# ขอบเขตการประมาณค่าด้วยการแจกแจงปกติสำหรับการสุ่มตัวอย่างแบบลาตินไฮเพอร์คิวบ์

นางสาวเพชรรัตน์ รัตนวงษ์



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2550

# BOUNDS ON A NORMAL APPROXIMATION FOR LATIN HYPERCUBE SAMPLING

Miss Petcharat Rattanawong

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Mathematics
Department of Mathematics
Faculty of Science
Chulalongkorn University
Academic Year 2007

| | |
|---|---|
| Thesis Title | : Bounds on a normal approximation for Latin hypercube sampling |
| By | : Miss Petcharat Rattanawong |
| Field of Study | : Mathematics |
| Thesis Advisor | : Professor Kritsana Neammanee, Ph.D. |

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctor's Degree

.............................. Dean of the Faculty of Science
(Professor Supot Hannongbua, Ph.D.)

THESIS COMMITTEE

.............................. Chairman
(Associate Professor Imchit Termwuttipong, Ph.D.)

.............................. Thesis Advisor
(Professor Kritsana Neammanee, Ph.D.)

.............................. Member
(Associate Professor Wicharn Lewkeeratiyutkul, Ph.D.)

.............................. Member
(Kittipat Wong, Ph.D.)

.............................. Member
(Associate Professor Virool Boonyasombat, Ph.D.)

เพชรรัตน์ รัตนวงษ์ : ขอบเขตการประมาณค่าด้วยการแจกแจงปกติสำหรับการสุ่มตัวอย่างแบบลาตินไฮเพอร์คิวบ์( BOUNDS ON A NORMAL APPROXIMATION FOR LATIN HYPERCUBE SAMPLING ) อ.ที่ปรึกษา : ศ. ดร. กฤษณะ เนียมมณี, 79 หน้า

ให้ $X$ เป็นเวกเตอร์สุ่มที่มีการแจกแจงแบบสม่ำเสมอบน $[0,1]^d$ และ $f$ เป็นฟังก์ชันจาก $[0,1]^d$ ไปยัง $\mathbb{R}$ ซึ่งสามารถหาปริพันธ์ได้ วัตถุประสงค์หนึ่งของการทดลองทางคอมพิวเตอร์คือประมาณค่า

$$\mu = Ef(X) = \int_{[0,1]^d} f(x)dx$$

ในบรรดาเทคนิคต่างๆที่ใช้ในการประมาณค่าอินทิเกรต วิธีมอนติคาร์โลเป็นวิธีที่มีประสิทธิภาพและนิยมใช้ในการประมาณค่าอินทิเกรตบนโดเมนที่มีหลายมิติ ตัวประมาณค่าโดยใช้วิธีมอนติ คาร์โล ของ $\mu$ คือ

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

โดยที่ $X_1, X_2, ..., X_n$ เป็นเวกเตอร์สุ่มบน $[0,1]^d$

แม็คเคย์ แบ็คแมน และโคโนเวอร์(ค.ศ. 1979)เสนอการสุ่มตัวอย่าง $X_1, X_2, ..., X_n$ แบบลาตินไฮเพอร์คิวบ์ให้เป็นวิธีหนึ่งในการสุ่มเลือก ในวิทยานิพนธ์ฉบับนี้เราจะหาขอบเขตการประมาณค่าด้วยการแจกแจงปกติสำหรับ $\hat{\mu}_n$ ที่ใช้การสุ่มตัวอย่างแบบลาตินไฮเพอร์คิวบ์




ภาควิชา ...**คณิตศาสตร์**... ลายมือชื่อนิสิต..............................

สาขาวิชา ...**คณิตศาสตร์**... ลายมือชื่ออาจารย์ที่ปรึกษา..............................

ปีการศึกษา ......2550...... ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.........-.........

## # 4773825223 : MAJOR MATHEMATICS
KEY WORDS : STEIN'S METHOD / LATIN HYPERCUBE SAMPLING / COMBINATORIAL CENTRAL LIMIT THEOREM

PETCHARAT RATTANAWONG : BOUNDS ON A NORMAL APPROXIMATION FOR LATIN HYPERCUBE SAMPLING. THESIS ADVISOR : PROF. KRITSANA NEAMMANEE, Ph.D., 79 pp.

Let $X$ be a random vector uniformly distributed on $[0,1]^d$ and let $f : [0,1]^d \to \mathbb{R}$ be an integrable function. An objective of many computer experiments is to estimate

$$\mu = E(f(X)) = \int_{[0,1]^d} f(x)dx.$$

Among numerical integration techniques, Monte Carlo methods are efficient and competitive for high-dimensional integration. The Monte Carlo's estimator for the integral $\mu$ is given by

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

where $X_1, X_2, ..., X_n$ are random vectors on $[0,1]^d$.

McKay, Beckman and Conover (1979) introduced Latin hypercube sampling(LHS) as an alternative method of generating $X_1, X_2, ..., X_n$. In this work, we investigate normal approximation of error bounds in the distribution of $\hat{\mu}_n$ based on a Latin hypercube sampling.

| | | | |
|---|---|---|---|
| Department | ...**Mathematics**... | Student's signature | Petcharat Rattanawong |
| Field of study | ...**Mathematics**... | Advisor's signature | Kritsana Neammanee |
| Academic year | .........**2007**............ | Co-advisor's signature | – |

# ACKNOWLEDGEMENTS

I am greatly indepted to Professor Dr. Kritsana Neammanee, my thesis advisor, for his willingness to sacrifice his time to suggest and advise me in preparing and writing this thesis. I would like to thank Associate Professor Dr. Imchit Termwuttipong, Associate Professor Dr. Wicharn Lewkeeratiyutkul, Dr. Kittipat Wong and Associate Professor Dr. Virool Boonyasombat, my thesis committee, for their suggestions to this thesis. I am also sincerely grateful to all of my teachers for my knowledge and skill.

In particular, thank you my dear friends for giving me good experiences at Chulalongkorn university.

Finally, I would like to express my deep gratitude to my beloved family for their love and encouragement throughout my graduate study.

# CONTENTS

# CHAPTER I
# INTRODUCTION

In many problems multi-dimensional integrals occur, which have to be evaluated numerically. We consider an integral over the $d$-dimensional hypercube $[0,1]^d$;

$$\mu = \int_{[0,1]^d} f(x)dx.$$

This is equivalent to finding $E(f(X))$, where the random vector $X$ has a uniform distribution on the unit hypercube $[0,1]^d$.

Among numerical integration techniques, Monte Carlo methods are especially useful and competitive for high-dimensional integration. The Monte Carlo estimate for the integral $\mu = E(f(X)) = \int_{[0,1]^d} f(x)dx$ is given by

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

where $X_1, ..., X_n$ is a random sampling on $[0,1]^d$. There are various alternative ways to select the points $X_i'$s. For examples, simple random sampling, lattice sampling(see Patterson[22]), Latin hypercube sampling(see McKay, Beckman, Conover[16], Owen[18], Loh[14]), the orthogonal arrays(see Loh[13], Owen[20], Tang[28], Laipaporn and Neammanee[12]) and scrambled net(see Owen[21], and [19]).

In 1979, McKay, Beckman and Conover[16] proposed Latin hypercube sampling(LHS) as an attractive alternative to generate $X_1, ..., X_n$. The main feature of Latin hypercube sampling is that, in contrast to simple random sampling, it simultaneously stratifies on all input dimensions. More precisely, let:

1. $\eta_k,\ 1 \le k \le d$ be random permutations of $\{1, ..., n\}$ each uniformly distributed over all the $n!$ possible permutations;
2. $U_{i_1,...,i_d,j}\ 1 \le i_1, ..., i_d \le n,\ 1 \le j \le d$, be $[0,1]$ uniform random variables;
3. the $U_{i_1,...,i_d,j}$'s and $\eta_k$'s all be stochastically independent.

A Latin hypercube sample of size $n$ (taken from the $d$-dimensional hypercube $[0,1]^d$) is defined to be $\{X(\eta_1(i), \eta_2(i), ..., \eta_d(i)) : 1 \le i \le n\}$, where for all $1 \le i_1, ..., i_d \le n$,

$1 \le j \le d$,

$$X_j(i_1, ..., i_d) = (i_j - U_{i_1,...,i_d,j})/n,$$

and $$X(i_1, ..., i_d) = (X_1(i_1, ..., i_d), ..., X_d(i_1, ..., i_d)).$$

Hence the estimator for $\mu$ that based on a Latin hypercube sampling is

$$\hat{\mu}_n = \frac{1}{n}\sum_{k=1}^{n} f(X(\eta_1(k), \eta_2(k), ..., \eta_d(k))).$$

Then $\hat{\mu}_n$ is an unbiased estimator for $\mu$.
Assume that $Var(\hat{\mu}_n) > 0$, we define

$$L = \frac{\hat{\mu}_n - \mu}{\sqrt{Var(\hat{\mu}_n)}}. \tag{1.1}$$

In 1996, Loh[14] used Stein's method to show that the distribution of $L$ can be approximated by normal distribution and gave a uniform bound on this approximation. The following is his result.

**Theorem 1.1.** *For $1 \le i_1, ..., i_d \le n$, let*

$$\mu(i_1, ..., i_d) = E(f(X(i_1, ..., i_d))),$$

$$\mu_{-k}(i_k) = \frac{1}{n^{d-1}}\sum_{j \ne k}\sum_{i_j=1}^{n} \mu(i_1, ..., i_d)$$

*and*

$$V(i_1, ..., i_d) = \frac{1}{n\sqrt{Var\hat{\mu}_n}}[f(X(i_1, ..., i_d)) - \sum_{k=1}^{d} \mu_{-k}(i_k) + (d-1)\mu].$$

*Then there exists a positive constant $C_d$ which depends only on d such that for sufficiently large $n$,*

$$\sup_{z\in\mathbb{R}} |P(L \le z) - \Phi(z)| \le C_d\beta_3$$

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

$$\beta_3 = \frac{1}{n^{d-1}}\sum_{i_1=1}^{n} \cdots \sum_{i_d=1}^{n} E|V(i_1, ..., i_d)|^3.$$

The following is an immediate corollary.

**Corollary 1.2.** *If $E|f(X)|^3 < \infty$, then*

$$\sup_{z\in\mathbb{R}} |P(L \le z) - \Phi(z)| \le \frac{C_d}{\sqrt{n}}.$$

In this work, we give a constant $C_d$ by using a concentration inequality approach of Stein's method and give a non-uniform bound of this approximation. These are our results.

**Theorem 1.3.** *(A uniform bound for LHS) For $1 \leq i_1, ..., i_d \leq n$, let*

$$\mu(i_1, ..., i_d) = E(f(X(i_1, ..., i_d))),$$
$$\mu_k(i_1, ..., i_d) = \frac{(-1)^k}{n^k} \sum_{1 \leq j_1 < j_2 < ... < j_k \leq d} \sum_{q_{j_1}=1}^{n} ... \sum_{q_{j_k}=1}^{n} \mu(l_1, ..., l_d),$$

*where*

$$l_p = \begin{cases} q_p & \text{if } p = j_1, ..., j_k, \\ i_p & \text{otherwise,} \end{cases}$$

*and*

$$V(i_1, ..., i_d) = \frac{1}{n\sqrt{Var(\hat{\mu}_n)}}[f(X(i_1, ..., i_d)) + \sum_{k=1}^{d-1} \mu_k(i_1, ..., i_d) + (-1)^d \mu].$$

*Then for $n \geq 6^d + 3$,*

$$\sup_{z \in \mathbb{R}} |P(L \leq z) - \Phi(z)| \leq \frac{11.946}{\sqrt{n}} + \frac{1.037\sqrt{d}\beta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + 8.314 d^{\frac{1}{4}} \beta_4 + 11.765\beta_4 + \frac{5.014 d \beta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}} + \frac{2\sqrt{2\pi}\beta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}$$

*where*

$$\beta_4 = \frac{1}{n^{d-\frac{3}{2}}} \sum_{i_1=1}^{n} ... \sum_{i_d=1}^{n} E|V(i_1, ..., i_d)|^4.$$

**Corollary 1.4.** *If $E|f(X)|^4 < \infty$, then*

$$\sup_{z \in \mathbb{R}} |P(L \leq z) - \Phi(z)| \leq \frac{C_d(28.725 + 1.037\sqrt{d} + 8.314 d^{\frac{1}{4}} + 5.014d)}{\sqrt{n}}.$$

*where $C_d = \frac{27}{C^2}[2 + (d-1)^3 \sum_{k=1}^{d-1} \binom{d}{k}]E|f(X)|^4$ for some constant $C$.*

**Theorem 1.5.** *(A non-uniform bound for LHS) Let $z \in \mathbb{R}$. Then there exists a positive constant $C$ which does not depend on $z$ such that*

$$\left|P(L \leq z) - \Phi(z)\right| \leq \frac{C}{1+|z|} \left\{ \frac{\beta_8^{\frac{1}{8}}}{n^{\frac{7}{16}}} + \frac{1}{n} + \beta_8 \right\}$$

*where*

$$\beta_8 = \frac{1}{n^{d-\frac{7}{2}}} \sum_{i_1=1}^{n} ... \sum_{i_d=1}^{n} E|V(i_1, ..., i_d)|^8.$$

**Corollary 1.6.** *If* $E|f(X)|^8 < \infty$, *then*

$$\left|P(L \leq z) - \Phi(z)\right| \leq \frac{C}{(1+|z|)\sqrt{n}}.$$

In order to prove Theorem 1.3 and Theorem 1.5, we shall first show that

$$L = \sum_{i=1}^{n} Y(i, \pi_1(i), ..., \pi_{d-1}(i)) \tag{1.2}$$

where $Y(i_1, ..., i_d)$, $1 \leq i_1, ..., i_d \leq n$, are random variables and $\pi_1, ..., \pi_{d-1}$ are random permutations of {1,...,n} such that $Y(i_1, ..., i_d)$'s and $\pi_k$'s are stochastically independent. In the case of $d = 1$, a theorem which has been proved under various conditions by Hoeffding([10]), Matoo([15]) and other authors states that the summation on the right-hand side of (1.2) is approximately standard normally distributed. This theorem is always called a combinatorial central limit theorem(CCLT). Thus from (1.2), it suffices to find a uniform bound and a non-uniform bound on normal approximation for the generalization of CCLT, the summation in the case of an arbitary $d$. These are our theorems.

**Theorem 1.7.** *(A uniform bound for the generalization of a CCLT)*

*Let $d$ be a positive integer and let:*

*1. $\pi_k$, $1 \leq k \leq d$, be random permutations of $\{1, ..., n\}$ each uniformly distributed over all the $n!$ possible permutations;*

*2. $Y(i_1, ..., i_{d+1})$, $1 \leq i_1, ..., i_{d+1} \leq n$, be random variables;*

*3. the $Y(i_1, ..., i_{d+1})$'s and $\pi_k$'s all be stochastically independent.*

*We define*

$$W = \sum_{i=1}^{n} Y(i, \pi_1(i), ..., \pi_d(i)). \tag{1.3}$$

*Suppose that*

$$EW = 0 \quad and \quad VarW = 1, \tag{1.4}$$

*and*

$$\sum_{i_j=1}^{n} \tilde{\mu}(i_1, ..., i_{d+1}) = 0 \quad for\ every \quad j \in \{1, ..., d+1\}$$

*where* $\tilde{\mu}(i_1, ..., i_{d+1}) = EY(i_1, ..., i_{d+1})$. *Then for* $n \geq 6^{d+1} + 3$,

$$\sup_{z\in\mathbb{R}} |P(W \leq z) - \Phi(z)| \leq \frac{11.946}{\sqrt{n}} + \frac{1.037\sqrt{d+1}\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + 8.314(d+1)^{\frac{1}{4}}\delta_4 + 11.765\delta_4 + 5.014(d+1)\frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}} + \frac{2\sqrt{2\pi}\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}$$

*where*

$$\delta_4 = \frac{1}{n^{d-\frac{1}{2}}} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} E|Y(i_1, ..., i_{d+1})|^4.$$

*Furthermore, if* $\delta_4 \sim n^{-1/2}$, *then*

$$\sup_{z\in\mathbb{R}} |P(W \leq z) - \Phi(z)| \leq \frac{28.725 + 1.037\sqrt{d+1} + 8.314(d+1)^{\frac{1}{4}} + 5.014(d+1)}{\sqrt{n}}. \quad (1.5)$$

Note that in the special case of $d = 1$, Neammanee and Suntornchost[17] yields the bound $\frac{216}{\sqrt{n}}$ while Theorem 1.7 gave a constant $C = 50.106$.

**Theorem 1.8.** *(A non-uniform bound for the generalization of a CCLT) Let* $z \in \mathbb{R}$. *With the notation and assumptions of Theorem 1.7, there exists a positive constant* $C$ *which does not depend on* $z$ *such that*

$$\left|P(W \leq z) - \Phi(z)\right| \leq \frac{C}{1+|z|}\left\{\frac{\delta_8^{\frac{1}{8}}}{n^{\frac{7}{16}}} + \frac{1}{n} + \delta_8\right\}$$

*where*

$$\delta_8 = \frac{1}{n^{d-\frac{5}{2}}} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} E|Y(i_1, ..., i_{d+1})|^8.$$

*Furthermore, if* $\delta_8 \sim n^{-1/2}$,

$$\left|P(W \leq z) - \Phi(z)\right| \leq \frac{C}{(1+|z|)\sqrt{n}}. \quad (1.6)$$

We observe that Corollary 1.4 and Corollary 1.6 are examples of (1.5) and (1.6), respectively.

In this thesis, we organize as follows. In chapter 2, we give some definitions in elementary probability theory, a background of Stein's method and some useful properties of Stein's solution. A proof of Theorem 1.5 is in chapter 3 while a proof of Theorem 1.6 is stated in chapter 4. A uniform bound and non-uniform bound for Latin hypercube sampling are proved in chapter 5 by applying Theorem 1.5 and Theorem 1.6 respectively.

# CHAPTER II
# PRELIMINARIES

In this chapter, we give some basic concepts in probability which will be used in our work and the idea of Stein's method. The proof is omited but can be found in [[1], [2], [23], [26]].

## 2.1 Probability Space and Random Variables

A **probability space** is a measure space $(\Omega, \mathcal{F}, P)$ for which $P(\Omega) = 1$. The measure $P$ is called a **probability measure.** The set $\Omega$ will be referred as a **sample space** and its elements are called **points** or **elementary events.** The elements of $\mathcal{F}$ are called **events.** For any event $A$, the value $P(A)$ is called the **probability of** $A$**.**

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A function $X : \Omega \to \mathbb{R}$ is called a **random variable** if for every Borel set $B$ in $\mathbb{R}$, $X^{-1}(B)$ belongs to $\mathcal{F}$. We shall use the notation $P(X \in B)$ in place of $P(\{\omega \in \Omega \mid X(\omega) \in B\})$. In the case where $B = (-\infty, a]$ or $[a, b]$, $P(X \in B)$ is denoted by $P(X \leq a)$ or $P(a \leq X \leq b)$, respectively.

Let $X$ be a random variable. A function $F : \mathbb{R} \to [0, 1]$ which is defined by

$$F(x) = P(X \leq x)$$

is called the **distribution function** of $X$.

Let $X$ be a random variable with the distribution function $F$. Then $X$ is said to be a **discrete random variable** if the image of $X$ is countable and $X$ is called a **continuous random variable** if $F$ can be written in the form

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

for some nonnegative integrable function $f$ on $\mathbb{R}$. In this case, we say that $f$ is the **probability density function** of $X$.

Now we will give some examples of random variables.

We say that $X$ is a **normal** random variable with parameter $\mu$ and $\sigma^2$, written as $X \sim N(\mu, \sigma^2)$, if its probability function is defined by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big).$$

Moreover, if $X \sim N(0,1)$ then $X$ is said to be a **standard normal** random variable.

We say that $X$ is a **uniform** random variable with parameter $n$ if there exist $x_1, x_2, ..., x_n$ such that $P(X = x_i) = \dfrac{1}{n}$ for any $i = 1, 2, ..., n$ and denoted by $X \sim U(n)$.

Furthermore, we say that $\pi$ is a **random permutation** if $\pi$ is a permutation-valued random variable.

## 2.2 Independence

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{F}_\alpha$ a sub $\sigma$-algebra of $\mathcal{F}$ for each $\alpha \in \Lambda$. We say that $\{\mathcal{F}_\alpha | \alpha \in \Lambda\}$ is **independent** if and only if for any subset $J = \{1, 2, ..., k\}$ of $\Lambda$,

$$P(\bigcap_{m=1}^{k} A_m) = \prod_{m=1}^{k} P(A_m)$$

where $A_m \in \mathcal{F}_m$ for $m = 1, ..., k$.

Let $\mathcal{E}_\alpha \subseteq \mathcal{F}$ for all $\alpha \in \Lambda$. We say that $\{\mathcal{E}_\alpha \mid \alpha \in \Lambda\}$ is **independent** if and only if $\{\sigma(\mathcal{E}_\alpha) \mid \alpha \in \Lambda\}$ is independent where $\sigma(\mathcal{E}_\alpha)$ is the smallest $\sigma$-algebra with $\mathcal{E}_\alpha \subseteq \sigma(\mathcal{E}_\alpha)$ .

We say that the set of random variables $\{X_\alpha \mid \alpha \in \Lambda\}$ is **independent** if $\{\sigma(X_\alpha) \mid \alpha \in \Lambda\}$ is independent, where $\sigma(X) = \{X^{-1}(B) \mid B$ is a Borel subset of $\mathbb{R}\}$.

**Theorem 2.1.** *Random variables $X_1, X_2, ..., X_n$ are* **independent** *if for any Borel sets $B_1, B_2, ..., B_n$ we have*

$$P(\bigcap_{i=1}^{n}\{X_i \in B_i\}) = \prod_{i=1}^{n} P(X_i \in B_i).$$

**Proposition 2.2.** *If $X_{ij}$ ; $i = 1, 2, ..., n,\ j = 1, 2, ..., m_i$ are independent and $f_i : \mathbb{R}^{m_i} \to \mathbb{R}$ are measurable, then $\{f_i(X_{i1}, X_{i2}, ..., X_{im_i}) \mid i = 1, 2, ..., n\}$ is independent.*

## 2.3 Expectation, Variance and Conditional Expectation

Let $X$ be any random variable on a probability space $(\Omega, \mathcal{F}, P)$.
If $\int_\Omega |X| dP < \infty$, then we define its **expected value** to be

$$E(X) = \int_\Omega X dP.$$

**Proposition 2.3.**

1. *If $X$ is a discrete random variable, then* $E(X) = \sum_{x \in ImX} xP(X = x)$.

2. *If $X$ is a continuous random variable with probability function $f$, then*

$$E(X) = \int_{\mathbb{R}} xf(x)dx.$$

**Proposition 2.4.** *Let $X$ and $Y$ be random variables such that $E(|X|) < \infty$ and $E(|Y|) < \infty$ and $a, b \in R$. Then we have the followings:*

1. $E(aX + bY) = aE(X) + bE(Y)$.

2. *If $X \leq Y$ , then $E(X) \leq E(Y)$.*

3. $|E(X)| \leq E(|X|)$.

4. *If $X$ and $Y$ are independent , then $E(XY) = E(X)E(Y)$.*

Let $X$ be a random variable which $E(|X|^k) < \infty$. Then $E(|X|^k)$ is called the **$k$-th moment** of $X$ about the origin and call $E[(X - E(X))^k]$ the **$k$-th moment** of $X$ about the mean.

We call the second moment of $X$ about the mean, the **variance** of $X$, denoted by $Var(X)$. Then

$$Var(X) = E[X - E(X)]^2.$$

We note that

1. $Var(X) \;=\; E(X^2) - E^2(X)$.

2. If $X \sim N(\mu, \sigma^2)$ then $E(X) = \mu$ and $Var(X) = \sigma^2$.

**Proposition 2.5.** *If $X_1, ..., X_n$ are independent and $E|X_i| < \infty$ for $i = 1, 2, ..., n$, then*

1. $E(X_1X_2...X_n) = E(X_1)E(X_2)...E(X_n)$,

2. $Var(a_1X_1 + \cdots + a_nX_n) \;=\; a_1^2Var(X_1) + \cdots + a_n^2Var(X_n)$ *for any real number* $a_1, ..., a_n$.

**Theorem 2.6.** *Let $(X_n)$ be an increasing sequence of randon variables on a probability space $(\Omega, \mathcal{F}, P)$ to $[0, \infty)$ and $\lim_{n\to\infty} X_n = X$ a.s. Then $\lim_{n\to\infty} E(X_n) = EX$.*

The following inequalities are useful in our work.

1. **Hölder's inequality :**

$$E(|XY|) \leq E^{\frac{1}{p}}(|X|^p)E^{\frac{1}{q}}(|Y|^q)$$

where $0 < p, q < 1$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$ and $E(|X|^p) < \infty$,$E(|Y|^q) < \infty$.

2. **Chebyshev's inequality :**

$$P(\{|X - E(X)| \geq \varepsilon\}) \leq \frac{Var(X)}{\varepsilon^2} \quad \text{for all} \quad \varepsilon > 0$$

where $E(X^2) < \infty$.

Let $X$ be a random variable with finite expected value on a probability space $(\Omega, \mathcal{F}, P)$ and $\mathcal{D}$ a sub $\sigma$-algebra of $\mathcal{F}$. Define a probability measure $P_{\mathcal{D}} : \mathcal{D} \to [0, 1]$ by

$$P_{\mathcal{D}}(E) = P(E)$$

and a signed measure $\mathcal{Q}_X : \mathcal{D} \to \mathbb{R}$ by

$$\mathcal{Q}_X(E) = \int_E X dP.$$

Then, by Radon-Nikodym theorem we have $\mathcal{Q}_X \ll P_{\mathcal{D}}$ and there exists a unique measurable function $E^{\mathcal{D}}(X)$ on $(\Omega, \mathcal{F}, P)$ such that

$$\int_E E^{\mathcal{D}}(X) dP_{\mathcal{D}} = \mathcal{Q}_X(E) = \int_E X dP \quad \text{for any } E \in \mathcal{D}.$$

We will say that $E^{\mathcal{D}}(X)$ is the **conditional expectation** of $X$ with respect to $\mathcal{D}$.

Moreover, for any random variables $X$ and $Y$ on the same probability space $(\Omega, \mathcal{F}, P)$ such that $E(|X|) < \infty$, we will denote $E^{\sigma(Y)}(X)$ by $E^Y(X)$.

**Theorem 2.7.** *Let $X$ be a random variable on a probability space $(\Omega, \mathcal{F}, P)$ such that $E(|X|) < \infty$, then the followings hold for any sub $\sigma$-algebra $\mathcal{D}$ of $\mathcal{F}$.*

1. *If $X$ is a random variable on $(\Omega, \mathcal{D}, P_{\mathcal{D}})$, then $E^{\mathcal{D}}(X) = X$ a.s.$[P_{\mathcal{D}}]$.*

2. *$E^{\mathcal{F}}(X) = X$ a.s.$[P]$.*

3. *If $\sigma(X)$ and $\mathcal{D}$ are independent, then $E^{\mathcal{D}}(X) = E(X)$ a.s.$[P_{\mathcal{D}}]$.*

**Theorem 2.8.** *Let $X$ and $Y$ be random variables on the same probability space $(\Omega, \mathcal{F}, P)$ such that $E(|X|)$ and $E(|Y|)$ are finite. Then for any sub $\sigma$-algebra $\mathcal{D}$ of $\mathcal{F}$ the followings hold.*

1. *If $X \leq Y$, then $E^{\mathcal{D}}(X) \leq E^{\mathcal{D}}(Y)$ a.s. $[P_{\mathcal{D}}]$.*
2. *$E^{\mathcal{D}}(aX + bY) = aE^{\mathcal{D}}(X) + bE^{\mathcal{D}}(X)$ a.s. $[P_{\mathcal{D}}]$ for any $a,\ b \in \mathbb{R}$.*
3. *$|E^{\mathcal{D}}(X)| \leq E^{\mathcal{D}}(|X|)$ a.s. $[P_{\mathcal{D}}]$.*

**Theorem 2.9.** *Let $X$ and $Y$ be random variables on the same probability space $(\Omega, \mathcal{F}, P)$ such that $E(|XY|)$ and $E(|Y|)$ are finite and $\mathcal{D}_1, \mathcal{D}_2$ be any sub $\sigma$-algebras of $\mathcal{F}$. If $X$ is a random variable with respect to $\mathcal{D}_1$, then*

1. *$E^{\mathcal{D}_1}(XY) = XE^{\mathcal{D}_1}(Y)$ a.s. $[P_{\mathcal{D}_1}]$,*
2. *$E^{\mathcal{D}_2}(XY) = E^{\mathcal{D}_2}(XE^{\mathcal{D}_1}(Y))$ a.s. $[P_{\mathcal{D}_2}]$.*

**Theorem 2.10.** *Let $(X_n)$ are increasing sequence of non-negative random variables on the same probability space $(\Omega, \mathcal{F}, P)$. If $X_n \to X$ a.s. and $E|X| \leq \infty$, then $0 \leq \lim\limits_{n\to\infty} E^{\mathcal{D}}X_n = E^{\mathcal{D}}X$ a.s. $[P_{\mathcal{D}}]$ for any sub $\sigma$-algebra of $\mathcal{F}$.*

**Theorem 2.11.** *Let $X$ be a random variable on the probability space $(\Omega, \mathcal{F}, P)$ such that $E(|X|)$ is finite and $\mathcal{D}_1, \mathcal{D}_2$ be any sub $\sigma$-algebra of $\mathcal{F}$. If $\sigma\{X, \mathcal{D}_1\}$ is independent of $\mathcal{D}_2$, then*

$$E^{\sigma\{\mathcal{D}_1, \mathcal{D}_2\}}(X) = E^{\mathcal{D}_1}(X) \text{ a.s. } [P].$$

## 2.4 Conditional Independence

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{D}$ a sub $\sigma$-algebra of $\mathcal{F}$. For any event $A$ on $\mathcal{F}$, we define the **conditional probability of $A$ given $\mathcal{D}$** by

$$P(A|\mathcal{D}) = E^{\mathcal{D}}\mathbb{I}(A)$$

where $\mathbb{I}(A)$ is an indicator function.

Let $X, Y$ and $Z$ be discrete random variables on a probability space $(\Omega, \mathcal{F}, P)$. If $P(X \leq x, Y \leq y \mid Z = z) = P(X \leq x \mid Z = z)P(Y \leq y \mid Z = z)$ for all $x, y, z \in \mathbb{R}$ such that $P(Z = z) > 0$, then we say that $X$ and $Y$ are **conditionally independent** given $Z$.

**Proposition 2.12.** *Let $X, Y$ and $Z$ be discrete random variables on a probability space $(\Omega, \mathcal{F}, P)$ such that $X$ and $Y$ are conditionally independent given $Z$. Then the followings hold.*

1. *$X^+$ and $Y^+$ are conditionally independent given $Z$,*
2. *$X^+$ and $Y^-$ are conditionally independent given $Z$,*
3. *$X^-$ and $Y^+$ are conditionally independent given $Z$,*
4. *$X^-$ and $Y^-$ are conditionally independent given $Z$*

*where for any $f : \Omega \to [-\infty, \infty]$, we defined the* **positive part** *of $f$ by $f^+ = \max\{f, 0\}$ and the* **negative part** *of $f$ by $f^- = -\min\{f, 0\}$.*

*Proof.* Assume that $X$ and $Y$ are conditionally independent given $Z$. Let $x, y, z \in \mathbb{R}$ be such that $P(Z = z) > 0$.
**case 1.** $\forall w \in \Omega, X(w) \geq 0$ or $\forall w \in \Omega, Y(w) \geq 0$.
WLOG, we assume that $\forall w \in \Omega, X(w) \geq 0$. If $y < 0$, then

$$P(X^+ \leq x, Y^+ \leq y \mid Z = z) = 0 = P(X^+ \leq x \mid Z = z)P(Y^+ \leq y \mid Z = z).$$

Let $y \geq 0$. We observe that

$$\begin{aligned} P(X^+ \leq x, Y^+ \leq y \mid Z = z) &= P(X \leq x, Y \leq y \mid Z = z) \\ &= P(X \leq x \mid Z = z)P(Y \leq y \mid Z = z) \\ &= P(X^+ \leq x \mid Z = z)P(Y^+ \leq y \mid Z = z). \end{aligned}$$

**case 2.** $\exists w \in \Omega, \quad X(w) < 0$ and $\exists w \in \Omega, Y(w) < 0$.
If $x < 0$ or $y < 0$,

$$P(X^+ \leq x, Y^+ \leq y \mid Z = z) = 0 = P(X^+ \leq x \mid Z = z)P(Y^+ \leq y \mid Z = z).$$

If $x \geq 0$ and $y \geq 0$, then

$$\begin{aligned} P(X^+ \leq x, Y^+ \leq y \mid Z = z) &= P(X \leq x, Y \leq y \mid Z = z) \\ &= P(X \leq x \mid Z = z)P(Y \leq y \mid Z = z) \\ &= P(X^+ \leq x \mid Z = z)P(Y^+ \leq y \mid Z = z). \end{aligned}$$

Thus, we prove 1. Similarly, we have 2,3 and 4. $\square$

**Proposition 2.13.** *Let $X, Y$ and $Z$ be discrete random variables on a probability space $(\Omega, \mathcal{F}, P)$ such that $X$ and $Y$ are conditionally independent given $Z$. Then*

$$EXY = E(E^Z X E^Z Y).$$

*Proof.* Assume that $X$ and $Y$ are conditionally independent given $Z$.

**Step 1.** $X$ and $Y$ are non-negative simple functions.

Let $X = \sum_{i=1}^{n} \alpha_i \mathbb{I}(E_i)$ and $Y = \sum_{j=1}^{m} \beta_j \mathbb{I}(F_j)$ where $\{E_1, E_2, ..., E_n\}$ and $\{F_1, F_2, ..., F_m\}$ are partitions of $\Omega$. Thus $XY = \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_i\beta_j \mathbb{I}(E_i \cap F_j)$ and $\{E_i \cap F_j, | i = 1, 2, ..., n$ and $j = 1, 2, ..., m\}$ is a partition of $\Omega$. Since $X$ and $Y$ are conditionally independent given $Z$, $P(X = x, Y = y \mid Z = z) = P(X = x \mid Z = z)P(Y = y \mid Z = z)$ for all $x, y, z \in \mathbb{R}$ such that $P(Z \leq z) > 0$. This implies

$$\begin{aligned}
P(E_i \cap F_j | Z) &= \sum_{k \in ImZ} P(E_i \cap F_j \mid Z = k)\mathbb{I}(Z = k) \\
&= \sum_{k \in ImZ} P(X = \alpha_i, Y = \beta_j \mid Z = k)\mathbb{I}(Z = k) \\
&= \sum_{k \in ImZ} P(X = \alpha_i \mid Z = k)P(Y = \beta_j \mid Z = k)\mathbb{I}(Z = k) \\
&= [\sum_{k \in ImZ} P(X = \alpha_i \mid Z = k)\mathbb{I}(Z = k)][\sum_{l \in ImZ} P(Y = \beta_j \mid Z = l)\mathbb{I}(Z = l)] \\
&= P(E_i|Z)P(F_j|Z)
\end{aligned}$$

for every $i = 1, 2, ..., n$ and $j = 1, 2, ..., m$.

Then

$$\begin{aligned}
EXY &= \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_i\beta_j E E^Z \mathbb{I}(E_i \cap F_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_i\beta_j E P(E_i \cap F_j | Z) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_i\beta_j E P(E_i|Z)P(F_j|Z) \\
&= E(\sum_{i=1}^{n} \alpha_i E^Z \mathbb{I}(E_i))(\sum_{j=1}^{m} \beta_j E^Z \mathbb{I}(F_j)) \\
&= E(E^Z X E^Z Y).
\end{aligned}$$

**Step 2.** $X$ and $Y$ are non-negative random variables.

Let $(X_n)$ and $(Y_n)$ are increasing sequences of non-negative simple fuctions such that $X_n \to X$ and $Y_n \to Y$. From step 1, Theorem 2.6 and Theorem 2.10, we have

$$EXY = \lim_{n\to\infty} E(X_nY_n) = \lim_{n\to\infty} E(E^ZX_nE^ZY_n) = E(E^ZXE^ZY).$$

**Step 3.** $X$ and $Y$ are random variables.

It follows from Proposition 2.12 and step 2 that

$$E(X^+Y^+) = E(E^ZX^+E^ZY^+), \quad E(X^+Y^-) = E(E^ZX^+E^ZY^-)$$
$$E(X^-Y^+) = E(E^ZX^-E^ZY^+), \quad E(X^-Y^-) = E(E^ZX^-E^ZY^-).$$

Thus

$$\begin{aligned} EXY &= E[(X^+ - X^-)(Y^+ - Y^-)] \\ &= E[X^+Y^+ - X^+Y^- - X^-Y^+ + X^-Y^-] \\ &= E(E^ZX^+E^ZY^+) - E(E^ZX^+E^ZY^-) - E(E^ZX^-E^ZY^+) + E(E^ZX^-E^ZY^-) \\ &= E[E^Z(X^+ - X^-)E^Z(Y^+ - Y^-)] \\ &= E(E^ZXE^ZY). \end{aligned}$$

□

## 2.5 Stein's Method for Normal Approximation

In 1972, Stein[25] introduced a powerful and general method for obtaining an explicit bound for the error in the normal approximation to the distribution of a sum of dependent random variables. The technique used was novel. Stein's technique is free of Fourier methods and relied instead on the elementary differential equation. This method was adapted and applied to the Poisson approximation by Chen in 1975[5]. Since then, Stein's method has stimulated an area of intensive research in combinatorics, probability and statistics.

There are at least three approaches to use Stein's method when the limit distribution is normal, i.e., namely a concentration inequality approach(see, for examples, Ho and Chen[9], and Chen and Shao[6]), an inductive approach (see, for example, Bolthausen[3]), and a coupling approach (see, for example, Stein[26]). In this work, we use the concentration inequality approach.

In this section we give basic results on the Stein's equation and its solution.

Let $Z$ be a standard normal distributed random variable and let $C_{bd}$ be the set of continuous and piecewise continuously differentiable finctions $g : \mathbb{R} \to \mathbb{R}$ with $E|g'(Z)| < \infty$.

For $g \in C_{bd}$ and any real valued function $I$ with $E|I(Z)| < \infty$, the equation

$$g'(w) - wg(w) = I(w) - EI(Z) \tag{2.1}$$

is called **Stein's equation**. The Stein's equation for normal distribution function is

$$g'(w) - wg(w) = \mathbb{I}(w \leq z) - \Phi(z), \tag{2.2}$$

for $z \in \mathbb{R}$, where $\mathbb{I}$ is an indicator function.
Hence

$$E(g'(W)) - EWg(W) = P(W \leq z) - \Phi(z) \tag{2.3}$$

for any random variable $W$ and the solution $g_z$ of (2.2) is given by

$$g_z(w) = \begin{cases} \sqrt{2\pi} e^{\frac{w^2}{2}} \Phi(w)[1 - \Phi(z)] & ,\text{if} \quad w \leq z, \\ \sqrt{2\pi} e^{\frac{w^2}{2}} \Phi(z)[1 - \Phi(w)] & ,\text{if} \quad w > z. \end{cases} \tag{2.4}$$

Thus, it suffices to find a bound for

$$E(g'_z(W)) - EWg_z(W)$$

instead of

$$P(W \leq z) - \Phi(z).$$

# CHAPTER III
# A UNIFORM BOUND FOR THE GENERALIZATION OF A COMBINATORIAL CENTRAL LIMIT THEOREM

Let $d$ be a positive integer and let:

1. $\pi_k,\ 1 \le k \le d$, be random permutations of $\{1, ..., n\}$ each uniformly distributed over all the $n!$ possible permutations;
2. $Y(i_1, ..., i_{d+1}),\ 1 \le i_1, ..., i_{d+1} \le n$, be random variables;
3. the $Y(i_1, ..., i_{d+1})$'s and $\pi_k$'s all be stochastically independent.

We define

$$W = \sum_{i=1}^{n} Y(i, \pi_1(i), ..., \pi_d(i)). \tag{3.1}$$

This chapter is concerned with the normal approximation to the distribution of $W$. In the special case of $d = 1$, this is called a combinatorial central limit theorem(see, for examples, Wald and Walfowitz[29], Hoeffding[10], Matoo[15], Hajek[7], Robinson[24], Kolchin and Chistyakov[11], Von Bahr[30], Ho and Chen[9], Bolthausen [3], Bolthausen and Gotze[4] and Neammanee and Suntornchost[17]). In this chapter, we establish a uniform bound for $W$ in the case of an arbitary $d$ by using technique from Ho and Chen[9] and Neammanee and Suntornchost[17]. This is our main result.

**Theorem 3.1.** *Suppose that*

$$EW = 0 \ \ \textit{and} \ \ VarW = 1, \tag{3.2}$$

*and*

$$\sum_{i_j=1}^{n} \tilde{\mu}(i_1, ..., i_{d+1}) = 0 \ \ \textit{for every} \ \ j \in \{1, ..., d+1\} \tag{3.3}$$

*where*

$$\tilde{\mu}(i_1, ..., i_{d+1}) = EY(i_1, ..., i_{d+1}).$$

*Then for* $n \geq 6^{(d+1)} + 3$,

$$\sup_{z\in\mathbb{R}} |P(W \leq z) - \Phi(z)| \leq \frac{11.946}{\sqrt{n}} + \frac{1.037\sqrt{d+1}\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + 8.314(d+1)^{\frac{1}{4}}\delta_4 + 11.765\delta_4 + 5.014(d+1)\frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}} + \frac{2\sqrt{2\pi}\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}$$

*where*

$$\delta_4 = \frac{1}{n^{d-\frac{1}{2}}} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} E|Y(i_1, ..., i_{d+1})|^4.$$

*Furthermore, if* $\delta_4 \sim n^{-1/2}$, *then*

$$\sup_{z\in\mathbb{R}} |P(W \leq z) - \Phi(z)| \leq \frac{28.725 + 1.037\sqrt{d+1} + 8.314(d+1)^{\frac{1}{4}} + 5.014(d+1)}{\sqrt{n}}. \quad (3.4)$$

Note that in the special case of $d = 1$, Neammanee and Suntornchost[17] yields the bound $\frac{216}{\sqrt{n}}$ while Theorem 3.1 gave a constant $C = 50.106$.

We remark that no generality is lost in this work by assuming conditions (3.2) and (3.3). In the case of $d = 1$, this is showed by Neammanee and Suntornchost[17] as follow.

Let $(X(i,j))$ be an $n \times n$ matrix of independent random variables and $\pi$ be a random permutation of $\{1, ..., n\}$, such that $\pi$ and $X(i,j)$'s are independent. Their work is concerned with the normal approximation to the distribution function of $W_n = \sum_{i=1}^{n} X(i, \pi(i))$.

For each $i, j \in \{1, 2, ..., n\}$, let $\mu_{ij}$ and $\sigma_{ij}^2$ be the mean and variance of $X_{ij}$, respectively and

$$\mu_{i.} = \frac{1}{n}\sum_j \mu_{ij}, \quad \mu_{.j} = \frac{1}{n}\sum_i \mu_{ij}, \quad \mu_{..} = \frac{1}{n^2}\sum_{i,j} \mu_{ij}$$

$$d^2 = \frac{1}{(n-1)}\sum_{i,j}(\mu_{ij} - \mu_{i.} - \mu_{j.} + \mu_{..})^2 \quad \text{and} \quad \sigma^2 = \frac{1}{n}\sum_{i,j}\sigma_{i,j}^2.$$

From Ho and Chen[9], $VarW_n = d^2 + \sigma^2$. Then

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

satisfies the conditions (3.2) and (3.3) where

$$Y(i,j) = \frac{1}{\sqrt{d^2 + \sigma^2}}(X(i,j) - \mu_{i.} - \mu_{.j} + \mu_{..})$$

(see[17],pp.560).

Next, we shall generalize this argument to the case of an arbitary $d$. Let $X(i_1,...,i_{d+1})$, $1 \le i_1,...,i_{d+1} \le n$, be random variables and $\pi_k,\ 1 \le k \le d$, be random permutations of $\{1,...,n\}$ each uniformly distributed over all the $n!$ possible permutations such that the $X(i_1,...,i_{d+1})$'s and $\pi_k$'s all are stochastically independent.

Assume that $\sigma^2 := Var(\sum_{i=1}^{n} X(i,\pi_1(i),...,\pi_d(i))) > 0$. We define

$$W = \sum_{i=1}^{n} Y(i,\pi_1(i),...,\pi_d(i)) \tag{3.5}$$

where

$$\begin{aligned}
Y(i_1,...,i_{d+1}) &= \frac{1}{\sigma}[X(i_1,...,i_{d+1}) + \sum_{k=1}^{d} \mu_k(i_1,...,i_{d+1}) + (-1)^{(d+1)}\frac{\mu}{n}],\\
\mu_k(i_1,...,i_{d+1}) &= \frac{(-1)^k}{n^k} \sum_{1\le j_1<j_2<...<j_k\le d+1} \sum_{q_{j_1}} ... \sum_{q_{j_k}} \tilde{\mu}(l_1,...,l_{d+1}),\\
l_p &= \begin{cases} q_p & \text{if } \ p = j_1,...,j_k,\\ i_p & \text{otherwise,} \end{cases} \qquad (3.6)\\
\tilde{\mu}(i_1,...,i_{d+1}) &= EX(i_1,...,i_{d+1})
\end{aligned}$$

and
$$\mu = \sum_{i=1}^{n} EX(i,\pi_1(i),...,\pi_d(i)).$$

Then $W$ satisfies conditions (3.2) and (3.3). The proofs are shown in the following proposition.

**Proposition 3.2.**

1.$EW = 0$ *and* $VarW = 1$

2. $\sum_{i_j=1}^{n} EY(i_1,...,i_{d+1}) = 0$ *for every* $j = 1,...,d+1$.

*Proof.* 1. Note that

$$\mu = \frac{1}{n^d}\sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} \tilde{\mu}(i_1,...,i_{d+1}).$$

By the fact that, for fixed $i$,

$$\sum_{i_{j_1}=1}^{n} ... \sum_{i_{j_k}=1}^{n} \tilde{\mu}(l_1,...,l_{d+1})$$

where $1 = j_1 < j_2 < ... < j_k \leq d+1$ and

$$l_p = \begin{cases} q_p & \text{if } \ p = 1, j_2, ..., j_k, \\ \pi_p(i) & \text{otherwise} \end{cases}$$

is equal to

$$\frac{1}{n^{d+1-k}} \sum_{k_1=1}^{n} ... \sum_{k_{d+1}=1}^{n} \tilde{\mu}(k_1, ..., k_{d+1})$$

and the fact that, for fixed $i$,

$$\sum_{q_{j_1}=1}^{n} ... \sum_{q_{j_k}=1}^{n} \tilde{\mu}(l_1, ..., l_{d+1})$$

where $1 < j_1 < j_2 < ... < j_k \leq d+1$ and

$$l_p = \begin{cases} i & \text{if } \ p = 1, \\ q_p & \text{if } \ p = j_1, j_2, ..., j_k, \\ \pi_p(i) & \text{otherwise} \end{cases}$$

is equal to

$$\frac{1}{n^{d-k}} \sum_{k_2=1}^{n} ... \sum_{k_{d+1}=1}^{n} \tilde{\mu}(i, k_2, ..., k_{d+1}),$$

we have

$$\begin{aligned} \mu_k(i, \pi_1(i), ..., \pi_d(i)) = \frac{(-1)^k}{n^{d+1}} & \sum_{1=j_1<j_2<...<j_k\leq d+1} \sum_{k_1=1}^{n} ... \sum_{k_{d+1}=1}^{n} \tilde{\mu}(k_1, ..., k_{d+1}) \\ & + \frac{(-1)^k}{n^d} \sum_{1<j_1<j_2<...<j_k\leq d+1} \sum_{k_2=1}^{n} ... \sum_{k_{d+1}=1}^{n} \tilde{\mu}(i, k_2, ..., k_{d+1}). \end{aligned}$$

Notice that

$$\begin{aligned} & \sum_{i=1}^{n} \sum_{k=1}^{d} \mu_k(i, \pi_1(i), ..., \pi_d(i)) + (-1)^{(d+1)}\mu \\ & = \sum_{k=1}^{d} \frac{(-1)^k}{n^d} \sum_{1=j_1<j_2<...<j_k\leq d+1} \sum_{k_1=1}^{n} ... \sum_{k_{d+1}=1}^{n} \tilde{\mu}(k_1, ..., k_{d+1}) \\ & + \sum_{k=1}^{d} \frac{(-1)^k}{n^d} \sum_{1<j_1<j_2<...<j_k\leq d+1} \sum_{i=1}^{n} \sum_{k_2=1}^{n} ... \sum_{k_{d+1}=1}^{n} \tilde{\mu}(i, k_2, ..., k_{d+1}) + (-1)^{(d+1)}\mu \\ & = \sum_{k=1}^{d} (-1)^k \sum_{1=j_1<j_2<...<j_k\leq d+1} \mu + \sum_{k=1}^{d} (-1)^k \sum_{1<j_1<j_2<...<j_k\leq d+1} \mu + (-1)^{(d+1)}\mu \end{aligned}$$

$$= \sum_{k=1}^{d}(-1)^k \sum_{1\le j_1<j_2<...<j_k\le d+1} \mu + (-1)^{(d+1)}\mu$$

$$= \sum_{k=1}^{d}(-1)^k \binom{d+1}{k}\mu + (-1)^{(d+1)}\mu$$

$$= \sum_{k=1}^{d+1}(-1)^k \binom{d+1}{k}\mu$$

$$= [(1-1)^{(d+1)} - 1]\mu$$

$$= -\mu.$$

Thus

$$W = \frac{1}{\sigma}[\sum_{i=1}^{n} X(i,\pi_1(i),...,\pi_d(i)) + \sum_{i=1}^{n}\sum_{k=1}^{d}\mu_k(i,\pi_1(i),...,\pi_d(i)) + (-1)^{(d+1)}\mu]$$

$$= \frac{1}{\sigma}[\sum_{i=1}^{n} X(i,\pi_1(i),...,\pi_d(i)) - \mu].$$

Hence $EW = 0$ and $VarW = 1$.

2. We shall first proof in the case of $j = 1$. Let $\mu_k(i_1,...,i_{d+1})$ and $l_1,...,l_{d+1}$ be defined as in (3.6) and $i_2,...,i_{d+1} \in \{1,...,n\}$. We observe that

$$\sum_{i_1=1}^{n}\sum_{k=1}^{d}\mu_k(i_1,...,i_{d+1})$$

$$= \sum_{i_1=1}^{n}\sum_{k=1}^{d}\frac{(-1)^k}{n^k}\sum_{1\le j_1<j_2<...<j_k\le d+1}\sum_{q_{j_1}}...\sum_{q_{j_k}}\tilde{\mu}(l_1,...,l_{d+1})$$

$$= \sum_{k=1}^{d}\frac{(-1)^k}{n^{k-1}}\sum_{1=j_1<j_2<...<j_k\le d+1}\sum_{q_1}\sum_{q_{j_2}}...\sum_{q_{j_k}}\tilde{\mu}(q_1,l_2,...,l_{d+1})$$

$$+ \sum_{k=1}^{d}\frac{(-1)^k}{n^k}\sum_{i_1=1}^{n}\sum_{1<j_1<j_2<...<j_k\le d+1}\sum_{q_{j_1}}...\sum_{q_{j_k}}\tilde{\mu}(i_1,l_2,...,l_{d+1})$$

$$= -\sum_{q_1=1}^{n}\tilde{\mu}(q_1,i_2,...,i_{d+1}) + \sum_{k=2}^{d}\frac{(-1)^k}{n^{k-1}}\sum_{1=j_1<j_2<...<j_k\le d+1}\sum_{q_1}\sum_{q_{j_2}}...\sum_{q_{j_k}}\tilde{\mu}(q_1,l_2,...,l_{d+1})$$

$$+ \sum_{k=1}^{d-1}\frac{(-1)^k}{n^k}\sum_{i_1=1}^{n}\sum_{1<j_1<j_2<...<j_k\le d+1}\sum_{q_{j_1}}...\sum_{q_{j_k}}\tilde{\mu}(i_1,l_2,...,l_{d+1})$$

$$+\frac{(-1)^d}{n^d}\sum_{i_1=1}^{n}\sum_{q_2}\cdots\sum_{q_{d+1}}\tilde{\mu}(i_1,q_2,...,q_{d+1})$$

$$=-\sum_{q_1=1}^{n}\tilde{\mu}(q_1,i_2,...,i_{d+1})+\sum_{k=2}^{d}\frac{(-1)^k}{n^{k-1}}\sum_{1=j_1<j_2<...<j_k\le d+1}\sum_{q_1}\sum_{q_{j_2}}\cdots\sum_{q_{j_k}}\tilde{\mu}(q_1,l_2,...,l_{d+1})$$

$$+\sum_{k=2}^{d}\frac{(-1)^{k-1}}{n^{k-1}}\sum_{1<j_1<j_2<...<j_{k-1}\le d+1}\sum_{i_1=1}^{n}\sum_{q_{j_1}}\cdots\sum_{q_{j_{k-1}}}\tilde{\mu}(i_1,l_2,...,l_{d+1})+(-1)^d\mu$$

$$=(-1)^d\mu-\sum_{q_1=1}^{n}\tilde{\mu}(q_1,i_2,...,i_{d+1}).$$

Thus

$$\sum_{i_1=1}^{n}EY(i_1,...,i_{d+1})$$

$$=\frac{1}{\sigma}\sum_{i_1=1}^{n}E[X(i_1,...,i_{d+1})+\sum_{k=1}^{d}\mu_k(i_1,...,i_{d+1})+(-1)^{(d+1)}\frac{\mu}{n}]$$

$$=\frac{1}{\sigma}[\sum_{i_1=1}^{n}\tilde{\mu}(i_1,...,i_{d+1})+\sum_{i_1=1}^{n}\sum_{k=1}^{d}\mu_k(i_1,...,i_{d+1})+(-1)^{(d+1)}\mu]$$

$$=0.$$

For $j=2,...,d+1$, we can prove this by using the same argument. □

## 3.1 Auxiliary results

In this section, we give auxiliary results for proving our main theorem(Theorem 3.1).

We shall first construct the following system which generalizes the result from Ho and Chen[9] naturally by extending to $d$-dimensions. Let $I,K,L_1,...,L_d,M_1,...,M_d$ be uniformly distributed random variables on $\{1,2,...,n\}$ and $\rho_1,...,\rho_d,\tau_1,...,\tau_d$ are random permutations of $\{1,2,...,n\}$. Assume that

$$\{I,K,L_1,...,L_d,M_1,...,M_d,\rho_1,...,\rho_d,\tau_1,...,\tau_d\}\ \text{ is independent of } Y(i_1,...,i_{d+1})\text{'s,} \tag{3.7}$$

$$(I,K),(L_1,M_1),...,(L_d,M_d)\ \text{ are uniformly distributed on } \{(i,k)|i,k=1,2,...,n\ \text{ and }\ i\neq k\}, \tag{3.8}$$

$$(I,K),(L_1,M_1),...,(L_d,M_d)\ \text{ and }\ \tau_1,...,\tau_d\ \text{ are mutually independent,} \tag{3.9}$$

$$(I, K) \text{ and } \rho_1, ..., \rho_d \text{ are mutually independent, and} \tag{3.10}$$

$$\rho_i(\alpha) = \begin{cases} \tau_i(\alpha) & \text{if} \quad \alpha \neq I, K, \tau_i^{-1}(L_i), \tau_i^{-1}(M_i) \\ L_i & \text{if} \quad \alpha = I, \\ M_i & \text{if} \quad \alpha = K, \\ \tau_i(I) & \text{if} \quad \alpha = \tau_i^{-1}(L_i) \\ \tau_i(K) & \text{if} \quad \alpha = \tau_i^{-1}(M_i), \end{cases} \tag{3.11}$$

where $\rho_i(\rho_i^{-1}(\alpha)) = \rho_i^{-1}(\rho_i(\alpha)) = \alpha,$ for $i = 1, ..., d$. In the case of $d = 1$, Ho[8] gave an example of random vectors $I, K, L, M, \pi, \rho, \tau$ and $Y(i, j), i, j = 1, 2, ..., n$, which satisfy the conditions (3.7)-(3.11). It easy to generalize his example naturally by extending to $d$-dimensions. Let

$$S(\rho) = \sum_{i=1}^{n} Y(i, \rho_1(i), ..., \rho_d(i)) \tag{3.12}$$

and

$$\widetilde{S}(\rho) = S(\rho) - S_1 - S_2 + S_3 + S_4$$

where

$$S_1 = Y(I, \rho_1(I), ..., \rho_d(I)), \quad S_2 = Y(K, \rho_1(K), ..., \rho_d(K))$$
$$S_3 = Y(I, \rho_1(K), .., \rho_d(K)), \quad S_4 = Y(K, \rho_1(I), ..., \rho_d(I)).$$

Clearly $S_1$ and $S_2$ have the same distribution and so do $S_3$ and $S_4$. We observe that

$$P(S_1 \leq a) = \frac{1}{n^{d+1}} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} P(Y(i_1, ..., i_{d+1}) \leq a) = P(S_3 \leq a) \quad \text{for all} \quad a \in \mathbb{R}.$$

Thus $S_1, S_2, S_3, S_4$ are identically distributed. (3.13)

**Lemma 3.3.** *Let $g : \mathbb{R} \to \mathbb{R}$ be a continuous and piecewise continuously differentiable function. Then*

$$ES(\rho)g(S(\rho)) = E\int_{-\infty}^{\infty} g'(S(\rho) + t)M(t)dt - \Delta g(S(\rho)) \tag{3.14}$$

*and*

$$|\Delta g(S(\rho))| \leq \frac{1}{n-1}[Eg^2(S(\rho))]^{1/2}, \tag{3.15}$$

*where*

$$M(t) = \frac{n}{4}(\widetilde{S}(\rho) - S(\rho))\{\mathbb{I}(0 \le t \le \widetilde{S}(\rho) - S(\rho)) - \mathbb{I}(\widetilde{S}(\rho) - S(\rho) \le t \le 0)\}$$

*and* $\mathbb{I}$ *is an indicator function.*

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

$$\{Y(i, \rho_1(i), ..., \rho_d(i)) : 1 \le i \le n\}.$$

By the same argument as in Neammanee and Suntornchost([17],pp.565), we have

$$2E\{g(S(\rho))E^{\mathcal{A}}(\widetilde{S}(\rho) - S(\rho))\} + E(\widetilde{S}(\rho) - S(\rho))[g(\widetilde{S}(\rho)) - g(S(\rho))] = 0. \tag{3.16}$$

We observe that for $i \neq k$ and $C \in \mathcal{A}$,

$$\begin{aligned}
\int_C Y(i, \rho_1(k),..., \rho_d(k))dP
&= \sum_{j_1=1}^{n} ... \sum_{j_d=1}^{n} \int_{\{\rho_1(k)=j_1,...,\rho_d(k)=j_d\}\cap C} Y(i, j_1, ..., j_d)dP \\
&= \sum_{j_1=1}^{n} ... \sum_{j_d=1}^{n} \int_{\Omega} Y(i, j_1, ..., j_d)\mathbb{I}(\{\rho_1(k) = j_1, ...\rho_d(k) = j_d\} \cap C)dP \\
&= \sum_{j_1=1}^{n} ... \sum_{j_d=1}^{n} EY(i, j_1, ..., j_d)\mathbb{I}(\{\rho_1(k) = j_1, ...\rho_d(k) = j_d\} \cap C) \\
&= \sum_{j_1=1}^{n} ... \sum_{j_d=1}^{n} \tilde{\mu}(i, j_1, ..., j_d)P(\{\rho_1(k) = j_1, ...\rho_d(k) = j_d\} \cap C) \\
&= \sum_{j_1=1}^{n} ... \sum_{j_d=1}^{n} \int_{\Omega} \tilde{\mu}(i, j_1, ..., j_d)\mathbb{I}(\{\rho_1(k) = j_1, ..., \rho_d(k) = j_d\} \cap C)dP \\
&= \sum_{j_1=1}^{n} ... \sum_{j_d=1}^{n} \int_{\{\rho_1(k)=j_1,...,\rho_d(k)=j_k\}\cap C} \tilde{\mu}(i, j_1, ..., j_d)dP \\
&= \int_C \tilde{\mu}(i, \rho_1(k), ..., \rho_d(k))dP.
\end{aligned}$$

Thus

$$E^{\mathcal{A}}Y(i, \rho_1(k), ..., \rho_d(k)) = E^{\mathcal{A}}\tilde{\mu}(i, \rho_1(k), ..., \rho_d(k)) \text{ for any } i \neq k. \tag{3.17}$$

By this fact, we have

$$\begin{aligned}
E^{\mathcal{A}}(\widetilde{S}(\rho) - S(\rho)) &= E^{\mathcal{A}}[-S_1 - S_2 + S_3 + S_4] \\
&= -E^{\mathcal{A}}Y(I, \rho_1(I), ..., \rho_d(I)) - E^{\mathcal{A}}Y(K, \rho_1(K), ..., \rho_d(K)) \\
&\quad + E^{\mathcal{A}}Y(I, \rho_1(K), ..., \rho_d(K)) + E^{\mathcal{A}}Y(K, \rho_1(I), ..., \rho_d(I)) \\
&= -\frac{2}{n}\sum_{i=1}^{n} Y(i, \rho_1(i), ..., \rho_d(i)) \\
&\quad + \frac{2}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{k \\ k\neq i}} E^{\mathcal{A}}Y(i, \rho_1(k), ..., \rho_d(k)) \\
&= -\frac{2}{n}S(\rho) + \frac{2}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{k \\ k\neq i}} E^{\mathcal{A}}\tilde{\mu}(i, \rho_1(k), ..., \rho_d(k)) \\
&= -\frac{2}{n}S(\rho) + \frac{2}{n(n-1)}E^{\mathcal{A}}\sum_{i=1}^{n}\{\sum_{k=1}^{n}\tilde{\mu}(i, \rho_1(k), ..., \rho_d(k)) \\
&\quad - \tilde{\mu}(i, \rho_1(i), ..., \rho_d(i))\} \\
&= -\frac{2}{n}S(\rho) + \frac{2}{n(n-1)}E^{\mathcal{A}}\sum_{i=1}^{n}\sum_{k=1}^{n}\tilde{\mu}(i, \rho_1(k), ..., \rho_d(k)) \\
&\quad - \frac{2}{n(n-1)}E^{\mathcal{A}}\sum_{i=1}^{n}\tilde{\mu}(i, \rho_1(i), ..., \rho_d(i)). \qquad (3.18)
\end{aligned}$$

We note from (3.3) that

$$\sum_{i=1}^{n}\sum_{k=1}^{n}\tilde{\mu}(i, \rho_1(k), ..., \rho_d(k)) = 0.$$

This implies the second term on the right-hand side of (3.18) is zero.
Hence

$$E^{\mathcal{A}}(\widetilde{S}(\rho) - S(\rho)) = -\frac{2}{n}S(\rho) - \frac{2}{n(n-1)}E^{\mathcal{A}}\sum_{i=1}^{n}\tilde{\mu}(i, \rho_1(i), ..., \rho_d(i)). \qquad (3.19)$$

We conclude from (3.16) and (3.19) that

$$\begin{aligned}
0 &= 2E\{g(S(\rho))[-\frac{2}{n}S(\rho) - \frac{2}{n(n-1)}E^{\mathcal{A}}\sum_{i=1}^{n}\tilde{\mu}(i, \rho_1(i), ..., \rho_d(i))]\} \\
&\quad + E(\widetilde{S}(\rho) - S(\rho))[g(\widetilde{S}(\rho)) - g(S(\rho))].
\end{aligned}$$

This implies

$$\begin{aligned}
ES(\rho)g(S(\rho)) &= \frac{n}{4}E(\widetilde{S}(\rho)-S(\rho))[g(\widetilde{S}(\rho))-g(S(\rho))] \\
&\quad -\frac{1}{n-1}E[g(S(\rho))E^{\mathcal{A}}\sum_{i=1}^{n}\tilde{\mu}(i,\rho_1(i),...,\rho_d(i))] \\
&= \frac{n}{4}E(\widetilde{S}(\rho)-S(\rho))[g(\widetilde{S}(\rho))-g(S(\rho))]-\Delta g(S(\rho)) \\
&= \frac{n}{4}E(\widetilde{S}(\rho)-S(\rho))\int_0^{\widetilde{S}(\rho)-S(\rho)} g'(S(\rho)+t)dt-\Delta g(S(\rho)) \\
&= E\int_{-\infty}^{\infty} g'(S(\rho)+t)M(t)dt-\Delta g(S(\rho))
\end{aligned}$$

where

$$\Delta g(S(\rho)) = \frac{1}{n-1}E[g(S(\rho))E^{\mathcal{A}}\sum_{i=1}^{n}\tilde{\mu}(i,\rho_1(i),...,\rho_d(i))].$$

To finish the proof, we need to bound $|\Delta g(S(\rho))|$. We observe from (3.2) and (3.3) that

$$\begin{aligned}
&E[\sum_{i=1}^{n}\tilde{\mu}(i,\rho_1(i),...,\rho_d(i))]^2 \\
&= \sum_{i=1}^{n}E\tilde{\mu}^2(i,\rho_1(i),...,\rho_d(i))+\sum_{i=1}^{n}\sum_{\substack{j\\ j\neq i}}E\tilde{\mu}(i,\rho_1(i),...,\rho_d(i))\tilde{\mu}(j,\rho_1(j),...,\rho_d(j)) \\
&= \frac{1}{n^d}\sum_{i_1=1}^{n}...\sum_{i_{d+1}=1}^{n}\tilde{\mu}^2(i_1,...,i_{d+1}) \\
&\quad +\frac{1}{(n(n-1))^d}\sum_{i_1=1}^{n}...\sum_{i_{d+1}=1}^{n}\tilde{\mu}(i_1,...,i_{d+1})\sum_{\substack{j_1\\ j_1\neq i_1}}...\sum_{\substack{j_{d+1}\\ j_{d+1}\neq i_{d+1}}}\tilde{\mu}(j_1,...,j_{d+1}) \\
&= \frac{1}{n^d}\sum_{i_1=1}^{n}...\sum_{i_{d+1}=1}^{n}\tilde{\mu}^2(i_1,...,i_{d+1})+\frac{(-1)^{d+1}}{(n(n-1))^d}\sum_{i_1=1}^{n}...\sum_{i_{d+1}=1}^{n}\tilde{\mu}^2(i_1,...,i_{d+1}) \\
&\le \frac{1}{n^d}\sum_{i_1=1}^{n}...\sum_{i_{d+1}=1}^{n}EY^2(i_1,...,i_{d+1})+\frac{(-1)^{(d+1)}}{(n(n-1))^d}\sum_{i_1=1}^{n}...\sum_{i_{d+1}=1}^{n}\tilde{\mu}^2(i_1,...,i_{d+1}) \\
&= \frac{1}{n^d}\sum_{i_1=1}^{n}...\sum_{i_{d+1}=1}^{n}EY^2(i_1,...,i_{d+1}) \\
&\quad +\frac{1}{(n(n-1))^d}\sum_{i_1=1}^{n}...\sum_{i_{d+1}=1}^{n}EY(i_1,...,i_{d+1})\sum_{\substack{j_1\\ j_1\neq i_1}}...\sum_{\substack{j_{d+1}\\ j_{d+1}\neq i_{d+1}}}EY(j_1,...,j_{d+1})
\end{aligned}$$

$$= \sum_{i=1}^{n} EY^2(i, \rho_1(i), ..., \rho_d(i)) + \sum_{i=1}^{n} \sum_{\substack{j \\ j \neq i}} EY(i, \rho_1(i), ..., \rho_d(i)) Y(j, \rho_1(j), ..., \rho_d(j))$$

$$= E[\sum_{i=1}^{n} Y(i, \rho_1(i), ..., \rho_d(i))]^2$$

$$= E(S(\rho))^2$$

$$= 1. \tag{3.20}$$

Hence, by Hölder's inequality and (3.20),

$$|\Delta g(S(\rho))| \leq \frac{1}{n-1} E[g^2(S(\rho))]^{\frac{1}{2}} \{E[\sum_{i=1}^{n} \tilde{\mu}(i, \rho_1(i), ..., \rho_d(i))]^2\}^{\frac{1}{2}}$$
$$\leq \frac{1}{n-1} E[g^2(S(\rho))]^{\frac{1}{2}}.$$

□

**Lemma 3.4.** *Let $S(\rho)$ be defined as in (3.12). Then in the special case of $d = 1$ and $n \geq 39$, we have*

$$ES^4(\rho) \leq 4.678\sqrt{n}\delta_4.$$

*Proof.* Note that

$$ES^4(\rho) = A_1 + A_2 + A_3 + A_4 + A_5 \tag{3.21}$$

where

$$A_1 = \sum_{i=1}^{n} EY^4(i, \rho_1(i))$$

$$A_2 = \sum_{i_1=1}^{n} \sum_{\substack{i_2 \\ i_2 \neq i_1}} EY^3(i_1, \rho_1(i_1)) Y(i_2, \rho_1(i_2))$$

$$A_3 = \sum_{i_1=1}^{n} \sum_{\substack{i_2 \\ i_2 \neq i_1}} EY^2(i_1, \rho_1(i_1)) Y^2(i_2, \rho_1(i_2))$$

$$A_4 = \sum_{i_1=1}^{n} \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{i_3 \\ i_3 \neq i_1, i_2}} EY^2(i_1, \rho_1(i_1)) Y(i_2, \rho_1(i_2)) Y(i_3, \rho_1(i_3))$$

$$A_5 = \sum_{i_1=1}^{n} \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{i_3 \\ i_3 \neq i_1, i_2}} \sum_{\substack{i_4 \\ i_4 \neq i_1, i_2, i_3}} EY(i_1, \rho_1(i_1)) Y(i_2, \rho_1(i_2)) Y(i_3, \rho_1(i_3)) Y(i_4, \rho_1(i_4)).$$

By the fact that $n \geq 39$ and (3.3), we have

$$A_1 = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} EY^4(i,j) = \frac{\delta_4}{\sqrt{n}} \leq 0.025\sqrt{n}\delta_4 \tag{3.22}$$

and

$$\begin{aligned}
A_2 &= \frac{1}{n(n-1)}\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY^3(i_1,j_1)\sum_{\substack{i_2\\ i_2\neq i_1}}\sum_{\substack{j_2\\ j_2\neq j_1}} EY(i_2,j_2)\\
&= \frac{1}{n(n-1)}\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY^3(i_1,j_1)EY(i_1,j_1)\\
&\leq \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n} EY^4(i,j)\\
&= \frac{\delta_4}{\sqrt{n}(n-1)}\\
&\leq 0.0006\sqrt{n}\delta_4. \qquad (3.23)
\end{aligned}$$

We observe that

$$\Big[\sum_{i=1}^{n}\sum_{j=1}^{n} EY^2(i,j)\Big]^2 \leq n^2\sum_{i=1}^{n}\sum_{j=1}^{n}[EY^2(i,j)]^2 \leq n^2\sum_{i=1}^{n}\sum_{j=1}^{n} EY^4(i,j) = n^2\sqrt{n}\delta_4. \tag{3.24}$$

By this fact,

$$\begin{aligned}
A_3 &= \frac{1}{n(n-1)}\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY^2(i_1,j_1)\sum_{\substack{i_2\\ i_2\neq i_1}}\sum_{\substack{j_2\\ j_2\neq j_1}} EY^2(i_2,j_2)\\
&\leq \frac{1}{n(n-1)}\Big[\sum_{i=1}^{n}\sum_{j=1}^{n} EY^2(i,j)\Big]^2\\
&\leq \frac{n\sqrt{n}}{n-1}\delta_4\\
&\leq 1.026\sqrt{n}\delta_4. \qquad (3.25)
\end{aligned}$$

Next, we will bound $A_4$. Note that

$$\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY^2(i_1,j_1)\sum_{\substack{i_2\\ i_2\neq i_1}}\sum_{\substack{j_2\\ j_2\neq j_1}} EY(i_2,j_2)\sum_{\substack{i_3\\ i_3\neq i_1,i_2}}\sum_{\substack{j_3\\ j_3\neq j_1,j_2}} EY(i_3,j_3) \tag{3.26}$$

consists of a sum of 4 terms each of the form

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

where $k = i_1$ or $i_2$ and $l = j_1$ or $j_2$. By (3.3), for fixed $i_1, j_1$,

$$\sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY(i_2, j_2)EY(k, l) \leq \sum_{i=1}^{n}\sum_{j=1}^{n}(EY(i,j))^2 \leq \sum_{i=1}^{n}\sum_{j=1}^{n} EY^2(i,j).$$

This implies

$$\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY^2(i_1, j_1) \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY(i_2, j_2) \sum_{\substack{i_3 \\ i_3 \neq i_1, i_2}} \sum_{\substack{j_3 \\ j_3 \neq j_1, j_2}} EY(i_3, j_3) \leq 4[\sum_{i=1}^{n}\sum_{j=1}^{n} EY^2(i,j)]^2.$$

From this fact, (3.24) and $n \geq 39$, we have

$$\begin{aligned}
A_4 &= \frac{1}{n(n-1)(n-2)} \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY^2(i_1, j_1) \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY(i_2, j_2) \sum_{\substack{i_3 \\ i_3 \neq i_1, i_2}} \sum_{\substack{j_3 \\ j_3 \neq j_1, j_2}} EY(i_3, j_3) \\
&\leq \frac{4}{n(n-1)(n-2)} [\sum_{i=1}^{n}\sum_{j=1}^{n} EY^2(i,j)]^2 \\
&\leq \frac{4n\sqrt{n}}{(n-1)(n-2)} \delta_4 \\
&= 4(\frac{n}{n-1})(\frac{1}{n-2})\sqrt{n}\delta_4 \\
&\leq 0.11\sqrt{n}\delta_4. \qquad (3.27)
\end{aligned}$$

It remains to bound $A_5$. Similarly to (3.26),

$$\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1, j_1) \sum_{\substack{i_2 \\ i_2 \neq i_1}} \sum_{\substack{j_2 \\ j_2 \neq j_1}} EY(i_2, j_2) \sum_{\substack{i_3 \\ i_3 \neq i_1, i_2}} \sum_{\substack{j_3 \\ j_3 \neq j_1, j_2}} EY(i_3, j_3) \sum_{\substack{i_4 \\ i_4 \neq i_1, i_2, i_3}} \sum_{\substack{j_4 \\ j_4 \neq j_1, j_2, j_3}} EY(i_4, j_4)$$

consists of a sum of 9 terms each of the form

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

where $k \in \{i_1, i_2, i_3\}$ and $l \in \{j_1, j_2, j_3\}$.

**case 1.** $(k, l) = (i_1, j_1)$

Note that

$$B(i_1, j_1) \leq A_{5,1} + A_{5,2} + A_{5,3} + A_{5,4}$$

where

$$A_{5,1} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^3 \sum_{\substack{i_2\\ i_2\neq i_1}}\sum_{\substack{j_2\\ j_2\neq j_1}} EY(i_2,j_2),$$

$$A_{5,2} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^2 \sum_{\substack{i_2\\ i_2\neq i_1}}\sum_{\substack{j_2\\ j_2\neq j_1}} EY(i_2,j_2)EY(i_1,j_2),$$

$$A_{5,3} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^2 \sum_{\substack{i_2\\ i_2\neq i_1}}\sum_{\substack{j_2\\ j_2\neq j_1}} EY(i_2,j_2)EY(i_2,j_1),$$

and
$$A_{5,4} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^2 \sum_{\substack{i_2\\ i_2\neq i_1}}\sum_{\substack{j_2\\ j_2\neq j_1}} (EY(i_2,j_2))^2.$$

By (3.3),

$$A_{5,1} = \sum_{i=1}^{n}\sum_{j=1}^{n}(EY(i,j))^4 \leq \sqrt{n}\delta_4,$$

$$A_{5,2} = -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^2 \sum_{\substack{j_2\\ j_2\neq j_1}} (EY(i_1,j_2))^2 < 0,$$

$$A_{5,3} = -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^2 \sum_{\substack{i_2\\ i_2\neq i_1}} (EY(i_2,j_1))^2 < 0.$$

and by (3.24),

$$A_{5,4} \leq [\sum_{i=1}^{n}\sum_{j=1}^{n} EY^2(i,j)]^2 \leq n^2\sqrt{n}\delta_4.$$

Thus

$$B(i_1,j_1) \leq \sqrt{n}\delta_4 + n^2\sqrt{n}\delta_4.$$

**case 2.** $(k,l) = (i_2,j_2)$

We observe that

$$B(i_2,j_2) = A_{5,4} + A_{5,5} + A_{5,6} + A_{5,7}$$

where

$$A_{5,5} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2\\ i_2\neq i_1}}\sum_{\substack{j_2\\ j_2\neq j_1}} (EY(i_2,j_2))^2 EY(i_1,j_2),$$

$$A_{5,6} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2\\ i_2\neq i_1}} \sum_{\substack{j_2\\ j_2\neq j_1}} (EY(i_2,j_2))^2 EY(i_2,j_1),$$

$$\text{and} \quad A_{5,7} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2\\ i_2\neq i_1}} \sum_{\substack{j_2\\ j_2\neq j_1}} (EY(i_2,j_2))^3.$$

By (3.3),

$$\begin{aligned}
A_{5,5} &= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{j_2\\ j_2\neq j_1}} EY(i_1,j_2) \sum_{i_2=1}^{n} (EY(i_2,j_2))^2 \\
&\quad - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{j_2\\ j_2\neq j_1}} (EY(i_1,j_2))^3 \\
&= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{j_2=1}^{n} EY(i_1,j_2) \sum_{i_2=1}^{n} (EY(i_2,j_2))^2 \\
&\quad - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} (EY(i_1,j_1))^2 \sum_{i_2=1}^{n} (EY(i_2,j_1))^2 - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{j_2=1}^{n} (EY(i_1,j_2))^3 \\
&\quad + \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} (EY(i_1,j_1))^4 \\
&\leq \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY^4(i_1,j_1) \\
&= \sqrt{n}\delta_4, \\
A_{5,6} &= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2\\ i_2\neq i_1}} EY(i_2,j_1) \sum_{j_2=1}^{n} (EY(i_2,j_2))^2 \\
&\quad - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2\\ i_2\neq i_1}} (EY(i_2,j_1))^3 \\
&= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{i_2=1}^{n} EY(i_2,j_1) \sum_{j_2=1}^{n} (EY(i_2,j_2))^2 \\
&\quad - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} (EY(i_1,j_1))^2 \sum_{j_2=1}^{n} (EY(i_1,j_2))^2 \\
&\quad - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{i_2=1}^{n} (EY(i_2,j_1))^3 + \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} (EY(i_1,j_1))^4
\end{aligned}$$

$$\leq \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY^4(i_1,j_1)$$
$$= \sqrt{n}\delta_4$$

and

$$A_{5,7} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{j_2=1}^{n} (EY(i_2,j_2))^3 - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} (EY(i_2,j_1))^3$$
$$= -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{i_2=1}^{n} (EY(i_2,j_1))^3 + \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} (EY(i_1,j_1))^4$$
$$\leq \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY^4(i_1,j_1)$$
$$= \sqrt{n}\delta_4.$$

Thus

$$B(i_2,j_2) \leq 3\sqrt{n}\delta_4 + n^2\sqrt{n}\delta_4.$$

**case 3.** $(k,l) = (i_3,j_3)$

We note that

$$B(i_3,j_3) = A_{5,1} + A_{5,7} + A_{5,8} + A_{5,9} + A_{5,10} + A_{5,11} + A_{5,12} + A_{5,13} + A_{5,14}$$

where

$$A_{5,8} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2) \sum_{i_3=1}^{n}\sum_{j_3=1}^{n} (EY(i_3,j_3))^2,$$

$$A_{5,9} = -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2) \sum_{j_3=1}^{n} (EY(i_1,j_3))^2,$$

$$A_{5,10} = -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2) \sum_{j_3=1}^{n} (EY(i_2,j_3))^2,$$

$$A_{5,11} = -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2) \sum_{i_3=1}^{n} (EY(i_3,j_1))^2,$$

$$A_{5,12} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2)(EY(i_2,j_1))^2,$$

$$A_{5,13} = -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2) \sum_{i_3=1}^{n}(EY(i_3,j_2))^2, \quad \text{and}$$

$$A_{5,14} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2)(EY(i_1,j_2))^2.$$

By (3.3) and (3.24),

$$\begin{aligned}
A_{5,8} &= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^2 \sum_{i_3=1}^{n}\sum_{j_3=1}^{n}(EY(i_3,j_3))^2 \\
&\leq [\sum_{i=1}^{n}\sum_{j=1}^{n} EY^2(i,j)]^2 \\
&\leq n^2\sqrt{n}\delta_4, \\
A_{5,9} &= -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^2 \sum_{j_3=1}^{n}(EY(i_1,j_3))^2 \\
&< 0, \\
A_{5,10} &= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{j_3=1}^{n}(EY(i_2,j_3))^2 EY(i_2,j_1) \\
&= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{i_2=1}^{n}\sum_{j_3=1}^{n}(EY(i_2,j_3))^2 EY(i_2,j_1) \\
&\quad - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^2 \sum_{j_3=1}^{n}(EY(i_1,j_3))^2 \\
&< 0, \\
A_{5,11} &= -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{j_2=1}^{n} EY(i_2,j_2) \sum_{i_3=1}^{n}(EY(i_3,j_1))^2 \\
&\quad + \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} EY(i_2,j_1) \sum_{i_3=1}^{n}(EY(i_3,j_1))^2 \\
&= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{i_2=1}^{n} EY(i_2,j_1) \sum_{i_3=1}^{n}(EY(i_3,j_1))^2 \\
&\quad - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^2 \sum_{i_3=1}^{n}(EY(i_3,j_1))^2 \\
&< 0,
\end{aligned}$$

$$
\begin{aligned}
A_{5,12} &= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{j_2=1}^{n} EY(i_2,j_2)(EY(i_2,j_1))^2 \\
&\quad - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} (EY(i_2,j_1))^3 \\
&= -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{i_2=1}^{n}(EY(i_2,j_1))^3 + \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^4 \\
&\leq \sqrt{n}\delta_4, \\
A_{5,13} &= -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{j_2=1}^{n} EY(i_2,j_2) \sum_{i_3=1}^{n}(EY(i_3,j_2))^2 \\
&\quad + \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} EY(i_2,j_1) \sum_{i_3=1}^{n}(EY(i_3,j_1))^2 \\
&= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{i_2=1}^{n} EY(i_2,j_1) \sum_{i_3=1}^{n}(EY(i_3,j_1))^2 \\
&\quad - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^2 \sum_{i_3=1}^{n}(EY(i_3,j_1))^2 \\
&< 0, \quad \text{and} \\
A_{5,14} &= \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{j_2=1}^{n} EY(i_2,j_2)(EY(i_1,j_2))^2 \\
&\quad - \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^3 \sum_{\substack{i_2 \\ i_2\neq i_1}} EY(i_2,j_1) \\
&= -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^3 \sum_{i_2=1}^{n} EY(i_2,j_1) + \sum_{i_1=1}^{n}\sum_{j_1=1}^{n}(EY(i_1,j_1))^4 \\
&\leq \sqrt{n}\delta_4.
\end{aligned}
$$

Thus

$$B(i_3,j_3) \leq n^2\sqrt{n}\delta_4 + 4\sqrt{n}\delta_4.$$

**case 4.** $(k,l) = (i_1,j_2)$

We note that

$$B(i_1,j_2) = A_{5,2} + A_{5,5} + A_{5,14} + A_{5,15}$$

where

$$A_{5,15} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}}\sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2)EY(i_1,j_2)EY(i_2,j_1).$$

Using the fact that $2ab \leq a^2+b^2$,

$$\begin{aligned}
A_{5,15} &\leq \frac{1}{2}\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} E|Y(i_1,j_1)| \sum_{\substack{i_2 \\ i_2\neq i_1}}\sum_{\substack{j_2 \\ j_2\neq j_1}} E|Y(i_2,j_2)|\{EY^2(i_1,j_2)+EY^2(i_2,j_1)\} \\
&\leq \frac{1}{2}\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} E|Y(i_1,j_1)| \sum_{i_2=1}^{n}\sum_{j_2=1}^{n} E|Y(i_2,j_2)| \sum_{i_3=1}^{n}\sum_{j_3=1}^{n} EY^2(i_3,j_3) \\
&= \frac{1}{2}\Big(\sum_{i=1}^{n}\sum_{j=1}^{n} E|Y(i,j)|\Big)^2 \sum_{i=1}^{n}\sum_{j=1}^{n} EY^2(i,j) \\
&\leq \frac{n^2}{2}\Big(\sum_{i=1}^{n}\sum_{j=1}^{n} EY^2(i,j)\Big)^2 \\
&\leq \frac{n^4}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} EY^4(i,j) \\
&= \frac{n^4\sqrt{n}}{2}\delta_4. \qquad (3.28)
\end{aligned}$$

Thus

$$B(i_1,j_2) \leq 2\sqrt{n}\delta_4 + \frac{n^4\sqrt{n}}{2}\delta_4.$$

**case 5.** $(k,l)=(i_1,j_3)$

We note that

$$B(i_1,j_3) = A_{5,1}+A_{5,3}+A_{5,5}+A_{5,9}+A_{5,14}+A_{5,16}$$

where

$$A_{5,16} = -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}}\sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2) \sum_{j_3=1}^{n} EY(i_2,j_3)EY(i_1,j_3).$$

By the same argument as (3.28), $A_{5,16} \leq \frac{n^4\sqrt{n}}{2}\delta_4$. Thus

$$B(i_1,j_3) \leq 3\sqrt{n}\delta_4 + \frac{n^4\sqrt{n}}{2}\delta_4.$$

**case 6.** $(k,l)=(i_2,j_1)$

Note that

$$B(i_2,j_1) \leq A_{5,3}+A_{5,6}+A_{5,12}+A_{5,15} \leq 2\sqrt{n}\delta_4 + \frac{n^4\sqrt{n}}{2}\delta_4.$$

**case 7.** $(k,l) = (i_2, j_3)$
Note that

$$B(i_2, j_3) \le A_{5,3} + A_{5,5} + A_{5,7} + A_{5,10} + A_{5,12} + A_{5,16} \le 3\sqrt{n}\delta_4 + \frac{n^4\sqrt{n}}{2}\delta_4.$$

**case 8.** $(k,l) = (i_3, j_1)$
We observe that

$$B(i_3, j_1) = A_{5,1} + A_{5,2} + A_{5,6} + A_{5,11} + A_{5,12} + A_{5,17}$$

where

$$A_{5,17} = -\sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2) \sum_{i_3=1}^{n} EY(i_3,j_1)EY(i_3,j_2).$$

We can use the same argument of (3.28), $A_{5,17} \le \frac{n^4\sqrt{n}}{2}\delta_4$. Thus

$$B(i_3, j_1) \le 3\sqrt{n}\delta_4 + \frac{n^4\sqrt{n}}{2}\delta_4.$$

**case 9.** $(k,l) = (i_3, j_2)$
Note that

$$B(i_3, j_2) = A_{5,2} + A_{5,6} + A_{5,7} + A_{5,13} + A_{5,14} + A_{5,17} \le 3\sqrt{n}\delta_4 + \frac{n^4\sqrt{n}}{2}\delta_4.$$

It follows from cases 1-9 and $n \ge 39$ that

$$\begin{aligned}
A_5 = {} & \frac{1}{n(n-1)(n-2)(n-3)} \sum_{i_1=1}^{n}\sum_{j_1=1}^{n} EY(i_1,j_1) \sum_{\substack{i_2 \\ i_2\neq i_1}} \sum_{\substack{j_2 \\ j_2\neq j_1}} EY(i_2,j_2) \sum_{\substack{i_3 \\ i_3\neq i_1,i_2}} \sum_{\substack{j_3 \\ j_3\neq j_1,j_2}} \\
& EY(i_3,j_3) \sum_{\substack{i_4 \\ i_4\neq i_1,i_2,i_3}} \sum_{\substack{j_4 \\ j_4\neq j_1,j_2,j_3}} EY(i_4,j_4) \\
= {} & \frac{1}{n(n-1)(n-2)(n-3)} \sum_{k=1}^{3}\sum_{l=1}^{3} B(i_k, j_l) \\
\le {} & \frac{\left(24\sqrt{n}\delta_4 + 3n^2\sqrt{n}\delta_4 + 3n^4\sqrt{n}\delta_4\right)}{n(n-1)(n-2)(n-3)} \\
= {} & \frac{24}{n(n-1)(n-2)(n-3)}\sqrt{n}\delta_4 + 3\left(\frac{n}{n-1}\right)\left(\frac{1}{n-2}\right)\left(\frac{1}{n-3}\right)\sqrt{n}\delta_4 \\
& + 3\left(\frac{n}{n-1}\right)\left(\frac{n}{n-2}\right)\left(\frac{n}{n-3}\right)\sqrt{n}\delta_4 \\
\le {} & 3.517\sqrt{n}\delta_4. \qquad (3.29)
\end{aligned}$$

Now, we conclude from (3.21), (3.22), (3.23), (3.25), (3.27) and (3.29) that

$$ES^4(\rho) \leq 4.678\sqrt{n}\delta_4.$$

□

**Lemma 3.5.** *Let $S(\rho)$ be defined as in (3.12). Then for $n \geq 6^{(d+1)} + 3$,*

$$ES^4(\rho) \leq 4.678\sqrt{n}\delta_4.$$

*Proof.* From Lemma 3.4, it suffices to proof the lemma in case of $d \geq 2$. Note that

$$ES^4(\rho) = A_1 + A_2 + A_3 + A_4 + A_5 \tag{3.30}$$

where

$$\begin{aligned}
A_1 &= \sum_{j=1}^{n} EY^4(j, \rho_1(j), ..., \rho_d(j)) \\
A_2 &= \sum_{j=1}^{n} \sum_{\substack{k \\ k \neq j}} EY^3(j, \rho_1(j), ..., \rho_d(j)) Y(k, \rho_1(k), ..., \rho_d(k)) \\
A_3 &= \sum_{j=1}^{n} \sum_{\substack{k \\ k \neq j}} EY^2(j, \rho_1(j), ..., \rho_d(j)) Y^2(k, \rho_1(k), ..., \rho_d(k)) \\
A_4 &= \sum_{j=1}^{n} \sum_{\substack{k \\ k \neq j}} \sum_{\substack{l \\ l \neq j,k}} EY^2(j, \rho_1(j), ..., \rho_d(j)) Y(k, \rho_1(k), ..., \rho_d(k)) \\
&\quad \times Y(l, \rho_1(l), ..., \rho_d(l)) \\
A_5 &= \sum_{j=1}^{n} \sum_{\substack{k \\ k \neq j}} \sum_{\substack{l \\ l \neq j,k}} \sum_{\substack{m \\ m \neq j,k,l}} EY(j, \rho_1(j), ..., \rho_d(j)) Y(k, \rho_1(k), ..., \rho_d(k)) \\
&\quad \times Y(l, \rho_1(l), ..., \rho_d(l)) Y(m, \rho_1(m), ..., \rho_d(m)).
\end{aligned} \tag{3.31}$$

We observe that

$$\begin{aligned}
(\frac{n}{n-1})^d &= (1 + \frac{1}{(n-1)})^d \\
&= 1 + \frac{d}{(n-1)} + \sum_{r=2}^{d} \binom{d}{r} \frac{1}{(n-1)^r} \\
&\leq 1 + \frac{d}{(n-1)} + \sum_{r=2}^{d} \frac{d^r}{r!(n-1)^r}
\end{aligned}$$

$$\leq 1+\frac{d}{(n-1)}+\frac{1}{2}\sum_{r=2}^{d}(\frac{d}{n-1})^r$$

$$\leq 1.03 \tag{3.32}$$

where we have used the fact that $\frac{d}{n-1}\leq\frac{d}{6^{(d+1)}+2}\leq\frac{1}{50}$ in the last inequality. By (3.32) and the fact that $n\geq 6^{(d+1)}+3\geq 219$, we can follow argument of Lemma 3.4 to show that

$$A_1\leq 0.004\sqrt{n}\delta_4,\ A_2\leq 0.000002\sqrt{n}\delta_4,\ A_3\leq 1.03\sqrt{n}\delta_4,\ A_4\leq 1.09\sqrt{n}\delta_4. \tag{3.33}$$

Hence we need to bound only $A_5$. Note that on expansion,

$$\sum_{j=1}^{n}\sum_{j_1=1}^{n}...\sum_{j_d=1}^{n}EY(j,j_1,...,j_d)\sum_{\substack{k\\k\neq j}}\sum_{\substack{k_1\\k_1\neq j_1}}...\sum_{\substack{k_d\\k_d\neq j_d}}EY(k,k_1,...,k_d)\sum_{\substack{l\\l\neq j,k}}\sum_{\substack{l_1\\l_1\neq j_1,k_1}}...\sum_{\substack{l_d\\l_d\neq j_d,k_d}}$$
$$EY(l,l_1,...,l_d)\sum_{\substack{m\\m\neq j,k,l}}\sum_{\substack{m_1\\m_1\neq j_1,k_1,l_1}}...\sum_{\substack{m_d\\m_d\neq j_d,k_d,l_d}}EY(m,m_1,...,m_d)$$

consists of a sum of $3^{d+1}$ terms each of the form

$$B(q_1,...,q_{d+1})=(-1)^{d+1}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}...\sum_{j_{d+1}=1}^{n}EY(j_1,j_2,..,j_{d+1})\sum_{\substack{k_1\\k_1\neq j_1}}\sum_{\substack{k_2\\k_2\neq j_2}}...\sum_{\substack{k_{d+1}\\k_{d+1}\neq j_{d+1}}}$$
$$EY(k_1,k_2,...,k_{d+1})\sum_{\substack{l_1\\l_1\neq j_1,k_1}}\sum_{\substack{l_2\\l_2\neq j_2,k_2}}...\sum_{\substack{l_{d+1}\\l_{d+1}\neq j_{d+1},k_{d+1}}}$$
$$EY(l_1,l_2...,l_{d+1})EY(q_1,q_2...,q_{d+1})$$

where $q_i\in\{j_i,k_i,l_i\}$ for $i=1,...,d+1$.

**case 1.** $\forall i\in\{1,...,d+1\},\ q_i=l_i$.

In this case, we have

$$B(l_1,...,l_{d+1})$$
$$\leq\Big(\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}...\sum_{j_{d+1}=1}^{n}E|Y(j_1,j_2,..,j_{d+1})|\Big)^2\sum_{l_1=1}^{n}\sum_{l_2=1}^{n}...\sum_{l_{d+1}=1}^{n}(EY(l_1,l_2...,l_{d+1}))^2$$
$$\leq n^{d+1}\Big(\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}...\sum_{j_{d+1}=1}^{n}EY^2(j_1,j_2...,j_{d+1})\Big)^2$$
$$\leq n^{2(d+1)}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}...\sum_{j_{d+1}=1}^{n}EY^4(j_1,j_2...,j_{d+1})$$
$$=n^{3d+1}\sqrt{n}\delta_4.$$

**case 2.** $\exists i \in \{1, ..., d+1\}$, $q_i \neq l_i$.

It suffices to prove only case $q_1 = l_1, ..., q_s = l_s, q_{s+1} \neq l_{s+1}, ..., q_{d+1} \neq l_{d+1}$ for some $s \in \{0, ..., d\}$.

We observe that

$$B(q_1, ..., q_{d+1}) = (-1)^{d+1} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_{d+1}=1}^{n} EY(j_1, j_2, ..., j_{d+1}) \sum_{\substack{k_1 \\ k_1 \neq j_1}} \sum_{\substack{k_2 \\ k_2 \neq j_2}} \cdots \sum_{\substack{k_{d+1} \\ k_{d+1} \neq j_{d+1}}} EY(k_1, k_2, ..., k_{d+1}) \sum_{\substack{l_1 \\ l_1 \neq j_1, k_1}} \cdots \sum_{\substack{l_s \\ l_s \neq j_s, k_s}} EY(l_1, ..., l_s, q_{s+1}, ..., q_{d+1}) \sum_{\substack{l_{s+1} \\ l_{s+1} \neq j_{s+1}, k_{s+1}}} \cdots \sum_{\substack{l_{d+1} \\ l_{d+1} \neq j_{d+1}, k_{d+1}}} EY(l_1, l_2, ..., l_{d+1}).$$

Note that for fixed $j_1, j_2, ..., j_{d+1}, k_1, k_2, ..., k_{d+1}$,

$$\sum_{\substack{l_1 \\ l_1 \neq j_1, k_1}} \cdots \sum_{\substack{l_s \\ l_s \neq j_s, k_s}} EY(l_1, ..., l_s, q_{s+1}, ..., q_{d+1}) \sum_{\substack{l_{s+1} \\ l_{s+1} \neq j_{s+1}, k_{s+1}}} \cdots \sum_{\substack{l_{d+1} \\ l_{d+1} \neq j_{d+1}, k_{d+1}}} EY(l_1, l_2, ..., l_{d+1})$$

consists of a sum of $2^{d-s+1}$ terms each of the form

$$C(r_{s+1}, ..., r_{d+1}) = (-1)^{d-s+1} \sum_{\substack{l_1 \\ l_1 \neq j_1, k_1}} \cdots \sum_{\substack{l_s \\ l_s \neq j_s, k_s}} EY(l_1, ..., l_s, q_{s+1}, ..., q_{d+1}) EY(l_1, ..., l_s, r_{s+1}, ..., r_{d+1})$$

where $r_i \in \{j_i, k_i\}$ for $i = s+1, ..., d+1$, and

$$\begin{aligned} &|C(r_{s+1}, ..., r_{d+1})| \\ &\leq \frac{1}{2} \sum_{\substack{l_1 \\ l_1 \neq j_1, k_1}} \cdots \sum_{\substack{l_s \\ l_s \neq j_s, k_s}} \{EY^2(l_1, ..., l_s, q_{s+1}, ..., q_{d+1}) + EY^2(l_1, ..., l_s, r_{s+1}, ..., r_{d+1})\} \\ &\leq \sum_{l_1=1}^{n} \cdots \sum_{l_{d+1}=1}^{n} EY^2(l_1, ..., l_{d+1}). \end{aligned}$$

By this fact and $s \geq 0$,

$$\begin{aligned}
B(q_1, ..., q_{d+1}) &\leq 2^{d-s+1} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} ... \sum_{j_{d+1}=1}^{n} E|Y(j_1, j_2, ..., j_{d+1})| \\
&\quad \sum_{\substack{k_1 \\ k_1 \neq j_1}} \sum_{\substack{k_2 \\ k_2 \neq j_2}} ... \sum_{\substack{k_{d+1} \\ k_{d+1} \neq j_{d+1}}} E|Y(k_1, k_2, ..., k_{d+1})| \\
&\quad \sum_{l_1=1}^{n} \sum_{l_2=1}^{n} ... \sum_{l_{d+1}=1}^{n} EY^2(l_1, l_2, ..., l_{d+1}) \\
&\leq 2^{d-s+1} \Big( \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} ... \sum_{j_{d+1}=1}^{n} E|Y(j_1, j_2, .., j_{d+1})| \Big)^2 \\
&\quad \sum_{l_1=1}^{n} \sum_{l_2=1}^{n} ... \sum_{l_{d+1}=1}^{n} EY^2(l_1, l_2 ..., l_{d+1}) \\
&\leq 2^{d+1} n^{d+1} \Big( \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} ... \sum_{j_{d+1}=1}^{n} EY^2(j_1, j_2, .., j_{d+1}) \Big)^2 \\
&\leq 2^{d+1} n^{2d+2} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} ... \sum_{j_{d+1}=1}^{n} EY^4(j_1, j_2, .., j_{d+1}) \\
&= 2^{d+1} n^{3d+1} \sqrt{n} \delta_4.
\end{aligned}$$

It follows from case 1. and case 2. that

$$\begin{aligned}
A_5 &\leq \frac{3^{d+1} 2^{d+1} n^{3d+1} \sqrt{n}}{(n(n-1)(n-2)(n-3))^d} \delta_4 \\
&= 6^{d+1} (\frac{n}{n-1})^d (\frac{n}{n-2})^d (\frac{n}{n-3}) \frac{1}{(n-3)^{d-1}} \sqrt{n} \delta_4 \\
&\leq 1.075 \sqrt{n} \delta_4 \qquad (3.34)
\end{aligned}$$

where we have used (3.32) and the fact that $n \geq 6^{(d+1)} + 3 \geq 219$ in the last inequality. Now we conclude from (3.30), (3.33) and (3.34) that, for $d \geq 2$,

$$ES^4(\rho) \leq 3.199 \sqrt{n} \delta_4 \leq 4.678 \sqrt{n} \delta_4.$$

□

**Lemma 3.6.** *Let*

$$S(\tau) = \sum_{i=1}^{n} Y(i, \tau_1(i), ..., \tau_d(i)).$$

*Then $S(\tau)$ and $G^2\mathbb{I}(A)$ are conditionally independent given $\tau_1, ..., \tau_d$ where*

$$A = \{\tau_i(I) \neq L_i, \tau_i(K) \neq M_i, \tau_i(I) \neq M_i, \tau_i(K) \neq L_i : i = 1, 2, ..., d\}$$

*and*

$$G = Y(I, M_1, ..., M_d) + Y(K, L_1, ..., L_d) - Y(I, L_1, ..., L_d) - Y(K, M_1, ..., M_d).$$

*Proof.* We shall first prove lemma in case of $d = 1$. Let $\alpha$ be a permutation of $\{1, 2, ..., n\}$ and $x, y \in \mathbb{R}$. We observe from (3.7) that

$$\begin{aligned}
&P(\sum_{j=1}^{n} Y(j, \tau_1(j)) \leq x, G^2\mathbb{I}(A) \leq y, \tau_1 = \alpha) \\
&= P(\sum_{j=1}^{n} Y(j, \tau_1(j)) \leq x, G^2\mathbb{I}(A) \leq y, \tau_1 = \alpha, \mathbb{I}(A) = 0) \\
&\quad + P(\sum_{j=1}^{n} Y(j, \tau_1(j)) \leq x, G^2\mathbb{I}(A) \leq y, \tau_1 = \alpha, \mathbb{I}(A) = 1) \\
&= P(\sum_{j=1}^{n} Y(j, \tau_1(j)) \leq x, 0 \leq y, \tau_1 = \alpha, (\tau_1(I) = L) \vee (\tau_1(K) = M) \vee (\tau_1(I) = M) \\
&\quad \vee (\tau_1(K) = L)) + P(\sum_{j=1}^{n} Y(j, \tau_1(j)) \leq x, [Y(I, M) + Y(K, L) - Y(I, L) \\
&\quad - Y(K, M)]^2 \leq y, \tau_1 = \alpha, \tau_1(I) \neq L, \tau_1(K) \neq M, \tau_1(I) \neq M, \tau_1(K) \neq L) \\
&= P(\sum_{j=1}^{n} Y(j, \alpha(j)) \leq x, 0 \leq y, \tau_1 = \alpha, (\tau_1(I) = L) \vee (\tau_1(K) = M) \vee (\tau_1(I) = M) \\
&\quad \vee (\tau_1(K) = L)) + \sum_{i=1}^{n} \sum_{\substack{k \\ k \neq i}} \sum_{l=1}^{n} \sum_{\substack{m \\ m \neq l}} P(\sum_{j=1}^{n} Y(j, \alpha(j)) \leq x, [Y(i, m) + Y(k, l) \\
&\quad - Y(i, l) - Y(k, m)]^2 \leq y, \tau_1 = \alpha, \alpha(i) \neq l, \alpha(k) \neq m, \alpha(i) \neq m, \alpha(k) \neq l, I = i, \\
&\quad K = k, L = l, M = m) \\
&= P(\sum_{j=1}^{n} Y(j, \alpha(j)) \leq x)\{P(0 \leq y, \tau_1 = \alpha, (\tau_1(I) = L) \vee (\tau_1(K) = M) \vee (\tau_1(I) = M) \\
&\quad \vee (\tau_1(K) = L)) + \sum_{i=1}^{n} \sum_{\substack{k \\ k \neq i}} \sum_{l=1}^{n} \sum_{\substack{m \\ m \neq l}} P([Y(i, m) + Y(k, l) - Y(i, l) - Y(k, m)]^2 \leq y, \\
&\quad \tau_1 = \alpha, \alpha(i) \neq l, \alpha(k) \neq m, \alpha(i) \neq m, \alpha(k) \neq l, I = i, K = k, L = l, M = m)\}
\end{aligned}$$

$$
\begin{aligned}
&= P(\sum_{j=1}^{n} Y(j, \alpha(j)) \leq x)\{P(0 \leq y, \tau_1 = \alpha, (\tau_1(I) = L) \vee (\tau_1(K) = M) \vee (\tau_1(I) = M) \\
&\quad \vee (\tau_1(K) = L)) + P([Y(I, M) + Y(K, L) - Y(I, L) - Y(K, M)]^2 \leq y, \tau_1 = \alpha, \\
&\quad \tau_1(I) \neq L, \tau_1(K) \neq M, \tau_1(I) \neq M, \tau_1(K) \neq L)\} \\
&= P(\sum_{j=1}^{n} Y(j, \alpha(j)) \leq x)\{P(G^2 \mathbb{I}(A) \leq y, \tau_1 = \alpha, \mathbb{I}(A) = 0) \\
&\quad + P(G^2 \mathbb{I}(A) \leq y, \tau_1 = \alpha, \mathbb{I}(A) = 1)\} \\
&= \frac{1}{P(\tau_1 = \alpha)} P(\sum_{j=1}^{n} Y(j, \alpha(j)) \leq x, \tau_1 = \alpha) P(G^2 \mathbb{I}(A) \leq y, \tau_1 = \alpha) \\
&= P(\sum_{j=1}^{n} Y(j, \tau_1(j)) \leq x, \tau_1 = \alpha) P(G^2 \mathbb{I}(A) \leq y | \tau_1 = \alpha).
\end{aligned}
$$

By this fact,

$$
\begin{aligned}
&P(\sum_{j=1}^{n} Y(j, \tau_1(j)) \leq x, G^2 \mathbb{I}(A) \leq y | \tau_1 = \alpha) \\
&= \frac{1}{P(\tau_1 = \alpha)} P(\sum_{j=1}^{n} Y(j, \tau_1(j)) \leq x, G^2 \mathbb{I}(A) \leq y, \tau_1 = \alpha) \\
&= \frac{1}{P(\tau_1 = \alpha)} P(\sum_{j=1}^{n} Y(j, \tau_1(j)) \leq x, \tau_1 = \alpha) P(G^2 \mathbb{I}(A) \leq y | \tau_1 = \alpha) \\
&= P(\sum_{j=1}^{n} Y(j, \tau_1(j)) \leq x | \tau_1 = \alpha) P(G^2 \mathbb{I}(A) \leq y | \tau_1 = \alpha).
\end{aligned}
$$

By the same argument, we can prove lemma in the case of an arbitary $d$. $\square$

## 3.2 Proof of Theorem 3.1

*Proof.* We shall prove Theorem 3.1 by using Stein's method. Note from (2.3) and Lemma 3.3 that

$$\begin{aligned}
\big|P\big(W \le z\big) - \Phi(z)\big| &= |Eg_z'(W) - EWg_z(W)| \\
&= |Eg_z'(S(\tau)) - ES(\rho)g_z(S(\rho))| \\
&= |Eg_z'(S(\tau)) - E\int_{-\infty}^{\infty} g_z'(S(\rho)+t)M(t)dt + \Delta g_z(S(\rho))| \\
&\le |Eg_z'(S(\tau))\int_{-\infty}^{\infty} M(t)dt - E\int_{-\infty}^{\infty} g_z'(S(\rho)+t)M(t)dt| \\
&\quad + |Eg_z'(S(\tau))E\int_{-\infty}^{\infty} M(t)dt - Eg_z'(S(\tau))\int_{-\infty}^{\infty} M(t)dt| \\
&\quad + |Eg_z'(S(\tau)) - Eg_z'(S(\tau))E\int_{-\infty}^{\infty} M(t)dt| \\
&\quad + |\Delta g_z(S(\rho))| \\
&= |T_1| + |T_2| + |T_3| + |T_4| \qquad (3.35)
\end{aligned}$$

where

$$\begin{aligned}
T_1 &= Eg_z'(S(\tau))\int_{-\infty}^{\infty} M(t)dt - E\int_{-\infty}^{\infty} g_z'(S(\rho)+t)M(t)dt, \\
T_2 &= Eg_z'(S(\tau))E\int_{-\infty}^{\infty} M(t)dt - Eg_z'(S(\tau))\int_{-\infty}^{\infty} M(t)dt, \\
T_3 &= Eg_z'(S(\tau)) - Eg_z'(S(\tau))E\int_{-\infty}^{\infty} M(t)dt \\
\text{and} \qquad T_4 &= \Delta g_z(S(\rho)).
\end{aligned}$$

By the fact that

$$g_z'(w+u) - g_z'(w+v) \le \begin{cases} 1 & \text{if } \ w+u<z, w+v>z, \\ \big(|w| + \frac{\sqrt{2\pi}}{4}\big)(|u|+|v|) & \text{if } \ u \ge v, \\ 0 & \text{ortherwise} \end{cases} \qquad (3.36)$$

(see, [6],pp.247), we have

$$\begin{aligned}
T_1 &= E\int_{-\infty}^{\infty} [g_z'(S(\tau)) - g_z'(S(\rho)+t)]M(t)dt \\
&= E\int_{-\infty}^{\infty} [g_z'(S(\rho)+\Delta S) - g_z'(S(\rho)+t)]M(t)dt \quad \text{where} \quad \Delta S = S(\tau) - S(\rho)
\end{aligned}$$

$$\leq E\int_{\substack{S(\rho)+\Delta S<z\\ S(\rho)+t>z}} M(t)dt + E\int_{t-\Delta S\leq 0}\big(|S(\rho)|+\frac{\sqrt{2\pi}}{4}\big)(|\Delta S|+|t|)M(t)dt$$
$$= M_1 + M_2$$

where

$$M_1 = E\int_{\substack{S(\rho)+\Delta S<z\\ S(\rho)+t>z}} M(t)dt$$

and

$$M_2 = E\int_{t-\Delta S\leq 0}\big(|S(\rho)|+\frac{\sqrt{2\pi}}{4}\big)(|\Delta S|+|t|)M(t)dt.$$

For each $\delta \geq 0$ and $a, b \in \mathbb{R}$ which is $a < b$, define a function $f_\delta$ by

$$f_\delta(t) = \begin{cases} -\frac{1}{2}(b-a)-\delta & \text{if } \ t < a-\delta, \\ -\frac{1}{2}(b+a)+t & \text{if } \ a-\delta \leq t \leq b+\delta, \\ \frac{1}{2}(b-a)+\delta & \text{if } \ b+\delta < t. \end{cases}$$

It is easy to see that

$$|f_\delta(t)| \leq \frac{1}{2}(b-a)+\delta \ \text{ for every } \ t\in\mathbb{R}, \tag{3.37}$$

and

$$\begin{aligned} E\int_{-\infty}^{\infty} f'_\delta(S(\rho)+t)M(t)dt &\geq E\int_{\substack{a\leq S(\rho)\leq b\\ |t|\leq\delta}} f'_\delta(S(\rho)+t)M(t)dt \\ &= E\int_{\substack{a\leq S(\rho)\leq b\\ |t|\leq\delta}} M(t)dt \\ &= E\int_{|t|\leq\delta}\mathbb{I}(a\leq S(\rho)\leq b)M(t)dt. \end{aligned} \tag{3.38}$$

By Lemma 3.3 and (3.38),

$$\begin{aligned} E\int_{|t|\leq\delta}\mathbb{I}(a\leq S(\rho)\leq b)M(t)dt &\leq E\int_{-\infty}^{\infty} f'_\delta(S(\rho)+t)M(t)dt \\ &\leq ES(\rho)f_\delta(S(\rho)) + \frac{1}{n-1}[Ef_\delta^2(S(\rho))]^{\frac{1}{2}}. \end{aligned} \tag{3.39}$$

We note that if $|t| > |\widetilde{S}(\rho)-S(\rho)|$,

$$M(t) = 0.$$

Thus we conclude from (3.37) and (3.39) that

$$\begin{aligned}
M_1 &= E\int_{t-\Delta S>0} \mathbb{I}(z-t<S(\rho)<z-\Delta S)M(t)dt \\
&\le E\int_{-\infty}^{\infty} \mathbb{I}(z-t<S(\rho)<z-\Delta S)M(t)dt \\
&= E\int_{|t|\le|\widetilde{S}(\rho)-S(\rho)|} \mathbb{I}(z-t<S(\rho)<z-\Delta S)M(t)dt \\
&\le E\int_{|t|\le|\widetilde{S}(\rho)-S(\rho)|} \mathbb{I}(z-|\widetilde{S}(\rho)-S(\rho)|<S(\rho)<z-\Delta S)M(t)dt \\
&\le ES(\rho)f_{|\widetilde{S}(\rho)-S(\rho)|}(S(\rho)) + \frac{1}{n-1}[Ef^2_{|\widetilde{S}(\rho)-S(\rho)|}(S(\rho))]^{\frac{1}{2}} \\
&\le E|S(\rho)|[\frac{3}{2}|\widetilde{S}(\rho)-S(\rho)| - \frac{\Delta S}{2}] + \frac{1}{n-1}\{E[\frac{3}{2}|\widetilde{S}(\rho)-S(\rho)| - \frac{\Delta S}{2}]^2\}^{\frac{1}{2}} \\
&\le \frac{3}{2}E|S(\rho)||\widetilde{S}(\rho)-S(\rho)| + \frac{1}{2}E|\Delta S||S(\rho)| \\
&\quad + \frac{1}{\sqrt{2}(n-1)}\{9E[\widetilde{S}(\rho)-S(\rho)]^2 + E(\Delta S)^2\}^{\frac{1}{2}}. \qquad (3.40)
\end{aligned}$$

By Lemma 3.3, when $g(w) = w$, we have

$$\begin{aligned}
E|\widetilde{S}(\rho)-S(\rho)|^2 &= \frac{4}{n}E\int_{-\infty}^{\infty} M(t)dt \\
&\le \frac{4}{n}(ES^2(\rho) + \frac{1}{n-1}[ES^2(\rho)]^{\frac{1}{2}}) \\
&= \frac{4}{n-1}. \qquad (3.41)
\end{aligned}$$

Since $ES^2(\rho) = 1$ and (3.41),

$$E|S(\rho)||\widetilde{S}(\rho)-S(\rho)| \le \sqrt{E|\widetilde{S}(\rho)-S(\rho)|^2} \le \frac{2}{\sqrt{n-1}}. \qquad (3.42)$$

Let

$$\delta_2 = \frac{1}{n^{d+\frac{1}{2}}}\sum_{i_1=1}^{n}\dots\sum_{i_{d+1}=1}^{n} E|Y(i_1,\dots,i_{d+1})|^2.$$

Since $E|\Delta S|^k$ is a sum of $4(d+1)$ terms each of the form $E|Y(I,\rho_1(I),\dots,\rho_d(I))|^k$,

$$E|\Delta S|^k \le \frac{4(d+1)\delta_k}{n^{\frac{k-1}{2}}} \qquad (3.43)$$

for $k = 2, 4$. By this fact

$$E|\Delta S||S(\rho)| \le \sqrt{E|\Delta S|^2} \le \frac{2\sqrt{d+1}\delta_2^{\frac{1}{2}}}{n^{\frac{1}{4}}}. \qquad (3.44)$$

From (3.40)-(3.44),

$$M_1 \le \frac{3}{\sqrt{n-1}} + \frac{\sqrt{d+1}\delta_2^{\frac{1}{2}}}{n^{\frac{1}{4}}} + \frac{1}{\sqrt{2}(n-1)}\left\{\frac{36}{n-1} + \frac{4(d+1)\delta_2}{\sqrt{n}}\right\}^{\frac{1}{2}}$$
$$\le \frac{3.202}{\sqrt{n}} + \frac{1.037\sqrt{d+1}\delta_2^{\frac{1}{2}}}{n^{\frac{1}{4}}}$$

where we have used the fact that $n \ge 39$ in the last inequality. In order to bound $M_2$, we need to bound $E|\widetilde{S}(\rho) - S(\rho)|^k,\ k = 3, 4$.

Let

$$\delta_3 = \frac{1}{n^d}\sum_{i_1=1}^{n} \dots \sum_{i_{d+1}=1}^{n} E|Y(i_1, ..., i_{d+1})|^3.$$

We note that for $k = 3, 4$,

$$E|S_1|^k = \frac{1}{n^{d+1}}\sum_{i=1}^{n}\sum_{i_1=1}^{n} \dots \sum_{i_d=1}^{n} |Y(i, i_1, ..., i_d)|^k = \frac{\delta_k}{n^{\frac{k-1}{2}}} \tag{3.45}$$

which implies

$$E|\widetilde{S}(\rho) - S(\rho)|^k = E|S_1 + S_2 - S_3 - S_4|^k \le 4^k E|S_1|^k = \frac{4^k\delta_k}{n^{\frac{k-1}{2}}} \tag{3.46}$$

where we have used the fact that $S_1, S_2, S_3$ and $S_4$ have the same distribution in the first inequlity. By Lemma 3.5, (3.43) and (3.46),

$$E\int_{\mathbb{R}} |S(\rho)||\Delta S|M(t)dt = \frac{n}{4}E|S(\rho)||\Delta S|(\widetilde{S}(\rho) - S(\rho))^2$$
$$\le \frac{n}{4}\{ES^2(\rho)(\Delta S)^2\}^{\frac{1}{2}}\{E|\widetilde{S}(\rho) - S(\rho)|^4\}^{\frac{1}{2}}$$
$$\le \frac{n}{4}\{ES^4(\rho)\}^{\frac{1}{4}}\{E(\Delta S)^4\}^{\frac{1}{4}}\{E|\widetilde{S}(\rho) - S(\rho)|^4\}^{\frac{1}{2}}$$
$$\le \frac{n}{4}\{4.678\sqrt{n}\delta_4\}^{\frac{1}{4}}\{\frac{4(d+1)\delta_4}{n\sqrt{n}}\}^{\frac{1}{4}}\{\frac{256\delta_4}{n\sqrt{n}}\}^{\frac{1}{2}}$$
$$= 8.314(d+1)^{\frac{1}{4}}\delta_4. \tag{3.47}$$

Hence by Lemma 3.5, (3.43), (3.46) and (3.47),

$$M_2 \le E\int_{\mathbb{R}} |S(\rho)||\Delta S|M(t)dt + E\int_{\mathbb{R}} |S(\rho)||t|M(t)dt$$
$$+ \frac{\sqrt{2\pi}}{4}E\int_{\mathbb{R}} |\Delta S|M(t)dt + \frac{\sqrt{2\pi}}{4}E\int_{\mathbb{R}} |t|M(t)dt$$

$$\leq 8.314(d+1)^{\frac{1}{4}}\delta_4 + \frac{n}{8}E|S(\rho)||\widetilde{S}(\rho) - S(\rho)|^3$$
$$+ \frac{\sqrt{2\pi}n}{16}E|\Delta S|(\widetilde{S}(\rho) - S(\rho))^2 + \frac{\sqrt{2\pi}n}{32}E|\widetilde{S}(\rho) - S(\rho)|^3$$
$$\leq 8.314(d+1)^{\frac{1}{4}}\delta_4 + \frac{n}{8}\{ES^4(\rho)\}^{\frac{1}{4}}\{E[\widetilde{S}(\rho) - S(\rho)]^4\}^{\frac{3}{4}}$$
$$+ \frac{\sqrt{2\pi}n}{16}\{E|\Delta S|^2\}^{\frac{1}{2}}\{E|\widetilde{S}(\rho) - S(\rho)|^4\}^{\frac{1}{2}} + \frac{\sqrt{2\pi}n}{32}\{E|\widetilde{S}(\rho) - S(\rho)|^4\}^{\frac{3}{4}}$$
$$\leq 8.314(d+1)^{\frac{1}{4}}\delta_4 + 11.765\delta_4 + 5.014(d+1)\delta_2^{\frac{1}{2}}\delta_4^{\frac{1}{2}} + \frac{2\sqrt{2\pi}\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}$$

which implies that

$$T_1 \leq \frac{3.202}{\sqrt{n}} + \frac{1.037\sqrt{d+1}\delta_2^{\frac{1}{2}}}{n^{\frac{1}{4}}} + 8.314(d+1)^{\frac{1}{4}}\delta_4 + 11.765\delta_4 + 5.014(d+1)\delta_2^{\frac{1}{2}}\delta_4^{\frac{1}{2}}$$
$$+ \frac{2\sqrt{2\pi}\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}. \tag{3.48}$$

By the same argument as in (3.48), by using the fact that

$$g_z'(w+u) - g_z'(w+v) \geq \begin{cases} -1 & \text{if } \; w+u > z, w+v < z, \\ -\left(|w| + \frac{\sqrt{2\pi}}{4}\right)(|u|+|v|) & \text{if } \; u < v, \\ 0 & \text{ortherwise} \end{cases}$$

(see[6],pp.247) instead of (3.36), we can show that

$$T_1 \geq -\frac{3.202}{\sqrt{n}} - \frac{1.037\sqrt{d+1}\delta_2^{\frac{1}{2}}}{n^{\frac{1}{4}}} - 8.314(d+1)^{\frac{1}{4}}\delta_4 - 11.765\delta_4 - 5.014(d+1)\delta_2^{\frac{1}{2}}\delta_4^{\frac{1}{2}}$$
$$- \frac{2\sqrt{2\pi}\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}.$$

Hence, we have

$$|T_1| \leq \frac{3.202}{\sqrt{n}} + \frac{1.037\sqrt{d+1}\delta_2^{\frac{1}{2}}}{n^{\frac{1}{4}}} + 8.314(d+1)^{\frac{1}{4}}\delta_4 + 11.765\delta_4 + 5.014(d+1)\delta_2^{\frac{1}{2}}\delta_4^{\frac{1}{2}}$$
$$+ \frac{2\sqrt{2\pi}\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}. \tag{3.49}$$

Next, we shall bound $T_2$ by using the technique from Lemma 4.10 of Ho and Chen([9],pp.245). Let $A$ and $G$ be defined as in Lemma 3.6, i.e.,

$$A = \{\tau_i(I) \neq L_i, \tau_i(K) \neq M_i, \tau_i(I) \neq M_i, \tau_i(K) \neq L_i : i = 1, 2, ..., d\}$$

and

$$G = Y(I, M_1, ..., M_d) + Y(K, L_1, ..., L_d) - Y(I, L_1, ..., L_d) - Y(K, M_1, ..., M_d).$$

Note that $EG^2 = E^{\tau_1,\dots,\tau_d}G^2$ by independence of $\tau_1,...\tau_d$ and $G$. From this fact and (3.11),

$$\begin{aligned}
|T_2| &= \frac{n}{4}|Eg_z'(S(\tau))E(\widetilde{S}(\rho)-S(\rho))^2 - Eg_z'(S(\tau))(\widetilde{S}(\rho)-S(\rho))^2| \\
&= \frac{n}{4}|Eg_z'(S(\tau))EG^2 - Eg_z'(S(\tau))G^2| \\
&= \frac{n}{4}|Eg_z'(S(\tau))E^{\tau_1,\dots\tau_d}G^2 - Eg_z'(S(\tau))G^2| \\
&\le \frac{n}{4}|Eg_z'(S(\tau))E^{\tau_1,\dots\tau_d}G^2\mathbb{I}(A) - Eg_z'(S(\tau))G^2\mathbb{I}(A)| \\
&\quad + \frac{n}{4}|Eg_z'(S(\tau))E^{\tau_1,\dots\tau_d}G^2\mathbb{I}(A^c) - Eg_z'(S(\tau))G^2\mathbb{I}(A^c)|. \qquad (3.50)
\end{aligned}$$

By Lemma 3.6 and Proposition 2.13,

$$\begin{aligned}
E[g_z'(S(\tau))G^2\mathbb{I}(A)] &= E\{[E^{\tau_1,\dots\tau_d}g_z'(S(\tau))][E^{\tau_1,\dots\tau_d}G^2\mathbb{I}(A)]\} \\
&= EE^{\tau_1,\dots\tau_d}[g_z'(S(\tau))E^{\tau_1,\dots\tau_d}G^2\mathbb{I}(A)] \\
&= E[g_z'(S(\tau))E^{\tau_1,\dots\tau_d}G^2\mathbb{I}(A)].
\end{aligned}$$

This implies the first term on the right-hand side of (3.50) is zero. Hence

$$|T_2| \le \frac{n}{4}|Eg_z'(S(\tau))E^{\tau_1,\dots\tau_d}G^2\mathbb{I}(A^c) - Eg_z'(S(\tau))G^2\mathbb{I}(A^c)|. \qquad (3.51)$$

Let $\mathcal{B}$ be the $\sigma$-algebra generated by

$$\{I, K, L_1, ..., L_d, M_1, ..., M_d, Y(i_1, ..., i_{d+1}) : 1 \le i_1, ..., i_{d+1} \le n\}.$$

Note from Theorem 2.11 that

$$E^{\mathcal{B}}\mathbb{I}(A) = E^{I,K,L_1,\dots,L_d,M_1,\dots,M_d}\mathbb{I}(A).$$

By this fact,

$$\begin{aligned}
E^{\mathcal{B}}\mathbb{I}(A^c) &= 1 - E^{\mathcal{B}}\mathbb{I}(A) \\
&= 1 - E^{I,K,L_1,\dots,L_d,M_1,\dots,M_d}\mathbb{I}(A) \\
&= 1 - \{\frac{(n-2)(n-3)[(n-2)]!}{n!}\}^d \\
&= 1 - [1 + \frac{6-4n}{n(n-1)}]^d \\
&\le \sum_{r=1}^{d}\binom{d}{r}(\frac{4n-6}{n(n-1)})^r
\end{aligned}$$

$$\leq \sum_{r=1}^{d} \binom{d}{r} (\frac{4}{n})^r$$

$$\leq \frac{4}{n} \sum_{r=1}^{d} \binom{d}{r}. \tag{3.52}$$

From (3.52) and the fact that $\sum_{r=1}^{d} \binom{d}{r} \leq \sqrt{n}$ for $n \geq 6^{(d+1)} + 3$, we have

$$E^{\mathcal{B}}\mathbb{I}(A^c) \leq \frac{4}{\sqrt{n}}. \tag{3.53}$$

By replacing $g(W)$ by $W$ in Lemma 3.3 we have

$$|E\int_{-\infty}^{\infty} M(t)dt - 1| \leq \frac{1}{n-1}[E(S(\rho))^2]^{\frac{1}{2}} = \frac{1}{n-1}. \tag{3.54}$$

Thus, we conclude from the fact that

$$|g_z'(w)| \leq 1 \tag{3.55}$$

(see[26],pp.23), (3.51), (3.53) and (3.54),

$$\begin{aligned} |T_2| &\leq \frac{n}{4}[E|g_z'(S(\tau))||E^{\tau_1,\cdots\tau_d}G^2\mathbb{I}(A^c)| + E|g_z'(S(\tau))||G^2\mathbb{I}(A^c)|] \\ &\leq \frac{n}{2}E[G^2\mathbb{I}(A^c)] \\ &= \frac{n}{2}E[G^2E^{\mathcal{B}}\mathbb{I}(A^c)] \\ &\leq 2\sqrt{n}EG^2 \\ &= \frac{8}{\sqrt{n}}E\int_{-\infty}^{\infty} M(t)dt \\ &\leq \frac{8.416}{\sqrt{n}} \end{aligned} \tag{3.56}$$

where we have used the fact that $n \geq 39$ in the last inequality. Next, we will find a bound for $T_3$ and $T_4$. From (3.54) and (3.55),

$$T_3 \leq \frac{1}{n-1} \tag{3.57}$$

and, from the fact that $0 \leq g_z(w) \leq 1$(see[26],pp.23) and (3.15),

$$T_4 \leq \frac{1}{n-1}. \tag{3.58}$$

Hence we conclude from (3.35), (3.49), (3.56), (3.57) and (3.58) that

$$\begin{aligned}\big|P\big(W \le z\big) - \Phi(z)\big| &\le \frac{11.946}{\sqrt{n}} + \frac{1.037\sqrt{d+1}\delta_2^{\frac{1}{2}}}{n^{\frac{1}{4}}} + 8.314(d+1)^{\frac{1}{4}}\delta_4 + 11.765\delta_4 \\ &\quad + 5.014(d+1)\delta_2^{\frac{1}{2}}\delta_4^{\frac{1}{2}} + \frac{2\sqrt{2\pi}\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}} \\ &\le \frac{11.946}{\sqrt{n}} + \frac{1.037\sqrt{d+1}\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + 8.314(d+1)^{\frac{1}{4}}\delta_4 + 11.765\delta_4 \\ &\quad + 5.014(d+1)\frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}} + \frac{2\sqrt{2\pi}\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}\end{aligned}$$

where we have used the fact that

$$\delta_2^2 \le \frac{1}{n^{d+1}} \sum_{i_1=1}^{n} \dots \sum_{i_{d+1}=1}^{n} EY^4(i_1, ..., i_{d+1}) = \frac{\delta_4}{\sqrt{n}} \tag{3.59}$$

in the last inequality. □

# CHAPTER IV
# A NON-UNIFORM BOUND FOR THE GENERALIZATION OF A COMBINATORIAL CENTRAL LIMIT THEOREM

In this chapter, we use the same notation as in the previous chapter. In chapter 3, we give a uniform bound on the generalization of a combinotorial central limit theorem(Theorem 3.1). In this chapter, we give a non-uniform bound for the approximation of $W$ by $\Phi$.

**Theorem 4.1.** *Let $z \in \mathbb{R}$. With the notation and assumptions of Theorem 3.1, there exists a positive constant $C$ which does not depend on $z$ such that*

$$\big|P(W \le z) - \Phi(z)\big| \le \frac{C}{1+|z|}\left\{\frac{\delta_8^{\frac{1}{8}}}{n^{\frac{7}{16}}} + \frac{1}{n} + \delta_8\right\}$$

*where*

$$\delta_8 = \frac{1}{n^{d-\frac{5}{2}}}\sum_{i_1=1}^{n}\dots\sum_{i_{d+1}=1}^{n} E|Y(i_1,...,i_{d+1})|^8.$$

*Furthermore, if $\delta_8 \sim n^{-1/2}$,*

$$\big|P(W \le z) - \Phi(z)\big| \le \frac{C}{(1+|z|)\sqrt{n}}.$$

This chapter is organized as follows. Auxiliary results are in section 4.1 while the proof of main result is given in section 4.2.

From now on, $C$ stands for a positive constant with possibly different values in different places.

## 4.1 Auxiliary Results

In this section, we shall give auxiliary results for proving our main theorem(Theorem 4.1).

**Lemma 4.2.** *For each* $i_1, ..., i_{d+1} \in \{1, ..., n\}$ *and* $z > 0$, *let*

$$Y_z(i_1, ..., i_{d+1}) = Y(i_1, ..., i_{d+1})\mathbb{I}(|Y(i_1, ..., i_{d+1})| > 1 + z),$$

$$\text{and} \qquad \widehat{Y}_z(i_1, ..., i_{d+1}) = Y(i_1, ..., i_{d+1})\mathbb{I}(|Y(i_1, ..., i_{d+1})| \leq 1 + z).$$

*Then*

$$E[\sum_{i=1}^{n}\sum_{k=1}^{n}\widehat{Y}_z(i, \rho_1(k), ..., \rho_d(k))]^2 \leq Cn^{\frac{5}{4}}\delta_4^{\frac{1}{2}} + Cn^2\delta_4^2.$$

*Proof.* We observe from (3.3) that

$$\begin{aligned}
&\sum_{i,k}\sum_{\substack{l,m\\(l,m)\neq(i,k)}} EY(i, \rho_1(k), ..., \rho_d(k))Y(l, \rho_1(m), ..., \rho_d(m))\\
&= \sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{\substack{l\\l\neq i}}\sum_{\substack{m\\m\neq k}} EY(i, \rho_1(k), ..., \rho_d(k))Y(l, \rho_1(m), ..., \rho_d(m))\\
&\quad + \sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{\substack{l\\l\neq i}} EY(i, \rho_1(k), ..., \rho_d(k))Y(l, \rho_1(k), ..., \rho_d(k))\\
&\quad + \sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{\substack{m\\m\neq k}} EY(i, \rho_1(k), ..., \rho_d(k))Y(i, \rho_1(m), ..., \rho_d(m))\\
&= \frac{1}{(n(n-1))^{d-1}}\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n} EY(k_1, ..., k_{d+1})\sum_{\substack{m_1\\m_1\neq k_1}}...\sum_{\substack{m_{d+1}\\m_{d+1}\neq k_{d+1}}} EY(m_1, ..., m_{d+1})\\
&\quad + \frac{1}{n^{d-1}}\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n} EY(k_1, ..., k_{d+1})\sum_{\substack{l\\l\neq k_1}} EY(l, k_2, ..., k_{d+1})\\
&\quad + \frac{1}{(n(n-1))^{d-1}}\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n} EY(k_1, ..., k_{d+1})\sum_{\substack{m_2\\m_2\neq k_2}}...\sum_{\substack{m_{d+1}\\m_{d+1}\neq k_{d+1}}} EY(k_1, m_2, ..., m_{d+1})\\
&\leq \frac{C}{n^{d-1}}\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n}[EY(k_1, ..., k_{d+1})]^2\\
&\leq Cn\sqrt{n}\delta_2.
\end{aligned}$$

Thus

$$E[\sum_{i=1}^{n}\sum_{k=1}^{n}Y(i,\rho_1(k),...,\rho_d(k))]^2 \le \sum_{i=1}^{n}\sum_{k=1}^{n}EY^2(i,\rho_1(k),...,\rho_d(k)) + Cn\sqrt{n}\delta_2$$
$$= \frac{1}{n^{d-1}}\sum_{i_1=1}^{n}...\sum_{i_{d+1}=1}^{n}EY^2(i_1,...,i_{d+1}) + Cn\sqrt{n}\delta_2$$
$$= Cn\sqrt{n}\delta_2. \tag{4.1}$$

Note that

$$E|Y^m(i_1,...,i_{d+1})Y_z^n(i_1,...,i_{d+1})|$$
$$\le E|Y^m(i_1,...,i_{d+1})Y_z^n(i_1,...,i_{d+1})|(\frac{|Y_z(i_1,...,i_{d+1})|}{1+z})^r$$
$$\le \frac{E|Y(i_1,...,i_{d+1})|^{m+n+r}}{(1+z)^r} \tag{4.2}$$

for any integers $m, n$ and $r$ which $m \ge 0, n, r > 0$. Using the fact that $2ab \le a^2 + b^2$ and (4.2), we have

$$\sum_{i,k}\sum_{\substack{l,m\\(l,m)\neq(i,k)}} EY_z(i,\rho_1(k),...,\rho_d(k))Y_z(l,\rho_1(m),...,\rho_d(m))$$
$$= \frac{1}{(n(n-1))^{d-1}}\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n}EY_z(k_1,...,k_{d+1})\sum_{\substack{m_1\\m_1\neq k_1}}...\sum_{\substack{m_{d+1}\\m_{d+1}\neq k_{d+1}}}EY_z(m_1,...,m_{d+1})$$
$$+ \frac{1}{n^{d-1}}\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n}EY_z(k_1,...,k_{d+1})\sum_{\substack{l\\l\neq k_1}}EY_z(l,k_2,...,k_{d+1})$$
$$+ \frac{1}{(n(n-1))^{d-1}}\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n}EY_z(k_1,...,k_{d+1})\sum_{\substack{m_2\\m_2\neq k_2}}...\sum_{\substack{m_{d+1}\\m_{d+1}\neq k_{d+1}}}EY_z(k_1,m_2,...,m_{d+1})$$
$$\le \frac{C}{n^{2(d-1)}}[\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n}E|Y_z(k_1,...,k_{d+1})|]^2$$
$$+ \frac{1}{2n^{d-1}}\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n}\sum_{\substack{l\\l\neq k_1}}\{EY_z^2(k_1,...,k_{d+1}) + EY_z^2(l,k_2,...,k_{d+1})\}$$
$$\le \frac{C}{n^{2(d-1)}}[\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n}E|Y_z(k_1,...,k_{d+1})|]^2 + \frac{C}{n^{d-2}}\sum_{k_1=1}^{n}...\sum_{k_{d+1}=1}^{n}EY_z^2(k_1,...,k_{d+1})$$

$$\leq \frac{C}{n^{2(d-1)}}[\sum_{k_1=1}^{n}\cdots\sum_{k_{d+1}=1}^{n}\frac{EY^4(k_1,...,k_{d+1})}{(1+z)^3}]^2+\frac{C}{n^{d-2}}\sum_{k_1=1}^{n}\cdots\sum_{k_{d+1}=1}^{n}\frac{EY^4(k_1,...,k_{d+1})}{(1+z)^2}$$
$$= C\{n\delta_4^2+n\sqrt{n}\delta_4\}.$$

Thus

$$\begin{aligned}E[\sum_{i=1}^{n}\sum_{k=1}^{n}Y_z(i,\rho_1(k),...,\rho_d(k))]^2 &\leq \sum_{i=1}^{n}\sum_{k=1}^{n}EY_z^2(i,\rho_1(k),...,\rho_d(k))+Cn\delta_4^2+Cn\sqrt{n}\delta_4\\ &= \frac{1}{n^{d-1}}\sum_{i_1=1}^{n}\cdots\sum_{i_{d+1}=1}^{n}EY_z^2(i_1,...,i_{d+1})+Cn\delta_4^2\\ &\quad +Cn\sqrt{n}\delta_4\\ &\leq C\{n\sqrt{n}\delta_2+n\delta_4^2+n\sqrt{n}\delta_4\}. \qquad (4.3)\end{aligned}$$

Now, we conclude from (4.1) and (4.3) that

$$\begin{aligned}&E[\sum_{i=1}^{n}\sum_{k=1}^{n}\widehat{Y}_z(i,\rho_1(k),...,\rho_d(k))]^2\\ &= E[\sum_{i=1}^{n}\sum_{k=1}^{n}Y(i,\rho_1(k),...,\rho_d(k))-\sum_{i=1}^{n}\sum_{k=1}^{n}Y_z(i,\rho_1(k),...,\rho_d(k))]^2\\ &\leq 2E[\sum_{i=1}^{n}\sum_{k=1}^{n}Y(i,\rho_1(k),...,\rho_d(k))]^2+2E[\sum_{i=1}^{n}\sum_{k=1}^{n}Y_z(i,\rho_1(k),...,\rho_d(k))]^2\\ &\leq C\{n\sqrt{n}\delta_2+n\delta_4^2+n\sqrt{n}\delta_4\}\\ &\leq C\{n^{\frac{5}{4}}\delta_4^{\frac{1}{2}}+n\delta_4^2+n\sqrt{n}\delta_4\} \qquad (4.4)\end{aligned}$$

where we have used the fact that $\delta_2^2 \leq \dfrac{\delta_4}{\sqrt{n}}$(see eq.(3.59)) in the last inequality. We observe that

$$n\sqrt{n}\delta_4 = (n^{\frac{5}{4}}\delta_4^{\frac{1}{2}})(n^{\frac{1}{4}}\delta_4^{\frac{1}{2}}) \leq n^{\frac{5}{4}}\delta_4^{\frac{1}{2}},$$

in case of $n^{\frac{1}{4}}\delta_4^{\frac{1}{2}} \leq 1$, and

$$n\sqrt{n}\delta_4 \leq (n^{\frac{5}{4}}\delta_4^{\frac{1}{2}})(n^{\frac{1}{4}}\delta_4^{\frac{1}{2}})^3 = n^2\delta_4^2,$$

in case of $n^{\frac{1}{4}}\delta_4^{\frac{1}{2}} > 1$. Thus

$$n\sqrt{n}\delta_4 \leq n^{\frac{5}{4}}\delta_4^{\frac{1}{2}}+n^2\delta_4^2. \qquad (4.5)$$

By this fact and (4.4),

$$E[\sum_{i=1}^{n}\sum_{k=1}^{n}\widehat{Y}_z(i,\rho_1(k),...,\rho_d(k))]^2 \leq C\{n^{\frac{5}{4}}\delta_4^{\frac{1}{2}}+n^2\delta_4^2\}.$$

□

Let $I, K, \rho_1, ..., \rho_d$ be defined as in previous chapter. We define

$$\widehat{Y}(\rho) = \sum_{i=1}^{n} \widehat{Y}_z(i, \rho_1(i), ..., \rho_d(i)) \tag{4.6}$$

and

$$\widetilde{Y}(\rho) = \widehat{Y}(\rho) - \widehat{S}_{1,z} - \widehat{S}_{2,z} + \widehat{S}_{3,z} + \widehat{S}_{4,z} \tag{4.7}$$

where

$$\widehat{S}_{1,z} = \widehat{Y}_z(I, \rho_1(I), ..., \rho_d(I)),\ \widehat{S}_{2,z} = \widehat{Y}_z(K, \rho_1(K), ..., \rho_d(K)),$$
$$\widehat{S}_{3,z} = \widehat{Y}_z(I, \rho_1(K), .., \rho_d(K)),\ \widehat{S}_{4,z} = \widehat{Y}_z(K, \rho_1(I), ..., \rho_d(I)).$$

Similarly to (3.13), $\widehat{S}_{i,z}$ for $i = 1, 2, 3, 4$ are identically distributed.

**Lemma 4.3.** *Let $g$ be a continuous and piecewise continuously differential function. Then*

$$E\widehat{Y}(\rho)g(\widehat{Y}(\rho)) = E\int_{-\infty}^{\infty} g'(\widehat{Y}(\rho) + t)K(t)dt + \Delta g(\widehat{Y}(\rho)) \tag{4.8}$$

*and*

$$|\Delta g(\widehat{Y}(\rho))| \leq C\left\{\frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + \delta_4\right\}\{Eg^2(\widehat{Y}(\rho))\}^{\frac{1}{2}} \tag{4.9}$$

*where*

$$K(t) = \frac{n-1}{4}(\widetilde{Y}(\rho) - \widehat{Y}(\rho))(\mathbb{I}(0 \leq t \leq \widetilde{Y}(\rho) - \widehat{Y}(\rho)) - \mathbb{I}(\widetilde{Y}(\rho) - \widehat{Y}(\rho) \leq t < 0)).$$

*Proof.* Let $\mathcal{C}$ be the $\sigma$-algebra generated by

$$\{\widehat{Y}_z(i, \rho_1(i), ..., \rho_d(i)) : 1 \leq i \leq n\}.$$

Similarly to (3.16),

$$2E\{g(\widehat{Y}(\rho))E^{\mathcal{C}}(\widetilde{Y}(\rho) - \widehat{Y}(\rho))\} + E(\widetilde{Y}(\rho) - \widehat{Y}(\rho))[g(\widetilde{Y})(\rho) - g(\widehat{Y}(\rho))] = 0. \tag{4.10}$$

Note that

$$
\begin{aligned}
&E^{\mathcal{C}}[\widetilde{Y}(\rho) - \widehat{Y}(\rho)] \\
&= -E^{\mathcal{C}}\widehat{Y}_z(I, \rho_1(I), ..., \rho_d(I)) - E^{\mathcal{C}}\widehat{Y}_z(K, \rho_1(K), ..., \rho_d(K)) \\
&\quad + E^{\mathcal{C}}\widehat{Y}_z(I, \rho_1(K), ..., \rho_d(K)) + E^{\mathcal{C}}\widehat{Y}_z(K, \rho_1(I), ..., \rho_d(I)) \\
&= -\frac{2}{n}\sum_{i=1}^{n}\widehat{Y}_z(i, \rho_1(i), ..., \rho_d(i)) + \frac{2}{n(n-1)}E^{\mathcal{C}}\sum_{i=1}^{n}\sum_{\substack{k \\ k\neq i}}\widehat{Y}_z(i, \rho_1(k), ..., \rho_d(k)) \\
&= -\frac{2}{n}\widehat{Y}(\rho) + \frac{2}{n(n-1)}E^{\mathcal{C}}\sum_{i=1}^{n}\{\sum_{k=1}^{n}\widehat{Y}_z(i, \rho_1(k), ..., \rho_d(k)) \\
&\quad - \widehat{Y}_z(i, \rho_1(i), ..., \rho_d(i))\} \\
&= -\frac{2}{n-1}\widehat{Y}(\rho) + \frac{2}{n(n-1)}E^{\mathcal{C}}\sum_{i=1}^{n}\sum_{k=1}^{n}\widehat{Y}_z(i, \rho_1(k), ..., \rho_d(k)).
\end{aligned}
$$

From this fact and (4.10), we have

$$
\begin{aligned}
E\widehat{Y}(\rho)g(\widehat{Y}(\rho)) &= \frac{n-1}{4}E(\widetilde{Y}(\rho) - \widehat{Y}(\rho))[g(\widetilde{Y}(\rho)) - g(\widehat{Y}(\rho))] \\
&\quad + \frac{1}{n}Eg(\widehat{Y}(\rho))\sum_{i=1}^{n}\sum_{k=1}^{n}\widehat{Y}_z(i, \rho_1(k), ..., \rho_d(k)) \\
&= \frac{n-1}{4}E(\widetilde{Y}(\rho) - \widehat{Y}(\rho))[g(\widetilde{Y}(\rho)) - g(\widehat{Y}(\rho))] + \Delta g(\widehat{Y}(\rho)) \\
&= \frac{n-1}{4}E(\widetilde{Y}(\rho) - \widehat{Y}(\rho))\int_0^{\widetilde{Y}(\rho)-\widehat{Y}(\rho)} g'(\widehat{Y}(\rho)+t)dt + \Delta g(\widehat{Y}(\rho)) \\
&= E\int_{-\infty}^{\infty} g'(\widehat{Y}(\rho)+t)K(t)dt + \Delta g(\widehat{Y}(\rho))
\end{aligned}
$$

where

$$\Delta g(\widehat{Y}(\rho)) = \frac{1}{n}Eg(\widehat{Y}(\rho))\sum_{i=1}^{n}\sum_{k=1}^{n}\widehat{Y}_z(i, \rho_1(k), ..., \rho_d(k)).$$

By Lemma 4.2,

$$
\begin{aligned}
|\Delta g(\widehat{Y}(\rho))| &\leq \frac{1}{n}\{Eg^2(\widehat{Y}(\rho))\}^{\frac{1}{2}}\{E[\sum_{i=1}^{n}\sum_{k=1}^{n}\widehat{Y}_z(i, \rho_1(k), ..., \rho_d(k))]^2\}^{\frac{1}{2}} \\
&\leq C\left\{\frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + \delta_4\right\}\{Eg^2(\widehat{Y}(\rho))\}^{\frac{1}{2}}.
\end{aligned}
$$

□

**Lemma 4.4.**

$$E\widehat{Y}^4(\rho) \leq C\{\sqrt{n}\delta_4 + n\delta_4^2\}.$$

*Proof.* We observe that

$$E(\sum_{i=1}^{n} Y_z(i,\rho_1(i),...,\rho_d(i)))^4 = M_1 + M_2 + M_3 + M_4 + M_5$$

where

$$M_1 = \sum_{i=1}^{n} EY_z^4(i,\rho_1(i),...,\rho_d(i)),$$

$$M_2 = \sum_{i=1}^{n} \sum_{\substack{j \\ j\neq i}} EY_z(i,\rho_1(i),...,\rho_d(i))Y_z^3(j,\rho_1(j),...,\rho_d(j)),$$

$$M_3 = \sum_{i=1}^{n} \sum_{\substack{j \\ j\neq i}} EY_z^2(i,\rho_1(i),...,\rho_d(i))Y_z^2(j,\rho_1(j),...,\rho_d(j)),$$

$$M_4 = \sum_{i=1}^{n} \sum_{\substack{j \\ j\neq i}} \sum_{\substack{k \\ k\neq i,j}} EY_z^2(i,\rho_1(i),...,\rho_d(i))Y_z(j,\rho_1(j),...,\rho_d(j))$$

$$Y_z(k,\rho_1(k),...,\rho_d(k)) \quad \text{and}$$

$$M_5 = \sum_{i=1}^{n} \sum_{\substack{j \\ j\neq i}} \sum_{\substack{k \\ k\neq i,j}} \sum_{\substack{l \\ l\neq i,j,k}} EY_z(i,\rho_1(i),...,\rho_d(i))Y_z(j,\rho_1(j),...,\rho_d(j))$$

$$Y_z(k,\rho_1(k),...,\rho_d(k))Y_z(l,\rho_1(l),...,\rho_d(l)).$$

We note that

$$|M_1| = \frac{1}{n^d} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} EY_z^4(i_1,...,i_{d+1}) = \frac{\delta_4}{\sqrt{n}}.$$

It follows from (4.2) that

$$|M_2| \leq \frac{C}{n^{2d}} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} E|Y_z(i_1,...,i_{d+1})| \sum_{j_1=1}^{n} ... \sum_{j_{d+1}=1}^{n} E|Y_z(j_1,...,j_{d+1})|^3$$

$$\leq \frac{C}{n^{2d}} [\sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} EY^4(i_1,...,i_{d+1})]^2$$

$$= \frac{C\delta_4^2}{n},$$

$$|M_3| \le \frac{C}{n^{2d}} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} EY_z^2(i_1, ..., i_{d+1}) \sum_{j_1=1}^{n} ... \sum_{j_{d+1}=1}^{n} EY_z^2(j_1, ..., j_{d+1})$$
$$\le \frac{C}{n^{2d}} [\sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} EY^4(i_1, ..., i_{d+1})]^2$$
$$= \frac{C\delta_4^2}{n},$$
$$|M_4| \le \frac{C}{n^{3d}} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} EY_z^2(i_1, ..., i_{d+1}) [\sum_{j_1=1}^{n} ... \sum_{j_{d+1}=1}^{n} E|Y_z(j_1, ..., j_{d+1})|]^2$$
$$\le \frac{C}{n^{3d}} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} EY^4(i_1, ..., i_{d+1}) [\sum_{j_1=1}^{n} ... \sum_{j_{d+1}=1}^{n} E|Y(j_1, ..., j_{d+1})|^2]^2$$
$$= C\sqrt{n}\delta_4\delta_2^2 \quad \text{and}$$
$$|M_5| \le \frac{C}{n^{4d}} [\sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} E|Y_z(i_1, ..., i_{d+1})|]^4$$
$$\le \frac{C}{n^{4d}} [\sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} EY^2(i_1, ..., i_{d+1})]^4$$
$$= Cn^2\delta_2^4.$$

Thus

$$E(\sum_{i=1}^{n} Y_z(i, \rho_1(i), ..., \rho_d(i)))^4 = \frac{\delta_4}{\sqrt{n}} + \frac{C\delta_4^2}{n} + C\sqrt{n}\delta_4\delta_2^2 + Cn^2\delta_2^4$$
$$\le C\left\{\frac{\delta_4}{\sqrt{n}} + \frac{\delta_4^2}{n} + \delta_4^2 + n\delta_4^2\right\}$$
$$\le C\left\{\frac{\delta_4}{\sqrt{n}} + n\delta_4^2\right\}$$

where we have used the fact that $\delta_2^2 \le \dfrac{\delta_4}{\sqrt{n}}$ in the first inequality. By this fact and Lemma 3.5,

$$E\widehat{Y}^4(\rho) = E(\sum_{i=1}^{n} Y(i, \rho_1(i), ..., \rho_d(i)) - \sum_{i=1}^{n} Y_z(i, \rho_1(i), ..., \rho_d(i)))^4$$
$$\le CEW^4 + CE(\sum_{i=1}^{n} Y_z(i, \rho_1(i), ..., \rho_d(i)))^4$$
$$\le C\{\sqrt{n}\delta_4 + n\delta_4^2\}.$$

□

**Lemma 4.5.**

$$E|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^2 = \frac{4}{n} + \Delta$$

*where* $|\Delta| \le C\left\{\frac{\delta_4^{\frac{1}{2}}}{n^{\frac{7}{4}}} + \frac{\delta_4}{n\sqrt{n}} + \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{9}{8}}}\right\}$.

*Proof.* Let

$$S_{1,z} = Y_z(I, \rho_1(I), ..., \rho_d(I)), \quad S_{2,z} = Y_z(K, \rho_1(K), ..., \rho_d(K)),$$
$$S_{3,z} = Y_z(I, \rho_1(K), ..., \rho_d(K)), \; S_{4,z} = Y_z(K, \rho_1(I), ..., \rho_d(I)).$$

By the same argument as in (3.13), $S_{1,z}, S_{2,z}, S_{3,z}, S_{4,z}$ have the same distribution. We observe that

$$\begin{aligned} E|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^2 &= E(-\hat{S}_{1,z} - \hat{S}_{2,z} + \hat{S}_{3,z} + \hat{S}_{4,z})^2 \\ &= E(-(S_1 - S_{1,z}) - (S_2 - S_{2,z}) + (S_3 - S_{3,z}) + (S_4 - S_{4,z}))^2 \\ &= \sum_{k=1}^{4} ES_k^2 + \Delta_1 \end{aligned} \tag{4.11}$$

where

$$|\Delta_1| \le \sum_{1\le i<j\le 4} |ES_iS_j| + 4\sum_{i=1}^{4} ES_{i,z}^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} E|S_iS_{j,z}|$$

and from (3.2) and (3.3),

$$\begin{aligned} ES_1^2 &= \frac{1}{n}\sum_{i=1}^{n} EY^2(i, \rho_1(i), ..., \rho_d(i)) \\ &= \frac{1}{n}[EW^2 - \sum_{i=1}^{n}\sum_{\substack{j \\ j\ne i}} EY(i, \rho_1(i), ..., \rho_d(i))Y(j, \rho_1(j), ..., \rho_d(j))] \\ &= \frac{1}{n} - \frac{1}{n^{d+1}(n-1)^d}\sum_{i_1=1}^{n}\cdots\sum_{i_{d+1}=1}^{n}\sum_{\substack{j_1 \\ j_1\ne i_1}}\cdots\sum_{\substack{j_{d+1} \\ j_{d+1}\ne i_{d+1}}} EY(i_1, ..., i_{d+1})EY(j_1, ..., j_{d+1}) \\ &= \frac{1}{n} - \frac{(-1)^{d+1}}{n^{d+1}(n-1)^d}\sum_{i_1=1}^{n}\cdots\sum_{i_{d+1}=1}^{n}[EY(i_1, ..., i_{d+1})]^2. \end{aligned} \tag{4.12}$$

By (4.11), (4.12) and the fact that $S_1, S_2, S_3, S_4$ have the same distribution,

$$E|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^2 = \frac{4}{n} + \Delta$$

where

$$|\Delta| \le \frac{C\delta_2}{n\sqrt{n}} + \sum_{1\le i<j\le 4} |ES_iS_j| + 4\sum_{i=1}^{4} ES_{i,z}^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} E|S_iS_{j,z}|. \tag{4.13}$$

To prove the lemma, it suffices to find appropriate bounds for the terms on the right-hand side of (4.13). Note from (3.3) that

$$\begin{aligned}
|ES_1S_2| &= |EY(I,\rho_1(I),...,\rho_d(I))Y(K,\rho_1(K),...,\rho_d(K))| \\
&= |\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{k\\ k\neq i}} EY(i,\rho_1(i),...,\rho_d(i))Y(k,\rho_1(k),...,\rho_d(k))| \\
&= |\frac{1}{(n(n-1))^{d+1}}\sum_{i_1=1}^{n}\cdots\sum_{i_{d+1}=1}^{n} EY(i_1,...,i_{d+1}) \sum_{\substack{k_1\\ k_1\neq i_1}}\cdots\sum_{\substack{k_{d+1}\\ k_{d+1}\neq i_{d+1}}} EY(k_1,...,k_{d+1})| \\
&= |\frac{1}{(n(n-1))^{d+1}}\sum_{i_1=1}^{n}\cdots\sum_{i_{d+1}=1}^{n} (EY(i_1,...,i_{d+1}))^2| \\
&\le \frac{C\delta_2}{n^2\sqrt{n}}.
\end{aligned}$$

By the same argument, we can show that

$$|ES_iS_j| \le \frac{C\delta_2}{n^2\sqrt{n}} \tag{4.14}$$

for $1 \le i < j \le 4$. We note from (4.2) that

$$\begin{aligned}
ES_{1,z}^2 &= \frac{1}{n}\sum_{i=1}^{n} EY_z^2(i,\rho_1(i),...,\rho_d(i)) \\
&= \frac{1}{n^{d+1}}\sum_{i_1=1}^{n}\cdots\sum_{i_{d+1}=1}^{n} EY_z^2(i_1,...,i_{d+1}) \\
&\le \frac{C}{n^{d+1}}\sum_{i_1=1}^{n}\cdots\sum_{i_{d+1}=1}^{n} EY^4(i_1,...,i_{d+1}) \\
&= \frac{C\delta_4}{n\sqrt{n}}.
\end{aligned} \tag{4.15}$$

By (3.45) and (4.15),

$$E|S_iS_{j,z}| \le \{ES_i^2\}^{\frac{1}{2}}\{ES_{j,z}^2\}^{\frac{1}{2}} \le \frac{C\delta_2^{\frac{1}{2}}\delta_4^{\frac{1}{2}}}{n} \tag{4.16}$$

for $i, j = 1, 2, 3, 4$. Now we conclude from (4.13)-(4.16) that

$$|\Delta| \leq C\left\{\frac{\delta_2}{n\sqrt{n}} + \frac{\delta_4}{n\sqrt{n}} + \frac{\delta_2^{\frac{1}{2}}\delta_4^{\frac{1}{2}}}{n}\right\}$$
$$\leq C\left\{\frac{\delta_4^{\frac{1}{2}}}{n^{\frac{7}{4}}} + \frac{\delta_4}{n\sqrt{n}} + \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{9}{8}}}\right\}$$

where we have used the fact that $\delta_2^2 \leq \dfrac{\delta_4}{\sqrt{n}}$ in the last inequality. □

We define

$$\widehat{Y}(\tau) = \sum_{i=1}^{n} \widehat{Y}_z(i, \tau_1(i), ..., \tau_d(i)).$$

**Lemma 4.6.** *Let $z > 0$ and $g_z$ be defined as in (2.4). Then*

$$|Eg_z'(\widehat{Y}(\tau))E\int_{-\infty}^{\infty} K(t)dt - Eg_z'(\widehat{Y}(\tau))\int_{-\infty}^{\infty} K(t)dt| \leq \frac{C}{1+z}\left\{\frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + \delta_4\right\}.$$

*Proof.* Let

$$\widehat{G} = \widehat{Y}_z(I, M_1, ..., M_d) + \widehat{Y}_z(K, L_1, ..., L_d) - \widehat{Y}_z(I, L_1, ..., L_d) - \widehat{Y}_z(K, M_1, ..., M_d)$$

and $A$ be defined as in Lemma 3.6.i.e.,

$$A = \{\tau_i(I) \neq L_i, \tau_i(K) \neq M_i, \tau_i(I) \neq M_i, \tau_i(K) \neq L_i, \ : \ i = 1, 2, ..., d\}.$$

By the same argument as (3.51),

$$|Eg_z'(\widehat{Y}(\tau))E\int_{-\infty}^{\infty} K(t)dt - E[g_z'(\widehat{Y}(\tau))\int_{-\infty}^{\infty} K(t)dt]| \leq M_1 + M_2 \tag{4.17}$$

where

$$M_1 = \frac{n-1}{4}|Eg_z'(\widehat{Y}(\tau))\mathbb{I}(\widehat{Y}(\tau) \leq \frac{z}{2})[E^{\tau_1,...,\tau_d}\widehat{G}^2\mathbb{I}(A^c) - \widehat{G}^2\mathbb{I}(A^c)]|$$

and $$M_2 = \frac{n-1}{4}|Eg_z'(\widehat{Y}(\tau))\mathbb{I}(\widehat{Y}(\tau) > \frac{z}{2})[E^{\tau_1,...,\tau_d}\widehat{G}^2\mathbb{I}(A^c) - \widehat{G}^2\mathbb{I}(A^c)]|.$$

To bound $M_1$, we observe that for $w \leq \frac{z}{2}$,

$$|g_z'(w)| = |[1-\Phi(z)][1+\sqrt{2\pi}we^{\frac{1}{2}w^2}\Phi(w)]|$$
$$\leq \{\frac{e^{-\frac{z^2}{2}}}{z}\}[1+\sqrt{2\pi}(\frac{z}{2})e^{\frac{z^2}{8}}\Phi(\frac{z}{2})]$$
$$\leq \frac{e^{-\frac{z^2}{2}}}{z} + \frac{\sqrt{2\pi}e^{-\frac{3z^2}{8}}}{2}$$
$$\leq \frac{C}{1+z} \tag{4.18}$$

where we have used the fact that $1-\Phi(z) \le \dfrac{e^{-\frac{z^2}{2}}}{z}$ for $z \ge 0$(see[1], pp.11) in the first inequality. Let $\mathcal{B}$ be defined as in chapter 3. By (3.52) and the fact that

$$E|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^k \le C\sum_{i=1}^{4} E|\widehat{S}_{i,z}|^k \le \frac{C\delta_k}{n^{\frac{k-1}{2}}} \tag{4.19}$$

for $k = 2, 4, 8$, we have

$$\begin{aligned} E[\widehat{G}^2\mathbb{I}(A^c)] &= E[\widehat{G}^2 E^{\mathcal{B}}\mathbb{I}(A^c)] \\ &\le \frac{CE\widehat{G}^2}{n} \\ &= \frac{CE|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^2}{n} \\ &\le \frac{C\delta_2}{n\sqrt{n}}. \end{aligned} \tag{4.20}$$

Then (4.18) and (4.20) yield

$$M_1 \le \frac{Cn}{1+z}E|E^{\tau_1,\dots\tau_d}\widehat{G}^2\mathbb{I}(A^c) - \widehat{G}^2\mathbb{I}(A^c)| \le \frac{Cn}{1+z}E[\widehat{G}^2\mathbb{I}(A^c)] \le \frac{C\delta_2}{(1+z)\sqrt{n}}. \tag{4.21}$$

To bound $M_2$, we note that

$$\begin{aligned} E\widehat{Y}^2(\tau) &= E[\sum_{i=1}^{n}\widehat{Y}_z(i,\tau_1(i),...,\tau_d(i))]^2 \\ &= E[\sum_{i=1}^{n} Y(i,\tau_1(i),...,\tau_d(i)) - \sum_{i=1}^{n} Y_z(i,\tau_1(i),...,\tau_d(i))]^2 \\ &\le 2E[\sum_{i=1}^{n} Y(i,\tau_1(i),...,\tau_d(i))]^2 + 2E[\sum_{i=1}^{n} Y_z(i,\tau_1(i),...,\tau_d(i))]^2 \\ &\le 2EW^2 + 2n\sum_{i=1}^{n} EY_z^2(i,\tau_1(i),...,\tau_d(i)) \\ &= C + \frac{C}{n^{d-1}}\sum_{i_1=1}^{n}\cdots\sum_{i=1_{d+1}}^{n} EY_z^2(i_1,...,i_{d+1}) \\ &\le C + \frac{C}{n^{d-1}}\sum_{i_1=1}^{n}\cdots\sum_{i=1_{d+1}}^{n} \frac{EY^4(i_1,...,i_{d+1})}{(1+z)^2} \\ &= C\{1+\sqrt{n}\delta_4\}. \end{aligned} \tag{4.22}$$

From this fact, (3.55) and Chebyshev's inequality, we have

$$\begin{aligned}M_2 &\le CnE\mathbb{I}(\widehat{Y}(\tau) > \frac{z}{2})|E^{\tau_1,\cdots\tau_d}\widehat{G}^2\mathbb{I}(A^c) - \widehat{G}^2\mathbb{I}(A^c)| \\ &\le Cn\{\frac{4E\widehat{Y}^2(\tau)}{z^2}\}^{\frac{1}{2}}\{E[E^{\tau_1,\cdots\tau_d}\widehat{G}^2\mathbb{I}(A^c) - \widehat{G}^2\mathbb{I}(A^c)]^2\}^{\frac{1}{2}} \\ &\le \frac{Cn}{1+z}\{C + C\sqrt{n}\delta_4\}^{\frac{1}{2}}\{E[E^{\tau_1,\cdots\tau_d}\widehat{G}^2\mathbb{I}(A^c) - \widehat{G}^2\mathbb{I}(A^c)]^2\}^{\frac{1}{2}}.\end{aligned}$$

To finish the proof, we need to bound $E[E^{\tau_1,\cdots\tau_d}\widehat{G}^2\mathbb{I}(A^c) - \widehat{G}^2\mathbb{I}(A^c)]^2$. We observe from (3.52) and (4.19) that

$$\begin{aligned}E[E^{\tau_1,\cdots\tau_d}\widehat{G}^2\mathbb{I}(A^c) - \widehat{G}^2\mathbb{I}(A^c)]^2 &\le 2E[E^{\tau_1,\cdots\tau_d}\widehat{G}^2\mathbb{I}(A^c)]^2 + 2E[\widehat{G}^2\mathbb{I}(A^c)]^2 \\ &= 2EE^{\tau_1,\cdots\tau_d}[\widehat{G}^2\mathbb{I}(A^c)E^{\tau_1,\cdots\tau_d}\widehat{G}^2\mathbb{I}(A^c)] + 2E\widehat{G}^4\mathbb{I}(A^c) \\ &\le 2E\widehat{G}^2\mathbb{I}(A^c)E^{\tau_1,\cdots\tau_d}\widehat{G}^2 + 2E\widehat{G}^4\mathbb{I}(A^c) \\ &= 2E\widehat{G}^2\mathbb{I}(A^c)E\widehat{G}^2 + 2E\widehat{G}^4\mathbb{I}(A^c) \\ &= 2E\widehat{G}^2E^{\mathcal{B}}\mathbb{I}(A^c)E\widehat{G}^2 + 2E\widehat{G}^4E^{\mathcal{B}}\mathbb{I}(A^c) \\ &\le \frac{CE\widehat{G}^4}{n} \\ &= \frac{CE|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^4}{n} \\ &\le \frac{C\delta_4}{n^2\sqrt{n}}.\end{aligned}$$

So,

$$M_2 \le \frac{C}{1+z}\left\{\frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + \delta_4\right\}. \tag{4.23}$$

Now, we conclude from (4.17), (4.21) and (4.23) that

$$\begin{aligned}|Eg_z'(\widehat{Y}(\tau))E\int_{-\infty}^{\infty}K(t)dt - Eg_z'(\widehat{Y}(\tau))\int_{-\infty}^{\infty}K(t)dt| &\le \frac{C}{1+z}\left\{\frac{\delta_2}{\sqrt{n}} + \frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + \delta_4\right\} \\ &\le \frac{C}{1+z}\left\{\frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + \delta_4\right\}\end{aligned}$$

where we have used the fact that $\delta_2^2 \le \frac{\delta_4}{\sqrt{n}}$ in the last inequality. □

**Lemma 4.7.** *Let $z > 0$ and let*

$$h(w) = (wg_z(w))'.$$

*Then*

$$E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| < \frac{z}{4}) \int_{-\infty}^{\infty}\int_{t}^{\Delta\widehat{Y}} h(\widehat{Y}(\rho)+u)K(t)dudt$$
$$\leq \frac{C}{1+z}\left\{\frac{\delta_8^{\frac{3}{8}}}{n^{\frac{5}{16}}} + \frac{\delta_8^{\frac{7}{8}}}{n^{\frac{1}{16}}}\right\}$$

*where* $\Delta\widehat{Y} = \widehat{Y}(\tau) - \widehat{Y}(\rho)$.

*Proof.* Since $E|\Delta\widehat{Y}|^k$ is a finite sum of terms each of the form $E|\widehat{Y}_z(I, \rho_1(I), ..., \rho_d(I))|^k$,

$$E|\Delta\widehat{Y}|^k \leq \frac{C\delta_k}{n^{\frac{k-1}{2}}} \tag{4.24}$$

for $k = 2, 4$. This implies

$$\begin{aligned}
&E(\int_{-\infty}^{\infty}\int_{t}^{\Delta\widehat{Y}} K(t)dudt)^k \\
&\leq E(\int_{-\infty}^{\infty}(|\Delta\widehat{Y}| + |t|)K(t)dt)^k \\
&= (\frac{n-1}{4})^k E(|\Delta\widehat{Y}||\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^2 + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^3)^k \\
&\leq Cn^k\{E|\Delta\widehat{Y}|^k|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^{2k} + E|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^{3k}\} \\
&\leq Cn^k\{E|\Delta\widehat{Y}|^{2k}\}^{\frac{1}{2}}\{E|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^{4k}\}^{\frac{1}{2}} + Cn^k\{E|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^{4k}\}^{\frac{3}{4}} \\
&\leq C\left\{\frac{\delta_{2k}^{\frac{1}{2}}\delta_{4k}^{\frac{1}{2}}}{n^{\frac{k-1}{2}}} + \frac{\delta_{4k}^{\frac{3}{4}}}{n^{\frac{4k-3}{8}}}\right\}
\end{aligned} \tag{4.25}$$

for $k = 1, 2$. By (4.25) and the fact that

$$h(w) \leq \begin{cases} C(1+z) & \text{if } \frac{z}{2} < w \leq z, \\ \dfrac{C}{(1+z)^2} & \text{if } w \leq \frac{z}{2} \text{ or } w > z \end{cases}$$

(see[12], pp.2357), we have

$$E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| < \frac{z}{4}) \int_{-\infty}^{\infty}\int_{t}^{\Delta\widehat{Y}} h(\widehat{Y}(\rho)+u)K(t)du\,dt$$

$$\leq E\int_{-\infty}^{\infty}\int_{t}^{\Delta\widehat{Y}} h(\widehat{Y}(\rho)+u)K(t)\mathbb{I}(\widehat{Y}(\rho)+u \leq \frac{z}{2} \text{ or } \widehat{Y}(\rho)+u > z)du\,dt$$

$$+ E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| < \frac{z}{4}) \int_{-\infty}^{\infty}\int_{t}^{\Delta\widehat{Y}} h(\widehat{Y}(\rho)+u)K(t)\mathbb{I}(\frac{z}{2} < \widehat{Y}(\rho)+u \leq z)du\,dt$$

$$\leq \frac{C}{(1+z)^2}\left\{\delta_2^{\frac{1}{2}}\delta_4^{\frac{1}{2}} + \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}\right\}$$

$$+ C(1+z)E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| < \frac{z}{4})\int_{-\infty}^{\infty}\int_t^{\Delta\widehat{Y}} K(t)\mathbb{I}(\widehat{Y}(\rho) + u > \frac{z}{2})du\, dt. \tag{4.26}$$

It remains to bound the second term on the right hand side of (4.1). By Lemma 4.4, (4.25) and the fact that $K(t) = 0$ for $|t| > |\widehat{Y}(\rho) - \widetilde{Y}(\rho)|$,

$$E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| < \frac{z}{4})\int_{-\infty}^{\infty}\int_t^{\Delta\widehat{Y}} K(t)\mathbb{I}(\widehat{Y}(\rho) + u > \frac{z}{2})du\, dt$$

$$\leq E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| < \frac{z}{4})\int_{-\infty}^{\infty}\int_t^{\Delta\widehat{Y}} K(t)\mathbb{I}(\widehat{Y}(\rho) + |\Delta\widehat{Y}|$$

$$+ |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| > \frac{z}{2})du\, dt$$

$$= E\int_{-\infty}^{\infty}\int_t^{\Delta\widehat{Y}} K(t)\mathbb{I}(\widehat{Y}(\rho) > \frac{z}{4})du\, dt$$

$$\leq \{P(\widehat{Y}(\rho) > \frac{z}{4})\}^{\frac{1}{2}}\{E(\int_{-\infty}^{\infty}\int_t^{\Delta\widehat{Y}} K(t)du\, dt)^2\}^{\frac{1}{2}}$$

$$\leq C\{\frac{E\widehat{Y}^4(\rho)}{z^4}\}^{\frac{1}{2}}\left\{\frac{C\delta_4^{\frac{1}{2}}\delta_8^{\frac{1}{2}}}{\sqrt{n}} + \frac{C\delta_8^{\frac{3}{4}}}{n^{\frac{5}{8}}}\right\}^{\frac{1}{2}}$$

$$\leq \frac{C}{z^2}\left\{\delta_4^{\frac{3}{4}}\delta_8^{\frac{1}{4}} + \frac{\delta_4^{\frac{1}{2}}\delta_8^{\frac{3}{8}}}{n^{\frac{1}{16}}} + n^{\frac{1}{4}}\delta_4^{\frac{5}{4}}\delta_8^{\frac{1}{4}} + n^{\frac{3}{16}}\delta_4\delta_8^{\frac{3}{8}}\right\}. \tag{4.27}$$

By (4.1) and (4.27),

$$E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| < \frac{z}{4})\int_{-\infty}^{\infty}\int_t^{\Delta\widehat{Y}} h(\widehat{Y}(\rho) + u)K(t)du\, dt$$

$$\leq \frac{C}{(1+z)^2}\left\{\delta_2^{\frac{1}{2}}\delta_4^{\frac{1}{2}} + \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}\right\} + \frac{C}{1+z}\left\{\delta_4^{\frac{3}{4}}\delta_8^{\frac{1}{4}} + \frac{\delta_4^{\frac{1}{2}}\delta_8^{\frac{3}{8}}}{n^{\frac{1}{16}}} + n^{\frac{1}{4}}\delta_4^{\frac{5}{4}}\delta_8^{\frac{1}{4}} + n^{\frac{3}{16}}\delta_4\delta_8^{\frac{3}{8}}\right\}$$

$$\leq \frac{C}{1+z}\left\{\frac{\delta_8^{\frac{3}{8}}}{n^{\frac{5}{16}}} + \frac{\delta_8^{\frac{5}{8}}}{n^{\frac{3}{16}}} + \frac{\delta_8^{\frac{7}{8}}}{n^{\frac{1}{16}}}\right\}$$

where we have used the fact that $\delta_2^2 \leq \frac{\delta_4}{\sqrt{n}}$ and $\delta_4^2 \leq \frac{\delta_8}{\sqrt{n}}$ in the last inequality. By the same argument as (4.5), we can show that $\frac{\delta_8^{\frac{5}{8}}}{n^{\frac{3}{16}}} \leq \frac{\delta_8^{\frac{3}{8}}}{n^{\frac{5}{16}}} + \frac{\delta_8^{\frac{7}{8}}}{n^{\frac{1}{16}}}$. Thus, we have the lemma. □

## 4.2 Proof of Theorem 4.1

*Proof.* To bound $|P(W \leq z) - \Phi(z)|$, it suffices to consider $z > 0$ as we have used the fact that $\Phi(z) = 1 - \Phi(-z)$ and apply the result to $-W$ when $z \leq 0$. So, from now on, we assume $z > 0$.

Let

$$\widehat{Y} = \sum_{i=1}^{n} \widehat{Y}_z(i, \pi_1(i), ..., \pi_d(i)).$$

Note that

$$|P(W \leq z) - \Phi(z)| \leq P(W \neq \widehat{Y}) + |P(\widehat{Y} \leq z) - \Phi(z)|. \tag{4.28}$$

We observe that

$$\begin{aligned}
P(W \neq \widehat{Y}) &= P(\sum_{i=1}^{n} \mathbb{I}(|Y(i, \pi_1(i), ..., \pi_d(i))| > 1 + z) \geq 1) \\
&\leq \sum_{i=1}^{n} E\mathbb{I}(|Y(i, \pi_1(i), ..., \pi_d(i))| > 1 + z) \\
&= \frac{1}{n^d} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} P(|Y(i_1, ..., i_{d+1})| > 1 + z) \\
&\leq \frac{1}{n^d} \sum_{i_1=1}^{n} ... \sum_{i_{d+1}=1}^{n} \frac{E|Y(i_1, ..., i_{d+1})|^4}{(1+z)^4} \\
&= \frac{\delta_4}{(1+z)^4\sqrt{n}}.
\end{aligned} \tag{4.29}$$

Next, we will bound the second term on the right hand side of (4.28). By the same argument as previous chapter, we can show that

$$\begin{aligned}
|P(\widehat{Y} \leq z) - \Phi(z)| &\leq |Eg_z'(\widehat{Y}(\tau)) \int_{-\infty}^{\infty} K(t)dt - E\int_{-\infty}^{\infty} g_z'(\widehat{Y}(\rho) + t)K(t)dt| \\
&\quad + |Eg_z'(\widehat{Y}(\tau))E\int_{-\infty}^{\infty} K(t)dt - Eg_z'(\widehat{Y}(\tau)) \int_{-\infty}^{\infty} K(t)dt| \\
&\quad + |Eg_z'(\widehat{Y}(\tau)) - Eg_z'(\widehat{Y}(\tau))E\int_{-\infty}^{\infty} K(t)dt| \\
&\quad + |\Delta g_z(\widehat{Y}(\rho))| \\
&= |T_1| + |T_2| + |T_3| + |T_4|
\end{aligned} \tag{4.30}$$

where

$$T_1 = Eg_z'(\widehat{Y}(\tau))\int_{-\infty}^{\infty} K(t)dt - E\int_{-\infty}^{\infty} g_z'(\widehat{Y}(\rho)+t)K(t)dt$$
$$T_2 = Eg_z'(\widehat{Y}(\tau))E\int_{-\infty}^{\infty} K(t)dt - Eg_z'(\widehat{Y}(\tau))\int_{-\infty}^{\infty} K(t)dt$$
$$T_3 = Eg_z'(\widehat{Y}(\tau)) - Eg_z'(\widehat{Y}(\tau))E\int_{-\infty}^{\infty} K(t)dt$$
$$T_4 = \Delta g_z(\widehat{Y}(\rho)).$$

By the fact that $0 \le g_z(w) \le \min(\frac{\sqrt{2\pi}}{4}, \frac{1}{|z|})$(see[6],pp.246) and (4.9),

$$|T_4| \le \left\{\frac{C\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + C\delta_4\right\}\{Eg_z^2(\widehat{Y}(\rho))\}^{\frac{1}{2}} \le \frac{C}{1+z}\left\{\frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + \delta_4\right\}. \tag{4.31}$$

Next, we will bound $T_3$. By the same argument as Chen and Shao([6],pp.248), we can show that $E|g_z'(\widehat{Y}(\tau))| \le \frac{C}{(1+z)^2}(1+\sqrt{n}\delta_4)$, for $z > 0$. From this fact and Lemma 4.5, we have

$$\begin{aligned}
|T_3| &\le E|g_z'(\widehat{Y}(\tau))||1 - E\int_{-\infty}^{\infty} K(t)dt| \\
&\le \frac{C}{(1+z)^2}(1+\sqrt{n}\delta_4)|1 - \frac{n-1}{4}E|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^2| \\
&\le \frac{C}{(1+z)^2}\left\{\frac{1}{n} + \frac{C\delta_4^{\frac{1}{2}}}{n^{\frac{3}{4}}} + \frac{C\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}} + \frac{C\delta_4}{\sqrt{n}} + \frac{C\delta_4^{\frac{3}{2}}}{n^{\frac{1}{4}}} + C\delta_4^{\frac{7}{4}}n^{\frac{3}{8}} + \delta_4^2\right\} \\
&\le \frac{C}{(1+z)^2}\left\{\frac{1}{n} + \frac{\delta_4^{\frac{1}{2}}}{n^{\frac{3}{4}}} + \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}} + \delta_4^2\right\}
\end{aligned} \tag{4.32}$$

where we have used the fact that $\frac{\delta_4}{\sqrt{n}}, \frac{\delta_4^{\frac{3}{2}}}{n^{\frac{1}{4}}}, \delta_4^{\frac{7}{4}}n^{\frac{3}{8}} \le \frac{\delta_4^{\frac{1}{2}}}{n^{\frac{3}{4}}} + \delta_4^2$ in the last inequality. By Lemma 4.6, (4.31) and (4.32),

$$\begin{aligned}
|T_2| + |T_3| + |T_4| &\le \frac{C}{1+z}\left\{\frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}} + \delta_4^2 + \frac{1}{n} + \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}} + \frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + \delta_4\right\} \\
&\le \frac{C}{1+z}\left\{\frac{1}{n} + \frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + \sqrt{n}\delta_4^2\right\}
\end{aligned} \tag{4.33}$$

where we have used the fact that $\frac{\delta_4^{\frac{1}{2}}}{n^{\frac{1}{4}}}, \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}, \delta_4 \leq \frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + \sqrt{n}\delta_4^2$ in the last inequality. To finish the proof of Theorem 4.1, it remains to bound $T_1$. By the fact that

$$|g_z'(w+u) - g_z'(w+v) - \int_v^u h(w+u)du| \leq \mathbb{I}(z - \max(u,v) < w \leq z - \min(u,v))$$

(see[6],pp.250), we have

$$\begin{aligned} T_1 &= E\int_{-\infty}^{\infty}[g_z'(\widehat{Y}(\tau)) - g_z'(\widehat{Y}(\rho)+t)]K(t)dt \\ &= E\int_{-\infty}^{\infty}[g_z'(\widehat{Y}(\rho)+\Delta\widehat{Y}) - g_z'(\widehat{Y}(\rho)+t)]K(t)dt \\ &= T_{11} + T_{12} + T_{13} \end{aligned} \tag{4.34}$$

where

$$\begin{aligned} T_{11} &= E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| < \frac{z}{4}) \\ &\quad \int_{-\infty}^{\infty}\mathbb{I}(z - max(\Delta\widehat{Y}, t) < \widehat{Y}(\rho) < z - min(\Delta\widehat{Y}, t))K(t)dt, \\ T_{12} &= E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| < \frac{z}{4})\int_{-\infty}^{\infty}\int_t^{\Delta\widehat{Y}} h(\widehat{Y}(\rho)+u)K(t)du\,dt, \\ T_{13} &= E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| \geq \frac{z}{4})\int_{-\infty}^{\infty}[g_z'(\widehat{Y}(\rho)+\Delta\widehat{Y}) - g_z'(\widehat{Y}(\rho)+t)]K(t)dt. \end{aligned}$$

First, we consider $T_{11}$. For $\delta = |\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)|$, we used the idea from Chen and Shao([2],pp.243) to define $f_\delta : \mathbb{R} \to \mathbb{R}$ by

$$f_\delta(t) = \begin{cases} 0 & \text{if } t < z - 2\delta, \\ (1+t+\delta)(t-z+2\delta) & \text{if } z - 2\delta \leq t \leq z + 2\delta, \\ 4\delta(1+t+\delta) & \text{if } t > z + 2\delta. \end{cases}$$

Note that

$$|f_\delta(t)| \leq 4\delta|1+t+\delta| \qquad \text{for every } t \in \mathbb{R} \tag{4.35}$$

and

$$f_\delta'(t) \geq \begin{cases} 1+z-\delta & \text{it } z - 2\delta \leq t \leq z + 2\delta, \\ 0 & \text{otherwise.} \end{cases} \tag{4.36}$$

By Lemma 4.3,

$$E\int_{-\infty}^{\infty} f'_\delta(\widehat{Y}(\rho)+t)K(t)dt = E\widehat{Y}(\rho)f_\delta(\widehat{Y}(\rho)) - \Delta f_\delta(\widehat{Y}(\rho)). \tag{4.37}$$

We conclude from Lemma 4.4, (4.19), (4.22), (4.24) and (4.35) that

$$\begin{aligned}
&E|\widehat{Y}(\rho)f_\delta(\widehat{Y}(\rho))| \\
&\le 4E|\widehat{Y}(\rho)|\delta|1+\widehat{Y}(\rho)+\delta| \\
&= C\{E\widehat{Y}^2(\rho)\}^{\frac{1}{2}}\{[E|\Delta\widehat{Y}|^2]^{\frac{1}{2}} + [E|\widetilde{Y}(\rho)-\widehat{Y}(\rho)|^2]^{\frac{1}{2}}\} + [E|\Delta\widehat{Y}|^4]^{\frac{1}{2}} \\
&\quad + [E|\widetilde{Y}(\rho)-\widehat{Y}(\rho)|^4]^{\frac{1}{2}} + C[E\widehat{Y}^4(\rho)]^{\frac{1}{2}}\{[E|\Delta\widehat{Y}|^2]^{\frac{1}{2}} + [E|\widetilde{Y}(\rho)-\widehat{Y}(\rho)|^2]^{\frac{1}{2}}\} \\
&\quad + 2[E\widehat{Y}^2(\rho)|\Delta\widehat{Y}|^2]^{\frac{1}{2}}[E|\widetilde{Y}(\rho)-\widehat{Y}(\rho)|^2]^{\frac{1}{2}} \\
&\le C\left\{\frac{\delta_2^{\frac{1}{2}}}{n^{\frac{1}{4}}} + \frac{\delta_4^{\frac{1}{2}}}{n^{\frac{3}{4}}} + \delta_2^{\frac{1}{2}}\delta_4^{\frac{1}{2}} + \frac{\delta_4}{\sqrt{n}} + n^{\frac{1}{4}}\delta_4\delta_2^{\frac{1}{2}} + \frac{\delta_4^{\frac{1}{2}}\delta_2^{\frac{1}{2}}}{n^{\frac{1}{2}}} + \frac{\delta_2^{\frac{1}{2}}\delta_4^{\frac{3}{4}}}{n^{\frac{3}{8}}}\right\} \\
&\le C\left\{\frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + n^{\frac{1}{8}}\delta_4^{\frac{5}{4}}\right\}.
\end{aligned} \tag{4.38}$$

By the same technique of (4.38), we can use Lemma 4.4, (4.9), (4.19), (4.22), (4.24) and (4.35) to show that

$$\Delta f_\delta(\widehat{Y}(\rho)) \le \left\{\frac{C\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + C\delta_4\right\}\{Ef_\delta^2(\widehat{Y}(\rho))\}^{\frac{1}{2}} \le C\left\{\frac{\delta_4^{\frac{3}{4}}}{n^{\frac{9}{8}}} + \frac{\delta_4^{\frac{5}{2}}}{n^{\frac{1}{4}}}\right\}. \tag{4.39}$$

Thus, by (4.34), (4.36), (4.37), (4.38) and (4.39),

$$\begin{aligned}
T_{11} &\le E\int_{|t|\le|\widetilde{Y}(\rho)-\widehat{Y}(\rho)|} \mathbb{I}(\delta<\frac{z}{4})\mathbb{I}(z-(|\Delta\widehat{Y}|+|t|)<\widehat{Y}(\rho)<z+(|\Delta\widehat{Y}|+|t|))K(t)dt \\
&\le E\int_{|t|\le|\widetilde{Y}(\rho)-\widehat{Y}(\rho)|} \mathbb{I}(\delta<\frac{z}{4})\mathbb{I}(z-\delta<\widehat{Y}(\rho)<z+\delta)K(t)dt \\
&\le \frac{C}{1+z}E\int_{|t|\le|\widetilde{Y}(\rho)-\widehat{Y}(\rho)|} \mathbb{I}(\delta<\frac{z}{4})(1+z-\delta)\mathbb{I}(z-\delta<\widehat{Y}(\rho)<z+\delta)K(t)dt \\
&\le \frac{C}{1+z}E\int_{|t|\le|\widetilde{Y}(\rho)-\widehat{Y}(\rho)|} f'_\delta(\widehat{Y}(\rho)+t)K(t)dt \\
&\le \frac{C}{1+z}[E\widehat{Y}(\rho)f_\delta(\widehat{Y}(\rho)) - \Delta f_\delta(\widehat{Y}(\rho))] \\
&\le \frac{C}{1+z}\left\{\frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + n^{\frac{1}{8}}\delta_4^{\frac{5}{4}} + \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{9}{8}}} + \frac{\delta_4^{\frac{5}{2}}}{n^{\frac{1}{4}}}\right\} \\
&\le \frac{C}{1+z}\left\{\frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + \frac{\delta_4^{\frac{5}{2}}}{n^{\frac{1}{4}}}\right\}.
\end{aligned} \tag{4.40}$$

By the fact that $|g_z'(w) - g_z'(v)| \leq 1$(see[6],pp.246), (4.19), and (4.24),

$$\begin{aligned}
T_{13} &\leq E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| \geq \frac{z}{4}) \int_{-\infty}^{\infty} K(t)dt \\
&= \frac{n-1}{4} E\mathbb{I}(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| \geq \frac{z}{4}) E|\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^2 \\
&\leq \frac{n-1}{4} \{P(|\Delta\widehat{Y}| + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)| \geq \frac{z}{4})\}^{\frac{1}{2}} \{E(\widetilde{Y}(\rho) - \widehat{Y}(\rho))^4\}^{\frac{1}{2}} \\
&\leq Cn\{\frac{E[|\Delta\widehat{Y}|^2 + |\widetilde{Y}(\rho) - \widehat{Y}(\rho)|^2]}{z^2}\}^{\frac{1}{2}} \{\frac{\delta_4}{n\sqrt{n}}\}^{\frac{1}{2}} \\
&= \frac{C}{1+z} \delta_2^{\frac{1}{2}} \delta_4^{\frac{1}{2}} \\
&= \frac{C}{1+z} \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}. \qquad (4.41)
\end{aligned}$$

Thus, by Lemma 4.7, (4.40) and (4.41),

$$\begin{aligned}
|T_1| &\leq \frac{C}{1+z} \left\{ \frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + \frac{\delta_4^{\frac{5}{2}}}{n^{\frac{1}{4}}} + \frac{\delta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}} + \frac{\delta_8^{\frac{3}{8}}}{n^{\frac{5}{16}}} + \frac{\delta_8^{\frac{7}{8}}}{n^{\frac{1}{16}}} \right\} \\
&\leq \frac{C}{1+z} \left\{ \frac{\delta_8^{\frac{1}{8}}}{n^{\frac{7}{16}}} + n^{\frac{1}{8}} \delta_8^{\frac{5}{4}} \right\}. \qquad (4.42)
\end{aligned}$$

So, by (4.28), (4.29), (4.30), (4.33) and (4.42),

$$\begin{aligned}
|P(W \leq z) - \Phi(z)| &\leq \frac{C}{1+z} \left\{ \frac{\delta_4}{\sqrt{n}} + \frac{\delta_8^{\frac{1}{8}}}{n^{\frac{7}{16}}} + n^{\frac{1}{8}} \delta_8^{\frac{5}{4}} + \frac{1}{n} + \frac{\delta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + \sqrt{n}\delta_4^2 \right\} \\
&\leq \frac{C}{1+z} \left\{ \frac{\delta_8^{\frac{1}{8}}}{n^{\frac{7}{16}}} + \frac{1}{n} + \delta_8 \right\}.
\end{aligned}$$

Now, we prove the main theorem. □

# CHAPTER V
# BOUNDS ON NORMAL APPROXIMATION FOR LATIN HYPERCUBE SAMPLING

In many problems, we are interested in estimating an integral over the $d$-dimensional hypercube $[0,1]^d$;

$$\mu = \int_{[0,1]^d} f(x)dx.$$

This is equivalent to finding $E(f(X))$, where $X$ is a random vector uniformly distributed on $[0,1]^d$.

Among numerical integration techniques, Monte Carlo methods are especially useful and competitive for high-dimensional integration. The Monte Carlo estimate for the integral $\mu = E(f(X)) = \int_{[0,1]^d} f(x)dx$ is given by

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

where $X_1, ..., X_n$ is a random sampling on $[0,1]^d$.

In 1979, McKay, Beckman and Conover[16] proposed Latin hypercube sampling(LHS) as an attractive alternative to generating $X_1, ..., X_n$. Let

1. $\eta_k,\ 1 \leq k \leq d$ be random permutations of $\{1, ..., n\}$ each uniformly distributed over all the $n!$ possible permutations;
2. $U_{i_1,...,i_d,j}$ $1 \leq i_1, ..., i_d \leq n,\ 1 \leq j \leq d$, be $[0,1]$ uniform random variables;
3. the $U_{i_1,...,i_d,j}$'s and $\eta_k$'s all be stochastically independent.

A Latin hypercube sample of size $n$ (taken from the $d$-dimensional hypercube $[0,1]^d$) is defined to be $\{X(\eta_1(i), \eta_2(i), ..., \eta_d(i)) : 1 \leq i \leq n\}$, where for all $1 \leq i_1, ..., i_d \leq n,\ 1 \leq j \leq d$,

$$X_j(i_1, ..., i_d) = (i_j - U_{i_1,...,i_d,j})/n,$$

and

$$X(i_1, ..., i_d) = (X_1(i_1, ..., i_d), ..., X_d(i_1, ..., i_d)).$$

Hence the estimator for $\mu$ that based on a Latin hypercube sampling is

$$\hat{\mu}_n = \frac{1}{n}\sum_{k=1}^{n} f\circ X(\eta_1(k),\eta_2(k),...,\eta_d(k)).$$

Then $\hat{\mu}_n$ is an unbiased estimator for $\mu$.
Assume that $Var(\hat{\mu}_n) > 0$, we define

$$L = \frac{\hat{\mu}_n - \mu}{\sqrt{Var(\hat{\mu}_n)}}. \tag{5.1}$$

Loh[14] used Stein's method to show that the distribution of $L$ can be approximated by normal distribution but he yield the convergence rate $\frac{C}{\sqrt{n}}$ under the finiteness of third moments without the value of $C$. In this chapter, we give a constant $C$ by using Stein's method. Furthermore, we give a non-uniform bound of this approximation. These are our main results.

**Theorem 5.1.** *(A uniform bound for LHS) For $1 \le i_1,...,i_d \le n$, let*

$$\mu(i_1,...,i_d) = Ef(X(i_1,...,i_d)),$$
$$\mu_k(i_1,...,i_d) = \frac{(-1)^k}{n^k}\sum_{1\le j_1<j_2<...<j_k\le d}\ \sum_{q_{j_1}=1}^{n}...\sum_{q_{j_k}=1}^{n}\mu(l_1,...,l_d),$$

*where*

$$l_p = \begin{cases} q_p & if\ \ p = j_1,...,j_k, \\ i_p & otherwise, \end{cases} \tag{5.2}$$

*and*

$$V(i_1,...,i_d) = \frac{1}{n\sqrt{var(\hat{\mu}_n)}}[f(X(i_1,...,i_d)) + \sum_{k=1}^{d-1}\mu_k(i_1,...,i_d) + (-1)^d\mu]. \tag{5.3}$$

*Then for $n \ge 6^d + 3$,*

$$\sup_{z\in\mathbb{R}}|P(L\le z) - \Phi(z)| \le \frac{11.946}{\sqrt{n}} + \frac{1.037\sqrt{d}\beta_4^{\frac{1}{4}}}{n^{\frac{3}{8}}} + 8.314d^{\frac{1}{4}}\beta_4 + 11.765\beta_4 + \frac{5.014d\beta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}} + \frac{2\sqrt{2\pi}\beta_4^{\frac{3}{4}}}{n^{\frac{1}{8}}}$$

*where*

$$\beta_4 = \frac{1}{n^{d-\frac{3}{2}}}\sum_{i_1=1}^{n}...\sum_{i_d=1}^{n}E|V(i_1,...,i_d)|^4.$$

**Corollary 5.2.** *If* $E|f(X)|^4 < \infty$*, then*

$$\sup_{z\in\mathbb{R}} |P(L \le z) - \Phi(z)| \le \frac{C_d(28.725 + 1.037\sqrt{d} + 8.314d^{\frac{1}{4}} + 5.014d)}{\sqrt{n}}$$

*where* $C_d = \frac{27}{C^2}[2 + (d-1)^3 \sum_{k=1}^{d-1} \binom{d}{k}]E|f(X)|^4$ *for some constant* $C$.

**Theorem 5.3.** *(A non-uniform bound for LHS) Let* $z \in \mathbb{R}$*. Then there exists a positive constant* $C$ *which does not depend on* $z$ *such that*

$$\big|P(L \le z) - \Phi(z)\big| \le \frac{C}{1+|z|}\left\{\frac{\beta_8^{\frac{1}{8}}}{n^{\frac{7}{16}}} + \frac{1}{n} + \beta_8\right\}$$

*where*

$$\beta_8 = \frac{1}{n^{d-\frac{7}{2}}} \sum_{i_1=1}^{n} ... \sum_{i_d=1}^{n} E|V(i_1, ..., i_d)|^8.$$

**Corollary 5.4.** *If* $E|f(X)|^8 < \infty$, *then*

$$\big|P(L \le z) - \Phi(z)\big| \le \frac{C}{(1+|z|)\sqrt{n}}.$$

## 5.1 Proof of Main Results

### 5.1.1 Proof of 5.1 and Theorem 5.3

We shall prove Theorem 5.1 and Theorem 5.3 by appling Theorem 3.1 and Theorem 4.1, respectively. Note from (5.1) that $EL = 0$ and $VarL = 1$. Hence (3.2) holds. In order to apply Theorem 3.1 and Theorem 4.1, we need to prove the following:

a)

$$L = \sum_{i=1}^{n} V(i, \pi_1(i), ..., \pi_{d-1}(i))$$

where $V(i_1, ..., i_d)$ is defined by (5.3) and $\pi_1, ..., \pi_{d-1}$ are random permutations of {1,...,n} such that $V(i_1, ..., i_d)$'s and $\pi_k$'s are stochastically independent.

b)

$$\sum_{i_j}^{n} EV(i_1, ..., i_d) = 0,$$

for every $j \in \{1, ..., d\}$. There are proved in Lemma 5.5 and Lemma 5.6.

**Lemma 5.5.** *There are random permutations* $\pi_1, ..., \pi_{d-1}$ *of* $\{1, ..., n\}$ *such that* $V(i_1, ..., i_d)$*'s and* $\pi_k$*'s are stochastically independent and*

$$L = \sum_{i=1}^{n} V(i, \pi_1(i), ..., \pi_{d-1}(i)).$$

*Proof.* Let $S_n$ be the set of all permutations of $\{1, 2, ..., n\}$. For $j \in \{1, ..., d-1\}$, we define

$$\pi_j(w) = \eta_{j+1}(w)(\eta_1(w)^{-1}).$$

Since $\eta_1, ..., \eta_d$ are independent,

$$\begin{aligned}
P(\pi_j = \gamma) &= P(\{w | \eta_{j+1}(w) = \gamma(\eta_1(w))\}) \\
&= P(\cup_{\alpha \in S_n}\{w | \eta_{j+1}(w) = \gamma(\eta_1(w)), \eta_1(w) = \alpha\}) \\
&= \sum_{\alpha \in S_n} P(\eta_{j+1} = \gamma(\alpha), \eta_1 = \alpha) \\
&= \sum_{\alpha \in S_n} P(\eta_{j+1} = \gamma(\alpha))P(\eta_1 = \alpha) \\
&= \frac{1}{n!}
\end{aligned}$$

for any $j \in \{1, ..., d-1\}$ and $\gamma \in S_n$. Hence $\pi_1, ..., \pi_{d-1}$ are random permutations. We observe that for $\gamma_1, ..., \gamma_{d-1} \in S_n$,

$$\begin{aligned}
&P(\pi_1 = \gamma_1, ..., \pi_{d-1} = \gamma_{d-1}) \\
&= P(\{w \mid \eta_2(w) = \gamma_1(\eta_1(w)), ..., \eta_d(w) = \gamma_{d-1}(\eta_1(w)))\\
&= P(\cup_{\alpha \in S_n}\{w \mid \eta_2(w) = \gamma_1(\eta_1(w)), ..., \eta_d(w) = \gamma_{d-1}(\eta_1(w)), \eta_1(w) = \alpha\}) \\
&= \sum_{\alpha \in S_n} P(\{\eta_2 = \gamma_1(\alpha), ..., \eta_d = \gamma_{d-1}(\alpha), \eta_1 = \alpha\}) \\
&= \frac{1}{(n!)^{d-1}} \\
&= P(\pi_1 = \gamma_1)...P(\pi_{d-1} = \gamma_{d-1}).
\end{aligned}$$

So, $\pi_j$'s are independent. It's easy to see that $V(i_1, ..., i_d)$'s and $\pi_j$'s are independent. By the fact that, for fixed $i$,

$$\sum_{i_{j_1}=1}^{n} ... \sum_{i_{j_k}=1}^{n} \mu(l_1, ..., l_d)$$

where $1 \leq j_1 < j_2 < ... < j_k \leq d$ and

$$l_p = \begin{cases} i_p & \text{if } \ p = j_1, ..., j_k, \\ \eta_p(i) & \text{otherwise} \end{cases}$$

is equal to

$$\frac{1}{n^{d-k}} \sum_{k_1=1}^{n} ... \sum_{k_d=1}^{n} \mu(k_1, ..., k_d)$$

and

$$\mu = E\hat{\mu}_n = \frac{1}{n^d} \sum_{k_1=1}^{n} ... \sum_{k_d=1}^{n} \mu(k_1, ..., k_d), \tag{5.4}$$

we have

$$\begin{aligned} \mu_k(\eta_1(i), ..., \eta_d(i)) &= \frac{(-1)^k}{n^d} \sum_{1 \leq j_1 < j_2 < ... < j_k \leq d} \sum_{k_1=1}^{n} ... \sum_{k_d=1}^{n} \mu(k_1, ..., k_d) \\ &= (-1)^k \sum_{1 \leq j_1 < j_2 < ... < j_k \leq d} \mu \\ &= (-1)^k \binom{d}{k} \mu. \end{aligned}$$

So,

$$\begin{aligned} L &= \frac{1}{\sqrt{Var(\hat{\mu}_n)}} (\frac{1}{n} \sum_{i=1}^{n} f(X(\eta_1(i), ..., \eta_d(i))) - \mu) \\ &= \frac{1}{\sqrt{Var(\hat{\mu}_n)}} (\frac{1}{n} \sum_{i=1}^{n} f(X(\eta_1(i), ..., \eta_d(i))) + \mu[(1-1)^d - 1]) \\ &= \frac{1}{\sqrt{Var(\hat{\mu}_n)}} (\frac{1}{n} \sum_{i=1}^{n} f(X(\eta_1(i), ..., \eta_d(i))) + \mu \sum_{k=1}^{d} (-1)^k \binom{d}{k}) \\ &= \frac{1}{\sqrt{Var(\hat{\mu}_n)}} (\frac{1}{n} \sum_{i=1}^{n} f(X(\eta_1(i), ..., \eta_d(i))) + \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{d-1} \mu_k(\eta_1(i), ..., \eta_d(i)) + (-1)^d \mu) \\ &= \sum_{i=1}^{n} V(\eta_1(i), ..., \eta_d(i)). \end{aligned} \tag{5.5}$$

By this fact and the fact that

$$\eta_{k+1}(i)(w) = (\eta_{k+1}(w))(i) = (\eta_{k+1}(w)(\eta_1(w)^{-1}))(j) = (\pi_k(w))(j) = \pi_k(j)(w)$$

for $(\eta_1(w))(i) = j$ and $k = 1, ...d-1$, we have

$$L = \sum_{i=1}^{n} V(\eta_1(i), ..., \eta_d(i)) = \sum_{j=1}^{n} V(j, \pi_1(j), ..., \pi_{d-1}(j)).$$

□

**Lemma 5.6.** *Let $V(i_1, ..., i_d)$ be defined as in (5.3). Then*

$$\sum_{i_j}^{n} EV(i_1, ..., i_d) = 0,$$

*for every $j \in \{1, ..., d\}$.*

*Proof.* First, we shall proof in the case of $j = 1$. For arbitary $j$, it follows from the same way. Let $\mu_k(i_1, ..., i_d)$ and $l_1, ..., l_d$ be defined as in (5.2). Then We observe that

$$\begin{aligned}
&\sum_{i_1=1}^{n}\sum_{k=1}^{d-1}\mu_k(i_1, ..., i_d)\\
&= \sum_{i_1=1}^{n}\sum_{k=1}^{d-1}\frac{(-1)^k}{n^k}\sum_{1\leq j_1<j_2<...<j_k\leq d}\sum_{q_{j_1}=1}^{n} ... \sum_{q_{j_k}=1}^{n}\mu(l_1, ..., l_d)\\
&= \sum_{k=1}^{d-1}\frac{(-1)^k}{n^{k-1}}\sum_{1=j_1<j_2<...<j_k\leq d}\sum_{q_1=1}^{n}\sum_{q_{j_2}=1}^{n} ... \sum_{q_{j_k}=1}^{n}\mu(q_1, l_2..., l_d)\\
&\quad + \sum_{k=1}^{d-1}\frac{(-1)^k}{n^k}\sum_{i_1=1}^{n}\sum_{2\leq j_1<...<j_k\leq d}\sum_{q_{j_1}=1}^{n} ... \sum_{q_{j_k}=1}^{n}\mu(i_1, l_2, ..., l_d)\\
&= -\sum_{q_1=1}^{n}\mu(q_1, i_2, ..., i_d) + \sum_{k=2}^{d-1}\frac{(-1)^k}{n^{k-1}}\sum_{1=j_1<j_2<...<j_k\leq d}\sum_{q_1=1}^{n}\sum_{q_{j_2}=1}^{n} ... \sum_{q_{j_k}=1}^{n}\mu(q_1, l_2, ..., l_d)\\
&\quad + \sum_{k=1}^{d-2}\frac{(-1)^k}{n^k}\sum_{i_1=1}^{n}\sum_{2\leq j_1<...<j_k\leq d}\sum_{q_{j_1}=1}^{n} ... \sum_{q_{j_k}=1}^{n}\mu(i_1, l_2, ..., l_d)\\
&\quad + \frac{(-1)^{(d-1)}}{n^{d-1}}\sum_{i_1=1}^{n}\sum_{q_2=1}^{n} ... \sum_{q_d=1}^{n}\mu(i_1, q_2..., q_d)\\
&= -\sum_{q_1=1}^{n}\mu(q_1, i_2, ..., i_d) + \sum_{k=2}^{d-1}\frac{(-1)^k}{n^{k-1}}\sum_{1=j_1<j_2<...<j_k\leq d}\sum_{q_1=1}^{n}\sum_{q_{j_2}=1}^{n} ... \sum_{q_{j_k}=1}^{n}\mu(q_1, l_2, ..., l_d)\\
&\quad + \sum_{k=2}^{d-1}\frac{(-1)^{k-1}}{n^{k-1}}\sum_{2\leq j_1<...<j_{(k-1)}\leq d}\sum_{i_1=1}^{n}\sum_{q_{j_1}=1}^{n} ... \sum_{q_{j_{(k-1)}}=1}^{n}\mu(i_1, l_2, ..., l_d) + (-1)^{d-1}n\mu\\
&= -\sum_{q_1=1}^{n}\mu(q_1, i_2, ..., i_d) + (-1)^{d-1}n\mu.
\end{aligned}$$

Hence for fixed $i_2, ..., i_d \in \{1, ..., n\}$,

$$
\begin{aligned}
\sum_{i_1=1}^{n} EV(i_1, i_2, ..., i_d) &= \frac{1}{n\sqrt{Var(\hat{\mu}_n)}} \sum_{i_1=1}^{n} E[f \circ X(i_1, ..., i_d) + \sum_{k=1}^{d-1} \mu_k(i_1, ..., i_d) + (-1)^d \mu] \\
&= \frac{1}{n\sqrt{Var(\hat{\mu}_n)}} \{\sum_{i_1=1}^{n} \mu(i_1, ..., i_d) + \sum_{i_1=1}^{n} \sum_{k=1}^{d-1} \mu_k(i_1, ..., i_d) + (-1)^d n\mu] \\
&= 0.
\end{aligned}
$$

□

### 5.1.2 Proof of Corollary 5.2 and Corollary 5.4

*Proof.* Let $\mu_k(i_1, ..., i_d)$ and $l_1, ..., l_d$ be defined as in (5.2). From Stein[27], $Var(\hat{\mu}_n) \geq \frac{C}{n}$ for some constant $C$. Thus

$$
\begin{aligned}
& E|V(i_1, ..., i_d)|^4 \\
&\leq \frac{1}{n^2C^2} E[f(X(i_1, ..., i_d)) + \sum_{k=1}^{d-1} \mu_k(i_1, ..., i_d) + (-1)^d \mu]^4 \\
&\leq \frac{27}{n^2C^2} [E|f(X(i_1, ..., i_d))|^4 + E(\sum_{k=1}^{d-1} \mu_k(i_1, ..., i_d))^4 + \mu^4] \\
&\leq \frac{27}{n^2C^2} [E|f(X)|^4 + (d-1)^3 \sum_{k=1}^{d-1} E\mu_k^4(i_1, ..., i_d) + E|f(X)|^4] \\
&= \frac{27}{n^2C^2} [2E|f(X)|^4 + (d-1)^3 \sum_{k=1}^{d-1} \frac{1}{n^{4k}} E( \sum_{1 \leq j_1 < j_2 < ... < j_k \leq d} \sum_{q_{j_1}} ... \sum_{q_{j_k}} \mu(l_1, ..., l_d))^4] \\
&\leq \frac{27}{n^2C^2} [2E|f(X)|^4 + (d-1)^3 \sum_{k=1}^{d-1} \frac{1}{n^{4k}} (\binom{d}{k} n^k)^4 E|f(X)|^4] \\
&= \frac{27}{n^2C^2} [2 + (d-1)^3 \sum_{k=1}^{d-1} \binom{d}{k}^4 ] E|f(X)|^4.
\end{aligned}
$$

So, $\beta_4 \leq \frac{27}{\sqrt{n}C^2}[2 + (d-1)^3 \sum_{k=1}^{d-1} \binom{d}{k}^4 ]E|f(X)|^4$. By the same argument,

$$
\beta_8 \leq \frac{2187}{\sqrt{n}C^4}[2 + (d-1)^7 \sum_{k=1}^{d-1} \binom{d}{k}^8 ]E|f(X)|^8.
$$

Corollary 5.2 and Corollary 5.4 are proved by applying this fact to Theorem 5.1 and Theorem 5.3, respectively. □

# REFERENCES


[1] Barbour, A.D. and Chen, L.H.Y. (2005). *An introduction to Stein's method.* Singapore University Press, Singapore.

[2] Bauer, H. (1996). *Probability Theory.* Walter de Gruyter, Berlin.

[3] Bolthausen E. (1984). An estimate of the remainder in a combinatorial central limiy theorem. *Z.Wahrsch. Verw. Gebiete* **66**:397-386.

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

[5] Chen, L.H.Y. (1975). Poisson approximation for dependent trials. *Ann.Prob.* **3**:534-545.

[6] Chen, L.H.Y. and Shao, Q.M. (2001). A non-uniform Berry-Esseen bound via Stein's method. *Probab. Theory Relat. Fields.* **120**:236-254.

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

[8] Ho, S.T. (1975). The remainders in the central limit theorem and in a generalization of Hoeffding's combinatorial limit theorem. M.Sc. thesis, Uni. of Singapore.

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

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

[11] Kolchin, V.F. and Chistyakov, V.P. (1973). On a combinatorial central limit theorem. *Theory. Probability Appl.* **18**:728-739.

[12] Laipaporn, K. and Neammanee, K. (2007). A non-uniform bound on normal approximation of randomized orthogonal array sampling designs. *International Mathematical Forum.* **48**:2347-2367.

[13] Loh, W.L. (1996). A combinatorial central limit theorem for randomized orthogonal array sampling designs. *Ann Statist.* **24**:1209-1224.

[14] Loh, W.L. (1996). On Latin hypercube sampling. *Ann. Stist.* **24**:2058-2080.

[15] Matoo, M. (1957). On the Hoeffding's combinatorial central limit theorem. *Ann. Inst. Statist. Math.* **8**:145-154.

[16] McKay, M.D., Conover, W.J. and Beckman, R.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics.* **21**:239-245.

[17] Neammanee, K. and Suntornchost, J. (2005). A uniform bound on a combinatorial central limit theorem. *Stochastic Analysis and Applications.* **23**:559-578.

[18] Owen, A.B. (1992). A central limit theorem for Latin hypercube sampling. *J.R.Statist. Soc. Ser.B.* **54**:541-551.

[19] Owen, A.B. (1997). Monte-Carlo variance of scrambled net quadrature. *SIAM J. Numer. Anal.* **34**:1884-1910.

[20] Owen, A.B. (1992). Orthogonal array for computer experiments, integration and visualization. *Statist. Sinica.* **2**:439-452.

[21] Owen, A.B. (1997). Scrambled net variance for integrals of smooth functions. *Ann. Statist.* **25**:1541-1562.

[22] Patterson, H.D. (1954). The errors of lattice sampling. *J.R. Statist. Soc. Ser. B.* **16**:140-149.

[23] Petrov, V.V. (1995). *Limit Theorem of Probability Theory; Sequences of Independent Random Variables.* Clarendon Press, Oxford.

[24] Robinson, J. (1972). A converse to a combinatorial limit theorem. *Ann. Math. Statist.* **43**:2053-2057.

[25] Stein, C.M. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent randon variables. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2**:583-602. Univ. California Press, Berkeley.

[26] Stein, C.M. (1986). *Approximate computation of expectation.* IMS, Hayward, CA.

[27] Stein, C.M. (1987). Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics.* **29** 143-151.

[28] Tang, B. (1993). Orthogonal array-based Latin hypercubes. *J. Amer. Statist. Assoc.* **88**:1392-1397.

[29] Wald, A. and Wolfowitz, J. (1944). Statistical tests on permutations of observations. *Ann. Math. Statist.* **15**:358-372.

[30] Von Bahr, B. (1976). Remainder term estimate in a combinatorial central limit theorem. *Z. Wahrsch. Verw. Gebiete* **35**:131-139.

# VITA

| | |
|---|---|
| Name | :Miss Petcharat Rattanawong |
| Date of Birth | :September 20, 1977 |
| Place of Birth | :Prachinburi, Thailand |
| Previous Studies | :B.Sc.(Mathematics), Chulalongkorn University, 1999, |
| | M.Sc.(Mathematics), Chulalongkorn University, 2001. |
| Work Experience | :Burapha University Sakeao Campus. |