The Annals of Statistics
1997, Vol. 25, No. 5, 2210–2227

ERROR BOUND IN A CENTRAL LIMIT THEOREM OF DOUBLE-INDEXED PERMUTATION STATISTICS

BY LINCHENG ZHAO[1], ZHIDONG BAI[2],
CHERN-CHING CHAO[2] AND WEN-QI LIANG[2]

University of Science and Technology of China,
National Sun Yat-Sen University,
Academia Sinica and Academia Sinica

An error bound in the normal approximation to the distribution of the double-indexed permutation statistics is derived. The derivation is based on Stein's method and on an extension of a combinatorial method of Bolthausen. The result can be applied to obtain the convergence rate of order $n^{-1/2}$ for some rank-related statistics, such as Kendall's tau, Spearman's rho and the Mann–Whitney–Wilcoxon statistic. Its applications to graph-related nonparametric statistics of multivariate observations are also mentioned.

1. Introduction. Let $\zeta(i, j, k, l)$, $i, j, k, l \in N = \{1, \dots, n\}$, be real numbers depending on n. We are interested in the double-indexed permutation statistics (DIPS) of the general form $\sum_{i,j} \zeta(i, j, \pi(i), \pi(j))$, where π is uniformly distributed on the set $\mathscr{P}_n$ of all permutations of N. The DIPS of the restricted form $\sum_{i,j} a_{ij} b_{\pi(i)\pi(j)}$ was first investigated by Daniels (1944) in the study of a generalized correlation coefficient with Kendall's tau and Spearman's rho being special cases. Daniels gave a set of sufficient conditions for their asymptotic normality as $n \to \infty$. Further investigations along this direction have been done by Bloemena (1964), Jogdeo (1968), Abe (1969), Shapiro and Hubert (1979), Barbour and Eagleson (1986) and Pham, Möcks and Sroka (1989). In these contexts, the so-called scores a_{ij} and b_{ij} are either symmetric ($a_{ij} = a_{ji}, b_{ij} = b_{ji}$) or skew-symmetric ($a_{ij} = -a_{ji}, b_{ij} = -b_{ji}$). The uses of DIPS have diversely been suggested by Friedman and Rafsky (1979, 1983) and Schilling (1986) in multivariate nonparametric tests, by Hubert and Schultz (1976) in clustering studies, by Mantel and Valand (1970) in biometry, and by Cliff and Ord (1981) in geography.

The purpose of this paper is to derive a bound for the error in the normal approximation to the distribution of the DIPS of the general form

Received October 1995; revised December 1996.

[1] Research partially supported by National Natural Science Foundation of China, Ph.D. Program Foundation of National Education Committee of China and Special Foundation of Academia Sinica.

[2] Research partially supported by National Science Council of the Republic of China.

AMS 1991 *subject classifications*. Primary 60F05, 62E20; secondary 62H20.

Key words and phrases. Asymptotic normality, correlation coefficient, graph theory, multivariate association, permutation statistics, Stein's method.

$\sum_{i,j}\zeta(i, j, \pi(i), \pi(j))$. This bound can be used to yield the convergence rate $O(n^{-1/2})$ for some well-known statistics, such as Kendall's tau, Spearman's rho and the Mann–Whitney–Wilcoxon statistic. In Section 2 the DIPS, $\sum_{i,j}\zeta(i, j, \pi(i), \pi(j))$, is converted to the form of $\sum_i a(i, \pi(i)) + n^{-1}\sum'_{i,j} b(i, j, \pi(i), \pi(j))$. (Throughout this paper, $\sum'_{i,j}$ denotes $\sum_{i,j,i\neq j}$.) A Berry–Esseen type of inequality for the latter is stated as Theorem 1. The result for DIPS, straightforwardly implied by Theorem 1, is stated as Theorem 2. In Section 3 the applications of Theorem 2 to Daniels' generalized correlation coefficient, the number of edges in the random intersection of two graphs and the Mann–Whitney–Wilcoxon statistic are demonstrated. The essential theoretic part of this paper, that is, the proof of Theorem 1, is presented in Section 4. Our derivations are based on Stein's method (1972) and an extension of the combinatorial method of Bolthausen (1984). Bolthausen successfully employed his combinatorial method combined with Stein's method to obtain a result on the convergence rate for the single-indexed permutation statistics of the form $\sum_i a(i, \pi(i))$. Our Theorem 1 reduces to Bolthausen's result when all $b(i, j, k, l) = 0$. These two methods were also used by Schneller (1989) to establish the Edgeworth expansion for general linear rank statistics.

There has been little success in establishing the Berry–Esseen bound of order $n^{-1/2}$ for general classes of statistics which are asymptotically normally distributed. For the importance of and the historic developments in the study of departures from normality, the reader is referred to an earlier survey paper by Bickel (1974). The possibility of applying Stein's method in such investigations is also pointed out therein.

2. Main results.

For each 4-tuple real array $(x(i, j, k, l))$ and each real matrix $(y(i, k))$, $i, j, k, l \in N$, the following notation is used:

$$x(i, j, k, \cdot) = n^{-1}\sum_l x(i, j, k, l), \qquad x(i, j, \cdot, \cdot) = n^{-2}\sum_{k,l} x(i, j, k, l),$$

$$x(i, \cdot, \cdot, \cdot) = n^{-3}\sum_{j,k,l} x(i, j, k, l), \qquad x(\cdot, \cdot, \cdot, \cdot) = n^{-4}\sum_{i,j,k,l} x(i, j, k, l);$$

$$y(i, \cdot) = n^{-1}\sum_k y(i, k), \qquad y(\cdot, \cdot) = n^{-2}\sum_{i,k} y(i, k),$$

and others defined similarly.

Let $A = (a(i, k))$, $i, k \in N$, be a given real matrix such that

$$a(i, \cdot) = a(\cdot, k) = 0, \tag{2.1}$$

$$\sum_{i,k} a^2(i, k) = n - 1. \tag{2.2}$$

Let $B = (b(i, j, k, l))$, $i, j, k, l \in N$, be a given 4-tuple real array such that

$$b(i, j, k, \cdot) = b(i, j, \cdot, l) = b(i, \cdot, k, l) = b(\cdot, j, k, l) = 0. \tag{2.3}$$

Consider the random variable

$$W = \sum_i a(i, \pi(i)) + n^{-1} \sum_{i,j}{}' b(i, j, \pi(i), \pi(j)),$$

where π is uniformly distributed on $\mathscr{P}_n$. We have the following result.

THEOREM 1. *There is an absolute constant $K > 0$ such that, for $n \geq 2$,*

$$\sup_x |P(W \leq x) - \Phi(x)| \leq K\left\{n^{-1} \sum_{i,k} |a(i,k)|^3 + n^{-3} \sum_{i,j,k,l} |b(i,j,k,l)|^3\right\},$$

where Φ is the standard normal distribution function.

Now, consider the asymptotic normality of the DIPS

$$D = \sum_{i,j} \zeta(i, j, \pi(i), \pi(j)).$$

For given $(\zeta(i,j,k,l))$, $i, j, k, l \in N$, let

$$\begin{aligned}
\zeta^*(i,j,k,l) = \zeta(i,j,k,l) &- [\zeta(i,j,k,\cdot) + \zeta(i,j,\cdot,l) \\
&\qquad + \zeta(i,\cdot,k,l) + \zeta(\cdot,j,k,l)] \\
&+ [\zeta(i,j,\cdot,\cdot) + \zeta(i,\cdot,k,\cdot) + \zeta(i,\cdot,\cdot,l) \\
&\qquad + \zeta(\cdot,j,k,\cdot) + \zeta(\cdot,j,\cdot,l) + \zeta(\cdot,\cdot,k,l)] \\
&- [\zeta(i,\cdot,\cdot,\cdot) + \zeta(\cdot,j,\cdot,\cdot) + \zeta(\cdot,\cdot,k,\cdot) + \zeta(\cdot,\cdot,\cdot,l)] \\
&+ \zeta(\cdot,\cdot,\cdot,\cdot).
\end{aligned}$$

Then

$$\zeta^*(i,j,k,\cdot) = \zeta^*(i,k,\cdot,l) = \zeta^*(i,\cdot,k,l) = \zeta^*(\cdot,j,k,l) = 0$$

and

$$\begin{aligned}
D &= \sum_{i,j} \zeta^*(i,j,\pi(i),\pi(j)) + n \sum_i \zeta(i,\cdot,\pi(i),\cdot) \\
&\quad + n \sum_j \zeta(\cdot,j,\cdot,\pi(j)) - n^2 \zeta(\cdot,\cdot,\cdot,\cdot) \\
&= \sum_{i,j}{}' \zeta^*(i,j,\pi(i),\pi(j)) + \sum_i a(i,\pi(i)) \\
&= \sum_{i,j}{}' \zeta^*(i,j,\pi(i),\pi(j)) + \sum_i a^*(i,\pi(i)) + n a(\cdot,\cdot),
\end{aligned}$$

where

$$a(i,k) = \zeta^*(i,i,k,k) + n\zeta(i,\cdot,k,\cdot) + n\zeta(\cdot,i,\cdot,k) - n\zeta(\cdot,\cdot,\cdot,\cdot)$$

and

$$a^*(i,k) = a(i,k) - a(i,\cdot) - a(\cdot,k) + a(\cdot,\cdot).$$

Note that $a^*(i,\cdot) = a^*(\cdot,k) = 0$. Defining and assuming that

$$\sigma^2 = \sum_{i,k} a^{*2}(i,k)/(n-1) > 0, \tag{2.4}$$

we have

$$\frac{D - na(\cdot,\cdot)}{\sigma} = \sum_i \frac{1}{\sigma} a^*(i, \pi(i)) + n^{-1} \sum_{i,j}{}' \frac{n}{\sigma} \zeta^*(i, j, \pi(i), \pi(j)).$$

Thus, we can apply Theorem 1 to obtain the following result.

THEOREM 2. *There is an absolute constant $K > 0$ such that, for $n \geq 2$,*

$$\sup_x \left| P\left(\frac{D - na(\cdot,\cdot)}{\sigma} \leq x \right) - \Phi(x) \right| \leq \frac{K}{\sigma^3} \left\{ n^{-1} \sum_{i,k} |a^*(i,k)|^3 + \sum_{i,j,k,l} |\zeta^*(i,j,k,l)|^3 \right\},$$

provided condition (2.4) *is satisfied.*

3. Applications. In this section the applications of Theorem 2 are demonstrated by three examples. In addition to those well-known testing statistics stated below, Theorem 2 reveals the potential for creating new nonparametric testing statistics, especially for multivariate observations, due to its generality.

EXAMPLE 1 (Mann–Whitney–Wilcoxon statistic). Let $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$, $n_1 + n_2 = n$, be independent univariate random samples from unknown continuous distributions F_X and F_Y, respectively. The Mann–Whitney–Wilcoxon statistic for testing the hypothesis H_0: $F_X = F_Y$ is defined to be the total number of pairs (x_i, y_j) for which $x_i < y_j$. Let $\pi(i)$, $i = 1, \ldots, n_1$, denote the rank of x_i and $\pi(n_1 + j)$, $j = 1, \ldots, n_2$, denote that of y_j in the combined sample. Then the Mann–Whitney–Wilcoxon statistic can be expressed as $\sum_{i,j} \zeta(i, j, \pi(i), \pi(j))$, where

$$\zeta(i,j,k,l) = \begin{cases} 1, & \text{if } 1 \leq i \leq n_1,\ n_1 + 1 \leq j \leq n \text{ and } 1 \leq k < l \leq n, \\ 0, & \text{otherwise;} \end{cases}$$

and π is uniformly distributed on $\mathscr{P}_n$ under H_0. Applying Theorem 2 with straightforward calculations, we obtain

$$\sup_x \left| P\left(\frac{\sum_{i,j} \zeta(i,j,\pi(i),\pi(j)) - \frac{1}{2} n_1 n_2}{\left(\frac{1}{12} n_1 n_2 (n+1)\right)^{1/2}} \leq x \right) - \Phi(x) \right| \leq K\left(n_1^{-1} + n_2^{-1}\right)^{1/2}.$$

The Mann–Whitney–Wilcoxon statistic is one of the members of U-statistics of degree two. The Berry–Esseen bounds and the Edgeworth expansions for U-statistics have been extensively studied; see Bickel, Götze and van Zwet (1986) and the references therein. For a systematic presentation of the theory of U-statistics, the reader is referred to Koroljuk and Borovskich (1994).

EXAMPLE 2 (Daniels' generalized correlation coefficient). Let $(d(i,j))$ and $(e(i,j))$, $i, j \in N$, be two real matrices. Daniels (1944) considers a generalized correlation coefficient $\sum_{i,j} d(i,j) e(\pi(i), \pi(j))$, where the scores $d(i,j)$ and $e(i,j)$ are skew-symmetric and π is uniformly distributed on $\mathscr{P}_n$. Applying

Theorem 2, we have

$$\begin{aligned}\sup_x \left| P\left(\frac{\sum_{i,j} d(i,j) e(\pi(i), \pi(j))}{\sigma} \le x \right) - \Phi(x) \right| \\ \le \frac{K}{\sigma^3} \Bigg\{ n^2 \sum_{i,k} |d(i,\cdot) e(k,\cdot)|^3 \\ + \sum_{i,j,k,l} |(d(i,j) - d(i,\cdot) - d(\cdot,j)) \\ \times (e(k,l) - e(k,\cdot) - e(\cdot,l))|^3 \Bigg\},\end{aligned} \tag{3.1}$$

where $\sigma^2 = 4n^2(n-1)^{-1}\sum_{i,k} d^2(i,\cdot)e^2(k,\cdot)$.

Consider ordered pairs of univariate observations (x_i, y_i), $i \in N$. Kendall's tau (letting $d(i,j) = \text{sign}(x_i - x_j)$ and $e(i,j) = \text{sign}(y_i - y_j)$) and Spearman's rho ($d(i,j) = \text{rank}(x_i) - \text{rank}(x_j)$ and $e(i,j) = \text{rank}(y_i) - \text{rank}(y_j)$) are two statistics for testing the hypothesis H_0: no correlation between X and Y. Applying (3.1), we conclude that the null distribution of both (standardized) statistics converges to $\Phi(x)$ with the rate $O(n^{-1/2})$.

EXAMPLE 3 (Number of edges in the random intersection of two graphs). Friedman and Rafsky (1983) extend the notion of association measures for univariate observations, such as Kendall's tau, to multivariate observations. The lack of ordering in multivariate observations is conquered by constructing interpoint-distance based graphs, such as the k minimal spanning tree and the k nearest-neighbor graph. Then, various measures for association or others can be defined in terms of the number of edges in the intersection of two graphs. The reader is referred to Friedman and Rafsky (1979, 1983) for details.

Now, let $G_1(N, E_1)$ and $G_2(N, E_2)$ be two graphs consisting of the same set of nodes, $N = \{1, \dots, n\}$, and sets of edges E_1 and E_2, respectively. The number of edges in the random intersection of G_1 and G_2 is defined as $\Gamma = \sum_{i,j} I_{\{(i,j)\in E_1\}} I_{\{(\pi(i),\pi(j))\in E_2\}}$. Here, I_D denotes the indicator of the set D. Let d_i denote the degree of node i in G_1, that is, the number of edges in E_1 that are incident to i. Let ρ_1 denote the total number of edges in E_1. Then $\rho_1 = \frac{1}{2}\sum_i d_i$. Similarly, define the degree d'_i of node i in G_2 and the total number ρ_2 of edges in E_2. Applying Theorem 2, we obtain

$$\begin{aligned}\sup_x \left| P\left(\frac{\Gamma - 4n^{-2}(1+n^{-1})\rho_1\rho_2}{\sigma} \le x \right) - \Phi(x) \right| \\ \le \frac{K}{\sigma^3} \Bigg\{ n^{-4} \sum_{i,k} \left| \left(d_i - 2n^{-1}\rho_1\right)\left(d'_k - 2n^{-1}\rho_2\right) \right|^3 \\ + \sum_{i,j,k,l} \Big| \left(I_{\{(i,j)\in E_1\}} - n^{-1}(d_i + d_j) + 2n^{-2}\rho_1 \right) \\ \times \left(I_{\{(k,l)\in E_2\}} - n^{-1}(d'_k + d'_l) + 2n^{-2}\rho_2 \right) \Big|^3 \Bigg\},\end{aligned}$$

where $\sigma^2 = 4n^{-3}\sum_{i,k}(d_i - 2n^{-1}\rho_1)^2(d'_k - 2n^{-1}\rho_2)^2$. Note that if the degrees d_i and d'_i of each node grow linearly with n, then the convergence rate reaches $O(n^{-1/2})$.

4. Proof of Theorem 1. In order to create sufficient independence needed in the proof of Theorem 1, we first extend Bolthausen's combinatorial method as follows.

Define a random vector $(I_1, J_1, I_2, J_2, K_1, L_1, K_2, L_2)$ in N^8 in the following way: first, let (I_1, J_1), (I_2, J_2) and (K_1, L_1) be independent and identically distributed with

$$P(I_1 = i, J_1 = j) = \begin{cases} 0, & \text{if } i = j, \\ \dfrac{1}{n(n-1)}, & \text{if } i \neq j. \end{cases}$$

Given these $I_1 \neq J_1$, $I_2 \neq J_2$ and $K_1 \neq L_1$, (K_2, L_2) and its conditional distribution are defined according to the following rules:

1. If $I_1 = I_2$ and $J_1 = J_2$, then $K_2 = K_1$ and $L_2 = L_1$.
2. If $I_1 = I_2$ and $J_1 \neq J_2$, then $K_2 = K_1$ and L_2 is uniformly distributed on $N - \{K_1, L_1\}$.
3. If $I_1 \neq I_2$ and $J_1 = J_2$, then $L_2 = L_1$ and K_2 is uniformly distributed on $N - \{K_1, L_1\}$.
4. If $I_1 \neq I_2$ and $J_1 \neq J_2$, then $K_2 \neq K_1$, $L_2 \neq L_1$, $K_2 \neq L_2$, and furthermore, (a) if $I_1 = J_2$ and $I_2 = J_1$, then $K_2 = L_1$ and $L_2 = K_1$; (b) if $I_1 = J_2$ and $I_2 \neq J_1$, then $K_2 = L_1$ and L_2 is uniformly distributed on $N - \{K_1, L_1\}$; (c) if $I_1 \neq J_2$ and $I_2 = J_1$, then $L_2 = K_1$ and K_2 is uniformly distributed on $N - \{K_1, L_1\}$; (d) if $I_1 \neq J_2$ and $I_2 \neq J_1$, then (K_2, L_2) is uniformly distributed on the set of all ordered pairs of distinct elements in $N - \{K_1, L_1\}$.

Next, let π_1 be a random permutation which is uniformly distributed on $\mathscr{P}_n$ and independent of $(I_1, J_1, I_2, J_2, K_1, L_1, K_2, L_2)$. Define

$$I_3 = \pi_1^{-1}(K_1),\ J_3 = \pi_1^{-1}(L_1),\ I_4 = \pi_1^{-1}(K_2),\ J_4 = \pi_1^{-1}(L_2),$$
$$K_3 = \pi_1(I_1),\ L_3 = \pi_1(J_1),\ K_4 = \pi_1(I_2),\ L_4 = \pi_1(J_2),$$

and denote $\underline{I} = (I_1, I_2, I_3, I_4)$, and similarly for $\underline{J}$, $\underline{K}$ and $\underline{L}$. Thus, $I_1 = I_2 \Leftrightarrow I_3 = I_4$, $J_1 = J_2 \Leftrightarrow J_3 = J_4$, $I_1 = J_2 \Leftrightarrow I_4 = J_3$ and $I_2 = J_1 \Leftrightarrow I_3 = J_4$. Let

$$M = \Big\{(\underline{i}, \underline{j}) \in N^8 : i_m \neq j_m,\ m = 1, \dots, 4, \\ \text{and satisfy the equivalence relations} \\ i_1 = i_2 \Leftrightarrow i_3 = i_4,\ j_1 = j_2 \Leftrightarrow j_3 = j_4, \\ i_1 = j_2 \Leftrightarrow i_4 = j_3 \text{ and } i_2 = j_1 \Leftrightarrow i_3 = j_4\Big\}.$$

For each $(\underline{i}, \underline{j}) \in M$, we fix once and for all permutations $t_1(\underline{i}, \underline{j})$ and $t_2(\underline{i}, \underline{j})$ of N with the properties described in Table 1.

TABLE 1
Definition of the permutations $t_1(\underline{i}, \underline{j})$ and $t_2(\underline{i}, \underline{j})$

	i_1	j_1	i_2	j_2	i_3	j_3	i_4	j_4	$N - \{i_1, \dots, i_4, j_1, \dots, j_4\}$
$t_1(\underline{i}, \underline{j})$	i_4	j_4	i_3	j_3	$\in \{i_1, \dots, i_4, j_1, \dots, j_4\}$				
$t_2(\underline{i}, \underline{j})$	i_2	j_2	$\in \{i_1, i_2, j_1, j_2\}$		Remain fixed				

Finally, define $\pi_2 = \pi_1 \circ t_1(\underline{I}, \underline{J})$ and $\pi_3 = \pi \circ t_2(\underline{I}, \underline{J})$. Summarizing the above definitions, we have Table 2 which shows how π_1, π_2 and π_3 map $\underline{I}$ and $\underline{J}$.

LEMMA 1. (i) *The terms π_1, π_2, and π_3 are independent of $(\underline{I}, \underline{J})$ and have the same law.*

(ii) *The term π_1 is independent of $(I_1, J_1, K_1, L_1, I_2, J_2, K_2, L_2)$ and π_2 is independent of (I_1, J_1, K_1, L_1).*

PROOF. (i) For given $\pi_0 \in \mathscr{P}_n$ and each $(\underline{i}, \underline{j}) \in M$, by the definition of π_1 and the independence of π_1 and $(I_m, J_m, K_m, L_m, m = 1, 2)$,

$$
\begin{aligned}
P\big(&\underline{I} = \underline{i}, \underline{J} = \underline{j}, \pi_1 = \pi_0\big) \\
&= P\big(I_m = i_m, J_m = j_m, K_m = \pi_0(i_{m+2}), L_m = \pi_0(j_{m+2}), m = 1, 2\big) \\
&\quad \times P(\pi_1 = \pi_0) \\
&= P\big((I_1, I_2, K_1, K_2) = \underline{i}, (J_1, J_2, L_1, L_2) = \underline{j}\big) P(\pi_1 = \pi_0).
\end{aligned}
$$

Summation over all $\pi_0 \in \mathscr{P}_n$ gives that $(\underline{I}, \underline{J})$ and $((I_1, I_2, K_1, K_2), (J_1, J_2, L_1, L_2))$ have the same law. Hence, π_1 is independent of $(\underline{I}, \underline{J})$. It then implies that

$$
\begin{aligned}
P\big(\underline{I} = \underline{i}, \underline{J} = \underline{j}, \pi_2 = \pi_0\big) &= P\big(\underline{I} = \underline{i}, \underline{J} = \underline{j}\big) P\big(\pi_1 = \pi_0 \circ t_1^{-1}(\underline{i}, \underline{j})\big) \\
&= \frac{1}{n!} P\big(\underline{I} = \underline{i}, \underline{J} = \underline{j}\big).
\end{aligned}
$$

The assertion for π_2 follows. The assertion for π_3 can be proved similarly.

(ii) Similarly to (i), by the independence of π_2 and $(\underline{I}, \underline{J})$, the assertion for π_2 follows. □

TABLE 2
Values of $\underline{I}$ and $\underline{J}$ under π_1, π_2, π_3

	I_1	J_1	I_2	J_2	I_3	J_3	I_4	J_4	$N - \{I_1, \dots, I_4, J_1, \dots, J_4\}$
π_1	K_3	L_3	K_4	L_4	K_1	L_1	K_2	L_2	$\in N - \{K_1, \dots, K_4, L_1, \dots, L_4\}$
π_3	K_2	L_2	K_1	L_1	$\in \{K_1, \dots, K_4, L_1, \dots, L_4\}$				Same as π_1
π_3	K_1	L_1	$\in \{K_1, K_2, L_1, L_2\}$		Same as π_2				Same as π_1

PROOF OF THEOREM 1. We use c to denote a positive constant which depends only on the formula where it appears. It may stand for different values even in consecutive inequalities.

Let $T = \sum_i a(i, \pi(i))$ and $S = \sum'_{i,j} b(i, j, \pi(i), \pi(j))$. Using (2.1), (2.2) and (2.3), we easily obtain

$$ET = 0, \qquad ES = \frac{1}{n(n-1)} \sum_{i,k} b(i, i, k, k), \tag{4.1}$$

$$ET^2 = 1 \quad \text{and} \quad ES^2 \le cn^{-2} \sum_{i,j,k,l} b^2(i, j, k, l). \tag{4.2}$$

Let $\alpha_A = \sum_{i,k} |a(i,k)|^3$ and $\beta_B = \sum_{i,j,k,l} |b(i,j,k,l)|^3$. Then, by (2.2) and Jensen's inequality,

$$\alpha_A \ge cn^{1/2}, \qquad n \ge 2, \tag{4.3}$$

and

$$n^{-3-1/2} \sum_{i,j,k,l} b^2(i,j,k,l) \le c\left(n^{-3}\beta_B + n^{-1/2}\right). \tag{4.4}$$

For arbitrary but fixed $n_0 \ge 8$ and $\varepsilon_0 > 0$, the statement of the theorem is true if $2 \le n \le n_0$ or $\alpha_A + n^{-2}\beta_B > \varepsilon_0 n$. Therefore, we assume that $n > n_0$ and $\alpha_A + n^{-2}\beta_B \le \varepsilon_0 n$, where n_0 and ε_0 will be specified later on but $n_0 > 8$.

For $\gamma > 0$, let

$$M_n(\gamma) = \left\{(A, B) \colon A \text{ and } B \text{ satisfy } (2.1), (2.2), (2.3) \text{ and } \alpha_A + n^{-2}\beta_B \le \gamma\right\}.$$

For large n, we may assume that $\gamma \ge 1$ due to (4.3). For $z, x \in \mathbb{R}$, $\lambda > 0$, define

$$h_{z,\lambda}(x) = ((1 + (z - x)/\lambda) \wedge 1) \vee 0 \quad \text{and} \quad h_{z,0}(x) = I_{(-\infty, z]}(x).$$

Let

$$\delta(\lambda, \gamma, n) = \sup\{|Eh_{z,\lambda}(W) - \Phi(h_{z,\lambda})| \colon z \in \mathbb{R}, (A, B) \in M_n(\gamma)\}$$

and $\delta(\gamma, n) = \delta(0, \gamma, n)$. Here, $\Phi(h_{z,\lambda})$ is the standard normal expectation of $h_{z,\lambda}$. Thus,

$$\delta(\gamma, n) \le \delta(\lambda, \gamma, n) + \lambda(2\pi)^{-1/2} \tag{4.5}$$

and what we aim to prove is

$$\sup\{n\delta(\gamma, n)/\gamma \colon \gamma \ge 1, n > n_0\} < \infty. \tag{4.6}$$

From now on, we write h instead of $h_{z,\lambda}$ for convenience. To use Stein's method, we let

$$f(x) = \exp(x^2/2) \int_{-\infty}^{x} (h(t) - \Phi(h)) \exp(-t^2/2)\, dt,$$

which satisfies the differential equation $f'(x) - xf(x) = h(x) - \Phi(h)$. Also, as stated in Bolthausen (1984), page 381,

$$|f(x)| \le 1, |xf(x)| \le 1, |f'(x)| \le 2 \quad \text{for all } x \in \mathbb{R} \tag{4.7}$$

and

$$(4.8)\quad |f'(x+y)-f'(x)| \le |y|\left(1+2|x|+\lambda^{-1}\int_0^1 I_{[z,z+\lambda]}(x+sy)\,ds\right).$$

To prove the theorem, we fix $(A, B) \in M_n(\gamma)$ and estimate

$$(4.9)\qquad |Eh(W)-\Phi(h)| = |Ef'(W)-EWf(W)|.$$

By the same truncation used in Bolthausen (1984), pages 381 and 382, we may assume

$$(4.10)\qquad |a(i,k)| \le 1 \quad \text{for all } i, k \in N.$$

Denote the set of these pairs $(A, B) \in M_n(\gamma)$ by $M_n^0(\gamma)$.

To prove Theorem 1, we need only estimate $|Ef'(W) - ETf(W)|$ and $|En^{-1}Sf(W)|$, which will be completed in Lemmas 2 and 3. However, in order to show the utility of the independence created in Lemma 1, parts of the proof of Lemma 2 are contained in the proof of Theorem 1. To this end, let π_m, $m = 1, 2, 3$, be defined as in Lemma 1 and define

$$W_m = T_m + n^{-1}S_m = \sum_i a(i,\pi_m(i)) + n^{-1}\sum_{i,j}{}' b(i,j,\pi_m(i),\pi_m(j))$$

and

$$\Delta T_m = T_{m+1} - T_m.$$

Then $\Delta T_1 \in \sigma(\underline{I}, \underline{J}, \underline{K}, \underline{L})$ and $\Delta T_2 \in \sigma(I_1, J_1, I_2, J_2, K_1, L_1, K_2, L_2)$, where $\sigma(\underline{X})$ denotes the σ-field generated by $\underline{X}$. Also, define

$$S^* = \sum_{i,j\in N-\{I_1,\dots,I_4,J_1,\dots,J_4\}}{}' b(i,j,\pi_1(i),\pi_1(j)),$$

which plays a vital bridge role in our derivations.

The independence of π_3 and I_1 [Lemma 1(i)] and $\pi_3(I_1) = K_1$ imply that

$$(4.11)\quad nEa(I_1,K_1)f(W_3) = nE\{E[a(I_1,\pi_3(I_1))f(W_3)|\pi_3]\} = ET_3 f(W_3).$$

We claim that

$$(4.12)\qquad \left|nEa(I_1,K_1)\big(f(W_3) - f(T_3 + n^{-1}S^*)\big)\right| \le cn^{-1}\gamma.$$

Using the mean-value theorem, $|f'(x)| \le 2$ of (4.7) and (2.2), we have

$$(4.13)\qquad \begin{aligned} &\left|nEa(I_1,K_1)\big(f(W_3) - f(T_3 + n^{-1}S^*)\big)\right| \\ &\quad\le 2E|a(I_1,K_1)(S_3 - S^*)| \\ &\quad\le 2\big(Ea^2(I_1,K_1)E(S_3-S^*)^2\big)^{1/2} \\ &\quad\le cn^{-1/2}\big(E(S_3-S^*)^2\big)^{1/2}. \end{aligned}$$

From Table 2, we see that $S_3 - S^*$ can be expressed as a sum of terms of the forms $\sum_{j\ne I_m} b(I_m, j, \pi_3(I_m), \pi_3(j))$, $\sum_{i\ne J_m} b(i, J_m, \pi_3(i), \pi_3(J_m))$ and $-b(I_m, J_{m'}, \pi_3(I_m), \pi_3(J_{m'}))I_{\{I_m\ne J_{m'}\}}$, and the number of these terms is bounded. Here, by Lemma 1(i) and (2.3), the second moment of the first one is $\le cn^{-3}\sum_{i,j,k,l} b^2(i,j,k,l)$ and that of the third one is $\le cn^{-4}\sum_{i,j,k,l} b^2(i,j,k,l)$. Thus, by (4.4), (4.12) can be proved.

Since, also, $|nEa(I_1, K_1)(f(T_3 + n^{-1}S_1) - f(T_3 + n^{-1}S^*))| \le cn^{-1}\gamma$, we have

$$\left|nEa(I_1, K_1)\big(f(W_3) - f\big(T_3 + n^{-1}S_1\big)\big)\right| \le cn^{-1}\gamma. \tag{4.14}$$

Similarly,

$$\left|nEa(I_1, K_1)\big(f(W_2) - f\big(T_2 + n^{-1}S_1\big)\big)\right| \le cn^{-1}\gamma. \tag{4.15}$$

Since π_2 and (I_1, K_1) are independent [Lemma 1(ii)], and using (2.1), (4.14) and (4.15), we obtain

$$\begin{aligned} &nEa(I_1, K_1) f(W_3) \\ &\quad = nEa(I_1, K_1)\big(f\big(T_3 + n^{-1}S_1\big) - f\big(T_2 + n^{-1}S_1\big)\big) + O\big(n^{-1}\gamma\big) \\ &\quad = nEa(I_1, K_1)\,\Delta T_2 \int_0^1 \big(f'\big(T_2 + n^{-1}S_1 + t\,\Delta T_2\big) - f'(W_1)\big)\,dt \\ &\qquad + nEa(I_1, K_1)\,\Delta T_2\, f'(W_1) + O\big(n^{-1}\gamma\big) \\ &\quad = H_{1n} + H_{2n} + O\big(n^{-1}\gamma\big) \quad \text{say.} \end{aligned} \tag{4.16}$$

By Lemma 1(ii), (4.1) and (4.2),

$$\begin{aligned} H_{2n} &= n\big(Ea(I_1, K_1)\,\Delta T_2\big)\big(Ef'(W_1)\big) \\ &= n\big(Ea(I_1, K_1)T_3\big)\big(Ef'(W_1)\big) = Ef'(W_1). \end{aligned} \tag{4.17}$$

Now, combining (4.11), (4.16) and (4.17), it remains to estimate H_{1n} and $|n^{-1}ES_3 f(W_3)|$. These include a series of inequalities on orders of magnitudes and complicated conditional arguments, which will be presented in the proofs of Lemmas 2 and 3. From those lemmas and (4.9),

$$|Eh(W) - \Phi(h)| \le cn^{-1}\gamma\Big(1 + (n\lambda)^{-1}\gamma + \lambda^{-1} \max_{2 \le m \le 8} \delta(c_1\gamma, n-m)\Big),$$

where c_1 is an absolute constant. By (4.5),

$$\delta(\gamma, n) \le c_2 n^{-1}\gamma\Big(1 + (n\lambda)^{-1}\gamma + \lambda^{-1} \max_{2 \le m \le 8} \delta(c_1\gamma, n-m)\Big) + \lambda(2\pi)^{-1/2}$$

for some absolute constant $c_2 > 0$. Taking $\lambda = 2c_1c_2n^{-1}\gamma$, we then have

$$\delta(\gamma, n) \le cn^{-1}\gamma + (2c_1)^{-1} \max_{2 \le m \le 8} \delta(c_1\gamma, n-m).$$

If $n \ge 16$, then

$$\sup_\gamma \{n\delta(\gamma, n)/\gamma\} \le c + \tfrac{1}{2} \max_{2 \le m \le 8} \sup_\gamma \{(n-m)\delta(\gamma, n-m)/\gamma\}.$$

This implies (4.6) and the theorem is proved. □

LEMMA 2. *There exists an absolute constant c_1 such that*

$$|Ef'(W_3) - ET_3 f(W_3)| \le cn^{-1}\gamma\Big(1 + (n\lambda)^{-1}\gamma + \lambda^{-1} \max_{2 \le m \le 8} \delta(c_1\gamma, n-m)\Big).$$

PROOF. We first claim that

$$\begin{aligned} H_{1n}^* &= \left|nEa(I_1, K_1)\,\Delta T_2\big(f'(W_1) - f'\big(T_1 + n^{-1}S^*\big)\big)\right| \\ &\le c\big(n^{1/2}\lambda\big)^{-1} n^{-1}\gamma. \end{aligned} \tag{4.18}$$

Since $|f''(x)| \le 2/\lambda$ [see Stein (1986), page 25], $H_{1n}^* \le 2\lambda^{-1}E|a(I_1, K_1)\Delta T_2(S_1 - S^*)|$. Here, ΔT_2 is a sum of $\pm a(u,v)$, $u \in \{I_1, J_1, I_2, J_2\}$ and $v \in \{K_1, L_1, K_2, L_2\}$; and $S_1 - S^*$ is similar to $S_3 - S^*$ except replacing π_3 by π_1. Thus, we need to estimate $\eta_m = E|a(u,v)a(I_1, K_1)\sum_{j \ne I_m} b(I_m, j, \pi_1(I_m), \pi_1(j))|$, $\eta_m' = E|a(u,v)a(I_1, K_1)\sum_{i \ne J_m} b(i, J_m, \pi_1(i), \pi_1(J_m))|$ and $\eta_{mm'} = E|a(u,v)a(I_1, K_1)b(I_m, J_{m'}, \pi_1(I_m), \pi_1(J_{m'}))I_{\{I_m \ne J_{m'}\}}|$, for $m, m' = 1, \dots, 4$. Let $\xi(I_m) = \sum_{j \ne I_m} b(I_m, j, \pi_1(I_m), \pi_1(j))$. Since $E|a(u,v)|^3 = E|a(I_1, K_1)|^3$, by Hölder's inequality,

$$\begin{aligned}\eta_m &\le \left(E|a(u,v)|^3\right)^{1/3}\left(E|a(I_1,K_1)|^3\right)^{1/6}\left(E|a(I_1,K_1)\xi^2(I_m)|\right)^{1/2}\\ &= \left(E|a(I_1,K_1)|^3\right)^{1/2}\left(E|a(I_1,K_1)\xi^2(I_m)|\right)^{1/2}.\end{aligned}$$

By the independence of π_1 and (I_1, K_1), using (2.3), we obtain

$$\begin{aligned}E|a(I_1,K_1)\xi^2(I_1)| &= E\big(|a(I_1,K_1)|E(\xi^2(I_1)\,|\,I_1,K_1)\big)\\ &\le cn^{-2}E|a(I_1,K_1)|\sum_{j\ne I_1}\sum_{k,l} b^2(I_1,j,k,l)\\ &\le c\left(nE|a(I_1,K_1)|^3\right)^{1/3}\left(n^{-3}\sum_{i,j,k,l}|b(i,j,k,l)|^3\right)^{2/3}.\end{aligned}$$

Therefore, $\eta_1 \le cn^{-3/2}\gamma$. Also, from the independence of π_1 and (I_1, K_1, I_2), and $\xi(I_3) = \sum_{l \ne K_1} b(\pi_1^{-1}(K_1), \pi_1^{-1}(l), K_1, l)$, similar arguments lead to $\eta_m \le cn^{-3/2}\gamma$, $m = 2, 3$. All η_m' can be estimated similarly and have the same bound. Also, $\eta_{mm'} \le cn^{-2}\gamma$. Therefore, (4.18) is proved.

Similar to (4.18), we have

$$\begin{aligned}\Big|nEa(I_1,K_1)\,\Delta T_2\big(f'(T_2 + n^{-1}S_1 + t\,\Delta T_2)\\ - f'(T_2 + n^{-1}S^* + t\,\Delta T_2)\big)\Big|\\ \le c(n^{1/2}\lambda)^{-1}n^{-1}\gamma \quad \text{for all } t \in [0,1].\end{aligned} \tag{4.19}$$

From (4.16), (4.18) and (4.19), we can write

$$H_{1n} = H_{3n} + O\big((n^{1/2}\lambda)^{-1}n^{-1}\gamma\big), \tag{4.20}$$

where

$$\begin{aligned}H_{3n} = nEa(I_1,K_1)\,\Delta T_2\int_0^1\big(f'(T_2 + n^{-1}S^* + t\,\Delta T_2)\\ -f'(T_1 + n^{-1}S^*)\big)\,dt.\end{aligned} \tag{4.21}$$

By (4.8), if we let $V_1 = |a(I_1, K_1)\,\Delta T_2|(|\Delta T_1| + |\Delta T_2|)$, then

$$\begin{aligned}|H_{3n}| \le nEV_1\Big(1 + 2|W_1 - n^{-1}(S_1 - S^*)|\\ + \lambda^{-1}\int_0^1\int_0^1 I_{[z,z+\lambda]}(T_1 + n^{-1}S^* + s\,\Delta T_1 + st\,\Delta T_2)\,ds\,dt\Big).\end{aligned} \tag{4.22}$$

Note that ΔT_1 is a sum of $\pm a(u_1, v_1)$, $u_1 \in \{I_1, \dots, I_4, J_1, \dots, J_4\}$ and $v_1 \in \{K_1, \dots, K_4, L_1, \dots, L_4\}$. Therefore,

$$nEV_1 \le cnE\big|a(I_1, K_1)\big|^3 \le cn^{-1}\gamma. \tag{4.23}$$

Also, ΔT_2 is a sum of $\pm a(u, v)$, $u \in \{I_1, I_2, J_1, J_2\}$, $v \in \{K_1, K_2, L_1, L_2\}$. From the independence of W_1 and $|a^2(I_1, K_1)\,\Delta T_2|$, (4.2) and $\alpha_A + n^{-2}\beta_B \le \varepsilon_0 n$, we have

$$\begin{aligned} nE|V_1 W_1| &\le n\Big(E(|\Delta T_1| + |\Delta T_2|)^2 |\Delta T_2|\Big)^{1/2}\Big(E\big|a^2(I_1, K_1)\,\Delta T_2\big|W_1^2\Big)^{1/2} \\ &\le n\Big(E\big|a(I_1, K_1)\big|^3\Big)^{1/2} \\ &\quad\times\Big\{\Big[E\big|a^2(I_1, K_1)\,\Delta T_2\big|\Big]\Big[E\big(T_1 - n^{-1}S_1\big)^2\Big]\Big\}^{1/2} \\ &\le cn\Big(E\big|a(I_1, K_1)\big|^3\Big)^{1/2} \\ &\quad\times\Big\{E\big|a(I_1, K_1)\big|^3\Big(1 + n^{-4}\sum_{i,j,k,l} b^2(i,j,k,l)\Big)\Big\}^{1/2} \\ &\le cn^{-1}\alpha_A\big(1 + n^{-1/2}\varepsilon_0\big)^{1/2} \le cn^{-1}\gamma. \end{aligned} \tag{4.24}$$

From (4.10) and the derivation of (4.18),

$$\begin{aligned} &nE\big|a(I_1, K_1)\,\Delta T_2\big|\big(|\Delta T_1| + |\Delta T_2|\big)\big|n^{-1}(S_1 - S^*)\big| \\ &\qquad\le cE\big|a(I_1, K_1)\,\Delta T_2(S_1 - S^*)\big| \le cn^{-1}\gamma. \end{aligned} \tag{4.25}$$

Now the only remaining part of H_{3n} to be estimated is

$$H_{4n} = n\lambda^{-1}EV_1\int_0^1\int_0^1 I_{[z,z+\lambda]}\big(T_1 + n^{-1}S^* + s\,\Delta T_1 + st\,\Delta T_2\big)\,ds\,dt. \tag{4.26}$$

Note that the conditional distribution of π_1 given $\underline{I} = \underline{i}$, $\underline{J} = \underline{j}$, $\underline{K} = \underline{k}$ and $\underline{L} = \underline{l}$ can be described as follows: π_1 takes each $\varphi \in \mathscr{P}_n$, which satisfies $\varphi(i_m) = k_{m+2}$ and $\varphi(j_m) = l_{m+2}$ for $m = 1, 2$, and $\varphi(i_m) = k_{m-2}$ and $\varphi(j_m) = l_{m-2}$ for $m = 3, 4$, with equal probability. For each 4-tuple B and given $i \in N$, define the I-row i of B,

$$B(i, N, N, N) = \big\{b(i,j,k,l) \colon (j,k,l) \in N^3\big\}.$$

The II-row j of B, $B(N, j, N, N)$, the I-column k of B, $B(N, N, k, N)$, and the II-column l of B, $B(N, N, N, l)$, are defined similarly. Let $\tilde{A}$ denote the matrix obtained from A by deleting the rows $i_1, \dots, i_4, j_1, \dots, j_4$, and the columns $k_1, \dots, k_4, l_1, \dots, l_4$. Let $\tilde{B}$ denote the 4-tuple obtained by deleting the I-rows and the II-rows $i_1, \dots, i_4, j_1, \dots, j_4$, and the I-columns and the II-columns $k_1, \dots, k_4, l_1, \dots, l_4$. Then $T_1 + n^{-1}S^*$, conditioned on $\underline{I} = \underline{i}$, $\underline{J} = \underline{j}$, $\underline{K} = \underline{k}$, $\underline{L} = \underline{l}$, has the same law as

$$\sum_{i \in \{i_1, \dots, i_4, j_1, \dots, j_4\}} a(i, \pi_1(i)) + \sum_{i=1}^{n-m} \tilde{a}(i, \tau(i)) + n^{-1}\sum_{i,j=1}^{n-m}{}' \tilde{b}(i, j, \tau(i), \tau(j)),$$

where m is the number of distinct elements of $\{i_1, \dots, i_4, j_1, \dots, j_4\}$ and τ is uniformly distributed on $\mathscr{P}_{n-m}$, $2 \le m \le 8$. Let $M_n^m(\gamma)$ be the set of all pairs $(\tilde{A}, \tilde{B})$, which can be obtained from $(A, B) \in M_n^0(\gamma)$ by deleting m rows and m columns of A, and m I-rows, m II-rows, m I-columns and m II-columns of B. Introducing

$$\alpha(\lambda, \gamma, n) = \sup\{\| P(T_1 + n^{-1}S^* \in [z, z+\lambda] \mid \underline{I}, \underline{J}, \underline{K}, \underline{L})\|_\infty: \\ z \in \mathbb{R}, (A, B) \in M_n^0(\gamma)\}, \tag{4.27}$$

where $\|g(\cdot)\|_\infty = \sup|g(\cdot)|$, we have

$$\alpha(\lambda, \gamma, n) \le \sup\Biggl\{ P\Biggl(\sum_{i=1}^{n-m} \tilde{a}(i, \tau(i)) \\ + n^{-1} \sum_{i,j=1}^{n-m}{}' \tilde{b}(i, j, \tau(i), \tau(j)) \in [z, z+\lambda] \Biggr): \\ z \in \mathbb{R}, (\tilde{A}, \tilde{B}) \in M_n^m(\gamma), 2 \le m \le 8 \Biggr\}. \tag{4.28}$$

Let $\sigma_{\tilde{A}}^2 = (n - m - 1)^{-1} \sum_{i,k=1}^{n-m} (\tilde{a}(i,k) - \tilde{a}(i,\cdot) - \tilde{a}(\cdot,k) + \tilde{a}(\cdot,\cdot))^2$ and $\tilde{a}^*(i,k) = (\tilde{a}(i,k) - \tilde{a}(i,\cdot) - \tilde{a}(\cdot,k) + \tilde{a}(\cdot,\cdot))/\sigma_{\tilde{A}}$. Since $(\tilde{A}, \tilde{B}) \in M_n^m(\gamma)$, $|\tilde{a}(i,\cdot)|$, $|\tilde{a}(\cdot,k)|$ and $|\tilde{a}(\cdot,\cdot)|$ are less than or equal to cn^{-1}. Therefore,

$$\left| \sigma_{\tilde{A}}^2 - \frac{1}{n-m-1} \sum_{i,k} a^2(i,k) \right| \le \frac{1}{n-m-1} \tilde{\sum} a^2(i,k) + o(1),$$

where $\tilde{\Sigma}$ is the sum over the deleted elements of A. Furthermore, if ε_0 is taken small enough and n_0 taken sufficiently large, then since $\tilde{\Sigma} a^2(i,k) \le cn\varepsilon_0^{2/3}$ when $n \ge n_0$ and $\alpha_A + n^{-2}\beta_B \le \varepsilon_0 n$, we have $|\sigma_{\tilde{A}}^2 - 1| \le 1/2$ and hence $\sigma_{\tilde{A}}^2 \ge 1/2$. Therefore, for sufficiently large n_0 and $n \ge n_0$,

$$\sup_z P\Biggl(\sum_{i=1}^{n-m} \tilde{a}(i, \tau(i)) \\ + n^{-1} \sum_{i,j=1}^{n-m}{}' \tilde{b}(i, j, \tau(i), \tau(j))) \in [z, z+\lambda] \Biggr) \\ \le \sup_z P\Biggl(\sum_{i=1}^{n-m} \tilde{a}^*(i, \tau(i)) \\ + \frac{1}{n-m} \sum_{i,j=1}^{n-m}{}' \tilde{b}(i, j, \tau(i), \tau(j)) / \sigma_{\tilde{A}} \in [z, z+2\lambda] \Biggr). \tag{4.29}$$

Let

$$\begin{aligned}
\tilde{b}_0 &= \tilde{b}(\cdot,\cdot,\cdot,\cdot),\\
\tilde{b}_1(i,\cdot,\cdot,\cdot) &= \tilde{b}(i,\cdot,\cdot,\cdot) - \tilde{b}_0,\\
&\ \vdots\\
\tilde{b}_2(i,j,\cdot,\cdot) &= \tilde{b}(i,j,\cdot,\cdot) - \big(\tilde{b}(i,\cdot,\cdot,\cdot) + \tilde{b}(\cdot,j,\cdot,\cdot)\big) + \tilde{b}_0,\\
&\ \vdots\\
\tilde{b}_3(i,j,k,\cdot) &= \tilde{b}(i,j,k,\cdot) - \big(\tilde{b}(i,j,\cdot,\cdot) + \tilde{b}(i,\cdot,k,\cdot) + \tilde{b}(\cdot,j,k,\cdot)\big)\\
&\quad + \big(\tilde{b}(i,\cdot,\cdot,\cdot) + \tilde{b}(\cdot,j,\cdot,\cdot) + \tilde{b}(\cdot,\cdot,k,\cdot)\big) - \tilde{b}_0,\\
&\ \vdots\\
\tilde{b}^*(i,j,k,l) &= \tilde{b}(i,j,k,l) - \big(\tilde{b}(i,j,k,\cdot) + \tilde{b}(i,j,\cdot,l)\\
&\qquad + \tilde{b}(i,\cdot,k,l) + \tilde{b}(\cdot,j,k,l)\big)\\
&\quad + \big(\tilde{b}(i,j,\cdot,\cdot) + \tilde{b}(i,\cdot,k,\cdot) + \tilde{b}(i,\cdot,\cdot,l)\\
&\quad + \tilde{b}(\cdot,j,k,\cdot) + \tilde{b}(\cdot,j,\cdot,l) + \tilde{b}(\cdot,\cdot,k,l)\big)\\
&\quad - \big(\tilde{b}(i,\cdot,\cdot,\cdot) + \tilde{b}(\cdot,j,\cdot,\cdot) + \tilde{b}(\cdot,\cdot,k,\cdot)\\
&\qquad + \tilde{b}(\cdot,\cdot,\cdot,l)\big) + \tilde{b}_0.
\end{aligned}$$

Straightforward calculations give

$$\begin{aligned}
\frac{1}{n-m}&\sum_{i,j=1}^{n-m}{}' \tilde{b}(i,j,\tau(i),\tau(j))\\
&= \frac{1}{n-m}\sum_{i,j=1}^{n-m}{}' \tilde{b}^*(i,j,\tau(i),\tau(j))\\
&\quad - \frac{1}{n-m}\sum_{i=1}^{n-m}\big\{\tilde{b}_3(i,i,\tau(i),\cdot) + \tilde{b}_3(i,i,\cdot,\tau(i))\\
&\qquad + \tilde{b}_3(i,\cdot,\tau(i),\tau(i)) + \tilde{b}_3(\cdot,i,\tau(i),\tau(i))\big\}\\
&\quad + \frac{n-m-1}{n-m}\sum_{i=1}^{n-m}\big\{\tilde{b}_2(i,\cdot,\tau(i),\cdot) + \tilde{b}_2(\cdot,i,\cdot,\tau(i))\big\}\\
&\quad - \frac{1}{n-m}\sum_{i=1}^{n-m}\big\{\tilde{b}_2(i,\cdot,\cdot,\tau(i)) + \tilde{b}_2(\cdot,i,\tau(i),\cdot)\big\}\\
&\quad + (n-m-1)\tilde{b}_0 - \frac{1}{n-m}\sum_{i=1}^{n-m}\big\{\tilde{b}_2(i,i,\cdot,\cdot) + \tilde{b}_2(\cdot,\cdot,i,i)\big\}\\
&= U_{n-m} + \Delta_3 + \Delta_2 + \Delta_2' + \Delta_0 \quad \text{say.}
\end{aligned} \tag{4.30}$$

Let N_1, N_2, N_3 and N_4 be the sets of indices of those m I-rows, m II-rows, m I-columns and m II-columns, respectively, that were deleted while forming $\tilde{B}$

from B. Also, let the index in B corresponding to the index i in $\tilde{B}$ be denoted as i'. We have

$$\tilde{b}(i,j,k,\cdot) = \frac{1}{n-m}\sum_{l\in N-N_4} b(i',j',k',l) = -\frac{1}{n-m}\sum_{l\in N_4} b(i',j',k',l),$$

$$\tilde{b}(i,\cdot,k,\cdot) = \frac{1}{(n-m)^2}\sum_{j\in N_2,\, l\in N_4} b(i',j,k',l),$$

etc. Hence, we get $E(\sum_{i=1}^{n-m}\tilde{b}_3(i,i,\tau(i),\cdot))^2 \le cn^{-3}\sum_{i,j,k,l}b^2(i,j,k,l)$, $E(\sum_{i=1}^{n-m}\tilde{b}_2(i,\cdot,\tau(i),\cdot))^2 \le cn^{-5}\sum_{i,j,k,l}b^2(i,j,k,l)$, and some other similar inequalities. These can be used to derive

$$P\big(|(\Delta_3+\Delta_2+\Delta_2')/\sigma_{\tilde{A}}| \ge n^{-1/2}\big) \le cnE\big\{(\Delta_3)^2+(\Delta_2)^2+(\Delta_2')^2\big\} \le cn^{-1}\gamma. \tag{4.31}$$

Therefore, from (4.29), (4.30) and (4.31),

$$\begin{aligned}
\sup_z P\Bigg(\sum_{i=1}^{n-m}\tilde{a}(i,\tau(i)) + n^{-1}\sideset{}{'}\sum_{i,j=1}^{n-m}\tilde{b}(i,j,\tau(i),\tau(j)) \in [z,z+\lambda]\Bigg)&\\
\le \sup_z P\Bigg(\sum_{i=1}^{n-m}\tilde{a}^*(i,\tau(i))&\\
+(U_{n-m}+\Delta_3+\Delta_2+\Delta_2')/\sigma_{\tilde{A}} \in [z,z+2\lambda]\Bigg)& \qquad (4.32)\\
\le \sup_z P\Bigg(\sum_{i=1}^{n-m}\tilde{a}^*(i,\tau(i))&\\
+\frac{1}{(n-m)\sigma_{\tilde{A}}}\sideset{}{'}\sum_{i,j=1}^{n-m}\tilde{b}^*(i,j,\tau(i),\tau(j))&\\
\in \left[z, z+2\left(\lambda+\frac{1}{\sqrt{n}}\right)\right]\Bigg) + cn^{-1}\gamma.&
\end{aligned}$$

Since $\sigma_{\tilde{A}}^2 \ge 1/2$, there exists an absolute constant $c_1 > 1$, such that

$$\sum_{i,k=1}^{n-m}|\tilde{a}^*(i,k)|^3 \le c_1\alpha_A \quad\text{and}\quad \sum_{i,j,k,l=1}^{n-m}\left|\sigma_{\tilde{A}}^{-1}\tilde{b}^*(i,j,k,l)\right|^3 \le c_1\beta_B. \tag{4.33}$$

From (4.28), (4.32), (4.33) and the definition of $\delta(\cdot,n)$, we have

$$\alpha(\lambda,\gamma,n) \le 2\max_{2\le m\le 8}\delta(c_1\gamma,n-m) + 2(\lambda+n^{-1/2})(2\pi)^{-1/2} + cn^{-1}\gamma. \tag{4.34}$$

From (4.26), (4.23) and (4.34), and noticing the definition (4.27) of $\alpha(\lambda, \gamma, n)$, we obtain

$$H_{4n} \le cn^{-1}\gamma\Big(1 + (n\lambda)^{-1}\gamma + \lambda^{-1}\max_{2\le m\le 8}\delta(c_1\gamma, n-m)\Big). \tag{4.35}$$

Therefore, tracing back to (4.11), (4.16), (4.17), (4.20)–(4.26) and (4.35), we complete the proof. □

LEMMA 3. *There exists an absolute constant c_1 such that*

$$\big|En^{-1}S_3 f(W_3)\big| \le cn^{-1}\gamma\Big(1 + (n\lambda)^{-1}\gamma + \lambda^{-1}\max_{2\le m\le 8}\delta(c_1\gamma, n-m)\Big).$$

PROOF. Since $|f'| \le 2$, using (4.2) and (4.4), we have

$$E\big|n^{-1}S_3(f(W_3) - f(T_3))\big| \le cn^{-1}\gamma. \tag{4.36}$$

By the independence of π_3 and (I_1, J_1), $(n-1)Eb(I_1, J_1, K_1, L_1)f(T_3) = n^{-1}ES_3 f(T_3)$. Since $|f| \le 1$, and with (4.1), $|(n-1)Eb(I_1, J_1, K_1, L_1)f(T_2)| \le cn^{-1}|ES_3| \le c(n^{-1/2} + n^{-4}\beta_B)$. Therefore,

$$\begin{aligned} n^{-1}&ES_3 f(T_3)\\ &= (n-1)Eb(I_1, J_1, K_1, L_1)(f(T_3) - f(T_2)) + O(n^{-1}\gamma)\\ &= (n-1)Eb(I_1, J_1, K_1, L_1)\,\Delta T_2\\ &\qquad \times \int_0^1 (f'(T_1 + \Delta T_1 + t\,\Delta T_2) - f'(T_1))\,dt\\ &\qquad + (n-1)Eb(I_1, J_1, K_1, L_1)\,\Delta T_2\, f'(T_1)\\ &\qquad + O(n^{-1}\gamma)\\ &= H_{5n} + H_{6n} + O(n^{-1}\gamma) \quad \text{say.} \end{aligned} \tag{4.37}$$

Using Lemma 1 and (2.1)–(2.3), we can prove that

$$|H_{6n}| \le 2n\big|Eb(I_1, J_1, K_1, L_1)\,\Delta T_2\big| \le cn^{-1}\gamma. \tag{4.38}$$

Denoting $V_2 = |b(I_1, J_1, K_1, L_1)\,\Delta T_2|(|\Delta T_1| + |\Delta T_2|)$ and using (4.8) and (4.37), we have

$$\begin{aligned} |H_{5n}| \le nEV_2\Big(&1 + 2|T_1|\\ &+ \lambda^{-1}\int_0^1\int_0^1 I_{[z, z+\lambda]}(T_1 + s\,\Delta T_1 + st\,\Delta T_2)\,ds\,dt\Big). \end{aligned} \tag{4.39}$$

Note that

$$\begin{aligned} nEV_2 &\le \big(E|\Delta T_2|^3\big)^{1/3}\Big(E(|\Delta T_1| + |\Delta T_2|)^3\Big)^{1/3}\Big(E\big|b(I_1, J_1, K_1, L_1)\big|^3\Big)^{1/3}\\ &\le cn^{-1}\gamma. \end{aligned} \tag{4.40}$$

By the independence of $b(I_1, J_1, K_1, L_1)\,\Delta T_2$ and T_1, and Hölder's inequality, using (4.40) and $ET_1^2 = 1$, we obtain

$$\begin{aligned} nE(V_2|T_1|) &\le n\Big(E\big(|\Delta T_1| + |\Delta T_2|\big)^3\Big)^{1/3} \\ &\qquad \times \Big[\Big(E\big|b(I_1, J_1, K_1, L_1)\,\Delta T_2\big|^{3/2}\Big)\Big(E|T_1|^{3/2}\Big)\Big]^{2/3} \\ &\le n\Big(E\big(|\Delta T_1| + |\Delta T_2|\big)^3\Big)^{1/3}\big(E|\Delta T_2|^3\big)^{1/3} \\ &\qquad \times \Big(E\big|b(I_1, J_1, K_1, L_1)\big|^3\Big)^{1/3}\big(ET_1^2\big)^{1/2} \\ &\le cn^{-1}\gamma. \end{aligned} \tag{4.41}$$

By similar but simpler derivations for estimating H_{4n}, that is, letting $b \equiv 0$, we can get

$$\begin{aligned} &n\lambda^{-1}EV_2\int_0^1\int_0^1 I_{[z, z+\lambda]}(T_1 + s\,\Delta T_1 + st\,\Delta T_2)\,ds\,dt \\ &\qquad \le cn^{-1}\gamma\Big(1 + (n\lambda)^{-1}\gamma + \lambda^{-1}\max_{2\le m\le 8}\delta(c_1\gamma, n-m)\Big). \end{aligned} \tag{4.42}$$

Therefore, (4.39)–(4.42) imply that

$$|H_{5n}| \le cn^{-1}\gamma\Big(1 + (n\lambda)^{-1}\gamma + \lambda^{-1}\max_{2\le m\le 8}\delta(c_1\gamma, n-m)\Big),$$

and hence, combining with (4.36)–(4.38), we complete the proof. □

Acknowledgment. The authors thank a referee for his valuable comments and suggestions, which greatly improved the presentation of the paper.

REFERENCES

ABE, O. (1969). A central limit theorem for the number of edges in the random intersection of two graphs. *Ann. Math. Statist.* **40** 144–151.

BARBOUR, A. D. and EAGLESON, G. K. (1986). Random association of symmetric arrays. *Stochastic Anal. Appl.* **4** 239–281.

BICKEL, P. J. (1974). Edgeworth expansions in nonparametric statistics. *Ann. Statist.* **2** 1–20.

BICKEL, P. J., GÖTZE, F. and VAN ZWET, W. R. (1986). The Edgeworth expansion for *U*-statistics of degree two. *Ann. Statist.* **14** 1463–1484.

BLOEMENA, A. R. (1964). *Sampling from a Graph. Math. Centre Tract* **2**. Math. Centrum, Amsterdam.

BOLTHAUSEN, E. (1984). An estimate of the remainder in a combinatorial central limit theorem. *Z. Wahrsch. Verw. Gebiete* **66** 379–386.

CLIFF, A. D. and ORD, J. K. (1981). *Spatial Processes; Models and Applications*. Pion, London.

DANIELS, H. E. (1944). The relation between measures of correlation in the universe of sample permutations. *Biometrika* **33** 129–135.

FRIEDMAN, J. H. and RAFSKY, L. C. (1979). Multivariate generalizations of the Wald–Wolfowitz and Smirnov two-sample tests. *Ann. Statist.* **7** 697–717.

FRIEDMAN, J. H. and RAFSKY, L. C. (1983). Graph-theoretic measures of multivariate association and prediction. *Ann. Statist.* **11** 377–391.

HUBERT, L. and SCHULTZ, J. (1976). Quadratic assignment as a general data analysis strategy. *British J. Math. Statist. Psych.* **29** 190–241.

JOGDEO, K. (1968). Asymptotic normality in nonparametric methods. *Ann. Math. Statist.* **39** 905–922.

KOROLJUK, V. S. and BOROVSKICH, YU. V. (1994). *Theory of U-Statistics*. Kluwer, Boston.

MANTEL, N. and VALAND, R. S. (1970). A technique of nonparametric multivariate analysis. *Biometrics* **26** 547–558.

PHAM, D. T., MÖCKS, J. and SROKA, L. (1989). Asymptotic normality of double-indexed linear permutation statistics. *Ann. Inst. Statist. Math.* **41** 415–427.

SCHILLING, M. F. (1986). Multivariate two-sample test based on nearest neighbors. *J. Amer. Statist. Assoc.* **81** 799–806.

SCHNELLER, W. (1989). Edgeworth expansions for linear rank statistics. *Ann. Statist.* **17** 1103–1123.

SHAPIRO, C. P. and HUBERT, L. (1979). Asymptotic normality of permutation statistics derived from weighted sum of bivariate functions. *Ann. Statist.* **7** 788–794.

STEIN, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2** 583–602. Univ. California Press, Berkeley.

STEIN, C. (1986). *Approximate Computation of Expectations*. IMS, Hayward, CA.

L. C. ZHAO
DEPARTMENT OF STATISTICS AND FINANCE
UNIVERSITY OF SCIENCE AND TECHNOLOGY
HEFEI, ANHUI 230026
CHINA

Z. D. BAI
INSTITUTE OF APPLIED MATHEMATICS
NATIONAL SUN YAT-SEN UNIVERSITY
KAOHSIUNG, TAIWAN
R.O.C.

C. C. CHAO
W. Q. LIANG
INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
TAIPEI 11529, TAIWAN
R.O.C.