



## A Constant on a Uniform Bound of a Combinatorial Central Limit Theorem

Kritsana Neammanee (Corresponding author)  
Department of Mathematics, Faculty of Science  
Chulalongkorn University, Bangkok 10330 Thailand  
E-mail: Kritsana.N@chula.ac.th

Petcharat Rattanawong  
Faculty of Science and Social Science  
Burapha University Sakeaw Campus, Sakeaw 27160, Thailand  
E-mail: petchar@buu.ac.th

### Abstract

Let  $n$  be a positive integer and  $Y(i, j)$ ,  $i, j = 1, \dots, n$ , be random variables with finite fourth moments. Let  $\pi$  be a random permutation on  $\{1, \dots, n\}$  which independent of  $Y(i, j)$ 's. In this paper, we use Stein's method and the technique from (Laipaporn, K., 2008) to give a uniform bound in a combinatorial central limit theorem of  $W = \sum_{i=1}^n Y(i, \pi(i))$ . For a sufficient large  $n$ , we yield the rate  $\frac{27.72}{\sqrt{n}}$ . This constant is better than the result in (Neammanee, K., 2005).

**Keywords:** Uniform bound, Combinatorial central limit theorem, Stein's method, Random permutation

### 1. Introduction and Main Results

Let  $n$  be a positive integer and let:

- (1).  $\pi$  be a random permutation of  $\{1, \dots, n\}$  uniformly distributed over all the  $n!$  possible permutations;
- (2).  $Y(i, j)$ ,  $i, j = 1, \dots, n$  be random variables;
- (3). the  $Y(i, j)$ 's and  $\pi$  all be stochastically independent.

We define

$$W = \sum_{i=1}^n Y(i, \pi(i)). \quad (1)$$

This paper is concerned with the normal approximation to the distribution of  $W$  which is always called a combinatorial central limit theorem. The literature concerning a combinatorial central limit theorem dates back to 1944 when Wald and Wolfowitz (Wald, A., 1944) first established the asymptotic normality of the statistic  $\eta_n = \sum_{i=1}^n a_i b_{\pi(i)}$  where  $a_i, b_i, i = 1, 2, \dots, n$ , are two sequences of real numbers and  $\pi$  is a random permutation of  $\{1, 2, \dots, n\}$ . This was extended by Hoeffding (Hoeffding, W., 1951) who considered  $W_n = \sum_{i=1}^n Y(i, \pi(i))$  where  $Y(i, j)$  are  $n^2$  real numbers. Matoo (Matoo, M., 1957) showed that a Lindeberg-type condition is sufficient for the asymptotic normality of  $W_n$ . The same condition was also shown to be necessary in the case of  $\eta_n$  by Hajek (Hajek, J., 1961). In 1972, Robinson (Robinson, J., 1972) obtained necessary and sufficient conditions for the moments of  $\eta_n$  to converge to those of a normal distribution and Kolchin and Chistyakov (Kolchin, V.F., 1973) considered a different  $\eta_n$  where  $\pi$  is no longer uniform but attributes equal probabilities to only those permutations with one cycle.

It seem so far that only limit theorem when  $Y(i, j)$ 's are real numbers has been proved. In the case when  $Y(i, j)$ 's are any random variables, the estimations have been obtained by Von Bahr (Von Bahr, B., 1976) and Ho and Chen (Ho,

S.T., 1978), and a Berry-Esseen-type bound was obtained by Bolthausen(Bolthausen, E., 1984) for univariate linear statistics and Bolthausen and Gotze(Bolthausen, E., 1993) for multivariate statistics. In 2005, Neammanee and Suntornchost(Neammanee, K., 2005) gave the uniform rate by using Stein’s method. In this work, we improve the constant by using the technique from (Laipaporn, K., 2008). Our constant is sharper than the result in (Neammanee, K., 2005)( $C \geq 198$ ). Our results are as follows:

**Theorem 1.1** Suppose that  $EY^4(i, j) < \infty, 1 \leq i, j \leq n,$

$$EW = 0 \text{ and } VarW = 1, \tag{2}$$

and

$$\sum_{i=1}^n \mu(i, j) = 0 \text{ and } \sum_{j=1}^n \mu(i, j) = 0 \tag{3}$$

where  $\mu(i, j) = EY(i, j).$  Then for  $n \geq 30,$

$$\sup_{w \in \mathbb{R}} |P(W \leq w) - \Phi(w)| \leq 11\delta_4 + 10.02 \frac{\delta_4^{3/4}}{n^{1/8}} + 1.467 \frac{\delta_4^{1/2}}{n^{1/4}} + \frac{5.228}{\sqrt{n}}$$

where  $\Phi$  is the standard normal distribution, and

$$\delta_4 = \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{j=1}^n EY^4(i, j).$$

**Corollary 1.2** If  $\delta_4 \sim n^{-1/2},$  then for  $n \geq 30,$

$$\sup_{w \in \mathbb{R}} |P(W \leq w) - \Phi(w)| \leq \frac{27.72}{\sqrt{n}}.$$

**2. Proof of Main Result (Theorem 1.1)**

*2.1 Auxillary results*

In this section, we give auxiliary results for proving our main theorem (Theorem 1.1). First, we need to bound  $EW^4.$

**Lemma 2.1** If  $W$  is defined as in (1), then for  $n \geq 30,$

$$EW^4 \leq 3.44 \sqrt{n}\delta_4.$$

**Proof:** Note that

$$EW^4 = A_1 + A_2 + A_3 + A_4 + A_5$$

where

$$\begin{aligned} A_1 &= \sum_{i=1}^n EY^4(i, \pi(i)) \\ A_2 &= \sum_{i_1=1}^n \sum_{\substack{i_2 \\ i_2 \neq i_1}} EY^3(i_1, \pi(i_1))Y(i_2, \pi(i_2)) \\ A_3 &= \sum_{i_1=1}^n \sum_{\substack{i_2 \\ i_2 \neq i_1}} EY^2(i_1, \pi(i_1))Y^2(i_2, \pi(i_2)) \\ A_4 &= \sum_{i_1=1}^n \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{i_3 \\ i_3 \neq i_1, i_2}} EY^2(i_1, \pi(i_1))Y(i_2, \pi(i_2))Y(i_3, \pi(i_3)) \\ A_5 &= \sum_{i_1=1}^n \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{i_3 \\ i_3 \neq i_1, i_2}} \sum_{\substack{i_4 \\ i_4 \neq i_1, i_2, i_3}} EY(i_1, \pi(i_1))Y(i_2, \pi(i_2))Y(i_3, \pi(i_3))Y(i_4, \pi(i_4)). \end{aligned}$$

By the fact that  $n \geq 30$  and (3) we have

$$A_1 = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n EY^4(i, j) = \frac{\delta_4}{\sqrt{n}} \leq 0.033 \sqrt{n}\delta_4 \tag{4}$$

and

$$\begin{aligned}
 A_2 &= \frac{1}{n(n-1)} \sum_{i_1=1}^n \sum_{j_1=1}^n EY^3(i_1, j_1) \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY(i_2, j_2) \\
 &\leq \frac{\delta_4}{\sqrt{n(n-1)}} \leq 0.001 \sqrt{n} \delta_4.
 \end{aligned} \tag{5}$$

We observe that

$$\left[ \sum_{i=1}^n \sum_{j=1}^n EY^2(i, j) \right]^2 \leq n^2 \sum_{i=1}^n \sum_{j=1}^n [EY^2(i, j)]^2 \leq n^2 \sum_{i=1}^n \sum_{j=1}^n EY^4(i, j) \leq n^2 \sqrt{n} \delta_4. \tag{6}$$

Thus

$$\begin{aligned}
 A_3 &= \frac{1}{n(n-1)} \sum_{i_1=1}^n \sum_{j_1=1}^n EY^2(i_1, j_1) \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY^2(i_2, j_2) \\
 &\leq \frac{1}{n(n-1)} \left[ \sum_{i=1}^n \sum_{j=1}^n EY^2(i, j) \right]^2 \\
 &\leq \frac{n \sqrt{n}}{n-1} \delta_4 \\
 &\leq 1.035 \sqrt{n} \delta_4.
 \end{aligned} \tag{7}$$

Next, we bound  $A_4$ . Note that

$$\begin{aligned}
 A_4 &= \frac{1}{n(n-1)(n-2)} \sum_{i_1=1}^n \sum_{j_1=1}^n EY^2(i_1, j_1) \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY(i_2, j_2) \\
 &\quad \times \sum_{\substack{i_3 \\ i_3 \neq i_1, i_2}} \sum_{\substack{j_3 \\ j_3 \neq j_1, j_2}} EY(i_3, j_3)
 \end{aligned}$$

consists of 4 finite summands of the form

$$\frac{1}{n(n-1)(n-2)} \sum_{i_1=1}^n \sum_{j_1=1}^n EY^2(i_1, j_1) \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY(i_2, j_2) EY(i_3, j_3)$$

where  $i_3 = i_1$  or  $i_2$  and  $j_3 = j_1$  or  $j_2$ .

Recall that, for fixed  $i_1, j_1$ ,

$$\sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY(i_2, j_2) EY(i_3, j_3) \leq \sum_{i_2=1}^n \sum_{j_2=1}^n (EY(i_2, j_2))^2,$$

and (6), we obtain

$$\begin{aligned}
 A_4 &\leq \frac{4}{n(n-1)(n-2)} \sum_{i_1=1}^n \sum_{j_1=1}^n EY^2(i_1, j_1) \sum_{i_2=1}^n \sum_{j_2=1}^n (EY(i_2, j_2))^2 \\
 &\leq \frac{4}{n(n-1)(n-2)} \left[ \sum_{i=1}^n \sum_{j=1}^n EY^2(i, j) \right]^2 \\
 &\leq \frac{4n \sqrt{n}}{(n-1)(n-2)} \delta_4 \\
 &= 4 \left( \frac{n}{n-1} \right) \left( \frac{1}{n-2} \right) \sqrt{n} \delta_4 \\
 &\leq 0.15 \sqrt{n} \delta_4
 \end{aligned} \tag{8}$$

in which we used the fact that  $n \geq 30$  in the last inequality. To bound  $A_5$ , we write

$$A_5 = \frac{1}{n(n-1)(n-2)(n-3)} \sum_{i_1=1}^n \sum_{j_1=1}^n EY(i_1, j_1) \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY(i_2, j_2) \\ \times \sum_{\substack{i_3 \\ i_3 \neq i_1, i_2}} \sum_{\substack{j_3 \\ j_3 \neq j_1, j_2}} EY(i_3, j_3) \sum_{\substack{i_4 \\ i_4 \neq i_1, i_2, i_3}} \sum_{\substack{j_4 \\ j_4 \neq j_1, j_2, j_3}} EY(i_4, j_4)$$

consisting of 9 summands of the form

$$\frac{1}{n(n-1)(n-2)(n-3)} B(m_1, m_2)$$

where

$$B(m_1, m_2) = \sum_{i_1=1}^n \sum_{j_1=1}^n EY(i_1, j_1) \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY(i_2, j_2) \\ \times \sum_{\substack{i_3 \\ i_3 \neq i_1, i_2}} \sum_{\substack{j_3 \\ j_3 \neq j_1, j_2}} EY(i_3, j_3) EY(m_1, m_2),$$

$m_1 = i_1, i_2, i_3$  and  $m_2 = j_1, j_2, j_3$ .

**Case 1.**  $(m_1, m_2) = (i_3, j_3)$

$$B(i_3, j_3) \leq \sum_{i_1=1}^n \sum_{j_1=1}^n \sum_{i_2=1}^n \sum_{j_2=1}^n \sum_{i_3=1}^n \sum_{j_3=1}^n E|Y(i_1, j_1)Y(i_2, j_2)Y^2(i_3, j_3)| \\ \leq n^4 \sum_{i=1}^n \sum_{j=1}^n EY^4(i, j) \\ \leq n^4 \sqrt{n}\delta_4,$$

where we have used the fact that

$$a_1 a_2 \dots a_k \leq p_1 a_1^{\frac{1}{p_1}} + p_2 a_2^{\frac{1}{p_2}} + \dots + p_k a_k^{\frac{1}{p_k}}$$

for  $a_i, p_i \geq 0$  and  $p_1 + p_2 + \dots + p_k = 1$  in the second inequality.

**Case 2.**  $(m_1, m_2) \neq (i_3, j_3)$

If  $m_1 = i_3$  and  $m_2 \neq j_3$ , then

$$B(i_3, m_2) \leq \sum_{i_1=1}^n \sum_{j_1=1}^n \sum_{i_2=1}^n \sum_{j_2=1}^n \sum_{i_3=1}^n E|Y(i_1, j_1)Y(i_2, j_2)Y(i_3, m_2)||Y(i_3, j_1) + Y(i_3, j_2)| \\ \leq 2n^3 \sum_{i=1}^n \sum_{j=1}^n E|Y^4(i, j)| \\ = 2n^3 \sqrt{n}\delta_4.$$

Similarly, we can show that

$$B(m_1, j_3) \leq 2n^3 \sqrt{n}\delta_4 \text{ and } B(m_1, m_2) \leq 4n^3 \sqrt{n}\delta_4 \text{ for } m_1 \neq i_3 \text{ and } m_2 \neq j_3.$$

From Cases 1 and 2,

$$A_5 = \frac{1}{n(n-1)(n-2)(n-3)} (n^4 \sqrt{n}\delta_4 + 24n^3 \sqrt{n}\delta_4) \\ \leq 2.22 \sqrt{n}\delta_4 \tag{9}$$

for  $n \geq 30$ . Hence, by (4), (5),(7),(8),(9),  $EW^4 \leq 3.44 \sqrt{n}\delta_4$ . □

In 1972, Stein(Stein, C.M., 1972) introduced a powerful and general method to obtain explicit bound for the error in the normal approximation to the distribution of a sum of dependent random variables. His technique was relied on elementary differential equations. For each  $z \in \mathbb{R}$ , the Stein's equation for normal distribution function is

$$g'_w(z) - zg_w(z) = \mathbb{I}(z \leq w) - \Phi(w) \tag{10}$$

where  $w \in \mathbb{R}$ . It is well-known that the solution  $g_w$  of (10) is of the form

$$g_w(z) = \begin{cases} \sqrt{2\pi}e^{\frac{1}{2}z^2}\Phi(z)[1 - \Phi(w)] & , \text{if } z \leq w, \\ \sqrt{2\pi}e^{\frac{1}{2}z^2}\Phi(w)[1 - \Phi(z)] & , \text{if } z > w, \end{cases}$$

with  $0 \leq g_w(z) \leq 1$  for all  $z \in \mathbb{R}$  (11)

and  $|g'_w(z)| \leq 1$  for all  $z \in \mathbb{R}$  (Stein, C.M., 1972, p.22-23). (12)

To apply Stein’s method, we construct the coupling  $\tilde{W}$  of  $W$  as follows.

Let  $I$  and  $K$  be uniformly distributed random variables on  $\{1, \dots, n\}$ ,  $(I, K)$  uniformly distributed on  $\{(i, k) | i, k = 1, \dots, n, i \neq k\}$  and assume that they are independent of  $\pi$  and  $Y(i, j)$ ’s.

Define

$$\tilde{W} = W - S_1 - S_2 + S_3 + S_4$$
 (13)

where

$$S_1 = Y(I, \pi(I)), \quad S_2 = Y(K, \pi(K))$$

$$S_3 = Y(I, \pi(K)), \quad S_4 = Y(K, \pi(I)).$$

Note that  $(\tilde{W}, W)$  is an exchangeable pair (Neammanee, K., 2005, p.261). Clearly,  $S_1$  and  $S_2$  have the same distribution and so do  $S_3$  and  $S_4$ . We observe that

$$P(S_1 \leq a) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n P(Y(i, j) \leq a) = P(S_3 \leq a) \text{ for all } a \in \mathbb{R}.$$

Thus  $S_1, S_2, S_3, S_4$  are identically distributed.

**Lemma 2.2** Let  $g : \mathbb{R} \rightarrow \mathbb{R}$  be a continuous and piecewise continuously differentiable function. Then

$$EWg(W) = E \int_{-\infty}^{\infty} g'(W + t)M(t)dt - \Delta g(W)$$
 (14)

and

$$|\Delta g(W)| \leq \frac{1}{n-1} [Eg^2(W)]^{1/2}$$
 (15)

where

$$M(t) = \frac{n}{4} (\tilde{W} - W) (\mathbb{I}(0 \leq t \leq \tilde{W} - W) - \mathbb{I}(\tilde{W} - W \leq t \leq 0))$$

and  $\mathbb{I}$  is the indicator function.

**Proof:** Let  $\mathcal{A}$  be the  $\sigma$ -algebra generated by

$$\{Y(i, \pi(i)) : 1 \leq i \leq n\}.$$

We can use the idea of (Neammanee, K., 2005) to show that

$$2E\{g(W)E^{\mathcal{A}}(\tilde{W} - W)\} + E(\tilde{W} - W)[g(\tilde{W}) - g(W)] = 0.$$

Let  $\tilde{\mu}(i, \pi(k)) = \mu(i, j)$  for  $\pi(k) = j$ . We observe that for  $i \neq k$

$$E^{\mathcal{A}} \sum_{i=1}^n \sum_{\substack{k \\ k \neq i}} Y(i, \pi(k)) = E^{\mathcal{A}} \sum_{i=1}^n \sum_{\substack{k \\ k \neq i}} \tilde{\mu}(i, \pi(k)).$$
 (16)

Then

$$\begin{aligned}
 E^{\mathcal{A}}(\tilde{W} - W) &= E^{\mathcal{A}}[-S_1 - S_2 + S_3 + S_4] \\
 &= -E^{\mathcal{A}}Y(I, \pi(I)) - E^{\mathcal{A}}Y(K, \pi(K)) + E^{\mathcal{A}}Y(I, \pi(K)) + E^{\mathcal{A}}Y(K, \pi(I)) \\
 &= -\frac{2}{n} \sum_{i=1}^n Y(i, \pi(i)) + \frac{2}{n(n-1)} E^{\mathcal{A}} \sum_{i=1}^n \sum_{\substack{k \\ k \neq i}}^n Y(i, \pi(k)) \\
 &= -\frac{2}{n} W + \frac{2}{n(n-1)} E^{\mathcal{A}} \sum_{i=1}^n \sum_{\substack{k \\ k \neq i}}^n \tilde{\mu}(i, \pi(k)) \\
 &= -\frac{2}{n} W + \frac{2}{n(n-1)} E^{\mathcal{A}} \sum_{i=1}^n \left\{ \sum_{k=1}^n \tilde{\mu}(i, \pi(k)) - \tilde{\mu}(i, \pi(i)) \right\} \\
 &= -\frac{2}{n} W + \frac{2}{n(n-1)} E^{\mathcal{A}} \sum_{i=1}^n \sum_{k=1}^n \tilde{\mu}(i, \pi(k)) - \frac{2}{n(n-1)} E^{\mathcal{A}} \sum_{i=1}^n \tilde{\mu}(i, \pi(i)).
 \end{aligned}$$

We note from (3) that

$$\sum_{i=1}^n \sum_{k=1}^n \tilde{\mu}(i, \pi(k)) = \sum_{i=1}^n \sum_{j=1}^n \mu(i, j) = 0.$$

Thus

$$E^{\mathcal{A}}(\tilde{W} - W) = -\frac{2}{n} W - \frac{2}{n(n-1)} E^{\mathcal{A}} \sum_{i=1}^n \tilde{\mu}(i, \pi(i)).$$

So, we can conclude that

$$0 = 2E\{g(W)\left[-\frac{2}{n} W - \frac{2}{n(n-1)} E^{\mathcal{A}} \sum_{i=1}^n \tilde{\mu}(i, \pi(i))\right]\} + E(\tilde{W} - W)[g(\tilde{W}) - g(W)],$$

which implies

$$\begin{aligned}
 EWg(W) &= \frac{n}{4} E(\tilde{W} - W)[g(\tilde{W}) - g(W)] - \frac{1}{n-1} E[g(W)E^{\mathcal{A}} \sum_{i=1}^n \tilde{\mu}(i, \pi(i))] \\
 &= \frac{n}{4} E(\tilde{W} - W)[g(\tilde{W}) - g(W)] - \Delta g(W) \\
 &= \frac{n}{4} E(\tilde{W} - W) \int_0^{\tilde{W}-W} g'(W+t) dt - \Delta g(W) \\
 &= E \int_{-\infty}^{\infty} g'(W+t)M(t) dt - \Delta g(W)
 \end{aligned}$$

where

$$\begin{aligned}
 |\Delta g(W)| &= \left| \frac{1}{n-1} E[g(W) \sum_{i=1}^n \tilde{\mu}(i, \pi(i))] \right| \\
 &\leq \frac{1}{n-1} [Eg^2(W)]^{1/2} \{E[\sum_{i=1}^n \tilde{\mu}(i, \pi(i))]^2\}^{1/2}.
 \end{aligned}$$

We note from (2) that

$$\begin{aligned}
 1 &= EW^2 \\
 &= E\left[\sum_{i=1}^n Y(i, \pi(i))\right]^2 \\
 &= \sum_{i=1}^n EY^2(i, \pi(i)) + \sum_{i_1=1}^n \sum_{\substack{i_2 \\ i_2 \neq i_1}}^n EY(i_1, \pi(i_1))Y(i_2, \pi(i_2)) \\
 &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n EY^2(i, j) + \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \mu^2(i, j).
 \end{aligned} \tag{17}$$

Thus

$$\begin{aligned}
 & E\left[\sum_{i=1}^n \tilde{\mu}(i, \pi(i))\right]^2 \\
 &= \sum_{i=1}^n E\tilde{\mu}^2(i, \pi(i)) + \sum_{i_1=1}^n \sum_{\substack{i_2 \\ i_2 \neq i_1}}^n E\tilde{\mu}(i_1, \pi(i_1))\tilde{\mu}(i_2, \pi(i_2)) \\
 &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \mu^2(i, j) + \frac{1}{n(n-1)} \sum_{i_1=1}^n \sum_{j_1=1}^n \mu(i_1, j_1) \sum_{\substack{i_2 \\ i_2 \neq i_1}}^n \sum_{\substack{j_2 \\ j_2 \neq j_1}}^n \mu(i_2, j_2) \\
 &\leq \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n EY^2(i, j) + \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \mu^2(i, j) \\
 &= 1.
 \end{aligned} \tag{18}$$

Hence

$$|\Delta g(W)| \leq \frac{1}{n-1} [Eg^2(W)]^{1/2}.$$

□

**Lemma 2.3** Let  $w \in \mathbb{R}$  and  $g_w$  be the unique bounded solution of (10). With the notation and assumptions of Theorem 1.1 and  $\delta_4 \leq 0.047$ , we have that for  $n \geq 30$ ,

$$|Eg'_w(W)E \int_{-\infty}^{\infty} M(t)dt - E[g'_w(W) \int_{-\infty}^{\infty} M(t)dt]| \leq 1.002 \frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + 0.465\delta_2 + \frac{1.612}{\sqrt{n-1}}$$

where  $\delta_2 = \frac{1}{n\sqrt{n}} \sum_{i=1}^n \sum_{j=1}^n EY^2(i, j)$ .

**Proof:** Let  $\mathcal{A}$  be defined as in Lemma 2.2. By the same argument of Lemma 5 in (Loh, W.Y., 1996), we have

$$\begin{aligned}
 & |Eg'_w(W)E \int_{-\infty}^{\infty} M(t)dt - E[g'_w(W) \int_{-\infty}^{\infty} M(t)dt]| \\
 &\leq \frac{1}{4} \sum_{k=1}^4 E|E^{\mathcal{A}}(nS_k^2 - 1)| + \frac{n}{2} \sum_{1 \leq j < k \leq 4} E|E^{\mathcal{A}}(S_j S_k)| + \frac{1}{n-1}.
 \end{aligned} \tag{19}$$

To prove the lemma, it suffices to find appropriate bounds for the terms on the right-hand side of (19). For the sake of clarity, we shall break the proof down into two steps.

Step 1. We will bound the first term on the right-hand side of (19). Hence, by (17)

$$\begin{aligned}
 & [E\{E^{\mathcal{A}}(nS_1^2 - 1)\}]^2 \\
 &= [E\{nE^{\mathcal{A}}Y^2(I, \pi(I)) - 1\}]^2 \\
 &= [E\{\sum_{i=1}^n Y^2(i, \pi(i)) - 1\}]^2 \\
 &\leq E\{\sum_{i=1}^n Y^2(i, \pi(i)) - 1\}^2 \\
 &= E\{\sum_{i=1}^n Y^2(i, \pi(i))\}^2 - 2E\sum_{i=1}^n Y^2(i, \pi(i)) + 1 \\
 &= E\sum_{i=1}^n Y^4(i, \pi(i)) + E\sum_{i=1}^n \sum_{\substack{k \\ k \neq i}} Y^2(i, \pi(i))Y^2(k, \pi(k)) - 2E\sum_{i=1}^n Y^2(i, \pi(i)) + 1 \\
 &= \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^n EY^4(i, k) + \frac{n}{n-1} (\frac{1}{n} \sum_{i=1}^n \sum_{k=1}^n EY^2(i, k))^2 - 2(\frac{1}{n} \sum_{i=1}^n \sum_{k=1}^n EY^2(i, k)) + 1 \\
 &= \frac{\delta_4}{\sqrt{n}} + (1 + \frac{1}{n-1})(1 - \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \mu^2(i, j))^2 \\
 &\quad - 2(1 - \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \mu^2(i, j)) + 1 \\
 &\leq \frac{\delta_4}{\sqrt{n}} + \frac{1}{(n(n-1))^2} [\sum_{i=1}^n \sum_{j=1}^n \mu^2(i, j)]^2 + \frac{1}{(n-1)} + \frac{1}{n^2(n-1)^3} [\sum_{i=1}^n \sum_{j=1}^n \mu^2(i, j)]^2 \\
 &\leq \frac{\delta_4}{\sqrt{n}} + \frac{n}{(n-1)^3} \sum_{i=1}^n \sum_{k=1}^n EY^4(i, k) + \frac{1}{n-1} \\
 &\leq 1.007 \frac{\delta_4}{\sqrt{n}} + \frac{1}{n-1}.
 \end{aligned} \tag{20}$$

By the symmetry of  $S_1$  and  $S_2$ ,

$$[E\{E^{\mathcal{A}}(nS_2^2 - 1)\}]^2 \leq 1.007 \frac{\delta_4}{\sqrt{n}} + \frac{1}{n-1}. \tag{21}$$

By the same technique of (20) we can show that

$$\begin{aligned}
 [E\{E^{\mathcal{A}}(nS_3^2 - 1)\}]^2 &\leq E\left(\frac{1}{n-1} \sum_{i=1}^n \sum_{\substack{k \\ k \neq i}} E^{\mathcal{A}}Y^2(i, \pi(k)) - 1\right)^2 \\
 &\leq 0.041 \frac{\delta_4}{\sqrt{n}} + \frac{2.034}{n-1},
 \end{aligned} \tag{22}$$

and from the symmetry between  $S_3$  and  $S_4$ ,

$$[E\{E^{\mathcal{A}}(nS_4^2 - 1)\}]^2 \leq 0.041 \frac{\delta_4}{\sqrt{n}} + \frac{2.034}{n-1}. \tag{23}$$

Hence we can conclude from (20), (21), (22) and (23) that

$$\frac{1}{4} \sum_{k=1}^4 E\{E^{\mathcal{A}}(nS_k^2 - 1)\} \leq [1.007 \frac{\delta_4}{\sqrt{n}} + \frac{2.034}{n-1}]^{1/2}. \tag{24}$$

Step 2. We will bound the second term on the right-hand side of (19).



Note that

$$\begin{aligned}
 nE|E^{\mathcal{A}}(S_1S_2)| &= \frac{1}{n-1}E|\sum_{i=1}^n \sum_{\substack{k \\ k \neq i}}^n Y(i, \pi(i))Y(k, \pi(k))| \\
 &= \frac{1}{n-1}E|\sum_{i=1}^n Y(i, \pi(i))\{\sum_{k=1}^n Y(k, \pi(k)) - Y(i, \pi(i))\}| \\
 &= \frac{1}{n-1}E|W^2 - \sum_{i=1}^n Y^2(i, \pi(i))| \\
 &\leq \frac{1}{n-1}EW^2 + \frac{1}{n-1} \sum_{i=1}^n EY^2(i, \pi(i)) \\
 &= \frac{1}{n-1} + \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n EY^2(i, j) \\
 &\leq \frac{2}{n-1},
 \end{aligned} \tag{25}$$

where we have used (17) in the last inequality.

By the fact that

$$E^{\mathcal{A}}Y(i, \pi(k)) = E^{\mathcal{A}}\tilde{\mu}(i, \pi(k))$$

and

$$E^{\mathcal{A}}Y(i, \pi(k))Y(k, \pi(i)) = E^{\mathcal{A}}\tilde{\mu}(i, \pi(k))\tilde{\mu}(k, \pi(i))$$

for  $i \neq k$ ,

$$\begin{aligned}
 nE|E^{\mathcal{A}}(S_3S_4)| &= \frac{1}{n-1}E|\sum_{i=1}^n \sum_{\substack{k \\ k \neq i}}^n E^{\mathcal{A}}Y(i, \pi(k))Y(k, \pi(i))| \\
 &= \frac{1}{n-1}E|\sum_{i=1}^n \sum_{\substack{k \\ k \neq i}}^n E^{\mathcal{A}}\tilde{\mu}(i, \pi(k))\tilde{\mu}(k, \pi(i))| \\
 &\leq \frac{1}{n-1}E|\sum_{i=1}^n \sum_{\substack{k \\ k \neq i}}^n \tilde{\mu}(i, \pi(k))\tilde{\mu}(k, \pi(i))| \\
 &\leq \frac{1}{n(n-1)^2} \sum_{j=1}^n \sum_{\substack{l \\ l \neq j}}^n |\sum_{i=1}^n \sum_{\substack{k \\ k \neq i}}^n \mu(i, j)\mu(k, l)| \\
 &= \frac{1}{n(n-1)^2} \sum_{j=1}^n \sum_{\substack{l \\ l \neq j}}^n |\sum_{i=1}^n \mu(i, j)\mu(i, l)| \\
 &\leq \frac{1}{2n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \mu^2(i, j) + \frac{1}{2(n-1)^2} \sum_{i=1}^n \sum_{l=1}^n \mu^2(i, l) \\
 &\leq \frac{\sqrt{n}}{2(n-1)}\delta_2 + \frac{n\sqrt{n}}{2(n-1)^2}\delta_2,
 \end{aligned} \tag{26}$$

and

$$\begin{aligned}
 nE|E^{\mathcal{A}}(S_1S_3)| &= \frac{1}{n-1}E\left|\sum_{i=1}^n\sum_{\substack{k \\ k\neq i}}^n Y(i,\pi(i))E^{\mathcal{A}}Y(i,\pi(k))\right| \\
 &\leq \frac{1}{n-1}E\left|\sum_{i=1}^n\sum_{\substack{k \\ k\neq i}}^n Y(i,\pi(i))\tilde{\mu}(i,\pi(k))\right| \\
 &= \frac{1}{n-1}E\left|\sum_{i=1}^n Y(i,\pi(i))\sum_{\substack{k \\ k\neq i}}^n \tilde{\mu}(i,\pi(k))\right| \\
 &= \frac{1}{n-1}E\left|\sum_{i=1}^n Y(i,\pi(i))\left\{\sum_{k=1}^n \tilde{\mu}(i,\pi(k)) - \tilde{\mu}(i,\pi(i))\right\}\right| \\
 &= \frac{1}{n-1}E\left|\sum_{i=1}^n Y(i,\pi(i))\left\{\sum_{k=1}^n \mu(i,k) - \tilde{\mu}(i,\pi(i))\right\}\right| \\
 &= \frac{1}{n-1}E\left|\sum_{i=1}^n Y(i,\pi(i))\tilde{\mu}(i,\pi(i))\right| \\
 &= \frac{1}{n(n-1)}\sum_{j=1}^n E\left|\sum_{i=1}^n Y(i,j)\mu(i,j)\right| \\
 &\leq \frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1}^n |\mu(i,j)|E|Y(i,j)| \\
 &\leq \frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1}^n EY^2(i,j) \\
 &= \frac{\sqrt{n}}{n-1}\delta_2.
 \end{aligned} \tag{27}$$

By the same argument of (27), we can show that

$$nE|E^{\mathcal{A}}(S_1S_4)| \leq \frac{\sqrt{n}}{n-1}\delta_2, \tag{28}$$

$$nE|E^{\mathcal{A}}(S_2S_3)| \leq \frac{\sqrt{n}}{n-1}\delta_2, \tag{29}$$

and

$$nE|E^{\mathcal{A}}(S_2S_4)| \leq \frac{\sqrt{n}}{n-1}\delta_2. \tag{30}$$

Therefore, by (25)-(30),

$$\frac{n}{2}\sum_{1\leq j < k \leq 4} E|E^{\mathcal{A}}(S_jS_k)| \leq 2.25\frac{\sqrt{n}}{n-1}\delta_2 + \frac{n\sqrt{n}}{4(n-1)^2}\delta_2 + \frac{1}{n-1}.$$

Now we conclude from (19), step 1. and step 2. that

$$\begin{aligned}
 &|Eg'_w(W)E\int_{-\infty}^{\infty} M(t)dt - E[g'_w(W)\int_{-\infty}^{\infty} M(t)dt]| \\
 &\leq [1.007\frac{\delta_4}{\sqrt{n}} + \frac{2.034}{n-1}]^{1/2} + 2.25\frac{\sqrt{n}}{n-1}\delta_2 + \frac{n\sqrt{n}}{4(n-1)^2}\delta_2 + \frac{1}{n-1} \\
 &\leq 1.002\frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + 0.465\delta_2 + \frac{1.612}{\sqrt{n-1}}.
 \end{aligned}$$

This proves the lemma. □

2.2 Proof of Theorem 1.1

**Case 1.**  $\delta_4 \leq 0.047$

We can use the Stein's method and technique from (Laipaporn, K., 2008, p252-254) to show that

$$|P(W \leq w) - \Phi(w)| \leq A_1 + A_2 + A_3 + A_4, \tag{31}$$

where

$$\begin{aligned} A_1 &= |E \int_{-\infty}^{\infty} \{g'_w(W) - g'_w(W+t)\}M(t)dt|, \\ A_2 &= |Eg'_w(W)E \int_{-\infty}^{\infty} M(t)dt - Eg'_w(W) \int_{-\infty}^{\infty} M(t)dt|, \\ A_3 &= |Eg'_w(W)| |1 - E \int_{-\infty}^{\infty} M(t)dt|, \\ A_4 &= \frac{1}{n-1} [Eg_w^2(W)]^{\frac{1}{2}}, \end{aligned}$$

and  $A_1 \leq A_{11} + A_{12} + A_{13}$  (32)

where

$$\begin{aligned} A_{11} &\leq (\frac{3}{2} + \frac{9}{4(n-1)})(E|\tilde{W} - W|^2)^{\frac{1}{2}}, \\ A_{12} &\leq \frac{n}{4} E|W| \frac{|\tilde{W} - W|^3}{2}, \\ A_{13} &\leq \frac{\sqrt{2\pi n}}{16} E|\tilde{W} - W|^3. \end{aligned}$$

From Lemma 2.2, with  $g(w) = w$ , we have

$$\frac{n}{4} E(\tilde{W} - W)^2 \leq EW^2 + \frac{1}{n-1} (EW^2)^{\frac{1}{2}}.$$

Thus

$$E(\tilde{W} - W)^2 \leq \frac{4.137}{n}. \tag{33}$$

Since  $S_1, S_2, S_3$  and  $S_4$  have the same distribution,

$$\begin{aligned} E|\tilde{W} - W|^4 &= E|S_1 + S_2 - S_3 - S_4|^4 \\ &\leq 256E|S_1|^4 \\ &= \frac{256}{n^2} \sum_{i=1}^n \sum_{j=1}^n E|Y(i, j)|^4 \\ &\leq \frac{256\delta_4}{n\sqrt{n}}. \end{aligned} \tag{34}$$

We can conclude from Lemma 2.1, (33) and (34) that

$$\begin{aligned} A_{11} &\leq (\frac{3}{2} + \frac{9}{4(n-1)})(\frac{4.137}{n})^{1/2} \leq \frac{3.21}{\sqrt{n}}, \\ A_{12} &\leq \frac{n}{8} (E|W|^4)^{1/4} (E|\tilde{W} - W|^4)^{3/4} \leq \frac{n}{8} (1.361n^{\frac{1}{8}}\delta_4^{\frac{1}{4}})(\frac{64\delta_4^{3/4}}{n^{9/8}}) \leq 10.881\delta_4, \\ \text{and } A_{13} &\leq \frac{\sqrt{2\pi n}}{16} (E|\tilde{W} - W|^4)^{\frac{3}{4}} \leq \frac{\sqrt{2\pi n}}{16} (\frac{64\delta_4^{\frac{3}{4}}}{n^{\frac{9}{8}}}) = 10.02 \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}, \end{aligned}$$

which implies

$$A_1 \leq \frac{3.21}{\sqrt{n}} + 10.881\delta_4 + 10.02 \frac{\delta_4^{3/4}}{n^{1/8}}. \tag{35}$$

It follows from Lemma 2.3 that

$$A_2 \leq 1.002 \frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + 0.465\delta_2 + \frac{1.612}{\sqrt{n-1}}. \tag{36}$$

Next, we will find a bound for  $A_3$ . From (12), it suffices to bound only

$$|E \int_{-\infty}^{\infty} M(t)dt - 1|.$$

By replacing  $g(W)$  by  $W$  in Lemma 2.2 we have

$$|E \int_{-\infty}^{\infty} M(t)dt - 1| \leq \frac{1}{n-1} (EW^2)^{\frac{1}{2}} = \frac{1}{n-1}.$$

Thus

$$A_3 \leq \frac{1}{n-1}. \tag{37}$$

We note from (11) that

$$A_4 \leq \frac{1}{n-1}. \tag{38}$$

Hence from (35), (36), (37) and (38), we have

$$\begin{aligned} & |P(W \leq w) - \Phi(w)| \\ & \leq 10.881\delta_4 + 10.02 \frac{\delta_4^{3/4}}{n^{1/8}} + 1.002 \frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + 0.465\delta_2 + \frac{5.228}{\sqrt{n}} \\ & \leq 10.881\delta_4 + 10.02 \frac{\delta_4^{3/4}}{n^{1/8}} + 1.467 \frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + \frac{5.228}{\sqrt{n}} \end{aligned}$$

where we have used the fact that

$$\delta_2^2 = \frac{1}{n^3} \left[ \sum_{i_1=1}^n \sum_{i_2=1}^n EY^2(i_1, i_2) \right]^2 \leq \frac{1}{n} \sum_{i_1=1}^n \sum_{i_2=1}^n EY^4(i_1, i_2) = \frac{\delta_4}{\sqrt{n}}$$

in the last inequality.

**Case 2.**  $\delta_4 > 0.047$

Note that for  $W$  such that  $EW = 0$  and  $EW^2 = 1$ ,

$$|P(W \leq w) - \Phi(w)| \leq 0.55$$

(Chen, L.H.Y, 2001). Hence the theorem follows from this fact and the fact that

$$10.881\delta_4 + 10.02 \frac{\delta_4^{3/4}}{n^{1/8}} + 1.467 \frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + \frac{5.228}{\sqrt{n}} > 11.992\delta_4 > 0.55.$$

□

**References**

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

Bolthausen, E. & Gotze, F. (1993). On the rate of convergence for multivariate sampling statistics. *Ann. Statist.*, 21, 1692-1710.

Chen, L.H.Y. & Shao, Q.M. (2001). A non-uniform Berry-Esseen bound via Stein's method. *Probab. Theory Related Fields.*, 120, 236-254.

Hajek, J. (1961). Some extensions of the Wald-Wolfowitz-Neother theorem. *Ann. Math. Statist.*, 15, 358-372.

Ho, S.T. & Chen L.H.Y. (1978). An  $L_p$  bound for the remainder in a combinatorial central limit theorem. *Ann. Probab.*, 6, 231-249.

Hoeffding, W. (1951). A combinatorial central limit theorem. *Ann. Math. Statist.*, 22, 558-566.

- Kolchin, V.F. & Chistyakov, V.P. (1973). On a combinatorial central limit theorem. *Theory. Probability Appl.*, 18, 728-739.
- Laipaporn, K. & Neammanee, K. (2008). A uniform bound on a combinatorial central limit theorem for randomized orthogonal array sampling designs. *Stochastic analysis and applications*, 26, 243-255.
- Loh, W.Y. (1996). A combinatorial central limit theorem for randomized orthogonal array sampling designs. *Annals of Statistics*, 23(3), 1209-1244.
- Matoo, M. (1957). On the Hoeffding's combinatorial central limit theorem. *Ann. Inst. Statist. Math*, 8, 145-154.
- Neammanee, K. & Suntornchost, J. (2005). A uniform bound on a combinatorial central limit theorem. *Stochastic Analysis and Applications.*, 23, 559-578.
- Robinson, J. (1972). A converse to a combinatorial limit theorem. *Ann. Math. Statist.*, 43, 2053-2057.
- Stein, C.M. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. mMath. Statist. Probab.*, 2, 583-602. Univ. California Press, Berkeley.
- Von Bahr, B. (1976). Remainder term estimate in a combinatorial central limit theorem. *Z. Wahrsch. Verw. Gebiete*, 35, 131-139.
- Wald, A. & Wolfowitz, J. (1944). Statistical tests on permutations of observations. *Ann. Math. Statist.*, 15, 358-372.