The Annals of Probability
2019, Vol. 47, No. 1, 61–108
https://doi.org/10.1214/18-AOP1255
© Institute of Mathematical Statistics, 2019

BERRY–ESSEEN BOUNDS OF NORMAL AND NONNORMAL APPROXIMATION FOR UNBOUNDED EXCHANGEABLE PAIRS[1]

BY QI-MAN SHAO AND ZHUO-SONG ZHANG

Chinese University of Hong Kong

An exchangeable pair approach is commonly taken in the normal and nonnormal approximation using Stein's method. It has been successfully used to identify the limiting distribution and provide an error of approximation. However, when the difference of the exchangeable pair is not bounded by a small deterministic constant, the error bound is often not optimal. In this paper, using the exchangeable pair approach of Stein's method, a new Berry–Esseen bound for an arbitrary random variable is established without a bound on the difference of the exchangeable pair. An optimal convergence rate for normal and nonnormal approximation is achieved when the result is applied to various examples including the quadratic forms, general Curie–Weiss model, mean field Heisenberg model and colored graph model.

1. Introduction. Let W_n be a sequence of random variables under study. Using the exchangeable pair approach of Stein's method, Chatterjee and Shao [10] and Shao and Zhang [27], provided a concrete tool to identify the limiting distribution of W_n as well as the L_1 bound (the Wasserstein distance) of the approximation. Our aim in this paper is to establish the Berry–Esseen-type bound for the approximation.

Write $W = W_n$ and let (W, W') be an exchangeable pair, that is, (W, W') and (W', W) have the same joint distribution. Put $\Delta = W - W'$. For the normal approximation, assume that

$$\mathrm{E}(\Delta \mid W) = \lambda(W + R).$$

Then, by Stein [28] (see also Proposition 2.4 in Chen, Goldstein and Shao [12]), for any absolutely continuous function h with $\|h'\| < \infty$,

$$\big|\mathrm{E}h(W) - \mathrm{E}h(Z)\big| \le 2\|h'\|\bigg(\mathrm{E}\bigg|1 - \frac{1}{2\lambda}\mathrm{E}(\Delta^2 \mid W)\bigg| + \frac{1}{\lambda}\mathrm{E}|\Delta|^3 + \mathrm{E}|R|\bigg).$$

Received November 2016; revised December 2017.

[1]Supported in part by Hong Kong RGC GRF 14302515 and 14304917.

MSC2010 subject classifications. Primary 60F05; secondary 60K35.

Key words and phrases. Stein's method, exchangeable pairs, Berry–Esseen bound, quadratic forms, simple random sampling, general Curie–Weiss model, mean field Heisenberg model, monochromatic edges.

Here and in the sequel, Z denotes the standard normal random variable. For the Berry–Esseen bound, we have

$$\sup_{z\in\mathbb{R}}\bigl|P(W\le z)-\Phi(z)\bigr|\le \mathrm{E}\left|1-\frac{1}{2\lambda}\mathrm{E}\bigl(\Delta^2\mid W\bigr)\right|+\mathrm{E}|R|+\left(\frac{\mathrm{E}|\Delta|^3}{\lambda}\right)^{1/2}, \tag{1.1}$$

where Φ is the standard normal distribution function. If in addition $|\Delta|\le\delta$ for some constant δ, then by Rinott and Rotar [24] (see also Shao and Su [25]),

$$\sup_{z\in\mathbb{R}}\bigl|P(W\le z)-\Phi(z)\bigr|\le \mathrm{E}\left|1-\frac{1}{2\lambda}\mathrm{E}\bigl(\Delta^2\mid W\bigr)\right|+\mathrm{E}|R|+1.5\delta+\delta^3/\lambda. \tag{1.2}$$

It is known that (1.1) usually fails to provide an optimal bound. Similarly, the bound in (1.2) may not be optimal unless δ is small enough. Hence, it would be interesting to seek an optimal Berry–Esseen bound for an unbounded Δ. To this end, Chen and Shao [13] established the following Berry–Esseen bound:

$$\begin{aligned}\sup_{z\in\mathbb{R}}&\bigl|P(W\le z)-\Phi(z)\bigr|\\ &\le \mathrm{E}|R|+\frac{1}{4\lambda}\mathrm{E}\bigl(|W|+1\bigr)\bigl|\Delta^3\bigr|\\ &\quad+(1+\tau^2)\biggl(4(1+\tau)\lambda^{1/2}+6\mathrm{E}\left|\frac{1}{2\lambda}\mathrm{E}(\Delta^2\mid W)-1\right|\\ &\quad+\frac{2}{\mathrm{E}[\Lambda]}\mathrm{E}\bigl|\Lambda-\mathrm{E}[\Lambda]\bigr|\biggr),\end{aligned} \tag{1.3}$$

where Λ is any random variable such that $\Lambda\ge \mathrm{E}(\Delta^4\mid W)$ and $\tau=\sqrt{\mathrm{E}(\Lambda)}/\lambda$. They obtained an optimal Berry–Esseen bound when the result was applied to an independence test by sums of squared sample correlation functions. However, (1.3) is still too complicated in general.

For the nonnormal approximation, Chatterjee and Shao [10] developed similar results for both the L_1 bound and Berry–Esseen bound.

The exchangeable pair approach of Stein's method has been widely used in the literature. For example, Chatterjee and Meckes [9], Reinert and Röllin [22] and Meckes [21] established the L_1 bounds for multivariate normal approximation, and Chatterjee [4] and Chatterjee and Dey [7] obtained the concentration inequalities. We refer to Chen, Goldstein and Shao [12] and Chatterjee [6] for recent developments on Stein's method.

In this paper, we establish a new Berry–Esseen-type bound for normal and nonnormal approximation via exchangeable pairs. The bound is as simple as

$$\mathrm{E}\left|1-\frac{1}{2\lambda}\mathrm{E}\bigl(\Delta^2\mid W\bigr)\right|+\mathrm{E}\bigl|\mathrm{E}\bigl(\Delta|\Delta|\mid W\bigr)\bigr|+\mathrm{E}|R|,$$

which yields an optimal bound in many applications.

The paper is organized as follows. The main results are presented in Section 2. Section 3 gives applications to the quadratic forms, general Curie–Weiss model, mean field Heisenberg model and colored graph model. The proof of the main results is given in Section 4. Other proofs of applications are postponed to Section 5.

2. Main results. In this section, we establish Berry–Esseen bounds for normal and nonnormal approximation via the exchangeable pair approach without the boundedness assumption.

2.1. *Normal approximation.* We first present a new Berry–Esseen bound for normal approximation, which is a refinement of (1.1), (1.2) and (1.3).

THEOREM 2.1. *Let (W, W') be an exchangeable pair satisfying*

$$\mathrm{E}(\Delta \mid W) = \lambda(W + R), \tag{2.1}$$

for some constant $\lambda \in (0, 1)$ and random variable R, where $\Delta = W - W'$. Then

$$\begin{aligned}\sup_{z\in\mathbb{R}}&\left|\mathrm{P}(W \le z) - \Phi(z)\right| \\ &\le \mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}(\Delta^2 \mid W)\right| + \mathrm{E}|R| + \frac{1}{\lambda}\mathrm{E}\left|\mathrm{E}(\Delta\Delta^* \mid W)\right|,\end{aligned}$$

where $\Delta^* := \Delta^*(W, W')$ is any random variable satisfying $\Delta^*(W, W') = \Delta^*(W', W)$ and $\Delta^* \ge |\Delta|$.

The following two corollaries may be useful.

COROLLARY 2.1. *If $|\Delta| \le \delta$ and $\mathrm{E}|W| \le 2$, then*

$$\sup_{z\in\mathbb{R}}\left|\mathrm{P}(W \le z) - \Phi(z)\right| \le \mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}(\Delta^2 \mid W)\right| + \mathrm{E}|R| + 3\delta.$$

Notice that the term δ^3/λ in (1.2) does not appear in the preceding corollary. One can check that under $|\Delta| \le \delta$,

$$\min\left(1, \mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}(\Delta^2 \mid W)\right| + \delta\right) \le 2\min\left(1, \mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}(\Delta^2 \mid W)\right| + \delta^3/\lambda\right).$$

Hence, Corollary 2.1 is an improvement of (1.2) at the cost of assuming $\mathrm{E}|W| \le 2$, which is easily satisfied.

It follows from the Cauchy inequality that for any $a > 0$,

$$|\Delta| \le a/2 + \Delta^2/(2a).$$

Thus, we can choose $\Delta^* = a/2 + \Delta^2/(2a)$ with a proper constant a and obtain the following corollary.

COROLLARY 2.2. *Assume that* $\mathrm{E}|W| \le 2$. *Then, under the condition of Theorem* 2.1,

$$\sup_{z\in\mathbb{R}}\bigl|\mathrm{P}(W \le z) - \Phi(z)\bigr| \le \mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}\bigl(\Delta^2 \mid W\bigr)\right| + \mathrm{E}|R| + 2\sqrt{\frac{\mathrm{E}|\mathrm{E}(\Delta^3 \mid W)|}{\lambda}}.$$

Clearly, $\mathrm{E}|\mathrm{E}(\Delta^3 \mid W)| \le \mathrm{E}|\Delta|^3$. Hence, Corollary 2.2 improves (1.1). In fact, Corollary 2.2 could yield an optimal bound while (1.1) may not.

2.2. *Nonnormal approximation.* In this subsection, we focus on the Berry–Esseen bound for nonnormal approximation.

Let W be a random variable satisfying $\mathrm{P}(a < W < b) = 1$ where $-\infty \le a < b \le \infty$. Let (W, W') be an exchangeable pair satisfying

$$\mathrm{E}\bigl(W - W' \mid W\bigr) = \lambda\bigl(g(W) + R\bigr), \tag{2.2}$$

where g is a measurable function with domain (a, b), $\lambda \in (0, 1)$ and R is a random variable.

Assume that g satisfies the following conditions:

(A1) g is nondecreasing, and there exists $w_0 \in (a, b)$ such that $(w - w_0)g(w) \ge 0$ for $w \in (a, b)$;
(A2) g' is continuous and $2(g'(w))^2 - g(w)g''(w) \ge 0$ for all $w \in (a, b)$; and
(A3) $\lim_{y\downarrow a} g(y)p(y) = \lim_{y\uparrow b} g(y)p(y) = 0$, where

$$p(y) = c_1 e^{-G(y)}, \qquad G(y) = \int_{w_0}^{y} g(t)\,dt, \tag{2.3}$$

and c_1 is the constant so that $\int_a^b p(y)\,dy = 1$.

Let Y be a random variable with the probability density function (p.d.f.) $p(y)$, and let $\Delta = W - W'$.

THEOREM 2.2. *We have*

$$\sup_{z\in\mathbb{R}}\bigl|\mathrm{P}(W \le z) - \mathrm{P}(Y \le z)\bigr| \le \mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}\bigl(\Delta^2 \mid W\bigr)\right| + \frac{1}{\lambda}\mathrm{E}\bigl|\mathrm{E}\bigl(\Delta\Delta^* \mid W\bigr)\bigr| + \frac{1}{c_1}\mathrm{E}|R|, \tag{2.4}$$

where $\Delta^* := \Delta^*(W, W')$ *is any random variable satisfying* $\Delta^*(W, W') = \Delta^*(W', W)$ *and* $\Delta^* \ge |\Delta|$.

To make the bound meaningful, one should choose $\lambda \sim (1/2)E(\Delta^2)$. It is easy to see that $g(w) = w$ satisfies conditions (A1)–(A3). More generally, (A1)–(A3) are also satisfied for $g(w) = w^{2k-1}$, where $k \ge 1$ is an integer.

3. Applications. In this section, we give some applications for our main result.

3.1. *Quadratic forms.* We first consider a classical example as a simple application. Suppose $X_1, X_2, \dots$ are i.i.d. random varaibles with a zero mean, unit variance and a finite fourth moment. Let $A = \{a_{ij}\}_{i,j=1}^n$ be a real symmetric matrix and let $W_n = \sum_{1 \le i \ne j \le n} a_{ij} X_i X_j$. The central limit theorem for W_n has been extensively discussed in the literature. For example, de Jong [14] used U-statistics and proved a central limit theorem for W_n when

$$\sigma_n^{-4} \operatorname{Tr}(A^4) \to 0 \quad \text{and} \quad \sigma_n^{-2} \max_{1 \le i \le n} \sum_{1 \le j \le n} a_{ij}^2 \to 0,$$

where $\sigma_n^2 = 2 \operatorname{Tr}(A^2) = \operatorname{Var}(W_n)$. An L_1 bound was given by Chatterjee [5] while Götze and Tikhomirov [19] gave a Kolmogorov distance with a convergence rate λ_1/σ_n, where λ_1 the largest absolute eigenvalue of A.

Here, we apply Theorem 2.1 and obtain the following result.

THEOREM 3.1. *Let $X_1, X_2, \dots$ be i.i.d. random variables with a zero mean, unit variance and a finite fourth moment. Let $A = (a_{ij})_{i,j=1}^n$ be a real symmetric matrix with $a_{ii} = 0$ for all $1 \le i \le n$ and $\sigma_n^2 = 2\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2$. Put $W_n = \frac{1}{\sigma_n} \sum_{i \ne j} a_{ij} X_i X_j$. Then*

$$\begin{aligned} &\sup_{x \in \mathbb{R}} \left| \mathrm{P}(W_n \le x) - \Phi(x) \right| \\ &\quad \le \frac{C \mathrm{E} X_1^4}{\sigma_n^2} \left(\sqrt{\sum_i \Big(\sum_j a_{ij}^2 \Big)^2} + \sqrt{\sum_{i,j} \Big(\sum_k a_{ik} a_{jk} \Big)^2} \right), \end{aligned} \tag{3.1}$$

where C is an absolute constant.

It is easy to check that

$$\sum_{i,j} \Big(\sum_k a_{ik} a_{jk} \Big)^2 = \operatorname{Tr}(A^4)$$

and

$$\sum_i \Big(\sum_j a_{ij}^2 \Big)^2 \le \max_{1 \le i \le n} \sum_j a_{ij}^2 \sigma_n^2 \le \lambda_1^2 \sigma_n^2,$$

which means that the first term in (3.1) is less than the bound λ_1/σ_n given in Theorem 1 of Götze and Tikhomirov [19]. However, comparing it with the L_1 bound given in Chatterjee [5], we conjecture that the bound in (3.1) can be improved to

$$\sup_{x\in\mathbb{R}}\bigl|\mathrm{P}(W_n \le x) - \Phi(x)\bigr| \le C\biggl(\frac{1}{\sigma_n^4}\sum_i\Bigl(\sum_j a_{ij}^2\Bigr)^2 + \frac{1}{\sigma_n^2}\sqrt{\sum_{i,j}\Bigl(\sum_k a_{ik}a_{jk}\Bigr)^2}\biggr).$$

3.2. *General Curie–Weiss model.* The Curie–Weiss model has been extensively discussed in the statistical physics field. The asymptotic behavior for the Curie–Weiss model was studied by Ellis and Newman [15–17]. Recently, Stein's method has been used to obtain the convergence rate of the Curie–Weiss model. For example, Chatterjee and Shao [10] used exchangeable pairs to get a Berry–Esseen bound at the critical temperature of the simplest Curie–Weiss model, where the magnetization was valued on $\{-1, 1\}$ with equal probability; and Chen, Fang and Shao [11] and Shao, Zhang and Zhang [26] established the Cramér type moderate deviation result for noncritical and critical temperature, respectively. More generally, when the magnetization was distributed as a measure ρ with a finite support, Chatterjee and Dey [7] obtained an exponential probability inequality. In this subsection, we apply Theorem 2.1 to establish a Berry–Esseen bound for the general Curie–Weiss model.

Let ρ be a probability measure satisfying

$$\int_{-\infty}^{\infty} x\,d\rho(x) = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} x^2\,d\rho(x) = 1. \tag{3.2}$$

ρ is said to be type k (an integer) with strength λ_ρ if

$$\int_{-\infty}^{\infty} x^j\,d\Phi(x) - \int_{-\infty}^{\infty} x^j\,d\rho(x) = \begin{cases} 0 & \text{for } j = 0, 1, \dots, 2k-1,\\ \lambda_\rho > 0 & \text{for } j = 2k,\end{cases}$$

where $\Phi(x)$ is the standard normal distribution function.

We define the Curie–Weiss model as follows. For a given measure ρ, let $(X_1, \dots, X_n)$ have the joint p.d.f.

$$dP_{n,\beta}(\mathbf{x}) = \frac{1}{Z_n}\exp\biggl(\frac{\beta(x_1 + \cdots + x_n)^2}{2n}\biggr)\prod_{i=1}^n d\rho(x_i), \tag{3.3}$$

where $\mathbf{x} = (x_1, \dots, x_n)$, $0 < \beta \le 1$ and Z_n is the normalizing constant.

Let ξ be a random variable with probability measure ρ. Moreover, assume that

(i) for $0 < \beta < 1$, there exists a constant $b > \beta$ such that

$$\mathrm{E}e^{t\xi} \le e^{\frac{t^2}{2b}} \qquad \text{for } -\infty < t < \infty; \tag{3.4}$$

(ii) for $\beta = 1$, there exist constants $b_0 > 0$, $b_1 > 0$ and $b_2 > 1$ such that

$$\mathrm{E}e^{t\xi} \le \begin{cases} \exp\bigl(t^2/2 - b_1 t^{2k}\bigr), & |t| \le b_0,\\ \exp\Bigl(\dfrac{t^2}{2b_2}\Bigr), & |t| > b_0.\end{cases} \tag{3.5}$$

Let $S_n = X_1 + \cdots + X_n$. Ellis and Newman [16], [17] showed that:

(i) if $0 < \beta < 1$, then $n^{-1/2} S_n$ converges to a normal distribution $\mathcal{N}(0, (1 - \beta)^{-1})$; and

(ii) if $\beta = 1$, and ρ is of type k, then $n^{-1+\frac{1}{2k}} S_n$ converges to a nonnormal distribution with p.d.f.

$$p(y) = c_1 e^{-c_2 y^{2k}},$$

where $c_2 > 0$ and c_1 is the normalizing constant.

The following theorem provides the rate of convergence.

THEOREM 3.2. *Let $(X_1, \ldots, X_n)$ follow the joint p.d.f.* (3.3), *where ρ satisfies* (3.2):

(i) *If $0 < \beta < 1$ and* (3.4) *is satisfied, then for $W_n = n^{-1/2} S_n$, we have*

$$\sup_{z \in \mathbb{R}} \bigl|P(W_n \le z) - P(Y_1 \le z)\bigr| \le C n^{-1/2}, \tag{3.6}$$

where $Y_1 \sim \mathcal{N}(0, \frac{1}{1-\beta})$ and C is a constant depending on b and β.

(ii) *If $\beta = 1$, ρ is of type k and* (3.5) *is satisfied, then for $W_n = n^{-1+\frac{1}{2k}} S_n$, we have*

$$\sup_{z \in \mathbb{R}} \bigl|P(W_n \le z) - P(Y_k \le z)\bigr| \le C n^{-\frac{1}{2k}}, \tag{3.7}$$

where C is a constant depending on b_0, b_1, b_2 and k; the density function of Y_k is given by

$$p(y) = c_1 e^{-c_2 y^{2k}}, \qquad c_2 = \frac{H^{(2k)}(0)}{(2k)!};$$

and c_1 is the normalizing constant and $H(s) = s^2/2 - \ln(\int_{-\infty}^{\infty} \exp(sx)\, d\rho(x))$.

3.3. *Mean field Heisenberg model.* The Heisenberg model is a statistical model for the phenomena of ferromagnetism and antiferromagnetism in the study of magnetism theory.

Let G_n be a finite complete graph with n vertices. At each site of the graph is a spin in $\mathbb{S}^2$, so the state space is $\Omega_n = (\mathbb{S}^2)^n$ with P_n the n-fold product of the uniform probability measure on $\mathbb{S}^2$. The mean field Hamiltonian energy of the Heisenberg model $H_n : \Omega_n \mapsto \mathbb{R}$ is

$$H_n(\sigma) = -\frac{1}{2n} \sum_{1 \le i, j \le n} \langle \sigma_i, \sigma_j \rangle,$$

where $\langle\cdot,\cdot\rangle$ is the inner product in $\mathbb{R}^3$. The Gibbs measure $P_{n,\beta}$ is given by the density function

$$dP_{n,\beta} = \frac{1}{Z_{n,\beta}} \exp\biggl(\frac{\beta}{2n} \sum_{1\le i,j\le n} \langle \sigma_i, \sigma_j\rangle\biggr) = \frac{1}{Z_{n,\beta}} \exp\bigl(-\beta H_n(\sigma)\bigr),$$

where $Z_{n,\beta} = \int_{\Omega_n} \exp(-\beta H_n(\sigma))\, dP_n$.

Consider the random variable

$$W_n = \sqrt{n}\biggl(\frac{\beta^2}{n^2\kappa^2}\biggl|\sum_{j=1}^n \sigma_j\biggr|^2 - 1\biggr), \tag{3.8}$$

where $|\cdot|$ is the Euclidean norm in $\mathbb{R}^3$ and κ is the solution to the equation

$$x/\beta = \bigl(\coth(x) - 1/x\bigr). \tag{3.9}$$

Let $\psi(x) = \coth(x) - 1/x$ and

$$B^2 = \frac{4\beta^2}{(1-\beta\psi'(\kappa))\kappa^2}\biggl(\frac{1}{\kappa^2} - \frac{1}{\sinh^2(\kappa)}\biggr). \tag{3.10}$$

Kirkpatrick and Meckes [20] showed that when $\beta > 3$, W_n/B converges to a standard normal distribution with an L_1 bound $O(\log(n) n^{-1/4})$. They also showed that when $\beta = 3$, the random variable $T_n = c_3 n^{-3/2} |\sum_j \sigma_j|^2$, where c_3 is a constant such that the variance of T_n is 1, converges in distribution to Y with the density function

$$p(y) = \begin{cases} Cy^5 e^{-3y^2/(5c_3)}, & y \ge 0, \\ 0, & y < 0, \end{cases}$$

where C is the normalizing constant.

The following theorem gives a Berry–Esseen bound for the case $\beta > 3$. The case $\beta = 3$ will be studied in another paper.

THEOREM 3.3. *Let W_n be the random variable defined as in* (3.8) *and B as in* (3.10) *with $\beta > 3$. Then we have*

$$\sup_{z\in\mathbb{R}} \bigl|\mathrm{P}(W_n/B \le z) - \Phi(z)\bigr| \le c_\beta n^{-1/2}, \tag{3.11}$$

where c_β is a constant depending on β.

3.4. *Counting monochromatic edges in uniformly colored graphs.* The study of monochromatic and heterochromatic subgraphs of an edge-colored graph dates back to the 1960s, and the last two decades has witnessed a significant development in the study of normal and Poisson approximation.

Barbour, Holst and Janson [2] used Stein's method to show that the number of monochromatic edges for the complete graph converges to a Poisson distribution. Arratia, Goldstein and Gordon [1] applied Stein's method to prove a Poisson approximation theorem for the number of monochromatic cliques in a uniform coloring of the complete graph. We refer to Chatterjee, Diaconis and Meckes [8] and Cerquetti and Fortini [3] for other related results.

In this subsection, we consider normal approximation for the counting of monochromatic edges in uniformly colored graphs. Let $G = \{V(G), E(G)\}$ be a simple undirected graph, where $V(G) = \{v_1, \dots, v_n\}$ is the vertex set and $E(G)$ is the edge set. For $1 \le i \le n$, let

$$A_i = \{1 \le j \le n, j \ne i, (v_i, v_j) \in E(G)\}$$

be the neighborhood of index i and $d_i = \#(A_i)$ be the number of edges connected to v_i. Denote the total number of edges in G by m_n, which is equal to $\sum_{i=1}^n d_i/2$. Each vertex is colored independently and uniformly with $c_n \ge 2$ colors, denoted by ξ_i the color of v_i. Let Y_n be the number of monochromatic edges in G_n. Rinott and Rotar [23] proved the central limit theorem for Y_n while Fang [18] obtained the Wasserstein distance with an order of $\sqrt{c_n/m_n} + c_n^{-1/2}$. The following theorem provides a Berry–Esseen bound.

THEOREM 3.4. *Let*

$$W_n = \frac{1}{2} \sum_{i=1}^n \sum_{j \in A_i} \frac{\mathbb{1}_{\{\xi_i = \xi_j\}} - \frac{1}{c_n}}{\sqrt{\frac{m_n}{c_n}(1 - \frac{1}{c_n})}}.$$

Then

$$\sup_{z \in \mathbb{R}} \bigl|\mathrm{P}(W_n \in z) - \Phi(z)\bigr| \le C\Bigl(\sqrt{1/c_n} + \sqrt{d_n^*/m_n} + \sqrt{c_n/m_n}\Bigr),$$

where C is an absolute constant and $d_n^* = \max\{d_i, 1 \le i \le n\}$.

4. Proof of main results.

As the normal approximation is a special case of the nonnormal approximation, we prove Theorem 2.2 only. The only difference for the normal approximation is that the Stein's solution can be bounded by 1 instead of $\sqrt{2\pi}$.

Let Y be the random variable with the p.d.f. $p(y)$ defined in (2.3). For a given z, let $f := f_z$ be the solution to the following Stein equation:

$$f'(w) - g(w)f(w) = \mathbb{1}_{\{w \le z\}} - F(z), \qquad z \in (a, b), \tag{4.1}$$

where F is the distribution function of Y. It is known (see, e.g., Chatterjee and Shao [10]) that

$$f_z(w) = \begin{cases} \dfrac{F(w)(1 - F(z))}{p(w)}, & w \le z, \\ \dfrac{F(z)(1 - F(w))}{p(w)}, & w > z. \end{cases} \tag{4.2}$$

We first prove some basic properties of f_z.

LEMMA 4.1. *Suppose that conditions* (A1)–(A3) *are satisfied. Then*

$$0 \le f_z(w) \le 1/c_1, \tag{4.3}$$

$$\|f_z'\| \le 1, \tag{4.4}$$

$$\|gf_z\| \le 1 \tag{4.5}$$

and

$$g(w) f_z(w) \text{ is nondecreasing.} \tag{4.6}$$

We remark that when $g(w) = w$, that is, for the normal approximation, it is known that $0 \le f_z(w) \le 1$ (see, e.g., Lemma 2.3 in Chen, Goldstein and Shao [12]).

PROOF. Without loss of generality, we assume that $a < 0 < b$ and $w_0 = 0$; thus, $p(0) = c_1$. For $w \le z$, define $H_z(w) = F(w)(1 - F(z)) - p(w)/c_1$. To prove (4.3), noting that $f_z(w) \ge 0$, it suffices to show that $\sup_{a<w<b} H_z(w) \le 0$. As $g(w)$ is nondecreasing, by the fact that $H_z'(w) = p(w)(1 - F(z) + g(w)/c_1)$,

$$\sup_{a<w\le z} H_z(w) = \max\{H_z(a), H_z(z)\}.$$

Clearly, $H_z(a) = -p(a)/c_1 \le 0$. Now we prove $\sup_{a<z<b} H_z(z) \le 0$. If $z \le 0$, define $H_1(z) = F(z) - p(z)/c_1$, and thus $H_1'(z) = p(z)(1 + g(z)/c_1)$. Note that $g(z) \le 0$ and $g(\cdot)$ is nondecreasing, then,

$$\sup_{a<z\le 0} H_z(z) \le \sup_{a<z\le 0} H_1(z) \le \max\{H_1(a), H_1(0)\} \le 0.$$

Using a similar argument, we also have $\sup_{0\le z<b} H_z(z) \le 0$. Therefore, $\sup_{a<z<b} H_z(z) \le 0$. This proves $\sup_{a<w\le z} f_z(w) \le 1/c_1$. Similarly, we have $\sup_{z<w<b} f_z(w) \le 1/c_1$.

A similar argument can be made for $w > z$. This completes the proof of (4.3).

We next show that gf_z is nondecreasing. For $w \le z$, by (4.2),

$$g(w) f_z(w) = \frac{g(w)F(w)(1 - F(z))}{p(w)},$$

and thus,

$$\big(g(w) f_z(w)\big)' = \big(1 - F(z)\big)\big(g(w) + \big(g'(w) + g^2(w)\big)F(w)/p(w)\big).$$

Let $\tau(w) = \frac{g(w)e^{-G(w)}}{g'(w)+g^2(w)}$. Then, by (A2),

$$-\tau'(w)e^{G(w)} = 1 - \left(\frac{2(g'(w))^2 - g''(w)g(w)}{(g'(w) + g^2(w))^2}\right) \le 1.$$

Hence,

$$e^{-G(w)} + \tau'(w) \ge 0$$

and

$$0 \le \int_a^w \big(\tau'(t) + e^{-G(t)}\big)\,dt = \tau(w) + \frac{1}{c_1}F(w) - \lim_{y\downarrow a}\tau(y).$$

By condition (A3), $\lim_{y\downarrow a}\tau(y) = 0$, and hence $\tau(w) + \frac{1}{c_1}F(w) \ge 0$. This proves that $(g(w)f_z(w))' \ge 0$ or $g(w)f_z(w)$ is nondecreasing for $w \le z$. Similarly, one can prove that $g(w)f_z(w)$ is nondecreasing for $w \ge z$. This proves (4.6).

To prove (4.5), by (A1), we have for $w \ge \max(z, 0)$,

$$\begin{aligned} g(w)f_z(w) &= \frac{F(z)g(w)\int_w^b p(t)\,dt}{p(w)} \\ &\le \frac{F(z)\int_w^b e^{-G(t)}g(t)\,dt}{e^{-G(w)}} \le F(z). \end{aligned}$$

Similarly, we have $g(w)f_z(w) \ge -(1 - F(z))$ for $w \le \min(0, z)$. Combining with (4.6) yields

$$F(z) - 1 \le g(w)f_z(w) \le F(z) \tag{4.7}$$

for all w. This proves (4.5).

Inequality (4.4) follows immediately from (4.1) and (4.7). □

PROOF OF THEOREM 2.2. Let $f = f_z$ be the solution to the Stein equation (4.1). Since (W, W') is an exchangeable pair, by (2.2), we have

$$\begin{aligned} 0 &= \mathrm{E}\big((W - W')(f(W) + f(W'))\big) \\ &= 2\mathrm{E}\big((W - W')f(W)\big) - \mathrm{E}\big((W - W')(f(W) - f(W'))\big) \\ &= 2\lambda\mathrm{E}\big(g(W)f(W)\big) + 2\lambda\mathrm{E}\big(Rf(W)\big) - \mathrm{E}\bigg(\Delta\int_{-\Delta}^0 f'(W+t)\,dt\bigg), \end{aligned}$$

and hence,

$$\mathrm{E}\big(g(W)f(W)\big) = \frac{1}{2\lambda}\mathrm{E}\bigg(\Delta\int_{-\Delta}^0 f'(W+t)\,dt\bigg) - \mathrm{E}\big(Rf(W)\big).$$

Thus,

$$\begin{aligned} &\mathrm{E}\big(f'(W) - g(W)f(W)\big) \\ &\quad = \mathrm{E}\bigg(f'(W)\Big(1 - \frac{1}{2\lambda}\mathrm{E}(\Delta^2 \mid W)\Big)\bigg) \\ &\qquad - \frac{1}{2\lambda}\mathrm{E}\bigg(\Delta\int_{-\Delta}^0 (f'(W+t) - f'(W))\,dt\bigg) + \mathrm{E}\big(Rf(W)\big). \end{aligned}$$

By (4.1), (4.3) and (4.4),

$$
\begin{aligned}
\left|\mathrm{P}(W \le z) - \mathrm{P}(Y \le z)\right| &= \left|\mathrm{E}\big(f'(W) - g(W)f(W)\big)\right| \\
&\le |I_1| + 2\mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}\big(\Delta^2 \mid W\big)\right| + \frac{1}{c_1}\mathrm{E}|R|,
\end{aligned} \tag{4.8}
$$

where

$$
I_1 = \frac{1}{2\lambda}\mathrm{E}\bigg(\Delta \int_{-\Delta}^{0} \big(f'(W+t) - f'(W)\big)\,dt\bigg).
$$

Recalling that f is the solution to (4.1), we have

$$
\begin{aligned}
I_1 &= \frac{1}{2\lambda}\mathrm{E}\bigg(\Delta \int_{-\Delta}^{0} \big(g(W+t)f(W+t) - g(W)f(W)\big)\,dt\bigg) \\
&\quad + \frac{1}{2\lambda}\mathrm{E}\bigg(\Delta \int_{-\Delta}^{0} (\mathbb{1}_{\{W+t\le z\}} - \mathbb{1}_{\{W\le z\}})\,dt\bigg).
\end{aligned} \tag{4.9}
$$

Noting that $g(w)f(w)$ is nondecreasing by Lemma 4.1 and that the indicator function $\mathbb{1}_{\{w\le z\}}$ is nonincreasing, we have

$$
\begin{aligned}
0 &\ge \int_{-\Delta}^{0} \big(g(W+t)f(W+t) - g(W)f(W)\big)\,dt \\
&\ge -\Delta\big(g(W)f(W) - g(W-\Delta)f(W-\Delta)\big)
\end{aligned}
$$

and

$$
0 \le \int_{-\Delta}^{0} (\mathbb{1}_{\{W+t\le z\}} - \mathbb{1}_{\{W\le z\}})\,dt \le \Delta(\mathbb{1}_{\{W-\Delta\le z\}} - \mathbb{1}_{\{W\le z\}}).
$$

Therefore,

$$
\begin{aligned}
I_1 &\le \frac{1}{2\lambda}E\big(-\Delta\mathbb{1}_{\{\Delta<0\}}\Delta\big(g(W)f(W) - g(W-\Delta)f(W-\Delta)\big)\big) \\
&\quad + \frac{1}{2\lambda}E\big(\Delta\mathbb{1}_{\{\Delta>0\}}\Delta(\mathbb{1}_{\{W-\Delta\le z\}} - \mathbb{1}_{\{W\le z\}})\big).
\end{aligned} \tag{4.10}
$$

Thus, for any $\Delta^* = \Delta^*(W, W') = \Delta^*(W', W) \ge |\Delta|$,

$$
\begin{aligned}
&\frac{1}{2\lambda}E\big(-\Delta\mathbb{1}_{\{\Delta<0\}}\Delta\big(g(W)f(W) - g(W-\Delta)f(W-\Delta)\big)\big) \\
&\quad \le \frac{1}{2\lambda}E\big(\Delta^*\mathbb{1}_{\{\Delta<0\}}\Delta\big(g(W)f(W) - g(W')f(W')\big)\big) \\
&\quad = \frac{1}{2\lambda}E\big(\Delta^*\Delta(\mathbb{1}_{\{\Delta<0\}} + \mathbb{1}_{\{\Delta>0\}})g(W)f(W)\big) \\
&\quad = \frac{1}{2\lambda}E\big(\Delta\Delta^* g(W)f(W)\big) \\
&\quad \le \frac{1}{2\lambda}\mathrm{E}\big|\mathrm{E}(\Delta\Delta^* \mid W)\big|,
\end{aligned} \tag{4.11}
$$

where $\mathrm{E}(\Delta^*\Delta\mathbb{1}_{\{\Delta<0\}}g(W')f(W')) = -\mathrm{E}(\Delta^*\Delta\mathbb{1}_{\{\Delta>0\}}g(W)f(W))$ because of the exchangeability of W and W' and $|g(w)f(w)| \le 1$ for all $w \in \mathbb{R}$. Similarly, we have

$$\frac{1}{2\lambda}E(\Delta\mathbb{1}_{\{\Delta>0\}}\Delta(\mathbb{1}_{\{W-\Delta\le z\}} - \mathbb{1}_{\{W\le z\}}) \le \frac{1}{2\lambda}\mathrm{E}\big|\mathrm{E}(\Delta\Delta^* \mid W)\big|. \tag{4.12}$$

Combining (4.10), (4.11) and (4.12) yields

$$I_1 \le \frac{1}{\lambda}\mathrm{E}\big|\mathrm{E}(\Delta\Delta^* \mid W)\big|. \tag{4.13}$$

Following the same argument, we also have

$$I_1 \ge -\frac{1}{\lambda}\mathrm{E}\big|\mathrm{E}(\Delta\Delta^* \mid W)\big|. \tag{4.14}$$

This proves (2.4), by (4.8), (4.13) and (4.14). □

5. Proofs of Theorems 3.1–3.4. In this section, we give proofs for the theorems in Section 3. The construction of an exchangeable pair is described as follows.

Let $\eta_1, \dots, \eta_n$ be a sequence of random variables and $W = h(\eta_1, \dots, \eta_n)$. For each $1 \le i \le n$, let η_i' have the conditional distribution of η_i given $\{\eta_j, 1 \le j \le n, j \ne i\}$, also, η_i' is conditionally independent of η_i given $\{\eta_j, 1 \le j \le n, j \ne i\}$. Let I be a random index uniformly distributed over $\{1, \dots, n\}$ independent of $\{\eta_i, \eta_i', 1 \le i \le n\}$. Set

$$W' = h(\eta_1, \dots, \eta_{I-1}, \eta_I', \eta_{I+1}, \dots, \eta_n).$$

Then (W, W') is an exchangeable pair. In particular, when η_i, $1 \le i \le n$ are independent, one can let $\{\eta_i', 1 \le i \le n\}$ be an independent copy of $\{\eta_i, 1 \le i \le n\}$. This sampling procedure is also called the Gibbs' sampler.

5.1. *Proof of Theorem* 3.1. Let $\mathscr{X} = \sigma(X_1, \dots, X_n)$, and $(X_1', X_2', \dots, X_n')$ be an independent copy of $(X_1, X_2, \dots, X_n)$. Let I be a random index uniformly distributed over $\{1, \dots, n\}$ independent of any other random variable. Write $W_n = h(X_1, \dots, X_n)$ and define $W_n' = h(X_1, \dots, X_I', \dots, X_n)$. Then (W_n, W_n') is an exchangeable pair. It is easy to see that

$$\Delta = W_n - W_n' = \frac{2}{\sigma_n}\sum_{j\ne I} a_{jI}X_j\big(X_I - X_I'\big)$$

and

$$\begin{aligned}\mathrm{E}(\Delta \mid \mathscr{X}) &= \frac{2}{\sigma_n}\sum_{i=1}^{n}\sum_{j\ne i}\mathrm{E}\big(a_{ji}X_j\big(X_i - X_i'\big) \mid \mathscr{X}\big)\\ &= \frac{2}{n}W_n.\end{aligned}$$

As such, condition (2.1) holds with $\lambda = 2/n$ and $R = 0$. Also,

$$\begin{aligned}\mathrm{E}(\Delta^2 \mid \mathscr{X}) &= \frac{4}{n\sigma_n^2}\sum_{i=1}^{n}\mathrm{E}\biggl(\biggl(\sum_{j\neq i} a_{ji}X_j(X_i - X_i')\biggr)^2 \biggm| \mathscr{X}\biggr)\\ &= \frac{4}{n\sigma_n^2}\sum_{i=1}^{n}(X_i^2+1)\biggl(\sum_{j=1}^{n} a_{ij}X_j\biggr)^2\end{aligned}$$

and

$$\frac{1}{2\lambda}\mathrm{E}(\Delta^2 \mid \mathscr{X}) = \frac{1}{\sigma_n^2}\sum_{i=1}^{n}(X_i^2+1)\biggl(\sum_{j=1}^{n} a_{ij}X_j\biggr)^2.$$

Note that by the assumptions $\sigma_n^2 = 2\sum_{i,j} a_{ij}^2$ and $a_{ii} = 0$,

$$\mathrm{E}\biggl(\frac{1}{2\lambda}\mathrm{E}(\Delta^2 \mid \mathscr{X})\biggr) = 1.$$

Then

$$\mathrm{E}\biggl|1 - \frac{1}{2\lambda}\mathrm{E}(\Delta^2 \mid W_n)\biggr|^2 \le \mathrm{Var}\biggl(\frac{1}{\sigma_n^2}\sum_{i=1}^{n}(X_i^2+1)\biggl(\sum_{j=1}^{n} a_{ij}X_j\biggr)^2\biggr).$$

Observe that

$$\begin{aligned}&\mathrm{Var}\biggl(\sum_{i=1}^{n}(X_i^2+1)\biggl(\sum_{j=1}^{n} a_{ij}X_j\biggr)^2\biggr)\\ &\quad = \sum_{i=1}^{n}\mathrm{Var}\biggl((X_i^2+1)\biggl(\sum_{j=1}^{n} a_{ij}X_j\biggr)^2\biggr)\\ &\qquad + \sum_{i\neq i'}\mathrm{Cov}\biggl((X_i^2+1)\biggl(\sum_{j=1}^{n} a_{ij}X_j\biggr)^2, (X_{i'}^2+1)\biggl(\sum_{k=1}^{n} a_{i'k}X_k\biggr)^2\biggr).\end{aligned}\tag{5.1}$$

For the first term, recalling that $a_{ii} = 0$ for all $1 \le i \le n$, we have

$$\begin{aligned}&\sum_{i=1}^{n}\mathrm{Var}\biggl((X_i^2+1)\biggl(\sum_{j=1}^{n} a_{ij}X_j\biggr)^2\biggr)\\ &\qquad \le \sum_{i=1}^{n}\mathrm{E}(X_i^2+1)^2\mathrm{E}\biggl(\sum_{j=1}^{n} a_{ij}X_j\biggr)^4\\ &\qquad \le C\sum_{i=1}^{n}(\mathrm{E}(X_1^4)+1)\mathrm{E}(X_1^4)\biggl(\sum_{j=1}^{n} a_{ij}^4 + \biggl(\sum_{j=1}^{n} a_{ij}^2\biggr)^2\biggr)\\ &\qquad \le C(\mathrm{E}(X_1^4))^2\sum_{i=1}^{n}\biggl(\sum_{j=1}^{n} a_{ij}^2\biggr)^2,\end{aligned}\tag{5.2}$$

where C is an absolute constant. To bound the second term of (5.1), for any $i \neq k$, define

$$M_i = (X_i^2 + 1)\biggl(\sum_{j=1}^n a_{ij} X_j\biggr)^2,$$

$$M_i^{(k)} = (X_i^2 + 1)\biggl(\sum_{j \neq k}^n a_{ij} X_j\biggr)^2.$$

For the second term of (5.1), for any $i \neq i'$, we have

$$\begin{aligned}
&\operatorname{Cov}\biggl((X_i^2+1)\biggl(\sum_{j=1}^n a_{ij}X_j\biggr)^2, (X_{i'}^2+1)\biggl(\sum_{k=1}^n a_{i'k}X_k\biggr)^2\biggr)\\
&\quad= \operatorname{Cov}(M_i, M_{i'})\\
&\quad= \operatorname{Cov}(M_i^{(i')}, M_{i'}) + \operatorname{Cov}(M_i, M_{i'}^{(i)})\\
&\qquad - \operatorname{Cov}(M_i^{(i')}, M_{i'}^{(i)}) + \operatorname{Cov}(M_i - M_i^{(i')}, M_{i'} - M_{i'}^{(i)}).
\end{aligned} \tag{5.3}$$

Given $\mathcal{F}_{ii'} := \sigma\{X_j, j \neq i, i'\}$, random variables $M_i^{(i')}$ and $M_{i'}^{(i)}$ are independent. Thus,

$$\begin{aligned}
&\operatorname{Cov}(M_i^{(i')}, M_{i'}^{(i)})\\
&\quad= \operatorname{Cov}\biggl(\mathrm{E}\biggl((X_i^2+1)\biggl(\sum_{j\neq i'}^n a_{ij}X_j\biggr)^2 \Bigm| \mathcal{F}_{ii'}\biggr),\\
&\qquad\qquad \mathrm{E}\biggl((X_{i'}^2+1)\biggl(\sum_{k\neq i}^n a_{i'j}X_k\biggr)^2 \Bigm| \mathcal{F}_{ii'}\biggr)\biggr)\\
&\quad= 4\operatorname{Cov}\biggl(\biggl(\sum_{j\neq i'}^n a_{ij}X_j\biggr)^2, \biggl(\sum_{k\neq i}^n a_{i'k}X_k\biggr)^2\biggr)\\
&\quad\leq C\sum_{j=1}^n a_{ij}^2 a_{i'j}^2 \mathrm{E}(X_1^4) + C\biggl(\sum_{k=1}^n a_{ik}a_{i'k}\biggr)^2.
\end{aligned}$$

Similar arguments hold for other terms of (5.3). Hence,

$$\begin{aligned}
&\sum_{i\neq i'}\operatorname{Cov}\biggl((X_i^2+1)\biggl(\sum_{j=1}^n a_{ij}X_j\biggr)^2, (X_{i'}^2+1)\biggl(\sum_{k=1}^n a_{i'k}X_k\biggr)^2\biggr)\\
&\quad\leq C\mathrm{E}(X_1^4)^2\biggl(\sum_{i=1}^n\biggl(\sum_{j=1}^n a_{ij}^2\biggr)^2 + \sum_{1\leq i,j\leq n}\biggl(\sum_{k=1}^n a_{ik}a_{jk}\biggr)^2\biggr).
\end{aligned} \tag{5.4}$$

It follows from (5.1), (5.2) and (5.4) that

$$\begin{aligned} &\mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}(\Delta^2 \mid W_n)\right| \\ &\quad \le C\sigma_n^{-2}\mathrm{E}(X_1^4)\left(\sqrt{\sum_i\Big(\sum_j a_{ij}^2\Big)^2} + \sqrt{\sum_{i,j}\Big(\sum_k a_{ik}a_{jk}\Big)^2}\right). \end{aligned} \tag{5.5}$$

Finally, it is sufficient to estimate the bound of $\mathrm{E}|\mathrm{E}(\Delta|\Delta| \mid W_n)|/\lambda$. In fact,

$$\begin{aligned} &\frac{1}{\lambda}\mathrm{E}(\Delta|\Delta| \mid \mathscr{X}) \\ &\quad = \frac{2}{\sigma_n^2}\sum_{i=1}^n \mathrm{E}\left(\Big(\sum_j a_{ij}X_j(X_i - X_i')\Big)\Big|\sum_j a_{ij}X_j(X_i - X_i')\Big| \;\Big|\; \mathscr{X}\right) \\ &\quad = \frac{2}{\sigma_n^2}\sum_{i=1}^n \Big(\sum_j a_{ij}X_j\Big)\Big|\sum_j a_{ij}X_j\Big| B_i, \end{aligned}$$

where $B_i = \mathrm{E}((X_i - X_i')|X_i - X_i'| \mid X_i)$.

For $i \ne i'$, define

$$\begin{aligned} K_i &= \Big(\sum_j a_{ij}X_j\Big)\Big|\sum_j a_{ij}X_j\Big| B_i, \\ K_i^{(i')} &= \Big(\sum_{j\ne i'} a_{ij}X_j\Big)\Big|\sum_{j\ne i'} a_{ij}X_j\Big| B_i \end{aligned}$$

and thus,

$$\mathrm{Var}\big((1/\lambda)\mathrm{E}(\Delta|\Delta| \mid \mathscr{X})\big) = \frac{4}{\sigma_n^4}\sum_{i=1}^n \mathrm{Var}(K_i) + \frac{4}{\sigma_n^4}\sum_{i\ne i'}\mathrm{Cov}(K_i, K_{i'}).$$

Similar to (5.2), we have

$$\sum_{i=1}^n \mathrm{Var}(K_i) \le C\big(\mathrm{E}(X_1^4)\big)^2 \sum_{i=1}^n\Big(\sum_{j=1}^n a_{ij}^2\Big)^2.$$

Recalling the definition of $\mathcal{F}_{ii'}$, given that $\mathcal{F}_{ii'}$, we have $K_i^{(i')}$ and $K_{i'}^{(i)}$ are conditionally independent, and thus

$$\mathrm{Cov}\big(K_i^{(i')}, K_{i'}^{(i)} \mid \mathcal{F}_{ii'}\big) = 0.$$

Moreover,

$$\mathrm{E}\big(K_i^{(i')} \mid \mathcal{F}_{ii'}\big) = \Big(\sum_{j\ne i'} a_{ij}\Big)\Big|\sum_{j\ne i'} a_{ij}X_j\Big|\mathrm{E}(B_i) = 0$$

because $\mathrm{E}(B_i)=0$. This proves $\mathrm{Cov}(K_i^{(i')}, K_{i'}^{(i)})=0$. Similarly, we have $\mathrm{Cov}(K_i^{(i')}, K_{i'})=0$ and $\mathrm{Cov}(K_i, K_{i'}^{(i)})=0$. Therefore,

$$\begin{aligned}\left|\mathrm{Cov}(K_i, K_{i'})\right| &= \mathrm{E}\left|\left(K_i - K_i^{(i')}\right)\left(K_{i'} - K_{i'}^{(i)}\right)\right| \\ &\le \frac{1}{2}\mathrm{E}\left(K_i - K_i^{(i')}\right)^2 + \frac{1}{2}\mathrm{E}\left(K_{i'} - K_{i'}^{(i)}\right)^2.\end{aligned}$$

Observe that

$$\left|K_i - K_i^{(i')}\right| \le |B_i|\left(2\left|a_{ii'}X_{i'}\sum_{j\ne i'} a_{ij}X_j\right| + a_{ii'}^2X_{i'}^2\right),$$

thus,

$$\begin{aligned}\mathrm{E}\left(K_i - K_i^{(i')}\right)^2 &\le C\mathrm{E}(B_i)^2\left(a_{ii'}^2\sum_j a_{ij}^2 + a_{ii'}^4\mathrm{E}(X_1^4)\right) \\ &\le C\left(\mathrm{E}(X_1^4)\right)^2\left(a_{ii'}^2\sum_j a_{ij}^2 + a_{ii'}^4\right).\end{aligned}$$

A similar result is true for $\mathrm{E}(K_{i'} - K_{i'}^{(i)})^2$. Combining the inequalities, we have

$$\mathrm{Var}\left(\mathrm{E}(\Delta|\Delta| \mid \mathscr{X})/\lambda\right) \le C\sigma_n^{-4}\left(\mathrm{E}(X_1^4)\right)^2\sum_{i=1}^n\left(\sum_{j=1}^n a_{ij}^2\right)^2.$$

By the Cauchy inequality, we have

$$\begin{aligned}&\frac{1}{\lambda}\mathrm{E}\left|\mathrm{E}(\Delta|\Delta| \mid W_n)\right| \\ &\quad\le C\sigma_n^{-2}\mathrm{E}(X_1^4)\left(\sqrt{\sum_i\left(\sum_j a_{ij}^2\right)^2} + \sqrt{\sum_{i,j}\left(\sum_k a_{ik}a_{jk}\right)^2}\right).\end{aligned} \tag{5.6}$$

This completes the proof of Theorem 3.1 by (5.5) and (5.6).

5.2. *Proof of Theorem* 3.2. Recall that $S_n=\sum_{i=1}^n X_i$. Let $\mathscr{X}=\sigma(X_1,\dots,X_n)$. We first construct an exchangeable pair (S_n, S_n') as follows. For each $1\le i\le n$, given $\{X_j, j\ne i\}$, let X_i' be conditionally independent of X_i with the same conditional distribution of X_i. Let I be a random index uniformly distributed over $\{1,\dots,n\}$ independent of any other random variable. Define $S_n' = S_n - X_I + X_I'$; then (S_n, S_n') is an exchangeable pair.

The proof of Theorem 3.2 is based on the following propositions. Let $\bar{X}=S_n/n$.

PROPOSITION 5.1. *Under the assumptions in Theorem* 3.2, *for* $\beta=1$, *we have*

$$\mathrm{E}\left(S_n - S_n' \mid \mathscr{X}\right) = \frac{H^{(2k)}(0)}{(2k-1)!}\bar{X}^{2k-1} + R_1, \tag{5.7}$$

where $\mathrm{E}|R_1|\le Cn^{-1}$ *with the constant* C *depending only on* b_0, b_1, b_3 *and* k.

For $0 < \beta < 1$, *we have*

$$\mathrm{E}\big(S_n - S_n' \mid \mathscr{X}\big) = (1-\beta)\bar{X} + R_2, \tag{5.8}$$

where $\mathrm{E}|R_2| \le Cn^{-1}$ *and* C *depends only on* β *and* b.

PROPOSITION 5.2. *Under the assumptions in Theorem* 3.2, *we have*

$$\mathrm{E}\big|\mathrm{E}\big((S_n - S_n')^2 \mid \mathscr{X}\big) - 2\big| \le Cn^{-1/2} \qquad \textit{for } 0 < \beta < 1, \tag{5.9}$$

and

$$\mathrm{E}\big|\mathrm{E}\big((S_n - S_n')^2 \mid \mathscr{X}\big) - 2\big| \le Cn^{-1/2k} \qquad \textit{for } \beta = 1. \tag{5.10}$$

PROPOSITION 5.3. *Under the assumptions in Theorem* 3.2, *we have for* $0 < \beta \le 1$,

$$\mathrm{E}\big|\mathrm{E}\big((S_n - S_n')|S_n - S_n'| \mid \mathscr{X}\big)\big| \le Cn^{-1/2}. \tag{5.11}$$

We now continue to prove Theorem 3.2.

(i) When $0 < \beta < 1$, define $W_n = S_n/\sqrt{n}$ and $W_n' = S_n'/\sqrt{n}$. Then (W_n, W_n') is an exchangeable pair, and by (5.8) in Proposition 5.1, we have

$$\mathrm{E}\big(W_n - W_n' \mid W_n\big) = \frac{1}{n}\big((1-\beta)W_n + \sqrt{n}R_2\big),$$

where $\mathrm{E}|R_2| \le Cn^{-1}$.

Moreover, taking $\lambda = 1/n$, by (5.9) and (5.11), we have

$$\mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}\big((W_n - W_n')^2 \mid W_n\big)\right| \le Cn^{-1/2}$$

and

$$\mathrm{E}\left|\frac{1}{2\lambda}\mathrm{E}\big((W_n - W_n')|W_n - W_n'| \mid W_n\big)\right| \le Cn^{-1/2}$$

for some constant C. This proves (3.6) by Theorem 2.2 with $g(w) = (1-\beta)w$.

(ii) When $\beta = 1$, define $W_n = n^{-1+\frac{1}{2k}}S_n$ and $W_n' = n^{-1+\frac{1}{2k}}S_n'$. Then (W_n, W_n') is an exchangeable pair, and by (5.7),

$$\mathrm{E}\big(W_n - W_n' \mid W_n\big) = n^{-2+1/k}\left(\frac{H^{2k}(0)}{(2k-1)!}W_n^{2k-1} + n^{-1+\frac{1}{2k}}R_1\right),$$

where $n^{-1+\frac{1}{2k}}\mathrm{E}|R_1| \le Cn^{-\frac{1}{2k}}$. Taking $\lambda = n^{-2+1/k}$ and by (5.10) and (5.11), we have

$$\mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}\big((W_n - W_n')^2 \mid W_n\big)\right| \le Cn^{-\frac{1}{2k}}$$

and

$$\mathrm{E}\left|\frac{1}{2\lambda}\mathrm{E}\big((W_n - W_n')|W_n - W_n'| \mid W_n\big)\right| \le Cn^{-1/2}.$$

This completes the proof of (3.7) by Theorem 2.2 with $g(w) = \frac{H^{2k}(0)}{(2k-1)!}w^{2k-1}$.

To prove Propositions 5.1 to 5.3, we need to prove some preliminary lemmas.

In what follows, we let $\xi, \xi_1, \xi_2, \dots$ be independent and identically distributed random variables with probability measure ρ satisfying (3.2), and (3.4) or (3.5).

LEMMA 5.1. *For any $z > 0$, under (3.4), we have*

$$\mathrm{P}\big(|\xi_1 + \cdots + \xi_n| > z\big) \le 2\exp\left(-\frac{bz^2}{2n}\right) \qquad \text{for } 0 < \beta < 1. \tag{5.12}$$

Under (3.5), and for $\beta = 1$,

$$\mathrm{P}\big(|\xi_1 + \cdots + \xi_n| > z\big) \le \begin{cases} 2\exp\left(-\dfrac{z^2}{2n} - \dfrac{b_1 z^{2k}}{n^{2k-1}}\right), & 0 < z \le b_0 n,\\ 2\exp\left(-\dfrac{b_2 z^2}{2n}\right), & z > b_0 n. \end{cases} \tag{5.13}$$

PROOF. (5.12) follows easily from (3.4) and Chebyshev's inequality.

As for (5.13), when $0 < z \le b_0 n$, set $t = z/n$. By the Chebyshev inequality, we have

$$\begin{aligned}\mathrm{P}(\xi_1 + \cdots + \xi_n > z) &\le e^{-tz}\mathrm{E}e^{t(\xi_1+\cdots+\xi_n)}\\ &= e^{-tz}\big(\mathrm{E}e^{t\xi}\big)^n\\ &\le e^{-tz}\exp\left(\frac{nt^2}{2} - nb_1t^{2k}\right)\\ &= \exp\left(-\frac{z^2}{2n} - \frac{b_1 z^{2k}}{n^{2k-1}}\right).\end{aligned}$$

Similarly,

$$\mathrm{P}(\xi_1 + \cdots + \xi_n < -z) \le \exp\left(-\frac{z^2}{2n} - \frac{b_1 z^{2k}}{n^{2k-1}}\right),$$

and hence

$$\mathrm{P}\big(|\xi_1 + \cdots + \xi_n| > z\big) \le 2\exp\left(-\frac{z^2}{2n} - \frac{b_1 z^{2k}}{n^{2k-1}}\right).$$

A similar argument for $z > b_0 n$ completes the proof of (5.13). □

LEMMA 5.2. *Under condition* (3.5) *and for* $\beta = 1$, *we have*

$$cn^{\frac{1}{2}-\frac{1}{2k}} \le \mathrm{E}\exp\left(\frac{1}{2n}(\xi_1 + \cdots + \xi_n)^2\right) \le Cn^{\frac{1}{2}-\frac{1}{2k}}, \tag{5.14}$$

where c *and* C *are constants such that* $0 < c < C < \infty$. *Under condition* (3.4), *for* $0 < \beta < 1$, *we have*

$$1 \le \mathrm{E}\exp\left(\frac{\beta}{2n}(\xi_1 + \cdots + \xi_n)^2\right) \le C, \tag{5.15}$$

where $C > 1$ *is a finite constant.*

PROOF. Noting that

$$e^{x^2/2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx - t^2/2}\, dt,$$

we have

$$\begin{aligned}
&\sqrt{2\pi}\mathrm{E}\exp\left(\frac{1}{2n}(\xi_1 + \cdots + \xi_n)^2\right)\\
&\quad = \int_{-\infty}^{\infty} \mathrm{E}\exp\left(\frac{t}{\sqrt{n}}(\xi_1 + \cdots + \xi_n) - \frac{t^2}{2}\right) dt\\
&\quad \le \int_{|t| \le b_0\sqrt{n}} e^{-b_1 t^{2k}/n^{k-1}}\, dt + \int_{|t| > b_0\sqrt{n}} e^{-\frac{t^2}{2}(1 - \frac{1}{b_2})}\, dt\\
&\quad \le Cn^{\frac{1}{2}-\frac{1}{2k}}
\end{aligned} \tag{5.16}$$

for some constant C.

For the lower bound of $\mathrm{E}e^{\frac{1}{2n}(\sum_{i=1}^n \xi_i)^2}$, as ρ is of type k with strength λ_ρ, then by the Taylor expansion, for $|t| \le b_0$,

$$\left|\frac{t^2}{2} - \log \mathrm{E}e^{t\xi}\right| \le C_\lambda t^{2k},$$

where $C_\lambda = \lambda_\rho + b_0 \sup_{|t| \le b_0} |H^{(2k+1)}(t)|$ is a constant. Thus, for $|t| \le b_0$,

$$\mathrm{E}e^{t\xi} \ge \exp\left(\frac{t^2}{2} - C_\lambda t^{2k}\right).$$

Similar to (5.16), we have

$$\begin{aligned}
\sqrt{2\pi}\mathrm{E}e^{\frac{1}{2n}(\sum_{i=1}^n \xi_i)^2} &\ge \mathrm{E}\int_{|t| \le b_0\sqrt{n}} e^{\frac{t}{\sqrt{n}}(\xi_1 + \cdots + \xi_n) - \frac{t^2}{2}}\, dt\\
&\ge cn^{\frac{1}{2}-\frac{1}{2k}}.
\end{aligned}$$

This proves (5.14).

Under condition (3.4) and similar to (5.16), we have

$$\mathrm{E}\exp\biggl(\frac{\beta}{2n}(\xi_1+\cdots+\xi_n)^2\biggr)\le C,$$

and by the Jensen inequality,

$$\begin{aligned}\mathrm{E}e^{\frac{\beta}{2n}(\sum_{i=1}^n\xi_i)^2}&\ge e^{\frac{\beta}{2n}\mathrm{E}((\xi_1+\cdots+\xi_n)^2)}\\&\ge 1.\end{aligned}$$

This completes the proof of (5.15). □

Let $X=(X_1,\dots,X_n)$ be a random vector following the Curie–Weiss distribution satisfying (3.3). We have the following inequalities.

LEMMA 5.3. *Under condition* (3.5), *we have*

$$\mathrm{E}\biggl(\frac{X_1+\cdots+X_n}{n^{1-\frac{1}{2k}}}\biggr)^{2k}\le C,\qquad \beta=1,\tag{5.17}$$

and under condition (3.4), *we have*

$$\mathrm{E}\biggl(\frac{X_1+\cdots+X_n}{\sqrt{n}}\biggr)^{2}\le C,\qquad 0<\beta<1.\tag{5.18}$$

PROOF. Let $M_n=\frac{1}{\sqrt{n}}(\xi_1+\cdots+\xi_n)$ and $Z_n=\mathrm{E}e^{\frac12 M_n^2}$. For $\beta=1$ and when (3.5) holds, by (3.3), we have

$$\begin{aligned}\mathrm{E}(S_n^{2k})&=\frac{n^k}{Z_n}\mathrm{E}\bigl(M_n^{2k}e^{\frac12 M_n^2}\bigr)\\&=\frac{n^k}{Z_n}\int_0^\infty\biggl(2kx^{2k-1}+\frac12x^{2k+1}\biggr)e^{\frac12x^2}\mathrm{P}\bigl(|M_n|\ge x\bigr)\,dx\\&=\frac{n^k}{Z_n}(I_1+I_2),\end{aligned}\tag{5.19}$$

where

$$\begin{aligned}I_1&=\int_0^{b_0\sqrt{n}}\biggl(2kx^{2k-1}+\frac12x^{2k+1}\biggr)e^{\frac12x^2}\mathrm{P}\bigl(|M_n|\ge x\bigr)\,dx,\\I_2&=\int_{b_0\sqrt{n}}^\infty\biggl(2kx^{2k-1}+\frac12x^{2k+1}\biggr)e^{\frac12x^2}\mathrm{P}\bigl(|M_n|\ge x\bigr)\,dx.\end{aligned}$$

For I_1, letting $D_n=[b_0\sqrt{n}]+1$, where $[a]$ is the integer part of a, we have

$$I_1\le\sum_{j=0}^{D_n}\int_j^{j+1}\biggl(2kx^{2k-1}+\frac12x^{2k+1}\biggr)e^{\frac12x^2}\mathrm{P}\bigl(|M_n|\ge x\bigr)\,dx$$

$$
\begin{aligned}
&\leq C\left(1+\sum_{j=1}^{D_n} j^{2k+1}\int_j^{j+1} e^{\frac{1}{2}x^2-jx+jx}\mathrm{P}\big(|M_n|\geq x\big)\,dx\right)\\
&\leq C\left(1+\sum_{j=1}^{D_n} j^{2k+1}e^{-\frac{j^2}{2}}\int_j^{j+1} e^{jx}\mathrm{P}\big(|M_n|\geq x\big)\,dx\right)\\
&\leq C\left(1+\sum_{j=1}^{D_n} j^{2k+1}e^{-\frac{j^2}{2}}\int_j^{j+1} e^{jx-\frac{x^2}{2}-\frac{b_1x^{2k}}{n^{k-1}}}\,dx\right) \qquad \text{by (5.13)}\\
&\leq C\left(1+\sum_{j=1}^{D_n} j^{2k+1}e^{-\frac{b_1j^{2k}}{n^{k-1}}}\,dx\right)\\
&\leq C\big(1+n^{(2k+1)(k-1)/2k}\big).
\end{aligned}
$$

A similar argument can be made for I_2. By (5.19) and (5.14), we have

$$
\mathrm{E}\big(S_n^{2k}\big)\leq Cn^{2k-1}. \tag{5.20}
$$

This completes the proof of (5.17). A similar argument holds for (5.18). This completes the proof of Lemma 5.3. □

LEMMA 5.4. *For $0<\beta\leq 1$, there exists a constant $b_3>\beta$ such that*

$$
\mathrm{E}e^{b_3\xi^2/2}\leq C. \tag{5.21}
$$

PROOF. When $0<\beta<1$, we choose b_3 such that $\beta<b_3<b$; then

$$
\begin{aligned}
\mathrm{E}e^{b_3\xi^2/2} &= \frac{1}{\sqrt{2\pi b_3}}\int_{-\infty}^{\infty}\mathrm{E}e^{t\xi-t^2/(2b_3)}\,dt\\
&\leq \frac{1}{\sqrt{2\pi b_3}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}(\frac{1}{b_3}-\frac{1}{b})}\,dt\\
&\leq C.
\end{aligned}
$$

When $\beta=1$, we choose b_3 such that $1<b_3<b_2$. Then

$$
\begin{aligned}
\mathrm{E}e^{b_3\xi^2/2} &= \frac{1}{\sqrt{2\pi b_3}}\int_{-\infty}^{\infty}\mathrm{E}e^{t\xi-t^2/2b_3}\,dt\\
&\leq \frac{1}{\sqrt{2\pi b_3}}\int_{|t|\leq b_0}\exp\left(\frac{t^2}{2}-b_1t^4-\frac{t^2}{2b_3}\right)dt\\
&\quad+\frac{1}{\sqrt{2\pi b_3}}\int_{|t|>b_0}\exp\left(\frac{t^2}{2b_2}-\frac{t^2}{2b_3}\right)dt\\
&\leq C.
\end{aligned}
$$

This proves (5.21). □

Let $\bar{X}_i = \frac{1}{n}(S_n - X_i)$.

LEMMA 5.5. *For* $0 < \beta \le 1$, *and for* $r \ge 1$, *we have*

$$\mathrm{E}\big(|X_i|^r \mid \bar{X}_i\big) \le Ce^{\beta \bar{X}_i^2}. \tag{5.22}$$

PROOF. Let ξ be a random variable with the probability measure ρ independent of $\bar{X}_i$. Then

$$\mathrm{E}\big(|X_i|^r \mid \bar{X}_i\big) = \frac{\mathrm{E}(|\xi|^r e^{\frac{\beta\xi^2}{2n} + \beta\bar{X}_i\xi} \mid \bar{X}_i)}{\mathrm{E}(e^{\frac{\beta\xi^2}{2n} + \beta\bar{X}_i\xi} \mid \bar{X}_i)}$$

and

$$\begin{aligned}
\mathrm{E}\big(e^{\frac{\beta\xi^2}{2n} + \beta\bar{X}_i\xi} \mid \bar{X}_i\big) &\ge \mathrm{E}\big(e^{\beta\bar{X}_i\xi} \mid \bar{X}_i\big) \\
&\ge e^{-\frac{\beta\bar{X}_i^2}{2}} \mathrm{E}\big(e^{-\beta\xi^2/2}\big) \\
&\ge e^{-\frac{\beta\bar{X}_i^2}{2}} e^{-\beta \mathrm{E}(\xi^2)/2} \\
&\ge e^{-\beta/2} e^{-\beta\bar{X}_i^2/2}.
\end{aligned} \tag{5.23}$$

By Lemma 5.4, given $t = b_3 y$, where b_3 depends on β, b and b_2, we have

$$\begin{aligned}
\mathrm{P}\big(|\xi| \ge y\big) &\le e^{-ty} \mathrm{E}\big(e^{t|\xi|}\big) \\
&\le e^{-ty} \mathrm{E}\big(e^{\frac{b_3\xi^2}{2} + \frac{t^2}{2b_3}}\big) \\
&\le Ce^{-b_3 y^2/2}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{E}\big(|\xi|^r e^{\frac{\beta\xi^2}{2n} + \frac{\beta\xi^2}{2}}\big) &\le \int_0^\infty (ry^{r-1} + 2\beta y^{r+1}) e^{\beta y^2(1+1/n)/2} \mathrm{P}\big(|\xi| \ge y\big)\,dy \\
&\le C\int_0^\infty (ry^{r-1} + 2\beta y^{r+1}) e^{\beta y^2(1+1/n)/2 - b_3 y^2/2}\,dy \\
&\le C,
\end{aligned}$$

and by the Cauchy inequality,

$$\begin{aligned}
\mathrm{E}\big(|\xi|^r e^{\frac{\beta\xi^2}{2n} + \beta\bar{X}_i\xi} \mid \bar{X}_i\big) &\le e^{\beta\bar{X}_i^2/2} \mathrm{E}\big(|\xi|^r e^{\frac{\beta\xi^2}{2n} + \frac{\beta\xi^2}{2}}\big) \\
&\le Ce^{\beta\bar{X}_i^2/2}.
\end{aligned} \tag{5.24}$$

This completes the proof of (5.22). $\square$

LEMMA 5.6. *If $0 < \beta < 1$ and (3.4) is satisfied, then for $r > 0$ and $\theta > 0$, we have*

$$\mathrm{E}\big(|\bar{X}_i|^r e^{\theta \bar{X}_i^2}\big) \le C n^{-r/2}. \tag{5.25}$$

If $\beta = 1$ and (3.5) is satisfied, then for $r \ge 0$ and $\theta > 0$, we have

$$\mathrm{E}\big(|\bar{X}_i|^r e^{\theta \bar{X}_i^2}\big) \le C n^{-\frac{r}{2k}}. \tag{5.26}$$

PROOF. Without loss of generality, assume $i = 1$. Observe that

$$\begin{aligned}\mathrm{E}\big(|\bar{X}_1|^r e^{\theta \bar{X}_1^2}\big) &= \frac{1}{n^r Z_n} \mathrm{E}|\xi_2 + \cdots + \xi_n|^r e^{\frac{\beta}{2n}(\xi_1 + \cdots + \xi_n)^2 + \frac{\theta}{n^2}(\xi_2 + \cdots + \xi_n)^2} \\ &\le \frac{1}{n^r Z_n} \mathrm{E} e^{\frac{\beta}{2}(1+1/n)\xi_1^2} \mathrm{E}|\xi_2 + \cdots + \xi_n|^r e^{(\frac{\beta}{2n} + \frac{\theta+\beta}{n^2})(\xi_2 + \cdots + \xi_n)^2}.\end{aligned}$$

When $0 < \beta < 1$ and (3.4) is satisfied, by (5.15), we have $Z_n \ge 1$. Also, similar to (5.12),

$$\mathrm{P}\big(|\xi_2 + \cdots + \xi_n| > y\big) \le 2 e^{-\frac{by^2}{2(n-1)}}.$$

Thus, for $r \ge 2$,

$$\begin{aligned}&\mathrm{E}|\xi_2 + \cdots + \xi_n|^r e^{(\frac{\beta}{2n} + \frac{\theta+\beta}{n^2})(\xi_2 + \cdots + \xi_n)^2} \\ &\quad \le C \int_0^\infty \big(r y^{r-1} + \big(\beta n^{-1} + 2(\theta + \beta) n^{-2}\big) y^{r+1}\big) e^{(\frac{\beta}{2n} + \frac{\theta+\beta}{n^2}) y^2 - \frac{b}{2(n-1)} y^2}\, dy \\ &\quad \le C n^{r/2}.\end{aligned}$$

This proves (5.25). Similarly, following the proof of (5.17), (5.26) holds for $r \ge 2$. When $r = 0$, similar to Lemma 5.2, we have

$$\mathrm{E}\big(e^{\theta \bar{X}_i^2}\big) \le C.$$

By the Cauchy inequality, (5.25) and (5.26) hold for $0 < r < 2$. This completes the proof of Lemma 5.6. $\square$

LEMMA 5.7. *For each $1 \le i < j \le n$, we have*

$$\begin{aligned}&\big|\mathrm{E}\big((X_i^2 - 1)(X_j^2 - 1)\big)\big| \\ &\quad \le \begin{cases} C n^{-1}, & 0 < \beta < 1, \text{ under (3.4)}, \\ C n^{-1/k}, & \beta = 1, \text{ under (3.5)}. \end{cases}\end{aligned} \tag{5.27}$$

PROOF. We consider $i = 1$, $j = 2$ only. Note that

$$\mathrm{E}\big((X_1^2 - 1)(X_2^2 - 1)\big) = \frac{1}{Z_n} \mathrm{E}(\xi_1^2 - 1)(\xi_2^2 - 1) \exp\bigg(\frac{\beta}{2n}(\xi_1 + \cdots + \xi_n)^2\bigg).$$

Set $m_{12} = \xi_3 + \cdots + \xi_n$. We first calculate the conditional expectation given $\xi_3, \dots, \xi_n$. In fact, for any s, we have

$$\begin{aligned}
&\mathrm{E}(\xi_1^2 - 1)(\xi_2^2 - 1)e^{\frac{\beta}{2n}(\xi_1+\xi_2)^2 + \frac{\beta}{n}(\xi_1+\xi_2)s} \\
&\quad = \frac{\sqrt{\beta}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{E}(\xi_1^2 - 1)(\xi_2^2 - 1) \\
&\qquad \times \exp\biggl(\frac{\beta t}{\sqrt{n}}(\xi_1 + \xi_2) + \frac{\beta s}{n}(\xi_1 + \xi_2) - \frac{\beta t^2}{2}\biggr)\,dt \\
&\quad = \frac{\sqrt{\beta}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \bigl(\mathrm{E}\bigl((\xi_1^2 - 1)e^{(\frac{\beta t}{\sqrt{n}} + \frac{\beta s}{n})\xi_1}\bigr)\bigr)^2 e^{-\beta t^2/2}\,dt.
\end{aligned}$$

Observe that

$$\begin{aligned}
&\biggl|\mathrm{E}(\xi_1^2 - 1)\exp\biggl(\frac{\beta t}{\sqrt{n}}\xi_1 + \frac{\beta s}{n}\xi_1\biggr)\biggr| \\
&\quad \le \biggl(\frac{\beta t}{\sqrt{n}} + \frac{\beta s}{n}\biggr)\mathrm{E}\bigl(|\xi_1|^3 + |\xi_1|\bigr)\exp\biggl(\frac{\beta t}{\sqrt{n}}|\xi_1| + \frac{\beta s}{n}|\xi_1|\biggr) \\
&\quad \le \biggl(\frac{\beta t}{\sqrt{n}} + \frac{\beta s}{n}\biggr)e^{\beta s^2/(2n^2) + \beta t^2/(2\sqrt{n})}\mathrm{E}\bigl(|\xi_1|^3 + |\xi_1|\bigr)e^{\frac{\beta \xi_1^2}{2}(1 + \frac{1}{\sqrt{n}})} \\
&\quad \le C\biggl(\frac{\beta t}{\sqrt{n}} + \frac{\beta s}{n}\biggr)e^{\beta s^2/(2n^2) + \beta t^2/(2\sqrt{n})}.
\end{aligned} \tag{5.28}$$

Therefore,

$$\begin{aligned}
&\bigl|\mathrm{E}(\xi_1^2 - 1)(\xi_2^2 - 1)e^{\frac{\beta}{2n}(\xi_1+\xi_2)^2 + \frac{\beta}{n}(\xi_1+\xi_2)s}\bigr| \\
&\quad \le C\int_{-\infty}^{\infty}\biggl(\frac{t^2}{n} + \frac{s^2}{n^2}\biggr)\exp\biggl(\frac{\beta t^2}{\sqrt{n}} + \frac{\beta s^2}{n^2} - \frac{\beta t^2}{2}\biggr)\,dt \\
&\quad \le C\biggl(\frac{1}{n} + \frac{s^2}{n^2}\biggr)e^{\beta s^2/n^2}.
\end{aligned}$$

Hence,

$$\bigl|\mathrm{E}\bigl((X_1^2 - 1)(X_2^2 - 1)\bigr)\bigr| \le C\mathrm{E}\biggl(\frac{1}{n} + \frac{m_{12}^2}{n^2}\biggr)e^{\beta m_{12}^2/n^2 + \beta m_{12}^2/(2n)}.$$

Similar to the proofs of Lemmas 5.3 and 5.6, for $0 < \beta < 1$,

$$\mathrm{E}\biggl(\frac{1}{n} + \frac{m_{12}^2}{n^2}\biggr)e^{\beta m_{12}^2/n^2 + \beta m_{12}^2/(2n)} \le Cn^{-1},$$

and for $\beta = 1$,

$$\mathrm{E}\biggl(\frac{1}{n} + \frac{m_{12}^2}{n^2}\biggr)e^{\beta m_{12}^2/n^2 + \beta m_{12}^2/(2n)} \le Cn^{-1/k}.$$

This completes the proof of (5.27). $\square$

For $1 \le i \le n$, let $\mathscr{X} = \sigma(X_1, \dots, X_n)$, and

$$Q_i = \mathrm{E}\big((X_i - X_i')|X_i - X_i'| \mid \mathscr{X}\big).$$

As defined at the beginning of this subsection, given $\{X_j, j \neq i\}$, X_i' and X_i are conditionally independent and have the same distribution.

LEMMA 5.8. *For $0 < \beta \le 1$, we have*

$$\mathrm{E}\big(Q_i^2\big) \le C, \tag{5.29}$$

$$\big|\mathrm{E}(Q_i Q_j)\big| \le C n^{-1}. \tag{5.30}$$

PROOF. By Lemmas 5.5 and 5.6,

$$\mathrm{E}\big(Q_i^2\big) \le \mathrm{E}\big(X_i - X_i'\big)^2 \le 4\mathrm{E}\big(X_i^2\big) \le C.$$

To prove (5.30), let

$$u(s,t) = (s-t)|s-t|.$$

Let $\xi, \xi_1, \dots, \xi_n$ be i.i.d. random variables with probability measure ρ, which are independent of $(X_1, \dots, X_n)$. We have

$$Q_i = \frac{\mathrm{E}(u(X_i, \xi)\exp(\frac{\beta\xi^2}{2n} + \beta\bar{X}_i\xi) \mid \mathscr{X})}{\mathrm{E}(\exp(\frac{\beta\xi^2}{2n} + \beta\bar{X}_i\xi) \mid \mathscr{X})}.$$

Without loss of generality, consider $i = 1,\ j = 2$. Define $\bar{X}_{12} = \frac{1}{n}(S_n - X_1 - X_2)$, and

$$Q_1' = \frac{\mathrm{E}(u(X_1, \xi)\exp(\beta\bar{X}_{12}\xi) \mid \mathscr{X})}{\mathrm{E}(\exp(\beta\bar{X}_{12}\xi) \mid \mathscr{X})},$$

$$Q_2' = \frac{\mathrm{E}(u(X_2, \xi)\exp(\beta\bar{X}_{12}\xi) \mid \mathscr{X})}{\mathrm{E}(\exp(\beta\bar{X}_{12}\xi) \mid \mathscr{X})}.$$

Again, let $m_{12} = (\xi_3 + \cdots + \xi_n)$. We have

$$\begin{aligned}
&\mathrm{E}\big(Q_1'Q_2'\big) \\
&\quad = \frac{1}{Z_n}\mathrm{E}\tilde{u}(\xi_1, m_{12})\tilde{u}(\xi_2, m_{12}) \\
&\qquad \times \exp\biggl(\frac{\beta}{2n}(\xi_1+\xi_2)^2 + \frac{\beta}{n}(\xi_1+\xi_2)m_{12} + \frac{\beta}{2n}m_{12}^2\biggr),
\end{aligned}$$

where

$$\tilde{u}(x,y) = \frac{\mathrm{E}(u(x,\xi)e^{\frac{\beta}{n}y\xi})}{\mathrm{E}(e^{\frac{\beta}{n}y\xi})}.$$

As $u(x, y)$ is antisymmetric, we have

$$\mathrm{E}\big(\tilde{u}(\xi_1, m_{12})\tilde{u}(\xi_2, m_{12})e^{\frac{\beta}{n}(\xi_1+\xi_2)m_{12}} \mid m_{12}\big) = 0.$$

Moreover,

$$\begin{aligned}&\mathrm{E}\big(|u(x,\xi)|e^{\frac{\beta y\xi}{n}}\big)\\&\quad\le C\big(x^2\mathrm{E}e^{\beta y\xi/n} + \mathrm{E}\big(\xi^2 e^{\beta y\xi/2}\big)\big)\\&\quad\le Ce^{Cy^2/n^2}(1+x^2+y^2/n^2).\end{aligned}$$

Similar to (5.23), $\mathrm{E}e^{\beta y\xi/n} \ge Ce^{-Cy^2/n^2}$, and thus,

$$\big|\tilde{u}(x, y)\big| \le Ce^{Cy^2/n^2}(1+x^2+y^2/n^2).$$

Therefore, similar to Lemmas 5.5 and 5.6,

$$\begin{aligned}&\big|\mathrm{E}(Q_1'Q_2')\big|\\&\quad\le \frac{\beta}{nZ_n}\mathrm{E}\big|\tilde{u}(\xi_1, m_{12})\tilde{u}(\xi_2, m_{12})\big|(\xi_1+\xi_2)^2 e^{\frac{\beta}{2n}(\xi_1+\cdots+\xi_n)^2}\\&\quad\le \frac{C}{nZ_n}\mathrm{E}\bigg(1+\xi_1^4+\xi_2^4+\frac{m_{12}^4}{n^4}\bigg)(\xi_1^2+\xi_2^2)e^{\frac{\beta}{2n}(\xi_1+\cdots+\xi_n)^2}\\&\quad\le \frac{C}{n}\mathrm{E}(1+\bar{X}_{12}^4)(1+X_1^6+X_2^6)e^{C\bar{X}_{12}^2}\\&\quad\le \frac{C}{n}.\end{aligned} \tag{5.31}$$

Next, we estimate $\mathrm{E}((Q_1 - Q_1')^2)$. Note that

$$\begin{aligned}&|Q_1 - Q_1'|\\&\quad\le \frac{|\mathrm{E}(u(X_1,\xi)e^{\beta\bar{X}_{12}\xi}(e^{\frac{\beta\xi^2}{2n}+\frac{\beta X_2}{n}}-1) \mid \mathscr{X})|}{\mathrm{E}\exp(\frac{\beta\xi^2}{2n}+\beta\bar{X}_1\xi)}\\&\qquad+\frac{\mathrm{E}(|u(X_1,\xi)|e^{\beta\bar{X}_{12}\xi} \mid \mathscr{X})\mathrm{E}(e^{\beta\bar{X}_{12}\xi}|e^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_{12}\xi}-1| \mid \mathscr{X})}{\mathrm{E}(e^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_{12}\xi} \mid \mathscr{X})\mathrm{E}(e^{\beta\bar{X}_{12}\xi} \mid \mathscr{X})}.\end{aligned}$$

Note also that $|u(s,t)| \le (s-t)^2$. Similar to Lemmas 5.5 and 5.6, we have

$$\mathrm{E}\big((Q_1 - Q_1')^2\big) \le Cn^{-2}. \tag{5.32}$$

Observe that

$$\begin{aligned}|\mathrm{E}(Q_1Q_2)| \le{}& |\mathrm{E}(Q_1'Q_2')| + |\mathrm{E}(Q_1(Q_2 - Q_2'))|\\&+|\mathrm{E}(Q_2(Q_1-Q_1'))| + |\mathrm{E}(Q_1-Q_1')(Q_2-Q_2')|.\end{aligned} \tag{5.33}$$

Then, by the Cauchy inequality and substituting (5.29), (5.31) and (5.32) into (5.33), we get the desired result. □

We are now ready to prove Propositions 5.1–5.3.

PROOF OF PROPOSITION 5.1. By the definition of S_n and S_n', we have

$$\begin{aligned}
\mathrm{E}(S_n - S_n' \mid \mathscr{X}) &= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i - X_i' \mid \mathscr{X}) \\
&= \bar{X} - \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i' \mid \mathscr{X}) \\
&= \bar{X} - \frac{1}{n}\sum_{i=1}^{n}\frac{\int_{-\infty}^{\infty} x e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)}.
\end{aligned}$$

Observe that for $0 < \beta \le 1$,

$$\frac{\int_{-\infty}^{\infty} x e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)} = h(\bar{X}_i) + r_{1i}, \tag{5.34}$$

where

$$h(s) = \frac{\int_{-\infty}^{\infty} x e^{\beta s x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\beta s x}\,d\rho(x)} \quad \text{and}$$

$$r_{1i} = \frac{\int_{-\infty}^{\infty} x e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)} - \frac{\int_{-\infty}^{\infty} x e^{\beta\bar{X}_i x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\beta\bar{X}_i x}\,d\rho(x)}.$$

We first give the bound of $\mathrm{E}|r_{1i}|$. Note that by (5.23) and (5.24),

$$\begin{aligned}
\mathrm{E}|r_{1i}| &\le \mathrm{E}\left|\frac{\int_{-\infty}^{\infty} x(e^{\frac{\beta x^2}{2n}}-1)e^{\beta\bar{X}_i x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\frac{\beta x^2}{2n}}e^{\beta\bar{X}_i x}\,d\rho(x)}\right| \\
&\quad + \mathrm{E}\left|\frac{\int_{-\infty}^{\infty}(e^{\frac{\beta x^2}{2n}}-1)e^{\beta\bar{X}_i x}\,d\rho(x)\int_{-\infty}^{\infty} x e^{\frac{\beta x^2}{2n}}e^{\beta\bar{X}_i x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\frac{\beta x^2}{2n}}e^{\beta\bar{X}_i x}\,d\rho(x)\int_{-\infty}^{\infty} e^{\beta\bar{X}_i x}\,d\rho(x)}\right| \\
&\le \frac{C}{n}\mathrm{E}\left|\frac{\int_{-\infty}^{\infty}|x|^3\exp(\frac{\beta x^2}{2n}+\beta\bar{X}_i x)\,d\rho(x)}{\int_{-\infty}^{\infty}\exp^{\beta\bar{X}_i x}\,d\rho(x)}\right| \qquad (5.35) \\
&\quad + \frac{C}{n}\mathrm{E}\left|\frac{\int_{-\infty}^{\infty}|x|^2 e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)\int_{-\infty}^{\infty}|x|e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)}{(\int_{-\infty}^{\infty} e^{\beta\bar{X}_i x}\,d\rho(x))^2}\right|
\end{aligned}$$

$$\leq Cn^{-1}\mathrm{E}e^{C\bar{X}_i^2}$$
$$\leq Cn^{-1}.$$

For $h(\bar{X}_i)$, we consider two cases.

Case 1. $\beta = 1$. As ρ is of type k, by the Taylor expansion,

$$h(\bar{X}_i) = \bar{X}_i + \frac{h^{(2k-1)}(0)}{(2k-1)!}\bar{X}_i^{2k-1} + \frac{1}{(2k-1)!}\int_0^{\bar{X}_i} h^{(2k)}(t)(\bar{X}_i - t)^{2k-1}\,dt$$

$$= \bar{X} - \frac{1}{n}X_i + \frac{h^{(2k-1)}(0)}{(2k-1)!}\bar{X}^{2k-1} + \frac{h^{(2k-1)}(0)}{(2k-1)!}(\bar{X}_i^{2k-1} - \bar{X}^{2k-1}) + \frac{1}{(2k-1)!}\int_0^{\bar{X}_i} h^{(2k)}(t)(\bar{X}_i - t)^{(2k-1)}\,dt. \tag{5.36}$$

Hence,

$$\mathrm{E}\big(S_n - S_n' \mid \mathscr{X}\big) = \frac{h^{2k-1}(0)}{(2k-1)!}\bar{X}^{2k-1} + R_1, \tag{5.37}$$

where

$$R_1 = -\frac{1}{n}\sum_{i=1}^n\biggl(h(\bar{X}_i) - \bar{X} - \frac{h^{(2k-1)}(0)}{(2k-1)!}\bar{X}^{2k-1}\biggr) - \frac{1}{n}\sum_{i=1}^n r_{1i},$$

and r_{1i} is given in (5.34) with $\beta = 1$.

Observe that by (5.36),

$$h(\bar{X}_i) - \bar{X} - \frac{h^{(2k-1)}(0)}{(2k-1)!}\bar{X}^{2k-1} = -\frac{1}{n}X_i + \frac{h^{(2k-1)}(0)}{(2k-1)!}(\bar{X}_i^{2k-1} - \bar{X}^{2k-1}) + \frac{1}{(2k-1)!}\int_0^{\bar{X}_i} h^{(2k)}(t)(\bar{X}_i - t)^{(2k-1)}\,dt. \tag{5.38}$$

For the first term of (5.38), it follows from Lemmas 5.5 and 5.6 that

$$\frac{1}{n}\mathrm{E}|X_i| \leq Cn^{-1}. \tag{5.39}$$

For the second term, by Lemmas 5.5 and 5.6 again,

$$\begin{aligned} &\frac{h^{(2k-1)}(0)}{(2k-1)!}\mathrm{E}|\bar{X}_i^{2k-1} - \bar{X}^{2k-1}| \\ &\quad\leq Cn^{-1}\mathrm{E}\big(|X_i|\big(|\bar{X}_i|^{2k-2} + \big(|X_i|/n\big)^{2k-2}\big)\big) \\ &\quad\leq Cn^{-1}\mathrm{E}\big(1 + |\bar{X}_i|^{2k-1}\big)e^{C|\bar{X}_i|^2} \\ &\quad\leq Cn^{-1}. \end{aligned} \tag{5.40}$$

To bound the last term, we first consider $h^{(2k)}(s)$. Recalling that

$$h(t) = \frac{\int_{-\infty}^{\infty} x e^{tx}\, d\rho(x)}{\int_{-\infty}^{\infty} e^{tx}\, d\rho(x)}$$

and observing that

$$\int_{-\infty}^{\infty} e^{tx}\, d\rho(x) \geq 1$$

and

$$\left|\frac{d^j}{dt^j}\int_{-\infty}^{\infty} e^{tx}\, d\rho(x)\right| = \left|\int_{-\infty}^{\infty} x^j e^{tx}\, d\rho(x)\right| \leq \int_{-\infty}^{\infty} (1+|x|^{2k+1}) e^{tx}\, d\rho(x)$$

for $j = 0, 1, \dots, 2k+1$, we have

$$\begin{aligned}|h^{(2k)}(t)| &\leq C\int_{-\infty}^{\infty} (1+|x|^{2k+1}) e^{tx}\, d\rho(x)\\ &\leq Ce^{t^2/2}.\end{aligned}$$

Thus, by (5.26),

$$\begin{aligned}&\frac{1}{(2k-1)!}\mathrm{E}\left|\int_0^{\bar{X}_i} h^{(2k)}(t)(\bar{X}_i - t)^{(2k-1)}\, dt\right| \\ &\quad \leq C\mathrm{E}\big(\bar{X}_i^{2k} e^{\bar{X}_i^2/2}\big) \leq Cn^{-1}.\end{aligned} \tag{5.41}$$

By (5.39), (5.40) and (5.41), (5.38) can be bounded by

$$\mathrm{E}\left|h(\bar{X}_i) - \bar{X} - \frac{h^{(2k-1)}(0)}{(2k-1)!}\bar{X}^{2k-1}\right| \leq Cn^{-1}. \tag{5.42}$$

Together with (5.34) and (5.35), we have

$$\mathrm{E}|R_1| \leq Cn^{-1}.$$

Case 2. For $\beta \in (0, 1)$, we have

$$\begin{aligned}h(\bar{X}_i) &= \beta\bar{X}_i + \int_0^{\bar{X}_i} h''(t)(\bar{X}_i - t)\, dt\\ &= \beta\bar{X} - \frac{\beta}{n}X_i + \int_0^{\bar{X}_i} h''(t)(\bar{X}_i - t)\, dt.\end{aligned}$$

Hence,

$$\mathrm{E}\big(S_n - S_n' \mid \mathscr{X}\big) = (1-\beta)\bar{X} + R_2,$$

where

$$R_2 = -\frac{1}{n}\sum_{i=1}^{n}\left(-\frac{\beta}{n}X_i + \int_0^{\bar{X}_i} h''(t)(\bar{X}_i - t)\, dt\right) - \frac{1}{n}\sum_{i=1}^{n} r_{1i}.$$

Similar to (5.42), we have

$$\mathrm{E}\left|-\frac{\beta}{n}X_i + \int_0^{\bar{X}_i} h''(t)(\bar{X}_i - t)\,dt\right| \le Cn^{-1}.$$

Together with (5.35), we have

$$\mathrm{E}|R_2| \le Cn^{-1}.$$

This completes the proof. □

PROOF OF PROPOSITION 5.2. Observe that

$$\begin{aligned}
\mathrm{E}((S_n - S_n')^2 \mid \mathscr{X}) &= \frac{1}{n}\sum_{i=1}^n \mathrm{E}\big(X_i^2 - 2X_iX_i' + (X_i')^2 \mid \mathscr{X}\big) \\
&= \frac{1}{n}\sum_{i=1}^n \Bigg(X_i^2 - \frac{2X_i \int_{-\infty}^{\infty} xe^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)} \\
&\qquad + \frac{\int_{-\infty}^{\infty} x^2 e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)}\Bigg) \\
&:= 2 + R_3 + R_4 + R_5,
\end{aligned}$$

where

$$R_3 = \frac{1}{n}\sum_{i=1}^n (X_i^2 - 1),$$

$$R_4 = -\frac{1}{n}\sum_{i=1}^n \frac{2X_i \int_{-\infty}^{\infty} xe^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)},$$

$$R_5 = \frac{1}{n}\sum_{i=1}^n \frac{\int_{-\infty}^{\infty} x^2 e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)}{\int_{-\infty}^{\infty} e^{\frac{\beta x^2}{2n}+\beta\bar{X}_i x}\,d\rho(x)} - 1.$$

By the Taylor expansion, and similar to the proof of $\mathrm{E}|R_1|$ and $\mathrm{E}|R_2|$, we have

$$\mathrm{E}|R_4| + \mathrm{E}|R_5| \le \begin{cases} Cn^{-1/2}, & 0 < \beta < 1, \\ Cn^{-\frac{1}{2k}}, & \beta = 1. \end{cases}$$

As for $\mathrm{E}|R_3|$, we have

$$\mathrm{E}(R_3^2) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{E}(X_i^2 - 1)^2 + \frac{1}{n^2}\sum_{i\ne j}\mathrm{E}(X_i^2 - 1)(X_j^2 - 1).$$

By Lemma 5.7, we have

$$\mathrm{E}(X_i^4) \le C, \qquad \left|\mathrm{E}(X_i^2 - 1)(X_j^2 - 1)\right| \le \begin{cases} Cn^{-1}, & 0 < \beta < 1, \\ Cn^{-1/k}, & \beta = 1. \end{cases}$$

Therefore,

$$\mathrm{E}|R_3| \le \begin{cases} Cn^{-1/2}, & 0 < \beta < 1, \\ Cn^{-\frac{1}{2k}}, & \beta = 1. \end{cases}$$

This proves (5.9) and (5.10). □

PROOF OF PROPOSITION 5.3. We have

$$\begin{aligned} &\mathrm{E}\big((S_n - S_n')|S_n - S_n'| \mid \mathscr{X}\big) \\ &\quad = \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\big((X_i - X_i')|X_i - X_i'| \mid \mathscr{X}\big). \end{aligned}$$

Then (5.11) follows from Lemma 5.8. □

5.3. *Proof of Theorem* 3.3. The Berry–Esseen bound (3.11) follows from Theorem 2.1 and Proposition 5.4 below.

PROPOSITION 5.4. *Let W_n be as defined in* (3.8) *and $\sigma' = \{\sigma_1', \dots, \sigma_n'\}$, where for each i, σ_i' is an independent copy of σ_i given $\{\sigma_j, j \ne i\}$. Let I be a random index independent of all others and uniformly distributed over $\{1, \dots, n\}$, and let $W_n' = \sqrt{n}(\frac{\beta^2}{n^2\kappa^2}|S_n'|^2 - 1)$, where $S_n' = \sum_{j=1}^n \sigma_j - \sigma_I + \sigma_I'$. Then (W_n, W_n') is an exchangeable pair and there exists a constant c_β depending on β only such that*

$$\mathrm{E}\big(W_n - W_n' \mid W_n\big) = \lambda(W_n - R_n) \quad \textit{and} \quad \mathrm{E}|R_n| \le c_\beta n^{-1/2}, \tag{5.43}$$

where $\lambda = \frac{1 - \beta\psi'(\kappa)}{n}$;

$$\mathrm{E}\left|B^2 - \frac{1}{2\lambda}\mathrm{E}\big((W_n - W_n')^2 \mid W_n\big)\right| \le c_\beta n^{-1/2}, \tag{5.44}$$

where B is defined in (3.10); *and*

$$\frac{1}{\lambda}\mathrm{E}\left|\mathrm{E}\big((W_n - W_n')|W_n - W_n'| \mid W_n\big)\right| \le c_\beta n^{-1/2}. \tag{5.45}$$

PROOF. Let $S_n = \sum_{i=1}^n \sigma_i$ and $\sigma^{(i)} = S_n - \sigma_i$. The proof is organized in the following three parts.

(i) *Proof of* (5.43). Let $\sigma = (\sigma_1, \dots, \sigma_n)$. As shown in Kirkpatrick and Meckes [20] [page 23, equation (12)], we have

$$\mathrm{E}\big(W_n - W_n' \mid \sigma\big) = \frac{2}{n} W_n + \frac{2}{\sqrt{n}} - \frac{2\beta}{n^{1/2}\kappa^2}(\beta|S_n|/n)\psi\big(\beta|S_n|/n\big) + R_1,$$

where $\psi(x) = \coth(x) - 1/x$ and $|R_1| \le Cn^{-3/2}$ for some constant C depending on β. The Taylor expansion yields

$$\big(\beta|S_n|/n\big)\psi\big(\beta|S_n|/n\big) = \kappa\psi(\kappa) + \big(\psi(\kappa) + \kappa\psi'(\kappa)\big)\bigg(\frac{\beta|S_n|}{n} - \kappa\bigg) + R_2,$$

where $|R_2| \le C(\beta|S_n|/n - \kappa)^2$ with C depending on β.

Moreover, by Kirkpatrick and Meckes [20] (page 25),

$$\frac{\beta|S_n|}{n} - \kappa = \frac{\kappa W_n}{2\sqrt{n}} + R_3,$$

where $|R_3| \le C|W_n|^2/n$. Recalling (3.9) and combining all of the preceding inequalities, we have

$$\mathrm{E}\big(W_n - W_n' \mid \sigma\big) = \frac{1 - \beta\psi'(\kappa)}{n}(W_n - R_n),$$

where $|R_n| \le CW_n^2/n^{1/2}$. It follows from Kirkpatrick and Meckes [20] (page 24) that there exists $\varepsilon_0 > 0$ such that for all $x \in (0, \varepsilon_0]$,

$$\mathrm{P}\bigg(\bigg|\frac{\beta|S_n|}{n} - \kappa\bigg| > x\bigg) \le e^{-K_\beta n x^2}$$

for some constant $K_\beta > 0$. Then

$$\begin{aligned} \mathrm{E}|\beta|S_n|/n - \kappa|^4 &\le 4\int_0^{\varepsilon_0} x^3 e^{-K_\beta n x^2}\,dx + C\mathrm{P}\bigg(\bigg|\frac{\beta|S_n|}{n} - \kappa\bigg| > \varepsilon_0\bigg) \\ &\le Cn^{-2} + Ce^{-K_\beta n \varepsilon_0} \qquad (5.46)\\ &\le Cn^{-2}. \end{aligned}$$

It follows from the definition of W_n that

$$\begin{aligned} \mathrm{E}|W_n|^2 &= n\mathrm{E}\bigg|\frac{\beta^2|S_n|^2}{n^2\kappa^2} - 1\bigg|^2 \\ &\le Cn\mathrm{E}\bigg|\frac{\beta|S_n|}{n\kappa} - 1\bigg|^2 \\ &\le C, \end{aligned}$$

where C depends on β. This proves (5.43).

(ii) *Proof of* (5.44). From Kirkpatrick and Meckes [20] [pages 25–27, equations (16) and (18)], we have

$$
\begin{aligned}
&\mathrm{E}\big((W_n - W_n')^2 \mid \sigma\big)\\
&\quad = \frac{4\beta^4}{n^4\kappa^4}\sum_{i=1}^n |\sigma^{(i)}|^2\big((1-2\psi(b_i)/b_i) - 2\psi(b_i)\cos\alpha_i + \cos^2\alpha_i\big)\\
&\quad = 2\lambda B^2 + \frac{4\beta^4}{n^4\kappa^4}\Biggl(\sum_{i=1}^n\Bigl(1-\frac{2}{\beta}\Bigr)\Bigl(|\sigma^{(i)}|^2 - \frac{(n-1)^2\kappa^2}{\beta^2}\Bigr)\\
&\qquad - \frac{2\kappa}{\beta}\sum_{i=1}^n\Bigl(|\sigma^{(i)}|^2\cos\alpha_i - \frac{n^2\kappa^3}{\beta^3}\Bigr)\\
&\qquad + \sum_{i=1}^n\Bigl(|\sigma^{(i)}|^2\cos^2\alpha_i - \Bigl(1-\frac{2}{\beta}\Bigr)\frac{(n-1)^2\kappa^2}{\beta^2}\Bigr)\Biggr)\\
&\qquad + \frac{4\beta^4}{n^4\kappa^4}\sum_{i=1}^n\Bigl(2|\sigma^{(i)}|^2\Bigl(\frac{\psi(b_i)}{b_i} - \frac{1}{\beta}\Bigr) - 2|\sigma^{(i)}|^2\cos\alpha_i\Bigl(\psi(b_i) - \frac{\kappa}{\beta}\Bigr)\Bigr),
\end{aligned}
$$

where $b_i = \beta|\sigma^{(i)}|/n$ and α_i is the angle between σ_i and $\sigma^{(i)}$. Therefore,

$$
\begin{aligned}
&\frac{1}{2\lambda}\mathrm{E}\big(\mathrm{E}\big((W_n - W_n')^2 \mid \sigma\big)\big) - B^2\\
&\quad = \frac{2\beta^4}{n^3\kappa^4(1-\beta\psi'(\kappa))}(R_4 + R_5 + R_6 + R_7),
\end{aligned}
\tag{5.47}
$$

where

$$
\begin{aligned}
R_4 &= \sum_{i=1}^n\Bigl(1-\frac{2}{\beta}\Bigr)\Bigl(|\sigma^{(i)}|^2 - \frac{(n-1)^2\kappa^2}{\beta^2}\Bigr),\\
R_5 &= \frac{2\kappa}{\beta}\sum_{i=1}^n\Bigl(|\sigma^{(i)}|^2\cos\alpha_i - \frac{n^2\kappa^3}{\beta^3}\Bigr),\\
R_6 &= \sum_{i=1}^n\Bigl(|\sigma^{(i)}|^2\cos^2\alpha_i - \Bigl(1-\frac{2}{\beta}\Bigr)\frac{(n-1)^2\kappa^2}{\beta^2}\Bigr),\\
R_7 &= \sum_{i=1}^n\Bigl(2|\sigma^{(i)}|^2\Bigl(\frac{\psi(b_i)}{b_i} - \frac{1}{\beta}\Bigr) - 2|\sigma^{(i)}|^2\cos\alpha_i\Bigl(\psi(b_i) - \frac{\kappa}{\beta}\Bigr)\Bigr).
\end{aligned}
$$

For R_4, note that $|\sigma^{(i)} - S_n| \le 1$; then, by (5.46),

$$
\mathrm{E}\Bigl|\frac{\beta|\sigma^{(i)}|}{n} - \kappa\Bigr|^4 \le 8\mathrm{E}\Bigl|\frac{\beta|S_n|}{n} - \kappa\Bigr|^4 + 8/n^4 \le Cn^{-2}.
\tag{5.48}
$$

Thus,

$$\begin{aligned}
\mathrm{E}|R_4| &\le \sum_{i=1}^{n} \mathrm{E}\left||\sigma^{(i)}|^2 - \frac{(n-1)^2\kappa^2}{\beta^2}\right|^2 \\
&\le Cn^2 \sum_{i=1}^{n} \mathrm{E}\left|\frac{\beta^2|\sigma^{(i)}|^2}{n^2} - \kappa^2\right|^2 \\
&\le Cn^2 \sum_{i=1}^{n} \mathrm{E}\left|\frac{\beta|\sigma^{(i)}|}{n} - \kappa\right| \\
&\le Cn^{5/2}.
\end{aligned} \tag{5.49}$$

For R_5, by Kirkpatrick and Meckes ([20], page 28) we have

$$\begin{aligned}
\mathrm{E}|R_5| &\le \mathrm{E}\left|\sum_{i=1}^{n} \frac{2\kappa}{\beta}\left(|S_n|\langle \sigma_i, S_n\rangle - \frac{n^2\kappa^3}{\beta^3}\right)\right| + 2\kappa n^2/\beta \\
&\le \frac{2\kappa}{\beta}\mathrm{E}\left||S_n|^3 - \frac{n^3\kappa^3}{\beta^3}\right| + 2\kappa n^2/\beta \\
&\le Cn^{5/2}.
\end{aligned} \tag{5.50}$$

For R_6, we shall prove shortly that

$$\mathrm{E}\left(\sum_{i=1}^{n}\left(\langle \sigma_i, \sigma^{(i)}\rangle^2 - \left(1 - \frac{2}{\beta}\right)\frac{(n-1)^2\kappa^2}{\beta^2}\right)\right)^2 \le Cn^5. \tag{5.51}$$

By (5.51) and the Cauchy inequality, we have

$$\mathrm{E}|R_6| \le Cn^{5/2}. \tag{5.52}$$

For R_7, as $\psi(\kappa)/\kappa = 1/\beta$, and by the smoothness of ψ, we have

$$\left|\frac{\psi(b_i)}{b_i} - \frac{\psi(\kappa)}{\kappa}\right| \le |b_i - \kappa|$$

and

$$\left|\psi(b_i) - \psi(\kappa)\right| \le |b_i - \kappa|.$$

Thus, by (5.48),

$$\mathrm{E}|R_7| \le Cn^2 \sum_{i=1}^{n} \mathrm{E}|b_i - \kappa| \le Cn^{5/2}. \tag{5.53}$$

Then (5.44) follows from (5.47)–(5.53).

(iii) *Proof of* (5.45). Similarly, we have

$$\mathrm{E}\big((W_n - W_n')|W_n - W_n'| \mid \sigma\big) = \frac{4\beta^4}{n^4\kappa^4}\sum_{i=1}^n M_i, \tag{5.54}$$

where

$$M_i = \mathrm{E}\big(\langle\sigma_i, \sigma^{(i)}\rangle|\langle\sigma_i, \sigma^{(i)}\rangle| - \langle\sigma_i', \sigma^{(i)}\rangle|\langle\sigma_i', \sigma^{(i)}\rangle| \mid \sigma\big).$$

We shall prove that

$$\mathrm{E}\Bigg(\sum_{i=1}^n M_i\Bigg)^2 \le Cn^5. \tag{5.55}$$

The proof of (5.55) is given at the end of this subsection.

By the definition of λ and (5.55), we have

$$\frac{1}{\lambda}\mathrm{E}\big|\mathrm{E}\big((W_n - W_n')|W_n - W_n'| \mid \sigma\big)\big| \le Cn^{-1/2}.$$

This proves (5.45). Thus, we complete the proof of Proposition 5.4. □

We now give the proofs of (5.51) and (5.55).

PROOF OF (5.51). Set $a = (1 - \frac{2}{\beta})\frac{(n-1)^2\kappa^2}{\beta^2}$. Given the symmetry, we have

$$\mathrm{E}\Bigg(\sum_{i=1}^n \big(\langle\sigma_i, \sigma^{(i)}\rangle^2 - a\big)\Bigg)^2 = H_1 + H_2, \tag{5.56}$$

where

$$H_1 = n\mathrm{E}\big(\langle\sigma_1, \sigma^{(1)}\rangle^2 - a\big)^2,$$
$$H_2 = n(n-1)\mathrm{E}\big(\langle\sigma_1, \sigma^{(1)}\rangle^2 - a\big)\big(\langle\sigma_2, \sigma^{(2)}\rangle^2 - a\big).$$

For H_1, as $|\sigma^{(1)}| \le n$, we have

$$H_1 \le Cn^5. \tag{5.57}$$

For H_2, we define $\sigma^{(1,2)} = S_n - \sigma_1 - \sigma_2$, and for $j = 1, 2$, we have

$$\big|\langle\sigma_j, \sigma^{(j)}\rangle^2 - \langle\sigma_j, \sigma^{(1,2)}\rangle^2\big| \le Cn.$$

Thus,

$$H_2 = H_3 + L_1, \tag{5.58}$$

where $|L_1| \le Cn^5$ and

$$H_3 = n(n-1)\mathrm{E}\big(\langle\sigma_1, \sigma^{(1,2)}\rangle^2 - a\big)\big(\langle\sigma_2, \sigma^{(1,2)}\rangle^2 - a\big).$$

For $i = 1, 2$, we define

$$V_i(\sigma^{(1,2)}) = \mathrm{E}(\langle \sigma_i, \sigma^{(1,2)} \rangle^2 \mid \sigma^{(1,2)}),$$

and thus,

$$\begin{aligned} &\mathrm{E}(\langle \sigma_1, \sigma^{(1,2)} \rangle^2 - a)(\langle \sigma_2, \sigma^{(1,2)} \rangle^2 - a) \\ &\quad = \mathrm{E}(\langle \sigma_1, \sigma^{(1,2)} \rangle^2 - V_1(\sigma^{(1,2)}))(\langle \sigma_2, \sigma^{(1,2)} \rangle^2 - V_2(\sigma^{(1,2)})) \\ &\qquad + \mathrm{E}(V_1(\sigma^{(1,2)}) - a)(V_2(\sigma^{(1,2)}) - a). \end{aligned} \tag{5.59}$$

Note that the conditional probability density function of (σ_1, σ_2) given $\sigma^{(1,2)}$ is

$$p_{12}(x, y) = \frac{1}{Z_n^{(1,2)}} \exp\left(\frac{\beta}{2n} \langle x, y \rangle^2 + \frac{\beta}{n} \langle x + y, \sigma^{(1,2)} \rangle \right), \tag{5.60}$$

where $x, y \in \mathbb{S}^2$ and

$$Z_n^{(1,2)} = \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \exp\left(\frac{\beta}{2n} \langle x, y \rangle^2 + \frac{\beta}{n} \langle x + y, \sigma^{(1,2)} \rangle \right) dP_n(x)\, dP_n(y).$$

Similarly, we define

$$\tilde{p}_{12}(x, y) = \frac{1}{\tilde{Z}_n^{(1,2)}} \exp\left(\frac{\beta}{n} \langle x + y, \sigma^{(1,2)} \rangle \right), \tag{5.61}$$

where $x, y \in \mathbb{S}^2$ and

$$\tilde{Z}_n^{(1,2)} = \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \exp\left(\frac{\beta}{n} \langle x + y, \sigma^{(1,2)} \rangle \right) dP_n(x)\, dP_n(y).$$

For any $x, y \in \mathbb{S}^2$, we have

$$|p_{12}(x, y) - \tilde{p}_{12}(x, y)| \le Cn^{-1}. \tag{5.62}$$

Let (ξ_1, ξ_2) be a random vector with conditional density function $\tilde{p}_{12}(x, y)$, given $\sigma^{(1,2)}$. Then, for the first term of (5.59), by (5.62), we have

$$\begin{aligned} &\mathrm{E}(\langle \sigma_1, \sigma^{(1,2)} \rangle^2 - V_1(\sigma^{(1,2)}))(\langle \sigma_2, \sigma^{(1,2)} \rangle^2 - V_2(\sigma^{(1,2)})) \\ &\quad = \mathrm{E}(\langle \xi_1, \sigma^{(1,2)} \rangle^2 - \tilde{V}_1(\sigma^{(1,2)}))(\langle \xi_2, \sigma^{(1,2)} \rangle^2 - \tilde{V}_2(\sigma^{(1,2)})) + L_2, \end{aligned} \tag{5.63}$$

where $|L_2| \le Cn^3$ and for $i = 1, 2$,

$$\begin{aligned} \tilde{V}_i(\sigma^{(1,2)}) &= \mathrm{E}(\langle \xi_i, \sigma^{(1,2)} \rangle^2 \mid \sigma^{(1,2)}) \\ &= |\sigma^{(1,2)}|^2 \left(1 - \frac{2\psi(b_{12})}{b_{12}} \right), \end{aligned} \tag{5.64}$$

$$b_{12} = \beta |\sigma^{(1,2)}| / n. \tag{5.65}$$

Observe that given $\sigma^{(1,2)}$, ξ_1 and ξ_2 are conditionally independent; then, the first term of (5.63) is 0, and thus,

$$\big|\mathrm{E}\big(\langle\sigma_1,\sigma^{(1,2)}\rangle^2 - V_1(\sigma^{(1,2)})\big)\big(\langle\sigma_2,\sigma^{(1,2)}\rangle^2 - V_2(\sigma^{(1,2)})\big)\big| \le Cn^3. \tag{5.66}$$

It suffices to bound the second term of (5.59). Again, by (5.62), we have

$$\begin{aligned}&\mathrm{E}\big(V_1(\sigma^{(1,2)}) - a\big)\big(V_2(\sigma^{(1,2)}) - a\big)\\ &\quad = \mathrm{E}\big(\tilde{V}_1(\sigma^{(1,2)}) - a\big)\big(\tilde{V}_2(\sigma^{(1,2)}) - a\big) + L_3,\end{aligned} \tag{5.67}$$

where $|L_3| \le Cn^3$. Recalling that $\beta\psi(\kappa) = \kappa$ and the definition of a, we obtain

$$\begin{aligned}&\big|\tilde{V}_1(\sigma^{(1,2)}) - a\big|\\ &\quad \le |\sigma^{(1,2)}|^2\left|\frac{\psi(b_{12})}{b_{12}} - \frac{\psi(\kappa)}{\kappa}\right| + \left(1 - \frac{2}{\beta}\right)\left||\sigma^{(1,2)}|^2 - \frac{(n-1)^2\kappa^2}{\beta^2}\right|\\ &\quad \le Cn^2|b_{12} - \kappa| + Cn.\end{aligned} \tag{5.68}$$

By (5.68) and similar to (5.48), we have

$$\begin{aligned}&\big|\mathrm{E}\big(\tilde{V}_1(\sigma^{(1,2)}) - a\big)\big(\tilde{V}_2(\sigma^{(1,2)}) - a\big)\big|\\ &\quad \le Cn^4\mathrm{E}|b_{12} - \kappa|^2 + Cn^3\\ &\quad \le Cn^3.\end{aligned} \tag{5.69}$$

It follows from (5.67) and (5.69) that

$$\big|\mathrm{E}\big(V_1(\sigma^{(1,2)}) - a\big)\big(V_2(\sigma^{(1,2)}) - a\big)\big| \le Cn^3. \tag{5.70}$$

The inequalities (5.58), (5.59), (5.66) and (5.70) yield $|H_2| \le Cn^5$, and this completes the proof together with (5.56) and (5.57). □

Next, we give the proof of (5.55).

PROOF OF (5.55). Given the symmetry, we have

$$\mathrm{E}\left(\sum_{i=1}^n M_i\right)^2 = n\mathrm{E}\big(M_1^2\big) + n(n-1)\mathrm{E}(M_1M_2). \tag{5.71}$$

As $|\sigma^{(1)}| \le n$, we have $\mathrm{E}(M_1^2) \le Cn^4$. For $\mathrm{E}(M_1M_2)$, we define

$$\begin{aligned} m_i &= \langle\sigma_i,\sigma^{(i)}\rangle|\langle\sigma_i,\sigma^{(i)}\rangle|,\\ m_i^{(1,2)} &= \langle\sigma_i,\sigma^{(1,2)}\rangle|\langle\sigma_i,\sigma^{(1,2)}\rangle|,\end{aligned}$$

where $i = 1, 2$. Then we have $|m_i - m_i^{(1,2)}| \le Cn$. Thus,

$$\mathrm{E}(M_1M_2) = \mathrm{E}\big(M_1^{(1,2)}M_2^{(1,2)}\big) + L_4, \tag{5.72}$$

where $|L_4| \le Cn^3$ and

$$M_i^{(1,2)} = m_i^{(1,2)} - \mathrm{E}\big(m_i^{(1,2)} \mid \sigma^{(1,2)}\big).$$

Let (ξ_1, ξ_2) be as defined in (5.63). By (5.60)–(5.62), we have

$$\big|\mathrm{E}\big(M_1^{(1,2)} M_2^{(1,2)}\big) - \mathrm{E}\big(\tilde{M}_1^{(1,2)} \tilde{M}_2^{(1,2)}\big)\big| \le Cn^3, \tag{5.73}$$

where for $i = 1, 2$,

$$\tilde{M}_i^{(1,2)} = \tilde{m}_i^{(1,2)} - \mathrm{E}\big(\tilde{m}_i^{(1,2)} \mid \sigma^{(1,2)}\big),$$
$$\tilde{m}_i^{(1,2)} = \langle \xi_i, \sigma^{(1,2)}\rangle |\langle \xi_i, \sigma^{(1,2)}\rangle|.$$

As ξ_1 and ξ_2 are conditionally independent given $\sigma^{(1,2)}$, we have

$$\mathrm{E}\big(\tilde{M}_1^{(1,2)} \tilde{M}_2^{(1,2)}\big) = 0,$$

and by (5.72) and (5.73) we have $|\mathrm{E}(M_1 M_2)| \le Cn^3$. Together with (5.71), we complete the proof of (5.55). □

5.4. *Proof of Theorem* 3.4. As the vertices are colored independently and uniformly, we can construct the exchangeable pair as follows. Let $\xi_i', \ldots, \xi_n'$ be independent copies of $\xi_1, \ldots, \xi_n$, and I be a random index independent of all others and uniformly distributed over $\{1, \ldots, n\}$. Recall that

$$W := W_n = \frac{1}{2}\sum_{i=1}^n \sum_{j \in A_i} \frac{\mathbb{1}_{\{\xi_i = \xi_j\}} - \frac{1}{c_n}}{\sqrt{\frac{m_n}{c_n}(1 - \frac{1}{c_n})}}.$$

We replace ξ_I with ξ_I' in W to obtain a new random variable W'; then (W, W') is an exchangeable pair. Let $\mathscr{X}$ be the sigma field generated by $\{\xi_1, \ldots, \xi_n\}$ and $\sigma^2 = \frac{m_n}{c_n}(1 - \frac{1}{c_n})$. We have

$$\begin{aligned}
\mathrm{E}(W - W' \mid \mathscr{X}) &= \frac{1}{n}\sum_{i=1}^n \sum_{j \in A_i} \frac{\mathbb{1}_{\{\xi_i = \xi_j\}} - \mathrm{E}(\mathbb{1}_{\{\xi_i' = \xi_j\}} \mid \mathscr{X})}{\sigma} \\
&= \frac{1}{n}\sum_{i=1}^n \sum_{j \in A_i} \frac{\mathbb{1}_{\{\xi_i = \xi_j\}} - 1/c_n}{\sigma} \\
&= \frac{2}{n} W.
\end{aligned}$$

Hence, (2.2) holds with $\lambda = \frac{2}{n}$ and $R_n = 0$. By Theorem 2.1, it suffices to prove

$$\begin{aligned}
&\mathrm{E}\left|1 - \frac{1}{2\lambda}\mathrm{E}\big((W - W')^2 \mid W\big)\right| \\
&\quad \le C\big(\sqrt{1/c_n} + \sqrt{d_n^*/m_n} + \sqrt{c_n/m_n}\big)
\end{aligned} \tag{5.74}$$

and

$$\frac{1}{\lambda}\mathrm{E}\big|\mathrm{E}\big((W-W')|W-W'|\mid W\big)\big| \leq C\big(\sqrt{d_n^*/m_n}+\sqrt{c_n/m_n}\big), \tag{5.75}$$

where C is an absolute constant and $d_n^* = \max\{d_i, 1 \leq i \leq n\}$.

PROOF OF (5.74). Observe that

$$\begin{aligned}
&\mathrm{E}\big((W-W')^2 \mid \mathscr{X}\big) \\
&\quad = \frac{1}{n\sigma^2}\sum_{i=1}^n \mathrm{E}\bigg(\bigg(\sum_{j\in A_i} \mathbb{1}_{\{\xi_i=\xi_j\}} - \mathbb{1}_{\{\xi_i'=\xi_j\}}\bigg)^2 \,\bigg|\, \mathscr{X}\bigg) \\
&\quad = \frac{1}{n\sigma^2}\sum_{i=1}^n \bigg(\bigg(\sum_{j\in A_i} (\mathbb{1}_{\{\xi_i=\xi_j\}} - 1/c_n)\bigg)^2 \\
&\qquad + \mathrm{E}\bigg(\bigg(\sum_{j\in A_i} \mathbb{1}_{\{\xi_i'=\xi_j\}} - 1/c_n\bigg)^2 \,\bigg|\, \mathscr{X}\bigg)\bigg) \\
&\quad = \frac{1}{n\sigma^2}\sum_{i=1}^n \bigg(\bigg(\sum_{j\in A_i} h(\xi_i,\xi_j)\bigg)^2 + \mathrm{E}\bigg(\bigg(\sum_{j\in A_i} h(\xi_i',\xi_j)\bigg)^2 \,\bigg|\, \mathscr{X}\bigg)\bigg),
\end{aligned} \tag{5.76}$$

where

$$h(x,y) = \mathbb{1}_{\{x=y\}} - 1/c_n.$$

By the law of total variance, we need only to bound the variance of the first term. Note that

$$\begin{aligned}
&\mathrm{Var}\bigg(\sum_{i=1}^n\bigg(\sum_{j\in A_i} h(\xi_i,\xi_j)\bigg)^2\bigg) \\
&\quad = \sum_{i=1}^n \mathrm{Var}\bigg(\sum_{j\in A_i} h(\xi_i,\xi_j)\bigg)^2 \\
&\qquad + \sum_{i\neq i'} \mathrm{Cov}\bigg(\bigg(\sum_{j\in A_i} h(\xi_i,\xi_j)\bigg)^2, \bigg(\sum_{l\in A_{i'}} h(\xi_{i'},\xi_l).\bigg)^2\bigg).
\end{aligned} \tag{5.77}$$

As

$$\bigg(\sum_{j\in A_i} h(\xi_i,\xi_j)\bigg)^2 = \sum_{j\in A_i} h^2(\xi_i,\xi_j) + \sum_{j\neq l\in A_i} h(\xi_i,\xi_j)h(\xi_i,\xi_l),$$

we have

$$\operatorname{Var}\biggl(\sum_{j\in A_i} h(\xi_i,\xi_j)\biggr)^2 \le 2\operatorname{Var}\biggl(\sum_{j\in A_i} h^2(\xi_i,\xi_j)\biggr) + 2\operatorname{Var}\biggl(\sum_{j\neq l\in A_i} h(\xi_i,\xi_j)h(\xi_i,\xi_l)\biggr).$$

Note that

$$\begin{aligned}
&\operatorname{Var}\biggl(\sum_{j\in A_i} h^2(\xi_i,\xi_j)\biggr)\\
&\quad = \mathrm{E}\biggl(\operatorname{Var}\biggl(\sum_{j\in A_i} h^2(\xi_i,\xi_j)\,\Big|\,\xi_i\biggr)\biggr) + \operatorname{Var}\biggl(\mathrm{E}\biggl(\sum_{j\in A_i} h^2(\xi_i,\xi_j)\,\Big|\,\xi_i\biggr)\biggr)\\
&\quad = d_i\biggl(\frac{1}{c_n}\biggl(1-\frac{1}{c_n}\biggr)\biggl(1-\frac{2}{c_n}+\frac{2}{c_n^2}\biggr)\biggr)\\
&\quad \le d_i/c_n,
\end{aligned}$$

where for every $i \neq j$,

$$\operatorname{Var}\bigl(h^2(\xi_i,\xi_j)\mid\xi_i\bigr) = (1/c_n)(1-1/c_n)\bigl(1-2/c_n+2/c_n^2\bigr) \tag{5.78}$$

and

$$\mathrm{E}\bigl(h^2(\xi_i,\xi_j)\mid\xi_i\bigr) = (1/c_n)(1-1/c_n). \tag{5.79}$$

Also, for $j \neq l \neq i$, $\mathrm{E}(h(\xi_i,\xi_j)h(\xi_i,\xi_l)) = 0$. Thus, we have

$$\begin{aligned}
&\operatorname{Var}\biggl(\sum_{j\neq l\in A_i} h(\xi_i,\xi_j)h(\xi_i,\xi_l)\biggr)\\
&\quad = \mathrm{E}\biggl(\sum_{j\neq l\in A_i} h(\xi_i,\xi_j)h(\xi_i,\xi_l)\biggr)^2\\
&\quad = 2d_i(d_i-1)\biggl(\frac{1}{c_n}\biggl(1-\frac{1}{c_n}\biggr)\biggr)^2\\
&\quad \le 2d_i^2/c_n^2.
\end{aligned}$$

Therefore,

$$\operatorname{Var}\biggl(\sum_{j\in A_i} h(\xi_i,\xi_j)\biggr)^2 \le 4d_i/c_n + 4d_i^2/c_n^2. \tag{5.80}$$

This gives the bound of the first term of (5.77). To bound the second term of (5.77), we let $\delta_{ii'} = \mathbb{1}_{\{(v_i,v_{i'})\in E\}}$ for $i \neq i'$, which indicates the connection between vertex

i and i'. We have

$$
\begin{aligned}
&\operatorname{Cov}\biggl(\Bigl(\sum_{j\in A_i} h(\xi_i,\xi_j)\Bigr)^2, \Bigl(\sum_{l\in A_{i'}} h(\xi_{i'},\xi_l)\Bigr)^2\biggr)\\
&\quad=\operatorname{Cov}\biggl(\sum_{j\in A_i} h^2(\xi_i,\xi_j)+\sum_{j\neq j'\in A_i} h(\xi_i,\xi_j)h(\xi_i,\xi_{j'}),\\
&\qquad \sum_{l\in A_{i'}} h^2(\xi_{i'},\xi_l)+\sum_{l\neq l'\in A_{i'}} h(\xi_{i'},\xi_l)h(\xi_{i'},\xi_{l'})\biggr)\\
&\quad=\sum_{j\in A_i}\sum_{l\in A_{i'}}\operatorname{Cov}\bigl(h^2(\xi_i,\xi_j),h^2(\xi_{i'},\xi_l)\bigr)\\
&\qquad+\sum_{j\neq j'\in A_i}\sum_{l\in A_{i'}}\operatorname{Cov}\bigl(h(\xi_i,\xi_j)h(\xi_i,\xi_{j'}),h^2(\xi_{i'},\xi_l)\bigr)\\
&\qquad+\sum_{j\in A_i}\sum_{l\neq l'\in A_{i'}}\operatorname{Cov}\bigl(h^2(\xi_i,\xi_j),h(\xi_{i'},\xi_l)h(\xi_{i'},\xi_{l'})\bigr)\\
&\qquad+\sum_{j\neq j'\in A_i}\sum_{l\neq l'\in A_{i'}}\operatorname{Cov}\bigl(h(\xi_i,\xi_j)h(\xi_i,\xi_{j'}),h(\xi_{i'},\xi_l)h(\xi_{i'},\xi_{l'})\bigr)\\
&\quad:=H_1+H_2+H_3+H_4.
\end{aligned}
\tag{5.81}
$$

Next, we compute the preceding covariances. For H_1, we have

$$
\begin{aligned}
H_1&=\delta_{ii'}\operatorname{Var}\bigl(h^2(\xi_i,\xi_{i'})\bigr)+\delta_{ii'}\sum_{j\in A_i\setminus\{i'\}}\operatorname{Cov}\bigl(h^2(\xi_i,\xi_{i'}),h^2(\xi_i,\xi_j)\bigr)\\
&\quad+\delta_{ii'}\sum_{l\in A_{i'}\setminus\{i\}}\operatorname{Cov}\bigl(h^2(\xi_i,\xi_{i'}),h^2(\xi_{i'},\xi_l)\bigr)\\
&\quad+\sum_{j\in A_i\setminus\{i'\}}\sum_{l\in A_{i'}\setminus\{i\}}\operatorname{Cov}\bigl(h^2(\xi_i,\xi_j),h^2(\xi_{i'},\xi_l)\bigr).
\end{aligned}
$$

For the first term, by (5.78) and (5.79), we have

$$
\operatorname{Var}\bigl(h^2(\xi_i,\xi_{i'})\bigr)\le 1/c_n.
$$

For $j\in A_i\setminus\{i'\}$, by (5.79), we have

$$
\begin{aligned}
\operatorname{Cov}\bigl(h^2(\xi_i,\xi_{i'}),h^2(\xi_i,\xi_j)\bigr)&=\operatorname{Cov}\bigl(\mathrm{E}\bigl(h^2(\xi_i,\xi_{i'})\mid\xi_i\bigr),\mathrm{E}\bigl(h^2(\xi_i,\xi_j)\mid\xi_i\bigr)\bigr)\\
&=0.
\end{aligned}
$$

Similarly, for $l\in A_{i'}\setminus\{i\}$, we have

$$
\operatorname{Cov}\bigl(h^2(\xi_{i'},\xi_i),h^2(\xi_{i'},\xi_l)\bigr)=0.
$$

For the last term, if $j \neq l \notin \{i, i'\}$, then $h(\xi_i, \xi_j)$ and $h(\xi_{i'}, \xi_l)$ are independent. If $j = l \notin \{i, i'\}$, by (5.79), we have

$$\operatorname{Cov}\big(h^2(\xi_i, \xi_j), h^2(\xi_{i'}, \xi_l)\big) = 0.$$

Therefore,

$$|H_1| \le \delta_{ii'}/c_n. \tag{5.82}$$

For H_2, we have

$$\begin{aligned} H_2 &= \delta_{ii'} \sum_{j \neq j' \in A_i} \operatorname{Cov}\big(h(\xi_i, \xi_j)h(\xi_i, \xi_{j'}), h^2(\xi_i, \xi_{i'})\big) \\ &\quad + \sum_{j \neq j' \in A_i} \sum_{l \in A_{i'} \setminus \{i\}} \operatorname{Cov}\big(h(\xi_i, \xi_j)h(\xi_i, \xi_{j'}), h^2(\xi_i, \xi_l)\big) \\ &= H_{21} + H_{22}. \end{aligned} \tag{5.83}$$

For H_{21}, if $j \neq i'$ or $j' \neq i'$, then

$$\begin{aligned} &\operatorname{Cov}\big(h(\xi_i, \xi_j)h(\xi_i, \xi_{j'}), h^2(\xi_i, \xi_{i'})\big) \\ &\qquad = \mathrm{E}\big(h(\xi_i, \xi_j)h(\xi_i, \xi_{j'})h^2(\xi_i, \xi_{i'})\big) \\ &\qquad = \mathrm{E}\big(\mathrm{E}\big(h(\xi_i, \xi_j)h(\xi_i, \xi_{j'})h^2(\xi_i, \xi_{i'}) \mid \xi_i, \xi_{i'}\big)\big) \\ &\qquad = 0. \end{aligned}$$

If $j = i'$ or $j' = i$, similarly,

$$\operatorname{Cov}\big(h(\xi_i, \xi_j)h(\xi_i, \xi_{j'}), h^2(\xi_i, \xi_{i'})\big) = 0.$$

Therefore,

$$H_{21} = 0. \tag{5.84}$$

For H_{22}, the covariance is not zero only if $\{j, j'\} = \{i', l\}$. Therefore,

$$\begin{aligned} H_{22} &= \sum_{l \in A_i \cap A_{i'}} \operatorname{Cov}\big(h(\xi_i, \xi_{i'})h(\xi_i, \xi_l), h^2(\xi_{i'}, \xi_l)\big) \\ &= \sum_{l \in A_i \cap A_{i'}} \mathrm{E}\big(\mathrm{E}\big(h(\xi_i, \xi_{i'})h(\xi_i, \xi_l), h^2(\xi_{i'}, \xi_l) \mid \xi_{i'}, \xi_l\big)\big) \\ &= \frac{1}{c_n} \sum_{l \in A_i \cap A_{i'}} \mathrm{E}\big(h^3(\xi_{i'}, \xi_l)\big) \\ &\le C(d_i \wedge d_{i'})/c_n^2. \end{aligned} \tag{5.85}$$

Similarly, $H_{22} \ge -C(d_i \wedge d_{i'})/c_n^2$. By (5.83)–(5.85),

$$|H_2| \le C(d_i \wedge d_{i'})/c_n^2. \tag{5.86}$$

Similarly,

$$|H_3| \le C(d_i \wedge d_{i'})/c_n^2. \tag{5.87}$$

For H_4, we have

$$\begin{aligned} H_4 &= 2\delta_{ii'} \sum_{j\in A_i\backslash\{i'\}} \sum_{l\neq l'\in A_{i'}} \operatorname{Cov}\bigl(h(\xi_i,\xi_{i'})h(\xi_i,\xi_j), h(\xi_{i'},\xi_l)h(\xi_{i'},\xi_{l'})\bigr) \\ &\quad + \sum_{j\neq j'\in A_i\backslash\{i'\}} \sum_{l\neq l'\in A_{i'}} \operatorname{Cov}\bigl(h(\xi_i,\xi_j)h(\xi_i,\xi_{j'}), h(\xi_{i'},\xi_l)h(\xi_{i'},\xi_{l'})\bigr) \\ &:= H_{41} + H_{42}. \end{aligned}$$

For H_{41}, the covariance is not zero only if $\{l,l'\} = \{i, j\}$. Thus,

$$\begin{aligned} |H_{41}| &= 4\delta_{ii'} \biggl| \sum_{j\in A_i\cap A_{i'}} \operatorname{Cov}\bigl(h(\xi_i,\xi_{i'})h(\xi_i,\xi_j), h(\xi_{i'},\xi_i)h(\xi_{i'},\xi_j)\bigr)\biggr| \\ &\le C\delta_{ii'}(d_i \wedge d_{i'})/c_n^2. \end{aligned}$$

For H_{42}, the covariance is not zero only if $\{j, j'\} = \{l.l'\}$:

$$\begin{aligned} H_{42} &= 2 \sum_{j\neq j'\in A_i\cap A_{i'}} \operatorname{Cov}\bigl(h(\xi_i,\xi_j)h(\xi_i,\xi_{j'}), h(\xi_{i'},\xi_j)h(\xi_{i'},\xi_{j'})\bigr) \\ &= 2 \sum_{j\neq j'\in A_i\cap A_{i'}} \operatorname{Cov}\bigl(\mathrm{E}\bigl(h(\xi_i,\xi_j)h(\xi_i,\xi_{j'}) \mid \xi_j,\xi_{j'}\bigr), \\ &\qquad \mathrm{E}\bigl(h(\xi_{i'},\xi_j)h(\xi_{i'},\xi_{j'}) \mid \xi_j,\xi_{j'}\bigr)\bigr) \\ &= \frac{2}{c_n^2} \sum_{j\neq j'\in A_i\cap A_{i'}} \operatorname{Var}\bigl(h(\xi_j,\xi_{j'})\bigr) \\ &\le C(d_i \wedge d_{i'})^2/c_n^3. \end{aligned}$$

Therefore,

$$|H_4| \le C\delta_{ii'}(d_i \wedge d_{i'})/c_n^2 + C(d_i \wedge d_{i'})^2/c_n^3. \tag{5.88}$$

Combining (5.81), (5.82), (5.86), (5.87) and (5.88) we have

$$\begin{aligned} &\operatorname{Cov}\biggl(\Bigl(\sum_{j\in A_i} h(\xi_i,\xi_j)\Bigr)^2, \Bigl(\sum_{l\in A_{i'}} h(\xi_{i'},\xi_l)\Bigr)^2\biggr) \\ &\qquad \le C\bigl(\delta_{ij}/c_n + (d_i \wedge d_{i'})/c_n^2 + (d_i \wedge d_{i'})^2/c_n^3\bigr). \end{aligned} \tag{5.89}$$

By (5.77), (5.80) and (5.89), we have

$$\begin{aligned} &\operatorname{Var}\biggl(\sum_{i=1}^n \Bigl(\sum_{j\in A_i} h(\xi_i,\xi_j)\Bigr)^2\biggr) \\ &\qquad \le C\bigl(d_n^* m_n/c_n^2 + m_n/c_n + m_n^2/c_n^3\bigr). \end{aligned}$$

The law of total variance yields

$$\operatorname{Var}\left(\sum_{i=1}^{n} \mathrm{E}\left(\left(\sum_{j\in A_i} h(\xi_i', \xi_j)\right)^2 \,\Big|\, \mathscr{X}\right)\right) \le C\left(\frac{d_n^* m_n}{c_n^2} + \frac{m_n}{c_n} + \frac{m_n^2}{c_n^3}\right),$$

and thus,

$$\begin{aligned}
&\operatorname{Var}\left(\frac{1}{2\lambda}\mathrm{E}((W - W')^2 \mid \mathscr{X})\right)\\
&\qquad\le \frac{C}{\sigma^4}\left(\frac{d_n^* m_n}{c_n^2} + \frac{m_n}{c_n} + \frac{m_n^2}{c_n^3}\right)\\
&\qquad\le C\big(d_n^*/m_n + c_n/m_n + 1/c_n\big).
\end{aligned}$$

This completes the proof of (5.74). $\square$

PROOF OF (5.75). This proof is slightly different from that of (5.74). Observe that

$$\begin{aligned}
&\mathrm{E}((W - W')|W - W'| \mid \mathscr{X})\\
&\quad = n\sigma^2 \sum_{i=1}^{n} \mathrm{E}\left(\left(\sum_{j\in A_i} \mathbb{1}_{\{\xi_i=\xi_j\}} - \mathbb{1}_{\{\xi_i'=\xi_j\}}\right)\left|\sum_{j\in A_i} \mathbb{1}_{\{\xi_i=\xi_j\}} - \mathbb{1}_{\{\xi_i'=\xi_j\}}\right| \,\Big|\, \mathscr{X}\right).
\end{aligned}$$

The variance of the preceding summation can be expanded to

$$\operatorname{Var}\left(\sum_{i=1}^{n} M_i\right) = \sum_{i=1}^{n} \operatorname{Var}(M_i) + \sum_{i\ne i'} \operatorname{Cov}(M_i, M_{i'}),$$

where

$$M_i = \mathrm{E}\left(\left(\sum_{j\in A_i} \mathbb{1}_{\{\xi_i=\xi_j\}} - \mathbb{1}_{\{\xi_i'=\xi_j\}}\right)\left|\sum_{j\in A_i} \mathbb{1}_{\{\xi_i=\xi_j\}} - \mathbb{1}_{\{\xi_i'=\xi_j\}}\right| \,\Big|\, \mathscr{X}\right).$$

Noting that $\mathrm{E}(M_i) = 0$, we have

$$\begin{aligned}
\operatorname{Var}(M_i) &= \mathrm{E}(M_i^2)\\
&\le \mathrm{E}\left(\left(\sum_{j\in A_i} \mathbb{1}_{\{\xi_i=\xi_j\}} - \mathbb{1}_{\{\xi_i'=\xi_j\}}\right)^4\right)\\
&\le C d_i\left(\frac{1}{c_n}\left(1 - \frac{1}{c_n}\right)\right)\left(2d_i\left(\frac{1}{c_n} - \frac{1}{c_n^2}\right) + 1\right).
\end{aligned}$$

To calculate the covariance term, for each $i \ne j$, let $\eta_{ij} = \mathbb{1}_{\{\xi_i=\xi_j\}} - \mathbb{1}_{\{\xi_i'=\xi_j\}}$,

$$T_i = \sum_{j\in A_i} \eta_{ij}, \quad \text{and} \quad T_i^{(i')} = \sum_{j\in A_i\setminus\{i'\}} \eta_{ij}.$$

Then $M_i = \mathrm{E}(T_i|T_i| \mid \mathscr{X})$.

Observe that for $i \neq i'$ and given that $\mathscr{X}$, $T_i|T_i|$ is a function of ξ_i' and $T_{i'}|T_{i'}|$ is a function of $\xi_{i'}'$; thus, $\mathrm{Cov}(T_i|T_i|, T_{i'}|T_{i'}| \mid \mathscr{X}) = 0$. By the total covariance formula, we have $\mathrm{Cov}(M_i, M_{i'}) = \mathrm{Cov}(T_i|T_i|, T_{i'}|T_{i'}|)$. As ξ_i and ξ_i' are independent and identically distributed, $T_i|T_i|$ and $-T_i|T_i|$ are also identically distributed. Therefore, $\mathrm{E}(T_i|T_i|) = 0$, and for some constant C, we have

$$\begin{aligned}
&\mathrm{Cov}(M_i, M_{i'}) \\
&\quad = \mathrm{E}\big(T_i|T_i|T_{i'}|T_{i'}|\big) \\
&\quad = \mathrm{E}\big(T_i^{(i')}|T_i^{(i')}|T_{i'}^{(i)}|T_{i'}^{(i)}|\big) + \mathrm{E}\big(T_i^{(i')}|T_i^{(i')}|\big(T_{i'}|T_{i'}| - \delta_{ii'}T_{i'}^{(i)}|T_{i'}^{(i)}|\big)\big) \\
&\qquad + \mathrm{E}\big(\big(T_i|T_i| - \delta_{ii'}T_i^{(i')}|T_i^{(i')}|\big)T_{i'}^{(i)}|T_{i'}^{(i)}|\big) \\
&\qquad + \mathrm{E}\big(\big(T_i|T_i| - \delta_{ii'}T_i^{(i')}|T_i^{(i')}|\big)\big(T_{i'}|T_{i'}| - \delta_{ii'}T_{i'}^{(i)}|T_{i'}^{(i)}|\big)\big).
\end{aligned}$$

Define $\mathcal{F}_i = \sigma\{\xi_j, j \neq i\}$. Given $\mathcal{F}_i$, $T_i|T_i|$ and $T_{i'}^{(i)}|T_{i'}^{(i)}|$ are conditionally independent,

$$\mathrm{E}\big(T_i|T_i|T_{i'}^{(i)}|T_{i'}^{(i)}|\big) = \mathrm{E}\big(T_{i'}^{(i)}|T_{i'}^{(i)}|\mathrm{E}(T_i|T_i| \mid \mathcal{F}_i)\big) = 0.$$

Similarly,

$$\mathrm{E}\big(T_{i'}|T_{i'}|T_i^{(i')}|T_i^{(i')}|\big) = 0$$

and

$$\mathrm{E}\big(T_i^{(i')}|T_i^{(i')}|T_{i'}^{(i)}|T_{i'}^{(i)}|\big) = 0.$$

Thus,

$$\begin{aligned}
&\mathrm{E}\big(T_i|T_i|T_{i'}|T_{i'}|\big) \\
&\quad = \mathrm{E}\big(\big(T_i|T_i| - \delta_{ii'}T_i^{(i')}|T_i^{(i')}|\big)\big(T_{i'}|T_{i'}| - \delta_{ii'}T_{i'}^{(i)}|T_{i'}^{(i)}|\big)\big).
\end{aligned} \tag{5.90}$$

Without loss of generality, we assume that $\delta_{ii'} = 1$. Note that

$$\begin{aligned}
&\big|T_i|T_i| - T_i^{(i')}|T_i^{(i')}|\big| \\
&\quad = \big|\big(T_i - T_i^{(i')}\big)|T_i| + T_i^{(i')}\big(|T_i| - |T_i^{(i)}|\big)\big| \\
&\quad \leq 2\big|\eta_{ii'}T_i^{(i')}\big| + |\eta_{ii'}^2|,
\end{aligned}$$

and thus,

$$\begin{aligned}
&\mathrm{E}\big(T_i|T_i| - T_i^{(i')}|T_i^{(i')}|\big)^2 \\
&\quad \leq C\mathrm{E}\big(\eta_{ii'}^2\big(T_i^{(i)}\big)^2\big) + C\mathrm{E}\big(\eta_{ii'}^4\big) \\
&\quad = C\bigg(\sum_{j \in A_i \setminus \{i'\}} \mathrm{E}\big(\eta_{ii'}^2\eta_{ij}^2\big) + \sum_{j \neq l \in A_i \setminus \{i'\}} \mathrm{E}\big(\eta_{ii'}^2\eta_{ij}\eta_{il}\big) + \mathrm{E}\big(\eta_{ii'}^4\big)\bigg) \\
&\quad \leq Cd_i/c_n^2 + C/c_n.
\end{aligned}$$

Similarly,

$$\mathrm{E}\big(T_{i'}|T_{i'}| - \delta_{ii'}T_{i'}^{(i)}|T_{i'}^{(i)}|\big)^2 \le C d_{i'}/c_n^2 + C/c_n.$$

By (5.90) and the Cauchy inequality, we finally have

$$\big|\mathrm{E}\big(T_i|T_i|T_{i'}|T_{i'}|\big)\big| \le C\sqrt{d_i d_{i'}}/c_n^2 + C/c_n.$$

Similar to the proof of (5.74), we obtain the bound (5.75). □

Acknowledgments. We thank the referee, the Associate Editor and the Editor for their helpful suggestions which led to a much improved presentation of the paper.

REFERENCES

[1] ARRATIA, R., GOLDSTEIN, L. and GORDON, L. (1990). Poisson approximation and the Chen–Stein method. *Statist. Sci.* **5** 403–424. MR1092983

[2] BARBOUR, A. D., HOLST, L. and JANSON, S. (1992). *Poisson Approximation. Oxford Studies in Probability* **2**. Oxford Univ. Press, New York. MR1163825

[3] CERQUETTI, A. and FORTINI, S. (2006). A Poisson approximation for coloured graphs under exchangeability. *Sankhyā* **68** 183–197. MR2303080

[4] CHATTERJEE, S. (2005). Concentration inequalities with exchangeable pairs. Ph.D. thesis, Stanford Univ., Stanford, CA.

[5] CHATTERJEE, S. (2008). A new method of normal approximation. *Ann. Probab.* **36** 1584–1610. MR2435859

[6] CHATTERJEE, S. (2014). A short survey of Stein's method. In *Proceedings of the International Congress of Mathematicians—Seoul* 2014 (S. Y. Jang, Y. R. Kim, D.-W. Lee and I. Yie, eds.) **IV** 1–24. MR3727600

[7] CHATTERJEE, S. and DEY, P. S. (2010). Applications of Stein's method for concentration inequalities. *Ann. Probab.* **38** 2443–2485. MR2683635

[8] CHATTERJEE, S., DIACONIS, P. and MECKES, E. (2005). Exchangeable pairs and Poisson approximation. *Probab. Surv.* **2** 64–106. MR2121796

[9] CHATTERJEE, S. and MECKES, E. (2008). Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 257–283. MR2453473

[10] CHATTERJEE, S. and SHAO, Q. M. (2011). Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model. *Ann. Appl. Probab.* **21** 464–483. MR2807964

[11] CHEN, L. H. Y., FANG, X. and SHAO, Q. M. (2013). Moderate deviations in Poisson approximation: A first attempt. *Statist. Sinica* **23** 1523–1540. MR3222808

[12] CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q. M. (2011). *Normal Approximation by Stein's Method*. Springer, Heidelberg. MR2732624

[13] CHEN, Y. and SHAO, Q. M. (2012). Berry–Esséen inequality for unbounded exchangeable pairs. In *Probability Approximations and Beyond* (A. Barbour, H. P. Chan and D. Siegmund, eds.). *Lecture Notes in Statistics* **205** 13–30. Springer, New York. MR3289373

[14] DE JONG, P. (1987). A central limit theorem for generalized quadratic forms. *Probab. Theory Related Fields* **75** 261–277. MR0885466

[15] ELLIS, R. and NEWMAN, C. (1978). Limit theorems for sums of dependent random variables occurring in statistical mechanics. *Z. Wahrsch. Verw. Gebiete* **44** 117–139. MR0503333

[16] ELLIS, R. S. and NEWMAN, C. M. (1978). The statistics of Curie–Weiss models. *J. Stat. Phys.* **19** 149–161. MR0503332

[17] ELLIS, R. S. and NEWMAN, C. M. (1978). Fluctuationes in Curie–Weiss exemplis. In *Mathematical Problems in Theoretical Physics* (*Proc*. *Internat*. *Conf*., *Univ. Rome*, *Rome*, 1977). *Lecture Notes in Physics* **80** 313–324. Springer, Berlin–New York. MR0518443

[18] FANG, X. (2015). A universal error bound in the CLT for counting monochromatic edges in uniformly colored graphs. *Electron*. *Commun*. *Probab.* **20** 1–6. MR3320409

[19] GÖTZE, F. and TIKHOMIROV, A. (2002). Asymptotic distribution of quadratic forms and applications. *J*. *Theoret*. *Probab.* **15** 423–475. MR1898815

[20] KIRKPATRICK, K. and MECKES, E. (2013). Asymptotics of the mean-field Heisenberg model. *J*. *Stat*. *Phys*. **152** 54–92. MR3067076

[21] MECKES, E. (2009). On Stein's method for multivariate normal approximation. In *High Dimensional Probability V*: *The Luminy Volume* 158–178. IMS, Beachwood, OH. MR2797946

[22] REINERT, G. and RÖLLIN, A. (2009). Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann*. *Probab.* **37** 2150–2173. MR2573554

[23] RINOTT, Y. and ROTAR, V. (1996). A multivariate CLT for local dependence with $n^{-1/2}\log n$ rate and applications to multivariate graph related statistics. *J*. *Multivariate Anal*. **56** 333–350. MR1379533

[24] RINOTT, Y. and ROTAR, V. (1997). On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics. *Ann*. *Appl*. *Probab.* **7** 1080–1105. MR1484798

[25] SHAO, Q. M. and SU, Z. G. (2006). The Berry–Esseen bound for character ratios. *Proc*. *Amer*. *Math*. *Soc*. **134** 2153–2159. MR2215787

[26] SHAO, Q. M., ZHANG, M. C. and ZHANG, Z. S. (2017). Cramér type moderate deviations for non-normal approximation. Available at arXiv:1809.07966.

[27] SHAO, Q. M. and ZHANG, Z. S. (2016). Identifying the limiting distribution by a general non-normal approximation of Stein's method. *Sci*. *China Math*. **59** 2379–2392.

[28] STEIN, C. (1986). *Approximate Computation of Expectations*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **7**. IMS, Hayward, CA. MR0882007

DEPARTMENT OF STATISTICS
CHINESE UNIVERSITY OF HONG KONG
SHATIN, NEW TERRITORIES
HONG KONG
SAR OF CHINA
E-MAIL: qmshao@sta.cuhk.edu.hk
zhuosongzhang@foxmail.com