*Submitted to Bernoulli*

# Cramér-type moderate deviation of normal approximation for unbounded exchangeable pairs

ZHUO-SONG ZHANG[1]

[1] *Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546.*
*E-mail:* zszhang.stat@gmail.com

In Stein's method, the exchangeable pair approach is commonly used to estimate the approximation errors in normal approximation. In this paper, we establish a Cramér-type moderate deviation theorem of normal approximation for unbounded exchangeable pairs. The results are applied to the sums of local statistics, subgraph counts in the Erdös–Rényi random graph and general Curie–Weiss model to obtain moderate deviation results with optimal convergence rates and ranges.

*MSC2020 subject classifications:* Primary 60F10; secondary 60G60
*Keywords:* Stein's method; exchangeable pair approach; Cramér-type moderate deviation; sums of local statistics; Erdös–Rényi random graph; general Curie–Weiss model

## 1. Introduction

The exchangeable pair approach of Stein's method is commonly used to estimate the convergence rates for distributional approximation. Using exchangeable pair approach, Chatterjee and Shao [7] and Shao and Zhang [21] provided a concrete tool to identify the limiting distribution of the target random variable as well as the $L_1$ bound of the approximation. We refer to Stein [24], Rinott and Rotar [19], Chatterjee, Diaconis and Meckes [5], Chatterjee and Meckes [6] and Meckes [16] for other related results of $L_1$ bound and Berry–Esseen bound for the exchangeable pair approach. Recently, Shao and Zhang [22] obtained a Berry–Esseen-type bound of normal and nonnormal approximation for unbounded exchangeable pairs. Specifically, let $W$ be the random variable of interest, and we say $(W, W')$ an *exchangeable pair* if $(W, W') \stackrel{d.}{=} (W, W')$. Let $\Delta = W - W'$. It is often to assume that there exists a constant $\lambda > 0$ and a random variable $R$ such that

$$\mathbb{E}\{\Delta \mid W\} = \lambda(W + R). \tag{1.1}$$

Shao and Zhang [22] proved that

$$\sup_{z\in\mathbb{R}}\big|\mathbb{P}(W\leqslant z)-\Phi(z)\big|\leqslant\mathbb{E}\bigg|1-\frac{1}{2\lambda}\,\mathbb{E}\left\{\Delta^2\,\middle|\,W\right\}\bigg|+\frac{1}{\lambda}\,\mathbb{E}\big|\mathbb{E}\left\{\Delta^*\Delta\,|\,W\right\}\big|+\mathbb{E}\,|R|, \tag{1.2}$$

where $\Phi(z)$ is the standard normal distribution function and where $\Delta^*:=\Delta^*(W,W')$ is any random variable satisfying that $\Delta^*(W,W')=\Delta^*(W',W)$ and $\Delta^*\geqslant|\Delta|$.

While the $L_1$ bound and Berry–Esseen-type bound describe the absolute error for the distributional approximation, the Cramér-type moderate deviation reflects the relative error in convergences in distribution. More precisely, let $\{Y_n, n\geqslant 1\}$ be a sequence of random variables that converge to $Y$ in distribution, the Cramér-type moderate deviation is

$$\frac{\mathbb{P}(Y_n>x)}{\mathbb{P}(Y>x)}=1+\text{ error term}\to 1$$

for $0\leqslant x\leqslant a_n$, where $a_n\to\infty$ as $n\to\infty$. Specially, for independent and identically distributed (i.i.d.) random variables $X_1,\dots,X_n$ with $\mathbb{E}\,X_1=0$, $\mathbb{E}\,X_1^2=1$ and $\mathbb{E}\,\mathrm{e}^{t_0\sqrt{|X_1|}}<\infty$, where $t_0>0$ is a constant, put $W_n=n^{-1/2}(X_1+\cdots+X_n)$. Then,

$$\frac{\mathbb{P}(W_n>x)}{1-\Phi(x)}=1+O(1)n^{-1/2}(1+x^3), \tag{1.3}$$

for $0\leqslant x\leqslant n^{1/6}$. We refer to Linnik [15] and Petrov [17] for details. The condition $\mathbb{E}\,\mathrm{e}^{t_0\sqrt{|X_1|}}<\infty$ is necessary, and the range $0\leqslant x\leqslant n^{1/6}$ and the order of the error term $n^{-1/2}(1+x^3)$ are optimal.

The proof of Cramér-type moderate deviation result (1.3) for independent random variables is based on the conjugate method and the Fourier transform. However, Stein's method usually performs better than the method of Fourier transform for dependent random variables. Since introduced by Stein [23] in 1972, Stein's method has been deeply developed in recent years, and shows its importance and power for estimating the approximation errors of normal and nonnormal approximation. We refer to Chen, Goldstein and Shao [11] and Chatterjee [4] for more details. In addition to the $L_1$ bound and Berry–Esseen bound, moderate deviation results can also be established by applying Stein's method in the literature. For instance, using Stein's method, Raič [18] proved the moderate deviation under certain local dependence structures. In the context of Poisson approximation, Barbour, Holst and Janson [3], Chen and Choi [8] and Barbour, Chen and Choi [1] applied Stein's method to prove moderate deviation results for sums of independent indicators, whereas Chen, Fang and Shao [9] studied sums of dependent indicators. Moreover, Chen, Fang and Shao [10] and Shao, Zhang and Zhang [20] obtained the general Cramér-type moderate deviation results of normal and nonnormal approximation for dependent random variables whose dependence structure is defined in terms of a Stein identity under a boundedness assumption on $|\Delta|$ (see Remark 2.2 for more details).

However, in practice, it may not be easy to check the condition (1.1) in general, and the boundedness assumption on $|\Delta|$ is also too strict in applications. In this paper, our aim is to apply Stein's method and the exchangeable pair approach to prove a Cramér-type moderate deviation result without assuming that $|\Delta|$ is bounded. The results are then applied to sums of local statistics, the subgraph counts in the Erdös–Rényi random graph and the general Curie–Weiss model to obtain the Cramér-type moderate deviation results with optimal ranges and convergence rates.

The rest of this paper is organized as follows. We present our main results in Section 2. In Section 3, we give some applications of our main result. The proofs of Theorems 2.1 and 2.2 and Corollary 2.1 are put in Section 4. The proofs of other results are postponed to Section 5.

## 2. Main results

Let $X$ be a field of random variables valued on a measurable space $\mathscr{X}$, and $W = \varphi(X) \in \mathbb{R}$ be the random variable of interest where $\varphi : \mathscr{X} \to \mathbb{R}$. We consider the following condition:

(D1) Let $(X, X')$ be an exchangeable pair. Assume that there exists $D := \Psi(X, X')$, where $\Psi : \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ is an anti-symmetric function, satisfying $\mathbb{E}\{D \,|\, X\} = \lambda(W + R)$ where $\lambda > 0$ is a constant and $R$ is a random variable.

***Remark* 2.1.** *The condition* (D1) *is a natural generalization of (1.1). Specially, if (1.1) is satisfied, we can simply choose $D = \Delta$. Under the condition* (D1)*, for any absolutely continuous function $f : \mathbb{R} \to \mathbb{R}$, by antisymmetry property of $D$, it follows that $\mathbb{E}\big\{D\big(f(W) + f(W')\big)\big\} = 0$ and a direct rearranging yields*

$$\begin{aligned}
0 &= 2\,\mathbb{E}\{Df(W)\} - \mathbb{E}\big\{D\big(f(W) - f(W')\big)\big\} \\
&= 2\lambda\,\mathbb{E}\{(W + R)f(W)\} - \mathbb{E}\bigg\{D\int_{-\Delta}^{0} f'(W + u)\,\mathrm{d}u\bigg\}.
\end{aligned}$$

*Then,*

$$\mathbb{E}\big\{Wf(W)\big\} = \frac{1}{2\lambda}\,\mathbb{E}\bigg\{D\int_{-\Delta}^{0} f'(W + u)\,\mathrm{d}u\bigg\} - \mathbb{E}\big\{Rf(W)\big\}. \tag{2.1}$$

Our main result Theorem 2.1 provides a Cramér-type moderate deviation result under the condition (D1) without the assumption that $|\Delta|$ is bounded:

**Theorem 2.1.** *Let $(X, X')$ be an exchangeable pair satisfying the condition* (D1)*, $W' = \varphi(X')$ and $\Delta = W - W'$. Let $D^* := D^*(X, X')$ be any random variable such that $D^*(X, X') = D^*(X, X')$ and $D^* \geqslant |D|$. Assume that there exists a constant $\tau_0 > 0$ and increasing functions $\delta_1(t), \delta_2(t)$ and $\delta_3(t)$ such that for all $0 \leqslant t \leqslant \tau_0$,*

(A1) $\mathbb{E}\big\{\big|1 - \frac{1}{2\lambda}\,\mathbb{E}\{D\Delta \,|\, X\}\big|\,\mathrm{e}^{tW}\big\} \leqslant \delta_1(t)\,\mathrm{e}^{t^2/2}$,

(A2) $\mathbb{E}\big\{\big|\frac{1}{2\lambda}\,\mathbb{E}\{D^*\Delta \,|\, X\}\big|\,\mathrm{e}^{tW}\big\} \leqslant \delta_2(t)\,\mathrm{e}^{t^2/2}$, *and*

(A3) $\mathbb{E}\big\{|R|\,\mathrm{e}^{tW}\big\} \leqslant \delta_3(t)\,\mathrm{e}^{t^2/2}$.

*Then,*

$$\Big|\frac{\mathbb{P}(W > z)}{1-\Phi(z)} - 1\Big| \leqslant 20\big((1+z^2)\big(\delta_1(z)+\delta_2(z)\big) + (1+z)\delta_3(z)\big), \tag{2.2}$$

*provided that* $0 \leqslant z \leqslant \tau_0$.

We also have the following corollary.

**Corollary 2.2.** *Let* $(X, X')$*,* $W$*,* $W'$*,* $\Delta$*, and* $D^*$ *be as in Theorem 2.1 and assume that condition* (D1) *is satisfied. Assume that there exists a constant* $\tau > 0$ *and increasing functions* $\delta_1(t), \delta_2(t)$ *and* $\delta_3(t)$ *such that for all* $0 \leqslant t \leqslant \tau$*,*

(B1) $\mathbb{E}\{\mathrm{e}^{tW}\} < \infty$,
(B2) $\mathbb{E}\big\{\big|1 - \frac{1}{2\lambda}\,\mathbb{E}\,\{D\Delta \,|\, X\}\big|\,\mathrm{e}^{tW}\big\} \leqslant \delta_1(t)\,\mathbb{E}\,\mathrm{e}^{tW}$,
(B3) $\mathbb{E}\big\{\big|\frac{1}{2\lambda}\,\mathbb{E}\,\{D^*\Delta \,|\, X\}\big|\,\mathrm{e}^{tW}\big\} \leqslant \delta_2(t)\,\mathbb{E}\,\mathrm{e}^{tW}$, *and*
(B4) $\mathbb{E}\big\{|R|\,\mathrm{e}^{tW}\big\} \leqslant \delta_3(t)\,\mathbb{E}\,\mathrm{e}^{tW}$.

*For* $\theta > 0$*, let*

$$\tau_0(\theta) := \max\left\{0 \leqslant t \leqslant \tau : \frac{t^2}{2}\big(\delta_1(t)+\delta_2(t)\big) + \delta_3(t)t \leqslant \theta\right\}.$$

*Then, for any* $\theta > 0$*,*

$$\Big|\frac{\mathbb{P}(W > z)}{1-\Phi(z)} - 1\Big| \leqslant 20\,\mathrm{e}^{\theta}\big((1+z^2)\big(\delta_1(z)+\delta_2(z)\big) + (1+z)\delta_3(z)\big), \tag{2.3}$$

*provided that* $0 \leqslant z \leqslant \tau_0(\theta)$.

***Remark 2.2.*** *We now make some remarks on our results and Chen, Fang and Shao [10]'s results. Chen, Fang and Shao [10] proved a moderate deviation result for a general Stein identity under some boundedness assumption. We now cite their results as follows (see, e.g., Chen, Fang and Shao [10, Theorem 3.1]). Assume that there exists a constant* $\delta > 0$*, a random function* $\widehat{K}(u) \geqslant 0$ *and a random variable* $\widehat{R}$ *such that for any absolutely continuous function* $f$*, the following Stein identity holds:*

$$\mathbb{E}\big\{Wf(W)\big\} = \mathbb{E}\bigg\{\int_{|t|\leqslant\delta} f'(W+u)\widehat{K}(u)\,\mathrm{d}u\bigg\} + \mathbb{E}\big\{\widehat{R}f(W)\big\}. \tag{2.4}$$

*Let* $K_1 := \int_{|t|\leqslant\delta}\widehat{K}(u)\,\mathrm{d}u$*, and assume that there exists constants* $d_0, \widehat{\delta}_1$ *and* $\widehat{\delta}_2$ *such that*

$$|K_1| \leqslant d_0, \quad |\,\mathbb{E}\,\{K_1 \,|\, W\} - 1| \leqslant \widehat{\delta}_1(1+|W|), \quad \big|\mathbb{E}\big\{\widehat{R} \,\big|\, W\big\}\big| \leqslant \widehat{\delta}_2(1+|W|). \tag{2.5}$$

*Chen, Fang and Shao [10, Theorem 3.1] proved that the random variable $W$ has the following moderate deviation result:*

$$\frac{\mathbb{P}(W>z)}{1-\Phi(z)} = 1 + O(1)d_0^3(1+z^3)(\delta+\widehat{\delta}_1+\widehat{\delta}_2), \tag{2.6}$$

*for $0 \leqslant z \leqslant d_0^{-1}(\delta^{-1/3}+\widehat{\delta}_1^{-1/3}+\widehat{\delta}_2^{-1/3})$ where $O(1)$ is bounded by a universal constant. Specially, for an exchangeable pair $(W,W')$ satisfying (1.1) and $|\Delta| \leqslant \delta$, we have (2.4) is satisfied with $\widehat{K}(u) = \Delta(\mathbf{1}_{\{-\Delta\leqslant t\leqslant 0\}} - \mathbf{1}_{\{0<t\leqslant -\Delta\}})/(2\lambda), K_1 = \Delta^2/(2\lambda)$ and $\widehat{R} = -R$.*

*Assume that the condition* (D1) *holds and there exists a constant $\delta > 0$ such that $|D| \leqslant \delta$ and $|\Delta| \leqslant \delta$. By (2.1), the equality (2.4) holds with*

$$\widehat{K}(u) = \frac{1}{2\lambda} D\big(\mathbf{1}_{\{-\Delta\leqslant t\leqslant 0\}} - \mathbf{1}_{\{0<t\leqslant -\Delta\}}\big),$$

*$\widehat{R} = -R$ and $K_1 = (D\Delta)/(2\lambda)$. Under the condition* (B1)*, it can be shown that (see, e.g. Chen, Fang and Shao [10, Lemma 5.1] and Shao, Zhang and Zhang [20, Lemma 4.4]) the condition (2.5) imply conditions* (B2)*–*(B4) *with $\delta_1(t) = C\widehat{\delta}_1(1+t)$, $\delta_2(t) = C\delta(1+\widehat{\delta}_2)(1+t)$ and $\delta_3(t) = C\widehat{\delta}_2(1+t)$, where $C > 0$ is a constant depending on $d_0$. Hence, under the exchangeable pair approach setting and under the assumptions $|\Delta| \leqslant \delta$ and (2.5), Corollary 2.2 implies (2.6).*

# 3. Applications

## 3.1. Sums of local statistics

Let $\mathcal{J}$ be an index set and $\xi = \{\xi_i, i \in \mathcal{J}\}$ a field of independent random variables where $\xi_i$ is valued on a measurable space $\mathscr{X}$. For any subset $J \subset \mathcal{J}$, define $\xi_J := \{\xi_j, j \in J\}$. For each $1 \leqslant i \leqslant n$, let $X_i = f_i(\xi_{J_i})$ where $f_i : \mathscr{X}^{|J_i|} \mapsto \mathbb{R}$ and $J_i \subset \mathcal{J}$. Assume that $\mathbb{E}\{X_i\} = 0$ and $\operatorname{Var}(X_i) < \infty$ for each $1 \leqslant i \leqslant n$. Let $W = \sum_{i=1}^n X_i$ be the random variable of interest such that $\operatorname{Var}(W) = 1$. Put $\mathcal{A}_i = \{j : J_i \cap J_j \neq \varnothing\}$.

Assume that there exist constants $\alpha > 0$ and $\beta \geqslant 1$ such that for each $1 \leqslant i \leqslant n$,

$$\mathbb{E}\big\{e^{\alpha|X_i|} \bigm| X_k, k \neq i\big\} \leqslant \beta, \quad \text{almost surely.} \tag{3.1}$$

We have the following theorem.

**Theorem 3.1.** *Let $d := \max\big\{|\mathcal{A}_i|, 1 \leqslant i \leqslant n\big\}$, and assume that (3.1) is satisfied, then for any $\theta > 0$,*

$$\left|\frac{\mathbb{P}(W>z)}{1-\Phi(z)} - 1\right| \leqslant 40\, e^{\theta}\, \delta(z)(1+z^2),$$

*for* $0 \leqslant z \leqslant \tau_0(\theta)$, *where*

$$\delta(t) = 24\beta^{5d/2}d^{3/2}\Big\{\sum_{i\in\mathcal{J}}\gamma_{4,i}(t)\Big\}^{1/2} + 192n^{1/2}\beta^{4d}d^2t\Big\{\sum_{i=1}^{n}\gamma_{6,i}(2t)\Big\}^{1/2}, \tag{3.2}$$

*and where* $\gamma_{p,m}(t) = \mathbb{E}\{|X_m|^p\,\mathrm{e}^{t|X_m|}\}$ *for* $p \geqslant 0$ *and* $\tau_0(\theta) = \max\{0 \leqslant t \leqslant \alpha : t^2\delta(t) \leqslant \theta\}$.

***Remark* 3.1.** *If* $d$ *is finite and* $|X_i| \leqslant Cn^{-1/2}$ *for some constant* $C > 0$, *then (3.1) is satisfied with* $\alpha = C^{-1}n^{1/2}$ *and* $\beta = \mathrm{e}$. *For* $0 \leqslant t \leqslant \alpha$, *we have* $\gamma_{4,i}(t) = O(n^{-2})$ *and* $\gamma_{6,t} = O(n^{-3})$. *In this case, Theorem 3.1 reduces to*

$$\frac{\mathbb{P}(W > z)}{1 - \Phi(z)} = 1 + O(n^{-1/2})(1 + z^3)$$

*for* $0 \leqslant z \leqslant n^{1/6}$. *This provides an same convergence rate and range as the i.i.d. case, and hence the result is optimal.*

***Remark* 3.2.** *Recently, Fang, Luo and Shao [14] proved a higher-order approximation relative error for the cases where* $|X_i|$*'s are bounded.*

## 3.2. Subgraph counts in the Erdös–Rényi random graph

For any integer $k \geqslant 1$, let $[n]_k := \{(i_1,\dots,i_k) : 1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n\}$. We say $\mathcal{G}(n,p)$ is an *Erdös–Rényi random graph* with vertices set $\mathcal{V} = \{1,\dots,n\}$ if each node pair $(i,j)$ is independently connected with probability $p$. For $(i,j) \in [n]_2$, define $\xi_{ij} = \mathbf{1}_{\{i \text{ and } j \text{ are connected in } \mathcal{G}(n,p)\}}$. Then, $\{\xi_{ij}, (i,j) \in [n]_2\}$ are independent and for each $(i,j) \in [n]_2$, $\mathbb{P}(\xi_{ij} = 1) = 1 - \mathbb{P}(\xi_{ij} = 0) = p$. For any graph $H$, let $v(H)$ and $e(H)$ denote the number of its vertices and edges, respectively.

Let $G$ be a given fixed graph with at least one edge. Let $S_n$ be the number of copies (not necessarily induced) of $G$ in $\mathcal{G}(n,p)$. Let $\mu_n = \mathbb{E}\{S_n\}$, $\sigma_n = \sqrt{\mathrm{Var}(S_n)}$ and $W_n = (S_n - \mu_n)/\sigma_n$.

**Theorem 3.2.** *Let* $\psi_n = \min\{n^{v(H)}p^{e(H)} : H \subset G, e(H) > 0\}$. *We have*

$$\frac{\mathbb{P}(W_n > z)}{1 - \Phi(z)} = 1 + O(1)(1 + z^2)b_n(p,z), \tag{3.3}$$

*for* $0 \leqslant z \leqslant (1-p)^{1/2}n^2p^{e(G)}\psi_n^{-1/2}$ *such that* $(1+z^2)b_n(p,z) \leqslant 1$, *where* $O(1)$ *is bounded by a constant depending only on* $G$ *and*

$$b_n(p,z) = \begin{cases} \psi_n^{-1/2}(1+z) & \text{if } 0 < p < 1/2, \\ n^{-1}(1-p)^{-1/2}\{1 + (1-p)^{-1/2}z\} & \text{if } 1/2 < p < 1. \end{cases}$$

***Remark* 3.3.** *After we finish this paper, we learned that Fang, Luo and Shao [14] proved a higher-order relative error using a different method.*

***Remark* 3.4.** *For fixed $p$ which is bounded away from 0 and 1, and independent of $n$, then for sufficiently large $n$, we have $\psi_n = O(n^2)$, $p = O(1)$ and the sample size is $N = n(n-1)/2$. In this case, (3.3) yields*

$$\frac{\mathbb{P}(W_n > z)}{1 - \Phi(z)} = 1 + O(1) N^{-1/2}(1 + z^3),$$

*for $z \in (0, N^{1/6})$. This shows that the convergence rate and range are as optimal as the i.i.d. case.*

### 3.3. The general Curie–Weiss model

The Curie–Weiss model of ferromagnetic interaction has been extensively studied in the past decades. Let $\rho$ be a probability measure on $\mathbb{R}$ such that

$$\int_{-\infty}^{\infty} x \,\mathrm{d}\rho(x) = \int_{-\infty}^{\infty} x^3 \,\mathrm{d}\rho(x) = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} x^2 \,\mathrm{d}\rho(x) = 1. \tag{3.4}$$

The *general Curie–Weiss model* $\mathrm{CW}(\rho)$ at inverse temperature $\beta$ is defined as the array of spin random variables $X = (X_1, \dots, X_n)$ with joint distribution

$$\mathrm{d}P_{n,\beta}(x) = Z_n^{-1} \exp\left(\frac{\beta}{2n}(x_1 + \cdots + x_n)^2\right) \prod_{i=1}^{n} \mathrm{d}\rho(x_i) \tag{3.5}$$

for $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ where $Z_n$ is the normalizing constant:

$$Z_n = \int \exp\left(\frac{\beta}{2n}(x_1 + \cdots + x_n)^2\right) \prod_{i=1}^{n} \mathrm{d}\rho(x_i).$$

The Curie–Weiss model $\mathrm{CW}(\rho)$ is called "at the critical temperature" if $\beta = 1$. The magnetization is defined by $m(X) = \frac{1}{n}\sum_{i=1}^{n} X_i$. The asymptotic behavior of the magnetization is well studied by Ellis and Newman [12, 13]. Stein's method can be applied to estimate the convergence rate, for example, using the exchangeable pair approach, Chatterjee and Shao [7] obtained a Berry–Esseen bound at the critical temperature of the simplest Curie–Weiss model, and Shao and Zhang [22] proved the Berry–Esseen bound of the general Curie–Weiss model for both noncritical and critical temperature. Moreover, Chen, Fang and Shao [10] and Shao, Zhang and Zhang [20] obtained the Cramér-type moderate deviation results for the cases where $\rho$ has a finite support. In this subsection, we establish the Cramér-type moderate deviation result for the general Curie–Weiss model at noncritical temperature with infinitely supported probability measure $\rho$. Let $(X_1, \dots, X_n)$ follow the joint distribution (3.5) with $0 < \beta < 1$ and $\rho$ satisfying (3.4) and

$$\int_{-\infty}^{\infty} \mathrm{e}^{tx} \, \mathrm{d}\rho(x) \leqslant \mathrm{e}^{t^2/2} \quad \text{for} \quad t \in \mathbb{R}. \tag{3.6}$$

Let $S_n = X_1 + \cdots + X_n$ and $W_n = n^{-1/2}(1-\beta)^{1/2} S_n$. We have the following theorem.

**Theorem 3.3.** *We have*

$$\left| \frac{\mathbb{P}(W_n > z)}{1 - \Phi(z)} - 1 \right| \leqslant C n^{-1/2}(1 + z^3) \quad \text{for } 0 \leqslant z \leqslant \sqrt{n}. \tag{3.7}$$

The Berry–Esseen bound was obtained by Shao and Zhang [22] with the convergence rate $O(n^{-1/2})$. For the simplest Curie–Weiss model, where the magnetization is valued on $\{-1, 1\}$ with equal probability, Chen, Fang and Shao [10] proved the same convergence rate as (3.7) with convergence range $[0, n^{1/6}]$. However, Theorem 3.3 provides a wider convergence range.

## 4. Proofs of Theorem 2.1 and Corollary 2.2

In this section, we give the proofs of our main results in Section 2. Before proving Theorem 2.1 and Corollary 2.2, we first present some preliminary lemmas. In the proofs, we use the techniques in Chen, Fang and Shao [10, Lemmas 5.1–5.2] and Shao and Zhang [22, pp. 71–73].

**Lemma 4.1.** *Let $f$ be a nondecreasing function. Then,*

$$\left| \mathbb{E}\left\{ D \int_{-\Delta}^{0} (f(W+u) - f(W)) \, \mathrm{d}u \right\} \right| \leqslant \mathbb{E}\{ D^* \Delta f(W) \},$$

*where $D^*$ is as defined in Theorem 2.1.*

***Proof of Lemma 4.1.*** Since $f(\cdot)$ is nondecreasing, it follows that

$$0 \geqslant \int_{-\Delta}^{0} \{ f(W+u) - f(W) \} \, \mathrm{d}u \geqslant -\Delta \{ f(W) - f(W') \},$$

which yields

$$\begin{aligned} -\mathbb{E}\big[ D \mathbf{1}_{\{D>0\}} \Delta \{ f(W) - f(W') \} \big] &\leqslant \mathbb{E}\left[ D \int_{-\Delta}^{0} \{ f(W+u) - f(W) \} \, \mathrm{d}u \right] \\ &\leqslant -\mathbb{E}\big[ D \mathbf{1}_{\{D<0\}} \Delta \{ f(W) - f(W') \} \big]. \end{aligned}$$

Recalling that $W = \varphi(X)$, $D = D(X, X')$ is antisymmetric and $D^* = D^*(X, X')$ is symmetric, as $(X, X')$ is exchangeable, we have

$$\mathbb{E}\big[ D \mathbf{1}_{\{D>0\}} \Delta \{ f(W) - f(W') \} \big] = -\mathbb{E}\big[ D \mathbf{1}_{\{D<0\}} \Delta \{ f(W) - f(W') \} \big],$$

and

$$\mathbb{E}\big\{D^*\mathbf{1}_{\{D>0\}}\Delta f(W)\big\} = -\,\mathbb{E}\big\{D^*\mathbf{1}_{\{D<0\}}\Delta f(W')\big\}.$$

Moreover, as $\mathbb{E}\,D^*\Delta\mathbf{1}_{\{D=0\}}\big\{f(W)-f(W')\big\}\geqslant 0$ and $\mathbb{E}\big\{D^*\mathbf{1}_{\{D=0\}}\Delta f(W)\big\} = -\,\mathbb{E}\big\{D^*\mathbf{1}_{\{D=0\}}\Delta f(W')\big\}$, it follows that $\mathbb{E}\,D^*\Delta\mathbf{1}_{\{D=0\}}f(W)\geqslant 0$. Therefore,

$$\begin{aligned}\left|\mathbb{E}\left[D\int_{-\Delta}^{0}\big\{f(W+u)-f(W)\big\}\,\mathrm{d}u\right]\right| &\leqslant -\,\mathbb{E}\,D\mathbf{1}_{\{D<0\}}\Delta\big\{f(W)-f(W')\big\}\\ &\leqslant \mathbb{E}\,D^*\mathbf{1}_{\{D<0\}}\Delta\big\{f(W)-f(W')\big\}\\ &= \mathbb{E}\,D^*\Delta\big(\mathbf{1}_{\{D>0\}}+\mathbf{1}_{\{D<0\}}\big)f(W)\\ &\leqslant \mathbb{E}\,D^*\Delta f(W),\end{aligned}$$

as desired. □

**Lemma 4.2.** *Under the conditions of Theorem 2.1, we have for $0\leqslant z\leqslant \tau_0$,*

$$\mathbb{E}\left\{\left|1-\frac{1}{2\lambda}\,\mathbb{E}\,\{D\Delta\,|\,W\}\right|W\,\mathrm{e}^{W^2/2}\,\mathbf{1}_{\{0\leqslant W\leqslant z\}}\right\}\leqslant 4(1+z^2)\delta_1(z), \tag{4.1}$$

$$\frac{1}{2\lambda}\,\mathbb{E}\left\{\big|\mathbb{E}\,\{D^*\Delta\,|\,W\}\big|W\,\mathrm{e}^{W^2/2}\,\mathbf{1}_{\{0\leqslant W\leqslant z\}}\right\}\leqslant 4(1+z^2)\delta_2(z), \tag{4.2}$$

$$\mathbb{E}\left\{|R|\,\mathrm{e}^{W^2/2}\,\mathbf{1}_{\{0\leqslant W\leqslant z\}}\right\}\leqslant 2(1+z)\delta_3(z). \tag{4.3}$$

***Proof of Lemma 4.2.*** We apply the idea of Chen, Fang and Shao [10, Lemma 5.2] in this proof. For $a\in\mathbb{R}_+$, denote $\lfloor a\rfloor=\max\{n\in\mathbb{N}:n\leqslant a\}$. By condition (A1), and recalling that the function $\delta_1(\cdot)$ is increasing, for any $0\leqslant x\leqslant z\leqslant\tau_0$,

$$\mathrm{e}^{-x^2/2}\,\mathbb{E}\left\{\left|1-\frac{1}{2\lambda}\,\mathbb{E}\,\{D\Delta\,|\,W\}\right|\mathrm{e}^{xW}\right\}\leqslant \delta_1(x)\leqslant\delta_1(z). \tag{4.4}$$

We have

$$\begin{aligned}&\mathbb{E}\left\{\left|1-\frac{1}{2\lambda}\,\mathbb{E}\,\{D\Delta\,|\,W\}\right|W\,\mathrm{e}^{W^2/2}\,\mathbf{1}_{\{0\leqslant W\leqslant z\}}\right\}\\ &\quad=\sum_{j=1}^{\lfloor z\rfloor}\mathbb{E}\left\{\left|1-\frac{1}{2\lambda}\,\mathbb{E}\,\{D\Delta\,|\,W\}\right|W\,\mathrm{e}^{W^2/2}\,\mathbf{1}_{\{j-1\leqslant W<j\}}\right\}\\ &\qquad+\mathbb{E}\left\{\left|1-\frac{1}{2\lambda}\,\mathbb{E}\,\{D\Delta\,|\,W\}\right|W\,\mathrm{e}^{W^2/2}\,\mathbf{1}_{\{\lfloor z\rfloor\leqslant W\leqslant z\}}\right\}\\ &\quad\leqslant\sum_{j=1}^{\lfloor z\rfloor} j\,\mathrm{e}^{(j-1)^2/2-j(j-1)}\,\mathbb{E}\left\{\left|1-\frac{1}{2\lambda}\,\mathbb{E}\,\{D\Delta\,|\,W\}\right|\mathrm{e}^{jW}\,\mathbf{1}_{\{j-1\leqslant W<j\}}\right\}\end{aligned}$$

$$
\begin{aligned}
&+ z\,\mathrm{e}^{\lfloor z\rfloor^2/2-\lfloor z\rfloor z}\,\mathbb{E}\left\{\left|1-\frac{1}{2\lambda}\,\mathbb{E}\left\{D\Delta\,|\,W\right\}\right|\mathrm{e}^{zW}\,\mathbf{1}_{\{\lfloor z\rfloor\leqslant W\leqslant z\}}\right\}\\
&\leqslant 2\sum_{j=1}^{\lfloor z\rfloor} j\,\mathrm{e}^{-j^2/2}\,\mathbb{E}\left\{\left|1-\frac{1}{2\lambda}\,\mathbb{E}\left\{D\Delta\,|\,W\right\}\right|\mathrm{e}^{jW}\,\mathbf{1}_{\{j-1\leqslant W<j\}}\right\}\\
&\quad+2z\,\mathrm{e}^{-z^2/2}\,\mathbb{E}\left\{\left|1-\frac{1}{2\lambda}\,\mathbb{E}\left\{D\Delta\,|\,W\right\}\right|\mathrm{e}^{zW}\,\mathbf{1}_{\{\lfloor z\rfloor\leqslant W\leqslant z\}}\right\}\\
&\leqslant 2\delta_1(z)\left(\sum_{j=1}^{\lfloor z\rfloor} j+z\right)\leqslant 4(1+z^2)\delta_1(z),
\end{aligned}
$$

where we used (4.4) in the last line. This proves (4.1). The inequalities (4.2) and (4.3) can be shown similarly. □

Now we are ready to give the proof of Theorem 2.1.

***Proof of Theorem 2.1.*** Let $z\geqslant 0$ be a fixed real number, and $f_z$ the solution to the Stein equation:

$$
f'(w)-wf(w)=\mathbf{1}_{\{w\leqslant z\}}-\Phi(z), \tag{4.5}
$$

where $\Phi(\cdot)$ is the distribution function of the standard normal distribution. It is well known that (see, e.g., Chen, Goldstein and Shao [11]) the solution to (4.5) is given by

$$
f_z(w)=\begin{cases}\dfrac{\Phi(w)\{1-\Phi(z)\}}{p(w)}, & w\leqslant z,\\[2ex] \dfrac{\Phi(z)\{1-\Phi(w)\}}{p(w)}, & w>z,\end{cases} \tag{4.6}
$$

where $p(w)=(2\pi)^{-1/2}\,\mathrm{e}^{-w^2/2}$ is the density function of the standard normal distribution.

By (2.1),

$$
\mathbb{E}\{Wf_z(W)\}=\frac{1}{2\lambda}\,\mathbb{E}\left\{D\int_{-\Delta}^{0} f_z'(W+t)\,\mathrm{d}t\right\}-\mathbb{E}\{Rf_z(W)\},
$$

and thus,

$$
\mathbb{P}(W>z)-\{1-\Phi(z)\}=\mathbb{E}\{f_z'(W)-Wf_z(W)\}:=J_1-J_2+J_3, \tag{4.7}
$$

where

$$
\begin{aligned}
J_1&=\mathbb{E}\left\{f_z'(W)\left(1-\frac{1}{2\lambda}\,\mathbb{E}\left\{D\Delta\,|\,W\right\}\right)\right\},\\
J_2&=\frac{1}{2\lambda}\,\mathbb{E}\left\{D\int_{-\Delta}^{0}\big(f_z'(W+u)-f_z'(W)\big)\,\mathrm{d}u\right\},\\
J_3&=\mathbb{E}\{Rf_z(W)\}.
\end{aligned}
$$

Without loss of generality, we only consider $J_2$, because $J_1$ and $J_3$ can be bounded similarly.

For $J_2$, observe that $f_z'(w) = wf(w) - \mathbf{1}_{\{w>z\}} + \{1-\Phi(z)\}$, and both $wf_z(w)$ and $\mathbf{1}_{\{w>z\}}$ are increasing functions (see, e.g. Chen, Goldstein and Shao [11, Lemma 2.3]), by Lemma 4.1,

$$\begin{aligned}|J_2| &\leqslant \frac{1}{2\lambda}\Big|\mathbb{E}\Big[D\int_{-\Delta}^{0}\{(W+u)f_z(W+u) - Wf_z'(W)\}\,\mathrm{d}u\Big]\Big| \\ &\quad + \frac{1}{2\lambda}\Big|\mathbb{E}\Big[D\int_{-\Delta}^{0}\Big\{\mathbf{1}_{\{W+u>z\}} - \mathbf{1}_{\{W>z\}}\Big\}\,\mathrm{d}u\Big]\Big| \\ &\leqslant \frac{1}{2\lambda}\,\mathbb{E}\big|\mathbb{E}\,\{D^*\Delta\,|\,W\}\big|\big(|Wf_z(W)| + \mathbf{1}_{\{W>z\}}\big) \coloneqq J_{21} + J_{22},\end{aligned} \tag{4.8}$$

where

$$\begin{aligned}J_{21} &= \frac{1}{2\lambda}\,\mathbb{E}\Big\{\big|\mathbb{E}\,\{D^*\Delta\,|\,W\}\big|\cdot\big|Wf_z(W)\big|\Big\}, \\ J_{22} &= \frac{1}{2\lambda}\,\mathbb{E}\Big\{\big|\mathbb{E}\,\{D^*\Delta\,|\,W\}\big|\mathbf{1}_{\{W>z\}}\Big\}.\end{aligned}$$

For any $w>0$, it is well known that

$$\frac{1-\Phi(w)}{p(w)} \leqslant \min\left\{\frac{1}{w}, \frac{\sqrt{2\pi}}{2}\right\}.$$

Then, for $w>z$,

$$\big|f_z(w)\big| \leqslant \frac{\sqrt{2\pi}}{2}\Phi(z), \quad \big|wf_z(w)\big| \leqslant \Phi(z), \tag{4.9}$$

and by symmetry, for $w<0$,

$$\big|f_z(w)\big| \leqslant \frac{\sqrt{2\pi}}{2}\{1-\Phi(z)\}, \quad \big|wf_z(w)\big| \leqslant 1-\Phi(z). \tag{4.10}$$

For $J_{21}$, by (4.6), (4.9) and (4.10), we have

$$\begin{aligned}J_{21} &\leqslant \frac{1}{2\lambda}\{1-\Phi(z)\}\,\mathbb{E}\Big\{\big|\mathbb{E}\,\{D^*\Delta\,|\,W\}\big|\mathbf{1}_{\{W<0\}}\Big\} \\ &\quad + \frac{\sqrt{2\pi}}{2\lambda}\{1-\Phi(z)\}\,\mathbb{E}\Big\{\big|\mathbb{E}\,\{D^*\Delta\,|\,W\}\big|W\,\mathrm{e}^{W^2/2}\,\mathbf{1}_{\{0\leqslant W\leqslant z\}}\Big\} \\ &\quad + \frac{1}{2\lambda}\,\mathbb{E}\Big\{\big|\mathbb{E}\,\{D^*\Delta\,|\,W\}\big|\mathbf{1}_{\{W>z\}}\Big\}.\end{aligned} \tag{4.11}$$

Thus, by (4.8) and (4.11),

$$\begin{aligned}|J_2| &\leqslant \frac{1}{2\lambda}\{1-\Phi(z)\}\,\mathbb{E}\Big\{\big|\mathbb{E}\,\{D^*\Delta\,|\,W\}\big|\mathbf{1}_{\{W<0\}}\Big\} \\ &\quad + \frac{\sqrt{2\pi}}{2\lambda}\{1-\Phi(z)\}\,\mathbb{E}\Big\{\big|\mathbb{E}\,\{D^*\Delta\,|\,W\}\big|W\,\mathrm{e}^{W^2/2}\,\mathbf{1}_{\{0\leqslant W\leqslant z\}}\Big\} \\ &\quad + \frac{1}{\lambda}\,\mathbb{E}\Big\{\big|\mathbb{E}\,\{D^*\Delta\,|\,W\}\big|\mathbf{1}_{\{W>z\}}\Big\}.\end{aligned} \tag{4.12}$$

For the first term of (4.12), by condition (A2) with $t = 0$, and noting that $\delta_2(\cdot)$ is increasing,

$$\frac{1}{2\lambda}\,\mathbb{E}\Big\{\big|\mathbb{E}\left\{D^*\Delta \,|\, W\right\}\big|\mathbf{1}_{\{W<0\}}\Big\} \leqslant \frac{1}{2\lambda}\,\mathbb{E}\Big\{\big|\mathbb{E}\left\{D^*\Delta \,|\, W\right\}\big|\Big\} \leqslant \delta_2(z). \tag{4.13}$$

For the second term of (4.12), by Lemma 4.2, we have

$$\frac{\sqrt{2\pi}}{2\lambda}\,\mathbb{E}\Big\{\big|\mathbb{E}\left\{D^*\Delta \,|\, W\right\}\big| W\,\mathrm{e}^{W^2/2}\,\mathbf{1}_{\{0\leqslant W\leqslant z\}}\Big\} \leqslant 4\sqrt{2\pi}(1+z^2)\delta_2(z). \tag{4.14}$$

It is well known that for $z > 0$,

$$\mathrm{e}^{-z^2/2} \leqslant \sqrt{2\pi}(1+z)\big\{1-\Phi(z)\big\} \leqslant \frac{3\sqrt{2\pi}}{2}(1+z^2)\big\{1-\Phi(z)\big\}.$$

For the third term of (4.12), by condition (A2), for $0 \leqslant z \leqslant \tau_0$,

$$\begin{aligned}\frac{1}{\lambda}\,\mathbb{E}\Big\{\big|\mathbb{E}\left\{D^*\Delta \,|\, W\right\}\big|\mathbf{1}_{\{W>z\}}\Big\} &\leqslant 2\delta_2(z)\,\mathrm{e}^{-z^2/2}\\ &\leqslant 3\sqrt{2\pi}(1+z^2)\delta_2(z)\big\{1-\Phi(z)\big\}.\end{aligned} \tag{4.15}$$

Therefore, combining (4.12)–(4.15), for $0 \leqslant z \leqslant \tau_0$, we have

$$|J_2| \leqslant (7\sqrt{2\pi}+1)(1+z^2)\delta_2(z)\big(1-\Phi(z)\big) \leqslant 20(1+z^2)\delta_2(z)\big(1-\Phi(z)\big).$$

Similarly,

$$|J_1| \leqslant 20(1+z^2)\delta_1(z)\big(1-\Phi(z)\big), \qquad |J_3| \leqslant 20(1+z)\delta_3(z)\big(1-\Phi(z)\big).$$

This completes the proof together with (4.7). □

To prove Corollary 2.2, we need to prove the following lemma, which provides a bound for the moment generating function of $W$.

**Lemma 4.3.** *Under the conditions of Corollary 2.2, for $0 \leqslant t \leqslant \tau$, we have*

$$\mathbb{E}\,\mathrm{e}^{tW} \leqslant \exp\Big\{\frac{t^2}{2}\big(1+\delta_1(t)+\delta_2(t)\big)+\delta_3(t)t\Big\}. \tag{4.16}$$

*For $\theta > 0$, let*

$$\tau_0(\theta) := \max\left\{0 \leqslant t \leqslant \tau : \frac{t^2}{2}\left\{\delta_1(t)+\delta_2(t)\right\}+\delta_3(t)t \leqslant \theta\right\}.$$

*Then, for $0 \leqslant t \leqslant \tau_0(\theta)$,*

$$\mathbb{E}\,\mathrm{e}^{tW} \leqslant \mathrm{e}^{t^2/2+\theta}. \tag{4.17}$$

***Proof of Lemma 4.3.*** We first prove (4.16). Let $h(t) = \mathbb{E}\,\mathrm{e}^{tW}$. Since $\mathbb{E}\,\mathrm{e}^{tW} < \infty$, and by the continuity of the exponential function, we have $h'(t) = \mathbb{E}\{W\,\mathrm{e}^{tW}\}$. Therefore,

$$\begin{aligned} h'(t) &= \frac{t}{2\lambda}\,\mathbb{E}\left\{D\int_{-\Delta}^{0}\mathrm{e}^{t(W+u)}\,\mathrm{d}u\right\} - \mathbb{E}\{R\,\mathrm{e}^{tW}\} \\ &\leqslant t\,\mathbb{E}\{\mathrm{e}^{tW}\} + \frac{t}{2\lambda}\,\mathbb{E}\left\{D\int_{-\Delta}^{0}\{\mathrm{e}^{t(W+u)} - \mathrm{e}^{tW}\}\,\mathrm{d}u\right\} \\ &\quad + t\,\mathbb{E}\left\{\left|1 - \frac{1}{2\lambda}\,\mathbb{E}\{D\Delta\,|\,W\}\right|\mathrm{e}^{tW}\right\} + \mathbb{E}\{|R|\,\mathrm{e}^{tW}\}. \end{aligned} \tag{4.18}$$

By condition (B3) and Lemma 4.1, we have for $0 \leqslant t \leqslant \tau$,

$$\frac{t}{2\lambda}\left|\mathbb{E}\left\{D\int_{-\Delta}^{0}\left(\mathrm{e}^{t(W+u)} - \mathrm{e}^{tW}\right)\mathrm{d}u\right\}\right| \leqslant t\delta_2(t)\,\mathbb{E}\,\mathrm{e}^{tW}. \tag{4.19}$$

By conditions (B2) and (B4), for $0 \leqslant t \leqslant \tau$,

$$t\,\mathbb{E}\left\{\left|1 - \frac{1}{2\lambda}\,\mathbb{E}\{D\Delta\,|\,W\}\right|\mathrm{e}^{tW}\right\} \leqslant t\delta_1(t)\,\mathbb{E}\,\mathrm{e}^{tW}, \qquad \mathbb{E}\{|R|\,\mathrm{e}^{tW}\} \leqslant \delta_3(t)\,\mathbb{E}\,\mathrm{e}^{tW}. \tag{4.20}$$

Combining (4.18)–(4.20), we have for $0 \leqslant t \leqslant \tau$,

$$h'(t) = \mathbb{E}\{W\,\mathrm{e}^{tW}\} \leqslant th(t) + \left\{t\left(\delta_1(t) + \delta_2(t)\right) + \delta_3(t)\right\}h(t).$$

Noting that $h(0) = 1$, and $\delta_1, \delta_2$ and $\delta_3$ are increasing, we complete the proof of (4.16) by solving the foregoing differential inequality. The inequality (4.17) follows from (4.16). □

***Proof of Corollary 2.2.*** By Lemma 4.3 and conditions (B2)–(B4), we have for $0 \leqslant t \leqslant \tau_0(\theta)$,

$$\begin{aligned} \mathbb{E}\left\{\left|1 - \frac{1}{2\lambda}\,\mathbb{E}\{D\Delta\,|\,X\}\right|\mathrm{e}^{tW}\right\} &\leqslant \mathrm{e}^{\theta}\,\delta_1(t)\,\mathrm{e}^{t^2/2}, \\ \mathbb{E}\left\{\left|\frac{1}{2\lambda}\,\mathbb{E}\{D^*\Delta\,|\,X\}\right|\mathrm{e}^{tW}\right\} &\leqslant \mathrm{e}^{\theta}\,\delta_2(t)\,\mathrm{e}^{t^2/2}, \\ \mathbb{E}\{|R|\,\mathrm{e}^{tW}\} &\leqslant \mathrm{e}^{\theta}\,\delta_3(t)\,\mathrm{e}^{t^2/2}. \end{aligned}$$

Then Corollary 2.2 follows from Theorem 2.1 by taking $\tau_0 = \tau_0(\theta)$. □

# 5. Proofs of other results

## 5.1. Proof of Theorem 3.1

We use Corollary 2.2 to prove this theorem. To this end, we need to construct an exchangeable pair and check the condition (D1). In order to construct the exchangeable pair, we first introduce some notations.

Let $\xi' = \{\xi'_i, i \in \mathcal{J}\}$ be an independent copy of $\{\xi_i, i \in \mathcal{J}\}$. For each $i \in \mathcal{J}$, define $\xi^{(i)} = \{\xi_j^{(i)}, j \in \mathcal{J}\}$ where

$$\xi_j^{(i)} = \begin{cases} \xi'_j, & \text{if } j \in J_i, \\ \xi_j, & \text{if } j \in \mathcal{J} \setminus J_i. \end{cases}$$

Let $I$ be a random index that is uniformly distributed over $\{1, 2, \ldots, n\}$ and independent of all other random variables. Then $(\xi, \xi^{(I)})$ is an exchangeable pair. For any $J \subset \mathcal{J}$, write $\xi_J^{(i)} = \{\xi_j^{(i)}, j \in \mathcal{J}\}$. Let $\mathcal{A}_i = \{j : J_i \cap J_j \neq \varnothing\}$, and

$$X_j^{(i)} = f_j(\xi_{J_j}^{(i)}), \quad D = X_I - X_I^{(I)}, \quad W' = \sum_{j=1}^n X_j^{(I)}, \quad \Delta = W - W' = \sum_{j \in \mathcal{A}_I} \bigl(X_j - X_j^{(I)}\bigr).$$

Let $\mathcal{F} = \sigma(\xi_i : i \in \mathcal{J})$ and $\mathcal{F}' = \sigma(\xi'_i : i \in \mathcal{J})$, and let $\mathcal{F} \vee \mathcal{F}'$ be the smallest $\sigma$-field containing $\mathcal{F}$ and $\mathcal{F}'$. Observe that for each $1 \leqslant i \leqslant n$, $X_i^{(i)}$ is independent of $\mathcal{F}$ and $\mathbb{E}\{X_i\} = 0$,

$$\mathbb{E}\{D \,|\, \mathcal{F}\} = \frac{1}{n} \sum_{i=1}^n \mathbb{E}\bigl\{X_i - X_i^{(i)} \bigm| \mathcal{F}\bigr\} = \frac{1}{n} W.$$

Then the condition (D1) is satisfied with $\lambda = 1/n$ and $R = 0$.

We now check conditions (B1)–(B4). By (2.3) with $f(w) = w$ and the assumption that $\mathbb{E}\{W^2\} = 1$,

$$\mathbb{E}\{D\Delta\} = \mathbb{E}\{D(W - W')\} = 2\,\mathbb{E}\{DW\} = 2\lambda\,\mathbb{E}\{W^2\} = 2\lambda.$$

Moreover,

$$\frac{1}{2\lambda} \mathbb{E}\{D\Delta \,|\, \mathcal{F} \vee \mathcal{F}'\} - 1 = \frac{1}{2} \sum_{i=1}^n \sum_{j \in \mathcal{A}_i} \bigl\{\bigl(X_i - X_i^{(i)}\bigr)\bigl(X_j - X_j^{(i)}\bigr) - \mathbb{E}\bigl(X_i - X_i^{(i)}\bigr)\bigl(X_j - X_j^{(i)}\bigr)\bigr\},$$

and with $D^* = |D|$,

$$\frac{1}{\lambda} \mathbb{E}\{D^*\Delta \,|\, \mathcal{F} \vee \mathcal{F}'\} = \sum_{i=1}^n \sum_{j \in \mathcal{A}_i} \bigl|X_i - X_i^{(i)}\bigr|\bigl(X_j - X_j^{(i)}\bigr).$$

Let $\chi_{ij} = \bigl\{\bigl(X_i - X_i^{(i)}\bigr)\bigl(X_j - X_j^{(i)}\bigr) - \mathbb{E}\bigl(X_i - X_i^{(i)}\bigr)\bigl(X_j - X_j^{(i)}\bigr)\bigr\}$ and $\zeta_{ij} = \bigl|X_i - X_i^{(i)}\bigr|\bigl(X_j - X_j^{(i)}\bigr)$. Let $\delta(t)$ be as in (3.2). We have the following proposition.

**Proposition 5.1.** *For $0 \leqslant t \leqslant \alpha$, we have $\mathbb{E}\,\mathrm{e}^{tW} < \infty$,*

$$\mathbb{E}\biggl\{\biggl|\sum_{i=1}^n \sum_{j \in \mathcal{A}_i} \chi_{ij}\biggr| \mathrm{e}^{tW}\biggr\} \leqslant \delta(t)\,\mathbb{E}\,\mathrm{e}^{tW}, \tag{5.1}$$

*and*

$$\mathbb{E}\bigg\{\bigg|\sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\zeta_{ij}\bigg|\,\mathrm{e}^{tW}\bigg\}\leqslant\delta(t)\,\mathbb{E}\,\mathrm{e}^{tW}. \tag{5.2}$$

Applying Corollary 2.2, we complete the proof of Theorem 3.1 by Proposition 5.1. Now, it suffices to prove Proposition 5.1. To this end, we first give some preliminary lemmas. Let $\mathcal{A}_{ij}=\mathcal{A}_i\cup\mathcal{A}_j$, $\mathcal{A}_{ijk}=\mathcal{A}_i\cup\mathcal{A}_j\cup\mathcal{A}_k$, and

$$W_{ij}=\sum_{l\in\mathcal{A}_{ij}}X_l,\quad W_{ij}^c=W-W_{ij},\quad W_{ijk}=\sum_{l\in\mathcal{A}_{ijk}}X_l,\quad W_{ijk}^c=W-W_{ijk}.$$

**Lemma 5.2.** *We have for* $0\leqslant t\leqslant\alpha$,

$$\mathbb{E}\Big|\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\Big|\leqslant 16\beta^{6d}\big(\gamma_{4,i}(t)+\gamma_{4,j}(t)+\gamma_{4,i'}(t)+\gamma_{4,j'}(t)\big)\,\mathbb{E}\,\mathrm{e}^{tW}, \tag{5.3}$$

*and*

$$\mathbb{E}\Big|\zeta_{ij}\zeta_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\Big|\leqslant 8\beta^{6d}\big(\gamma_{4,i}(t)+\gamma_{4,j}(t)+\gamma_{4,i'}(t)+\gamma_{4,j'}(t)\big)\,\mathbb{E}\,\mathrm{e}^{tW}. \tag{5.4}$$

**Lemma 5.3.** *For any* $i,j,i',j',k$ *and* $l$, *we have for* $0\leqslant t\leqslant\alpha$,

$$\mathbb{E}\big|X_kX_l\xi_{ij}\xi_{i'j'}e^{tW_{ijk}^c}\big|\leqslant 88\beta^{7d}\bigg(\sum_{m\in\{i,j,i',j',k,l\}}\gamma_{6,m}(t)\bigg)\,\mathbb{E}\,e^{tW},$$

$$\mathbb{E}\big|X_kX_l\xi_{ij}\xi_{i'j'}e^{tW_{ij}^c}\big|\leqslant 88\beta^{6d}\bigg(\sum_{m\in\{i,j,i',j',k,l\}}\gamma_{6,m}(t)\bigg)\,\mathbb{E}\,e^{tW}.$$

**Lemma 5.4.** *For any* $i,j,k$ *and* $i',j'\notin\mathcal{A}_{ijk}$, *we have for* $0\leqslant t\leqslant\alpha$,

$$\begin{aligned}&t\big|\mathbb{E}\big\{X_k\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ijk}^c}\big\}\big|\\&\leqslant 176\beta^{9d}\,t^2\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{m\in\mathcal{A}_{i'j'}}\big(\gamma_{3,i}(t)+\gamma_{3,j}(t)+\gamma_{3,k}(t)\big)\big(\gamma_{3,i'}(t)+\gamma_{3,j'}(t)+\gamma_{3,m}(t)\big).\end{aligned} \tag{5.5}$$

*For any* $i,j,i',j'$ *and* $k\in\mathcal{A}_{ij}$ *such that* $\{i',j'\}\cap\mathcal{A}_{ijk}\neq\varnothing$, *we have for* $0\leqslant t\leqslant\alpha$,

$$\begin{aligned}t\big|\mathbb{E}\{X_k\chi_{ij}\chi_{i'j'}e^{tW_{ijk}^c}\}\big|\leqslant{}&\frac{8}{|\mathcal{A}_{ij}|}\beta^{6d}\big\{\gamma_{4,i}(t)+\gamma_{4,j}(t)+\gamma_{4,i'}(t)+\gamma_{4,j'}(t)\big\}\,\mathbb{E}\,\mathrm{e}^{tW}\\&+44t^2\,|\mathcal{A}_{ij}|\beta^{6d}\bigg\{\sum_{m\in\{i,j,i',j',k\}}\gamma_{6,m}(t)\bigg\}\,\mathbb{E}\,\mathrm{e}^{tW}\\&+176t^2\sum_{l\in\mathcal{A}_{ijk}}\beta^{7d}\bigg\{\sum_{m\in\{i,j,i',j',k,l\}}\gamma_{6,m}(t)\bigg\}\,\mathbb{E}\,\mathrm{e}^{tW}.\end{aligned} \tag{5.6}$$

The proofs of Lemmas 5.2–5.4 are postponed to Appendix A.1. Now we are ready to give the proof of Proposition 5.1.

***Proof of Proposition 5.1.*** Let $t \in [0, \alpha]$ be a fixed real number. For fixed $n$, by (3.1), we have $\mathbb{E}\,\mathrm{e}^{tW} < \infty$. It suffices to prove (5.1), because (5.2) can be shown similarly. By the Cauchy inequality,

$$\biggl(\mathbb{E}\biggl\{\biggl|\sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\chi_{ij}\biggr|\,\mathrm{e}^{tW}\biggr\}\biggr)^2 \leqslant \mathbb{E}\,\mathrm{e}^{tW}\,\mathbb{E}\biggl\{\biggl(\sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\chi_{ij}\biggr)^2\mathrm{e}^{tW}\biggr\}. \tag{5.7}$$

Now we bound the square term. Let $\mathcal{A}_{ij} = \mathcal{A}_i \cup \mathcal{A}_j$, and then $\{X_i, X_i^{(i)}, X_j, X_j^{(i)}\}$ is independent of $\{X_j, j \in \mathcal{A}_{ij}^c\}$. Let $W_{ij} = \sum_{k\in\mathcal{A}_{ij}} X_k$ and $W_{ij}^c = W - W_{ij}$. Expanding the square term, we have

$$\mathbb{E}\biggl\{\biggl(\sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\chi_{ij}\biggr)^2\mathrm{e}^{tW}\biggr\} := H_1 + H_2 + H_3, \tag{5.8}$$

where

$$\begin{aligned}
H_1 &= \sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\sum_{i'=1}^{n}\sum_{j'\in\mathcal{A}_{i'}}\mathbb{E}\{\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ij}}\}\\
H_2 &= \sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\sum_{i'=1}^{n}\sum_{j'\in\mathcal{A}_{i'}}\mathbb{E}\{\chi_{ij}\chi_{i'j'}W_{ij}\,\mathrm{e}^{tW_{ij}^c}\}\\
H_3 &= \sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\sum_{i'=1}^{n}\sum_{j'\in\mathcal{A}_{i'}}\mathbb{E}\{\chi_{ij}\chi_{i'j'}\nu(tW_{ij})\,\mathrm{e}^{tW_{ij}^c}\},
\end{aligned}$$

and where $\nu(x) = \mathrm{e}^x - 1 - x$.

For $H_1$, if $i' \in \mathcal{A}_{ij}^c$ and $j' \in \mathcal{A}_{ij}^c$, then $\chi_{ij}$ is independent of $\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}$. Therefore,

$$\begin{aligned}
&\sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\sum_{i'=1}^{n}\sum_{j'\in\mathcal{A}_{i'}}\mathbb{E}\Bigl\{\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\,\mathbf{1}_{\{i'\in\mathcal{A}_{ij}^c, j'\in\mathcal{A}_{ij}^c\}}\Bigr\}\\
&\quad= \sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\sum_{i'=1}^{n}\sum_{j'\in\mathcal{A}_{i'}}\mathbb{E}\,\chi_{ij}\,\mathbb{E}\Bigl\{\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\,\mathbf{1}_{\{i'\in\mathcal{A}_{ij}^c, j'\in\mathcal{A}_{ij}^c\}}\Bigr\} = 0.
\end{aligned}$$

Let $\widetilde{\mathcal{A}}_j = \{i : j \in \mathcal{A}_i\}$. By the definition, we have $\widetilde{\mathcal{A}}_j = \mathcal{A}_j$. Hence,

$$\begin{aligned}
H_1 &= \sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\sum_{i'=1}^{n}\sum_{j'\in\mathcal{A}_{i'}}\mathbb{E}\Bigl\{\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\,\mathbf{1}_{\{i'\in\mathcal{A}_{ij}\text{ or } j'\in\mathcal{A}_{ij}\}}\Bigr\}\\
&\leqslant \sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{A}_{ij}}\sum_{j'\in\mathcal{A}_{i'}}\mathbb{E}\Bigl\{|\chi_{ij}\chi_{i'j'}|\,\mathrm{e}^{tW_{ij}^c}\Bigr\} + \sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\sum_{j'\in\mathcal{A}_{ij}}\sum_{i'\in\mathcal{A}_{j'}}\mathbb{E}\Bigl\{|\chi_{ij}\chi_{i'j'}|\,\mathrm{e}^{tW_{ij}^c}\Bigr\}.
\end{aligned}$$

By Lemma 5.2, and recalling that $|\mathcal{A}_i| \leqslant d$ and $|\mathcal{A}_{ij}| \leqslant 2d$, we have

$$\begin{aligned} H_1 &\leqslant 8\beta^{6d}\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{A}_{ij}}\sum_{j'\in\mathcal{A}_{i'}}\bigl\{\gamma_{4,i}(t)+\gamma_{4,j}(t)+\gamma_{4,i'}(t)+\gamma_{4,j'}(t)\bigr\} \\ &\leqslant 64\beta^{6d}d^3\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i=1}^{n}\gamma_{4,i}(t). \end{aligned} \tag{5.9}$$

Now we move to bound $H_2$. Let $\widetilde{W}_{ijk} = \sum_{l\in\mathcal{A}_{ijk}\setminus\mathcal{A}_{ij}} X_l$. Observe that

$$\begin{aligned} &\mathbb{E}\bigl\{\chi_{ij}\chi_{i'j'}W_{ij}\,\mathrm{e}^{tW_{ij}^c}\bigr\} \\ &\quad= \sum_{k\in\mathcal{A}_{ij}}\mathbb{E}\bigl\{\chi_{ij}\chi_{i'j'}X_k\,\mathrm{e}^{tW_{ij}^c}\bigr\} \\ &\quad= \sum_{k\in\mathcal{A}_{ij}}\mathbb{E}\bigl\{\chi_{ij}\chi_{i'j'}X_k\,\mathrm{e}^{tW_{ijk}^c}\bigr\} + \sum_{k\in\mathcal{A}_{ij}}\mathbb{E}\Bigl\{\chi_{ij}\chi_{i'j'}X_k\,\mathrm{e}^{tW_{ij}^c}\Bigl(\mathrm{e}^{t\widetilde{W}_{ijk}}-1\Bigr)\Bigr\}, \end{aligned}$$

and

$$\bigl|\mathrm{e}^{t\widetilde{W}_{ijk}}-1\bigr| \leqslant \sum_{l\in\mathcal{A}_{ijk}\setminus\mathcal{A}_{ij}} t|X_l|(1+\mathrm{e}^{t\widetilde{W}_{ijk}}).$$

Thus,

$$H_2 \leqslant H_{21}+H_{22}+H_{23},$$

where

$$\begin{aligned} H_{21} &= \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{J}}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_{ij}} t\bigl|\mathbb{E}\bigl\{X_k\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ijk}^c}\bigr\}\bigr|, \\ H_{22} &= \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{J}}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_{ij}}\sum_{l\in\mathcal{A}_{ijk}\setminus\mathcal{A}_{ij}} t^2\,\mathbb{E}\bigl|X_kX_l\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ijk}^c}\bigr|, \\ H_{23} &= \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{J}}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_{ij}}\sum_{l\in\mathcal{A}_{ijk}\setminus\mathcal{A}_{ij}} t^2\,\mathbb{E}\bigl|X_kX_l\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\bigr|. \end{aligned}$$

We now consider $H_{22}$ and $H_{23}$. By Lemma 5.3,

$$\begin{aligned} |H_{22}| &\leqslant 88\,\beta^{7d}t^2\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{J}}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_{ij}}\sum_{l\in\mathcal{A}_{ijk}\setminus\mathcal{A}_{ij}}\Biggl\{\sum_{m\in\{i,j,i',j',k,l\}}\gamma_{6,m}(t)\Biggr\}\mathbb{E}\,\mathrm{e}^{tW} \\ &\leqslant 1056n\beta^{7d}d^4t^2\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i=1}^{n}\gamma_{6,i}(t), \end{aligned} \tag{5.10}$$

and similarly,

$$|H_{23}| \leqslant 1056n\beta^{6d}d^4t^2\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i=1}^{n}\gamma_{6,i}(t). \tag{5.11}$$

Now, it remains to bound $H_{21}$. Observe that

$$H_{21} \leqslant H_{211} + H_{212},$$

where

$$\begin{aligned} H_{211} &= \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{J}}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_{ij}} t\big|\mathbb{E}\big\{X_k\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ijk}^c}\big\}\big|\mathbf{1}_{\{i'\in\mathcal{A}_{ijk}^c, j'\in\mathcal{A}_{ijk}^c\}}, \\ H_{212} &= \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{J}}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_{ij}} t\big|\mathbb{E}\big\{X_k\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ijk}^c}\big\}\big|\mathbf{1}_{\{\{i',j'\}\cap\mathcal{A}_{ijk}\neq\varnothing\}}. \end{aligned}$$

By Lemma 5.4,

$$\begin{aligned} H_{211} &\leqslant 176\beta^{9d}t^2\Big[\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{A}_i}\sum_{k\in\mathcal{A}_{ij}}\big\{\gamma_{3,i}(t)+\gamma_{3,j}(t)+\gamma_{3,k}(t)\big\}\Big]^2\,\mathbb{E}\,\mathrm{e}^{tW} \\ &\leqslant 6336\beta^{9d}d^4t^2\Big\{\sum_{i=1}^{n}\gamma_{3,i}(t)\Big\}^2\,\mathbb{E}\,\mathrm{e}^{tW}, \end{aligned} \tag{5.12}$$

and

$$\begin{aligned} H_{212} &\leqslant 8\beta^{6d}\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{A}_i}\sum_{k\in\mathcal{A}_{ij}}\sum_{i'\in\mathcal{A}_{ijk}}\sum_{j\in\mathcal{A}_{i'}}|\mathcal{A}_{ij}|^{-1}\big\{\gamma_{4,i}(t)+\gamma_{4,j}(t)+\gamma_{4,i'}(t)+\gamma_{4,j'}(t)\big\} \\ &\quad + 220\beta^{7d}\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{J}}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_{ij}}\sum_{l\in\mathcal{A}_{ijk}}\Big\{\sum_{m\in\{i,j,i',j',k,l\}}\gamma_{6,m}(t)\Big\} \\ &\leqslant 192\beta^{6d}d^3\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i\in\mathcal{J}}\gamma_{4,i}(t) + 7920n\beta^{7d}d^4t^2\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i=1}^{n}\gamma_{6,i}(t). \end{aligned} \tag{5.13}$$

By (5.10)–(5.13), we have

$$\begin{aligned} H_2 &\leqslant 6336\beta^{9d}d^4t^2\,\mathbb{E}\,\mathrm{e}^{tW}\Big\{\sum_{i=1}^{n}\gamma_{3,i}(t)\Big\}^2 \\ &\quad + 192\beta^{6d}d^3\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i\in\mathcal{J}}\gamma_{4,i}(t) + 10032n\beta^{7d}d^4t^2\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i=1}^{n}\gamma_{6,i}(t). \end{aligned} \tag{5.14}$$

For $H_3$, by (5.8), noting that $|\nu(x)| \leqslant \frac{1}{2}x^2(1+\mathrm{e}^x)$, similar to $H_{22}$, we have

$$|H_3| \leqslant 2112n\beta^{7d}d^4t^2\,\mathbb{E}\,\mathrm{e}^{tW}\sum_{i=1}^{n}\gamma_{6,i}(t). \tag{5.15}$$

By (5.7), (5.9), (5.14) and (5.15), we have

$$\mathbb{E}\left\{\left|\sum_{i=1}^{n}\sum_{j\in\mathcal{A}_i}\chi_{ij}\right|\mathrm{e}^{tW}\right\}\leqslant 80\beta^{5d}d^2t\,\mathbb{E}\,\mathrm{e}^{tW}\left\{\sum_{i=1}^{n}\gamma_{3,i}(t)\right\}+24\beta^{3d}d^{3/2}\,\mathbb{E}\,\mathrm{e}^{tW}\left\{\sum_{i\in\mathcal{J}}\gamma_{4,i}(t)\right\}^{1/2}$$
$$+112n^{1/2}\beta^{4d}d^2t\,\mathbb{E}\,\mathrm{e}^{tW}\left\{\sum_{i=1}^{n}\gamma_{6,i}(t)\right\}^{1/2}.$$

By the Cauchy inequality,

$$\sum_{i=1}^{n}\gamma_{3,i}(t)\leqslant n^{1/2}\left\{\sum_{i=1}^{n}\gamma_{6,i}(2t)\right\}^{1/2},$$

and this completes the proof. □

## 5.2. Proof of Theorem 3.2

The proof is organized as follows: we first introduce some notation, then construct the exchangeable pair, and finally check the conditions (B1)–(B4).

In this subsection, the constants $C$'s depend only on the fixed graph $G$, which may take different values in different places. Let $N=n(n-1)/2$ and let $\{e_1,\ldots,e_N\}$ be the ordered node pairs in the graph $\mathcal{G}(n,p)$. Define

$$\mathcal{I}_n=\big\{i=\{i_1,\ldots,i_{e(G)}\}:1\leqslant i_1<i_2<\cdots<i_{e(G)}\leqslant N,$$
$$G_i:=\{e_{i_1},\ldots,e_{i_e}\}\text{ is isomorphic to }G\big\}.$$

Let

$$X_i=\frac{1}{\sigma_n}\left(\prod_{l=1}^{e(G)}\varepsilon_{i_l}-p^{e(G)}\right),\quad W_n=\sum_{i\in\mathcal{I}_n}X_i,$$

where $\sigma_n^2=\mathrm{Var}(S_n)$ and $\varepsilon_{i_l}$ is the indicator of the event that the node pair $e_{i_l}$ is connected in $\mathcal{G}(n,p)$. It is known that (see, e.g., Barbour, Karoński and Ruciński [2, p. 132])

$$\sigma_n^2\geqslant C(1-p)n^{2v(G)}p^{2e(G)}\psi_n^{-1}. \tag{5.16}$$

Now we construct $\mathcal{A}_i,\mathcal{A}_{ij}$ and $\mathcal{A}_{ijk}$, which are useful in constructing the exchangeable pair. Let

$$\mathcal{A}_i=\{j\in\mathcal{I}_n:|i\cap j|>0\},\quad i\in\mathcal{I}_n,$$
$$\mathcal{A}_{ij}=\big\{k\in\mathcal{I}_n:\big|k\cap(i\cup j)\big|>0\big\},\quad i\in\mathcal{I}_n,j\in\mathcal{A}_i,$$

and

$$\mathcal{A}_{ijk} = \left\{ l \in \mathcal{I}_n : \left| l \cap (i \cup j \cup k) \right| > 0 \right\}, \quad i \in \mathcal{I}_n, j \in \mathcal{A}_i, k \in \mathcal{A}_{ij}.$$

Here, $|\cdot|$ is the cardinality. It follows that $\mathcal{A}_{ij} = \mathcal{A}_i \cup \mathcal{A}_j$ and $\mathcal{A}_{ijk} = \mathcal{A}_i \cup \mathcal{A}_j \cup \mathcal{A}_k$. Also, $|\mathcal{A}_i| \leqslant Cn^{v-2}$ for each $i \in \mathcal{I}_n$.

We now construct the exchangeable pair for $X = (X_i)_{i \in \mathcal{I}_n}$. Let $\{\varepsilon_l' : 1 \leqslant l \leqslant N\}$ be an independent copy of $\{\varepsilon_l : 1 \leqslant l \leqslant N\}$. For each $i = \left\{i_1, \dots, i_{e(G)}\right\} \in \mathcal{I}_n$, define $X^{(i)} = (X_j^{(i)})_{j \in \mathcal{I}_n}$, where

$$X_j^{(i)} = \begin{cases} \frac{1}{\sigma_n} \left( \prod_{l=1}^{e(G)} \varepsilon_{i_l}' - p^{e(G)} \right) & \text{if } j = i, \\ \frac{1}{\sigma_n} \left( \prod_{k \in i \cap j} \varepsilon_k' \prod_{l \in j \cap i^c} \varepsilon_l - p^{e(G)} \right) & \text{if } j \in \mathcal{A}_i, \\ X_j & \text{otherwise.} \end{cases}$$

Let $I$ be random index uniformly distributed over $\mathcal{I}_n$ which is independent of all others. Then, $(X, X^{(I)})$ is an exchangeable pair. Let

$$W^{(I)} = \sum_{j \notin \mathcal{A}_I} X_j + \sum_{j \in \mathcal{A}_I} X_j^{(I)}, \quad D = X_I - X_I^{(I)}, \quad \Delta = W - W^{(I)} = \sum_{j \in \mathcal{A}_I} (X_j - X_j^{(I)}).$$

Then, $(W, W^{(I)})$ is also an exchangeable pair and $D$ is antisymmetric with respect to $X$ and $X^{(I)}$. Let $\mathcal{F} = \sigma\left\{\varepsilon_i, 1 \leqslant i \leqslant N\right\}$. It follows that $\mathbb{E}\left\{D \,|\, \mathcal{F}\right\} = W/|\mathcal{I}_n|$. This implies that condition (D1) is satisfied with $\lambda = 1/|\mathcal{I}_n|$ and $R = 0$.

Now, we move to check conditions (B1)–(B4). Note that by (2.1) with $f(w) = w$ and recall that $\mathbb{E} W^2 = 1$, it follows that $\mathbb{E}\{D\Delta\} = 2\lambda$. Moreover,

$$\begin{aligned} &\mathbb{E}\left\{ (X_i - X_i^{(i)})(X_j - X_j^{(i)}) \,\middle|\, \mathcal{F} \right\} \\ &= \frac{1}{\sigma_n^2}\left(1 - p^{|i \cap j|}\right) \prod_{k \in i \cup j} \varepsilon_k - \frac{1}{\sigma_n^2} p^{e(G)} \prod_{k \in j} \varepsilon_k + \frac{1}{\sigma_n^2} p^{e(G) + |i \cap j|} \prod_{k \in j \cap i^c} \varepsilon_k := \nu_{ij}, \end{aligned}$$

and

$$\mathbb{E}\left\{ (X_i - X_i^{(i)})(X_j - X_j^{(i)}) \right\} = \frac{1}{\sigma_n^2} p^{|i \cup j|}\left(1 - p^{|i \cap j|}\right) := \bar{\nu}_{ij}.$$

Also, with

$$\mu_{ij} := \mathbb{E}\left\{ |X_i - X_i^{(i)}|(X_j - X_j^{(i)}) \,|\, \mathcal{F} \right\},$$

we have $\mathbb{E}\, \mu_{ij} = 0$ by exchangeability. Then,

$$\frac{1}{2\lambda} \mathbb{E}\left\{ D\Delta \,|\, \mathcal{F} \right\} - 1 = \frac{1}{2} \sum_{i \in \mathcal{I}_n} \sum_{j \in \mathcal{A}_i} \left( \nu_{ij} - \bar{\nu}_{ij} \right), \qquad \frac{1}{\lambda} \mathbb{E}\left\{ |D|\Delta \,|\, \mathcal{F} \right\} = \sum_{i \in \mathcal{I}_n} \sum_{j \in \mathcal{A}_i} \mu_{ij}.$$

We have the following proposition.

**Proposition 5.5.** *For $0 \leqslant t \leqslant (1-p)^{1/2} n^2 p^{e(G)} \psi_n^{-1/2}$,*

$$\mathbb{E}\bigg\{\bigg(\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\big(\nu_{ij}-\bar{\nu}_{ij}\big)\bigg)^2 \mathrm{e}^{tW}\bigg\} \leqslant \begin{cases} C\psi_n^{-1}(1+t^2)\,\mathbb{E}\,\mathrm{e}^{tW}, & \text{if } 0<p\leqslant 1/2,\\ \dfrac{C}{n^2(1-p)}\bigg(1+\dfrac{t^2}{1-p}\bigg)\mathbb{E}\,\mathrm{e}^{tW}, & \text{if } 1/2<p<1,\end{cases} \tag{5.17}$$

*and*

$$\mathbb{E}\bigg\{\bigg(\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\mu_{ij}\bigg)^2 \mathrm{e}^{tW}\bigg\} \leqslant \begin{cases} C\psi_n^{-1}(1+t^2)\,\mathbb{E}\,\mathrm{e}^{tW}, & \text{if } 0<p\leqslant 1/2,\\ \dfrac{C}{n^2(1-p)}\bigg(1+\dfrac{t^2}{1-p}\bigg)\mathbb{E}\,\mathrm{e}^{tW}, & \text{if } 1/2<p<1,\end{cases} \tag{5.18}$$

*where $C$ is a constant depending only on the fixed graph $G$.*

Note that for fixed $n$, we have $|W_n| \leqslant n^{v(G)}/\sigma_n$, and then $\mathbb{E}\,\mathrm{e}^{tW_n} < \infty$. By Proposition 5.5, conditions (B1)–(B4) are satisfied with

$$\delta_1(t) = \delta_2(t) = \begin{cases} C\psi_n^{-1/2}(1+t), & 0<p\leqslant 1/2,\\ Cn^{-1}(1-p)^{-1/2}\big\{1+(1-p)^{-1/2}t\big\}, & 1/2<p<1.\end{cases}$$

Applying Corollary 2.2 yields the moderate deviation (3.3), as desired.

Now it suffices to prove Proposition 5.5. We first prove some preliminary lemmas, which will be used in the proof of Proposition 5.5.

**Lemma 5.6.** *We have*

$$\begin{aligned}&\mathbb{E}\left\{|X_i - X_i^{(i)}||X_j - X_j^{(i)}|\,\Big|\,\mathcal{F}\right\}\\ &\quad\leqslant \frac{C}{\sigma_n^2}\min\bigg\{\big(1+p^{|i\cap j|}\big)\prod_{k\in i\cup j}\varepsilon_k + p^{e(G)}\prod_{k\in j}\varepsilon_k + p^{e(G)+|i\cap j|}\prod_{k\in j\cap i^c}\varepsilon_k,\\ &\qquad\qquad\qquad\bigg|1-\prod_{k\in i}\varepsilon_k\bigg| + \mathbb{E}\bigg|1-\prod_{k\in i}\varepsilon_k\bigg|\bigg\}.\end{aligned}$$

***Proof.*** On one hand,

$$\begin{aligned}&\mathbb{E}\left\{|X_i - X_i^{(i)}||X_j - X_j^{(i)}|\,\Big|\,\mathcal{F}\right\}\\ &\quad\leqslant \frac{1}{\sigma_n^2}\bigg\{\big(1+p^{|i\cap j|}\big)\prod_{k\in i\cup j}\varepsilon_k + p^{e(G)}\prod_{k\in j}\varepsilon_k + p^{e+|i\cap j|}\prod_{k\in j\cap i^c}\varepsilon_k\bigg\}.\end{aligned} \tag{5.19}$$

On the other hand, by Barbour, Karoński and Ruciński [2, p. 132], we have $\sigma_n|X_i| \leqslant 1$ and

$$\sigma_n|X_i| \leqslant \bigg|1-\prod_{k\in i}\varepsilon_k\bigg| + \mathbb{E}\bigg|1-\prod_{k\in i}\varepsilon_k\bigg|.$$

Thus,

$$\mathbb{E}\left\{|X_i - X_i^{(i)}||X_j - X_j^{(i)}|\,\Big|\,\mathcal{F}\right\} \leqslant \frac{C}{\sigma_n^2}\Big(\Big|1-\prod_{k\in i}\varepsilon_k\Big| + \mathbb{E}\Big|1-\prod_{k\in i}\varepsilon_k\Big|\Big). \tag{5.20}$$

Combining (5.19) and (5.20) yields the desired result. □

**Lemma 5.7.** *For $1 < p < 1/2$, we have*

$$\sum_{i\in\mathcal{I}_n}\sum_{j:|i\cap j|\geqslant 1}\sum_{i':|i'\cap(i\cup j)|\geqslant 1} p^{3e(G)-|i\cap j|-|i'\cap(i\cup j)|} \leqslant C\sigma_n^2\psi_n^{-1}n^{v(G)}p^{e(G)}, \tag{5.21}$$

$$\sum_{i\in\mathcal{I}_n}\sum_{j:|i\cap j|\geqslant 1}\sum_{i':|i'\cap(i\cup j)|\geqslant 1}\sum_{j':|j'\cap i'|\geqslant 1} p^{4e(G)-|i\cap j|-|i'\cap j'|-|i'\cap(i\cap j)|} \tag{5.22}$$
$$\leqslant C\sigma_n^2(\psi_n^{-1}n^{v(G)}p^{e(G)})^2,$$

$$\sum_{\substack{i\in\mathcal{I}_n\\ j:|i\cap j|\geqslant 1}}\sum_{\substack{i':|i'\cap(i\cup j)|\geqslant 1\\ j':|j'\cap i'|\geqslant 1}}\sum_{k:|k\cap(i'\cup j')|\geqslant 1} p^{5e(G)-|i\cap j|-|i'\cap j'|-|i'\cap(i\cap j)|-|k\cap(i'\cup j')|} \tag{5.23}$$
$$\leqslant C\sigma_n^2(\psi_n^{-1}n^{v(G)}p^{e(G)})^3,$$

*and*

$$\sum_{\substack{i\in\mathcal{I}_n\\ j:|i\cap j|\geqslant 1}}\sum_{\substack{i'\in\mathcal{I}_n\\ j':|j'\cap i'|\geqslant 1}} p^{6e(G)-|i\cap j|-|i'\cap j'|}\Bigg(\sum_{k:|k\cap(i\cup j\cup i'\cup j')|\geqslant 1} p^{-|k\cap(i\cup j\cup i'\cup j')|}\Bigg)^2$$
$$\leqslant C\sigma_n^4(\psi_n^{-1}n^{v(G)}p^{e(G)})^2. \tag{5.24}$$

The proof of Lemma 5.7 is given in Appendix A.2.

***Proof of Proposition 5.5.*** Without loss of generality, we only prove (5.17), because (5.18) can be shown similarly. The proof is organized as follows: we first introduce some notation, then expand the left hand side of (5.17) into several terms, and after that, we give the bound of each term separately.

Now we introduce some notation. For any $i, j, i', j', k, q \in \mathcal{I}_n$, write

$$W_{iji'j'} = \sum_{l\in\mathcal{I}_n} X_l\mathbf{1}_{\{l\in\mathcal{A}_{i,j}\cup\mathcal{A}_{i',j'}\}}, \qquad W^c_{iji'j'} = \sum_{l\in\mathcal{I}_n} X_l\mathbf{1}_{\{l\notin\mathcal{A}_{i,j}\cup\mathcal{A}_{i',j'}\}},$$

$$W_{iji'j'k} = \sum_{l\in\mathcal{I}_n} X_l\mathbf{1}_{\{l\in\mathcal{A}_{i,j}\cup\mathcal{A}_{i',j'}\cup\mathcal{A}_k\}}, \qquad W^c_{iji'j'k} = \sum_{l\in\mathcal{I}_n} X_l\mathbf{1}_{\{l\notin\mathcal{A}_{i,j}\cup\mathcal{A}_{i',j'}\cup\mathcal{A}_k\}},$$

$$W_{iji'j'kq} = \sum_{l\in\mathcal{I}_n} X_l\mathbf{1}_{\{l\in\mathcal{A}_{i,j}\cup\mathcal{A}_{i',j'}\cup\mathcal{A}_k\cap\mathcal{A}_q\}}, \quad W^c_{iji'j'kq} = \sum_{l\in\mathcal{I}_n} X_l\mathbf{1}_{\{l\notin\mathcal{A}_{i,j}\cup\mathcal{A}_{i',j'}\cup\mathcal{A}_k\cap\mathcal{A}_q\}}.$$

For any $\mathcal{T} \subset \mathcal{I}_n$, write $W_{\mathcal{T}} = \sum_{j \in \mathcal{T}} X_j$. Note that $|X_j| \leqslant \sigma_n^{-1}$ for each $j \in \mathcal{I}_n$, it follows that $|W - W_{\mathcal{T}}| \leqslant \sigma_n^{-1} |\mathcal{I}_n \setminus \mathcal{T}|$, a.s., and thus

$$\mathrm{e}^{tW} = \mathrm{e}^{tW_{\mathcal{T}}} \times \mathrm{e}^{tW - tW_{\mathcal{T}}} \geqslant \mathrm{e}^{-t|\mathcal{I}_n \setminus \mathcal{T}|\sigma_n^{-1}} \mathrm{e}^{tW_{\mathcal{T}}}.$$

Recalling that $\sigma_n \geqslant C(1-p)^{1/2} n^{v(G)} p^{e(G)} \psi_n^{-1/2}$ and $|\mathcal{A}_i| \leqslant C n^{v(G)-2}$, then, for $0 \leqslant t \leqslant (1-p)^{1/2} n^2 p^{e(G)} \psi_n^{-1/2}$, we have $t|\mathcal{A}_i|\sigma_n^{-1} \leqslant C$ and

$$\max\{\mathrm{e}^{tW_{iji'j'}}, \mathrm{e}^{tW_{iji'j'k}}, \mathrm{e}^{tW_{iji'j'kq}}\} \leqslant C\,\mathrm{e}^{tW}. \tag{5.25}$$

It is well known that

$$|\mathrm{e}^x - 1 - x| \leqslant \frac{1}{2} x^2 (1 + \mathrm{e}^x), \text{ for } x \in \mathbb{R}. \tag{5.26}$$

Expanding the squared term and by (5.26), we have

$$\begin{aligned}\mathbb{E}\Big\{\Big(\sum_{i \in \mathcal{I}_n} \sum_{j \in \mathcal{A}_i} (\nu_{ij} - \bar{\nu}_{ij})\Big)^2 \mathrm{e}^{tW}\Big\} &= \sum_{i \in \mathcal{I}_n} \sum_{j \in \mathcal{A}_i} \sum_{i' \in \mathcal{I}_n} \sum_{j' \in \mathcal{A}_{i'}} \mathbb{E}\{(\nu_{ij} - \bar{\nu}_{ij})(\nu_{i'j'} - \bar{\nu}_{i'j'})\,\mathrm{e}^{tW}\}\\ &\leqslant Q_1 + Q_2 + Q_3 + Q_4,\end{aligned}$$

where

$$\begin{aligned}
Q_1 &= \sum_{i \in \mathcal{I}_n} \sum_{j \in \mathcal{A}_i} \sum_{i' \in \mathcal{I}_n} \sum_{j' \in \mathcal{A}_{i'}} \big|\mathbb{E}\{(\nu_{ij} - \bar{\nu}_{ij})(\nu_{i'j'} - \bar{\nu}_{i'j'})\,\mathrm{e}^{tW^c_{iji'j'}}\}\big|,\\
Q_2 &= \sum_{i \in \mathcal{I}_n} \sum_{j \in \mathcal{A}_i} \sum_{i' \in \mathcal{I}_n} \sum_{j' \in \mathcal{A}_{i'}} t\big|\mathbb{E}\{W_{iji'j'}(\nu_{ij} - \bar{\nu}_{ij})(\nu_{i'j'} - \bar{\nu}_{i'j'})\,\mathrm{e}^{tW^c_{iji'j'}}\}\big|,\\
Q_3 &= \frac{1}{2} \sum_{i \in \mathcal{I}_n} \sum_{j \in \mathcal{A}_i} \sum_{i' \in \mathcal{I}_n} \sum_{j' \in \mathcal{A}_{i'}} t^2\,\mathbb{E}\{W^2_{iji'j'} |\nu_{ij} - \bar{\nu}_{ij}| |\nu_{i'j'} - \bar{\nu}_{i'j'}|\,\mathrm{e}^{tW^c_{iji'j'}}\},\\
Q_4 &= \frac{1}{2} \sum_{i \in \mathcal{I}_n} \sum_{j \in \mathcal{A}_i} \sum_{i' \in \mathcal{I}_n} \sum_{j' \in \mathcal{A}_{i'}} t^2\,\mathbb{E}\{W^2_{iji'j'} |\nu_{ij} - \bar{\nu}_{ij}| |\nu_{i'j'} - \bar{\nu}_{i'j'}|\,\mathrm{e}^{tW}\}.
\end{aligned}$$

For $Q_1$, observe that $(\nu_{ij} - \bar{\nu}_{ij})(\nu_{i'j'} - \bar{\nu}_{i'j'})$ and $W^c_{iji'j'}$ are independent, then

$$\mathbb{E}\{(\nu_{ij} - \bar{\nu}_{ij})(\nu_{i'j'} - \bar{\nu}_{i'j'})\,\mathrm{e}^{tW^c_{iji'j'}}\} = \mathbb{E}\{(\nu_{ij} - \bar{\nu}_{ij})(\nu_{i'j'} - \bar{\nu}_{i'j'})\}\,\mathbb{E}\,\mathrm{e}^{tW^c_{iji'j'}}.$$

If $i', j' \in \mathcal{A}^c_{ij}$, then $\nu_{ij}$ and $\nu_{i'j'}$ are independent, and thus $\mathbb{E}\{(\nu_{ij} - \bar{\nu}_{ij})(\nu_{i'j'} - \bar{\nu}_{i'j'})\} = 0$. If $|i \cap j| = m_1$, $|i' \cap j'| = m_2$ and $|(i \cup j) \cap (i' \cup j')| = m_3$, where $1 \leqslant m_1, m_2 \leqslant e(G)$, and $1 \leqslant m_3 \leqslant 2e - 1$, then, by Lemma 5.6, it follows that

$$\big|\mathbb{E}\{(\nu_{ij} - \bar{\nu}_{ij})(\nu_{i'j'} - \bar{\nu}_{i'j'})\}\big| \leqslant \begin{cases} C\sigma_n^{-4} p^{4e(G) - m_1 - m_2 - m_3}, & 0 < p \leqslant 1/2,\\ C\sigma_n^{-4}(1-p), & 1/2 < p < 1.\end{cases}$$

For any $i,j\in\mathcal{I}_n$, denote by $G(i)$ the graph generated by the edge set $i$ and $G(i)\cap G(j)$ the common part of $G(i)$ and $G(j)$. Therefore, for $0<p<1/2$, we have

$$\begin{aligned}
&\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{I}_n}\sum_{j'\in\mathcal{A}_{i'}}\big|\mathbb{E}\big\{\big(\nu_{ij}-\bar\nu_{ij}\big)\big(\nu_{i'j'}-\bar\nu_{i'j'}\big)\big\}\big|\\
&=\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{I}_n}\sum_{j'\in\mathcal{A}_{i'}}\big|\mathbb{E}\big\{\big(\nu_{ij}-\bar\nu_{ij}\big)\big(\nu_{i'j'}-\bar\nu_{i'j'}\big)\big\}\big|\mathbf{1}_{\{|(i\cup j)\cap(i'\cup j')|\geqslant 1\}}\\
&\leqslant C\sigma_n^{-4}\sum_{i\in\mathcal{I}_n}\sum_{j:|i\cap j|\geqslant 1}\sum_{i':|i'\cap(i\cup j)|\geqslant 1}\sum_{j':|j'\cap i'|\geqslant 1}p^{4e(G)-|i\cap j|-|i'\cap j'|-|i'\cap(i\cap j)|}\\
&\quad+C\sigma_n^{-4}\sum_{i\in\mathcal{I}_n}\sum_{j:|i\cap j|\geqslant 1}\sum_{j':|j'\cap(i\cup j)|\geqslant 1}\sum_{i':|i'\cap j'|\geqslant 1}p^{4e(G)-|i\cap j|-|i'\cap j'|-|j'\cap(i\cap j)|}\\
&\leqslant C\sigma_n^{-2}(\psi_n^{-1}n^{v(G)}p^{e(G)})^2\leqslant C\psi_n^{-1},
\end{aligned}\tag{5.27}$$

where we used (5.16) and (5.21) and in the last line. For $1/2<p<1$, by (5.16) again, we have

$$\begin{aligned}
\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{I}_n}\sum_{j'\in\mathcal{A}_{i'}}&\big|\mathbb{E}\big\{\big(\nu_{ij}-\bar\nu_{ij}\big)\big(\nu_{i'j'}-\bar\nu_{i'j'}\big)\big\}\big|\\
&\leqslant C\sigma_n^{-4}n^{4v(G)-6}(1-p)\leqslant Cn^{-2}(1-p)^{-1}.
\end{aligned}\tag{5.28}$$

Then, for $0\leqslant t\leqslant(1-p)^{1/2}n^2p^{e(G)}\psi_n^{-1/2}$, by (5.16), (5.25), (5.27) and (5.28), we have

$$|Q_1|\leqslant\begin{cases}C\psi_n^{-1}\,\mathbb{E}\,\mathrm{e}^{tW}, & 0<p\leqslant 1/2,\\ Cn^{-2}(1-p)^{-1}\,\mathbb{E}\,\mathrm{e}^{tW}, & 1/2<p<1.\end{cases}\tag{5.29}$$

For $Q_2$, we have

$$Q_2\leqslant Q_{21}+Q_{22},\tag{5.30}$$

where

$$\begin{aligned}
Q_{21}&=\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{I}_n}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_i\cup\mathcal{A}_j}t\big|\mathbb{E}\big\{X_k\big(\nu_{ij}-\bar\nu_{ij}\big)\big(\nu_{i'j'}-\bar\nu_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'}}\big\}\big|,\\
Q_{22}&=\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{I}_n}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_{i'}\cup\mathcal{A}_{j'}}t\big|\mathbb{E}\big\{X_k\big(\nu_{ij}-\bar\nu_{ij}\big)\big(\nu_{i'j'}-\bar\nu_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'}}\big\}\big|.
\end{aligned}$$

By symmetry, we only consider $Q_{21}$. By the inequality $|\mathrm{e}^{x}-1| \leqslant |x|(1+\mathrm{e}^{x})$, it follows that

$$
\begin{aligned}
&\big|\mathbb{E}\big\{X_k\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'}}\big\}\big| \\
&\leqslant \big|\mathbb{E}\big\{X_k\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'k}}\big\}\big| \\
&\quad + t\big|\mathbb{E}\big\{X_k(W_{iji'j'k}-W_{iji'j'})\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'k}}\big\}\big| \\
&\quad + t\big|\mathbb{E}\big\{X_k(W_{iji'j'k}-W_{iji'j'})\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'}}\big\}\big|.
\end{aligned} \tag{5.31}
$$

Note that $X_k\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)$ is independent of $W^c_{iji'j'k}$. Then,

$$
\mathbb{E}\big\{X_k\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'k}}\big\} = \mathbb{E}\big\{X_k\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\big\}\,\mathbb{E}\,\mathrm{e}^{tW^c_{iji'j'k}}.
$$

If $\big|\big(i'\cup j'\big)\cap\big(\mathcal{A}_i\cup\mathcal{A}_j\cup\mathcal{A}_k\big)\big| = 0$, then $X_k\big(\nu_{ij}-\bar{\nu}_{ij}\big)$ is independent of $\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)$, and thus $\mathbb{E}\{X_k\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\} = 0$. Denote $|i\cap j| = m_1$, $|i'\cap j'| = m_2$, $k\cap\big(i\cup j\big) = m_4$, and $\big|\big(i\cup j\big)\cap\big(i'\cup j'\big)\big| = m_3$, where $1\leqslant m_1, m_2, m_4 \leqslant e(G)$, and $1\leqslant m_3\leqslant 2e-1$, then, by Lemma 5.6,

$$
\big|\mathbb{E}\big\{X_k\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\big\}\big| \leqslant \begin{cases} C\sigma_n^{-5}p^{5e(G)-m_1-m_2-m_3-m_4}, & 0<p\leqslant 1/2,\\ C\sigma_n^{-5}(1-p), & 1/2<p<1. \end{cases}
$$

Then, for $0<p\leqslant 1/2$, by (5.23) in Lemma 5.13, we have

$$
\begin{aligned}
&\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{I}_n}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_i\cup\mathcal{A}_j}\big|\mathbb{E}\big\{X_k\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\big\}\big| \\
&\leqslant C\sigma_n^{-5}\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{\substack{i'\in\mathcal{I}_n, j'\in\mathcal{A}_{i'}\\ |(i\cup j)\cap|i'\cup j')|\geqslant 1}}\sum_{k\in\mathcal{A}_i\cup\mathcal{A}_j} p^{5e-|i\cap j|-|i'\cap j'|-|(i\cup j)\cap(i'\cup j')|-|k\cap(i\cup j)|} \\
&\leqslant C\sigma_n^{-3}\big(\psi_n^{-1}n^{v(G)}p^{e(G)}\big)^3 \leqslant C\psi_n^{-3/2}.
\end{aligned} \tag{5.32}
$$

For $1/2<p<1$,

$$
\begin{aligned}
&\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{I}_n}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_i\cup\mathcal{A}_j}\big|\mathbb{E}\big\{X_k\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\big\}\big| \\
&\qquad\leqslant C\sigma_n^{-5}n^{5v(G)-8}(1-p) \leqslant Cn^{-3}(1-p)^{-3/2}.
\end{aligned} \tag{5.33}
$$

Note that $|X_k|\leqslant\sigma_n^{-1}$ and $|\mathcal{A}_k|\leqslant Cn^{v-2}$, and we have $\mathrm{e}^{tW_{iji'j'k}}\leqslant C\,\mathrm{e}^{tW_{iji'j'kq}}$ for $0\leqslant t\leqslant (1-p)^{1/2}n^2p^{e(G)}\psi_n^{-1/2}$. For $|i\cap j| = m_1$, $|i'\cap j'| = m_2$, $\big|\big(i\cup j\big)\cap\big(i'\cup j'\big)\big| = m_3$, $\big|k\cap\big(i\cup j\cup i'\cup j'\big)\big| = m_4$ and $\big|q\cap\big(i\cup j\cup i'\cup j'\cup k\big)\big| = m_5$, where $1\leqslant m_1,m_2\leqslant e(G)$, $1\leqslant m_4,m_5\leqslant e(G)$, and $0\leqslant m_3\leqslant 2e-1$, then for $0\leqslant t\leqslant (1-p)^{1/2}n^2p^{e(G)}\psi_n^{-1/2}$, by Lemma 5.6, we have

$$
\big|\mathbb{E}\big\{X_k(W_{iji'j'k}-W_{iji'j'})\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'k}}\big\}\big|
$$

$$\begin{aligned}
&\leqslant \sum_{q\in\mathcal{A}_k} \mathbb{E}\big|X_k X_q \big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'k}}\big| \\
&\leqslant C \sum_{q\in\mathcal{A}_k} \mathbb{E}\big|X_k X_q \big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\big|\,\mathbb{E}\,\mathrm{e}^{tW^c_{iji'j'kq}} \\
&\leqslant \begin{cases} C\sum_{q\in\mathcal{A}_k} \sigma_n^{-6} p^{6e(G)-m_1-m_2-m_3-m_4-m_5}\,\mathbb{E}\,\mathrm{e}^{tW}, & 0<p\leqslant 1/2, \\ Cn^{v(G)-2}\sigma_n^{-6}(1-p)\,\mathbb{E}\,\mathrm{e}^{tW}, & 1/2<p<1, \end{cases}
\end{aligned}$$

where we used (5.25) in the last line.

For $0<p\leqslant 1/2$ and for $0\leqslant t\leqslant (1-p)^{1/2}n^2p^{e(G)}\psi_n^{-1/2}$, it follows from (5.24) that

$$\begin{aligned}
&\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{I}_n}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_i\cup\mathcal{A}_j} \big|\mathbb{E}\big\{X_k(W_{iji'j'k}-W_{iji'j'})\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'k}}\big\}\big| \\
&\leqslant C\sigma_n^6\,\mathbb{E}\,\mathrm{e}^{tW} \sum_{\substack{i\in\mathcal{I}_n\\ j:|i\cap j|\geqslant 1}} \sum_{\substack{i'\in\mathcal{I}_n\\ j':|j'\cap i'|\geqslant 1}} p^{6e(G)-|i\cap j|-|i'\cap j'|}\Bigg(\sum_{k:|k\cap(i\cup j\cup i'\cup j')|\geqslant 1} p^{-|k\cap(i\cup j\cup i'\cup j')|}\Bigg)^2 \\
&\leqslant C\,\mathbb{E}\,\mathrm{e}^{tW}\,\sigma_n^{-2}(\psi_n^{-1}n^{v(G)}p^{e(G)})^2 \leqslant C\psi_n^{-1}\,\mathbb{E}\,\mathrm{e}^{tW}\,. \qquad (5.34)
\end{aligned}$$

For $1/2<p<1$ and for $0\leqslant t\leqslant (1-p)^{1/2}n^2p^{e(G)}\psi_n^{-1/2}$, we have

$$\begin{aligned}
&\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{I}_n}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_i\cup\mathcal{A}_j} \big|\mathbb{E}\big\{X_k(W_{iji'j'k}-W_{iji'j'})\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'k}}\big\}\big| \\
&\leqslant Cn^{-2}(1-p)^{-2}\,\mathbb{E}\,\mathrm{e}^{tW}\,. \qquad (5.35)
\end{aligned}$$

Similar to (5.34) and (5.35), for $0\leqslant t\leqslant (1-p)^{1/2}n^2p^{e(G)}\psi_n^{-1/2}$, it follows that

$$\begin{aligned}
&\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i}\sum_{i'\in\mathcal{I}_n}\sum_{j'\in\mathcal{A}_{i'}}\sum_{k\in\mathcal{A}_i\cup\mathcal{A}_j} \big|\mathbb{E}\big\{X_k(W_{iji'j'k}-W_{iji'j'})\big(\nu_{ij}-\bar{\nu}_{ij}\big)\big(\nu_{i'j'}-\bar{\nu}_{i'j'}\big)\,\mathrm{e}^{tW^c_{iji'j'}}\big\}\big| \\
&\leqslant \begin{cases} C\psi_n^{-1}\,\mathbb{E}\,\mathrm{e}^{tW}, & 0<p\leqslant 1/2, \\ Cn^{-2}(1-p)^{-2}\,\mathbb{E}\,\mathrm{e}^{tW}, & 1/2<p<1. \end{cases} \qquad (5.36)
\end{aligned}$$

Substituting (5.25) and (5.32)–(5.36) to (5.30) and (5.31), for $0\leqslant t\leqslant (1-p)^{1/2}n^2p^{e(G)}\psi_n^{-1/2}$, we have

$$Q_2 \leqslant \begin{cases} C(t\psi_n^{-3/2}+t^2\psi_n^{-1})\,\mathbb{E}\,\mathrm{e}^{tW}, & 0<p\leqslant 1/2, \\ C(tn^{-3}(1-p)^{-3/2}+t^2n^{-2}(1-p)^{-2})\,\mathbb{E}\,\mathrm{e}^{tW}, & 1/2<p<1. \end{cases} \qquad (5.37)$$

For any $H\subset G$ such that $e(H)>0$, we have $v(H)\geqslant 2$ and $e(H)\leqslant e(G)$, and it follows that

$$n^{v(H)}p^{e(H)} \geqslant n^2p^{e(G)}.$$

Thus, for $0 < p \leqslant 1/2$ and $0 \leqslant t \leqslant (1-p)^{1/2} n^2 p^{e(G)} \psi_n^{-1/2}$, it follows that $0 \leqslant t\psi_n^{-1/2} \leqslant n^2 p^{e(G)} \psi_n^{-1} \leqslant 1$. Hence, (5.37) becomes

$$Q_2 \leqslant \begin{cases} C\psi_n^{-1}(1+t^2)\,\mathbb{E}\,\mathrm{e}^{tW}, & 0<p\leqslant 1/2, \\ C(tn^{-3}(1-p)^{-3/2} + t^2 n^{-2}(1-p)^{-2})\,\mathbb{E}\,\mathrm{e}^{tW}, & 1/2<p<1. \end{cases} \tag{5.38}$$

Similar to (5.34)–(5.36), we have for $0 \leqslant t \leqslant (1-p)^{1/2} n^2 p^{e(G)} \psi_n^{-1/2}$,

$$Q_3 + Q_4 \leqslant \begin{cases} Ct^2\psi_n^{-1}\,\mathbb{E}\,\mathrm{e}^{tW}, & 0<p\leqslant 1/2, \\ Ct^2 n^{-2}(1-p)^{-2}\,\mathbb{E}\,\mathrm{e}^{tW}, & 1/2<p<1. \end{cases}$$

This proves (5.17) together with (5.29) and (5.38). □

## 5.3. Proof of Theorem 3.3

In this proof, we denote by $C$ a general constant that depends only on $\beta$, where $0<\beta<1$. Let $\mathcal{X}$ be the sigma field generated by $(X_1, \ldots, X_n)$. For each $1 \leqslant i \leqslant n$, given $\{X_j, j \neq i\}$, let $X_i'$ be conditionally independent of $X_i$ with the conditional distribution of $X_i$. Let $I$ be a random index that is uniformly distributed over $\{1, \ldots, n\}$ and independent of any other random variable. Define $S_n' = S_n - X_I + X_I'$; then $(S_n, S_n')$ is an exchangeable pair. Recall that $S_n = X_1 + \cdots + X_n$ and $W := W_n = n^{-1/2}(1-\beta)^{1/2} S_n$. Let $\xi, \xi_1, \ldots, \xi_n$ be i.i.d. random variables with the probability measure $\rho$. Let $\bar{X} = S_n/n, \bar{X}_i = \frac{1}{n}(S_n - X_i)$ and $\bar{X}_{ij} = \frac{1}{n}(S_n - X_i - X_j)$ for $1 \leqslant i \neq j \leqslant n$.

For $n \leqslant 40\beta/(1-\beta)$, and for $0 \leqslant z \leqslant \sqrt{n}$, we have $z \leqslant z_\beta := \sqrt{40\beta/(1-\beta)}$. By Shao and Zhang [22, Theorem 3.2], for $0 \leqslant z \leqslant z_\beta$,

$$\big|\mathbb{P}(W>z) - (1-\Phi(z))\big| \leqslant Cn^{-1/2} \leqslant Cn^{-1/2}(1-\Phi(z_\beta)) \leqslant Cn^{-1/2}(1-\Phi(z)).$$

Hence (3.7) holds. For $n > 40\beta(1-\beta)$, we apply Theorem 2.1 to prove the moderate deviation result. To this end, we give the following propositions.

**Proposition 5.8.** *We have*

$$\mathbb{E}\left\{S_n - S_n' \,\middle|\, \mathcal{X}\right\} = (1-\beta)\bar{X} + R_1, \tag{5.39}$$

*where $R_1$ is a random variable such that for $n > 40\beta/(1-\beta)$ and $0 \leqslant t \leqslant \sqrt{n}$,*

$$\mathbb{E}\{|R_1|\,\mathrm{e}^{tW}\} \leqslant Cn^{-1}\,\mathrm{e}^{t^2/2}. \tag{5.40}$$

**Proposition 5.9.** *We have for $n > 40\beta/(1-\beta)$ and $0 \leqslant t \leqslant \sqrt{n}$,*

$$\mathbb{E}\big\{\big|\mathbb{E}\{(S_n - S_n')^2 \mid \mathcal{X}\} - 2\big|\,\mathrm{e}^{tW}\big\} \leqslant Cn^{-1/2}(1+t)\,\mathrm{e}^{t^2/2}, \tag{5.41}$$

*and*

$$\mathbb{E}\{\big|\mathbb{E}\{(S_n - S_n')|S_n - S_n' \mid \mathcal{X}\}\big| \,\mathrm{e}^{tW}\} \leqslant Cn^{-1/2}(1+t)\,\mathrm{e}^{t^2/2}. \tag{5.42}$$

With the foregoing propositions, we can check the conditions (A1)–(A3) immediately. Let $W' = n^{-1/2}(1-\beta)S_n'$. Observe that

$$\mathbb{E}\left\{W - W' \,|\, \mathcal{X}\right\} = n^{-1/2}(1-\beta)^{1/2}\,\mathbb{E}\left\{S_n - S_n' \,|\, \mathcal{X}\right\} = \lambda(W+R),$$

where $\lambda = (1-\beta)/n$ and $R = n^{1/2}(1-\beta)^{1/2}R_1$. Moreover,

$$\frac{1}{2\lambda}\,\mathbb{E}\{(W-W')^2 \mid \mathcal{X}\} - 1 = \frac{1}{2}\Big(\mathbb{E}\{(S_n - S_n')^2 \mid \mathcal{X}\} - 2\Big),$$

and

$$\frac{1}{\lambda}\,\mathbb{E}\{(W-W')|W-W'| \mid \mathcal{X}\} = \mathbb{E}\{(S_n - S_n')|S_n - S_n' \mid \mathcal{X}\}.$$

Hence, by Propositions 5.8 and 5.9, conditions (A1)–(A3) are satisfied with $\tau_0 = \sqrt{n}, \delta_1(t) = \delta_2(t) = Cn^{-1/2}(1+t)$ and $\delta_3(t) = Cn^{-1/2}$. This completes the proof by Theorem 2.1.

It suffices to give the proofs of Propositions 5.8 and 5.9, to this end, we need to prove some preliminary lemmas.

**Lemma 5.10.** *For $0 \leqslant \theta < 1$ and $z > 0$,*

$$\mathbb{E}\,\mathrm{e}^{\theta\xi^2/2} \leqslant C_\theta, \tag{5.43}$$

*and for $r > 1$,*

$$\mathbb{E}\{|\xi|^r\,\mathrm{e}^{\theta\xi^2/2}\} \leqslant C_{\theta,r}, \tag{5.44}$$

*where $C_\theta > 0$ is a constant depending on $\theta$ and $C_{\theta,r} > 0$ is a constant depending on $\theta$ and $r$. Also,*

$$\mathbb{P}\big(\big|\xi_1 + \cdots + \xi_n\big| > z\big) \leqslant 2\,\mathrm{e}^{z^2/(2n)}. \tag{5.45}$$

*Moreover, for any $s \in \mathbb{R}$ and $0 < \beta < 1$,*

$$\mathbb{E}\,\mathrm{e}^{\frac{\beta\xi^2}{2n} + \beta s\xi} \geqslant \mathrm{e}^{-\beta(1+s^2)/2}. \tag{5.46}$$

Let $T_n = \xi_1 + \cdots + \xi_n$.

**Lemma 5.11.** *Let $\alpha_n = n^{-1/2}(1-\beta)^{1/2}t$. We have for $\theta_0 > 0, 0 < \beta < 1, 0 \leqslant \theta \leqslant \min\{n(1-\beta)/4, \theta_0\}$ and $0 \leqslant t \leqslant \sqrt{n}$,*

$$\mathbb{E}\exp\bigg(\bigg(\frac{\beta}{2n} + \frac{\theta}{n^2}\bigg)T_n^2 + \alpha_n T_n\bigg) \leqslant C_0\,\mathrm{e}^{t^2/2}, \tag{5.47}$$

*where $C_0 > 0$ is a constant depending only on $\beta$ and $\theta_0$. For $r > 1$,*

$$\mathbb{E}\left\{|T_n|^r \exp\left(\left(\frac{\beta}{2n} + \frac{\theta}{n^2}\right)T_n^2 + \alpha_n T_n\right)\right\} \leqslant C_{0,r} n^{r/2} \mathrm{e}^{t^2/2}, \tag{5.48}$$

*where $C_{0,r} > 0$ is a constant depending only on $\beta$, $\theta_0$ and $r$.*

Recall that for each $1 \leqslant i \leqslant n$, given $\{X_j, j \neq i\}$, $X_i'$ is conditionally independent of $X_i$ with the conditional distribution of $X_i$. Also, recall that the normalizing constant $Z_n = \mathbb{E}\exp\{\beta(\xi_1 + \cdots + \xi_n)^2/(2n)\}$.

**Lemma 5.12.** *For $0 < \beta < 1$, we have*

$$1 \leqslant Z_n \leqslant C, \quad \mathbb{E}|S_n|^2 \leqslant Cn, \tag{5.49}$$

*and for $n > 4\beta/(1-\beta)$ and $0 \leqslant t \leqslant \sqrt{n}$,*

$$\mathbb{E}\,\mathrm{e}^{tW} \leqslant C\,\mathrm{e}^{t^2/2}, \tag{5.50}$$

$$\mathbb{E}\{|X_i|^6\,\mathrm{e}^{tW}\} \leqslant C\,\mathrm{e}^{t^2/2}, \tag{5.51}$$

$$\mathbb{E}\{|X_i'|^6\,\mathrm{e}^{tW}\} \leqslant C\,\mathrm{e}^{t^2/2}. \tag{5.52}$$

The proofs of Lemmas 5.10–5.12 are put in Appendix A.3.

**Lemma 5.13.** *For $i \neq j$, we have for $n > 8\beta/(1-\beta)$ and $0 \leqslant t \leqslant \sqrt{n}$,*

$$\left|\mathbb{E}\{(X_i^2 - 1)(X_j^2 - 1)\,\mathrm{e}^{tW}\}\right| \leqslant Cn^{-1}\,\mathrm{e}^{t^2/2}.$$

***Proof of Lemma 5.13.*** The proof is similar to Lemma 5.7 of Shao and Zhang [22]. We only consider $i = 1$ and $j = 2$. Let $M_{12} = \xi_3 + \cdots + \xi_n$, and $\alpha_n = n^{-1/2}(1-\beta)^{1/2}t$,

$$\begin{aligned}
&\mathbb{E}\{(X_i^2 - 1)(X_j^2 - 1)\,\mathrm{e}^{tW}\} \\
&= \frac{1}{Z_n}\mathbb{E}\left\{(\xi_1^2 - 1)(\xi_2^2 - 1)\exp\left(\frac{\beta}{2n}(\xi_1 + \cdots + \xi_n)^2 + n^{-1/2}(1-\beta)^{1/2}t(\xi_1 + \cdots + \xi_n)\right)\right\} \\
&= \frac{1}{Z_n}\mathbb{E}\left\{(\xi_1^2 - 1)(\xi_2^2 - 1)\exp\left(\frac{\beta}{2n}(\xi_1 + \xi_2)^2 + \left(\frac{\beta}{n}M_{12} + \alpha_n\right)(\xi_1 + \xi_2) + \frac{\beta}{2n}M_{12}^2 + \alpha_n M_{12}\right)\right\}.
\end{aligned}$$

By Shao and Zhang [22, Eq. (5.28)],

$$\left|\mathbb{E}\left\{(\xi_1^2 - 1)(\xi_2^2 - 1)\exp\left(\frac{\beta}{2n}(\xi_1 + \xi_2)^2 + s(\xi_1 + \xi_2)\right)\right\}\right| \leqslant C\left(\frac{1}{n} + s^2\right)\mathrm{e}^{\beta s^2},$$

and then for $0 \leqslant t \leqslant \sqrt{n}$, by (5.49) and Lemma 5.11 with $\theta_0 = 2\beta$, we have for $n > 8\beta/(1-\beta)$,

$$\begin{aligned}
&\bigl|\mathbb{E}\bigl\{(X_i^2-1)(X_j^2-1)\,e^{tW}\bigr\}\bigr| \\
&\quad\leqslant Cn^{-1}\,\mathbb{E}\biggl\{\Bigl(1+t^2+\frac{M_{12}^2}{n}\Bigr)\exp\Bigl(\Bigl(\frac{\beta}{2n}+\frac{\beta}{n^2}\Bigr)M_{12}^2+\alpha_n M_{12}\Bigr)\biggr\} \\
&\quad\leqslant Cn^{-1}(1+t^2)\,e^{t^2/2}.
\end{aligned}$$

This completes the proof. □

Recall that $\mathcal{F} = \sigma(X_1, \ldots, X_n)$ and let

$$Q_i = \mathbb{E}\bigl\{(X_i - X_i')|X_i - X_i'| \bigm| \mathcal{F}\bigr\}.$$

**Lemma 5.14.** *We have for $n > 16\beta/(1-\beta)$ and $0 \leqslant t \leqslant \sqrt{n}$,*

$$\mathbb{E}\bigl\{Q_i^2\,e^{tW}\bigr\} \leqslant C\,e^{t^2/2}, \tag{5.53}$$

$$\bigl|\mathbb{E}\bigl\{Q_iQ_j\,e^{tW}\bigr\}\bigr| \leqslant Cn^{-1}(1+t^2)\,e^{t^2/2}. \tag{5.54}$$

***Proof of Lemma 5.14.*** Note that

$$\mathbb{E}\bigl\{Q_i^2\,e^{tW}\bigr\} \leqslant \mathbb{E}\{(X_i - X_i')^2\,e^{tW}\} \leqslant 2\,\mathbb{E}\{X_i^2\,e^{tW}\} + 2\,\mathbb{E}\{(X_i')^2\,e^{tW}\}.$$

Then (5.53) follows from (5.51) and (5.52).

Now, we prove (5.54). Recall that $\bar{X}_i = \frac{1}{n}(S_n - X_i)$. Let $g(s,t) = (s-t)|s-t|$. Let $\xi, \xi_1, \ldots, \xi_n$ be i.i.d. random variables with the probability measure $\rho$, which are independent of $(X_1, \ldots, X_n)$. Let $\mathbb{E}_\xi$ denote the expectation with respect to $\xi$ conditional on all other random variables, then we can rewrite $Q_i$ as

$$Q_i = \frac{\mathbb{E}_\xi\bigl\{g(X_i,\xi)\exp(\beta\xi^2/(2n) + \beta\bar{X}_i\xi)\bigr\}}{\mathbb{E}_\xi\bigl\{\exp(\beta\xi^2/(2n) + \beta\bar{X}_i\xi)\bigr\}}.$$

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

$$Q_j' = \frac{\mathbb{E}_\xi\bigl\{g(X_j,\xi)\,e^{\beta\bar{X}_{12}\xi}\bigr\}}{\mathbb{E}_\xi\bigl\{e^{\beta\bar{X}_{12}\xi}\bigr\}}, \quad j = 1, 2.$$

Let $M_{12} = (\xi_3 + \cdots + \xi_n)$ and recall that $\alpha_n = n^{-1/2}t(1-\beta)^{1/2}$, then

$$\begin{aligned}
&\mathbb{E}\bigl\{Q_1'Q_2'\,e^{tW}\bigr\} \\
&= \frac{1}{Z_n}\,\mathbb{E}\biggl\{h(\xi_1)h(\xi_2)\exp\Bigl(\frac{\beta}{2n}(\xi_1+\xi_2+M_{12})^2 + \alpha_n(\xi_1+\xi_2+M_{12})\Bigr)\biggr\}
\end{aligned}$$

$$= \frac{1}{Z_n}\mathbb{E}\bigg\{h(\xi_1)h(\xi_2)\exp\bigg(\frac{\beta}{2n}(\xi_1+\xi_2)^2+\bigg(\frac{\beta}{n}M_{12}+\alpha_n\bigg)(\xi_1+\xi_2)+\frac{\beta}{2n}M_{12}^2+\alpha_n M_{12}\bigg)\bigg\}, \tag{5.55}$$

where

$$h(x) = \frac{\mathbb{E}_\xi\{g(x,\xi)\exp(\beta M_{12}\xi/n)\}}{\mathbb{E}_\xi \exp(\beta M_{12}\xi/n)}.$$

By (5.23) of Shao and Zhang [22], we have $\mathbb{E}\,\mathrm{e}^{\beta s\xi/n} \geqslant C\,\mathrm{e}^{-\beta s^2/(2n^2)}$. Also, since $|g(s,t)| \leqslant (s-t)^2$, it follows that

$$|h(x)| \leqslant C\,\mathrm{e}^{\beta M_{12}^2/n^2}(1+x^2+M_{12}^2/n^2). \tag{5.56}$$

Noting that

$$\mathrm{e}^{\beta x^2/2} = \sqrt{\frac{\beta}{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\beta t x-\beta t^2/2}\,\mathrm{d}t,$$

we have with $m = \beta M_{12}/n + \alpha_n + \beta u/\sqrt{n}$,

$$\begin{aligned}
&\mathbb{E}\left\{h(\xi_1)h(\xi_2)\exp\bigg(\frac{\beta}{2n}(\xi_1+\xi_2)^2+\Big(\frac{\beta M_{12}}{n}+\alpha_n\Big)(\xi_1+\xi_2)\bigg)\,\bigg|\,M_{12}\right\}\\
&= \sqrt{\frac{\beta}{2\pi}}\int_{-\infty}^{\infty}\mathbb{E}\big\{h(\xi_1)h(\xi_2)\exp\big(m(\xi_1+\xi_2)-\beta u^2/2\big)\,\big|\,M_{12}\big\}\,\mathrm{d}u.
\end{aligned} \tag{5.57}$$

Since $(\xi_1,\xi_2)$ is independent of $M_{12}$, and $\xi_1$ is independent of and identically distributed as $\xi_2$,

$$\mathbb{E}\left\{h(\xi_1)h(\xi_2)\exp\big(m(\xi_1+\xi_2)\big)\,\big|\,M_{12}\right\} = (\mathbb{E}\left\{h(\xi_1)\exp\big(m\xi_1\big)\,\big|\,M_{12}\right\})^2. \tag{5.58}$$

Since $g(u,t)$ is antisymmetric,

$$\mathbb{E}\left\{h(\xi_1)\,|\,M_{12}\right\} = 0.$$

Also, by Lemma 5.10,

$$\mathbb{E}\{(1+|\xi_1|^3)\exp(\beta\xi_1^2/2)\} \leqslant C.$$

By the Taylor expansion and (5.56),

$$\begin{aligned}
|\mathbb{E}\{h(\xi_1)e^{m\xi_1}|M_{12}\}| &= |\mathbb{E}\{h(\xi_1)\}+\mathbb{E}\{h(\xi_1)(e^{m\xi_1}-1)|M_{12}\}|\\
&\leqslant m\,\mathbb{E}\{|\xi_1 h(\xi_1)|e^{|m\xi_1|}|M_{12}\}\\
&\leqslant Cm(1+M_{12}^2/n^2)e^{m^2/(2\beta)}\,\mathbb{E}\{(1+|\xi_1|^3)e^{\beta|\xi_1|/2}\}\\
&\leqslant Cm(1+M_{12}^2/n^2)\exp\big(m^2/(2\beta)\big).
\end{aligned} \tag{5.59}$$

Substituting (5.58) and (5.59) into (5.57), for $0 \leqslant t \leqslant \sqrt{n}$,

$$\biggl|\mathbb{E}\biggl\{h(\xi_1)h(\xi_2)\exp\biggl\{\frac{\beta}{2n}(\xi_1+\xi_2)^2+\Bigl(\frac{\beta M_{12}}{n}+\alpha_n\Bigr)(\xi_1+\xi_2)\biggr\}\biggm| M_{12}\biggr\}\biggr| \leqslant Cn^{-1}\Bigl(1+t^2+\frac{M_{12}^2}{n}\Bigr)\Bigl(1+\frac{M_{12}^4}{n^4}\Bigr)\,\mathrm{e}^{\frac{4\beta M_{12}^2}{n^2}}. \quad (5.60)$$

Combining (5.55) and (5.60), and by (5.47) and (5.48), we have for $n \geqslant 16\beta/(1-\beta)$,

$$\bigl|\mathbb{E}\{Q_1'Q_2'\,\mathrm{e}^{tW}\}\bigr| \leqslant Cn^{-1}(1+t^2)\,\mathrm{e}^{t^2/2}. \quad (5.61)$$

Next, we estimate $\mathbb{E}\{(Q_1-Q_1')^2\,\mathrm{e}^{tW}\}$. We have

$$|Q_1-Q_1'| \leqslant \frac{|\,\mathbb{E}_\xi\{g(X_1,\xi)\,\mathrm{e}^{\beta\bar{X}_{12}\xi}(\mathrm{e}^{\frac{\beta\xi^2}{2n}+\frac{\beta X_2\xi}{n}}-1)\}|}{\mathbb{E}_\xi\exp\{\beta\xi^2/(2n)+\beta\bar{X}_1\xi\}} + \frac{\mathbb{E}_\xi\{|g(X_1,\xi)|\,\mathrm{e}^{\beta\bar{X}_{12}\xi}\}\,\mathbb{E}_\xi\{\mathrm{e}^{\beta\bar{X}_{12}\xi}|\mathrm{e}^{\frac{\beta\xi^2}{2n}+\frac{\beta X_2\xi}{n}}-1|\}}{\mathbb{E}_\xi\exp\{\beta\xi^2/(2n)+\beta\bar{X}_1\xi\}\,\mathbb{E}_\xi\{\mathrm{e}^{\beta\bar{X}_{12}\xi}\}}.$$

For the first term, as $|g(s,t)| \leqslant (s-t)^2$,

$$\begin{aligned}&\Bigl|\mathbb{E}_\xi\Bigl\{g(X_1,\xi)\,\mathrm{e}^{\beta\bar{X}_{12}\xi}\bigl(\mathrm{e}^{\frac{\beta\bar{X}_{12}\xi}{2n}+\frac{\beta X_2\xi}{n}}-1\bigr)\Bigr\}\Bigr|\\ &\leqslant Cn^{-1}(1+|X_1|^3+|X_2|^3)\exp\biggl\{\frac{\beta(|X_2|+S_{12})^2}{2n^2}\biggr\}\,\mathbb{E}\bigl\{(1+\xi^4)\,\mathrm{e}^{\frac{\beta\xi^2}{2n}+\frac{\beta\xi^2}{2}}\bigr\}\\ &\leqslant Cn^{-1}(1+|X_1|^3+|X_2|^3)\exp\biggl\{\frac{\beta(|X_2|+S_{12})^2}{2n^2}\biggr\},\end{aligned}$$

where $S_{12} = X_3+\cdots+X_n$ and we used Lemma 5.10 in the last line. Applying Lemma 5.10 again,

$$\mathbb{E}_\xi\exp\{\beta\xi^2/(2n)+\beta\bar{X}_1\xi\} \geqslant C\exp\biggl\{-\frac{\beta(|X_2|+S_{12})^2}{2n^2}\biggr\}.$$

Thus,

$$\frac{|\,\mathbb{E}_\xi\{g(X_1,\xi)\,\mathrm{e}^{\beta\bar{X}_{12}\xi}(\mathrm{e}^{\frac{\beta\xi^2}{2n}+\frac{\beta X_2\xi}{n}}-1)\}|}{\mathbb{E}_\xi\exp\{\beta\xi^2/(2n)+\beta\bar{X}_1\xi\}} \leqslant Cn^{-1}(1+|X_1|^3+|X_2|^3)\exp\biggl\{\frac{\beta(|X_2|+S_{12})^2}{n^2}\biggr\}.$$

Similarly,

$$\frac{\mathbb{E}_\xi\{|g(X_1,\xi)|\,\mathrm{e}^{\beta\bar{X}_{12}\xi}\}\,\mathbb{E}_\xi\{\mathrm{e}^{\beta\bar{X}_{12}\xi}|\mathrm{e}^{\frac{\beta\xi^2}{2n}+\frac{\beta X_2\xi}{n}}-1|\}}{\mathbb{E}_\xi\exp\{\beta\xi^2/(2n)+\beta\bar{X}_1\xi\}\,\mathbb{E}_\xi\{\mathrm{e}^{\beta\bar{X}_{12}\xi}\}} \leqslant Cn^{-1}(1+|X_1|^3+|X_2|^3)\exp\biggl\{\frac{2\beta(|X_2|+S_{12})^2}{n^2}\biggr\}.$$

Thus,

$$\left|Q_1 - Q_1'\right| \leqslant Cn^{-1}(1 + |X_1|^3 + |X_2|^3)\exp\left\{\frac{2\beta(|X_2| + S_{12})^2}{n^2}\right\},$$

and then with $M_{12} = (\xi_3 + \cdots + \xi_n)$, by Lemma 5.11 with $\theta_0 = 10\beta$, we have for $n > 40\beta/(1-\beta)$,

$$\begin{aligned}
&\mathbb{E}\left\{\left|Q_1 - Q_1'\right|^2 e^{tW}\right\} \\
&\leqslant Cn^{-2}\,\mathbb{E}\bigg\{(1 + |\xi_1|^6 + |\xi_2|^6) \\
&\qquad\qquad \times \exp\left(\frac{4\beta(|\xi_2| + M_{12})^2}{n^2} + \frac{\beta}{2n}(\xi_1 + \xi_2 + M_{12})^2 + \alpha_n(\xi_1 + \xi_2 + M_{12})\right)\bigg\} \\
&\leqslant Cn^{-2}\,\mathbb{E}\exp\left\{\left(\frac{\beta}{2n} + \frac{10\beta}{n^2}\right)M_{12}^2 + \alpha_n M_{12}\right\} \\
&\leqslant Cn^{-2}\,e^{t^2/2}.
\end{aligned}$$

Similarly, we have

$$\mathbb{E}\left\{|Q_2 - Q_2'|^2 e^{tW}\right\} \leqslant Cn^{-2}(1 + t^2)\,e^{t^2/2}.$$

Now, observe that

$$\begin{aligned}
|\mathbb{E}\{Q_1 Q_2 e^{tW}\}| &\leqslant |\mathbb{E}\{Q_1' Q_2' e^{tW}\}| + |\mathbb{E}\{Q_1(Q_2 - Q_2')\,e^{tW}\}| \\
&\quad + |\mathbb{E}\{Q_2(Q_1 - Q_1')\,e^{tW}\}| + \mathbb{E}\{|(Q_1 - Q_1')(Q_2 - Q_2')|\,e^{tW}\}.
\end{aligned}$$

By (5.60) and (5.61) and the Cauchy inequality, we have for $n \geqslant 40\beta/(1-\beta)$ and $0 \leqslant t \leqslant \sqrt{n}$,

$$\left|\mathbb{E}\left\{Q_1 Q_2 e^{tW}\right\}\right| \leqslant Cn^{-1}(1 + t^2)\,e^{t^2/2}.$$

This completes the proof. □

Now we are ready to prove Propositions 5.8 and 5.9. In what follows, we fix $n > 40\beta(1-\beta)$ and $0 \leqslant t \leqslant \sqrt{n}$. Again, let $\alpha_n = n^{-1/2}(1-\beta)^{1/2}t$.

***Proof of Proposition 5.8.*** Let $\xi, \xi_1, \ldots, \xi_n$ be i.i.d. random variables with probability measure $\rho$. Let $\mathbb{E}_\xi$ denote the expectation with respect to $\xi$ conditional on other random variables. Let $\bar{X} = (X_1 + \cdots + X_n)/n$ and $\bar{X}_i = \bar{X} - X_i/n$. By the definition of $(S_n, S_n')$, we have

$$\mathbb{E}\left\{S_n - S_n' \,\middle|\, \mathcal{X}\right\} = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left\{X_i - X_i' \,\middle|\, \mathcal{X}\right\} = \bar{X} - \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{E}_\xi\{\xi\, e^{\frac{\beta\xi^2}{2n} + \beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{e^{\frac{\beta\xi^2}{2n} + \beta\bar{X}_i\xi}\}}. \tag{5.62}$$

Observe that

$$\frac{\mathbb{E}_\xi\{\xi\,\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}} = h(\bar{X}_i) + r_{1i}, \tag{5.63}$$

where

$$h(s) = \frac{\mathbb{E}\{\xi\,\mathrm{e}^{\beta s\xi}\}}{\mathbb{E}\,\mathrm{e}^{\beta s\xi}}, \qquad r_{1i} = \frac{\mathbb{E}_\xi\{\xi\,\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}} - \frac{\mathbb{E}_\xi\{\xi\,\mathrm{e}^{\beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{\mathrm{e}^{\beta\bar{X}_i\xi}\}}.$$

By Lemma 5.10, we have

$$|r_{1i}| \leqslant Cn^{-1}\,\mathrm{e}^{2\beta\bar{X}_i^2}.$$

By the Taylor expansion,

$$h(\bar{X}_i) = \beta\bar{X} - \frac{\beta}{n}X_i + \int_0^{\bar{X}_i} h''(t)(\bar{X}_i - t)\,\mathrm{d}t. \tag{5.64}$$

By Shao and Zhang [22, Eq. (5.41)],

$$\left|\int_0^{\bar{X}_i} h''(t)(\bar{X}_i - t)\,\mathrm{d}t\right| \leqslant C|\bar{X}_i|^2\,\mathrm{e}^{\beta\bar{X}_i^2/2}. \tag{5.65}$$

It follows from (5.62)–(5.65) that (5.39) is satisfied with

$$|R_1| \leqslant \frac{C}{n}\sum_{i=1}^n\left\{\beta n^{-1}|X_i| + |\bar{X}_i|^2\,\mathrm{e}^{\beta\bar{X}_i^2/2} + |r_{1i}|\right\}, \tag{5.66}$$

Next we prove the bound of $\mathbb{E}|R_1|\,\mathrm{e}^{tW}$. By Lemmas 5.11 and 5.12, for $n > 8\beta/(1-\beta)$ and $0 \leqslant t \leqslant \sqrt{n}$,

$$\mathbb{E}\{|r_{1i}|\,\mathrm{e}^{tW}\} \leqslant Cn^{-1}\,\mathrm{e}^{t^2/2}. \tag{5.67}$$

By Lemmas 5.11 and 5.12, with $M_1 = \xi_2 + \cdots + \xi_n$, by symmetry,

$$\begin{aligned}
&\mathbb{E}\{|\bar{X}_i|^2\,\mathrm{e}^{\beta\bar{X}_i^2/2}\,\mathrm{e}^{tW}\}\\
&\leqslant \frac{1}{n^2 Z_n}\,\mathbb{E}\left\{M_1^2\exp\left(\left(\frac{\beta}{2n} + \frac{\beta}{n^2}\right)M_1^2 + \frac{\beta}{2}\xi_1^2 + \alpha_n(\xi_1 + M_1)\right)\right\}\\
&\leqslant Cn^{-1}\,\mathrm{e}^{t^2/2}.
\end{aligned} \tag{5.68}$$

By (5.51),

$$\mathbb{E}\{|X_i|\,\mathrm{e}^{tW}\} \leqslant C\,\mathrm{e}^{t^2/2}. \tag{5.69}$$

Combining (5.67)–(5.69), we complete the proof of Proposition 5.8. □

***Proof of Proposition 5.9.*** Observe that

$$\mathbb{E}\big\{(S_n - S_n')^2 \bigm| \mathcal{X}\big\} := 2 + R_2 + R_3 + R_4, \tag{5.70}$$

where

$$R_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 1),$$

$$R_3 = -\frac{1}{n}\sum_{i=1}^{n}\frac{2X_i\,\mathbb{E}_\xi\{\xi\,\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}},$$

$$R_4 = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{E}_\xi\{\xi^2\,\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}} - 1.$$

It follows from Lemmas 5.11–5.13 and the Cauchy inequality that

$$\mathbb{E}\{|R_2|\,\mathrm{e}^{tW}\} \leqslant \big(\mathbb{E}\,\mathrm{e}^{tW}\big)^{1/2}\big(\mathbb{E}\{|R_2|^2\,\mathrm{e}^{tW}\}\big)^{1/2} \leqslant Cn^{-1/2}(1+t)\,\mathrm{e}^{t^2/2}. \tag{5.71}$$

Note that by (5.63),

$$R_3 = -\frac{2}{n}\sum_{i=1}^{n}X_i\big\{h(\bar{X}_i) + r_{1i}\big\},$$

and similar to the proof of (5.40), we have

$$\mathbb{E}\,|R_3| \leqslant Cn^{-1}\,\mathrm{e}^{t^2/2}. \tag{5.72}$$

For $R_4$, note that

$$\frac{\mathbb{E}_\xi\{\xi^2\,\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}} - 1 = \frac{\mathbb{E}_\xi\{(\xi^2-1)\,\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}} = \frac{\mathbb{E}_\xi\{(\xi^2-1)\,\mathrm{e}^{\beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{\mathrm{e}^{\beta\bar{X}_i\xi}\}} + r_{2i},$$

where

$$r_{2i} = \frac{\mathbb{E}_\xi\{(\xi^2-1)\,\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{\mathrm{e}^{\frac{\beta\xi^2}{2n}+\beta\bar{X}_i\xi}\}} - \frac{\mathbb{E}_\xi\{(\xi^2-1)\,\mathrm{e}^{\beta\bar{X}_i\xi}\}}{\mathbb{E}_\xi\{\mathrm{e}^{\beta\bar{X}_i\xi}\}}.$$

Similar to (5.67), we have

$$\mathbb{E}\{|r_{2i}|\,\mathrm{e}^{tW}\} \leqslant Cn^{-1}\,\mathrm{e}^{t^2/2}.$$

By the symmetry property of $\rho$, we know $\mathbb{E}\,\xi^3 = \mathbb{E}\,\xi = 0$, and then

$$\big|\mathbb{E}_\xi\{(\xi^2-1)\,\mathrm{e}^{\beta\bar{X}_i\xi}\}\big| \leqslant \big|\mathbb{E}\{\xi^2-1\}\big| + \big|\beta\bar{X}_i\,\mathbb{E}\{\xi(\xi^2-1)\}\big| + C\bar{X}_i^2\,\mathbb{E}_\xi\{|(\xi^2-1)\xi_i^2|\,\mathrm{e}^{\beta|\bar{X}_i\xi|}\}$$

$$\leqslant C \bar{X}_i^2 \, e^{\beta \bar{X}_i^2/2}.$$

By Lemma 5.10, $\mathbb{E}_\xi\{e^{\beta \bar{X}_i \xi}\} \geqslant C \, e^{-\beta \bar{X}_i^2/2}$. Hence, similar to (5.68),

$$\mathbb{E}\left| \frac{\mathbb{E}_\xi\{(\xi^2 - 1) \, e^{\beta \bar{X}_i \xi}\}}{\mathbb{E}_\xi\{e^{\beta \bar{X}_i \xi}\}} \right| e^{tW} \leqslant C \, \mathbb{E}\{\bar{X}_i^2 \, e^{\beta \bar{X}_i^2}\} \leqslant C n^{-1} \, e^{t^2/2}.$$

Therefore,

$$\mathbb{E}\{|R_4| \, e^{tW}\} \leqslant C n^{-1} \, e^{t^2/2}. \tag{5.73}$$

This completes the proof of (5.41) by combining (5.70)–(5.73) For (5.42), we have

$$\mathbb{E}\big\{(S_n - S_n')|S_n - S_n'| \,\big|\, \mathcal{X}\big\} = \frac{1}{n} \sum_{i=1}^n Q_i,$$

where

$$Q_i = \mathbb{E}\big\{(X_i - X_i')|X_i - X_i'| \,\big|\, \mathcal{X}\big\}.$$

By Lemmas 5.12 and 5.14 and the Cauchy inequality, we have

$$\begin{aligned} \mathbb{E}\left\{ \left| \frac{1}{n} \sum_{i=1}^n Q_i \right| e^{tW} \right\} &\leqslant \big(\mathbb{E}\, e^{tW}\big)^{1/2} \left( \mathbb{E}\left\{ \left| \frac{1}{n} \sum_{i=1}^n Q_i \right|^2 e^{tW} \right\} \right)^{1/2} \\ &\leqslant C n^{-1/2} (1 + t) \, e^{t^2/2}. \end{aligned}$$

This completes the proof of (5.42). □

## Acknowledgement

The author would like to thank Qi-Man Shao for comments that lead to an improvement of this paper. Part of the paper was completed during the period of my visit at the Chinese University of Hong Kong. It was partially supported by Hong Kong Research Grants Council GRF 14304917. This research was also partially supported by the Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers CE140100049 and by Singapore Ministry of Education Academic Research Fund MOE 2018-T2-076.

## References

[1] BARBOUR, A., CHEN, L. H. and CHOI, K. (1995). Poisson approximation for unbounded functions, I: Independent summands. *Stat. Sin.* **5** 749–766.

[2] BARBOUR, A., KAROŃSKI, M. and RUCIŃSKI, A. (1989). A central limit theorem for decomposable random variables with applications to random graphs. *J. Comb. Theory B*. **47** 125–145.

[3] BARBOUR, A. D., HOLST, L. and JANSON, S. (1992). *Poisson approximation*. **2**. The Clarendon Press, Oxford University Press, New York. Oxford Science Publications.

[4] 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.) **4** 1–24.

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

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

[7] 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.

[8] CHEN, L. H. Y. and CHOI, K. P. (1992). Some asymptotic and large deviation results in poisson approximation. *Ann. Probab.* **20** 1867–1876.

[9] CHEN, L. H. Y., FANG, X. and SHAO, Q.-M. (2013). Moderate deviations in Poisson approximation: a first attempt. *Stat. Sin.* **23** 1523–1540.

[10] CHEN, L. H. Y., FANG, X. and SHAO, Q.-M. (2013). From Stein identities to moderate deviations. *Ann. Probab.* **41** 262–293.

[11] CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q.-M. (2011). *Normal Approximation by Stein's Method*. Probability and its Applications. Springer Berlin Heidelberg, New York.

[12] ELLIS, R. S. and NEWMAN, C. M. (1978). Limit theorems for sums of dependent random variables occurring in statistical mechanics. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*. **44** 117–139.

[13] ELLIS, R. S. and NEWMAN, C. M. (1978). The statistics of curie-weiss models. *Journal of Statistical Physics*. **19** 149–161.

[14] FANG, X., LUO, L. and SHAO, Q.-M. (2020). A refined Cramér-type moderate deviation for sums of local statistics. *Bernoulli*. **26** 2319–2352.

[15] LINNIK, Y. V. (1961). On the probability of large deviations for the sums of independent variables. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* **2** 289–306. Univ of California Press.

[16] MECKES, E. (2009). On Stein's Method for Multivariate Normal Approximation. In *High Dimensional Probability V: The Luminy Volume* **5** 153–178. IMS, Beachwood, OH.

[17] PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer Berlin Heidelberg.

[18] RAIČ, M. (2007). CLT-related large deviation bounds based on stein's method. *Adv. Appl. Probab.*. **39** 731–752.

[19] 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.

[20] SHAO, Q.-M., ZHANG, M. and ZHANG, Z.-S. (2020). Cramér-type moderate deviation theorems for nonnormal approximation. To appear in *Ann. Appl. Probab*.

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

[22] SHAO, Q.-M. and ZHANG, Z.-S. (2019). Berry–Esseen bounds of normal and nonnormal approximation for unbounded exchangeable pairs. *Ann. Probab.* **47** 61–108.

[23] STEIN, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* **2** 583–602. University of California Press, Berkeley.

[24] STEIN, C. (1986). *Approximate Computation of Expectations*. **7**. IMS, Hayward, CA.

# Appendix A: Proof of some lemmas in Section 5

## A.1. Proof of Lemmas 5.2–5.4

We first introduces some preliminary lemmas.

**Lemma A.1.** *Under the conditions of Theorem 3.1, for any $\mathcal{I} \subset \mathcal{T} \subset \{1, \ldots, n\}$ and for $0 \leqslant t \leqslant \alpha$, we have*

$$\beta^{-|\mathcal{I}|} \, \mathbb{E} \, \mathrm{e}^{tW_{\mathcal{T}}} \leqslant \mathbb{E} \, \mathrm{e}^{tW_{\mathcal{T} \setminus \mathcal{I}}} \leqslant \beta^{|\mathcal{I}|} \, \mathbb{E} \, \mathrm{e}^{tW_{\mathcal{T}}} .$$

***Proof of Lemma A.1.*** Let $\mathcal{G}_{\mathcal{T} \setminus \mathcal{I}} = \sigma \left\{ X_j, j \in \mathcal{T} \setminus \mathcal{I} \right\}$. By the total expectation formula,

$$\mathbb{E} \, \mathrm{e}^{tW_{\mathcal{T}}} = \mathbb{E} \Big\{ \mathrm{e}^{tW_{\mathcal{T} \setminus \mathcal{I}}} \, \mathbb{E} \left\{ \mathrm{e}^{tW_{\mathcal{I}}} \,\middle|\, \mathcal{G}_{\mathcal{T} \setminus \mathcal{I}} \right\} \Big\}.$$

By (3.1), it follows that, for $0 \leqslant t \leqslant \alpha$,

$$\beta^{-|\mathcal{I}|} \leqslant \mathbb{E} \left\{ \mathrm{e}^{tW_{\mathcal{I}}} \,\middle|\, \mathcal{G}_{\mathcal{T} \setminus \mathcal{I}} \right\} \leqslant \beta^{|\mathcal{I}|}.$$

Then the desired result follows. □

Let $\gamma_{p,m}(t) = \mathbb{E} \, |X_m|^p \, \mathrm{e}^{t|X_m|}$ for $p \geqslant 0$ and $t \geqslant 0$.

**Lemma A.2.** *Let $\mathcal{T}$ be a subset of $\mathcal{J}$, and let $W_{\mathcal{T}} = \sum_{j\in\mathcal{T}} X_j$. Under the conditions of Theorem 3.1, for any $i \in \mathcal{J}$ and $j \in \mathcal{A}_i$, we have for $0 \leqslant t \leqslant \alpha$ and for $q \geqslant 0$,*

$$\mathbb{E}\big\{|X_i|^q \,\mathrm{e}^{tW_{\mathcal{T}}}\big\} \leqslant \beta^{2d}\gamma_{q,i}(t)\,\mathbb{E}\,\mathrm{e}^{tW_{\mathcal{T}}}, \tag{A.1}$$

$$\mathbb{E}\big\{|X_j^{(i)}|^q \,\mathrm{e}^{tW_{\mathcal{T}}}\big\} \leqslant \beta^{4d}\gamma_{q,j}(t)\,\mathbb{E}\,\mathrm{e}^{tW_{\mathcal{T}}}. \tag{A.2}$$

*and for $q \geqslant 1$,*

$$\mathbb{E}\big\{\big|\chi_{ij}\big|^q \,\mathrm{e}^{tW_{\mathcal{T}}}\big\} \leqslant 2^{3q-1}\beta^{4d}\big(\gamma_{2q,i}(t) + \gamma_{2q,j}(t)\big)\,\mathbb{E}\,\mathrm{e}^{tW_{\mathcal{T}}}, \tag{A.3}$$

$$\mathbb{E}\big\{\big|\zeta_{ij}\big|^q \,\mathrm{e}^{tW_{\mathcal{T}}}\big\} \leqslant 2^{2q-1}\beta^{4d}\big(\gamma_{2q,i}(t) + \gamma_{2q,j}(t)\big)\,\mathbb{E}\,\mathrm{e}^{tW_{\mathcal{T}}}, \tag{A.4}$$

***Proof of Lemma A.2.*** For each $i \in \mathcal{J}$, let $\mathcal{T}_i = \mathcal{T} \cap \mathcal{A}_i$, $\mathcal{T}_i^c = \mathcal{T} \cap \mathcal{A}_i^c$, $W_{\mathcal{T}_i} = \sum_{j\in\mathcal{T}_i} X_j$ and $W_{\mathcal{T}_i^c} = \sum_{j\in\mathcal{T}_i^c} X_j$. Thus, $\mathcal{T} = \mathcal{T}_i \cup \mathcal{T}_i^c$ and $W_{\mathcal{T}} = W_{\mathcal{T}_i} + W_{\mathcal{T}_i^c}$. Let $\mathcal{F}_{\mathcal{A}_i^c} = \sigma\big(X_j, j \in \mathcal{A}_i^c\big)$ and $W_{\mathcal{T}_i\setminus\{j\}} = \sum_{k\in\mathcal{T}_i\setminus\{j\}} X_k$.

We now prove (A.1). Recall that $|\mathcal{T}_i| \leqslant |\mathcal{A}_i| \leqslant d$. By (3.1),

$$\begin{aligned}
\mathbb{E}\big\{|X_i|^q \,\mathrm{e}^{tW_{\mathcal{T}}}\big\} &= \mathbb{E}\left\{\mathbb{E}\left\{|X_i|^q \,\mathrm{e}^{tW_{\mathcal{T}}} \,\middle|\, \mathcal{F}_{\mathcal{A}_i^c}\right\}\right\} \\
&= \mathbb{E}\left\{\mathrm{e}^{tW_{\mathcal{T}_i^c}}\,\mathbb{E}\left\{|X_i|^q \,\mathrm{e}^{tW_{\mathcal{T}_i}} \,\middle|\, \mathcal{F}_{\mathcal{A}_i^c}\right\}\right\} \\
&\leqslant \beta^d\,\mathbb{E}\big\{|X_i|^q\,\mathrm{e}^{t|X_i|}\big\}\,\mathbb{E}\big\{\mathrm{e}^{tW_{\mathcal{T}_i^c}}\big\} \\
&\leqslant \beta^{2d}\,\mathbb{E}\big\{|X_i|^q\,\mathrm{e}^{t|X_i|}\big\}\,\mathbb{E}\big\{\mathrm{e}^{tW_{\mathcal{T}}}\big\},
\end{aligned} \tag{A.5}$$

where we used Lemma A.1 in the last line.

Next, we prove (A.2). Let $Y$ be a random variable such that for any $x \in \mathbb{R}$,

$$\mathbb{P}\big(Y \leqslant x \,\big|\, \mathcal{F}_{\mathcal{A}_i^c}\big) = \frac{\mathbb{E}\left\{\mathbf{1}_{\{X_j\leqslant x\}}\,\mathrm{e}^{tW_{\mathcal{T}_i\setminus\{j\}}} \,\middle|\, \mathcal{F}_{\mathcal{A}_i^c}\right\}}{\mathbb{E}\left\{\mathrm{e}^{tW_{\mathcal{T}_i\setminus\{j\}}} \,\middle|\, \mathcal{F}_{\mathcal{A}_i^c}\right\}}.$$

Since $|x|^{2q}$ and $\mathrm{e}^{t|x|}$ are both increasing functions of $|x|$, by the Kimball inequality, we have

$$\mathbb{E}\left\{|Y|^{2q} \,\middle|\, \mathcal{F}_{\mathcal{A}_i^c}\right\}\mathbb{E}\big\{\mathrm{e}^{t|Y|} \,\big|\, \mathcal{F}_{\mathcal{A}_i^c}\big\} \leqslant \mathbb{E}\big\{|Y|^{2q}\,\mathrm{e}^{t|Y|} \,\big|\, \mathcal{F}_{\mathcal{A}_i^c}\big\}.$$

By (3.1), we have

$$\mathbb{E}\big\{\mathrm{e}^{tW_{\mathcal{T}_i\setminus\{j\}}} \,\big|\, \mathcal{F}_{\mathcal{A}_i^c}\big\} \leqslant \beta^d.$$

Thus,

$$\begin{aligned}
&\mathbb{E}\big\{|X_j|^{2q}\,\mathrm{e}^{tW_{\mathcal{T}_i\setminus\{j\}}} \,\big|\, \mathcal{F}_{\mathcal{A}_i^c}\big\}\mathbb{E}\big\{\mathrm{e}^{t|X_j|+tW_{\mathcal{T}_i\setminus\{j\}}} \,\big|\, \mathcal{F}_{\mathcal{A}_i^c}\big\} \\
&\quad\leqslant \mathbb{E}\big\{|X_j|^{2q}\,\mathrm{e}^{t|X_j|+tW_{\mathcal{T}_i\setminus\{j\}}} \,\big|\, \mathcal{F}_{\mathcal{A}_i^c}\big\}\,\mathbb{E}\big\{\mathrm{e}^{tW_{\mathcal{T}_i\setminus\{j\}}} \,\big|\, \mathcal{F}_{\mathcal{A}_i^c}\big\} \\
&\quad\leqslant \beta^d\,\mathbb{E}\big\{|X_j|^{2q}\,\mathrm{e}^{t|X_j|+tW_{\mathcal{T}_i\setminus\{j\}}} \,\big|\, \mathcal{F}_{\mathcal{A}_i^c}\big\}.
\end{aligned} \tag{A.6}$$

Now, as $X_j^{(i)}$ is conditionally independent of $\{X_j, j \in \mathcal{A}_i\}$ and has the same conditional distribution as $X_j$ given $\mathcal{F}_{\mathcal{A}_i^c}$, we have

$$\begin{aligned}
&\mathbb{E}\big\{\big|X_j^{(i)}\big|^{2q} \mathrm{e}^{tW_{\mathcal{T}}} \bigm| \mathcal{F}_{\mathcal{A}_i^c}\big\} \\
&\leqslant \mathrm{e}^{tW_{\mathcal{T}_i^c}} \mathbb{E}\big\{\big|X_j\big|^{2q} \bigm| \mathcal{F}_{\mathcal{A}_i^c}\big\}\mathbb{E}\big\{\mathrm{e}^{t|X_j|+tW_{\mathcal{T}_i\setminus\{j\}}} \bigm| \mathcal{F}_{\mathcal{A}_i^c}\big\} \\
&\leqslant \beta^d \mathrm{e}^{tW_{\mathcal{T}_i^c}} \mathbb{E}\big\{\big|X_j\big|^{2q}\mathrm{e}^{tW_{\mathcal{T}_i\setminus\{j\}}} \bigm| \mathcal{F}_{\mathcal{A}_i^c}\big\}\mathbb{E}\big\{\mathrm{e}^{t|X_j|+tW_{\mathcal{T}_i\setminus\{j\}}} \bigm| \mathcal{F}_{\mathcal{A}_i^c}\big\} \\
&\leqslant \beta^{2d} \mathrm{e}^{tW_{\mathcal{T}_i^c}} \mathbb{E}\big\{|X_j|^{2q} \mathrm{e}^{t|X_j|+tW_{\mathcal{T}_i\setminus\{j\}}} \bigm| \mathcal{F}_{\mathcal{A}_i^c}\big\},
\end{aligned} \tag{A.7}$$

where we used (A.6) in the last inequality. Following the same argument as (A.5),

$$\mathbb{E}\big\{|X_j|^{2q} \mathrm{e}^{t|X_j|+t\sum_{k\in\mathcal{T}\setminus\{j\}} X_k}\big\} \leqslant \beta^{2d} \mathbb{E}\big\{|X_j|^{2q} \mathrm{e}^{t|X_j|}\big\} \mathbb{E}\,\mathrm{e}^{tW_{\mathcal{T}}},$$

which, together with (A.7), proves (A.2).

We now move to prove (A.3). By the basic inequality that $(x+y)^q \leqslant 2^{q-1}(x^q+y^q)$ for $x \geqslant 0, y \geqslant 0, q \geqslant 1$, we have

$$\begin{aligned}
\mathbb{E}\big\{\big|\chi_{ij}\big|^q \mathrm{e}^{tW_{\mathcal{T}}}\big\} \leqslant 2^{q-1}\Big(&\mathbb{E}\big\{\big|\big(X_i - X_i^{(i)}\big)\big(X_j - X_j^{(i)}\big)\big|^q \mathrm{e}^{tW_{\mathcal{T}}}\big\} \\
&+ \mathbb{E}\big\{\big|\big(X_i - X_i^{(i)}\big)\big(X_j - X_j^{(i)}\big)\big|^q\big\} \mathbb{E}\,\mathrm{e}^{tW_{\mathcal{T}}}\Big).
\end{aligned} \tag{A.8}$$

By the Cauchy inequality,

$$\begin{aligned}
&\mathbb{E}\big\{\big|\big(X_i - X_i^{(i)}\big)\big(X_j - X_j^{(i)}\big)\big|^q \mathrm{e}^{tW_{\mathcal{T}}}\big\} \\
&\leqslant \frac{1}{2}\mathbb{E}\big\{\big|X_i - X_i^{(i)}\big|^{2q} \mathrm{e}^{tW_{\mathcal{T}}}\big\} + \frac{1}{2}\mathbb{E}\big\{\big|X_j - X_j^{(i)}\big|^{2q} \mathrm{e}^{tW_{\mathcal{T}}}\big\} \\
&\leqslant 2^{2q-2}\Big(\mathbb{E}\big\{\big|X_i\big|^{2q} \mathrm{e}^{tW_{\mathcal{T}}}\big\} + \mathbb{E}\big\{\big|X_i^{(i)}\big|^{2q} \mathrm{e}^{tW_{\mathcal{T}}}\big\} \\
&\qquad + \mathbb{E}\big\{\big|X_j\big|^{2q} \mathrm{e}^{tW_{\mathcal{T}}}\big\} + \mathbb{E}\big\{\big|X_j^{(i)}\big|^{2q} \mathrm{e}^{tW_{\mathcal{T}}}\big\}\Big).
\end{aligned} \tag{A.9}$$

As for $\mathbb{E}\big\{\big|X_i^{(i)}\big|^{2q} \mathrm{e}^{tW_{\mathcal{T}}}\big\}$, since $X_i^{(i)}$ is independent of $W_{\mathcal{T}}$, it follows that

$$\mathbb{E}\big\{\big|X_i^{(i)}\big|^{2q} \mathrm{e}^{tW_{\mathcal{T}}}\big\} = \mathbb{E}\big|X_i\big|^{2q} \mathbb{E}\,\mathrm{e}^{tW_{\mathcal{T}}}. \tag{A.10}$$

Substituting (A.1), (A.2) and (A.10) to (A.9), and recalling that $\beta \geqslant 1$, we have

$$\begin{aligned}
&\mathbb{E}\big\{\big|\big(X_i - X_i^{(i)}\big)\big(X_j - X_j^{(i)}\big)\big|^q \mathrm{e}^{tW_{\mathcal{T}}}\big\} \\
&\leqslant 2^{2q-1}\beta^{4d}\Big(\mathbb{E}\big\{|X_i|^{2q} \mathrm{e}^{t|X_i|}\big\} + \mathbb{E}\big\{|X_j|^{2q} \mathrm{e}^{t|X_j|}\big\}\Big) \mathbb{E}\,\mathrm{e}^{tW_{\mathcal{T}}}.
\end{aligned} \tag{A.11}$$

When $t = 0$, by a similar argument, we have

$$\mathbb{E}\big\{\big|\big(X_i - X_i^{(i)}\big)\big(X_j - X_j^{(i)}\big)\big|^q\big\} \leqslant 2^{2q-1}\Big(\mathbb{E}\,|X_i|^{2q} + \mathbb{E}\,|X_j|^{2q}\Big). \tag{A.12}$$

By (A.8), (A.11) and (A.12), we have

$$\mathbb{E}\big\{\big|\chi_{ij}\big|^q \,\mathrm{e}^{tW_{\mathcal{T}}}\big\} \leqslant 2^{3q-1}\beta^{4d}\Big(\mathbb{E}\big\{|X_i|^{2q}\,\mathrm{e}^{t|X_i|}\big\} + \mathbb{E}\big\{|X_j|^{2q}\,\mathrm{e}^{t|X_j|}\big\}\Big)\,\mathbb{E}\,\mathrm{e}^{tW_{\mathcal{T}}}\,.$$

This proves (A.3). The inequality (A.4) follows from a similar argument. □

***Proof of Lemma 5.2.*** Without loss of generality, we only prove (5.3), because (5.4) can be shown similarly. By the Cauchy inequality and recalling that $\chi_{ij}$ is independent of $W_{ij}^c$, we have

$$\begin{aligned}\mathbb{E}\Big|\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\Big| &\leqslant \frac{1}{2}\,\mathbb{E}\big\{\chi_{ij}^2\,\mathrm{e}^{tW_{ij}^c}\big\} + \frac{1}{2}\,\mathbb{E}\big\{\chi_{i'j'}^2\,\mathrm{e}^{tW_{ij}^c}\big\}\\ &= \frac{1}{2}\,\mathbb{E}\,\chi_{ij}^2\,\mathbb{E}\,\mathrm{e}^{tW_{ij}^c} + \frac{1}{2}\,\mathbb{E}\big\{\chi_{i'j'}^2\,\mathrm{e}^{tW_{ij}^c}\big\}.\end{aligned}$$

For the first term, by (A.12) again with $q=2$, it follows that

$$\mathbb{E}\,\chi_{ij}^2 \leqslant 8\big(\mathbb{E}\,|X_i|^4 + \mathbb{E}\,|X_j|^4\big).$$

For the second term, by Lemma A.1 and (A.3), we have

$$\begin{aligned}\mathbb{E}\big\{\chi_{i'j'}^2\,\mathrm{e}^{tW_{ij}^c}\big\} &\leqslant 32\beta^{4d}\big(\gamma_{4,i'}(t) + \gamma_{4,j'}(t)\big)\,\mathbb{E}\,\mathrm{e}^{tW_{ij}^c}\\ &\leqslant 32\beta^{6d}\big(\gamma_{4,i'}(t) + \gamma_{4,j'}(t)\big)\,\mathbb{E}\,\mathrm{e}^{tW}\,.\end{aligned}$$

This completes the proof of (5.3). □

***Proof of Lemma 5.3.*** By Lemma A.1 and recalling that $|\mathcal{A}_{ijk}| \leqslant 3d$, we have for $0 \leqslant t \leqslant \alpha$,

$$\mathbb{E}\,\mathrm{e}^{tW_{ijk}^c} \leqslant \beta^{3d}\,\mathbb{E}\,\mathrm{e}^{tW}\,. \tag{A.13}$$

For any $i, j, i', j', k$ and $l$, by the Hölder inequality and (A.1), (A.3) and (A.13), we have

$$\begin{aligned}&\mathbb{E}\big\{\big|X_k X_l \chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ijk}^c}\big|\big\}\\ &\leqslant \frac{1}{6}\,\mathbb{E}\{(|X_k|^6 + |X_l|^6 + 2|\chi_{ij}|^3 + 2|\chi_{i'j'}|^3)e^{tW_{ijk}^c}\}\\ &\leqslant 88\,\beta^{4d}\Bigg\{\sum_{m\in\{i,j,i',j',k,l\}}\gamma_{6,m}(t)\Bigg\}\,\mathbb{E}\,\mathrm{e}^{tW_{ijk}^c}\\ &\leqslant 88\,\beta^{7d}\Bigg\{\sum_{m\in\{i,j,i',j',k,l\}}\gamma_{6,m}(t)\Bigg\}\,\mathbb{E}\,\mathrm{e}^{tW},\end{aligned}$$

where we used (A.13) in the last inequality. This proves the first inequality, and the second one can be shown similarly. □

***Proof of Lemma 5.4.*** We first prove (5.5). If $i' \in \mathcal{A}_{ijk}^c$ and $j' \in \mathcal{A}_{ijk}^c$, then $X_k\chi_{ij}$ is independent of $\chi_{i'j'}$ and $W_{ijk}^c$. Thus,

$$\mathbb{E}\{X_k\chi_{ij}\chi_{i'j'}e^{tW_{ijk}^c}\} = \mathbb{E}\{X_k\chi_{ij}\}\,\mathbb{E}\{\chi_{i'j'}e^{tW_{ijk}^c}\}. \tag{A.14}$$

Now we calculate $\mathbb{E}\{X_k\chi_{ij}\}$ and $\mathbb{E}\{\chi_{i'j'}e^{tW_{ijk}^c}\}$ separately.

Note that for $i' \in \mathcal{A}_{ijk}^c$ and $j' \in \mathcal{A}_{ijk}^c$, we have

$$\begin{aligned}
&\mathbb{E}\{\chi_{i'j'}\,\mathrm{e}^{tW_{ijk}^c}\}\\
&= \mathbb{E}\Big\{\chi_{i'j'}\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\setminus\mathcal{A}_{i'j'}} X_l\Big)\Big\} &\text{(A.15)}\\
&\quad + \mathbb{E}\Big\{\chi_{i'j'}\Big(\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\cap\mathcal{A}_{i'j'}} X_l\Big)-1\Big)\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\setminus\mathcal{A}_{i'j'}} X_l\Big)\Big\}. &\text{(A.16)}
\end{aligned}$$

As $\chi_{i'j'}$ and $\{X_l, l \in \mathcal{A}_{ijk}^c \setminus \mathcal{A}_{i'j'}\}$ are independent and $\mathbb{E}\,\chi_{i'j'} = 0$, we have (A.15) is

$$\mathbb{E}\Big\{\chi_{i'j'}\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\setminus\mathcal{A}_{i'j'}} X_l\Big)\Big\} = 0. \tag{A.17}$$

As for (A.16), by the inequality that $|e^x - 1| \leqslant |x|(1+e^x)$, we have

$$\begin{aligned}
&\Big|\mathbb{E}\Big\{\chi_{i'j'}\Big(\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\cap\mathcal{A}_{i'j'}} X_l\Big)-1\Big)\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\setminus\mathcal{A}_{i'j'}} X_l\Big)\Big\}\Big|\\
&\leqslant t\sum_{m\in\mathcal{A}_{ijk}^c\cap\mathcal{A}_{i'j'}} \mathbb{E}\Big\{\big|\chi_{i'j'}X_m\big|\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\setminus\mathcal{A}_{i'j'}} X_l\Big)\Big\}\\
&\qquad + t\sum_{m\in\mathcal{A}_{ijk}^c\cap\mathcal{A}_{i'j'}} \mathbb{E}\Big\{\big|\chi_{i'j'}X_m\big|\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c} X_l\Big)\Big\}.
\end{aligned}$$

By Lemma A.1, we have

$$\mathbb{E}\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\setminus\mathcal{A}_{i'j'}} X_l\Big) \leqslant \beta^{5d}\,\mathbb{E}\,e^{tW}. \tag{A.18}$$

By (A.1), (A.3) and (A.18) and the Cauchy inequality, we have

$$\mathbb{E}\Big\{\big|\chi_{i'j'}\big|^{3/2}\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\setminus\mathcal{A}_{i'j'}} X_l\Big)\Big\} \leqslant 12\beta^{9d}\big\{\gamma_{3,i'}(t) + \gamma_{3,j'}(t)\big\}\,\mathbb{E}\,\mathrm{e}^{tW},$$

and

$$\mathbb{E}\Big\{\big|X_m\big|^3\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\setminus\mathcal{A}_{i'j'}} X_l\Big)\Big\} \leqslant \beta^{9d}\gamma_{3,m}(t)\,\mathbb{E}\,\mathrm{e}^{tW}.$$

Then, by (A.18) and the inequality $|xy| \leqslant (2/3)|x|^{3/2} + (1/3)|y|^3$,

$$\begin{aligned}
&\mathbb{E}\Big\{\big|\chi_{i'j'}X_m\big|\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c\setminus\mathcal{A}_{i'j'}} X_l\Big)\Big\}\\
&\qquad \leqslant 8\,\beta^{9d}\big\{\gamma_{3,i'}(t) + \gamma_{3,j'}(t) + \gamma_{3,m}(t)\big\}\,\mathbb{E}\,\mathrm{e}^{tW},
\end{aligned} \tag{A.19}$$

and similarly,

$$\begin{aligned}&\mathbb{E}\Big\{\big|\chi_{i'j'}X_m\big|\exp\Big(t\textstyle\sum_{l\in\mathcal{A}_{ijk}^c}X_l\Big)\Big\}\\&\qquad\leqslant 8\,\beta^{9d}\big\{\gamma_{3,i'}(t)+\gamma_{3,j'}(t)+\gamma_{3,m}(t)\big\}\,\mathbb{E}\,\mathrm{e}^{tW}.\end{aligned}\tag{A.20}$$

By (A.15)–(A.17), (A.19) and (A.20),

$$\big|\mathbb{E}\big\{\chi_{i'j'}\,\mathrm{e}^{tW_{ijk}^c}\big\}\big|\leqslant 16\beta^{9d}\,t\sum_{m\in\mathcal{A}_{i'j'}}\big\{\gamma_{3,i'}(t)+\gamma_{3,j'}(t)+\gamma_{3,m}(t)\big\}\,\mathbb{E}\,\mathrm{e}^{tW}.\tag{A.21}$$

By the Cauchy inequality and by the monotonicity of $\gamma_{3,j}$, we have for $t\geqslant 0$,

$$\begin{aligned}\mathbb{E}\,|X_k\chi_{ij}|&\leqslant 11\big\{\gamma_{3,i}(0)+\gamma_{3,j}(0)+\gamma_{3,k}(0)\big\}\\&\leqslant 11\big\{\gamma_{3,i}(t)+\gamma_{3,j}(t)+\gamma_{3,k}(t)\big\}.\end{aligned}\tag{A.22}$$

By (A.14), (A.21) and (A.22), we have (5.5) is proved.

Next, we prove (5.6). For any $i,j,i',j'$ and $k$ such that $k\in\mathcal{A}_{ij}$ and $\{i',j'\}\cap\mathcal{A}_{ijk}\neq\varnothing$, by Lemma A.1 and (A.1) and (A.3) and the Cauchy inequality, we have

$$\begin{aligned}&t\big|\mathbb{E}\big\{X_k\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ijk}^c}\big\}\big|\\&\leqslant t\big|\mathbb{E}\big\{X_k\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\big\}\big|+t\big|\mathbb{E}\big\{X_k\chi_{ij}\chi_{i'j'}\big(\mathrm{e}^{tW_{ij}^c}-\mathrm{e}^{tW_{ijk}^c}\big)\big\}\big|\\&\leqslant\frac{1}{2|\mathcal{A}_{ij}|}\,\mathbb{E}\big|\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\big|+\frac{1}{2}|\mathcal{A}_{ij}|t^2\,\mathbb{E}\big|X_k^2\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\big|\\&\quad+\sum_{l\in\mathcal{A}_{ijk}}t^2\,\mathbb{E}\big|X_kX_l\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ij}^c}\big|+\sum_{l\in\mathcal{A}_{ijk}}t^2\,\mathbb{E}\big|X_kX_l\chi_{ij}\chi_{i'j'}\,\mathrm{e}^{tW_{ijk}^c}\big|.\end{aligned}$$

By (5.3) and Lemma 5.3, we complete the proof of (5.6). □

## A.2. Proof of Lemma 5.7

***Proof of Lemma 5.7.*** The first inequality (5.21) was shown by Barbour, Karoński and Ruciński [2, Eq. (3.10)]. We will apply their ideas to prove the other inequalities of this lemma. In what follows, for each $i=(i_1,\dots,i_{e(G)})\in\mathcal{I}_n$, we denote by $G(i)$ by the graph generated by the edges $\{e_{i_1},\dots,e_{i_{e(G)}}\}$ and for any $i,j\in\mathcal{I}_n$, we denote by $G(i)\cap G(j)$ the graph generated by $\{e_{i_1},\dots,e_{i_{e(G)}}\}\cap\{e_{j_1},\dots,e_{j_{e(G)}}\}$.

As for (5.22), we have

$$\begin{aligned}&\sum_{i\in\mathcal{I}_n}\sum_{j:|i\cap j|\geqslant 1}\sum_{i':|i'\cap(i\cup j)|\geqslant 1}\sum_{i':|j'\cap i'|\geqslant 1}p^{4e(G)-|i\cap j|-|i'\cap j'|-|i'\cap(i\cup j)|}\\&\quad\leqslant C\sum_{i\in\mathcal{I}_n}\sum_{j:|i\cap j|\geqslant 1}\sum_{i':|i'\cap(i\cup j)|\geqslant 1}\sum_{\substack{H\subset G(i')\\ e(H)\geqslant 1}}\sum_{i':G(i')\cap G(j')=H}p^{4e(G)-|i\cap j|-|i'\cap(i\cup j)|-e(H)}\end{aligned}$$

$$
\begin{aligned}
&\leqslant C \sum_{i\in\mathcal{I}_n}\sum_{j:|i\cap j|\geqslant 1}\sum_{i':|i'\cap(i\cup j)|\geqslant 1}\sum_{\substack{H\subset G(i')\\ e(H)\geqslant 1}} n^{v(G)-v(H)} p^{4e(G)-|i\cap j|-|i'\cap(i\cup j)|-e(H)}\\
&\leqslant C\psi_n^{-1} n^{v(G)}p^{e(G)} \sum_{i\in\mathcal{I}_n}\sum_{j:|i\cap j|\geqslant 1}\sum_{i':|i'\cap(i\cup j)|\geqslant 1} p^{3e(G)-|i\cap j|-|i'\cap(i\cup j)|}\\
&\leqslant C\sigma_n^2(\psi_n^{-1}n^{v(G)}p^{e(G)})^2,
\end{aligned}
$$

where we used (5.21) again in the last line. The inequality (5.23) follows from a similar argument.

Now, we consider (5.24). Note that for any fixed $i, i' \in \mathcal{I}_n$ and $j \in \mathcal{A}_i, j' \in \mathcal{I}_n$, with $\widetilde{G} = G(i) \cup G(j) \cup G(i') \cup G(j')$, we have

$$
\begin{aligned}
\sum_{k:|k\cap(i\cup j\cup i'\cup j')|\geqslant 1} p^{-|k\cap(i\cup j\cup i'\cup j')|} &\leqslant C \sum_{H:H\subset\widetilde{G},e(H)\geqslant 1}\sum_{k:G(k)\cap\widetilde{G}=H} p^{-e(H)}\\
&\leqslant C \sum_{H:H\subset\widetilde{G},e(H)\geqslant 1} n^{v(G)-v(H)}p^{-e(H)}\\
&\leqslant C n^{v(G)}\psi_n^{-1}.
\end{aligned}
$$

Moreover, observing that for $0 < p < 1/2$,

$$
\sigma_n^2 = \sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i} p^{2e(G)-|i\cap j|}(1-p^{|i\cap j|}) \geqslant \frac{1}{2}\sum_{i\in\mathcal{I}_n}\sum_{j\in\mathcal{A}_i} p^{2e(G)-|i\cap j|},
$$

we have

$$
\text{LHS of (5.24)} \leqslant C\sigma_n^4(\psi_n^{-1}n^{v(G)}p^{e(G)})^2.
$$

This proves (5.24) and hence completes the proof of this lemma. □

## A.3. Proof of Lemmas 5.10–5.12

***Proof of Lemma 5.10.*** By (3.6) and

$$
\mathrm{e}^{\theta x^2/2} = \frac{1}{\sqrt{2\pi\theta}}\int_{-\infty}^{\infty} \mathrm{e}^{tx-t^2/(2\theta)}\,\mathrm{d}t,
$$

it follows that

$$
\mathbb{E}\,\mathrm{e}^{\theta\xi^2/2} \leqslant \frac{1}{\sqrt{2\pi\theta}}\int_{-\infty}^{\infty} \mathrm{e}^{-t^2/2(1/\theta-1)}\,\mathrm{d}t \leqslant C_\theta.
$$

This proves (5.43). For (5.44), by the Chebyshev inequality and (5.43), taking $\theta' = (\theta+1)/2$,

$$
\mathbb{P}(|\xi| > z) \leqslant C_\theta\,\mathrm{e}^{-\theta' z^2/2}.
$$

Hence,

$$\begin{aligned}\mathbb{E}\,|\xi|^r\,\mathrm{e}^{\theta\xi^2/2} &= \int_0^\infty (ry^{r-1}+y^{r+1})\,\mathrm{e}^{\theta y^2/2}\,\mathbb{P}(|\xi|>y)\,\mathrm{d}y \\ &\leqslant C_\theta \int_0^\infty (ry^{r-1}+y^{r+1})\,\mathrm{e}^{-(1-\theta)y^2/4}\,\mathrm{d}y \\ &\leqslant C_{\theta,r}.\end{aligned}$$

This proves (5.44). The inequalities (5.45) and (5.46) follow from Shao and Zhang [22, Eqs. (5.12) and (5.23)]. □

***Proof of Lemma 5.11.*** Let $\lambda_n = (\beta/n + 2\theta/n^2)^{1/2}$. By (3.6), it follows that

$$\mathbb{E}\{e^{tT_n}\} \leqslant e^{nt^2/2}.$$

Therefore,

$$\begin{aligned}&\mathbb{E}\exp\Big(\Big(\frac{\beta}{2n}+\frac{\theta}{n^2}\Big)T_n^2+\alpha_n T_n\Big) \\ &\quad= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathbb{E}\exp\big((\lambda_n u+\alpha_n)T_n - u^2/2\big)\,\mathrm{d}u \\ &\quad\leqslant \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\big(n(\lambda_n u+\alpha_n)^2/2 - u^2/2\big)\,\mathrm{d}u \\ &\quad= \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{n\alpha_n^2/2}\int_{-\infty}^{\infty}\exp\Big(-\frac{1}{2}\Big(1-\beta-2\theta n^{-1}\Big)u^2 + n\lambda_n\alpha_n u\Big)\,\mathrm{d}u \\ &\quad\leqslant \frac{1}{\sqrt{2\pi(1-\beta-2\theta n^{-1})}}\exp\Big(\frac{n\alpha_n^2}{2}+\frac{n^2\lambda_n^2\alpha_n^2}{2(1-\beta-2\theta n^{-1})}\Big).\end{aligned}$$

Since $0\leqslant\theta<n(1-\beta)/4$, it follows that $(1-\beta-2\theta n^{-1})>(1-\beta)/2$. Hence, for $0\leqslant t\leqslant\sqrt{n}$,

$$\begin{aligned}&\mathbb{E}\exp\Big(\Big(\frac{\beta}{2n}+\frac{\theta}{n^2}\Big)T_n^2+\alpha_n T_n\Big) \\ &\quad\leqslant \frac{1}{\sqrt{\pi(1-\beta)}}\exp\Big(\frac{t^2}{2}+\frac{3n^{-1}\theta_0 t^2}{1-\beta}\Big) \qquad \text{(A.23)} \\ &\quad\leqslant C_0\,\mathrm{e}^{t^2/2}.\end{aligned}$$

By (5.45), we have

$$\begin{aligned}&\mathbb{E}\Big\{|T_n|^r\exp\Big(\Big(\frac{\beta}{2n}+\frac{\theta}{n^2}\Big)T_n^2+\alpha_n T_n\Big)\Big\} \\ &\quad= \int_0^\infty (ry^{r-1}+\lambda_n^2 y^{r+1}+\alpha_n y^r)\exp\{\lambda_n^2 y^2/2+\alpha_n y\}\,\mathbb{P}(|T_n|>y)\,\mathrm{d}y\end{aligned}$$

$$\leqslant 2\int_0^\infty \big(ry^{r-1}+\lambda_n^2 y^{r+1}+\alpha_n y^r\big)\exp\big\{\lambda_n^2y^2/2+\alpha_n y-y^2/(2n)\big\}\,\mathrm{d}y.$$

Similar to (A.23), we complete the proof of (5.48). □

***Proof of Lemma 5.12.*** The inequality (5.49) follows from Shao and Zhang [22, Eqs. (5.15) and (5.18)].

Note that with $\alpha_n = n^{-1/2}(1-\beta)^{1/2}t$ and $T_n = \xi_1+\cdots+\xi_n$, by (5.49) and (5.47), we have

$$\mathbb{E}\,\mathrm{e}^{tW} = \frac{1}{Z_n}\,\mathbb{E}\exp\Big(\frac{\beta}{2n}T_n^2+\alpha_n T_n\Big)\leqslant C\,\mathrm{e}^{t^2/2}.$$

This proves (5.50). For (5.51), it suffices to consider $i=1$. Let $M_1=\xi_2+\cdots+\xi_n$. As $n>4\beta/(1-\beta)$, it follows that $\beta(1+1/n)<(1+\beta)/2$. Then, by (5.44),

$$\mathbb{E}\big\{|\xi_1|^6\exp\big(\beta(1+1/n)\xi_1^2/2\big)\big\}\leqslant C.$$

For $0\leqslant t\leqslant\sqrt{n}$, we have $\alpha_n^2\leqslant 1$.

$$\begin{aligned}&\mathbb{E}\big\{|X_1|^6\,\mathrm{e}^{tW}\big\}\\&=\frac{1}{Z_n}\,\mathbb{E}\big\{|\xi_1|^6\,\mathrm{e}^{\frac{\beta}{2n}(\xi_1+M_1)^2+\alpha_n(\xi_1+M_1)}\big\}\\&\leqslant\frac{1}{Z_n}\,\mathbb{E}\Big\{|\xi_1|^6\exp\Big(\frac{\beta}{2n}\xi_1^2+\Big(\frac{\beta}{n}M_1+\alpha_n\Big)|\xi_1|+\frac{\beta}{2n}M_1^2+\alpha_nM_1\Big)\Big\}\end{aligned}\tag{A.24}$$

Applying the Young inequality, for $a,b\geqslant 0$, we have

$$ab\leqslant\frac{\beta(1+1/n)}{2}a+\frac{1}{2\beta(1+1/n)}b^2\leqslant\frac{\beta(1+1/n)}{2}a^2+\frac{1}{2\beta}b^2.$$

Thus, with $a=|\xi_1|$ and $b=\beta M_1/n+\alpha_n$, we have the right hand side of (A.24) can be bounded by

$$\begin{aligned}&C\,\mathbb{E}\Big\{|\xi_1|^6\,\mathrm{e}^{\frac{\beta}{2}(1+1/n)\xi_1^2}\Big\}\,\mathbb{E}\exp\Big\{\frac{\beta}{2n}M_1^2+\alpha_nM_1+\frac{1}{2\beta}\Big(\frac{\beta}{n}M_1+\alpha_n\Big)^2\Big\}\\&\qquad\leqslant C\,\mathbb{E}\exp\Big\{\Big(\frac{\beta}{2n}+\frac{\beta}{n^2}\Big)M_1^2+\alpha_nM_1\Big\}\leqslant C\,\mathrm{e}^{t^2/2}.\end{aligned}$$

where the last inequality follows from (5.47) with $\theta_0=\beta$. This proves (5.51).

For (5.52), by the definition of $X_i'$, by the Kimball inequality,

$$\begin{aligned}\mathbb{E}\,\Big\{|X_1'|^6\,\mathrm{e}^{tX_1}\,\Big|\,X_j,j\neq 1\Big\}&\leqslant\mathbb{E}\,\Big\{|X_1|^6\,\Big|\,X_j,j\neq1\Big\}\,\mathbb{E}\big\{\mathrm{e}^{t|X_1|}\,\big|\,X_j,j\neq1\big\}\\&\leqslant\mathbb{E}\big\{|X_1|^6\,\mathrm{e}^{t|X_1|}\,\big|\,X_j,j\neq1\big\}.\end{aligned}$$

This completes the proof of (5.52) by applying (5.51). □