

3.2A. *Convergence of the graduate empirical processes.*

PROPOSITION 3.4. *Under the conditions* (a) *and* (c) *or* (b) *and* (c) *of Theorem* 2.1, *W_n converges weakly in the special Skorohod topology to a Gaussian process W_∞ with trajectories a.s. in C_{k+2}. Under the conditions* (a′) *of Theorem* 2.2, *$\tilde{W}_n = W_n(1, \mathbf{t})$ converges in the Skorohod topology to a Gaussian process $\tilde{W}_\infty$ with trajectories a.s. in C_{k+1}.*

3.2B. *Convergence of the graduate rank process.*

PROPOSITION 3.5. *Under the conditions* (a), (c) *and* (d) *or* (b), (c) *and* (d) *of Theorem* 2.1, *L_n converges weakly in uniform topology to a Gaussian process L_∞ with trajectories a.s. in C_{k+2}. Under the conditions* (a′) *and* (b′) *of Theorem* 2.2, *$\tilde{L}_n = L_n(1, \mathbf{t})$ converges weakly in uniform topology to a Gaussian process $\tilde{L}_\infty$ with trajectories a.s. in C_{k+1}.*

4. Proofs of Propositions 3.4 and 3.5. Our proofs of the Propositions 3.4 and 3.5 are based on the ideas of Balacheff and Dupont (1980), who considered the asymptotic normality of the truncated empirical processes under φ-mixing with rates $\sum_{m=1}^{\infty} m\varphi^{1/2}(m) < \infty$. Here in this paper, we consider the rates (1.2) and (1.3), which are slower than the one considered by them. In addition, we also derive results under strong mixing (1.4) which have not been considered in the literature. To establish their result, Balacheff and Dupont (1980) used a slight modification of an inequality due to Rüschendorf (1974) which is not applicable in our situation. Our proofs are based on the following two lemmas.

LEMMA 4.1. *Let the sequence $\{X_{n,i}\}$ of real-valued random variables (centered at its expectation) be φ-mixing with rates $\sum_{m=1}^{\infty} m^{-1}\varphi^{1/2q}(m) < \infty$, where q is an integer. Denote by N_n the number of indices i, $1 \leq i \leq n$, for which $X_{n,i}$ is not identical to zero. Set $S_n = \sum_{i=1}^{n} X_{n,i}$ and $\|X_{n,i}\|_l =$*

$(\int |X_{n,i}|^{2l}\, dP_n)^{1/2l}$. *Then, for any* $q \geq 1$, *there exists a constant* $C_q(\varphi)$ *depending only on* q *and* φ *such that*

$$E\big(S_n^{2q}\big) \leq C_q(\varphi) \sum_{l=1}^{q} N_n^{q/l}\Big(\sup_{1\leq j\leq n} \|X_{n,j}\|_l\Big)^{2q}. \tag{4.1}$$

The proof is a slight modification of Theorem 2.1 of Neumann (1982) and is sketched briefly in the Appendix.

LEMMA 4.2. *Let the sequence* $\{X_{n,i}\}$ *of real-valued random variables (centered at its expectation) be strong-mixing with rates* $\sum_{m=1}^{\infty} m^{2q-2}\alpha^{\varepsilon}(m) < \infty$, $\varepsilon \in (0, 1/2q)$ *and* $|X_{n,i}| \leq 1$, $1 \leq i \leq n$, $n \geq 1$, *where* q *is an integer. Let* N_n *be the number of indexes* i, $1 \leq i \leq n$, *for which* $X_{n,i}$ *is not identical to zero. Set* $S_n = \sum_{i=1}^{n} X_{n,i}$ *and* $\|X_{n,i}\|_{\varepsilon} = (\int |X_{n,i}|^{2/(1-\varepsilon)}\, dP_n)^{1-\varepsilon}$. *Then, for any* $q \geq 1$, *there exists a constant* $C_q(\alpha)$ *depending only on* q *and* α *such that*

$$E\big(S_n^{2q}\big) \leq C_q(\alpha) \sum_{l=1}^{q} N_n^{l}\Big(\sup_{1\leq i\leq n} \|X_{n,i}\|_{\varepsilon}\Big)^{l}. \tag{4.2}$$

The proof is similar to that of Theorem II.10 of Doukhan and Portal (1987) and is therefore omitted.

LEMMA 4.3 (Neumann, 1982). *Let* $\{Y_i,\ i \geq 1\}$ *be a stationary sequence of real-valued random variables centered at its expectation and with finite second moment. Assume that the sequence is* φ*-mixing with rates* $\sum_{m=1}^{\infty} m^{-1}\varphi^{1/2}(m) < \infty$. *Then, there exists a positive constant* K *such that* $n^{-1}E(\sum_{i=1}^{n} Y_i)^2 \to K^2$ *as* $n \to \infty$.

Since the reference Neumann (1982) is not readily available, we have (at the suggestion of one of the referees) given the proof in the Appendix.

PROOF OF PROPOSITION 3.4. Consider a sequence $Z_{m,i}$, $1 \leq i \leq m$, $m \geq 1$ of $\mathbb{R}^{k+1}$-valued random variables defined by $Z_{m,i} = (X_{m+k,i}, \dots, X_{m+k,i+k}) = (Z_{m,i}^{(1)}, \dots, Z_{m,i}^{(k+1)})$, $1 \leq i \leq m$, $m \geq 1$. Then the $(k+1)$-variate truncated empirical process $\tilde{W}_m$ associated with this sequence is given by

$$\tilde{W}_m(t_0, \mathbf{t}) = m^{-1/2} \sum_{i=1}^{[(m+k)t_0]-k} \left[\prod_{j=1}^{k+1} I_{\{F_{m+k}(Z_{m,i}^{(j)}) \leq t_j\}} - H_{m+k}(\mathbf{t})\right] \tag{4.3}$$

and this is the same as the graduate process W_n defined in (2.2). Now the process W_n defines a probability measure Q_n on $(D_{k+2}, \mathscr{D}_{k+2})$.

To prove this proposition we have to verify (3.3) and (3.4). Following Withers [(1975), Corollary 1], it can be shown that $\phi_U(Q_n)$ converges weakly to a Gaussian measure Q_U if (i) $K_n \to$ some function K, (ii) $\sum_{m\geq 1}\alpha(m) < \infty$ and (iii) $m^{1-a}\alpha([m^b]) \to 0$ as $m \to \infty$, where $0 < 2b < a < 1 - b$. Now, in our situation (i) holds by assumption (c), (ii) follows from (1.2) or (1.4) and (iii) follows from (1.2) or (1.4) by taking $a = 3/4 - \varepsilon/8$, $b = 1/4$ and ε sufficiently small. (Since taking $\alpha(m) = m^{-1-\varepsilon}$, $m^{1-a}\alpha([m^b]) \le Am^{-\varepsilon/8} \to 0$ as $m \to \infty$.) Thus, (3.3) holds whenever conditions (a) and (c) or (b) and (c) of Theorem 2.1 are satisfied.

Now suppose that the condition (a′) of Theorem 2.2 holds with $X_{n,i} \equiv X_i$. Then, for any $p \in \mathbb{N}^*$, any $\mathbf{t}^{(l)} \in [0,1]^{k+1}$ and any $\lambda_l \in \mathbb{R}$, $1 \le l \le p$, let $g_i^{(l)}(X_i)$ and $g_i(X_i)$ be the random variables defined by

$$g_i^{(l)}(X_i) = \prod_{j=1}^{k+1} \left[I_{\{F(X_{i+j-k-1}) \le t_j\}} - H(\mathbf{t}^{(l)}) \right] \quad \text{and} \quad g_i(X_i) = \sum_{l=1}^{p} \lambda_l g_i^{(l)}(X_i),$$

where $X_i = (X_{i-k}, X_{i-k+1}, \dots, X_i)$. Then, we have $\sum_{l=1}^{p} \lambda_l \tilde{W}_n(\mathbf{t}^{(l)}) = (n-k)^{-1/2} \sum_{i=k+1}^{n} g_i(X_i)$, and so (3.3) also holds by Lemma 4.3 and the central limit theorem for the stationary and φ-mixing case [cf. Ibragimov and Linnik (1971), Theorem 18.5.1 and Lemma 4.3]. Now, to prove (3.4), we shall use Proposition 3.3 and verify (3.5) [which will imply (3.4)].

Let $T_n = \{i/m;\ 0 \le i \le m\}^{k+2}$ be a sequence of grids with $n = m + k$. T_n is asymptotically dense in $[0,1]^{k+2}$ and we prove that T_n accompanies Q_n. Now for every $\mathbf{t} \in (0,1)^{k+1}$, let $(\underline{\mathbf{t}}, \bar{\mathbf{t}})$ be the points of $\pi(T_n)$, where π is the projection defined by $\pi(t) = \mathbf{t}$ such that $\underline{\mathbf{t}} \le \mathbf{t} \le \bar{\mathbf{t}}$ and $\|\bar{\mathbf{t}} - \underline{\mathbf{t}}\| \le 1/m$. Let us write $\underline{t}_0 = [nt_0]/n$ for every $t_0 \in [0,1]$. As the marginals of H_n are uniform, we obtain (after some computations) that

$$|W_n(t_0, \mathbf{t}) - W_n(t_0', \mathbf{t}')| \le \frac{2K}{\sqrt{m}} + |W_n(\underline{t}_0, \bar{\mathbf{t}}) - W_n(t_0', \underline{\mathbf{t}}')|,$$

$\forall\ (t_0, \mathbf{t}) \in [0,1]^{k+2}$ and $\forall\ (t_0', \mathbf{t}') \in [0,1]^{k+2}$. Consequently, $\forall\ \delta \in (0, \frac{1}{2}]$, we have $\omega(W_n, \delta) \le 2k/\sqrt{m} + \omega_{T_n}(W_n, 2\delta)$. It follows that T_n accompanies Q_n. It remains to show that Q_n satisfies (3.5).

Let $\sum_{m=1}^{\infty} m^{-1} \varphi^{1/2(3+k)}(m) < \infty$ [see (1.3)], and let B_n be a block of T_n defined in Section 3.1.b. Using Lemma 4.1 with $q = k+3$, we obtain [see (3.2)]

$$\begin{aligned} E\left[e_{T_n}^{(B_n)}(W_n)\right]^{2(k+3)} &\le C_{k+3}(\varphi) \sum_{l=1}^{k+3} m^{-(k+3)} \left[(m+k)(t_0 - t_0')\right]^{(k+3)/l} \\ &\quad \times \left[\prod_{j=1}^{k+1} (t_j - t_j')\right]^{(k+3)/(k+1)l} \\ &\le C_{k+3}(\varphi) \sum_{l=1}^{k+3} m^{-(k+3)} \times (m+k)^{(k+3)/l} \\ &\quad \times \left[\prod_{j=0}^{k+1} (t_j - t_j')\right]^{(k+3)/(k+2)l} \\ &\le C_{k+3}(\varphi)(k+3) \left[\prod_{j=0}^{k+1} (t_j - t_j')\right]^{(k+3)/(k+2)}. \end{aligned}$$

Now let $\nu = (C_{k+3}(\varphi)(k+3))^{1/\beta}U$, where U is a uniform measure on $[0,1]^{k+2}$ and $\beta = (k+3)/(k+2)$. Then, by the Markov inequality, we obtain [see (3.5)]

$$Q_n\Big[f \in D_{k+2};\ \big|e_{T_n}^{(B_n)}(f)\big| \geq \lambda\Big] \leq \lambda^{-2(k+3)}(\tilde{\nu}_n(B_n))^{\beta},$$

which implies (3.6) for the φ-mixing rates (1.3) [and so also for (1.2)]. For the strong-mixing case with rates (1.4), we use Lemma 4.2 with $\varepsilon < (2(k+3))^{-1}$ and obtain from (4.2),

$$E\Big[e_{T_n}^{(B_n)}(W_n)\Big]^{2(k+3)}$$

$$\leq C_{k+3}(\alpha)\sum_{l=1}^{k+3} m^{-(k+3-l)}(t_0 - t_0')^l\left(\prod_{j=1}^{k+1}(t_j - t_j')\right)^{l(1-\varepsilon)/(k+1)},$$

which [with $\beta = ((k+2)(1-\varepsilon)+1)/k+2$] implies (3.5) and hence (3.6). We derive the convergence with respect to the special Skorohod topology because W_n is measurable with respect to this topology and we use Proposition 3.1 to the first coordinate (of W_n). □

PROOF OF PROPOSITION 3.5. The main line of proof is as follows: We consider a map $G_n\colon \mathscr{V}_n \to D_{k+2}$, where $\mathscr{V}_n$ is a subset of D_{k+2} and is such that $L_n = G_n \circ W_n$, $n \geq 1$. We show that $G_n\colon (\mathscr{V}_n, d) \to (D_{k+2}, \rho)$ is a continuous map, where d is the special Skorohod metric and ρ is the uniform metric.

Let $\mathscr{V}$ be a subset of D_{k+2} such that for any $v \in \mathscr{V}$, v equals zero on the lower boundary of $[0,1]^{k+2}$ and also for $\mathbf{t} = (1,\dots,1)$. It will be noted that $\mathscr{V}_n \subset \mathscr{V}$ for $\forall\ n \geq 1$.

Let $G\colon \mathscr{V} \to D_{k+2}$ be a map defined by

$$G(v)(t) = v(t) - t_0\sum_{j=1}^{k+1}\big[v(1,\dots,t_j,\dots,1) \times l_j(t_1,\dots,t_{k+1})\big], \tag{4.4}$$

where l_j is the limit of $\partial H_n/\partial t_j$ as $n \to \infty$. We will show that $\forall\ (v_n)_{n\in\mathbb{N}^*} \in (\Pi_{n\in\mathbb{N}^*}\mathscr{V}_n)$ and $\forall\ v \in \mathscr{V} \cap C_{k+2}$, $v_n \to_d v \Rightarrow G_n(v_n) \to_\rho G(v)$ as $n \to \infty$. Now, using Lemma 3 of Balacheff and Dupont (1980), we get the desired convergence.

Let $\mathscr{Y}_n = \{y \in [0,1]^n\colon (y^{(1)},\dots,y^{(n)})$ are distinct points of $(0,1)\}$. We define $Y_n\colon [0,1]^n \to D_{k+2}$ by setting

$$Y_n(y)(t) = (n-k)^{-1/2}\sum_{i=k+1}^{[nt_0]}\left[\prod_{j=1}^{k+1} I_{\{y^{(i+j-k-1)} \leq t_j\}} - H_n(\mathbf{t})\right],$$

for all $y = (y^{(1)},\dots,y^{(n)}) \in \mathscr{Y}_n$ and $t = (t_0,\mathbf{t}) \in [0,1]^{k+2}$.

We define the space $\mathscr{V}_n$ by $\mathscr{V}_n = Y_n(\mathscr{Y}_n)$. For any $j \in \{1,\dots,k+1\}$ we define an operator $\tau_j\colon \mathscr{V}_n \to D_1$ as follows.

Let $y_{(1)} < \cdots < y_{(n)}$ be the order values of $(y^{(1)}, \dots, y^{(n)})$ (by convention, $y_{(0)} = 0$, $y_{(n+1)} = 1$), and let $v_n = Y_n(y)$. Then,

$$\tau_j(v_n)(t_j) \tag{4.5}$$
$$= \begin{cases} y^{(l)}, & \text{where } y^{(l)} = \max\{y^{(m)}; m \in \{j, \dots, j+n-k-1\}\} \\ & \text{if } t_j = 1, \\ y^{(q)}, & \text{where } y^{(q)} = \max\{y^{(m)} \le y_{(i)}; \\ & \qquad m \in \{0, j, \dots, j+n-k-1\}\} \\ & \qquad \text{if } t_j \in \left[\dfrac{i}{n}, \dfrac{i+1}{n}\right), \end{cases}$$

where $i = \{0, 1, \dots, n-1\}$.

Now the map $G_n\colon \mathscr{V}_n \to D_{k+2}$ is given by

$$\begin{aligned} G_n(v_n)(t) &= v_n(t_0, \tau_1(v_n)(t_1), \dots, \tau_{k+1}(v_n)(t_{k+1})) \\ &\quad + (n-k)^{-1/2} \sum_{i=k+1}^{[nt_0]} \big[H_n(\tau_1(v_n)(t_1), \dots, \tau_{k+1}(v_n)(t_{k+1})) \\ &\qquad - H_n(t_1, \dots, t_{k+1})\big]. \end{aligned} \tag{4.6}$$

We now give the formal proof.

The first thing we have to show is that G_n is continuous for every n.

Let $\{v_{n,l}\} n \ge 1$, $l \ge 1$, be a sequence of functions in $\mathscr{V}_n$ and let $v_{n,l} \to v_n (\in \mathscr{V}_n)$ with respect to special Skorohod topology. We show that $G_n(v_{n,l}) \to G_n(v_n)$ in uniform topology. From the definition of the special Skorohod topology, we have a sequence $\{\lambda_{j,l}\}_{1 \le j \le k+1, l \ge 1} \in \Lambda^{k+1}$ such that $\forall\ \varepsilon > 0$, $\exists\ l_\varepsilon \in \mathbb{N}$ such that $\max_{1 \le j \le k+1} |\lambda_{j,l}(t_j) - t_j| \le \varepsilon$ and

$$|v_{n,l}(t) - v_n(t_0, \lambda_{1,l}(t_1), \dots, \lambda_{k+1,l}(t_{k+1}))| < \varepsilon \qquad \forall\ l \le l_\varepsilon \text{ and } \forall\ t \in [0,1]^{k+2}, \tag{4.7}$$

where Λ denotes the space of maps $h\colon [0,1] \to [0,1]$ which are nondecreasing, continuous and bijective, and Λ^{k+1} denotes the space of maps $\lambda\colon [0,1]^{k+1} \to [0,1]^{k+1}$, where $\lambda(t_1, \dots, t_{k+1}) = (\lambda_1(t_1), \dots, \lambda_{k+1}(t_{k+1}))$, $\lambda_j \in \Lambda$, $1 \le j \le k+1$. Then, we have the following lemma.

Lemma 4.4. *$\exists\ l_0 > 0$ such that $\forall\ l \ge l_0$, $\forall\ j \in \{1, \dots, k+1\}$ and $\forall\ t_j \in [0,1]$,*

$$\lambda_{j,l}\big(\tau_j(v_{n,l})(t_j)\big) = \tau_j(v_n)(t_j).$$

Proof. For fixed j, let $(y^{l,1}, \dots, y^{l,n-k})$ be a nondecreasing sequence of discontinuity points of $\tau_j(v_{n,l})$, and let $(y^{0,1}, \dots, y^{0,n-k})$ be a (nondecreasing) sequence of discontinuity points of $\tau_j(v_n)$. (By convention, $y^{l,0} = y^{0,0} = 0$, $y^{l,n-k+1} = y^{0,n-k+1} = 1$.) For $i \in \{0, 1, \dots, n-k+1\}$, let $t_j \in [y^{l,i}, y^{l,i+1})$. Then,

$$(4.8)\quad (n-k)^{-1/2}\big[v_{n,l}(1,\dots,t_j,\dots,1) + H_n(1,\dots,t_j,\dots,1)\big] = i(n-k)^{-1}.$$

Let $h \in \{0, 1, \dots, n-k+1\} \to \lambda_{j,l}(t_j) \in [y^{0,h}, y^{0,h+1})$. Then, we have

$$(4.9)\quad \begin{aligned} &(n-k)^{-1/2}\big[v_n\big(1,\dots,\lambda_{j,l}(t_j),\dots,1\big) + H_n\big(1,\dots,\lambda_{j,l}(t_j),\dots,1\big)\big] \\ &\quad = h(n-k)^{-1}. \end{aligned}$$

From (4.8) and (4.9), we deduce

$$(4.10)\quad \begin{aligned} \left|\frac{h}{n-k} - \frac{i}{n-k}\right| &\le |t_j - \lambda_{j,l}(t_j)| + (n-k)^{-1/2} \\ &\quad \times \big|v_{n,l}(1,\dots,t_j,\dots,1) - v_n\big(1,\dots,\lambda_{j,l}(t_j),\dots,1\big)\big| \\ &< \frac{1}{n-k} \qquad \forall\, l \ge \text{some } l_j. \end{aligned}$$

Thus, $|h/(n-k) - i/(n-k)| < 1/(n-k)$ and this implies that $h = i$. Now let $l_0 = \max_{1 \le j \le k} l_j$. Then, $\forall\ l \ge l_0$ and $\forall\ t_j \in [y^{l,i}, y^{l,i+1})$, we have $\lambda_{j,l}(t_j) \in [y^{0,i}, y^{0,i+1})$. Since the functions $\lambda_{j,l}$ are continuous and strictly nondecreasing, the proof follows. □

We now decompose G_n defined in (4.7) as $G_n = \gamma_n + \delta_n$, where $\gamma_n(v_n)(t) = v_n(t_0, \tau_1(v_n)(t_1), \dots, \tau_{k+1}(v_n)(t_{k+1}))$ and $\delta_n = G_n - \gamma_n$.

LEMMA 4.5. (a) $\gamma_n: (\mathscr{V}_n, d) \to (D_{k+2}, \rho)$ *is continuous.*
(b) $\delta_n: (\mathscr{V}_n, d) \to (D_{k+2}, \rho)$ *is continuous.*

PROOF. For $t \in [0,1]^{k+2}$, for $\forall\ \varepsilon > 0$, $\exists\ l_\varepsilon \ni \forall\ l \ge l_\varepsilon$, we have (using Lemma 4.4)

$$\begin{aligned} &|v_{n,l}(t_0, \tau_1(v_{n,l})(t_1), \dots, \tau_{k+1}(v_{n,l})(t_{k+1})) \\ &\quad - v_n(t_0, \tau_1(v_n)(t_1), \dots, \tau_{k+1}(v_n)(t_{k+1}))| < \varepsilon. \end{aligned}$$

The proof follows. Part (b) follows analogously, noting that H_n has uniform marginals. □

We now prove the convergence of the sequence $\{G_n\}$.

Let $v_n \in \mathscr{V}_n$, $n \in \mathbb{N}^*$, and suppose that $v_n \to_d v \in C_{k+2}$ and $v = 0$ on the lower boundary of $[0,1]^{k+2}$ and also when $\mathbf{t} = (1, \dots, 1)$. We have to prove that $G_n(v_n) \to_\rho G(v)$. The proof is based on the following lemmas.

LEMMA 4.6. $\forall\, j \in \{1, \dots, k+1\}$:

(a) $\tau_j(v_n) \to \mathrm{id}_{[0,1]}$ *in uniform topology.*
(b) $(n-k)^{1/2}(\tau_j(v_n) - \mathrm{id}_{[0,1]}) \to -v(1, \dots, \mathrm{id}_{[0,1]}, \dots, 1)$ *in uniform topology, where* $\mathrm{id}_{[0,1]}$ *is the identity function on* $[0,1]$.

PROOF. Note that $\forall\, v_n$, $\exists\, y_n = (y_n^{(1)}, \dots, y_n^{(n)})$ such that $v_n = Y_n(y_n)$. Now, for fixed j and for each $n \geq k+1$, define $v_n^{(j)}(t_j)$ as $v_n^{(j)}(t_j) = n^{-1/2}\sum_{i=1}^{n}\{I_{\{y_n^{(i)} \leq t_j\}} - t_j\}$ and note that $v_n^{(j)}(t_j)$ can also be written as

$$v_n^{(j)}(t_j) = ((n-k)/n)^{1/2} v_n(1, \dots, t_j, \dots, 1) + n^{-1/2} \sum_{i=1}^{k-j} \left[I_{\{y_n^{(i)} \leq t_j\}} - t_j \right] + n^{-1/2} \sum_{i=n-j}^{n} \left[I_{\{y_n^{(i)} \leq t_j\}} - t_j \right].$$

Since $v_n \to_d v$ (which implies $v_n \to_\rho v$), it follows that $v_n^{(j)}(t_j) \to_\rho v(1, \dots, t_j, \dots, 1)$. Thus, we can write

$$\begin{aligned} |\tau_j(v_n)(t_j) - t_j| &= \left| n^{-1} \sum_{i=1}^{n} \{ I_{\{y_n^{(i)} \leq \tau_j(v_n)t_j)\}} - t_j \} - n^{-1/2} v_n^{(j)}\big(\tau_j(v_n)(t_j)\big) \right| \\ &\leq \frac{k}{n} + n^{-1/2}\big[\rho\big(v_n^{(j)}, v(1, \dots, \cdot\,, \dots, 1)\big) \\ &\quad + \rho\big(v(1, \dots, \cdot\,, \dots, 1), g\big)\big] \to 0 \quad \text{as} \quad n \to \infty, \text{ where } g \equiv 0. \end{aligned}$$

This proves part (a). The proof of part (b) is similar. □

LEMMA 4.7. *$\gamma_n(v_n) \to v$ in uniform topology.*

PROOF. Follows by definition and Lemma 4.6(a). □

LEMMA 4.8. *$\delta_n(v_n) \to \delta(v) = G(v) - v$ in uniform topology.*

PROOF. For $t \in [0,1]^{k+2}$, we have

$$\delta_n(v_n)(t) = \frac{[nt_0] - k}{(n-k)^{1/2}} \times \{ H_n(\tau_1(v_n)(t_1), \dots, \tau_{k+1}(v_n)(t_{k+1})) - H_n(t_1, \dots, t_{k+1}) \}.$$

If there exists a $j \in \{1, \dots, k+1\}$, $t_j < n^{-1}$, then

$$\delta_n(v_n)(t) = \frac{k - [nt_0]}{(n-k)^{1/2}} H_n(t_1, \dots, t_{k+1}) \leq \frac{k - [nt_0]}{(n-k)^{1/2}} \frac{1}{n},$$

and so $\delta_n(v_n) \to 0$ as $n \to \infty$.

If $\forall\, j \in \{1, \dots, k+1\}$, $t_j \geq n^{-1}$, then by the Taylor expansion,

$$\delta_n(v_n)(t) = \frac{[nt_0] - k}{(n-k)^{1/2}} \sum_{j=1}^{k+1} \{ \tau_j(v_n)(t_j) - t_j \} \times \frac{\partial}{\partial t_j} H_n(\theta_{n,1}(t_1), \dots, \theta_{n,k+1}(t_{k+1})),$$

where $\theta_{n,j}(t_j) \in [t_j \wedge \tau_j(v_n)(t_j), t_j \vee \tau_j(v_n)(t_j)]$. Since $\{H_n\}$ satisfies the differentiability condition, we deduce from Lemma 4.6, the desired result.

Now, since $G_n(v_n) = \gamma_n(v_n) + \delta_n(v_n)$, we obtain (using Lemmas 4.7 and 4.8) that $G_n(v_n) \to G(v) = v + \delta(v)$. The proof of Proposition 3.5 follows. □

5. Proof of Theorem 2.1. First, we show $\mathscr{S}_n$ can be written as

$$\mathscr{S}_n = (n-k)^{-1/2}\left[\int_{[0,1]^{k+2}} L_n(t)\lambda_n(dt)\right] + b_n, \tag{5.1}$$

where λ_n is a signed measure on $[0,1]^{k+2}$ and b_n is the centering constant defined in Section 2.

$$\begin{aligned}
&(n-k)^{1/2}(\mathscr{S}_n - b_n)\\
&= (n-k)^{-1/2}\left[\sum_{i=k+1}^{n} c_{n,i}a_n(R_{n,i-k},\dots,R_{n,i}) - b_n\right]\\
&= (n-k)^{-1/2}\Bigg[\Bigg(\sum_{i=k+1}^{n} c_{n,i}\Bigg(\sum_A a_n(i_1,\dots,i_{k+1})\prod_{j=1}^{k+1} I_{[R_{n,i+j-k-1}=i_j]}\Bigg)\Bigg)\\
&\qquad\qquad - \sum_B \lambda_n\left(\frac{i_0}{n},\dots,\frac{i_{k+1}}{n}\right)\hat{H}_n\left(\frac{i_0}{n},\dots,\frac{i_{k+1}}{n}\right)\Bigg]\\
&= (n-k)^{-1/2}\Bigg[\sum_B \lambda_n\left(\frac{i_0}{n},\dots,\frac{i_{k+1}}{n}\right)\\
&\qquad\qquad \times\Bigg(\sum_{i=k+1}^{i_0}\Bigg(\prod_{j=1}^{k+1} I_{[R_{n,i+j-k-1}\le i_j]} - \hat{H}_n\left(\frac{i_0}{n},\dots,\frac{i_{k+1}}{n}\right)\Bigg)\Bigg)\Bigg]\\
&= \sum_B \lambda_n\left(\frac{i_0}{n},\dots,\frac{i_{k+1}}{n}\right)L_n\left(\frac{i_0}{n},\dots,\frac{i_{k+1}}{n}\right)\\
&= \int_{[0,1]^{k+2}} L_n(t)\lambda_n(dt),
\end{aligned}$$

where Σ_A is the sum over all $(i_1,\dots,i_{k+1})$ in $\{1,\dots,n\}^{k+1}$ and Σ_B is the sum over all $(i_0,\dots,i_{k+1})$ in $\{1,\dots,n\}^{k+2}$, where λ_n is defined in Section 2 and L_n is given by (2.3). We now prove that

$$\int_{[0,1]^{k+2}} L_n(t)\lambda_n(dt) \to \int_{[0,1]^{k+2}} L_\infty(t)\lambda_0(dt) \quad \text{as } n\to\infty. \tag{5.2}$$

Let $h_n\colon D_{k+2} \to \mathbb{R}$ be defined as

$$h_n(f) = \int_{[0,1]^{k+2}} f\lambda_n(dt), \qquad n \ge 0. \tag{5.3}$$

Let $\{f_n, n \geq 1\}$ be a sequence of functions in D_{k+2}, and suppose that $f_n \to f_0$ in uniform topology, where $f_0 \in C_{k+2}$. We show that

$$h_n(f_n) \to h_0(f_0). \tag{5.4}$$

We have

$$\begin{aligned} &\left| \int_{[0,1]^{k+2}} f_n \lambda_n(dt) - \int_{[0,1]^{k+2}} f_0 \lambda_0(dt) \right| \\ &\quad \leq \left| \int_{[0,1]^{k+2}} |f_n - f_0| \lambda_n(dt) \right| + \left| \int_{[0,1]^{k+2}} f_0(\lambda_n - \lambda_0)(dt) \right| \\ &\quad \leq \sup_{t \in [0,1]^{k+2}} |f_n(t) - f(t)| \left| \int_{[0,1]^{k+2}} \lambda_n(dt) \right| + \left| \int_{[0,1]^{k+2}} f_0(\lambda_n - \lambda_0)(dt) \right|. \end{aligned}$$

(5.4) follows using (2.5), (2.6) and (5.3), and (5.2) follows using Billingsley [(1968), Theorem 5.5)] and Proposition 3.5.

Now we prove that condition (2.6) of Theorem 2.1 is satisfied. By using (5.1) and (5.2) we deduce that

$$\sigma^2 = \int_{[0,1]^{k+2}} \int_{[0,1]^{k+2}} E[L_\infty(t) L_\infty(t')] \lambda_0(dt) \lambda_0(dt'). \tag{5.5}$$

We have [see (4.4)]

$$L_\infty(t) = W_\infty(t) - \sum_{j=1}^{k+1} t_0 W_\infty(1, \dots, t_j, \dots, 1) l_j(\mathbf{t}). \tag{5.6}$$

From (5.5) and (5.6) the equality in (2.6) holds.

It remains to show that $\sigma^2 < \infty$. By assumption (d) of Proposition 3.4,

$$\begin{aligned} \lim_{n \to \infty} &\left| E\left[\left\{ W_n(t) - \sum_{j=1}^{k+1} t_0 W_n(1, \dots, t_j, \dots, 1) l_j(\mathbf{t}) \right\} \right.\right. \\ &\qquad \left.\left. \times \left\{ W_n(t') - \sum_{j=1}^{k+1} t'_0 W_n(1, \dots, t'_j, \dots, 1) l_j(\mathbf{t}) \right\} \right] \right| \\ &= |E[L_\infty(t) L_\infty(t')]| \\ &\leq \lim_{n \to \infty} \left[E\left\{ W_n(t) - \sum_{j=1}^{k+1} t_0 W_n(1, \dots, t_j, \dots, 1) l_j(\mathbf{t}) \right\}^2 \right]^{1/2} \\ &\qquad \times \left[E\left\{ W_n(t') - \sum_{j=1}^{k+1} t'_0 W_n(1, \dots, t'_j, \dots, 1) l_j(\mathbf{t}') \right\}^2 \right]^{1/2} \\ &= \lim_{n \to \infty} [A_n B_n], \end{aligned} \tag{5.7}$$

by the Schwarz inequality.

Let now $\{X_{n,i}\}$ be φ-mixing with rates (1.2) or (1.3). Then from Lemma 4.1 with $q = 1$, we obtain

$$\begin{aligned} A_n &\le \Bigg[E\Bigg\{W_n^2(t) + 2|W_n(t)|t_0\sum_{j=1}^{k+1}|W_n(1,\dots,t_j,\dots,1)|l_j(\mathbf{t}) \\ &\qquad + t_0^2\sum_{s=1}^{k+1}\sum_{j=1}^{k+1} l_j(\mathbf{t})l_s(\mathbf{t})|W_n(1,\dots,t_j,\dots,1)W_n(1,\dots,t_s,\dots,1)|\Bigg\}\Bigg]^{1/2} \\ &\le C_1\Bigg[\Bigg(\prod_{m=1}^{k+1}t_m\Bigg)^{1/(k+2)} + 2t_0\Bigg(\prod_{m=1}^{k+1}t_m\Bigg)^{1/2(k+1)} l_j(\mathbf{t}) \\ &\qquad + t_0^2\sum_{s=1}^{k+1}\sum_{j=1}^{k+1} l_j(\mathbf{t})l_s(\mathbf{t})t_j^{1/2}t_s^{1/2}\Bigg]^{1/2}, \end{aligned} \tag{5.8}$$

where $C_1 > 0$ is some constant.

Similarly, B_n is less than or equal to some inequality with t's changed to t's. Thus $|E[L_0(t)L_0(t')]|$ is bounded by a function which is $\lambda_0 \times \lambda_0$ integrable, and so $|E[L_0(t)L_0(t')]|$ is also $\lambda_0 \times \lambda_0$ integrable.

Let now $\{X_{n,i}\}$ be strong-mixing with rates (1.4). Then, using Lemma 4.2 with $q = 2$, we obtain

$$\begin{aligned} A_n \le C_2\Bigg[&\Bigg(\prod_{m=0}^{k+1}t_m\Bigg)^{(1-\varepsilon)/(k+2)} + 2t_0\Bigg(\prod_{m=1}^{k+1}t_m\Bigg)^{(1-\varepsilon)/2(k+1)} l_j(\mathbf{t}) \\ &+ t_0^2\sum_{s=1}^{k+1}\sum_{j=1}^{k+1} l_j(\mathbf{t})l_s(\mathbf{t})t_j^{(1-\varepsilon)/2}t_s^{(1-\varepsilon)/2}\Bigg]^{1/2} \end{aligned}$$

and a similar inequality for B_n, and the result follows as in the case of φ-mixing.

The proof of Theorem 2.2 follows analogously.

6. Applications to Markov processes and ARMA processes.

6.1. *Markov processes.* Consider a sequence $\{X_{t,n};\ n \in \mathbb{Z}\}$ of $\mathbb{R}$-valued processes such that for all $t \in \mathbb{N}^*$, $\{X_{t,n}\}$ is a k-Markov process with stationary transition probabilities $P_t(x_1,\dots,x_k; A)$, where $A \in \mathscr{B}$, $\mathscr{B}$ is the Borel σ-field of $\mathbb{R}$ and $(x_1,\dots,x_k) \in \mathbb{R}^k$.

We say that the Markov process is *ergodic* if there exists a unique probability measure μ_t on $\mathbb{R}^k$ with marginals Π_t on $\mathbb{R}$ such that

$$\Pi_t(A) = \int_{\mathbb{R}^k} P_t(x_1,\dots,x_k; A)\mu_t(dx_1,\dots,dx_k) \quad \text{for all } A \in \mathscr{B}.$$

We denote by P_t^m the m-step transition probability defined by

$$P_t^{m+l}(x_1,\dots,x_k;A) = \int_{\mathbb{R}} P_t^l(x_2,\dots,x_k,y;A)\,P_t^m(x_1,\dots,x_k;dy),$$

for all $A \in \mathscr{B}$ and $(x_1,\dots,x_k) \in \mathbb{R}^k$.

For a transition probability $P_t(\cdots,\cdot)$ and invariant measure μ_t and marginal Π_t, we denote by $P_t^*(\cdot,\cdot)$ the transition probability defined by

$$\int_{\prod_{i=1}^k(-\infty,y_i]} P_t(u_1,\dots,u_k;(-\infty,y_{k+1}])\mu_t(du_1,\dots,du_k)$$
$$= \int_{-\infty}^{y_{k+1}} P_t^*\left(u_{k+1};\prod_{i=1}^k(-\infty,y_i]\right)\Pi_t(du_{k+1}).$$

We say that the Markov process is *geometrically ergodic* if it is ergodic and if there exists $0<\rho_t<1$ such that

$$\|P_t^m(x_1,\dots,x_k;\cdot) - \Pi_t(\cdot)\| = O(\rho_t^m) \quad \text{for all a.s. } (x_1,\dots,x_k)\in\mathbb{R}^k,$$

where $\|\cdot\|$ denotes the norm of total variation and ρ_t is called the rate.

The Markov process is *Harris recurrent* if there exists a σ-finite measure ν_t on $\mathbb{R}$ with $\nu_t(\mathbb{R})>0$ such that $\nu_t(A)>0$ implies $P_t(x_1,\dots,x_k;X_{t,n}\in A \text{ i.o.}) = 1$ for all $(x_1,\dots,x_k)\in\mathbb{R}^k$.

Finally, the Markov process is *Doeblin recurrent* if it is ergodic and there exists a finite measure ν_t on $\mathbb{R}$ with $\nu_t(\mathbb{R})>0$, an $m\ge 1$ and $\varepsilon>0$ such that $P_t^m(x_1,\dots,x_k;A)\le 1-\varepsilon$ if $\nu_t(A)\le\varepsilon$ for all $(x_1,\dots,x_k)\in\mathbb{R}^k$ and $A\in\mathscr{B}$.

Let us denote $\forall\, j\in\{1,\dots,k+1\}$ and $\forall\, M>0$,

$$R_j(M) = (-\infty,+\infty)^{j-1}\times[-M,M]\times(-\infty,+\infty)^{k-j+1}.$$

Then we have the following theorem.

THEOREM 6.1. *Let $\{X_{t,n},\ n\in\mathbb{Z}\}$ be a Markov process such that for every $t\in N^*$, $\{X_{t,n}\}$ is either* (a) *aperiodic, Harris recurrent and geometrically ergodic with rates $0<\rho_t<\rho_0$, $\rho_0\in(0,1)$ or* (b) *aperiodic and Doeblin recurrent. Suppose there exists a probability μ_0 on $\mathbb{R}^k$ and a transition probability $P_0(\cdots;\cdot)$ such that*

$$\sup_{A\in\mathscr{B}^k}|\mu_t(A)-\mu_0(A)| = O(t^{-\alpha}),\qquad \alpha>0, \tag{6.1}$$

$$\sup|P_t(x_1,\dots,x_k;A) - P_0(x_1,\dots,x_k;A)| \to 0 \quad \text{as } t\to\infty, \tag{6.2}$$

where sup *is over $A\in\mathscr{B}$ and $(x_1,\dots,x_k)\in R_j(M)$, for every $j\in\{1,\dots,k+1\}$, $\forall\, M>0$, and*

$$\sup|P_t^*(x_{k+1};A_k) - P_0^*(x_{k+1};A_k)| \to 0 \quad \text{as } t\to\infty, \tag{6.3}$$

where sup *is over $|x_{k+1}|\le M$ and $A_k\in\mathscr{B}^k$.*

Then, under the assumptions (3.2) *and* (3.3), $(n-k)^{1/2}(\mathscr{S}_n - b_n)$ *converges in law to a normal distribution with mean* 0 *and variance σ^2, where b_n*

and σ^2 *are given by* (2.4) *and* (2.7), *respectively*. (It is assumed that P_t, P_0, P_t^*, P_0^* have densities *continuous* in x's and μ_0 have densities).

PROOF. (i) Suppose (a) holds. First, we show that the process is geometrically strong mixing. It is well known [see Nummelin and Tuominen (1982)] that if a Markov chain is aperiodic, Harris recurrent and geometrically ergodic with rate ρ_t, then

$$\int \|P_t^m(x_1,\dots,x_k;\cdot) - \Pi_t(\cdot)\| \mu_t(dx_1,\dots,dx_k) = O(\rho_t^m)$$

and this property is equivalent to strong mixing with rate ρ_t^m [see Rosenblatt (1971), page 199]. Next, we show that the covariance functions of the associated graduate empirical process (2.2), converge to a function K, but this is a consequence of Lemma 6 of Rüschendorf (1974) which remains true for strong-mixing conditions with a geometric rate.

Let G_t be the distribution function of the $k+1$ successive random variables of $\{X_{tn}\}$ and let H_t be the measure on $[0,1]^{k+1}$ defined by $H_t(y_1,\dots,y_{k+1}) = G_t(\Pi_t^{-1}(y_1),\dots,\Pi_t^{-1}(y_{k+1}))$, where Π_t is the marginal of μ_t for all $(y_1,\dots,y_{k+1}) \in [0,1]^{k+1}$ and $t \geq 0$ (note that we also denote by Π_t the distribution function associated with the measure Π_t).

We have to show that $\{H_t\}_{t>0}$ satisfies the differentiability condition (given in Section 2).

Set $l_t^{(j)} = \partial H_t/\partial t_j$, and let $F_t^{(j)}$ be the conditional distribution function defined as

$$\begin{aligned} &F_t^{(j)}(u_j; y_1,\dots,y_{j-1},y_{j+1},\dots,y_{k+1}) \\ &\qquad = \int_R P_t(u_1,\dots,u_k;(-\infty,y_{k+1}])\mu_t^j(du_1,\dots,du_{j-1},du_{j+1},du_k), \end{aligned}$$

where $R = \Pi_{l=1}^{j-1}(-\infty,y_l]\Pi_{l=j+1}^{k}(-\infty,y_l]$, μ_t^j is the measure associated with the distribution function $H_t(u_1,\dots,u_{j-1},1,u_{j+1},\dots,u_k,1)$ if $j \leq k$, and $F_t^{(k+1)}$ is the conditional distribution function associated with P_t^*. We have

$$\begin{aligned} &l_t^{(j)}(y_1,\dots,y_{k+1}) \\ &\qquad = F_t^{(j)}\big(\Pi_t^{-1}(y_j);\Pi_t^{-1}(y_1),\dots,\Pi_t^{-1}(y_{j-1}),\Pi_t^{-1}(y_{j+1}),\dots,\Pi_t^{-1}(y_{k+1})\big). \end{aligned}$$

Also,

$$\begin{aligned} &l_t^{(j)}(y_1,\dots,y_k) - l_0^{(j)}(y_1,\dots,y_k) \\ &\quad = \big[F_t^{(j)}(\Pi_t^{-1}(y_j);\Pi_t^{-1}(y_1),\dots,\Pi_t^{-1}(y_{j-1}),\Pi_t^{-1}(y_{j+1}),\dots,\Pi_t^{-1}(y_{k+1})) \\ &\qquad -F_0^{(j)}(\Pi_t^{-1}(y_j);\Pi_t^{-1}(y_1),\dots,\Pi_t^{-1}(y_{j-1}),\Pi_t^{-1}(y_{j+1}),\dots,\Pi_t^{-1}(y_{k+1}))\big] \\ &\qquad +\big[F_0^{(j)}(\Pi_t^{-1}(y_j);\Pi_t^{-1}(y_1),\dots,\Pi_t^{-1}(y_{j-1}),\Pi_t^{-1}(y_{j+1}),\dots,\Pi_t^{-1}(y_{k+1})) \\ &\qquad\quad -F_0^{(j)}(\Pi_0^{-1}(y_j);\Pi_0^{-1}(y_1),\dots,\Pi_0^{-1}(y_{j-1}),\Pi_0^{-1}(y_{j+1}),\dots,\Pi_0^{-1}(y_{k+1}))\big] \\ &\quad = A + B. \end{aligned}$$

To simplify the notation, take $j = 1$ (for $j \neq 1$, the method is exactly the same). Then, we have

$$\begin{aligned}
A &= F_t^{(1)}\big(\Pi_t^{-1}(y_1); \Pi_t^{-1}(y_2), \dots, \Pi_t^{-1}(y_{k+1})\big) \\
&\quad - F_0^{(1)}\big(\Pi_t^{-1}(y_1); \Pi_t^{-1}(y_2), \dots, \Pi_t^{-1}(y_{k+1})\big) \\
&= \int_{\Pi_{l=2}^k(-\infty, x_l]} P_t(x_1; u_2, \dots, u_k; (-\infty, x_{k+1}])\mu_t^1(du_2, \dots, du_k) \\
&\quad - \int_{\Pi_{l=2}^k(-\infty, x_l]} P_0(x_1, u_2, \dots, u_k; (-\infty, x_{k-1}])\mu_0^1(du_2, \dots, du_k) \\
&\qquad\qquad \big[\text{for } \Pi_t^{-1}(y_l) = x_l,\ l \in \{1, \dots, k+1\}\big] \\
&\le \varepsilon + \left| \int_{\Pi_{l=2}^k(-\infty, x_l]} \mu_t^1(du_2, \dots, u_k) - \mu_0^1(du_2, \dots, du_k) \right| \le 2\varepsilon
\end{aligned}$$

for all $(x_1, \dots, x_{k+1}) \in R_1(M)$.

We also have

$$\begin{aligned}
B &= F_0^{(1)}\big(\Pi_t^{-1}(y_1); \Pi_t^{-1}(y_2), \dots, \Pi_t^{-1}(y_{k+1})\big) \\
&\quad - F_0^{(1)}\big(\Pi_0^{-1}(y_1); \Pi_0^{-1}(y_2), \dots, \Pi_0^{-1}(y_{k+1})\big) \\
&= F_0^{(1)}\big(\Pi_0^{-1} \circ \Pi_0 \circ \Pi_t^{-1}(y_1); \Pi_0^{-1} \circ \Pi_0 \circ \Pi_t^{-1}(y_2), \dots, \Pi_0^{-1} \circ \Pi_0 \circ \Pi_t^{-1}(y_{k+1})\big) \\
&\quad - F_0^{(1)}\big(\Pi_0^{-1} \circ \Pi_0 \circ \Pi_0^{-1}(y_1); \Pi_0^{-1} \circ \Pi_0 \circ \Pi_0^{-1}(y_2), \dots, \Pi_0^{-1} \circ \Pi_0 \circ \Pi_0^{-1}(y_{k+1}).
\end{aligned}$$

Noting that

$$\begin{aligned}
\sup_{y_1 \in [0,1]} |\Pi_0 \circ \Pi_t^{-1}(y_1) - y_1| &= \sup_{y_1 \in [0,1]} |y_1 - \Pi_t \circ \Pi_0^{-1}(y_1)| \\
&= \sup_{y_1 \in [0,1]} |\Pi_0 \circ \Pi_0^{-1}(y_1) - \Pi_t \circ \Pi_0^{-1}(y_1)|,
\end{aligned}$$

we find [using (6.1)] that $B < \varepsilon$ for sufficiently large t. Thus, $l_t^{(j)}(y_1, \dots, y_{k+1}) \to l_0^{(j)}(y_1, \dots, y_{k+1})$ as $t \to \infty$ uniformly in $(y_1, \dots, y_k) \in R_j(M)$ for any $M > 0$, and so $\{H_t\}_{t>0}$ satisfies the differentiability condition.

(ii) Suppose (b) is satisfied. Then the proof follows from Davydov (1973), who proved that a Markov process which is Doeblin recurrent and aperiodic is geometrically φ-mixing. □

EXAMPLE 6.1. Consider the process $\{X_n,\ n \in \mathbb{Z}\}$, where $X_{n+1} = a_1 X_n + a_2 X_n \varepsilon_{n+1} + a_3 \varepsilon_{n+1} + a_4 \varepsilon_{n+1}^2 + a_5$, where the a's are real numbers and $\{\varepsilon_n,\ n \in \mathbb{Z}\}$ is a white noise with strictly positive density. Then Mokkadem (1985) has shown that if $a_1^2 + a_2^2 E(\varepsilon_1^2) < 1$ and $E(\varepsilon_1^4) < \infty$, then the process $\{X_n,\ n \in \mathbb{Z}\}$ is geometrically ergodic and geometrically strong mixing. Thus, the asymptotic normality of the statistic $\mathscr{S}_n$ based on the ranks of $\{X_n\}$ follows.

EXAMPLE 6.2. Consider the process $\{X_n,\ n \in \mathbb{Z}\}$, where $X_{n+1} = f(X_n) + \varepsilon_{n+1}$, where the ε's are independent and identically distributed random variables with strictly positive density, and $f\colon \mathbb{R} \to \mathbb{R}$ is bounded, nondecreasing and continuous. [This model was studied by Collomb and Doukhan (1983).] It is easy to check that this model is Doeblin recurrent and aperiodic, and we deduce that $\{X_n\}$ is geometrically φ-mixing and we can apply Theorem 6.1.

6.2. *ARMA processes.* Consider a sequence of ARMA (k_1, k_2) processes,

$$\prod_{j=1}^{k_1}\left(1 - a_j^{(n)}U\right)X_{n,i} = Q_{k_2}^{(n)}(U)\varepsilon_i, \qquad i \in \mathbb{Z},\ n \in \mathbb{N}^*, \tag{6.4}$$

where $U\lambda_i = \lambda_{i-1}$, $Q_{k_2}^{(n)}(U) = \sum_{l=0}^{k_2} b_l^{(n)}U^l$, $b_0^{(n)} = 1$ and $\{\varepsilon_i,\ i \in \mathbb{Z}\}$ is a sequence of independent random variables such that $E(\varepsilon_i) = 0$ and ε_i has a density $g_i(x)$, $i \in \mathbb{Z}$.

Then we have the following lemma.

LEMMA 6.1 [Gorodetskii (1977), Withers (1981)]. *Let the sequence $\{X_{n,i}$, $i \in \mathbb{Z}\}$ satisfy the following conditions*:

$$\sup_{i\in\mathbb{Z}} \int_{-\infty}^{\infty} |g_i(x + \beta) - g_i(x)|\, dx \le c_1|\beta|, \quad \forall\, \beta \text{ and some } c_1 > 0; \tag{6.5}$$

$$\sup_{i\in\mathbb{Z}} E|\varepsilon_i| < c_2 < \infty \quad \text{and} \quad \sup_{n\in\mathbb{N}}\ \sup_{1\le j\le k_1} |a_j^{(n)}| < \rho < 1, \tag{6.6}$$

where c_2 and ρ are some constants. Then for any $n \in \mathbb{N}^*$, the process $\{X_{n,i}$; $i \in \mathbb{Z}\}$ is strong-mixing with rate $\alpha(m) = O(\rho_0^{m/2})$ for each $\rho_0 > \rho$.

THEOREM 6.2. *Let the sequence $\{X_{n,i},\ i \in \mathbb{Z}\}$ of ARMA (k_1, k_2) process given by* (6.4) *satisfy the following conditions*:

(6.7) *$\{\varepsilon_i,\ i \in \mathbb{Z}\}$ is a sequence of independent and identically distributed random variables, each having $\mathcal{N}(0, \sigma^{*2})$ distribution.*

(6.8) *$\forall\ j \in \{1,\dots,k_1\}$, $\exists\ \alpha > 0$ and $a_j \in (-1,1)$, $a_j \ne 0$, such that $|a_j^{(n)} - a_j| = O(n^{-\alpha})$, and $\forall\ l \in \{1,\dots,k_2\}$, $\exists\ \beta > 0$, and $b_l \in \mathbb{R}$ such that $|b_l^{(n)} - b_l| = O(n^{-\beta})$.*

Then for the rank statistic $\mathscr{S}_n$ associated with the sequence $\{X_{n,1},\dots,X_{n,n}\}$ and the score functions satisfying the assumptions of Theorem 2.1, *$(n - k)^{1/2}(\mathscr{S}_n - b_n)$ converges in law to the $\mathcal{N}(0, \sigma^2)$ distribution, where b_n and σ^2 are given by* (2.4) *and* (2.7), *respectively.*

PROOF. To prove this theorem, we first note, using Lemma 6.1, that the sequence $\{X_{n,i}\}$ is geometrically strong-mixing. Now, let F_n be the distribution function of $X_{n,i}$ and F_0 the distribution function of a stationary random variable X_{oi} defined by an ARMA (k_1, k_2) process with coefficients a_j, $1 \le j \le k_1$ and b_l, $1 \le l \le k_2$. Now we prove the differentiability condition for $H_n(\mathbf{t})$ defined in (2.1) be verifying (6.1), (6.2) and (6.3).

Let P_n^j be the transition distribution function of $X_{n,1},\dots,X_{n,j-1}$, $X_{n,j+1},\dots,X_{n,k_1}$, and G_n^j the distribution function of $(X_{n,1},\dots,X_{n,k_1-1})$,

$n \geq 0$. Then (6.1), (6.2) and (6.3) are satisfied in view of the following well-known result.

LEMMA 6.2. *Let $\{G_n,\ n \geq 0\}$ be a sequence of k-dimensional normal distribution functions each with mean vector* $\mathbf{0}$. *Let the covariance matrices of G_n and G_0 be $\Sigma_n = ((\sigma_{il}^{(n)}))$ and $\Sigma_0 = ((\sigma_{il}^{*}))$ and assume that $|\sigma_{il}^{(n)} - \sigma_{il}^{*}| = O(n^{-\alpha})$ for each $i, l = 1, \dots, k$. Then G_n converges uniformly to G_0.*

7. Appendix.

7A. *Proof of Proposition* 3.1. (a) and (b) are sufficient conditions. They follow immediately from Proposition 3.2 by using a result from Billingsley [(1968), page 151, line 15].

We have only to prove that (a) and (b) are necessary conditions.

Let $\mathscr{U}_{k+2}^{*}$ be the σ-field generated by the uniform topology on C_{k+2}.

As P is concentrated on a separable space $(C_{k+2}, \mathscr{U}_{k+2}^{*})$, it follows from Wichura [(1970), Theorem 1] that there exists a probability space $(\Omega, \mathscr{A}, \mu)$ and a sequence of random variables $\{Y_n^*\}$, $n \in \mathbb{N}^*$, and a random variable Y^* such that $\mu(Y_n^*) = P_n$, $\mu(Y^*) = P$ and $Y_n^* \to Y^*$ a.s. μ. For any δ (> 0), we consider the map T_δ: $D_{k+2} \to \mathbb{R}$ defined by

$$T_\delta(f) = \sup\{|f(t) - f(t'); \|t - t'\| \leq \delta\}.$$

Then, T_δ is a continuous map for the uniform topology on $\mathscr{U}_{k+2}$.

Now, consider a sequence of random variables $\{Z_{n,\delta}\}$, $n \in \mathbb{N}^*$, and a random variable Z_δ defined as

$$Z_{n,\delta} = T_\delta \circ Y_n^*, \qquad Z_\delta = T_\delta \circ Y^*.$$

As Y_n^* converges a.s. to Y^*, it follows that $\forall\ \varepsilon > 0$, $\exists\ N_0 \in \mathbb{N}$ such that

$$(7.1) \qquad \mu\{|Z_{n,\delta} - Z| > \varepsilon/2\} < \varepsilon/2, \qquad \forall\ n \geq N_0.$$

As Y^* is concentrated on C_{k+2}, we have also $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ such that

$$(7.2) \qquad \mu\{|Z_\delta| > \varepsilon/2\} < \varepsilon/2.$$

(7.1) and (7.2) imply

$$\mu\{|Z_{n,\delta}| > \varepsilon\} < \varepsilon$$

or

$$(7.3) \qquad P_n[f; \omega(f,\delta) \geq \varepsilon] = \mu\{|Z_{n,\delta}| > \varepsilon\} < \varepsilon,$$

and from (7.3) we obtain condition (b). Condition (a) is immediate. Proposition 3.1. is proved.

7B. *Proof of Lemma* 4.1. Without loss of generality, we can take $N_n = n$. First, we prove that for any p, $1 \leq p \leq n$,

$$(7.4) \qquad E\left(\sum_{i=1}^{p} X_{n,i}\right)^{2q} \leq C_q \sum_{l=1}^{q} p^{q/l}\left(\sup_{1\leq j\leq n} \|X_{n,j}\|_l\right)^{2q}(h(p,l))^{2q},$$

where C_q is a constant depending only on q and φ and

$$h(p,l) = \exp\left\{\sum_{l=1}^{s} \varphi^{1/2q}([2^{1/2(l+1)}])\right\} \quad \text{for } 2^s \le p < 2^{s+1}.$$

For any (l,p), $1 \le l \le p \le n$, we define $S(l,p)$ by

$$S(l,p) = \sum_{i=l\wedge(n+1)}^{(l+p-1)\wedge(n+1)} X_{n,i}, \quad \text{where, by convention, } X_{n,n+1} \equiv 0.$$

Denote $S = S(l,p)$, $S' = S(l+p+r,p)$, $R = S(l+p,r) - S(2p+l,r)$ for $r \ge 1$,

$$a(p,q) = \sup_{l\ge 1}\left[E\left(\sum_{i=l\wedge(n+1)}^{(l+p-1)\wedge(n+1)} X_{n,i}\right)^{2q}\right]^{1/2q}, \quad \text{and} \quad m_l = \sup_{1\le j\le n} \|X_{n,j}\|_l.$$

Then, after some computations, we obtain the inequality,

$$E(S+S')^{2q} \le 2(a(p,q))^{2q} \exp\{(2q\varphi(r))^{1/2q}\} + (2a(p,q-1))^{2q}. \tag{7.5}$$

From the Minkowski inequality, it follows that

$$\left\|\sum_{l=1}^{(l+2p-1)\wedge(n+1)} X_{n,i}\right\|_{2q} = \|S+S'+R\|_{2q} \le \|S+S'\|_{2q} + 2rm_q$$
$$\le 2^{1/2q}a(p,q)\exp\{\varphi(r)^{1/2q}\} + 2a(p,q-1) + 2rm_q.$$

Now take $p = 2^s$, $s \ge 1$, and put $r = r(s,q) = [2^{s/2(q+1)}]$, $\varphi(s,q) = (\varphi(r))^{1/2q}$.
Then, from (7.5), we can write

$$a(2^s,q) \le 2^{1/2q}a(2^{s-1},q)\exp\{\psi(s,q)\} + 2a(2^{s-1},q-1) + 2r(s,q)m_q,$$

$$\begin{aligned} a(2^s,q) &\le 2^{s/2q}\left(1 + 2\sum_{i=1}^{s} 2^{-i/2q}r(i,q)\right)m_q \exp\left\{\sum_{j=1}^{s}\psi(j,q)\right\} \\ &\quad + 2\sum_{i=1}^{s} 2^{(s-i)/2q}a(2^{i-1},q-1)\exp\left\{\sum_{j=i+1}^{s}\psi(j,q)\right\}, \end{aligned} \tag{7.6}$$

where, by convention, $\sum_{j=s+1}^{s}\psi(j,q) = 0$.
For $q = 1$, we have

$$a(2^s,1) \le K_1 2^{s/2}h(2^s,1)m_1,$$

where K_1 is a positive constant.
We give a proof by recurrence on q. Suppose that, for all $q \ge 2$ and $p = 2^s$, we have

$$a(2^s,q-1) \le K_{q-1}\sum_{l=1}^{q-1} 2^{s/2l}h(2^s,l)m_l,$$

where K_{q-1} is a positive constant. From (7.6), we deduce

$$a(2^s,q) \le 2^{s/2q}\left(\exp\left\{\sum_{j=1}^{s}\psi(j,q)\right\}\right)\left(1+2\sum_{j=1}^{s}2^{-1/2(j/q-j/(q+1))}\right)m_q$$
$$+\,2A_qK_{q-1}\sum_{l=1}^{q-1}2^{s/2l}h(2^s,l)m_l,$$

where A_q is a constant depending only on q and φ. That is,

$$a(2^s,q) \le K_q\sum_{l=1}^{q}2^{s/2l}h(2^s,l)m_l. \tag{7.7}$$

Finally, for each $p \le n$, we can write the binary decomposition as

$$p=\sum_{i=0}^{s}v_i2^i, \qquad v_i\in\{0,1\}.$$

From the equality $h(p,l)=h(2^s,l)$ for $2^s\le p<2^{s+1}$ and (7.7), it follows that

$$a(p,q)\le\sum_{i=0}^{s}v_ia(2^i,q)\le\sum_{i=0}^{s}a(2^i,q),$$

$$a(p,q)\le K_q\sum_{i=0}^{s}\sum_{l=1}^{q}2^{i/2l}h(2^i,l)m_l, \tag{7.8}$$

$$(a(p,q))^{2q}\le C_q\sum_{l=1}^{q}p^{q/l}(h(p,l)m_l)^{2q},$$

and (7.4) is proved.

(4.1) now follows by putting $p=n=N_n$ and by using the relation

$$\sum_{l=0}^{\infty}(\varphi(2^l))^{1/2q}<+\infty \quad\Leftrightarrow\quad \sum_{m=1}^{\infty}m^{-1}(\varphi(m))^{1/2q}<+\infty. \tag{7.9}$$

Lemma 4.1 is proved.

7C. *Proof of Lemma* 4.3. For every $p, N, r\in\mathbb{N}$, we define $S_N=\sum_{i=1}^{N}Y_i$, $T_{N,j}=\sum_{i=1}^{N}Y_{j(N+r)+i}$, $R_{N,j}=\sum_{i=1}^{r}(Y_{j(N+r)+N+i}-Y_{pN+jr+i})$ for $j=0,\dots,p-1$. For every $l\in\mathbb{N}$, we denote $K_l^2=E(\sum_{i=1}^{l}Y_i)^2$.

From the property of stationarity, we have

$$K_N^2=E(S_N^2)=E(T_{N,j})^2 \quad \text{for } j=0,\dots,p-1$$

and

$$\left|E(S_N)^2-\frac{1}{p}E\left(\sum_{j=0}^{p-1}T_{N,j}\right)^2\right|\le p(\varphi(r))^{1/2}K_N^2. \tag{7.10}$$

We have

$$S_{pN} = \sum_{i=1}^{pN} Y_i = \sum_{j=0}^{p-1} T_{N,j} + \sum_{j=0}^{p-1} R_{N,j}.$$

We deduce from (7.10)

$$(7.11) \qquad \left| E(S_N)^2 - \frac{1}{p} E(S_{pN})^2 \right| \le p(\varphi(r))^{1/2} K_N^2 + 4pK_r K_N.$$

From Lemma 4.1, there exists a constant C depending only on φ such that

$$K_r^2 \le CrK_1^2 \quad \text{and} \quad K_N^2 \le CNK_1^2.$$

Now, taking $r = r(N) = [N^{1/2}]$ and using (7.10) for $p = 2$, we obtain

$$\left| \frac{1}{2^s} E(S_{2^s})^2 - \frac{1}{2^{s+m}} E(S_{2^{s+m}})^2 \right| \le \frac{1}{2^s} \sum_{k=0}^{m-1} \frac{1}{2^k} \left| E(S_{2^{s+k}})^2 - \frac{1}{2} E(S_{2^{s+k+1}})^2 \right|$$

$$\le 8CK_1^2 \sum_{k=s}^{s+m-1} \left((\varphi([2^{k/2}]) + 2^{-k/4} \right).$$

It follows that $(1/2^s)E(S_{2^s})^2$ is a Cauchy sequence. Hence there exists a constant K such that

$$\frac{1}{2^s} E(S_{2^s})^2 \to K^2 \quad \text{as } s \to \infty.$$

We deduce that for every $p_0 \in \mathbb{N}$,

$$\sup_{p \ge p_0} \left| K^2 - \frac{1}{p2^s} E(S_{p2^s})^2 \right| \to 0 \quad \text{as } s \to \infty.$$

Let l be such that $pq \le l \le (p+1)q$. Then

$$\left| \frac{1}{pq} E(S_{pq})^2 - \frac{1}{l} E(S_l)^2 \right| \le 2CK_1^2 \left(\frac{1}{p} + \frac{1}{p^{1/2}} \right) \quad \text{for all } q \ge 1.$$

Consequently, $\forall\, \varepsilon > 0$, $\exists\, p_0 \in \mathbb{N}$, such that

$$\sup_{l \ge p} \left| \frac{1}{p[1/p]} E(S_{p[1/p]})^2 - \frac{1}{l} E(S_l)^2 \right| < \frac{\varepsilon}{2}, \qquad \forall\, p \ge p_0.$$

Now if we choose s_0 such that $2^{s_0} \ge p_0$ and

$$\sup_{p \le 2p_0} \left| K^2 - \frac{1}{p2^s} E(S_{p2^s})^2 \right| < \frac{\varepsilon}{2} \quad \text{for all } s \ge s_0,$$

then there exists for every $n \ge n_0 = p_0 2^{s_0}$, an $s \ge s_0$ and a p for which $p_0 \le p < 2p_0$ and $p2^s \le n < (p+1)2^s$.

We deduce that $\forall\ n \geq n_0$,

$$\left|K^2 - \frac{1}{n}E(S_n)^2\right| \leq \left|K^2 - \frac{1}{p2^s}E(S_{p2^s})^2\right| + \left|\frac{1}{p2^s}E(S_{p2^s})^2 - \frac{1}{n}E(S_n)^2\right|$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which implies

$$\frac{1}{n}E(S_n)^2 \to K^2 \quad \text{as } n \to \infty$$

and Lemma 4.3 is proved.

Acknowledgments. The authors would like to express their sincere gratitude to Professor Peter Ney, the Associate Editor, and the referees for their critical examination of the first draft. Their constructive criticisms and suggestions for improvements are gratefully acknowledged.

REFERENCES

BALACHEFF, S. and DUPONT, G. (1980). Normalité asymptotique des processus empiriques tronqués et des processus de rang. *Statistique non Paramétrique Asymptotique. Lecture Notes in Math.* **821** 19–45. Springer, Berlin.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

COLLOMB, G. and DOUKHAN, P. (1983). Estimation de la fonction d'autoregression d'un processus stationnaire et φ mélangeant: Risque quadratique pour la méthode du noyau. *C. R. Acad. Sci. Paris Sér. I Math.* **296** 859–862.

DAVYDOV, YU. A. (1973). Mixing conditions for Markov chains. *Theory Probab. Appl.* **18** 312–328.

DOUKHAN, P. and PORTAL, F. (1987). Principe d'invariance faible pour la fonction de répartition empirique dans un cadre multidimensionnel et mélangeant. *Probab. Math. Statist.* **8** 117–132.

DUDLEY, R. M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* **6** 899–929.

GAENSSLER, P. and STUTE, W. (1987). *Seminar on Empirical Processes*. Birkäuser, Boston.

GORODETSKII, V. V. (1977). On the strong mixing property for linear sequences. *Theory Probab. Appl.* **22** 411–413.

HALLIN, M., INGENBLEEK, J.-F. and PURI, M. L. (1985). Linear serial rank tests for randomness against ARMA alternatives. *Ann. Statist.* **13** 1156–1181.

HAREL, M. (1980). Convergence en loi pour la topologie de Skorohod du processus empirique multidimensionnel normalisé tronqué et semi-corrigé. *Statistique non Paramétrique Asymptotique. Lecture Notes in Math.* **821** 46–85. Springer, Berlin.

HAREL, M. and PURI, M. L. (1989a). Limiting behavior of U-statistics, V-statistics and one-sample rank order statistics for non-stationary absolutely regular processes. *J. Multivariate Anal.* **31** 181–204.

HAREL, M. and PURI, M. L. (1989b). Weak convergence of the U-statistic and weak invariance of the one-sample rank order statistic for Markov processes and ARMA models. *J. Multivariate Anal.* **31** 259–265.

HAREL, M. and PURI, M. L. (1990). Weak invariance of generalized U-statistics for nonstationary absolutely regular processes. *Stochastic Process. Appl.* **34** 341–360.

IBRAGIMOV, J. A. and LINNIK, YU. V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Nordhoff, Groningen.

MOKKADEM, A. (1985). Le modèle non linéaire AR(1) général. Ergodicité et ergodicité géométrique. *C. R. Acad. Sci. Paris Sér. I Math.* **301** 889–892.

NEUHAUS, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** 1285–1295.

NEUMANN, N. (1982). Ein schwaches Invarianzprinzip für den gewichten empirischen Prozess von gleichmässig mischenden Zufallsvariablen. Ph.D. dissertation, Göttingen.

NUMMELIN, E. and TUOMINEN, P. (1982). Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. *Stochastic Process. Appl.* **12** 187–202.

ROSENBLATT, M. (1971). *Markov Processes: Structure and Asymptotic Behavior*. Springer, Berlin.

RÜSCHENDORF, L. (1974). On the empirical process of multivariate, dependent random variables. *J. Multivariate Anal.* **4** 469–478.

RÜSCHENDORF, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Ann. Statist.* **4** 912–923.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

SKOROHOD, A. V. (1956). Limit theorems for stochastic processes. *Theory Probab. Appl.* **1** 261–290.

WICHURA, M. J. (1970). On the construction of almost uniformly convergent random variables with given weakly convergent image laws. *Ann. Math. Statist.* **41** 1101–1115.

WITHERS, C. S. (1975). Convergence of empirical processes of mixing rv's on [0, 1]. *Ann. Statist.* **3** 1101–1108.

WITHERS, C. S. (1981). Conditions for linear processes to be strong mixing. *Z. Wahrsch. Verw. Gebiete* **57** 477–480.

DÉPARTEMENT TECHNIQUE DE COMMERCIALISATION
INSTITUT UNIVERSITAIRE DE TECHNOLOGIE DE LIMOGES
ALLÉE ANDRÉ MAUROIS
87065 LIMOGES CEDEX
FRANCE

DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47405