A study of serial ranks via random graphs

ERICH HAEUSLER[1], DAVID M. MASON[2] and TATYANA S. TUROVA[3]

[1]*Mathematics Institute, University of Giessen, Arndtstr. 2, D-35392 Giessen, Germany. E-mail: erich.haeusler@math.uni-giessen.de*
[2]*Department of Mathematical Sciences, 501 Ewing Hall, University of Delaware, Newark DE 19716, USA. E-mail: davidm@math.udel.edu*
[3]*Department of Mathematical Statistics, University of Lund, Box 118, S-221 00 Lund, Sweden. E-mail: tatyana@maths.lth.se*

Serial ranks have long been used as the basis for nonparametric tests of independence in time series analysis. We shall study the underlying graph structure of serial ranks. This will lead us to a basic martingale which will allow us to construct a weighted approximation to a serial rank process. To show the applicability of this approximation, we will use it to prove two very general central limit theorems for Wald–Wolfowitz-type serial rank statistics.

Keywords: martingales; random graphs; rank statistics; serial ranks; weighted approximations

1. Introduction and statement of results

Let $X_1, \ldots, X_n$ be independent random variables with a common continuous distribution function. For any $i = 1, \ldots, n$, let $R(i)$ denote the rank of X_i among $X_1, \ldots, X_n$. Then

$$P\{(R(1), \ldots, R(n)) = (\pi(1), \ldots, \pi(n))\} = \frac{1}{n!}$$

for any permutation $(\pi(1), \ldots, \pi(n)) \in \mathscr{P}$, where $\mathscr{P}$ denotes the set of all permutations of $(1, \ldots, n)$. Let $A = (A_1, \ldots, A_n)$ denote the vector of *anti-ranks* of $(R(1), \ldots, R(n))$ – that is, A is the inverse permutation of $(R(1), \ldots, R(n))$ defined by

$$R(A_j) = j, \qquad j = 1, \ldots, n.$$

Clearly, for any permutation $(\pi(1), \ldots, \pi(n)) \in \mathscr{P}$,

$$P\{(A_1, \ldots, A_n) = (\pi(1), \ldots, \pi(n))\} = \frac{1}{n!}.$$

For $n \geqslant 2$, let $c_n(1), \ldots, c_n(n)$ be real constants such that

$$\sum_{i=1}^{n} c_n(i) = 0. \tag{1.1}$$

Denote, for $n \geqslant 2$,

$$S_n(j) = \sum_{i=1}^{j} c_n(R(i))/(n-j), \qquad j = 1, \ldots, n-1;$$

and define, for each $j = 1, \ldots, n-1$, the σ-field

$$\mathscr{G}_n(j) = \sigma(R(1), \ldots, R(j)).$$

Let $\mathscr{G}_n(0)$ be the trivial σ-field. It is readily verified that

$$\{(S_n(j), \mathscr{G}_n(j)) : 1 \leqslant j \leqslant n-1\} \tag{1.2}$$

is a martingale; consult, in particular, Shorack and Wellner (1986). Now, for $n \geqslant 3$, define the rank process

$$C_n(t) = \sum_{i \leqslant nt} c_n(R(i)), \qquad 0 \leqslant t \leqslant 1. \tag{1.3}$$

For the constants $c_n(1), \ldots, c_n(n)$, introduce the following two conditions: there exist finite positive constants $\mathscr{Y}_1$ and $\mathscr{Y}_2$ such that, for all $n \geqslant 2$,

$$\sigma_n^2(c) = \sum_{i=1}^{n} c_n^2(i) \leqslant \mathscr{Y}_1 \tag{1.4}$$

and

$$\mu_n = \sum_{i=1}^{n} c_n^4(i) \leqslant n^{-1} \mathscr{Y}_2. \tag{1.5}$$

(Of course, (1.4) follows from (1.5) by the Cauchy–Schwarz inequality.) Shorack (1991), using the martingale property (1.2) in combination with the Skorohod embedding, established the following weighted approximation for the rank process.

Theorem A. *Assume that* (1.1), (1.4) *and* (1.5) *hold. Then on a rich enough probability space there exist a sequence of versions* $(\tilde{C}_n)_{n \geqslant 1}$ *of* $(C_n)_{n \geqslant 1}$ *(i.e.* $\tilde{C}_n \overset{\mathscr{L}}{=} C_n$ *for every n) and a standard Brownian bridge B such that, for all* $0 \leqslant \nu < \frac{1}{4}$ *and* $\rho > 0$,

$$\sup_{\rho/n \leqslant t \leqslant 1-\rho/n} \frac{|\tilde{C}_n(t) - \sigma_n(c)B(t)|}{(t(1-t))^{1/2-\nu}} = O_p(n^{-\nu}). \tag{1.6}$$

Einmahl and Mason (1992) generalized Shorack's result to a class of exchangeable processes. Weighted approximations to exchangeable processes and the uniform empirical and quantile processes, among other processes, have proved to be tremendously useful in establishing the asymptotic distribution of complicated statistics for which traditional weak convergence methods are difficult to apply. See, for example, Part II of the proceedings volume edited by Hahn *et al.* (1991) and the monograph by Csörgő and Horváth (1993), and the many references therein. For weighted approximations to certain sequences of continuous-time martingales, refer to Haeusler and Mason (1999).

One of our intentions here is to obtain the analogue of Shorack's result for a serial rank type process. To motivate our investigations, consider the nonparametric testing problem

$$H_0 : X_1, \ldots, X_n \text{ are independent } \quad \text{versus} \quad H_1 : X_1, \ldots, X_n \text{ are dependent,}$$

where $X_1, \ldots, X_n$ are considered to be observations of a time series at times $i = 1, \ldots, n$. Introduce the notion of the pairs of *ranks/serial ranks* at lag r, $r \geqslant 1$, based upon the sample $X_1, \ldots, X_n$ taken in the given time order:

$$(R(1), R((1-r)_{\mathrm{mod}n})), \ldots, (R(n), R((n-r)_{\mathrm{mod}n})), \tag{1.7}$$

where for any $-n + 1 \leqslant M \leqslant n - 1$,

$$M_{\mathrm{mod}n} = \begin{cases} M + n, & M \leqslant 0 \\ M, & M > 0. \end{cases}$$

The classical test statistics for H_0 versus H_1 are often based on these pairs of *ranks/ serial ranks* at some lag r, $r \geqslant 1$. For instance, Spearman's rho is formed by choosing $r = 1$ in (1.7) and taking the sample correlation of these pairs. Consult Hallin and Puri (1988; 1991; 1992a; 1992b; 1994; 1995) and the references therein to see how to construct optimal tests based on serial ranks against various time series parametric alternatives.

Towards defining a serial rank process analogous to the rank process given in (1.3), for any $n \geqslant 3$ and $1 \leqslant r \leqslant n - 1$, we set

$$(f_r(1), \ldots, f_r(n)) = (R((A_1 - r)_{\mathrm{mod}n}), \ldots, R((A_n - r)_{\mathrm{mod}n})). \tag{1.8}$$

Notice that $(f_r(1), \ldots, f_r(n))$, for $1 \leqslant r \leqslant n - 1$, is a random permutation of $(1, \ldots, n)$, and observe, for instance, that

$$T_n(r) := \sum_{i=1}^{n} a_n(i) a_n(f_r(i)) = \sum_{i=1}^{n} a_n(R(i)) a_n(R(i-r)_{\mathrm{mod}n}),$$

where $a_n(i)$, $i = 1, \ldots, n$, is any triangular array of constants. The statistic $T_n(r)$ is the classic serial rank test of Wald and Wolfowitz (1943) for testing $H_0 : X_1, \ldots, X_n$ are independent versus $H_1 : X_1, \ldots, X_n$ are dependent.

For $n \geqslant 3$, let $c_n(1), \ldots, c_n(n)$ be real constants such that (1.1) holds, and for $1 \leqslant r \leqslant n - 1$, define the serial rank process at lag r

$$C_n^{(r)}(t) = \sum_{i \leqslant nt} c_n(f_r(i)), \qquad 0 \leqslant t \leqslant 1.$$

One of our aims in this paper is to establish the following weighted approximation for the serial rank process.

Theorem 1.1. *Let r be fixed and assume that (1.1), (1.4) and (1.5) hold along the subsequence $\{n_i\}_{i \geqslant 1}$ of $\{n\}$ of all integers $n \geqslant 3$ such that n_i and r are relatively prime. Then on a rich enough probability space there exist a sequence of versions $(\tilde{C}_{n_i}^{(r)})_{i \geqslant 1}$ of $(C_{n_i}^{(r)})_{i \geqslant 1}$ (i.e. $\tilde{C}_{n_i}^{(r)} \overset{\mathscr{L}}{=} C_{n_i}^{(r)}$ for every i) and a standard Brownian bridge B such that, for all $0 \leqslant \nu < 1/4$ and $\rho > 0$,*

$$\sup_{\rho/n_i \leqslant t \leqslant 1 - \rho/n_i} \frac{|\tilde{C}_{n_i}^{(r)}(t) - \sigma_{n_i}(c)B(t)|}{(t(1-t))^{1/2 - \nu}} = O_p(n_i^{-\nu}). \tag{1.9}$$

We remark that a straightforward modification and simplification of the proof of Theorem 1.1 yields the following weak convergence result for $C_n^{(r)}$.

Theorem 1.2. *Let r be fixed and assume (1.1) and (1.4), and replace (1.5) by the assumption that*

$$\max_{1 \leqslant j \leqslant n_i} c_{n_i}^2(j) \to 0$$

holds along the subsequence $\{n_i\}_{i \geqslant 1}$ of $\{n\}$ of all integers $n \geqslant 3$ such that n_i and r are relatively prime. Then on a rich enough probability space there exist a sequence of versions $(\tilde{C}_{n_i}^{(r)})_{i \geqslant 1}$ of $(C_{n_i}^{(r)})_{i \geqslant 1}$ (i.e. $\tilde{C}_{n_i}^{(r)} \overset{\mathscr{L}}{=} C_{n_i}^{(r)}$ for every i) and a standard Brownian bridge B such that

$$\sup_{0 \leqslant t \leqslant 1} |\tilde{C}_{n_i}^{(r)}(t) - \sigma_{n_i}(c)B(t)| = o_p(1). \tag{1.10}$$

We will not give the routine details of the proof here. However, refer to Remark 3.2 below.

To show the applicability of our weighted approximation for $C_n^{(r)}$ we shall establish two general results on the asymptotic normality of the following class of Wald–Wolfowitz statistics. For $n \geqslant 3$, let $a_n(1), \ldots, a_n(n)$ and $b_n(1), \ldots, b_n(n)$ be constants, and choose $1 \leqslant r \leqslant n - 1$. Consider the statistic

$$W_n(r) := \frac{\sqrt{n} \sum_{i=1}^n (a_n(i) - \overline{a}_n) b_n(f_r(i))}{\sigma_n(a)\sigma_n(b)},$$

where $\sigma_n^2(a) = \sum_{i=1}^n (a_n(i) - \overline{a}_n)^2$, $\overline{a}_n = n^{-1} \sum_{i=1}^n a_n(i)$ and $\sigma_n^2(b) = \sum_{i=1}^n b_n^2(i)$.

For our first central limit theorem, introduce the condition on the constants $a_n(1), \ldots, a_n(n)$, $n \geqslant 3$: for $i = 1, \ldots, n$,

$$a_n(i) = J_1\left(\frac{i}{n+1}\right) - J_2\left(\frac{i}{n+1}\right), \tag{1.11}$$

where J_1 and J_2 are non-increasing functions on $(0, 1)$ satisfying

$$\sum_{i=1}^n \frac{J_1^2(i/(n+1)) + J_2^2(i/(n+1))}{\sigma_n^2(a)} = O(1) \tag{1.12}$$

and

$$\max_{1 \leqslant i \leqslant n} \frac{J_1^2(i/(n+1)) + J_2^2(i/(n+1))}{\sigma_n^2(a)} = o(1). \tag{1.13}$$

For our second central limit theorem we need the following condition on the constants $a_n(1), \ldots, a_n(n)$, $n \geqslant 3$. For some measurable function φ on $(0, 1)$ satisfying $0 < \text{var}\,\varphi(U) < \infty$, where U is Uniform$(0, 1)$, assume

$$\lim_{n \to \infty} \int_0^1 (a_n(1 + [un]) - \varphi(u))^2 \, \mathrm{d}u = 0. \tag{1.14}$$

Furthermore, impose on the constants $b_n(1), \ldots, b_n(n)$, $n \geq 3$, the condition

$$\lim_{\tau \to \infty} \limsup_{n \to \infty} \sum_{i=1}^{n} 1\{b_n^2(i)/\sigma_n^2(b) > \tau/n\} b_n^2(i)/\sigma_n^2(b) = 0. \tag{1.15}$$

For instance, it is readily checked that (1.15) is satisfied if (1.14) holds with $b_n(i)$ replacing $a_n(i)$.

We will establish the following central limit theorem for $W_n(r)$. In the course of its proof we will see how to adapt the approximation in Theorem 1.1 to determine the limiting distribution of functionals of the serial rank process when n and r are not necessarily relatively prime.

Corollary 1.1. *Let $a_n(1), \ldots, a_n(n)$, $n \geq 3$, be a triangular array of constants satisfying $\sigma_n^2(a) > 0$, $\sum_{i=1}^{n} a_n(i) = 0$ and (1.11), (1.12) and (1.13). Further, let $b_n(1), \ldots, b_n(n)$, $n \geq 3$, be a triangular array of constants such that $\sigma_n^2(b) > 0$, $\sum_{i=1}^{n} b_n(i) = 0$ and (1.15) hold. Then for any $r \geq 1$, we have*

$$W_n(r) \xrightarrow{\mathscr{L}} Z \qquad as \ n \to \infty, \tag{1.16}$$

where Z is a standard normal random variable.

Remark 1.1 Notice that the functions J_1 and J_2 need not be square-integrable to satisfy (1.12) and (1.13). For instance, the functions $J_1(u) = J_2(u) = u^{-1/2}$ fulfil (1.12) and (1.13).

Corollary 1.2. *Replace (1.11), (1.12) and (1.13) by (1.14), and keep the rest of the assumptions of Corollary 1.1. Then the conclusion (1.16) remains true.*

Corollaries 1.1 and 1.2 are related to the central limit theorem of Hájek and Šidák (1967) for simple linear rank statistics (see their Theorem V.1.6). As a special case of Corollary 1.1 we get the central limit theorem for the Wald–Wolfowitz statistic recently established by Hallin and Vermandele (1996) (refer, in particular, to their Proposition 11).

Surprisingly, proofs of central limit theorems for Wald–Wolfowitz statistics of the form $W_n(r)$ under minimal conditions turn out to be quite intricate and involved. The proof that Hallin and Vermandele (1996) provided for their central limit theorem is based on a simple technique developed by Lombard (1986). However, their proof is long and highly technical. The proof of our central limit theorem, though significantly shorter than theirs, remains very technical. This is largely due to the problem of moving from the case when r and n are relatively prime to the case when they are not. This problem was recognized by Wald and Wolfowitz (1943), who in their paper only established a central limit theorem along the subsequence $\{n_i\}_{i \geq 1}$ of $\{n\}$ for all integers $n \geq 3$ such that n_i and r are relatively prime. In practical situations when the sample size n is large and r and n are not relatively prime, they suggested reducing to the relatively prime case by tossing out a small number r' less than r observations to make r and $n' = n - r'$ relatively prime and then basing the statistical test on the slightly smaller set of n' observations.

In the process of establishing Theorem 1.1 we will uncover the random graph structure

that lies behind the *ranks/serial ranks* in (1.8). This will guide us to the construction of the basic martingale which will lead through the Skorohod embedding to Theorem 1.1. As far as we know, Hallin *et al.* (1992) were the first to analyse serial rank tests using notions from graph theory. We should also mention the survey article on combinatorics and statistics by Kolchin and Chistyakov (1974). The random graph structure will be derived in Section 2. In Section 3 we will construct the basic martingale and prove Theorem 1.1. The proofs of Corollaries 1.1 and 1.2 will be detailed in Section 4.

2. The graph structure of serial ranks

We shall make use of notions from graph theory. Let us introduce *directed graphs* on the set $\{1, \ldots, n\}$. Any directed graph consists of the set of vertices $V = \{1, \ldots, n\}$, and the set of ordered pairs $\{(i_k, j_k) : k = 1, \ldots, N\}$, $N \geqslant 1$, where $i_k, j_k \in V$ and $i_k \neq j_k$ for each $k = 1, \ldots, N$. The pair (i_k, j_k) represents the arc from the vertex i_k to the vertex j_k. Thus the positive integer N denotes the number of arcs in the directed graph. We say that the vertex i is connected to the vertex j if there is a path of arcs from i to j. Let W be the set of all directed graphs for which $N = n$ and

$$\{i_1, \ldots, i_n\} = \{j_1, \ldots, j_n\} = \{1, \ldots, n\}. \tag{2.1}$$

Hence a directed graph $\Gamma \in W$ is defined uniquely by the set of its arcs, which enables us to use the following notation for the elements of W:

$$\Gamma = \{(i_k, j_k) : k = 1, \ldots, n\}. \tag{2.2}$$

It follows from (2.1) that any directed graph $\Gamma \in W$ has the property that for any vertex i there exist exactly one incoming arc (j, i) for some $j \neq i$ and exactly one outcoming arc (i, l) for some $l \neq i$.

For any $2 \leqslant k \leqslant n$, we call a k-cycle any subset of the arcs of the form

$$\{(i_1, i_2), (i_2, i_3), \ldots, (i_k, i_1)\}, \tag{2.3}$$

where $i_1, i_2, \ldots, i_k$ are distinct.

For any $m > 1$ and $a \in \mathbb{N}$ such that $n = am$, define $W(a, m) \subset W$ to be the set of the directed graphs on $\{1, \ldots, n\}$, which consist of exactly a different m-cycles.

The following result is crucial for our work since it reveals the structure of the permutations $(f_r(1), \ldots, f_r(n))$ defined in (1.8).

Proposition 2.1. *For any $n \geqslant 3$ and $1 \leqslant r \leqslant n - 1$, such that $r = ab$, $n = am$, where a is the greatest common divisor of r and n,*

$$\Omega_{r,n} := \{\{(1, f_r(1)), \ldots, (n, f_r(n))\} : A \in \mathscr{P}\} = W(a, m). \tag{2.4}$$

Proof. First we will show that

$$\Omega_{r,n} \subseteq W(a, m). \tag{2.5}$$

Let us arbitrarily fix $A \in \mathscr{P}$ and $1 \leqslant r \leqslant n - 1$. Consider

$$f_r(1) = R((A_1 - r)_{\text{mod}\,n}), \ldots, f_r(n) = R((A_n - r)_{\text{mod}\,n}). \tag{2.6}$$

We will show that

$$\Gamma := \{(1, f_r(1)), \ldots, (n, f_r(n))\} \in W(a, m). \tag{2.7}$$

Let $i_{1,k}$ be such that $A_{i_{1,k}} = k$, for $k = 1, \ldots, a$, and set

$$A_{i_{s,k}} = (A_{i_{1,k}} - (s-1)r)_{\text{mod}\,n}, \qquad \text{for } s = 2, \ldots, m. \tag{2.8}$$

Assume that, for some $1 \leqslant i \leqslant j \leqslant a$ and $0 \leqslant k_1, k_2 \leqslant m - 1$, we have

$$(j - k_1 r)_{\text{mod}\,n} = (i - k_2 r)_{\text{mod}\,n}. \tag{2.9}$$

This implies that for some $l \in \mathbb{Z}$

$$0 \leqslant j - i = a(lm + (k_2 - k_1)b) < a. \tag{2.10}$$

Obviously, the inequalities in (2.10) can be true only if $j = i$, which in turn forces $k_1 = k_2$ since b and m are relatively prime. Hence (2.9) holds only if both $j = i$ and $k_1 = k_2$. This shows that the

$$A_{i_{s,k}} = (A_{i_{1,k}} - (s-1)r)_{\text{mod}\,n}, \qquad \text{for } s = 1, \ldots, m, \ k = 1, \ldots, a,$$

are n distinct values $\{1, \ldots, n\}$, as are

$$i_{s,k} = R(A_{i_{s,k}}), \qquad \text{for } s = 1, \ldots, m, \ k = 1, \ldots, a. \tag{2.11}$$

Now according to formulae (1.8) and (2.8) for $s = 1, \ldots, m - 1$, $k = 1, \ldots, a$,

$$f_r(i_{s,k}) = R((A_{i_{s,k}} - r)_{\text{mod}\,n}) = R((A_{i_{1,k}} - sr)_{\text{mod}\,n}) = R(A_{i_{s+1,k}}) = i_{s+1,k}$$

and

$$f_r(i_{m,k}) = R((A_{i_{1,k}} - mr)_{\text{mod}\,n}) = R(A_{i_{1,k}}) = i_{1,k}.$$

Clearly each set

$$C_k := \{(i_{s,k}, f_r(i_{s,k})) : s = 1, \ldots, m\} = \{(i_{1,k}, i_{2,k}), (i_{2,k}, i_{3,k}), \ldots, (i_{m,k}, i_{1,k})\}$$

is an m-cycle with set of vertices $\mathscr{V}_k = \{i_{1,k}, \ldots, i_{m,k}\}$, $k = 1, \ldots, a$. Thus

$$\Gamma = \{(i_{s,k}, f_r(i_{s,k})) : s = 1, \ldots, m, \ k = 1, \ldots, a\} = \cup_{k=1}^a C_k \in W(a, m),$$

which confirms (2.5).

Next we will show that

$$W(a, m) \subseteq \Omega_{r,n}. \tag{2.12}$$

Suppose that $\Gamma \in W(a, m)$. Then clearly

$$\Gamma = \cup_{k=1}^a C_k,$$

where

$$C_k = \{(i_{1,k}, i_{2,k}), (i_{2,k}, i_{3,k}), \ldots, (i_{m,k}, i_{1,k})\}, \qquad \text{for } k = 1, \ldots, a.$$

It is easy to show that, for an appropriate permutation $A \in \mathscr{P}$,

$$\Gamma = \{(1, f_r(1)), \ldots, (n, f_r(n))\}. \tag{2.13}$$

To see how this works, set

$$A_{i_{1,k}} = k, \qquad \text{for } k = 1, \ldots, a,$$

and define

$$A_{i_{s,k}} = (A_{i_{1,k}} - (s-1)r)_{\bmod n}, \qquad \text{for } s = 2, \ldots, m,$$

and

$$A_{i_{m+1,k}} = A_{i_{1,k}} = k.$$

Now set

$$f_r(i_{s,k}) = R((A_{i_{1,k}} - sr)_{\bmod n}) = R(A_{i_{s+1,k}}) = i_{s+1,k}, \qquad \text{for } s = 1, \ldots, m-1,$$

and

$$f_r(i_{m,k}) = R(A_{i_{m+1,k}}) = R(A_{i_{1,k}}) = i_{1,k}.$$

Thus

$$\{(i_{1,k}, f_r(i_{1,k})), \ldots, (i_{m,k}, f_r(i_{m,k}))\} = C_k, \qquad \text{for } k = 1, \ldots, a,$$

which yields (2.13) and therefore (2.12). Clearly, (2.5) and (2.12) imply the assertion of our proposition. □

Next we derive some useful corollaries from this proposition.

For any permutation $(f_r(1), \ldots, f_r(n))$ of the set $(1, \ldots, n)$, define its inverse

$$(f_r^{-1}(f_r(1)), \ldots, f_r^{-1}(f_r(n))) = (1, \ldots, n).$$

Then it follows that

$$f_r^{-1}(j) = i \text{ if and only if } (i, j) \in \Gamma, \tag{2.14}$$

where Γ is defined as in (2.7). Set also $f_r^{-k}(j) = f_r^{-1}(f_r^{-k+1}(j))$ for all $k > 1$ and $j \in \{1, \ldots, n\}$. For any $l \in \{1, \ldots, n\}$, $M \subseteq \{1, \ldots, n\}$ and $\Gamma(M) = \{(j, f_r(j)), j \in M\} \subseteq \Gamma \in W(a, m)$, define the length of the longest connected path of the arcs of the graph $\Gamma(M)$ to the vertex l as follows:

$$k(l, \Gamma(M)) := \begin{cases} 0, & \text{if } l \notin V(\Gamma(M)), \\ \max\{0 \leq K < m : (f_r^{-k}(l), f_r^{-k+1}(l)) \in \Gamma(M) & \text{otherwise,} \\ \quad \text{for all } 0 \leq k \leq K\}, \end{cases} \tag{2.15}$$

where $V(\Gamma(M))$ is the set of vertices of the graph $\Gamma(M)$, that is, $V(\Gamma(M)) = \{j, f_r(j) : j \in M\}$. We shall introduce the following notation. Assuming

$$\Gamma = \cup_{k=1}^a C_k = \{(i_{s,k}, \bar{f}_r(i_{s,k})) : s = 1, \ldots, m, k = 1, \ldots, a\} \in W(a, m),$$

write

$$\mathscr{V}_k = \{i_{1,k}, \ldots, i_{m,k}\} = \{s_{1,k}, \ldots, s_{m,k}\}, \qquad \text{for } k = 1, \ldots, a, \tag{2.16}$$

to denote the set of the vertices of the cycle C_k, where

$$1 \leqslant s_{1,k} < \ldots < s_{m,k} \leqslant n$$

for $k = 1, \ldots, a$. In this notation we have, for each $k = 1, \ldots, a$,

$$C_k = \{(s_{1,k}, \bar{f}_r(s_{1,k})), \ldots, (s_{m,k}, \bar{f}_r(s_{m,k}))\}.$$

Further, for each $1 \leqslant j \leqslant m$ and $i = 1, \ldots, a$, set

$$n_i(j) = \#\{t \geqslant 1 : s_{t,i} \leqslant j\}$$

and

$$\Gamma_{ij} := \{(s_{1,i}, \bar{f}_r(s_{1,i})), \ldots, (s_{n_i(j),i}, \bar{f}_r(s_{n_i(j),i}))\}.$$

Corollary 2.1. *Assume that* $(f_r(1), \ldots, f_r(n))$, $n \geqslant 3$, *is a random permutation as in (1.8). Choose any* $1 \leqslant r \leqslant n-1$, *and set* $r = ab$ *and* $n = am$, *where* a *is the greatest common divisor of* r *and* n. *Suppose*

$$\Gamma_j := \{(1, \bar{f}_r(1)), \ldots, (j, \bar{f}_r(j))\} \subset \Gamma, \qquad \text{for some } \Gamma \in W(a, m), \qquad (2.17)$$

where $1 \leqslant j < n$ *is arbitrary but fixed and* $j + 1 \in \mathscr{V}_i$, *for some* $1 \leqslant i \leqslant a$. *Then*

$$P\{f_r(j+1) = \bar{f}_r(j+1) | f_r(1) = \bar{f}_r(1), \ldots, f_r(j) = \bar{f}_r(j), \mathscr{V}_1, \ldots, \mathscr{V}_a\}$$

$$= \begin{cases} \dfrac{1}{m - n_i(j+1)}, & \text{if } 1 \leqslant n_i(j+1) < m, \bar{f}_r(j+1) \in \mathscr{V}_i \text{ and} \\ & \bar{f}_r(j+1) \notin \{\bar{f}_r^{-k(j+1,\Gamma_{ij})}(j+1), \bar{f}_r(s_{1,i}), \ldots, \bar{f}_r(s_{n_i(j),i})\}, \\ 1, & \text{if } n_i(j+1) = m \text{ and } \bar{f}_r(j+1) = \bar{f}_r^{-m+1}(j+1), \\ 0, & \text{otherwise.} \end{cases} \qquad (2.18)$$

Before we proceed with a proof, we point out the important special case of Corollary 2 when r and n are relatively prime.

Corollary 2.2. *Assume that* $(f_r(1), \ldots, f_r(n))$, $n \geqslant 3$, *is a random permutation as in (1.8). Choose any* $1 \leqslant r \leqslant n-1$ *to be relatively prime to* n. *Suppose*

$$\Gamma_j := \{(1, \bar{f}_r(1)), \ldots, (j, \bar{f}_r(j))\} \subset \Gamma, \qquad \text{for some } \Gamma \in W(1, n), \qquad (2.19)$$

where $1 \leqslant j < n$ *is arbitrary but fixed. Then*

$$P\{f_r(j+1) = \bar{f}_r(j+1) | f_r(1) = \bar{f}_r(1), \ldots, f_r(j) = \bar{f}_r(j)\}$$

$$= \begin{cases} \dfrac{1}{n-j-1}, & \text{if } j+1 < n \text{ and} \\[2mm] & \bar{f}_r(j+1) \notin \{\bar{f}_r^{-k(j+1,\Gamma_j)}(j+1), \bar{f}_r(1), \ldots, \bar{f}_r(j)\}, \\[2mm] 1, & \text{if } j+1 = n \text{ and } \bar{f}_r(n) = \bar{f}_r^{-n+1}(n), \\[2mm] 0, & \text{otherwise.} \end{cases} \qquad (2.20)$$

Proof of Corollary 2.1. Note that condition (2.17) entails

$$P\{f_r(1) = \bar{f}_r(1), \ldots, f_r(j) = \bar{f}_r(j)\} > 0.$$

Further, according to Proposition 2.1, the probability on the left-hand side of (2.18) equals zero unless

$$\Gamma_{j+1} = \Gamma_j \cup \{(j+1, \bar{f}_r(j+1))\} \qquad (2.21)$$

$$= \{(1, \bar{f}_r(1)), \ldots, (j, \bar{f}_r(j)), (j+1, \bar{f}_r(j+1))\} \subset \Gamma,$$

for some $\Gamma \in W(a, m)$.

Consider four cases.

   (i) Suppose $n_i(j+1) = 1$. Then condition (2.21) holds if and only if $\bar{f}_r(j+1) \neq j+1$.

   (ii) Suppose $j+1 \notin \{\bar{f}_r(s_{1,i}), \ldots, \bar{f}_r(s_{n_i(j),i})\}$ and $2 \leq n_i(j+1) < m$. Then, clearly, condition (2.21) is satisfied if and only if

$$\bar{f}_r(j+1) \notin \{j+1, \bar{f}_r(s_{1,i}), \ldots, \bar{f}_r(s_{n_i(j),i})\}. \qquad (2.22)$$

   (iii) Suppose $j+1 \in \{\bar{f}_r(s_{1,i}), \ldots, \bar{f}_r(s_{n_i(j),i})\}$ and $2 \leq n_i(j+1) < m$. Then, clearly condition (2.21) is fulfilled if

$$\bar{f}_r(j+1) \notin \{\bar{f}_r^{-k(j+1,\Gamma_{ij})}(j+1), \bar{f}_r(s_{1,i}), \ldots, \bar{f}_r(s_{n_i(j),i})\}. \qquad (2.23)$$

Now assume that (2.23) does not hold, that is,

$$\bar{f}_r(j+1) \in \{\bar{f}_r^{-k(j+1,\Gamma_{ij})}(j+1), \bar{f}_r(s_{1,i}), \ldots, \bar{f}_r(s_{n_i(j),i})\}.$$

Then the set of the arcs

$$\{(s_{1,i}, \bar{f}_r(s_{1,i})), \ldots, (s_{n_i(j),i}, \bar{f}_r(s_{n_i(j),i})), (j+1, \bar{f}_r(j+1))\}$$

makes one k-cycle where obviously $k \leq j+1 < m$, which contradicts (2.21). Hence, in case (iii) condition (2.21) is equivalent to (2.23).

   (iv) Finally, assume $n_i(j+1) = m$. To satisfy condition (2.21) in this case there is only one possible value for $\bar{f}_r(j+1)$, namely

$$\bar{f}_r(j+1) = \bar{f}_r^{-k(j+1,\Gamma_{ij})}(j+1), \qquad (2.24)$$

where, clearly, $k(j+1, \Gamma_{ij}) = m - 1$.

Combining (i), (ii), (iii) and (iv), and taking into account that any value satisfying (2.17) or (2.21) is equally likely, we conclude (2.18). □

3. The basic martingale and proof of Theorem 1.1

3.1. The basic martingale

First we shall introduce a martingale that will be crucial for our approximation.

Assume that $n \geq 3$ and $1 \leq r \leq n - 1$ are relatively prime. Let $b(i)$, for $i = 1, \ldots, n$, denote a sequence of constants satisfying

$$\sum_{j=1}^{n} b(j) = 0.$$

We shall write, for $j = 1, \ldots, n - 2$,

$$K_{n,j} = K_{n,j}(f_r(1), \ldots, f_r(j)) := f_r^{-k(j+1, \Gamma_j)}(j + 1),$$

with $k(j + 1, \Gamma_j)$ defined as in (2.15). For $1 \leq j \leq n - 2$, introduce the centring constants

$$\overline{b}_j = \begin{cases} \dfrac{1}{n-j}(\sum_{i=1}^{j-1} b(f_r(i)) + b(K_{n,j-1})), & \text{if } 2 \leq j \leq n - 2, \\ 0, & \text{if } j = 1, \end{cases}$$

and the σ-fields

$$\mathscr{F}_n(j) = \sigma(f_r(1), \ldots, f_r(j)),$$

with $\mathscr{F}_n(0)$ being the trivial σ-field. Now let $a(i)$, for $i = 1, \ldots, n$, be constants and set, for $j = 1, \ldots, n - 2$,

$$M_n(j) = \sum_{i=1}^{j} a(i)\{b(f_r(i)) + \overline{b}_i\}.$$

Before we proceed we need the following lemma.

Lemma 3.1. *For any choice of $n \geq 3$, constants $c(i)$, $1 \leq i \leq n$, and $1 \leq r \leq n - 1$,*

$$\mathrm{E}(c(f_r(i))|\mathscr{F}_n(i-1)) = \frac{\sum_{l=1}^{n} c(l) - \sum_{l=1}^{i-1} c(f_r(l)) - c(K_{n,i-1})}{n - i} \tag{3.1}$$

and

$$\mathrm{E}c(K_{n,i-1}) = \frac{\sum_{k=1}^{i-1} c(k)}{n - 1} + \frac{(n - i)c(i)}{n - 1}. \tag{3.2}$$

Proof. Notice that by Corollary 2.2 we have (3.1). Taking the expectation of both sides, we obtain

$$\frac{\sum_{l=1}^{n} c(l) - c(i)}{n-1} = \frac{\sum_{l=1}^{n} c(l)}{n-1} + \frac{\sum_{l=1}^{i-1} c(l)}{(n-1)(n-i)} - \frac{Ec(K_{n,i-1})}{n-i}.$$

The proof is finished by solving for $Ec(K_{n,i-1})$ in this last equation.   □

Lemma 3.1 immediately yields the following proposition.

Proposition 3.1. *Let $n \geq 3$ and $1 \leq r \leq n-1$ be integers such that r and n are relatively prime. Then*

$$\{(M_n(j), \mathscr{F}_n(j)) : 1 \leq j \leq n-2\}$$

is a martingale.

By choosing

$$a(i) = \frac{1}{n-i-1} \quad \text{and} \quad b(i) = c_n(i), \qquad \text{for } i = 1, \ldots, n-2,$$

we see after a little calculation that

$$M_n(j) = \frac{\sum_{i=1}^{j} c_n(f_r(i))}{n-j-1} + \sum_{i=2}^{j} \frac{c_n(K_{n,i-1})}{(n-i-1)(n-i)}. \tag{3.3}$$

This will be the basic martingale upon which we will base the proof of Theorem 1.1.

3.2. Proof of Theorem 1.1

For $n \geq 3$, we set

$$D_n^{(r)}(t) = \begin{cases} \displaystyle\sum_{i \leq nt} c_n(f_r(i)) + (n - [nt] - 1) \sum_{2 \leq i \leq nt} \frac{c_n(K_{n,i-1})}{(n-i-1)(n-i)}, & 0 \leq t \leq 1 - \frac{2}{n}, \\[4mm] D_n^{(r)}\left(1 - \frac{2}{n}\right), & 1 - \frac{2}{n} < t \leq 1. \end{cases} \tag{3.4}$$

Then, for $j = 1, \ldots, n-2$,

$$\frac{D_n^{(r)}(j/n)}{n-j-1} = M_n(j). \tag{3.5}$$

For the proof of Theorem 1.1, let r be fixed, and let $\{n_i\}_{i \geq 1}$ denote the subsequence of $\{n\}$ of all integers $n \geq 3$ such that n_i and r are relatively prime.

We will use Proposition 3.1 and the Skorohod embedding for martingales to construct a sequence of versions $(\tilde{D}_{n_i}^{(r)})_{i \geq 1}$ of $(D_{n_i}^{(r)})_{i \geq 1}$ and a standard Brownian bridge B such that, for all $0 \leq \nu < \frac{1}{4}$,

$$\sup_{t \in J} \frac{|\tilde{D}_{n_i}^{(r)}(t) - \sigma_{n_i}(c)B(t)|}{(t(1-t))^{1/2-\nu}} = O_p(n_i^{-\nu}), \tag{3.6}$$

where J is any of the following intervals: $[2/n_i, 1 - 2/n_i]$, $[\rho/n_i, 2/n_i]$, $[1 - 2/n_i, 1 - \rho/n_i]$ with $0 < \rho < 2$, and thus (3.6) holds for any interval J of the form $[\rho'/n_i, 1 - \rho'/n_i]$, with $\rho' > 0$.

From (3.6) we can construct the $\tilde{C}_{n_i}^{(r)}$ of Theorem 1.1 using Lemma A1 of Berkes and Philipp (1979), provided that we can show that, for all $0 < \rho < 1$,

$$\sup_{\rho/n_i \leqslant t \leqslant 1 - \rho/n_i} \frac{|C_{n_i}^{(r)}(t) - D_{n_i}^{(r)}(t)|}{(t(1-t))^{1/2-\nu}} = O_p(n_i^{-\nu}). \tag{3.7}$$

Throughout this proof κ_i, $i \geqslant 1$, will denote finite positive constants independent of $n \geqslant 3$, $\mathscr{Y}_1$ and $\mathscr{Y}_2$. Also, to simplify notation, throughout the rest of the proof we will write n instead of n_i and σ_n instead of $\sigma_n(c)$.

Step 1. First we prove (3.7). Note that $C_n^{(r)}(t) = D_n^{(r)}(t)$ for $0 \leqslant t < 2/n$. Thus, for all $0 < \rho < 1$, we see that the supremum in (3.7) is bounded above by

$$\sup_{2/n \leqslant t < 1 - 1/n} \frac{|C_n^{(r)}(t) - D_n^{(r)}(t)|}{(t(1-t))^{1/2-\nu}} + \sup_{1 - 1/n \leqslant t \leqslant 1 - \rho/n} \frac{|C_n^{(r)}(t) - D_n^{(r)}(t)|}{(t(1-t))^{1/2-\nu}} =: I_n + II_n.$$

By definition of $C_n^{(r)}$ and $D_n^{(r)}$ and some easy computations, we obtain

$$\mathrm{E}I_n \leqslant \mathrm{E}\left(\sup_{2/n \leqslant t < 1 - 1/n} \frac{n - [nt] - 1}{(t(1-t))^{1/2-\nu}} \sum_{j=2}^{[nt]} \frac{|c_n(K_{n,j-1})|}{(n-j-1)(n-j)}\right)$$

$$\leqslant \kappa_1 \mathrm{E}\left(\max_{2 \leqslant i \leqslant n-1} \left(\frac{i}{n}\left(1 - \frac{i}{n}\right)\right)^{-1/2+\nu} (n-i) \sum_{j=2}^{i} \frac{|c_n(K_{n,j-1})|}{(n-j)^2}\right),$$

which, by the elementary fact that, for any non-negative a_i, $i = 1, \ldots, n$, and $0 \leqslant b_n \leqslant \ldots \leqslant b_1$, $n \geqslant 1$,

$$\max_{1 \leqslant i \leqslant n} b_i \sum_{j=1}^{i} a_j \leqslant \sum_{i=1}^{n} a_i b_i, \tag{3.8}$$

and noting that $i^{-1/2+\nu}(n-i)^{1/2+\nu}$ is non-increasing in i, in turn, gives

$$\mathrm{E}I_n \leqslant \kappa_1 n^{1-2\nu} \sum_{i=2}^{n-1} i^{-1/2+\nu}(n-i)^{-3/2+\nu} E|c_n(K_{n,i-1})|.$$

Now by Lemma 3.1,

$$\mathrm{E}\big|c_n(K_{n,i-1})\big| \leqslant 2\frac{\sum_{k=1}^{n-1}|c_n(k)|}{n} + 2\frac{(n-i)|c_n(i)|}{n},$$

$$\leqslant 2n^{-1/2}\sigma_n + 2\frac{(n-i)|c_n(i)|}{n}$$

by the Cauchy–Schwarz inequality. Thus

$$\mathrm{E}I_n \leqslant 2\kappa_1 n^{1/2-2\nu}\sum_{i=2}^{n-1} i^{-1/2+\nu}(n-i)^{-3/2+\nu}\sigma_n + 2\kappa_1 n^{-2\nu}\sum_{i=2}^{n-1} i^{-1/2+\nu}(n-i)^{-1/2+\nu}|c_n(i)|.$$

Using assumption (1.4), it is readily shown that this last bound is $O(n^{-\nu})$, which implies that $I_n = O_p(n^{-\nu})$.

Now, for any $1 - 1/n \leqslant t < 1$, we have

$$\big|C_n^{(r)}(t) - D_n^{(r)}(t)\big| \leqslant |c_n(f_r(n-1))| + \sum_{i=2}^{n-2}\frac{|c_n(K_{n,i-1})|}{(n-i-1)(n-i)}.$$

Therefore, by a similar argument to that for I_n,

$$II_n = O_p(n^{-\nu}).$$

This completes the proof of step 1.

Step 2. Next we shall construct $\tilde{D}_{n_i}^{(r)}$ such that (3.6) holds. Notice that, for any sequence of versions $\tilde{D}_n^{(r)}$ of $D_n^{(r)}$, we have

$$\max_{2 \leqslant j \leqslant n-2}\ \sup_{j/n \leqslant t < (j+1)/n}\big|\tilde{D}_n^{(r)}(t) - \tilde{D}_n^{(r)}(j/n)\big| = 0,$$

so that the supremum in (3.6) when $J = [2/n_i, 1 - 2/n_i]$ is bounded by

$$\max_{2 \leqslant j \leqslant n-2}\ \sup_{j/n \leqslant t \leqslant (j+1)/n}\frac{\big|\tilde{D}_n^{(r)}(j/n) - \sigma_n B(j/n)\big|}{(t(1-t))^{1/2-\nu}} + \max_{2 \leqslant j \leqslant n-2}\ \sup_{j/n \leqslant t \leqslant (j+1)/n}\frac{\big|\sigma_n B(j/n) - \sigma_n B(t)\big|}{(t(1-t))^{1/2-\nu}}$$

$$=: \tilde{I}_n + \widetilde{II}_n.$$

By (2.25) in Einmahl and Mason (1992), we have $\widetilde{II}_n = O_p(n^{-\nu})$. Observe that, by smoothness properties of $(t(1-t))^{1/2-\nu}$, we can derive $\tilde{I}_n = O_p(n^{-\nu})$ from the statement

$$\max_{2 \leqslant j \leqslant n-2}\frac{\big|\tilde{D}_n^{(r)}(j/n) - \sigma_n B(j/n)\big|}{((j/n)(1-j/n))^{1/2-\nu}} = O_p(n^{-\nu}). \tag{3.9}$$

Therefore it is enough to construct $\tilde{D}_n^{(r)}$ such that (3.9) holds in order to verify (3.6) when $J = [2/n_i, 1 - 2/n_i]$.

This construction will be performed by an application of the Skorohod embedding. Notice that the martingale

$$\frac{nD_n^{(r)}\left(\dfrac{j}{n}\right)}{n-j-1} = nM_n(j), \qquad 1 \leqslant j \leqslant n-2,$$

is not centred, but has expectation

$$m_n = \mathrm{E}(nM_n(1)) = \frac{n\,\mathrm{E}(c_n(f_r(1)))}{n-2} = \frac{-nc_n(1)}{(n-1)(n-2)}.$$

Hence we must apply the Skorohod embedding to the centred martingale

$$nZ_n(j) = nM_n(j) - m_n, \qquad 1 \leqslant j \leqslant n-2,$$

whose difference sequence $\xi_n(1), \ldots, \xi_n(n-2)$ is given by

$$\xi_n(j) = \begin{cases} nM_n(1) - m_n, & j = 1, \\ n(M_n(j) - M_n(j-1)), & 2 \leqslant j \leqslant n-2. \end{cases}$$

For later use, we record that by (3.1) of Lemma 3.1, we have, for $j = 1, \ldots, n-2$,

$$\xi_n(j) = \frac{n}{n-j-1}(c_n(f_r(j)) - \mathrm{E}\{c_n(f_r(j))|\mathscr{F}_n(j-1)\}), \tag{3.10}$$

which, by the way, clearly shows that nZ_n is a centred martingale.

Let $(\Omega, \mathscr{A}, P)$ be a probability space which carries a standard Wiener process W. For each $n \geqslant 3$, define the Wiener process

$$W_n(t) := \sigma_n W(t/\sigma_n^2), \qquad 0 \leqslant t < \infty.$$

Then there exist stopping times $T_n(j) \leqslant T_n(j+1)$, $1 \leqslant j \leqslant n-2$, with respect to W_n such that the random variables

$$n\zeta_n(j) := W_n(T_n(j)), \qquad 1 \leqslant j \leqslant n-2, \tag{3.11}$$

satisfy

$$(n\zeta_n(j))_{1 \leqslant j \leqslant n-2} \overset{\mathscr{L}}{=} (nZ_n(j))_{1 \leqslant j \leqslant n-2}. \tag{3.12}$$

Moreover, the following is true. Let $\mathscr{B}[n, t] = \sigma(W_n(s) : 0 \leqslant s \leqslant t)$ for $0 \leqslant t < \infty$, and let $\mathscr{B}_n(j)$ be the σ-field of the events in the filtration which are known at time $T_n(j)$ for $1 \leqslant j \leqslant n-2$. Let $\tau_n(j)$, $1 \leqslant j \leqslant n-2$, be defined by $\tau_n(1) = T_n(1)$ and

$$\tau_n(i) = T_n(j) - T_n(j-1), \qquad \text{for } 2 \leqslant j \leqslant n-2. \tag{3.13}$$

Then each $\tau_n(j)$ is non-negative and measurable with respect to $\mathscr{B}_n(j)$, and for the difference sequence $\eta_n(1), \ldots, \eta_n(n-2)$ of the martingale $n\zeta_n(1), \ldots, n\zeta_n(n-2)$ we have, for $1 \leqslant j \leqslant n-2$,

$$\mathrm{E}(\tau_n(j)|\mathscr{B}_n(j-1)) = \mathrm{E}(\eta_n^2(j)|\mathscr{B}_n(j-1))$$

$$= \mathrm{E}(\eta_n^2(j)|\eta_n(1), \ldots, \eta_n(j-1)), \tag{3.14}$$

where $\mathscr{B}_n(0)$ is the trivial σ-field. Finally, for any $1 < p < \infty$, there exists a constant $C_p < \infty$ depending only on p such that, for all $1 \le j \le n - 2$,

$$\mathrm{E}(\tau_n^p(j)|\mathscr{B}_n(j-1)) \le C_p\mathrm{E}(|\eta_n(j)|^{2p}|\mathscr{B}_n(j-1))$$

$$= C_p\mathrm{E}(|\eta_n(j)|^{2p}|\eta_n(1), \ldots, \eta_n(j-1)) \tag{3.15}$$

(see Hall and Heyde 1980; Scott and Huggins 1983). Statement (3.9) will be derived later on from

$$\max_{2 \le j \le n-2} \frac{|n\zeta_n(j) - W_n(j\sigma_n^2/(n-j))|}{(j/n)^{1/2-\nu}(1 - j/n)^{-1/2-\nu}} = O_p(n^{-\nu}). \tag{3.16}$$

Statement (3.16), itself, is derived from the following proposition by exactly the same arguments as employed in the proof of Theorem 1 in Einmahl and Mason (1992, pp. 113–115). We will not detail these arguments here but concentrate on the proof of the proposition.

Proposition 3.2. *For* $0 \le \delta < \frac{1}{4}$, *we have, along the sequence of all integers n such that n and r are relatively prime,*

$$\max_{2 \le j \le n-2} j^{2\delta-1}(n-j)^{1+2\delta}\left|T_n(j) - \frac{j\sigma_n^2}{n-j}\right| = O_p(n^{2\delta}).$$

Proof. We write, using (3.14),

$$\max_{2 \le j \le n-2} j^{2\delta-1}(n-j)^{1+2\delta}\left|T_n(j) - \frac{j\sigma_n^2}{n-j}\right|$$

$$\le \max_{2 \le j \le n-2} j^{2\delta-1}(n-j)^{1+2\delta}\left|\sum_{i=1}^j [\tau_n(i) - \mathrm{E}(\tau_n(i)|\mathscr{B}_n(i-1))]\right|$$

$$+ \max_{2 \le j \le n-2} j^{2\delta-1}(n-j)^{1+2\delta}\left|\sum_{i=1}^j \mathrm{E}(\eta_n^2(i)|\eta_n(1), \ldots, \eta_n(i-1)) - \frac{j\sigma_n^2}{n-j}\right|$$

$$=: \bar{I}_n + \bar{II}_n.$$

Consider $\bar{I}_n$ first. By the Hájek–Rényi inequality we obtain, for $0 < c < \infty$,

$$p_n(c) = P(n^{-2\delta}\bar{I}_n \ge c) \tag{3.17}$$

$$\le \kappa_2 c^{-2} n^{-4\delta} \sum_{i=1}^{n-2} i^{4\delta-2}(n-i)^{2+4\delta}\mathrm{E}([\tau_n(i) - \mathrm{E}(\tau_n(i)|\mathscr{B}_n(i-1))]^2).$$

Clearly, by (3.15), for $1 \le i \le n - 2$,

$$\mathrm{E}([\tau_n(i) - \mathrm{E}(\tau_n(i)|\mathscr{B}_n(i-1))]^2) \le \mathrm{E}(\tau_n^2(i)) \le C_2\mathrm{E}(\eta_n^4(i)) = C_2\mathrm{E}(\xi_n^4(i)), \tag{3.18}$$

so that

$$p_n(c) \leqslant C_2 \kappa_2 c^{-2} n^{-4\delta} \sum_{i=1}^{n-2} i^{4\delta-2} (n-i)^{2+4\delta} \mathrm{E}(\xi_n^4(i)). \qquad (3.19)$$

We shall bound $\mathrm{E}(\xi_n^4(i))$. Notice that by (3.10), the c_r-inequality and Jensen's inequality in combination with Corollary 2.2, we obtain, for $i = 1, \ldots, n,$

$$\mathrm{E}(\xi_n^4(i)) \leqslant 32 \left(\frac{n}{n-i-1} \right)^4 \mathrm{E} c_n^4 (f_r(i)) \leqslant 64 \left(\frac{n}{n-i-1} \right)^4 \frac{\mu_n}{n}, \qquad (3.20)$$

where μ_n is as in (1.5). Combining (3.19) and (3.20), we find

$$p_n(c) \leqslant 64 \kappa_3 C_2 c^{-2} n^{3-4\delta} \mu_n \sum_{i=1}^{n-2} i^{4\delta-2} (n-i)^{-2+4\delta}.$$

It can be readily shown that

$$n^{2-4\delta} \sum_{i=1}^{n-2} i^{4\delta-2} (n-i)^{-2+4\delta} = O(1).$$

Thus, in view of assumption (1.5),

$$p_n(c) \leqslant 64 \kappa_3 C_2 c^{-2} n \mu_n O(1) = O(c^{-2}),$$

which yields $\bar{I}_n = O_p(n^{2\delta})$.

For the proof of $\bar{II}_n = O_p(n^{2\delta})$, notice that

$$\bar{II}_n \stackrel{\mathscr{L}}{=} \max_{2 \leqslant j \leqslant n-2} j^{2\delta-1} (n-j)^{1+2\delta} \left| \sum_{i=1}^{j} \mathrm{E}(\xi_n^2(i) | \xi_n(1), \ldots, \xi_n(j-1)) - \frac{j\sigma_n^2}{n-j} \right|$$

$$\leqslant \max_{2 \leqslant j \leqslant n-2} j^{2\delta-1} (n-j)^{1+2\delta} \left| \sum_{i=1}^{j} [\mathrm{E}(\xi_n^2(i) | \xi_n(1), \ldots, \xi_n(j-1)) - \xi_n^2(i)] \right|$$

$$+ \max_{2 \leqslant j \leqslant n-2} j^{2\delta-1} (n-j)^{1+2\delta} \left| \sum_{i=1}^{j} [\xi_n^2(i) - \mathrm{E}(\xi_n^2 | \mathscr{F}_n(i-1))] \right|$$

$$+ \max_{2 \leqslant j \leqslant n-2} j^{2\delta-1} (n-j)^{1+2\delta} \left| \sum_{i=1}^{j} \mathrm{E}(\xi_n^2(i) | \mathscr{F}_n(i-1)) - \frac{j\sigma_n^2}{n-j} \right|$$

$$=: R_{n,1} + R_{n,2} + R_{n,3}.$$

Note that the summands appearing in $R_{n,1}$ and $R_{n,2}$ are martingale differences. Therefore, the Hájek-Rényi inequality can be applied as in (3.17) to obtain, for $0 < c < \infty$ and $j = 1, 2,$

$$P(n^{-2\delta} R_{n,j} \geqslant c) \leqslant \kappa_4 c^{-2} n^{-4\delta} \sum_{i=1}^{n-2} i^{4\delta-2} (n-i)^{2+4\delta} \mathrm{E}(\xi_n^4(i)).$$

This is the same bound as in (3.19). Therefore, we have $R_{n,j} = O_p(n^{2\delta})$, for $j = 1, 2$. It remains to consider $R_{n,3}$. Note that

$$\frac{j}{n-j} = \sum_{i=1}^{j} \frac{n}{(n-i)(n-i+1)},$$

so that we can write after a few steps

$$R_{n,3} \leq \max_{2 \leq j \leq n-2} j^{2\delta-1}(n-j)^{1+2\delta} \left| \sum_{i=2}^{j} \left[E(\xi_n^2(i)|\mathscr{F}_n(i-1)) - \frac{n(n-i+1)\sigma_n^2}{(n-i)(n-i-1)^2} \right] \right|$$

$$+ n^{1+2\delta} \left| \left(\frac{n}{n-2} \right)^2 E(c_n^2(f_n(1))) - m_n^2 - \frac{\sigma_n^2}{n-1} \right| + O(n^{2\delta})$$

$$=: R'_{n,3} + R''_{n,3} + O(n^{2\delta}).$$

Then it is easy to check that

$$R''_{n,3} \leq \kappa_5 \sigma_n^2 n^{2\delta},$$

which by assumption (1.4) gives $R''_{n,3} = O(n^{2\delta})$.

It remains to treat $R'_{n,3}$. For this, we need to compute $E(\xi_n^2(i)|\mathscr{F}_n(i-1))$, for $2 \leq i \leq n-2$. Using (3.1) in Lemma 3.1 along with (3.10), we obtain

$$E(\xi_n^2(i)|\mathscr{F}_n(i-1))$$

$$= \frac{n^2}{(n-i-1)^2(n-i)} \left\{ \sum_{l=i}^{n} c_n^2(f_r(l)) - c_n^2(K_{n,i-1}) - \frac{1}{n-i} \left(\sum_{l=1}^{i-1} c_n(f_r(l)) + c_n(K_{n,i-1}) \right)^2 \right\}.$$

$$(3.21)$$

Using this representation, we can write

$$R'_{n,3} \leq \max_{2 \leq j \leq n-2} j^{2\delta-1}(n-j)^{1+2\delta} \sum_{i=2}^{j} \frac{n^2}{(n-i-1)^2(n-i)} \left| \sum_{l=i}^{n} (c_n^2(f_r(l)) - \sigma_n^2/n) \right|$$

$$+ \max_{2 \leq j \leq n-2} j^{2\delta-1}(n-j)^{1+2\delta} \sum_{i=2}^{j} \frac{n^2}{(n-i-1)^2(n-i)^2} \left(\sum_{l=i}^{n} c_n(f_r(l)) - c_n(K_{n,i-1}) \right)^2$$

$$+ \max_{2 \leq j \leq n-2} j^{2\delta-1}(n-j)^{1+2\delta} \sum_{i=2}^{j} \frac{n^2 c_n^2(K_{n,i-1})}{(n-i-1)^2(n-i)},$$

that is,

$$R'_{n,3} \leqslant \kappa_6 \max_{2 \leqslant j \leqslant n-2} j^{2\delta-1}(n-j)^{1+2\delta} \sum_{i=2}^{j} \frac{n^2}{(n-i)^3} \left| \sum_{l=i}^{n} (c_n^2(f_r(l)) - \sigma_n^2/n) \right|$$

$$+ \kappa_6 \max_{2 \leqslant j \leqslant n-2} j^{2\delta-1}(n-j)^{1+2\delta} \sum_{i=2}^{j} \frac{n^2}{(n-i)^4} \left(\sum_{l=i}^{n} c_n(f_r(l)) \right)^2$$

$$+ \kappa_6 \max_{2 \leqslant j \leqslant n-2} j^{2\delta-1}(n-j)^{1+2\delta} \sum_{i=2}^{j} \frac{n^2 c_n^2(K_{n,i-1})}{(n-i)^3}. \tag{3.22}$$

Now by applying the elementary fact (3.8) and noticing that $j^{2\delta-1}(n-j)^{1+2\delta}$ is non-increasing in j, we obtain

$$R'_{n,3} \leqslant \kappa_6 n^2 \sum_{i=2}^{n-2} i^{2\delta-1}(n-i)^{-2+2\delta} \left| \sum_{l=i}^{n} d_n(f_r(l)) \right|$$

$$+ \kappa_6 n^2 \sum_{i=2}^{n-2} i^{2\delta-1}(n-i)^{-3+2\delta} \left(\sum_{l=i}^{n} c_n(f_n(l)) \right)^2$$

$$+ \kappa_6 n^2 \sum_{i=2}^{n-2} i^{2\delta-1}(n-i)^{-2+2\delta} c_n^2(K_{n,i-1})$$

$$=: r_{n,1} + r_{n,2} + r_{n,3},$$

where $d_n(k) := c_n^2(k) - \sigma_n^2/n$ for $k = 1, \ldots, n$.

To bound $r_{n,1}$ and $r_{n,2}$ we need the following lemma, which we state in a more general form than we need here, but which will be used in this form later on.

Lemma 3.2. *Let $a(i)$, $b(i)$, $1 \leqslant i \leqslant n$, $n \geqslant 3$, be constants such that $\sum_{i=1}^{n} a(i) = 0$. Let a be the greatest common divisor of n and $1 \leqslant r \leqslant n-1$. Then whenever $m > 2$, with $n = ma$, or r is relatively prime to n,*

$$E \left(\sum_{i=1}^{n} a(i)b(f_r(i)) \right)^2 \leqslant \frac{2}{n} \sum_{i=1}^{n} a^2(i) \sum_{i=1}^{n} b^2(i). \tag{3.23}$$

Proof. It is easy to see, taking into account Proposition 2.1, that for any $i \neq j$

$$Eb(f_r(i))b(f_r(j)) = \frac{1}{(n-1)(n-2)} \sum{}^{1} b(k)b(l), \tag{3.24}$$

where $\sum^{1}$ runs over all $1 \leqslant k, l \leqslant n$, such that $k \neq l$, $k \neq i$, $l \neq j$ and $(k, l) \neq (j, i)$. This, together with Proposition 2.1, gives us after some computation

$$E\left(\sum_{i=1}^{n} a(i)b(f_r(i))\right)^2 = \frac{\sum_{i=1}^{n} a^2(i)\sum_{i=1}^{n} b^2(i)}{n-2} - \frac{n\sum_{i=1}^{n} a^2(i)b^2(i)}{(n-1)(n-2)}$$

$$\leq \frac{2}{n}\sum_{i=1}^{n} a^2(i)\sum_{i=1}^{n} b^2(i),$$

where we used the condition $\sum_{i=1}^{n} a(i) = 0$.   □

From Lemma 3.2 with

$$a(l) = \begin{cases} 1 - \dfrac{i-1}{n}, & \text{for } 1 \leq l \leq i-1, \\[2mm] -\dfrac{i-1}{n}, & \text{for } i-1 < l \leq n, \end{cases}$$

and $b(i) = c_n(i)$, we obtain that, for any $1 < i \leq n$,

$$E\left(\sum_{l=1}^{i-1} c_n(f_r(l))\right)^2 \leq 2\left(\frac{i-1}{n}\right)\left(1 - \frac{i-1}{n}\right)\sum_{l=1}^{n} c_n^2(l). \qquad (3.25)$$

Further, by noting that $\sum_{k=1}^{n} d_n(k) = 0$, we obtain that

$$E\left(\sum_{l=1}^{i-1} d_n(f_r(l))\right)^2 \leq 2\left(\frac{i-1}{n}\right)\left(1 - \frac{i-1}{n}\right)\sum_{l=1}^{n} d_n^2(l)$$

$$\leq 2\left(\frac{i-1}{n}\right)\left(1 - \frac{i-1}{n}\right)\mu_n. \qquad (3.26)$$

We now see, using (3.26), that

$$Er_{n,1} \leq \kappa_7 n \sum_{i=2}^{n-2} i^{2\delta-1/2}(n-i)^{-3/2+2\delta}\mu_n^{1/2} = O(n^{2\delta})$$

and, applying (3.25), that

$$Er_{n,2} \leq \kappa_8 \sum_{i=2}^{n-2} i^{2\delta}(n-i)^{-2+2\delta}\sigma_n^2 = O(n^{2\delta}).$$

Further, by Lemma 3.1,

$$\mathrm{E}r_{n,3} = \kappa_9 n^2 \sum_{i=2}^{n-2} i^{2\delta-1}(n-i)^{-2+2\delta}\left(\frac{\sum_{k=1}^{i-1}c_n^2(k)}{n-1} + \frac{(n-i)c_n^2(i)}{n-1}\right)$$

$$\leqslant \kappa_9 n \sum_{i=2}^{n-2} i^{2\delta-1}(n-i)^{-2+2\delta}\sigma_n^2 + \kappa_9 n \sum_{i=2}^{n-2} i^{2\delta-1}(n-i)^{-1+2\delta}c_n^2(i)$$

$$\leqslant \kappa_{10} n^{2\delta}\sigma_n^2 + \kappa_9 n\sqrt{\mu_n}\left(\sum_{i=2}^{n-2} i^{4\delta-2}(n-i)^{-2+4\delta}\right)^{1/2},$$

which, by (1.4) and (1.5), is $O(n^{2\delta})$. This completes the proof of Proposition 3.2. $\square$

We can now define the vector $(\tilde{D}_n^{(r)}(2/n), \ldots, \tilde{D}_n^{(r)}((n-2)/n))$ and then verify (3.9). Set

$$\frac{n\tilde{D}_n^{(r)}(j/n)}{n-j-1} = n\zeta_n(j) + m_n, \qquad 2 \leqslant j \leqslant n-2.$$

Then we have

$$\frac{n\tilde{D}_n^{(r)}(j/n)}{n-j-1} \overset{\mathscr{L}}{=} nZ_n(j) + m_n = nM_n(j) = \frac{nD_n^{(r)}(j/n)}{n-j-1},$$

where, in view of (3.12), the equality in distribution holds uniformly in $2 \leqslant j \leqslant n-2$. This shows

$$(\tilde{D}_n^{(r)}(2/n), \ldots, \tilde{D}_n^{(r)}((n-2)/n)) \overset{\mathscr{L}}{=} (D_n^{(r)}(2/n), \ldots, D_n^{(r)}((n-2)/n)).$$

For the proof of (3.9) we proceed as in Einmahl and Mason (1992, p. 113) and define the required Brownian bridge by

$$B(t) = (1-t)W\left(\frac{t}{1-t}\right), \qquad 0 \leqslant t < 1,$$

and $B(1) = 0$. Notice that

$$\sigma_n B(t) = (1-t)W_n\left(\frac{t\sigma_n^2}{1-t}\right).$$

Then, after a little algebra, we obtain

$$\max_{2\leqslant j\leqslant n-2} \frac{|\tilde{D}_n^{(r)}(j/n) - \sigma_n B(j/n)|}{((j/n)(1-j/n))^{1/2-\nu}} \leqslant \max_{2\leqslant j\leqslant n-2} \frac{|n\zeta_n(j) - W_n(j\sigma_n^2/(n-j))|}{(j/n)^{1/2-\nu}(1-j/n)^{-1/2-\nu}}$$

$$+ \max_{2\leqslant j\leqslant n-2} \frac{n^{-1}|W_n(j\sigma_n^2/(n-j))|}{(j/n)^{1/2-\nu}(1-j/n)^{1/2-\nu}} + |m_n|n^{1/2-\nu}$$

$$:= I_n' + II_n' + III_n'.$$

The first summand, I_n', is $O_p(n^{-\nu})$ by (3.16). Further, the easily established facts that

$$\sum_{j=1}^{n-1} \frac{\mathrm{E}\,W_n^2(j\sigma_n^2/(n-j))}{n^2(j/n)^{1-2\nu}(1-j/n)^{1-2\nu}} = O(n^{-2\nu})$$

and

$$|m_n| \leqslant 4\sigma_n n^{-1/2}, \qquad \text{for } n \geqslant 3,$$

along with $0 \leqslant \nu < \frac{1}{4}$ and assumption (1.4), imply $II_n' = O_p(n^{-\nu})$ and $III_n' = O(n^{-\nu})$. Hence we have achieved the construction of the versions $\tilde{D}_n^{(r)}$ such that (3.9) holds. Thus, by the arguments just preceding (3.9), we see that (3.6) is satisfied when $J = [2/n_i, 1 - 2/n_i]$.

It remains to verify that (3.6) also holds when $J = [\rho/n_i, 2/n_i]$ and $[1 - 2/n_i, 1 - \rho/n_i]$, with $0 < \rho < 2$. For this, note that for all $0 < \rho \leqslant 2$ we have

$$\sup_{\rho/n \leqslant t \leqslant 2/n} \frac{\sigma_n |B(t)|}{(t(1-t))^{1/2-\nu}} + \sup_{1-2/n \leqslant t \leqslant 1-\rho/n} \frac{\sigma_n |B(t)|}{(t(1-t))^{1/2-\nu}} = O_p(n^{-\nu});$$

see Einmahl and Mason (1992, p. 115). Therefore, it is enough to show that versions $\tilde{D}_n^{(r)}$ of $D_n^{(r)}$ can be constructed so that also

$$\sup_{\rho/n \leqslant t < 2/n} \frac{|\tilde{D}_n^{(r)}(t)|}{(t(1-t))^{1/2-\nu}} + \sup_{1-2/n < t \leqslant 1-\rho/n} \frac{|\tilde{D}_n^{(r)}(t)|}{(t(1-t))^{1/2-\nu}} = O_p(n^{-\nu}). \tag{3.27}$$

Since we can construct $\tilde{D}_n^{(r)}$ on $(\rho/n, 2/n) \cup (1 - 2/n, 1 - \rho/n)$ from $\tilde{D}_n^{(r)}$ on $[2/n, 1 - 2/n]$, it suffices to show that the processes $D_n^{(r)}$ which we want to mimic have the desired behaviour, that is, it remains to verify (3.27) for $D_n^{(r)}$ instead of $\tilde{D}_n^{(r)}$. This, however, is easily done employing the definition of $D_n^{(r)}$, which gives

$$D_n^{(r)}(t) = \begin{cases} 0, & 0 \leqslant t < 1/n \\ c_n(f_r(1)), & 1/n \leqslant t < 2/n \\ D_n^{(r)}(1 - 2/n), & 1 - 2/n \leqslant t \leqslant 1. \end{cases}$$

We will not give the details of these elementary arguments. This completes the proof of step 2. Clearly steps 1 and 2 imply Theorem 1.1. $\qquad\square$

Remark 3.1. For reference later on, we note that an examination of the proof of Theorem 1.1 shows that if γ_i denotes the $O_P(n_i^{-\nu})$ term in (1.6), then for all $x > 0$ and for some function ρ

$$\sup_{i \geqslant 1} P(n_i^\nu \gamma_i > x) \leqslant \rho(x),$$

satisfying $\rho(x) \to 0$, as $x \to \infty$, where the function ρ depends only on the $\mathscr{Y}_1$ and $\mathscr{Y}_2$ appearing in (1.4) and (1.5) and $0 \leqslant \nu < \frac{1}{4}$.

Remark 3.2. Using the trivial inequality

$$\mu_n = \sum_{i=1}^n c_n^4(i) \leqslant \max_{1 \leqslant j \leqslant n} c_n^2(j)\sigma_n^2(c),$$

it is routine to modify the proof of Theorem 1.1 to yield Theorem 1.2.

4. Proofs of Corollaries 1.1 and 1.2

4.1. Proof of Corollary 1.1

Choose a $\gamma > 0$ and set, for $i = 1, \ldots, n$,

$$c_n(i) = b_n(i)/\sigma_n(b),$$

$$c^t_{\gamma,n}(i) = c_n(i) \text{ or } 0, \text{ according as } c^2_n(i) \leq \gamma/n \text{ or not,}$$

$$c_{\gamma,n}(i) = c^t_{\gamma,n}(i) - n^{-1} \sum_{k=1}^{n} c^t_{\gamma,n}(k).$$

Notice that

$$W_n(r) = \frac{\sqrt{n} \sum_{i=1}^{n} a_n(i) c_n(f_r(i))}{\sigma_n(a)}$$

$$= \frac{\sqrt{n} \sum_{i=1}^{n} a_n(i) c_{\gamma,n}(f_r(i))}{\sigma_n(a)} + \frac{\sqrt{n} \sum_{i=1}^{n} a_n(i)(c_n(f_r(i)) - c_{\gamma,n}(f_r(i)))}{\sigma_n(a)}$$

$$=: W_{\gamma,n}(r) + W^t_{\gamma,n}(r).$$

Let $n = ma$ and $r = sa$, where a is the greatest common divisor of r and n, and $C_1, \ldots, C_a$ be the m-cycles as in Proposition 2.1. Further, for $k = 1, \ldots, a$, let

$$\mathscr{V}_k = \{i_{1,k}, \ldots, i_{m,k}\} = \{s_{1,k}, \ldots, s_{m,k}\}$$

as in (2.16) denote the set of the vertices of the cycle C_k, where

$$1 \leq s_{1,k} < \ldots < s_{m,k} \leq n.$$

Write, for $k = 1, \ldots, a$,

$$\bar{a}_{k,m} = m^{-1} \sum_{i \in \mathscr{V}_k} a_n(i) \quad \text{and} \quad \bar{c}_{\gamma,k,m} = m^{-1} \sum_{i \in \mathscr{V}_k} c_{\gamma,n}(i).$$

With this notation we can write

$$W_{\gamma,n}(r) = \frac{\sqrt{n} \sum_{k=1}^{a} \sum_{i \in \mathscr{V}_k} a_n(i)(c_{\gamma,n}(f_r(i)) - \bar{c}_{\gamma,k,m})}{\sigma_n(a)} + \frac{\sqrt{n} m \sum_{k=1}^{a} \bar{a}_{k,m} \bar{c}_{\gamma,k,m}}{\sigma_n(a)}$$

$$=: \sum_{k=1}^{a} T_{\gamma,k,n}(s) + \Delta_{\gamma,n}(r) =: V_{\gamma,n}(r) + \Delta_{\gamma,n}(r).$$

Set, for $k = 1, \ldots, a$, and $i = 1, \ldots, m$,

$$a_{k,m}(i) = a_n(s_{i,k}) - \bar{a}_{k,m}, \qquad c_{\gamma,k,m}(i) = c_{\gamma,n}(s_{i,k}) - \bar{c}_{\gamma,k,m}$$

and

$$f_{k,s}(i) \text{ to be the index } j \text{ such that } s_{j,k} = f_r(s_{i,k}).$$

Further, let

$$C_{\gamma,k,m}^{(s)}(t) = \sum_{i \leq mt} c_{\gamma,k,m}(f_{k,s}(i)), \qquad 0 \leq t \leq 1.$$

Using this notation we can write, for each $k = 1, \ldots, a$,

$$T_{\gamma,k,n}(s) = \frac{\sqrt{n} \sum_{i=1}^m a_{k,m}(i) c_{\gamma,k,m}(f_{k,s}(i))}{\sigma_n(a)}.$$

For each $\gamma > 0$ and $k = 1, \ldots, a$, let

$$\sigma_{k,m}^2(a) = \sum_{i=1}^m a_{k,m}^2(i) \quad \text{and} \quad \sigma_{\gamma,k,m}^2(c) = \sum_{i=1}^m c_{\gamma,k,m}^2(i).$$

The following proposition is crucial to our proof.

Proposition 4.1. *With the above assumptions and notation, for each $\gamma > 0$, $n \geq 3$, $1 \leq r \leq n-1$ and $k = 1, \ldots, a$, there exist on a rich enough probability space versions $\tilde{T}_{\gamma,k,n}(s)$ of $T_{\gamma,k,n}(s)$, that is, $\tilde{T}_{\gamma,k,n}(s) \overset{\mathscr{L}}{=} T_{\gamma,k,n}(s)$, jointly in $k = 1, \ldots, a$, and a standard normal random variable Z such that*

$$\tilde{V}_{\gamma,n}(r) := \sum_{k=1}^a \tilde{T}_{\gamma,k,n}(s) = s_n(\gamma) Z + o_p(1), \tag{4.1}$$

where

$$s_n^2(\gamma) = \sum_{k=1}^a a \sigma_{k,m}^2(a) \sigma_{\gamma,k,m}^2(c) / \sigma_n^2(a). \tag{4.2}$$

Proof. For any $\gamma > 0$ and $k = 1, \ldots, a$, we see from their definitions

$$\sum_{i=1}^m c_{\gamma,k,m}(i) = 0, \qquad \sum_{i=1}^m c_{\gamma,k,m}^2(i) \leq \gamma, \qquad \sum_{i=1}^m c_{\gamma,k,m}^4(i) \leq m\gamma^2/n^2.$$

Thus conditioned on $\mathscr{V}_k$, the process $C_{\gamma,k,m}^{(s)}$ satisfies the assumptions of Theorem 1.1. Hence conditioned on $\mathscr{V}_k$, for every $k = 1, \ldots, a$, there exist versions $\tilde{C}_{\gamma,k,m}^{(s)}$ of $C_{\gamma,k,m}^{(s)}$ and a standard Brownian bridge B such that for all $0 \leq \nu < \frac{1}{4}$

$$\sup_{1/m \leq t \leq 1-1/m} \frac{m^\nu |\tilde{C}_{\gamma,k,m}^{(s)}(t) - \sigma_{\gamma,k,m}(c) B(t)|}{(t(1-t))^{1/2-\nu}} =: \delta_{k,m}, \tag{4.3}$$

where $\delta_{k,m}$ is a random variable such that for all for $x > 0$, uniformly in $n \geq 3$ and $1 \leq k \leq a$,

$$P(\delta_{k,m} > x | \mathscr{V}_k) \leq \rho(x), \tag{4.4}$$

for some function ρ satisfying $\rho(x) \to 0$ as $x \to \infty$. (See Remark 3.1.)

Next let $A_{m,k}$ (or $|A_{m,k}|$) be the random atomic measure that places mass

$$\frac{a_{k,m}(i) - a_{k,m}(i+1)}{\sigma_n(a)} \quad \left(\text{or} \quad \frac{|a_{k,m}(i) - a_{k,m}(i+1)|}{\sigma_n(a)} \right)$$

on each i/m, for $i = 1, \ldots, m$, where $a_{k,m}(m+1) := 0$. Define each $\tilde{T}_{\gamma,k,n}(s)$, $k = 1, \ldots, a$, to be the integral

$$\tilde{T}_{\gamma,k,n}(s) = \sqrt{n} \int_{1/m}^{1-1/m} \tilde{C}_{\gamma,k,m}^{(s)}(t)\, \mathrm{d}A_{m,k}(t),$$

which one readily checks to have the same distribution as $T_{\gamma,k,n}(s)$. Introduce the random variables

$$Z_{\gamma,k,n} = \sqrt{n}\,\sigma_{\gamma,k,m}(c) \int_{1/m}^{1-1/m} B(t)\, \mathrm{d}A_{m,k}(t).$$

Now by (4.3) we have, for any $0 \leqslant \nu < \frac{1}{4}$,

$$|\tilde{T}_{\gamma,k,n}(s) - Z_{\gamma,k,n}| \leqslant a^\nu n^{1/2-\nu} \int_{1/m}^{1-1/m} (t(1-t))^{1/2-\nu}\, \mathrm{d}|A_{m,k}|(t)\delta_{k,m},$$

$$\leqslant D r_{n,k}(\nu)\delta_{k,m},$$

for some constant $D > 0$, where

$$r_{n,k}(\nu) = m^{1/2-\nu} \sum_{i=1}^{m} \left(\frac{i}{m}\left(1 - \frac{i}{m}\right) \right)^{1/2-\nu} |a_{k,m}(i) - a_{k,m}(i+1)|/\sigma_n(a).$$

Notice that

$$a_{k,m}(i) = J_1\left(\frac{s_{i,k}}{n+1} \right) - J_2\left(\frac{s_{i,k}}{n+1} \right), \qquad i = 1, \ldots, m.$$

Thus by setting

$$d_n(i) = \left| J_1\left(\frac{i}{n+1} \right) \right| + \left| J_2\left(\frac{i}{n+1} \right) \right|, \qquad i = 1, \ldots, n,$$

we conclude by using the monotonicity of J_1 and J_2 that, both conditioned on $\mathscr{V}_k$ and unconditionally, we obtain after summation by parts, that for some $D_1 > 0$

$$r_{n,k}(\nu) \leqslant D_1 \sum_{i=1}^{[n/2]+1} i^{-1/2-\nu}(d_n(i) + d_n(n+1-i))/\sigma_n(a) =: r_n.$$

Since by (1.13), for each $j \geqslant 1$,

$$\max_{1 \leqslant i \leqslant j} |d_n(i) + d_n(n+1-i)|/\sigma_n(a) \to 0,$$

we see that, for each fixed $j \geqslant 1$,

$$r_n = o(1) + D_1 \sum_{i=j+1}^{[n/2]+1} i^{-1/2-\nu}(d_n(i) + d_n(n+1-i))/\sigma_n(a)$$

$$\leqslant o(1) + D_2 j^{-\nu}$$

by using the Cauchy–Schwarz inequality and (1.12), for some $D_2 > 0$. Noting that $j \geqslant 1$ can be made arbitrarily large, we conclude

$$r_n \to 0 \qquad \text{as } n \to \infty. \tag{4.5}$$

Thus we obtain, conditioned on $\mathscr{V}_k$, that

$$|\tilde{T}_{\gamma,k,n}(s) - Z_{\gamma,k,n}| \leqslant Dr_n \delta_{k,m}, \tag{4.6}$$

where r_n satisfies (4.5). Notice that, conditioned on $\mathscr{V}_k$, $Z_{\gamma,k,n}$ is a mean-zero normal random variable with

$$n^{-1}\,\mathrm{var}(Z_{\gamma,k,n}|\mathscr{V}_k)$$

$$= \sigma_n^{-2}(a)\sigma_{\gamma,k,m}^2(c) \sum_{i=1}^{m}\sum_{j=1}^{m}\left(\frac{i \wedge j}{m} - \frac{ij}{m^2}\right)(a_{k,m}(i) - a_{k,m}(i+1))(a_{k,m}(j) - a_{k,m}(j+1)).$$

At this point we need the following easily verified fact:

Fact 1. *Choose $m \geqslant 2$ and let $d(1), \ldots, d(m+1)$ be constants such that $d(m+1) = 0$. Then*

$$\sum_{i=1}^{m}\sum_{j=1}^{m}\left(\frac{i \wedge j}{m} - \frac{ij}{m^2}\right)(d(i) - d(i+1))(d(j) - d(j+1)) = \sum_{i=1}^{m}\frac{d^2(i)}{m} - (\overline{d}_m)^2,$$

where $\overline{d}_m = m^{-1}\sum_{i=1}^{m}d(i)$.

From this fact we get that

$$\mathrm{var}(Z_{\gamma,k,n}|\mathscr{V}_k) = \frac{a\sigma_{k,m}^2(a)\sigma_{\gamma,k,m}^2(c)}{\sigma_n^2(a)} =: s_{k,m}^2(\gamma).$$

Thus, conditioned on $\mathscr{V}_k$, we can write each

$$Z_{\gamma,k,n} =: s_{k,m}(\gamma)Z_k,$$

where Z_k is a standard normal random variable. Clearly, since the $T_{\gamma,k,n}(s)$, $k = 1, \ldots, a$, are independent given $\mathscr{V}_1, \ldots, \mathscr{V}_a$, the $Z_1, \ldots, Z_a$ can be constructed to be independent conditioned on $\mathscr{V}_1, \ldots, \mathscr{V}_a$, so that, conditioned on $\mathscr{V}_1, \ldots, \mathscr{V}_a$,

$$Z := s_n^{-1}(\gamma) \sum_{k=1}^{a} s_{k,m}(\gamma)Z_k,$$

with $s_n(\gamma)$ as in (4.2), is a standard normal random variable. Putting everything together, we have constructed $\tilde{T}_{\gamma,k,n}(s)$, $k = 1, \ldots, a$, and Z, conditionally on $\mathscr{V}_1, \ldots, \mathscr{V}_a$, such that

$$\tilde{V}_{\gamma,n}(r) = \sum_{k=1}^{a} \tilde{T}_{\gamma,k,n}(s) = s_n(\gamma)Z + Dr_n\delta_m, \tag{4.7}$$

where $\delta_m = \sum_{k=1}^{a}\delta_{k,m}$. However, since there are only finitely many possible $\mathscr{V}_1, \ldots, \mathscr{V}_a$ for any fixed n and r, without any measurability problems, we can construct these random variables so that (4.7) holds unconditionally. Finally, notice by (4.4) that $\delta_m = O_P(1)$, which, when combined with (4.5), completes the proof of Proposition 4.1. $\qquad\square$

Lemma A1 of Berkes and Phillipp (1979) allows the appropriate probability spaces on which $(s_n(\gamma)Z, \tilde{V}_{\gamma,n}(r))$ and $(V_{\gamma,n}(r), W^t_{\gamma,n}(r) + \Delta_{\gamma,n}(r))$ sit to be coupled to form the triple

$$(s_n(\gamma)\bar{Z}, \bar{V}_{\gamma,n}(r), \bar{W}^t_{\gamma,n}(r) + \bar{\Delta}_{\gamma,n}(r)),$$

where

$$(s_n(\gamma)Z, \tilde{V}_{\gamma,n}(r)) \overset{\mathscr{L}}{=} (s_n(\gamma)\bar{Z}, \bar{V}_{\gamma,n}(r))$$

and

$$(\bar{V}_{\gamma,n}(r), \bar{W}^t_{\gamma,n}(r) + \bar{\Delta}_{\gamma,n}(r)) \overset{\mathscr{L}}{=} (V_{\gamma,n}(r), W^t_{\gamma,n}(r) + \Delta_{\gamma,n}(r)).$$

Thus

$$W_n(r) \overset{\mathscr{L}}{=} \bar{V}_{\gamma,n}(r) + \bar{W}^t_{\gamma,n}(r) + \bar{\Delta}_{\gamma,n}(r). \tag{4.8}$$

Clearly now, on account of (4.8) and (4.1), using the fact that $\bar{V}_{\gamma,n}(r)$ and $\tilde{V}_{\gamma,n}(r)$ are equal in distribution, the proof of Corollary 1.1 will be complete as soon as we have shown that

$$\lim_{\gamma \to \infty} \limsup_{n \to \infty} \mathrm{E}(W^t_{\gamma,n}(r))^2 = 0, \tag{4.9}$$

for each $\gamma > 0$,

$$\Delta_{\gamma,n}(r) = o_p(1), \tag{4.10}$$

and, for each $\gamma > 0$ and for all $\varepsilon > 0$,

$$\lim_{\gamma \to \infty} \limsup_{n \to \infty} P\{|1 - s_n^2(\gamma)| > \varepsilon\} = 0. \tag{4.11}$$

First consider (4.9). Note that, by Lemma 3.2,

$$\mathrm{E}(W^t_{\gamma,n}(r))^2 \leqslant 2 \sum_{j=1}^{n} \mathbf{1}\{b_n^2(j)/\sigma_n^2(b) > \gamma/n\} b_n^2(j)/\sigma_n^2(b).$$

Therefore by assumption (1.15), we have (4.9).

Next we shall show (4.10). For this we need the following elementary finite-sampling fact:

Fact 2. *Choose $n \geqslant 2$, $1 \leqslant m \leqslant n$ and constants $c(1), \ldots, c(n)$ such that $\sum_{i=1}^{n} c(i) = 0$. Further, let V_m be a random subset of size m chosen from $\{1, \ldots, n\}$ with probability $1/\binom{n}{m}$.*

Then

$$\mathrm{E}\left(\sum_{i \in V_m} c(i)\right)^2 = \frac{m}{n}\left(1 - \frac{m-1}{n-1}\right)\sum_{i=1}^{n} c^2(i).$$

Applying Fact 2, we get that, for each $k = 1, \ldots, a$,

$$\mathrm{E}\bar{c}_{\gamma,k,m}^2 \leqslant m^{-2}\gamma \quad \text{and} \quad \mathrm{E}(\bar{a}_{k,m}^2) \leqslant m^{-2}\sigma_n^2(a),$$

from which we obtain

$$\bar{c}_{\gamma,k,m} = O_p(m^{-1}) \quad \text{and} \quad \bar{a}_{k,m}/\sigma_n(a) = O_p(m^{-1}). \tag{4.12}$$

Thus for each $k = 1, \ldots, a$,

$$\frac{\sqrt{n}\,m\bar{a}_{k,m}\bar{c}_{\gamma,k,m}}{\sigma_n(a)} = O_p(m^{-1/2}), \tag{4.13}$$

which implies (4.10).

Finally we turn to (4.11). Notice that each

$$\sigma_{k,m}^2(a) = \sum_{i \in \mathscr{V}_k} a_n^2(i) - m\bar{a}_{k,m}^2.$$

From (4.12) we have, uniformly in $1 \leqslant k \leqslant a$,

$$m\bar{a}_{k,m}^2/\sigma_n^2(a) = o_p(1).$$

This immediately yields that, uniformly in $1 \leqslant k \leqslant a$,

$$\sigma_{k,m}^2(a)/\sigma_n^2(a) = \sum_{i \in \mathscr{V}_k} a_n^2(i)/\sigma_n^2(a) + o_p(1).$$

In the same way, it can be argued that, uniformly in $1 \leqslant k \leqslant a$,

$$\sigma_{\gamma,k,m}^2(c)/\sigma_n^2(c) = \sum_{i \in \mathscr{V}_k} c_{\gamma,n}^2(i)/\sigma_n^2(c) + o_p(1).$$

Notice that

$$\mathrm{E}\left[\sum_{i \in \mathscr{V}_k} a_n^2(i)\right] = \frac{m\sigma_n^2(a)}{n}$$

and

$$\mathrm{E}\left[\sum_{i \in \mathscr{V}_k} c_{\gamma,n}^2(i)\right] = \frac{m}{n}\sum_{i=1}^{n} c_{\gamma,n}^2(i) =: \frac{m}{n}\sigma_{\gamma,n}^2(c).$$

Applying Fact 2, we obtain that

$$\mathrm{E}\left[\sum_{i\in\mathscr{T}_k} a_n^2(i) - \frac{m\sigma_n^2(a)}{n}\right]^2 \leqslant \sum_{i=1}^n \left[a_n^2(i) - \frac{\sigma_n^2(a)}{n}\right]^2.$$

Now, using (1.13), it is easy to show that

$$\max_{1\leqslant i\leqslant n} \{a_n^2(i)/\sigma_n^2(a)\} = o(1).$$

Thus one sees by using Chebyshev's inequality that, uniformly in $1 \leqslant k \leqslant a$,

$$\sum_{i\in\mathscr{T}_k} a_n^2(i)/\sigma_n^2(a) - \frac{m}{n} = o_p(1).$$

Similarly, one proves that, uniformly in $1 \leqslant k \leqslant a$,

$$\sum_{i\in\mathscr{T}_k} c_{\gamma,n}^2(i)/\sigma_n^2(c) - m\sigma_{\gamma,n}^2(c)/(n\sigma_n^2(c)) = o_p(1).$$

Using assumption (1.15), it is now straightforward to verify that (4.11) holds. This completes the proof of Corollary 1.1.

4.2. Proof of Corollary 1.2

To show that (1.16) remains true under assumption (1.14), note that, given any φ as in (1.14) and $\varepsilon > 0$, we can find a polynomial ϕ on $(0, 1)$ such that

$$\int_0^1 (\varphi(u) - \phi(u))^2 \, \mathrm{d}u < \varepsilon. \tag{4.14}$$

The argument just given shows that asymptotic normality holds for the statistic $\hat{W}_n(r)$ formed when the $a_n(i)$, $i = 1, \ldots, n$, in $W_n(r)$ are replaced by

$$\phi(i/(n+1)), \qquad i = 1, \ldots, n,$$

since any polynomial ϕ satisfies (1.12) and (1.13). Next, by applying Lemma 3.2 in combination with (1.14), we readily check that we can make the mean squared difference between the statistics $\hat{W}_n(r)$ and $W_n(r)$ arbitrarily close, as $n \to \infty$, by choosing $\varepsilon > 0$ small enough in (4.14) and ϕ appropriately. This finishes the proof of Corollary 1.2.

Acknowledgements

References

Berkes, I. and Philipp, W. (1979) Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.*, **7**, 29–54.

Csörgő, M. and Horváth, L. (1993) *Weighted Approximations in Probability and Statistics*. Chichester: Wiley.

Einmahl, U. and Mason, D.M. (1992) Approximations to permutation and exchangeable processes. *J. Theoret. Probab.*, **5**, 101–126.

Haeusler, E. and Mason, D.M. (1999) Weighted approximations to continuous time martingales with applications. *Scand. J. Statist.*, **26**, 281–295.

Hahn, M., Mason, D.M. and Weiner, D. (eds) (1991) *Sums, Trimmed Sums and Extremes*. Boston: Birkhäuser.

Hájek, J. and Šidák, Z. (1967) *Theory of Rank Tests*. New York: Academic Press.

Hall, P. and Heyde, C.C. (1980) *Martingale Limit Theory and Its Application.* New York: Academic Press.

Hallin, M. and Puri, M.L. (1988) Optimal rank-based procedures for time series analysis: testing an ARMA model against other ARMA models. *Ann. Statist.*, **16,** 402–432.

Hallin, M. and Puri, M.L. (1991) Time series analysis via rank order theory: signed-rank tests for ARMA models. *J. Multivariate Anal.*, **39**, 1–29.

Hallin, M. and Puri, M.L. (1992a) Rank tests for time series analysis: a survey. In D. Brillinger, P. Caines, J. Geweke, E. Parzen, M. Rosenblatt and M.S. Taqqu (eds), *New Directions in Time Series Analysis, Part I*, IMA Vol. Math. Appl. 45, pp. 111–153. New York: Springer-Verlag.

Hallin, M. and Puri, M.L. (1992b) Some asymptotic results for a broad class of nonparametric statistics. *J. Statist. Plann. Inference*, **32**, 165–196.

Hallin, M. and Puri, M.L. (1994) Aligned rank tests for linear models with autocorrelated error terms. *J. Multivariate Anal.*, **50**, 175–237.

Hallin, M. and Puri, M.L. (1995) A multivariate Wald–Wolfowitz rank test against serial dependence. *Canad. J. Statist.*, **23**, 55–65.

Hallin, M. and Vermandele, C. (1996) An elementary proof of asymptotic normality for serial rank statistics. In E. Brunner and M. Denker (eds), *Research Developments in Probability and Statistics*, pp. 163–191. Utrecht: VSP.

Hallin, M., Mélard, G. and Milhaud, X. (1992) Permutational extreme values of autocorrelation coefficients and a Pitman test against serial dependence. *Ann. Statist.*, **20**, 523–534.

Kolchin, V.F. and Chistyakov, V.P. (1974) Combinatorial problems of probability theory. *J. Soviet Math.*, **4**, 217–243.

Lombard, F. (1986) An elementary proof of asymptotic normality for linear rank statistics. *South African Statist. J.*, **20**, 29–35.

Scott, D.J. and Huggins, R.M. (1983) On the embedding of processes in Brownian motion and the law of the iterated logarithm for reverse martingales. *Bull. Austral. Math. Soc.*, **27**, 443–459.

Shorack, G. (1991) Embedding the sampling process at a rate. *Ann. Probab.*, **19**, 826–842.

Shorack, G. and Wellner, J. (1986) *Empirical Processes with Applications to Statistics*. New York: Wiley.

Wald, A. and Wolfowitz, J. (1943) An exact test for randomness in the nonparametric case based on serial correlation. *Ann. Math. Statist.*, **14**, 378–388.