CRAMÉR-TYPE MODERATE DEVIATIONS VIA STEIN IDENTITIES WITH APPLICATIONS TO LOCAL DEPENDENCE AND COMBINATORIAL CENTRAL LIMIT THEOREMS

Song-Hao Liu Zhuo-Song Zhang

The Chinese University of Hong Kong and National University of Singapore

Abstract

We establish a Cramér-type moderate deviation theorem for a general class of dependent random variables by using Stein identities. Our main result improves the existing results by relaxing the boundedness assumption. Optimal Cramér-type moderate deviations for sums of local dependent random variables and combinatorial central limit theorems are obtained as applications. The proof of our main result is based on Stein's method and a recursive method.

1 INTRODUCTION

Moderate deviations estimates the relative errors for distributional approximations. Since Cramér (1938) proved a moderate deviation result for tail probabilities for sums of independent random variables, Cramér-type moderate deviation theorems have been widely applied to estimate rare event probabilities. For independent and identically distributed random variables $X_1, \dots, X_n$ with zero mean and variance 1 satisfying that $\mathbb{E}e^{t_0|X_1|} \leqslant c$ for some $t_0 > 0$, it follows that

$$\left| \frac{\mathbb{P}(W_n > x)}{1 - \Phi(x)} - 1 \right| \leqslant A n^{-1/2}(1 + x^3)$$

for $0 \leqslant x \leqslant a n^{1/6}$, where $W_n = (X_1 + \cdots + X_n)/\sqrt{n}$, $\Phi(x)$ is the standard normal distribution function, and A and a are positive constants depending only on t_0 and c. We remark that the range $0 \leqslant x \leqslant a n^{1/6}$ and the error term $n^{-1/2}(1+x^3)$ are optimal for independent and identically distributed random variables. For other results on Cramér-type moderate deviation results, we refer to Linnik (1961) and Petrov (1975) .

Classical proof of Cramér-type moderate deviation is based on the conjugate method and Fourier transforms, which performs well when dealing with independent random variables. However, in practice, we need to consider dependent random fields, and in this case, Stein's method is a powerful tool to estimate convergence rates for both normal and nonnoraml approximations. Since introduced by Stein (1972), Stein's method has been widely used to prove optimal Berry–Esseen bounds and L_1 bounds with explicit constant factors for many distributional approximations. For a comprehensive survey, we refer to Chen, Goldstein and Shao (2010) and Chatterjee (2014).

Stein's method has also been applied to develop Cramér-type moderate deviation theorems. For example, Chen, Fang and Shao (2013) first applied Stein's method to prove Cramér-type moderate deviation results for normal approximation, and recently, Shao, Zhang and Zhang (2021) generalized their result to obtain a Cramér-type moderate devaition result for nonnormal approximation. In both papers, the authors made a boundedness assumption about the random variables of interest. Applying Stein's method of exchangeable approach approach, Zhang (2019) developed a Cramér-type moderate deviation result for unbounded random variables. Besides the exchangeable pair approach, other techniques in Stein's method including zero-bias couplings and size-bias couplings are also widely applied, and all of them can be expressed as a Stein identity, which is defined as follows.

Throughout this paper, let W be the random variable of interest satisfying

$$\mathbb{E}W = 0 \text{ and } \mathbb{E}W^2 = 1, \tag{1.1}$$

and suppose that there exists a random function $\hat{K}(u)$ and a random variable R such that for all absolutely continuous functions f,

$$\mathbb{E}\{Wf(W)\} = \mathbb{E}\left\{\int_{-\infty}^{\infty} f'(W+u)\hat{K}(u)du\right\} + \mathbb{E}\{Rf(W)\}. \tag{1.2}$$

This setting was first introduced in Chen, Goldstein and Shao (2010). Under this setting, both L_1 bounds and Berry–Esseen bounds have been well studied in the literature. For moderate deviations, Chen, Fang and Shao (2013) made some conditions as follows: there exists $\delta_0, \delta_1, \delta_2$ and θ such that

$$\begin{aligned} &\hat{K}(u) = 0 \text{ for } |u| > \delta_0\ , \quad |\mathbb{E}\{\hat{K}_1|W\} - 1| \leqslant \delta_1(1+|W|),\\ &\mathbb{E}\{\hat{K}_1|W\} \leqslant \theta, \qquad\qquad\qquad |\mathbb{E}\{R|W\}| \leqslant \delta_2(1+|W|). \end{aligned} \tag{1.3}$$

However, the conditions may be restrict to apply in some applications. First, the random function $\hat{K}(u)$ is assumed to be positive and supported on a bounded interval $[-\delta_0, \delta_0]$, where the constant δ_0 is of order $\mathrm{O}(n^{-1/2})$ in some typical applications. Second, the conditional expectations may not be easy to calculate if we know few on the distribution of W. Therefore, it would be interesting if we could derive a new Cramér-type moderate deviation result under a general Stein's identity setting.

In this paper, we establish a moderate deviation result for W under some mild conditions (see Theorem 2.1 below). There are several advantages of our result. First, we relaxing the boundedness assumption and thus our main result can be applied to a much wider class of statistics. Second, optimal error bounds and optimal ranges can be obtained when applying to many examples. For instance, we apply our main result to sums of local independent random variables and combinatorial central limit theorems. The details are given in Sections 3 and 4.

The idea of the proof is a combination of Stein's method and a recursive method. The recursive method has been applied to obtain optimal Berry–Esseen bounds for both 1-dimensional and multivariate normal approximations, see Raic (2003), Raič (2019) and Chen, Röllin and Xia (2020) for example. Another novel idea in the proof is that we use a truncated exponential function to prove tail probabilities. It is known that exponential-type tail probabilities play an crucial role in the proof of Cramér-type moderate deviations. In Chen, Fang and Shao (2013) and Shao, Zhang and Zhang (2021), the authors use exponential functions to prove upper bounds for such tail probabilities. In the present paper, a key observation of this paper is that the exponential function can be replaced by a smoothing truncated exponential function $\Psi_{\beta,t}$ (see Section 2) when proving exponential-type tail probabilities, and the function $\Psi_{\beta,t}$ plays an important role in relaxing the boundedness assumption when applying our main results.

The rest of this paper is organized as follows. Our main result is given in Section 2. Applications to local dependent random variables and combinatorial central limit theorems are discussed in Sections 3 and 4. We prove our main result in Section 5. Finally, the proofs of other results are postponed to Sections 6 and 7. Some supplementary materials are given in the appendix.

2 MAIN RESULTS

Let W be the random variable of interest satisfying the Stein identity (1.2) with a random function $\hat{K}(u)$ and a random variable R. For $t \geqslant 0$ and $u \in \mathbb{R}$, let

$$K(u) = \mathbb{E}\{\hat{K}(u)\}, \quad \hat{K}_1 = \int_{-\infty}^{\infty} \hat{K}(u)du, \tag{2.1}$$

$$\hat{K}_{2,t} = \int_{-\infty}^{\infty} |u|e^{t|u|}|\hat{K}(u)|du, \tag{2.2}$$

$$\hat{K}_{3,t} = \int_{|u|\leqslant 1} e^{2t|u|}\big(\hat{K}(u) - K(u)\big)^2 du, \tag{2.3}$$

$$\hat{K}_{4,t} = \int_{|u|\leqslant 1} |u|e^{2t|u|}\big(\hat{K}(u) - K(u)\big)^2 du, \tag{2.4}$$

and

$$M_t = \int_{|u|\leqslant 1} e^{t|u|}|K(u)|du. \tag{2.5}$$

For any $\beta \geqslant 0$ and $t \geqslant 0$, let

$$\Psi_{\beta,t}(w) = \begin{cases} e^{tw} + 1 & \text{if } w \leqslant \beta, \\ 2e^{t\beta} - e^{t(2\beta - w)} + 1 & \text{if } w > \beta. \end{cases} \tag{2.6}$$

We remark that the function $\Psi_{\beta,t}$ is a smoothing version of the truncated exponential function, which plays an important role in relaxing the boundedness assumption in applications.

Our main result is based on the following conditions:

(A1) Assume that there exist constants $m_0 > 0, \rho > 0$ and $r_j \geqslant 0, \tau_j \geqslant 0$ for $j = 0, 1, \ldots, 4$ such that for all $\beta, t \in [0, m_0]$,

$$\mathbb{E}\{|R|\Psi_{\beta,t}(W)\} \leqslant r_0(1+t^{\tau_0})\,\mathbb{E}\{\Psi_{\beta,t}(W)\}, \tag{2.7}$$

$$\mathbb{E}\{|\mathbb{E}\{\hat{K}_1|W\}-1|\Psi_{\beta,t}(W)\} \leqslant r_1(1+t^{\tau_1})\,\mathbb{E}\{\Psi_{\beta,t}(W)\}, \tag{2.8}$$

$$\mathbb{E}\{\hat{K}_{2,t}\Psi_{\beta,t}(W)\} \leqslant r_2(1+t^{\tau_2})\,\mathbb{E}\{\Psi_{\beta,t}(W)\}, \tag{2.9}$$

$$\mathbb{E}\{\hat{K}_{3,t}\Psi_{\beta,t}(W)\} \leqslant r_3(1+t^{\tau_3})\,\mathbb{E}\{\Psi_{\beta,t}(W)\}, \tag{2.10}$$

$$\mathbb{E}\{\hat{K}_{4,t}\Psi_{\beta,t}(W)\} \leqslant r_4(1+t^{\tau_4})\,\mathbb{E}\{\Psi_{\beta,t}(W)\} \tag{2.11}$$

and

$$\sup_{0\leqslant t\leqslant m_0} M_t \leqslant \rho. \tag{2.12}$$

Now we state our main result.

Theorem 2.1. *Under condition (A1). Let $\tau = \max\{\tau_0+1, \tau_1+2, \tau_2+3, \tau_3+1, \tau_4+1\}$ and let*

$$z_0 = \min\bigl\{m_0, 0.02e^{-\tau/2}(r_0^{1/(\tau_0+1)} + r_1^{1/(\tau_1+2)} + r_2^{1/(\tau_2+3)})^{-1}\bigr\}. \tag{2.13}$$

We have

$$\left|\frac{\mathbb{P}[W\geqslant z]}{1-\Phi(z)} - 1\right| \leqslant \left(\frac{4}{\delta(m_0)} + C(150^{\tau}+\rho)e^{\tau^2/2}\right)\delta(z) \tag{2.14}$$

for $0 \leqslant z \leqslant z_0$, where $\Phi(z)$ is the standard normal distribution function and C is an absolute constant, and

$$\begin{aligned}\delta(z) = r_0(1+z^{\tau_0+1}) + r_1(1+z^{\tau_1+2}) + r_2(1+z^{\tau_2+3})&\\ + r_3(1+z^{\tau_3+1}) + r_4^{1/2}(1+z^{\tau_4+1}).&\end{aligned} \tag{2.15}$$

We give some remarks on our main result.

Remark 2.1. Chen, Fang and Shao (2013) proved a moderate deviation for Stein identities under a boundedness assumption (1.3). Then, for $0 \leqslant t \leqslant \delta_0^{-1}$, it can be shown that (2.9) is satisfied with $r_2 = 3\theta\delta_0$. Moreover, for all $t, \beta \in (0, \delta_0^{-1})$, one can verify (see, e.g., (5.5) and (5.6) of Chen, Fang and Shao (2013)) that there exists a constant $C > 0$ depending only on θ such that

$$\mathbb{E}\bigl\{|W|\Psi_{\beta,t}(W)\bigr\} \leqslant C(1+t)\,\mathbb{E}\{\Psi_{\beta,t}(W)\},$$

then we have (2.7) and (2.8) are satisfied with $r_0 = C'\delta_2, r_1 = C'\delta_1$ and $\tau_0 = \tau_1 = 1$, where $C' > 0$ is an absolute constant. Therefore, by Theorem 2.1, we have (2.14) holds with $m_0 = \delta_0^{-1}$, $\rho = \theta$, $\tau = 3$, $\tau_0 = \tau_1 = 1$, $\tau_2 = \tau_3 = \tau_4 = 0$, $r_0 = C'\delta_2$, $r_1 = C'\delta_1$, $r_2 = 3\theta\delta_0$ and

$$r_3 = 8 \int_{|u| \leqslant \delta_0} \mathbb{E}\{(\hat{K}(u) - K(u))^2\} du, \quad r_4 = 8\delta_0 \int_{|u| \leqslant \delta_0} \mathbb{E}\{(\hat{K}(u) - K(u))^2\} du.$$

Note that this result involves two terms r_3 and $r_4^{1/2}$ that did not appear in Chen, Fang and Shao (2013). However, in many applications, both terms r_3 and $r_4^{1/2}$ has the same order as δ_0. This shows that our result Theorem 2.1 covers Theorem 3.1 of Chen, Fang and Shao (2013) with the cost of two additional terms.

3 AN APPLICATION TO SUMS OF LOCALLY DEPENDENT RANDOM VARIABLES

In this section, we prove a Cramér-type moderate deviation theorem for sums of local dependent random variables. A family of local dependent random variables means that certain subset of the random variables are independent of those outside their respective neighborhoods, which is a generalization of m-dependence. The concept of local dependence is widely applied in the literature. For example, in the graph dependency models, one may assume that the random variables is indexed by the vertices of a graph such that $\{\xi_i, i \in I\}$ and $\{\xi_j, j \in J\}$ are independent as long as $I \cap J = \varnothing$. For more examples, we refer to Baldi and Rinott (1989), Baldi, Rinott and Stein (1989), Rinott (1994), Dembo and Rinott (1996), Chen and Shao (2004) and Fang (2019).

We follow the notation in Chen and Shao (2004). Let $\mathcal{J}$ be an index set and let $\{X_i, i \in \mathcal{J}\}$ be a field of random variables with zero means and finite variances. Let $W = \sum_{i\in\mathcal{J}} X_i$ and assume that $\mathrm{Var}(W) = 1$. For $A \subset \mathcal{J}$, write $X_A = \{X_i, i \in A\}$, $A^c = \{j \in \mathcal{J} : j \notin A\}$ and denote by $|A|$ the cardinality of A.

Now we introduce the following dependence conditions:

(LD1) For each $i \in \mathcal{J}$, there exists $A_i \subset \mathcal{J}$ such that X_i is independent of $X_{A_i^c}$.

(LD2) For each $i \in \mathcal{J}$, there exists $B_i \subset \mathcal{J}$ such that $B_i \supset A_i$ and X_{A_i} is independent of $X_{B_i^c}$.

These local dependence conditions were firstly introduced by Chen and Shao (2004), and we refer to other types of local dependence structures in Baldi and Rinott (1989), Baldi, Rinott and Stein (1989), Rinott (1994), Dembo and Rinott (1996), Fang (2019). Absolute error bounds such as L_1 bounds and Berry–Esseen bounds for local dependent random variables have also been well studied in the literature. For example, in Section 4.7 of Chen, Goldstein and Shao (2010), an L_1 bound was established under (LD1) and (LD2). Chen and Shao (2004) proved several sharp Berry–Esseen bounds under different local dependence conditions and some polynomial moment conditions. Recently, Fang (2019) proved a bound for the Wasserstein-2 distances under some different dependence conditions. Although Cramér-type moderate deviations have been proved for m-dependent random variables (see, e.g., Heinrich (1982)), however, as far as we know, no Cramér-type moderate deviation results were obtained for local dependent random variables even for bounded cases.

For each $i \in \mathcal{J}$, let $Y_i = \sum_{j\in A_i} X_i$. Further, define

$$\hat{K}_i(u) = X_i\{\mathbf{1}(-Y_i \leqslant u < 0) - \mathbf{1}(0 \leqslant u \leqslant -Y_i)\}, \quad \hat{K}(u) = \sum_{i\in\mathcal{J}} \hat{K}_i(u), \tag{3.1}$$

and define

$$K_i(u) = \mathbb{E}\{\hat{K}_i(u)\}, \quad K(u) = \sum_{i\in\mathcal{J}} K_i(u). \tag{3.2}$$

Note that X_i is mean zero and that X_i and $W - Y_i$ are independent, and thus $\mathbb{E}\{X_i f(W - Y_i)\} = 0$. Therefore,

$$\begin{aligned} \mathbb{E}\{Wf(W)\} &= \sum_{i\in\mathcal{J}} \mathbb{E}\{X_i(f(W) - f(W - Y_i))\} \\ &= \mathbb{E}\int_{-\infty}^{\infty} f'(W + t)\hat{K}(t)dt. \end{aligned}$$

Hence, it follows that (1.2) holds with $R = 0$ and $\hat{K}(t)$ defined as in (3.1). Let $N_i = \{j \in \mathcal{J} : B_i \cap B_j \neq \varnothing\}$ and let $\kappa := \max_{i\in\mathcal{J}}|N_i|$. Let $n = |\mathcal{J}|$. Assume that there exist $a_n \geqslant 1$ and $b \geqslant 1$ such that for all $i \in \mathcal{J}$,

$$\mathbb{E}\Big\{\exp\Big(a_n \sum_{j\in B_i} |X_j|\Big)\Big\} \leqslant b. \tag{3.3}$$

We have the following theorem.

Theorem 3.1. *Under* (LD1) *and* (LD2), *and assume that* (3.3) *holds. Then*

$$\left|\frac{\mathbb{P}[W \geqslant z]}{1 - \Phi(z)} - 1\right| \leqslant C\delta_n(1 + z^3) \tag{3.4}$$

for $0 \leqslant z \leqslant ca_n^{1/3}\min\{1, \kappa^{-1/3}(1 + \theta_n)^{-2/3}\}$, *where* C *and* c *are absolute constants and* $\delta_n = \kappa^2 a_n^{-1}(1 + \theta_n^6)$ *and* $\theta_n = b^{1/2}n^{1/2}a_n^{-1}$.

Remark 3.1. When a_n is of order $\mathrm{O}(n^{1/2})$ and κ and b are of order $\mathrm{O}(1)$, we have $\theta_n = \mathrm{O}(1)$ and $\delta_n = \mathrm{O}(n^{-1/2})$. Therefore, the error bound in (3.4) is of order $(1+z^3)/\sqrt{n}$ and the range is $0 \leqslant z \leqslant \mathrm{O}(1)n^{1/6}$. Specially, for i.i.d. random variables $\xi_1, \ldots, \xi_n$ satisfying that $\mathbb{E}\xi_1 = 0$, $\mathrm{Var}(\xi_1) = 1/n$ and $\mathbb{E}e^{\sqrt{n}|\xi_1|} \leqslant b_0$ for some $b_0 > 0$, we have (3.3) holds with $a_n = \sqrt{n}$ and $b = b_0$. Hence, Theorem 3.1 reduces to

$$\left|\frac{\mathbb{P}(\sum_{i=1}^n \xi_i \geqslant z)}{1 - \Phi(z)} - 1\right| \leqslant Cn^{-1/2}(1 + z^3) \text{ for } 0 \leqslant z \leqslant cn^{1/6},$$

where c, C are constants depending only on b_0. Thus, Theorem 3.1 is optimal in the sense that it provides optimal error bounds and ranges for i.i.d. random variables.

To illustrate our result gives optimal error bounds and ranges for other settings, we consider the following corollary for m-dependent random fields. Let $d \geqslant 1$ and let $\mathbb{Z}^d$ denote the d-dimensional space of positive integers. For any $i = (i_1, \ldots, i_d), j = (j_1, \ldots, j_d) \in \mathbb{Z}^d$, we define the distance by $|i-j| := \max_{1 \leqslant k \leqslant d} |i_k - j_k|$, and for $A, B \subset \mathbb{Z}^d$, we define the distance between A and B by $\rho(A, B) = \inf\{|i-j| : i \in A, j \in B\}$. Let $\mathcal{J}$ be a subset of $\mathbb{Z}^d$, and we say a field of random variables $\{X_i : i \in \mathcal{J}\}$ is an *m-dependent random field* if $\{X_i, i \in A\}$ and $\{X_j, j \in B\}$ are independent whenever $\rho(A, B) > m$ for any $A, B \subset \mathcal{J}$. Choose $A_i = \{j \in \mathcal{J} : |i-j| \leqslant m\}$, $B_i = \{j \in \mathcal{J} : |i-j| \leqslant 2m\}$. We have $\kappa = (8m+1)^d$. Then, Theorem 3.1 reduces to the following corollary.

Corollary 3.2. *Let $\{X_i : i \in \mathcal{J}\}$ be an m-dependent random field with $\mathbb{E}\{X_i\} = 0$, $W = \sum_{i \in \mathcal{J}} X_i$ and $\mathrm{Var}(W) = 1$. Assume that (3.3) holds. Then, (3.4) holds with $\kappa = (8m+1)^d$.*

4 AN APPLICATION TO COMBINATORIAL CENTRAL LIMIT THEOREMS

Let $n \geqslant 1$, and let $\boldsymbol{X} := \{X_{i,j} : 1 \leqslant i, j \leqslant n\}$ be an $n \times n$ array of independent random variables with $\mathbb{E}\{X_{i,j}\} = a_{i,j}$ and $\mathrm{Var}(X_{i,j}) = c_{i,j}^2$. Moreover, assume that

$$\sum_{i=1}^{n} a_{i,j} = 0 \quad \text{for all } 1 \leqslant j \leqslant n, \qquad \sum_{j=1}^{n} a_{i,j} = 0 \quad \text{for all } 1 \leqslant i \leqslant n, \tag{4.1}$$

and

$$\frac{1}{n-1} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j}^2 + \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{i,j}^2 = 1. \tag{4.2}$$

Let $\mathcal{S}_n$ be the collection of all permutations over $[n] := \{1, 2, \ldots, n\}$ and let π be a random permutation chosen uniformly from $\mathcal{S}_n$ dependent of $\boldsymbol{X}$. Let

$$W = \sum_{i=1}^{n} X_{i,\pi(i)}. \tag{4.3}$$

Combinatorial central limit theorems for $\tilde{W} := \sum_{i=1}^{n} a_{i,\pi(i)}$, which is a special case of W, was firstly introduced by Hoeffding (1951). Goldstein (2005) proved a Berry-Esseen theorem for $\tilde{W}$ by Stein's method and zero bias coupling, and Chen, Fang and Shao (2013) also gives the moderate deviation result of the normal approximation for this special case, and the convergence rate and range depend on $\max_{i,j} |a_{ij}|$. Hu, Robinson and Wang (2007) proved a moderate deviation result for the simple random sample problem, which is a special case of the combinatorial central limit theorems. The generalized case was firstly studied by Ho and Chen (1978) who proved a Berry–Esseen bound of the central limit theorem for W using the concentration inequality approach, and Chen and Fang (2015) proved that the Kolmogorov distance is bounded by $451 n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbb{E}|X_{i,j}|^3$. In recent years, Frolov (2019) gave a Cramér-type moderate deviation result for general combinatorial central limit theorems under some Bernstein conditions. However, they didn't provide the error bounds. The following theorem provides a Cramér-type moderate deviation for W.

Theorem 4.1. *Assume that there exist $\alpha_n \geqslant 1$ and $b \geqslant 1$ such that*

$$\max_{1\leqslant i,j\leqslant n} \mathbb{E}\{\exp(\alpha_n|X_{i,j}|)\} \leqslant b. \tag{4.4}$$

Then

$$\left|\frac{\mathbb{P}(W \geqslant z)}{1-\Phi(z)} - 1\right| \leqslant C\delta_n(1+z^3), \tag{4.5}$$

for $0 \leqslant z \leqslant c\alpha_n^{1/3}\min\{1, b^{-1}(\theta_n^{-1/2}+\theta_n)^{-1}\}$, where C and c are absolute constants, $\theta_n = n^{1/2}\alpha_n^{-1}$ and $\delta_n = b^2(\alpha_n^{-1}+n^{-1/2})(\theta_n^{-2}+\theta_n^6)$.

Remark 4.1. When $X_{i,j}$ is of order $\mathrm{O}(n^{-1/2})$, $\alpha_n = \mathrm{O}(n^{1/2})$ and $b = \mathrm{O}(1)$, (4.5) reduces to

$$\left|\frac{\mathbb{P}(W \geqslant z)}{1-\Phi(z)} - 1\right| \leqslant Cn^{-1/2}(1+z^3)$$

for $0 \leqslant z \leqslant cn^{1/6}$ for some constants $c, C > 0$.

5 PROOF OF MAIN RESULT

In this section, we provide the proof of Theorem 2.1. We need to develop two preliminary lemmas: Lemmas 5.1 and 5.2, whose proofs are postponed to Section 5.3. In Lemma 5.1, we establish an upper bound of the ratio for a smoothing indicator function, and we provide a Berry–Esseen bound under (1.2) in Lemma 5.2. The proof of Theorem 2.1 is given in Section 5.2, where we mainly apply Lemma 5.1 and a smoothing inequality.

5.1 Preliminary lemmas

Let $Z \sim N(0,1)$, $\phi(w) = (1/\sqrt{2\pi})e^{-w^2/2}$ and $\Phi(w) = \int_{-\infty}^{w}\phi(t)dt$. In what follows, we write $Nh = \mathbb{E}\{h(Z)\}$ for any function h. For any $z \geqslant 0$ and $\varepsilon > 0$, let

$$h_{z,\varepsilon}(w) = \begin{cases} 1 & \text{if } w \leqslant z, \\ 0 & \text{if } w > z+\varepsilon, \\ 1+\varepsilon^{-1}(z-w) & \text{if } z < w \leqslant z+\varepsilon. \end{cases}$$

Let

$$C_0 = \sup_{0\leqslant z\leqslant z_0}\left|\frac{\mathbb{P}[W > z] - (1-\Phi(z))}{\delta(z)(1-\Phi(z))}\right|, \tag{5.1}$$

where $\delta(z)$ is given in (2.15). The following lemma gives a relative error for the test function $h_{z,\varepsilon}$.

Lemma 5.1. *Assume that condition (A1) holds and z_0 in (2.13) satisfies $z_0 \geqslant 8$. Let z be a fixed real number satisfying that $8 \leqslant z \leqslant z_0$, and let $\varepsilon := \varepsilon(z) = 40e^{\tau/2}r_2(1+z^{\tau_2})$. We have*

$$\frac{|\mathbb{E}\{h_{z,\varepsilon}(W)\} - Nh_{z,\varepsilon}|}{\delta(z)(1-\Phi(z))} \leqslant 0.75\Big(C_0 + \frac{1}{\delta(m_0)}\Big) + (184+2\rho)e^{\tau/2} + (150e^{\tau/2})^{\tau}. \tag{5.2}$$

We also need to develop a Berry–Esseen bound to prove Theorem 2.1. The following lemma is a slight modification of Theorem 2.1 in Chen, Röllin and Xia (2020), because the Stein identity (1.2) in our paper involves an additional error term $\mathbb{E}\{Rf(W)\}$ compared with that in Chen, Röllin and Xia (2020). The proof is put in Subsection 5.3, where we used a similar argument as the proof of Theorem 2.1 in Chen, Röllin and Xia (2020).

Lemma 5.2. *Let W be a random variable satisfying that $\mathbb{E}W = 0$ and $\mathbb{E}W^2 = 1$. Assume that (1.2) and (A1) hold. Then,*

$$\sup_{z\in\mathbb{R}}|\mathbb{P}(W \leqslant z) - \Phi(z)| \leqslant 4r_0 + 4r_1 + 28r_2 + 20r_3 + 13r_4^{1/2}.$$

5.2 Proof of Theorem 2.1

Proof of Theorem 2.1. It follows from Lemma 5.2 that

$$\sup_{0\leqslant z\leqslant 9}\left|\frac{\mathbb{P}(W \leqslant z) - \Phi(z)}{\delta(z)(1 - \Phi(z))}\right| \leqslant \frac{28}{1 - \Phi(9)}. \tag{5.3}$$

Now, it suffices to prove (2.14) for the case $9 \leqslant z \leqslant z_0$. When $z_0 \geqslant 8$, from (2.13), we have

$$0.02e^{-\tau/2}\min\{r_0^{-1/(\tau_0+1)}, r_1^{-1/(\tau_1+2)}, r_2^{-1/(\tau_2+3)}\} \geqslant 8, \tag{5.4}$$

and one can verify that

$$\max\{r_0, r_1, r_2\} \leqslant 0.02e^{-\tau/2}. \tag{5.5}$$

Next, we use a smoothing inequality to prove the upper bound for the case $9 \leqslant z \leqslant z_0$. Let $\varepsilon := \varepsilon(z) = 40e^{\tau/2}r_2(1 + z^{\tau_2})$. By the following well-known inequality:

$$\frac{1}{1+x}\phi(x) \leqslant 1 - \Phi(x) \leqslant \frac{1}{x}\phi(x) \quad \text{for all } x \geqslant 1, \tag{5.6}$$

we have

$$\phi(z - \varepsilon) \leqslant e^{z\varepsilon}\phi(z) \leqslant e^{z\varepsilon}(1 + z)(1 - \Phi(z)). \tag{5.7}$$

By (2.13), we have $z_0 \leqslant 0.02e^{-\tau/2}r_2^{-1/(\tau_2+3)}$, in other word

$$r_2 \leqslant \Big(\frac{0.02e^{-\tau/2}}{z_0}\Big)^{\tau_2+3}, \tag{5.8}$$

then it follows that for $z_0 \geqslant 8$,

$$\begin{aligned} z_0\varepsilon &\leqslant 40e^{\tau/2}r_2z_0 + 40e^{\tau/2}r_2z_0^{1+\tau_2} \\ &\leqslant 40e^{\tau/2}\Big(\frac{0.02e^{-\tau/2}}{z_0}\Big)^{\tau_2+3}z_0 + 40e^{\tau/2}\Big(\frac{0.02e^{-\tau/2}}{z_0}\Big)^{\tau_2+3}z_0^{1+\tau_2} \\ &\leqslant \frac{80e^{\tau/2}(0.02e^{-\tau/2})^3}{z_0^2} \\ &\leqslant 0.03, \end{aligned}$$

where we use the fact that $\tau \geqslant 3$. Thus, it follows that

$$e^{z_0\varepsilon} \leqslant 1.05. \tag{5.9}$$

Also, note that

$$(1+z^{\tau_2}) \leqslant 2z^{\tau_2} \leqslant \frac{1}{256} z^{\tau_2+3} \leqslant 0.004(1+z^{\tau_2+3}), \tag{5.10}$$

and by (5.8),

$$r_2(1+z^{\tau_2+3}) \leqslant (0.02e^{-\tau/2})^{\tau_2+3} \frac{1+z_0^{\tau_2+3}}{z_0^{\tau_2+3}} \leqslant 0.04e^{-\tau/2} \tag{5.11}$$

for $8 \leqslant z \leqslant z_0$. Then, by (5.10) and (5.11), we have

$$\begin{aligned} \varepsilon(z) &= 40e^{\tau/2} r_2(1+z^{\tau_2}) \\ &\leqslant 0.16e^{\tau/2} r_2(1+z^{\tau_2+3}) \\ &\leqslant 0.1 \end{aligned} \tag{5.12}$$

for $8 \leqslant z \leqslant z_0$. Noting that $1+z \leqslant 1.25z_0$ for all $8 \leqslant z \leqslant z_0$, we have

$$\begin{aligned} \Phi(z) - \Phi(z-\varepsilon) &\leqslant \varepsilon\phi(z-\varepsilon) \\ &\leqslant \varepsilon e^{z\varepsilon}(1+z)(1-\Phi(z)) \\ &\leqslant 1.25z_0\varepsilon e^{z_0\varepsilon}(1-\Phi(z)) \\ &\leqslant 0.05(1-\Phi(z)), \end{aligned}$$

and thus,

$$(1-\Phi(z-\varepsilon)) \leqslant (1-\Phi(z)) + (\Phi(z)-\Phi(z-\varepsilon)) \leqslant 1.05(1-\Phi(z)). \tag{5.13}$$

On the other hand, for $8 \leqslant z \leqslant z_0$, we have

$$(1+z)(1+z^{\tau_2}) \leqslant 1.25^2 z^{\tau_2+1} \leqslant \frac{1.25^2}{64} z^{\tau_2+3} \leqslant 0.04(1+z^{\tau_2+3}). \tag{5.14}$$

Then, we have

$$\begin{aligned} \Phi(z) - \Phi(z-\varepsilon) &\leqslant \varepsilon(1+z)e^{z\varepsilon}(1-\Phi(z)) \\ &\leqslant 40r_2 e^{\tau/2} e^{z\varepsilon}(1+z^{\tau_2})(1+z)(1-\Phi(z)) \\ &\leqslant 2r_2 e^{\tau/2}(1+z^{\tau_2+3})(1-\Phi(z)), \end{aligned} \tag{5.15}$$

and similarly,

$$\Phi(z+\varepsilon) - \Phi(z) \leqslant 2r_2 e^{\tau/2}(1+z^{\tau_2+3})(1-\Phi(z)). \tag{5.16}$$

Recall that C_0 is defined as in (5.1). By (5.2) and (5.16), we have for $8 \leqslant z \leqslant z_0$,

$$\begin{aligned} &\mathbb{P}[W \leqslant z] - \Phi(z) \\ &\quad \leqslant \mathbb{E}\{h_{z,\varepsilon}(W) - Nh_{z,\varepsilon}\} + \Phi(z+\varepsilon) - \Phi(z) \\ &\quad \leqslant \big(0.75(C_0 + \delta(m_0)^{-1}) + (186+2\rho)e^{\tau/2} + (150e^{\tau/2})^{\tau}\big)\delta(z)(1-\Phi(z)). \end{aligned} \tag{5.17}$$

Let $\varepsilon' = 40e^{\tau/2}r_2(1+(z-\varepsilon)^{\tau_2})$. By (5.2) with replacing z and ε by $z-\varepsilon$ and ε', respectively, and by (5.13), we have for $9 \leqslant z \leqslant z_0$,

$$\begin{aligned} &|\,\mathbb{E}\{h_{z-\varepsilon,\varepsilon'}(W) - Nh_{z-\varepsilon,\varepsilon'}\}| \\ &\quad \leqslant \big(0.75(C_0+\delta(m_0)^{-1}) + (184+2\rho)e^{\tau/2} + (150e^{\tau/2})^{\tau}\big)\delta(z)(1-\Phi(z-\varepsilon)) \\ &\quad \leqslant \big(0.8(C_0+\delta(m_0)^{-1}) + (194+2.1\rho)e^{\tau/2} + 1.05(150e^{\tau/2})^{\tau}\big)\delta(z)(1-\Phi(z)) \end{aligned} \tag{5.18}$$

Thus, by (5.15) and (5.18), we have for $9 \leqslant z \leqslant z_0$,

$$\begin{aligned} &\mathbb{P}[W \leqslant z] - \Phi(z) \\ &\quad \geqslant \mathbb{E}\{h_{z-\varepsilon,\varepsilon'}(W)\} - Nh_{z-\varepsilon,\varepsilon'} - \big(\Phi(z)-\Phi(z-\varepsilon)\big) \\ &\quad \geqslant -\big(0.8(C_0+\delta(m_0)^{-1}) + (196+2.1\rho)e^{\tau/2} + 1.05(150e^{\tau/2})^{\tau}\big)\delta(z)(1-\Phi(z)). \end{aligned} \tag{5.19}$$

By (5.17) and (5.19), we have for $9 \leqslant z \leqslant z_0$,

$$\begin{aligned} &\big|\mathbb{P}(W \leqslant z) - \Phi(z)\big| \\ &\quad \leqslant \big(0.8(C_0+\delta(m_0)^{-1}) + (196+2.1\rho)e^{\tau/2} + 1.05(150e^{\tau/2})^{\tau}\big)\delta(z)(1-\Phi(z)). \end{aligned} \tag{5.20}$$

Moving $\delta(z)(1-\Phi(z))$ in (5.20) to the LHS and Taking supremum over $9 \leqslant z \leqslant z_0$, and by (5.3), we have

$$C_0 \leqslant \big(0.8(C_0+\delta(m_0)^{-1}) + (530+2.1\rho)e^{\tau/2} + 1.05(150e^{\tau/2})^{\tau}\big) + \frac{28}{1-\Phi(9)}. \tag{5.21}$$

Solving the recursive inequality (5.21), we obtain

$$C_0 \leqslant 4\delta(m_0)^{-1} + C(150^{\tau}+\rho)e^{\tau^2/2},$$

which proves (2.14). □

5.3 Proofs of Lemma 5.1 and 5.2

We first prove the following lemmas, which are useful in the proof of Lemma 5.1. The first lemma gives an upper bound for the truncated exponential moment $\mathbb{E}\{\Psi_{\beta,t}(W)\}$.

Lemma 5.3 (Exponential bound)**.** *Assume that conditions* (2.7)–(2.9) *hold and z_0 in* (2.13) *satisfies $z_0 \geqslant 8$. We have*

$$\mathbb{E}\{\Psi_{\beta,t}(W)\} \leqslant 4e^{t^2/2} \quad \textit{for } 0 \leqslant t \leqslant z_0. \tag{5.22}$$

Proof. Since $z_0 \geqslant 8$, we have (5.5) holds. Let

$$\delta_1(t) = r_0(1+t^{\tau_0+1}) + r_1(1+t^{\tau_1+2}) + r_2(1+t^{\tau_2+3}) \quad \text{for } t \geqslant 0.$$

Then, recalling that $\tau \geqslant 3$ and that z_0 is defined in (2.13), by the similar argument as (5.11), we have $\delta_1(z_0) \leqslant 0.1e^{-\tau/2} \leqslant 0.1$. We prove a more general result as follows: for all $0 \leqslant t \leqslant m_0$,

$$\mathbb{E}\{\Psi_{\beta,t}(W)\} \leqslant 2e^{t^2/2+2\delta_1(t)}, \tag{5.23}$$

which, together with the fact that $\delta_1(z_0) \leqslant 0.1$, implies (5.22) immediately.

Now it suffices to prove (5.23). For $t \geqslant 0$, let $h(t) = \mathbb{E}\{\Psi_{\beta,t}(W)\}$. As $\Psi_{\beta,t}(w) \leqslant 2e^{t\beta} + 1$, then $h(t) < \infty$ for all $0 \leqslant t \leqslant m_0$. Write

$$\Psi'_{\beta,t}(w) = \frac{\partial}{\partial w}\Psi_{\beta,t}(w), \quad \Psi''_{\beta,t}(w) = \frac{\partial^2}{\partial w^2}\Psi_{\beta,t}(w).$$

By the definition of $\Psi_{\beta,t}(W)$,

$$\frac{\partial}{\partial t}\Psi_{\beta,t}(w) = \begin{cases} we^{tw} & \text{if } w \leqslant \beta, \\ 2\beta e^{t\beta} - (2\beta - w)e^{t(2\beta-w)} & \text{if } w > \beta, \end{cases} \tag{5.24}$$

$$\Psi'_{\beta,t}(w) = \begin{cases} te^{tw} & \text{if } w \leqslant \beta, \\ te^{t(2\beta-w)} & \text{if } w > \beta, \end{cases} \tag{5.25}$$

and

$$\Psi''_{\beta,t}(w) = \begin{cases} t^2e^{tw} & \text{if } w \leqslant \beta, \\ -t^2e^{t(2\beta-w)} & \text{if } w > \beta. \end{cases} \tag{5.26}$$

Also,

$$\frac{\partial}{\partial t}\Psi_{\beta,t}(w) \leqslant w(\Psi_{\beta,t}(w) - 1), \quad \Psi'_{\beta,t}(w) \leqslant t\Psi_{\beta,t}(w). \tag{5.27}$$

By (5.24) and the first inequality of (5.27), it follows that

$$h'(t) = \mathbb{E}\Big\{\frac{\partial}{\partial t}\Psi_{\beta,t}(W)\Big\} \leqslant \mathbb{E}\big\{W(\Psi_{\beta,t}(W) - 1)\big\}. \tag{5.28}$$

By (1.2) and (5.27) and noting that $\mathbb{E}W = 0$ and $|\Psi_{\beta,t}(w) - 1| \leqslant \Psi_{\beta,t}(w)$ for all $w \in \mathbb{R}$, we have

$$\begin{aligned} \mathbb{E}\big\{W(\Psi_{\beta,t}(W) - 1)\big\} &= \mathbb{E}\Big\{\int_{-\infty}^{\infty}\Psi'_{\beta,t}(W+u)\hat{K}(u)du\Big\} + \mathbb{E}\big\{R(\Psi_{\beta,t}(W) - 1)\big\} \\ &= \mathbb{E}\Big\{\int_{-\infty}^{\infty}\Psi'_{\beta,t}(W)\hat{K}(u)du\Big\} + \mathbb{E}\big\{R(\Psi_{\beta,t}(W) - 1)\big\} \\ &\quad + \mathbb{E}\Big\{\int_{-\infty}^{\infty}\big(\Psi'_{\beta,t}(W+u) - \Psi'_{\beta,t}(W)\big)\hat{K}(u)du\Big\} \\ &\leqslant th(t) + t\,\mathbb{E}\{|\,\mathbb{E}\{\hat{K}_1|W\} - 1|\Psi_{\beta,t}(W)\} + \mathbb{E}\{|R\Psi_{\beta,t}(W)|\} \\ &\quad + \mathbb{E}\Big\{\int_{-\infty}^{\infty}\big(\Psi'_{\beta,t}(W+u) - \Psi'_{\beta,t}(W)\big)\hat{K}(u)du\Big\}. \end{aligned} \tag{5.29}$$

By (5.26), for all $w \in \mathbb{R}$, we have

$$\begin{aligned} \big|\Psi'_{\beta,t}(w+u) - \Psi'_{\beta,t}(w)\big| &\leqslant |u| \sup_{s\leqslant|u|}\big|\Psi''_{\beta,t}(w+s)\big| \\ &\leqslant |u|t^2 \sup_{s\leqslant|u|}\Psi_{\beta,t}(w+s) \\ &\leqslant |u|t^2e^{t|u|}\Psi_{\beta,t}(w). \end{aligned} \tag{5.30}$$

By (2.2), (2.9) and (5.30), we have the last term of (5.29) can be bounded by

$$\begin{aligned} & \Big|\mathbb{E}\Big\{\int_{-\infty}^{\infty}\big(\Psi'_{\beta,t}(W+u)-\Psi'_{\beta,t}(W)\big)\hat{K}(u)du\Big\}\Big| \\ & \qquad \leqslant t^2\,\mathbb{E}\{\hat{K}_{2,t}\Psi_{\beta,t}(W)\} \leqslant r_2t^2(1+t^{\tau_2})\,\mathbb{E}\{\Psi_{\beta,t}(W)\}. \end{aligned} \tag{5.31}$$

Substituting (2.7), (2.8), (5.29) and (5.31) to (5.28) yields

$$h'(t) \leqslant th(t) + r_0(1+t^{\tau_0})h(t) + r_1t(1+t^{\tau_1})h(t) + r_2t^2(1+t^{\tau_2})h(t).$$

Solving the differential inequality yields

$$h(t) \leqslant 2\exp\Big\{\frac{t^2}{2} + r_0\Big(t+\frac{t^{\tau_0+1}}{\tau_0+1}\Big) + r_1\Big(\frac{t^2}{2}+\frac{t^{\tau_1+2}}{\tau_1+2}\Big) + r_2\Big(\frac{t^3}{3}+\frac{t^{\tau_2+3}}{\tau_2+3}\Big)\Big\}. \tag{5.32}$$

Since $\tau_0, \tau_1, \tau_2 \geqslant 0$, by Young's inequality, we have

$$\begin{aligned} t + \frac{t^{\tau_0+1}}{\tau_0+1} &\leqslant 2(1+t^{\tau_0+1}), \\ \frac{t^2}{2} + \frac{t^{\tau_1+2}}{\tau_1+2} &\leqslant (1+t^{\tau_1+2}), \\ \frac{t^3}{3} + \frac{t^{\tau_2+3}}{\tau_2+3} &\leqslant (1+t^{\tau_2+3}). \end{aligned} \tag{5.33}$$

Combining (5.32) and (5.33) yields (5.23), as desired. □

The following lemma gives an error bound for the difference of tail probabilities of W and Z.

Lemma 5.4. *Assume that conditions in Lemma 5.1 hold. For $8 \leqslant z \leqslant z_0$, $0 \leqslant \varepsilon \leqslant 2$, $|u| \leqslant 1$ and $u \wedge 0 \leqslant s \leqslant u \vee 0$, we have*

$$|\mathbb{P}[W+s>z] - \mathbb{P}[Z+s>z]| \leqslant 2e^{\tau/2}e^{z|u|}\delta(z)(1-\Phi(z))\big(C_0+c_0\big) \tag{5.34}$$

and

$$|\mathbb{P}[W+s>z+\varepsilon] - \mathbb{P}[Z+s>z+\varepsilon]| \leqslant 2e^{\tau/2}e^{z|u|+z\varepsilon}\delta(z)(1-\Phi(z))\big(C_0+c_0\big), \tag{5.35}$$

where C_0 is defined as in (5.1), τ is as in Theorem 2.1, and $c_0 := 1/\delta(m_0) + (150e^{\tau/2})^{\tau}$.

Proof of Lemma 5.4. We first introduce some inequalities. For $z \geqslant 8$ and $0 \leqslant a \leqslant 3$, we have

$$\begin{aligned} 1-\Phi(z-a) &\leqslant \frac{1}{z-a}\phi(z-a) \leqslant \frac{e^{za}}{z-a}\phi(z) \\ &\leqslant \frac{(1+z^2)e^{za}}{z(z-a)}(1-\Phi(z)) \leqslant 2e^{za}(1-\Phi(z)). \end{aligned} \tag{5.36}$$

Moreover, noting that

$$(1+(z+3)^\ell) \leqslant 1.1(z+3)^\ell \leqslant 1.1\times 1.375^\ell z^\ell \leqslant e^{\tau/2}(1+z^\ell) \quad \text{for all } 1\leqslant \ell\leqslant\tau \text{ and } z\geqslant 8,$$

and by the definition of $\delta(z)$ as in Theorem 2.1, we have

$$\delta(z+a)\leqslant e^{\tau/2}\delta(z) \quad \text{for all } 0\leqslant a\leqslant 3 \text{ and } z\geqslant 8. \tag{5.37}$$

We first prove (5.34). To this end, we consider three cases.

(1). If $s>0$, then by (5.36) and noting that $|s|\leqslant|u|\leqslant 1$, we have

$$\begin{aligned} |\mathbb{P}[W+s>z]-\mathbb{P}[Z+s>z]| &\leqslant C_0\delta(z-s)(1-\Phi(z-s)) \\ &\leqslant 2C_0e^{z|u|}\delta(z)(1-\Phi(z)). \end{aligned} \tag{5.38}$$

(2). If $s<0$ and $z-s\leqslant z_0$, by (5.37) and noting that $|s|\leqslant 1$,

$$\begin{aligned} |\mathbb{P}[W+s>z]-\mathbb{P}[Z+s>z]| &\leqslant C_0\delta(z-s)(1-\Phi(z-s)) \\ &\leqslant C_0e^{\tau/2}\delta(z)(1-\Phi(z)). \end{aligned} \tag{5.39}$$

(3). If $s<0$ but $z-s>z_0$, then it follows that $z_0\geqslant z\geqslant 8$ and $|z-z_0|\leqslant|s|\leqslant|u|\leqslant 1$. By (5.1), (5.36) and (5.37),

$$\begin{aligned} |\,\mathbb{P}[W+s>z]-\mathbb{P}[Z+s>z]| &\leqslant \mathbb{P}[W>z_0]+\mathbb{P}[Z>z_0] \\ &\leqslant (1-\Phi(z_0))+C_0\delta(z_0)(1-\Phi(z_0))+1-\Phi(z_0) \\ &\leqslant 2e^{z|u|}(1-\Phi(z))(1+e^{\tau/2}C_0\delta(z)). \end{aligned}$$

By (2.13), we have

$$z_0\geqslant 0.02e^{-\tau/2}\min\Big\{50e^{\tau/2}m_0, \frac{1}{3}r_0^{-1/(\tau_0+1)}, \frac{1}{3}r_1^{-1/(\tau_1+2)}, \frac{1}{3}r_2^{-1/(\tau_2+3)}\Big\};$$

hence, by (2.15) and recalling that $c_0=1/\delta(m_0)+(150e^{\tau/2})^\tau$, we have

$$\frac{1}{\delta(z_0)}\leqslant c_0.$$

Now, by (5.37) and the fact that $z\leqslant z_0\leqslant z+1$, we have

$$1=\frac{\delta(z)}{\delta(z)}\leqslant e^{\tau/2}\delta(z)\frac{1}{\delta(z_0)}\leqslant c_0e^{\tau/2}\delta(z).$$

Therefore, it follows that

$$|\,\mathbb{P}[W+s>z]-\mathbb{P}[Z+s>z]|\leqslant 2(C_0+c_0)e^{\tau/2}e^{z|u|}\delta(z)(1-\Phi(z)). \tag{5.40}$$

Combining (5.38)–(5.40) yields (5.34). The inequality (5.35) can be shown similarly. □

Now we are ready to give the proof of Lemma 5.1.

Proof of Lemma 5.1. We first introduce some notation and inequalities. We fix $8 \leqslant z \leqslant z_0$ in this proof. Since $z_0 \geqslant 8$, we have (5.5) holds. Also, choose $\beta := z_0$ in the function $\Psi_{\beta,t}(w)$, and let $\varepsilon = 40e^{\tau/2}r_2(1+z^{\tau_2})$. By (5.9) and (5.12), we have $e^{\beta\varepsilon} \leqslant 1.05$ and $\varepsilon \leqslant 0.1$.

Now, consider the Stein equation

$$f'(w) - wf(w) = h_{z,\varepsilon}(w) - \mathbb{E}\{h_{z,\varepsilon}(Z)\}, \tag{5.41}$$

and let $f := f_{z,\varepsilon}$ be its solution. Let $g(w) = wf(w)$ and let $v(w) = (2\pi)^{-1/2}\int_0^w se^{-(z+\varepsilon-\varepsilon s)^2/2}ds$. Recall that $Nh = \mathbb{E}h(Z)$, $Z \sim N(0,1)$, $\phi(\cdot)$ is the standard normal probability density function and $\Phi(\cdot)$ is the standard normal distribution function. It can be shown that (see, e.g., Lemma 5.3 of Chen and Shao (2004))

$$Nh_{z,\varepsilon} = \Phi(z) + \varepsilon v(1) = \Phi(z) + \int_z^{z+\varepsilon}\Big(1 + \frac{z-s}{\varepsilon}\Big)\phi(s)ds, \tag{5.42}$$

$$f(w) = \begin{cases} \dfrac{\Phi(w)}{\phi(w)}\big(1 - Nh_{z,\varepsilon}\big) & \text{if } w \leqslant z, \\ \dfrac{1-\Phi(w)}{\phi(w)}Nh_{z,\varepsilon} - \dfrac{\varepsilon}{\phi(w)}v\Big(1 + \dfrac{z-w}{\varepsilon}\Big) & \text{if } z < w \leqslant z+\varepsilon, \\ \dfrac{1-\Phi(w)}{\phi(w)}Nh_{z,\varepsilon} & \text{if } w > z+\varepsilon, \end{cases} \tag{5.43}$$

and

$$g'(w) = \begin{cases} \Big(\dfrac{(1+w^2)\Phi(w)}{\phi(w)} + w\Big)\big(1 - Nh_{z,\varepsilon}\big) & \text{if } w \leqslant z, \\ \Big(\dfrac{(1+w^2)(1-\Phi(w))}{\phi(w)} - w\Big)Nh_{z,\varepsilon} & \\ \quad -\dfrac{\varepsilon(1+w^2)}{\phi(w)}v\Big(1 + \dfrac{z-w}{\varepsilon}\Big) + \dfrac{w(z-w+\varepsilon)}{\varepsilon} & \text{if } z < w \leqslant z+\varepsilon, \\ \Big(\dfrac{(1+w^2)(1-\Phi(w))}{\phi(w)} - w\Big)Nh_{z,\varepsilon} & \text{if } w > z+\varepsilon. \end{cases} \tag{5.44}$$

Thus, by (1.2) and (5.41),

$$\begin{aligned} &\big|\mathbb{E}\{h_{z,\varepsilon}(W)\} - Nh_{z,\varepsilon}\big| \\ &\quad = \big|\mathbb{E}\{f'(W) - Wf(W)\}\big| \\ &\quad = \bigg|\mathbb{E}\{f'(W)\} - \mathbb{E}\bigg\{\int_{-\infty}^{\infty} f'(W+u)\hat{K}(u)du\bigg\} - \mathbb{E}\{Rf(W)\}\bigg| \\ &\quad \leqslant |I_1| + |I_2| + |I_3|, \end{aligned} \tag{5.45}$$

where

$$I_1 = \mathbb{E}\bigg\{\int_{-\infty}^{\infty} f'(W+u) - f'(W)\hat{K}(u)du\bigg\}, \quad I_2 = \mathbb{E}\{f'(W)(1-\hat{K}_1)\}, \quad I_3 = \mathbb{E}\{Rf(W)\}.$$

For I_1, by (5.41), we have

$$I_1 = I_{11} + I_{12} + I_{13} + I_{14}, \tag{5.46}$$

where

$$\begin{aligned}
I_{11} &= \mathbb{E}\Big\{\int_{-\infty}^{\infty}\big(g(W+u) - g(W)\big)\hat{K}(u)du\Big\},\\
I_{12} &= \mathbb{E}\Big\{\int_{|u|>1}\big(h_{z,\varepsilon}(W+u) - h_{z,\varepsilon}(W)\big)\hat{K}(u)du\Big\},\\
I_{13} &= \mathbb{E}\Big\{\int_{|u|\leqslant 1}\big(h_{z,\varepsilon}(W+u) - h_{z,\varepsilon}(W)\big)K(u)du\Big\},\\
I_{14} &= \mathbb{E}\Big\{\int_{|u|\leqslant 1}\big(h_{z,\varepsilon}(W+u) - h_{z,\varepsilon}(W)\big)(\hat{K}(u) - K(u))du\Big\}.
\end{aligned}$$

In what follows, we prove the following inequalities:

$$|I_{11}| \leqslant 41 r_2(1 + z^{\tau_2+3})(1 - \Phi(z)), \tag{5.47}$$

$$|I_{12}| \leqslant r_2(1 + z^{\tau_2+3})(1 - \Phi(z)), \tag{5.48}$$

$$|I_{13}| \leqslant 0.31(C_0 + c_0)\delta(z)(1 - \Phi(z)) + (1 + 2\rho e^{\tau/2})r_2(1 + z^{\tau_2+3})(1 - \Phi(z)), \tag{5.49}$$

$$|I_{14}| \leqslant (0.44C_0 + 0.44c_0 + 100e^{\tau/2})\delta(z)(1 - \Phi(z)), \tag{5.50}$$

$$|I_2| \leqslant 66 r_1(1 + z^{\tau_1+2})(1 - \Phi(z)), \tag{5.51}$$

$$|I_3| \leqslant 82 r_0(1 + z^{\tau_0+1})(1 - \Phi(z)). \tag{5.52}$$

Combining the (5.47)–(5.52), we complete the proof of Lemma 5.1. Now, it suffices to prove (5.47)–(5.52). We remark that we use a recursive method in the proofs of (5.49) and (5.50), and the proofs for I_{11}, I_{12}, I_2 and I_3 are routine.

(i) Proof of (5.47). For I_{11}, we have

$$|I_{11}| \leqslant \Big|\mathbb{E}\Big\{\int_{-\infty}^{\infty}\int_0^u g'(W+s)\hat{K}(u)dsdu\Big\}\Big| \leqslant I_{111} + I_{112} + I_{113},$$

where

$$\begin{aligned}
I_{111} &= \Big|\mathbb{E}\Big\{\int_{-\infty}^{\infty}\int_0^u g'(W+s)\,\mathbf{1}(W+s\leqslant 0)\hat{K}(u)dsdu\Big\}\Big|,\\
I_{112} &= \Big|\mathbb{E}\Big\{\int_{-\infty}^{\infty}\int_0^u g'(W+s)\,\mathbf{1}(0 < W+s\leqslant z)\hat{K}(u)dsdu\Big\}\Big|,\\
I_{113} &= \Big|\mathbb{E}\Big\{\int_{-\infty}^{\infty}\int_0^u g'(W+s)\,\mathbf{1}(W+s > z)\hat{K}(u)dsdu\Big\}\Big|.
\end{aligned}$$

Now, we bound these terms separately.

(1) Bound of I_{111}. Observe that

$$0 \leqslant \frac{(1+w^2)\Phi(w)}{\phi(w)} + w \leqslant 2 \quad \text{for } w \leqslant 0,$$

and thus, by (5.44), $|g'(w)|\,\mathbf{1}[w \leqslant 0] \leqslant 2(1 - Nh_{z,\varepsilon})$. By (2.9) with $t = 0$, noting that $\Psi_{\beta,0}(w) \equiv 2$, we have

$$\int_{-\infty}^{\infty} \mathbb{E}\{|u\hat{K}(u)|\} du \leqslant 2r_2. \tag{5.53}$$

Moreover,

$$1 - Nh_{z,\varepsilon} \leqslant 1 - \Phi(z). \tag{5.54}$$

Thus,

$$|I_{111}| \leqslant 2(1 - Nh_{z,\varepsilon})\,\mathbb{E}\left\{\int_{-\infty}^{\infty} |u\hat{K}(u)|du\right\} \leqslant 2r_2(1 - \Phi(z)). \tag{5.55}$$

(2) Bound of I_{112}. Observe that

$$0 \leqslant \frac{(1+w^2)}{\phi(w)} + w \leqslant 3(1+w^2)e^{w^2/2} \quad \text{for } 0 \leqslant w \leqslant z. \tag{5.56}$$

For any $0 \leqslant a \leqslant b \leqslant z$ and for any $u \wedge 0 \leqslant s \leqslant u \vee 0$, we have

$$\begin{aligned}
&\mathbb{E}\{(1 + (W+s)^2)e^{(W+s)^2/2}|\hat{K}(u)|\,\mathbf{1}(a \leqslant W+s \leqslant b)\} \\
&\quad \leqslant (1+b^2)\,\mathbb{E}\{|\hat{K}(u)|e^{(W+s)^2/2 - b(W+s) + b(W+s)}\,\mathbf{1}(a \leqslant W+s \leqslant b)\} \\
&\quad \leqslant (1+b^2)e^{a^2/2 - ab}\,\mathbb{E}\{|\hat{K}(u)|e^{b(W+s)}\,\mathbf{1}(a \leqslant W+s \leqslant b)\} \\
&\quad \leqslant (1+b^2)e^{(b-a)^2/2}e^{-b^2/2}\,\mathbb{E}\{|\hat{K}(u)|e^{b|u|}\Psi_{z_0,b}(W)\}.
\end{aligned} \tag{5.57}$$

Noting that for $u \wedge 0 \leqslant s \leqslant u \vee 0$, by (5.44), (5.54) and (5.56) and applying (5.57) with $a = j-1, b = j$ and $a = \lfloor z \rfloor, b = z$, respectively, we have

$$\begin{aligned}
&|\,\mathbb{E}\{g'(W+s)\hat{K}(u)\,\mathbf{1}(0 \leqslant W+s \leqslant z)\}| \\
&\quad \leqslant 3(1-\Phi(z))\,\mathbb{E}\{(1+(W+s)^2)e^{(W+s)^2/2}|\hat{K}(u)|\,\mathbf{1}(0 \leqslant W+s \leqslant z)\} \\
&\quad \leqslant 3(1-\Phi(z))\sum_{j=1}^{\lfloor z \rfloor}\mathbb{E}\{(1+(W+s)^2)e^{(W+s)^2/2}|\hat{K}(u)|\,\mathbf{1}(j-1 \leqslant W+s \leqslant j)\} \\
&\qquad + 3(1-\Phi(z))\,\mathbb{E}\{(1+(W+s)^2)e^{(W+s)^2/2}|\hat{K}(u)|\,\mathbf{1}(\lfloor z \rfloor \leqslant W+s \leqslant z)\} \\
&\quad \leqslant 3e^{1/2}(1-\Phi(z))\sum_{j=1}^{\lfloor z \rfloor}(1+j^2)e^{-j^2/2}\,\mathbb{E}\{|\hat{K}(u)|e^{j|u|}\Psi_{z_0,j}(W)\} \\
&\qquad + 3e^{1/2}(1-\Phi(z))(1+z^2)e^{-z^2/2}\,\mathbb{E}\{|\hat{K}(u)|e^{z|u|}\Psi_{z_0,z}(W)\}.
\end{aligned}$$

Thus, by the definition of I_{112}, and by (2.9) and Lemma 5.3, we have for $0 \leqslant z \leqslant z_0$,

$$\begin{aligned} I_{112} &\leqslant 5(1-\Phi(z)) \sum_{j=1}^{\lfloor z \rfloor} (1+j^2) e^{-j^2/2}\, \mathbb{E}\big\{\hat{K}_{2,j} \Psi_{z_0,j}(W)\big\} \\ &\quad + 5(1-\Phi(z))(1+z^2) e^{-z^2/2}\, \mathbb{E}\big\{\hat{K}_{2,z} \Psi_{z_0,z}(W)\big\} \\ &\leqslant 20 r_2 (1-\Phi(z)) \bigg(\sum_{j=1}^{\lfloor z \rfloor} (1+j^2)(1+j^{\tau_2}) + (1+z^2)(1+z^{\tau_2}) \bigg). \end{aligned}$$

For all $\ell \geqslant 0$ and $z \geqslant 8$, it can be shown that

$$\begin{aligned} &\sum_{j=1}^{[z]} (1+j^2)(1+j^\ell) \\ &\qquad \leqslant \sum_{j=1}^{[z]} 1 + j^2 + j^\ell + j^{2+\ell} \leqslant z + \frac{z^3}{3} + \frac{z^2}{2} + \frac{z}{6} + \sum_{j=1}^{[z]} (j^\ell + j^{2+\ell}) \\ &\qquad \leqslant 0.51 z^{\ell+3} + \sum_{j=1}^{[z]} (j^\ell + j^{2+\ell}) \end{aligned} \tag{5.58}$$

and for any $m \geqslant 0$ and $n \geqslant 1$ we have

$$\sum_{j=1}^{n} j^m \leqslant n^m + \sum_{j=1}^{n-1} \int_j^{j+1} x^m dx = n^m + \frac{1}{m+1} n^{m+1} \leqslant \Big(\frac{1}{m+1} + \frac{1}{n} \Big) n^{m+1} \tag{5.59}$$

Recalling that $z \geqslant 8$ and $\ell \geqslant 0$, by (5.59) with $n = [z]$ and $m = \ell$ or $\ell + 2$, we have

$$\begin{aligned} \sum_{j=1}^{[z]} (1+j^2)(1+j^\ell) &\leqslant 0.51 z^{\ell+3} + \Big(\frac{1}{\ell+1} + \frac{1}{8} \Big) z^{\ell+1} + \Big(\frac{1}{\ell+1} + \frac{1}{8} \Big) z^{\ell+3} \\ &\leqslant \Big(0.51 + \frac{1}{(\ell+1) * 8^2} + \frac{1}{8^3} + \frac{1}{\ell+3} + \frac{1}{8} \Big) z^{\ell+3} \leqslant z^{\ell+3}. \end{aligned} \tag{5.60}$$

Then, for all $\ell \geqslant 0$ and $z \geqslant 8$,

$$\begin{aligned} &\sum_{j=1}^{[z]} (1+j^2)(1+j^\ell) \leqslant (1+z^{\ell+3}), \\ &(1+z^2)(1+z^\ell) \leqslant 2\Big(1 + \frac{1}{64}\Big) z^{\ell+2} \leqslant \frac{2.032}{8} z^{\ell+3} \leqslant 0.26(1+z^{\ell+3}). \end{aligned} \tag{5.61}$$

Thus,

$$I_{112} \leqslant 26 r_2 (1 + z^{\tau_2+3})(1-\Phi(z)). \tag{5.62}$$

(3) Bound of I_{113}. According to Eqs. (4.5) and (4.6) in Chen and Shao (2004) , we have $|f(w)| \leqslant 1$ and $|f'(w)| \leqslant 1$ for $w \in \mathbb{R}$. Thus, recalling that $g(w) = wf(w)$ and by the fact that $\varepsilon \leqslant 1$, we have

$$|g'(w)| \leqslant |f(w) + wf'(w)| \leqslant 1 + z + \varepsilon \leqslant 4(z+1) \quad \text{if } z + \varepsilon \geqslant w \geqslant z. \tag{5.63}$$

By (5.44) and (5.63) and the fact that

$$\Big| \frac{(1+w^2)(1-\Phi(w))}{\phi(w)} - w \Big| \leqslant 1 \text{ for } w \geqslant 8, \tag{5.64}$$

we have

$$|g'(w)| \leqslant 4(z+1) \quad \text{if } w \geqslant z. \tag{5.65}$$

For any $\ell \geqslant 0$ and $z \geqslant 8$, we have

$$(1+z)^2(1+z^{\ell}) \leqslant 2 \times 1.125^2 z^{\ell+2} \leqslant 0.32(1+z^{\ell+3}). \tag{5.66}$$

By (2.9), (5.6), (5.22), (5.65) and (5.66) and the Markov's inequality,

$$\begin{aligned}
I_{113} &\leqslant 4(1+z)\,\mathbb{E}\Big\{ \int_{-\infty}^{\infty} \int_{0\wedge u}^{0\vee u} \mathbf{1}(W+s>z)|\hat{K}(u)|du \Big\} \\
&\leqslant 4(1+z)\Psi_{z_0,z}(z)^{-1}\,\mathbb{E}\Big\{ \int_{-\infty}^{\infty} |u\hat{K}(u)|\Psi_{z_0,z}(W+|u|)du \Big\} \\
&\leqslant 4(1+z)e^{-z^2}\,\mathbb{E}\Big\{ \int_{-\infty}^{\infty} e^{z|u|}|u\hat{K}(u)|\Psi_{z_0,z}(W)du \Big\} \\
&\leqslant 16(2\pi)^{1/2} r_2(1+z)(1+z^{\tau_2})\phi(z) \\
&\leqslant 40.2 r_2(1+z)^2(1+z^{\tau_2})(1-\Phi(z)) \\
&\leqslant 13 r_2(1+z^{\tau_2+3})(1-\Phi(z)).
\end{aligned} \tag{5.67}$$

Therefore, (5.47) follows from (5.55), (5.62) and (5.67).

(ii) Proof of (5.48). By the Markov inequality,

$$\begin{aligned}
|I_{12}| &\leqslant \mathbb{E}\Big\{ \int_{|u|>1} \mathbf{1}(W+u>z)|\hat{K}(u)|du \Big\} \\
&\quad + \mathbb{E}\Big\{ \int_{|u|>1} \mathbf{1}(W>z)|\hat{K}(u)|du \Big\} \\
&\leqslant \mathbb{E}\Big\{ \int_{-\infty}^{\infty} |u|e^{-z^2}\Psi_{z_0,z}(W+|u|)|\hat{K}(u)|du \Big\} \\
&\quad + \mathbb{E}\Big\{ \int_{-\infty}^{\infty} |u|e^{-z^2}\Psi_{z_0,z}(W)|\hat{K}(u)|du \Big\} \\
&\leqslant 2\,\mathbb{E}\Big\{ \int_{-\infty}^{\infty} e^{-z^2}\Psi_{z_0,z}(W)|u|e^{z|u|}|\hat{K}(u)|du \Big\} \\
&\leqslant 2e^{-z^2}\,\mathbb{E}\{\hat{K}_{2,z}\Psi_{z_0,z}(W)\} \\
&\leqslant 8 r_2(1+z^{\tau_2})e^{-z^2/2},
\end{aligned}$$

where we used (2.9) and (5.22) in the last line. By (5.6), (5.9) and (5.14), we have

$$\begin{aligned}|I_{12}| &\leqslant 8(2\pi)^{1/2} r_2(1+z^{\tau_2})\phi(z)\\ &\leqslant 21 r_2(1+z)(1+z^{\tau_2})(1-\Phi(z))\\ &\leqslant r_2(1+z^{\tau_2+3})(1-\Phi(z)),\end{aligned}$$

which proves (5.48).

(iii) Proof of (5.49). Observe that (see also (2.5) of Chen, Röllin and Xia (2020))

$$\begin{aligned}|h_{z,\varepsilon}(w+u)-h_{z,\varepsilon}(w)| &\leqslant \frac{1}{\varepsilon}\int_{u\wedge 0}^{u\vee 0}\mathbf{1}[z<w+s\leqslant z+\varepsilon]ds\\ &\leqslant \mathbf{1}(z-u\vee 0<w\leqslant z-u\wedge 0+\varepsilon).\end{aligned}\tag{5.68}$$

Recall that $8\leqslant z\leqslant z_0$, and by (5.68) and Fubini's theorem,

$$\begin{aligned}|I_{13}| &\leqslant \int_{|u|\leqslant 1}\mathbb{E}\{|h_{z,\varepsilon}(W+u)-h_{z,\varepsilon}(W)|\}|K(u)|du\\ &\leqslant \frac{1}{\varepsilon}\int_{|u|\leqslant 1}\int_{u\wedge 0}^{u\vee 0}\mathbb{P}[z<W+s\leqslant z+\varepsilon]|K(u)|dsdu\\ &\leqslant I_{131}+I_{132}+I_{133},\end{aligned}\tag{5.69}$$

where

$$\begin{aligned}I_{131} &:= \int_{|u|\leqslant 1}\big(\Phi(z-0\wedge u+\varepsilon)-\Phi(z-0\vee u)\big)|K(u)|du,\\ I_{132} &:= \frac{1}{\varepsilon}\left|\int_{|u|\leqslant 1}\int_{u\wedge 0}^{u\vee 0}(\mathbb{P}[W+s>z]-\mathbb{P}[Z+s>z])|K(u)|dsdu\right|,\\ I_{133} &:= \frac{1}{\varepsilon}\left|\int_{|u|\leqslant 1}\int_{u\wedge 0}^{u\vee 0}(\mathbb{P}[W+s>z+\varepsilon]-\mathbb{P}[Z+s>z+\varepsilon])|K(u)|dsdu\right|.\end{aligned}\tag{5.70}$$

One can easily verify that $(1+z^{\tau_2})(1+z^3)/(1+z^{\tau_2+3})$ is a decreasing function for $z\geqslant 1$ and $\tau_2\geqslant 0$, and

$$\sup_{z\geqslant 2}\frac{(1+z^{\tau_2})(1+z^3)}{1+z^{\tau_2+3}}\leqslant 2.\tag{5.71}$$

Thus, for $z\geqslant 8$,

$$(1+z)\leqslant 1.125z\leqslant 0.02(1+z^3),\tag{5.72}$$

$$(1+z^{\tau_2})(1+z^3)\leqslant 2(1+z^{\tau_2+3}).\tag{5.73}$$

Moreover, as $1\leqslant \Psi_{\beta,t}(w)\leqslant 3$ for all w and t if $\beta=0$, by (2.9) with $\beta=0$, we have

$$\int_{|u|\leqslant 1}e^{z|u|}|uK(u)|du\leqslant \mathbb{E}\{\hat{K}_{2,z}\}\leqslant 3r_2(1+z^{\tau_2}).\tag{5.74}$$

For I_{131}, by (5.6), (5.7) and (5.72), for $z \geqslant 8$ and $|u| \leqslant 1$.

$$
\begin{aligned}
\Phi(z - 0 \wedge u + \varepsilon) - \Phi(z - 0 \vee u) &\leqslant (|u| + \varepsilon)\phi(z - 0 \vee u) \\
&\leqslant (|u| + \varepsilon)e^{z|u|}\phi(z) \\
&\leqslant (|u| + \varepsilon)e^{z|u|}(1 + z)(1 - \Phi(z)) \\
&\leqslant (0.02|u| + 0.02\varepsilon)e^{z|u|}(1 + z^3)(1 - \Phi(z)),
\end{aligned} \tag{5.75}
$$

and then, recalling that $\varepsilon = 40e^{\tau/2}r_2(1 + z^{\tau_2})$, by (2.12), (5.73) and (5.74),

$$
\begin{aligned}
I_{131} &\leqslant 0.02(1 + z^3)(1 - \Phi(z)) \int_{|u|\leqslant 1} (|u| + \varepsilon)e^{z|u|}|K(u)|du \\
&\leqslant (0.06 + 0.8\rho e^{\tau/2})r_2(1 + z^3)(1 + z^{\tau_2})(1 - \Phi(z)) \\
&\leqslant (0.12 + 1.6\rho e^{\tau/2})r_2(1 + z^{\tau_2+3})(1 - \Phi(z)).
\end{aligned} \tag{5.76}
$$

As for I_{132}, by Lemma 5.4 and (5.74) and recalling that $\varepsilon = 40e^{\tau/2}r_2(1 + z^{\tau_2})$, we have for $8 \leqslant z \leqslant z_0$ and $|u| \leqslant 1$,

$$
\begin{aligned}
I_{132} &\leqslant 2e^{\tau/2}\varepsilon^{-1}(C_0 + c_0)\delta(z)(1 - \Phi(z)) \int_{|u|\leqslant 1} e^{z|u|}|uK(u)|du \\
&\leqslant 6e^{\tau/2}r_2\varepsilon^{-1}(C_0 + c_0)\delta(z)(1 - \Phi(z))(1 + z^{\tau_2}) \\
&\leqslant 0.15(C_0 + c_0)\delta(z)(1 - \Phi(z)).
\end{aligned} \tag{5.77}
$$

As for I_{133}, as $8 \leqslant z \leqslant z_0$ and $0 \leqslant \varepsilon \leqslant 1$, by Lemma 5.4 and (5.9) again, we have

$$
\begin{aligned}
I_{133} &\leqslant 0.15e^{z_0\varepsilon}(C_0 + c_0)\delta(z)(1 - \Phi(z)) \\
&\leqslant 0.16(C_0 + c_0)\delta(z)(1 - \Phi(z)).
\end{aligned} \tag{5.78}
$$

By (5.70) and (5.76)–(5.78), we have

$$
|I_{13}| \leqslant 0.31(C_0 + c_0)\delta(z)(1 - \Phi(z)) + (0.12 + 1.6\rho e^{\tau/2})r_2(1 + z^{\tau_1+3})(1 - \Phi(z)).
$$

This proves (5.49).

(iv) Proof of (5.50). Without loss of generality, we assume that $r_4 > 0$, otherwise the proof is even easier. Note that by (5.68),

$$
|I_{14}| \leqslant \mathbb{E}\Big\{\int_{|u|\leqslant 1} \mathbf{1}[z - 0 \vee u < W \leqslant z - 0 \wedge t + \varepsilon]|\hat{K}(u) - K(u)|du\Big\}.
$$

Recall Young's inequality

$$
ab \leqslant \frac{a^2}{2c} + \frac{cb^2}{2} \quad \text{for } a, b \geqslant 0 \text{ and } c > 0.
$$

Applying Young's inequality with $a = \mathbf{1}[z - 0 \vee u < W \leqslant z - 0 \wedge t + \varepsilon], b = |\hat{K}(u) - K(u)|\,\mathbf{1}(W > z - 0 \vee u)$ and

$$c = \frac{e^{z|u|}}{99e^{\tau/2}}\big(e^{\tau/2}(4.1C_0 + 4.1c_0 + 1.6) + r_4^{-1/2}|u|\big),$$

we have

$$\begin{aligned} |I_{14}| &\leqslant \frac{1}{2}\int_{|u|\leqslant 1} c^{-1}\mathbb{P}[z - 0 \vee u < W \leqslant z - 0 \wedge u + \varepsilon]du \\ &\quad + \frac{1}{2}\,\mathbb{E}\Big\{\int_{|u|\leqslant 1} c\big(\hat{K}(u) - K(u)\big)^2\,\mathbf{1}[W > z - 0 \vee u]du\Big\} \\ &:= I_{141} + I_{142}. \end{aligned} \tag{5.79}$$

Using a similar argument to (5.75), and by (5.73), we have

$$\begin{aligned} &\mathbb{P}[z - 0 \vee u < Z \leqslant z - 0 \wedge u + \varepsilon] \\ &\quad\leqslant |u|e^{z|u|}(1 + z^{\tau_4+1})(1 - \Phi(z)) + 0.8e^{\tau/2}r_2(1 + z^{\tau_2})(1 + z^3)(1 - \Phi(z)) \\ &\quad\leqslant r_4^{-1/2}|u|e^{z|u|}\delta(z)(1 - \Phi(z)) + 1.6e^{\tau/2}\delta(z)(1 - \Phi(z)). \end{aligned} \tag{5.80}$$

By (5.80) and Lemma 5.4, and noting that $e^{z_0\varepsilon} \leqslant 1.05$ in (5.9),

$$\begin{aligned} &\mathbb{P}[z - 0 \vee u < W \leqslant z - 0 \wedge u + \varepsilon] \\ &\leqslant \mathbb{P}[z - 0 \vee u < Z \leqslant z - 0 \wedge u + \varepsilon] \\ &\quad + |\mathbb{P}[W > z - 0 \vee u] - \mathbb{P}[Z > z - 0 \vee u]| \\ &\quad + |\mathbb{P}[W > z - 0 \wedge u + \varepsilon] - \mathbb{P}[Z > z - 0 \wedge u + \varepsilon]| \\ &\leqslant \big(e^{\tau/2}(4.1C_0 + 4.1c_0 + 1.6) + r_4^{-1/2}|u|\big)e^{z|u|}\delta(z)(1 - \Phi(z)). \end{aligned} \tag{5.81}$$

Then, we have

$$I_{141} \leqslant 99e^{\tau/2}\delta(z)(1 - \Phi(z)). \tag{5.82}$$

On the other hand, as $z \geqslant 8$, by the Markov inequality and by Lemma 5.3,

$$\begin{aligned} &\mathbb{E}\Big\{\int_{|u|\leqslant 1} |u|e^{z|u|}\big(\hat{K}(u) - K(u)\big)^2\,\mathbf{1}(W > z - 0 \vee u)du\Big\} \\ &\quad\leqslant e^{-z^2}\,\mathbb{E}\Big\{\int_{|u|\leqslant 1} |u|e^{2z|u|}\big(\hat{K}(u) - K(u)\big)^2\Psi_{z_0,z}(W)du\Big\} \\ &\quad\leqslant r_4(1 + z^{\tau_4})e^{-z^2}\,\mathbb{E}\{\Psi_{z_0,z}(W)\} \\ &\quad\leqslant 4(2\pi)^{1/2}r_4(1 + z^{\tau_4})(1 + z)(1 - \Phi(z)) \\ &\quad\leqslant 21r_4(1 + z^{\tau_4+1})(1 - \Phi(z)), \end{aligned} \tag{5.83}$$

where in the last line we used the inequality that

$$(1 + z)(1 + z^{\tau_4}) \leqslant 2(1 + z^{\tau_4+1}) \quad \text{for } z \geqslant 8.$$

Similarly,

$$
\begin{aligned}
&\mathbb{E}\Big\{\int_{|u|\leqslant 1} e^{z|u|}\big(\hat{K}(u)-K(u)\big)^2\,\mathbf{1}(W>z-0\vee u)du\Big\}\\
&\qquad\leqslant 4(2\pi)^{1/2}r_3(1+z)(1+z^{\tau_3})(1-\Phi(z))\\
&\qquad\leqslant 21r_3(1+z^{\tau_3+1})(1-\Phi(z)).
\end{aligned}
\tag{5.84}
$$

Then, by (5.83) and (5.84), we have

$$
\begin{aligned}
I_{142} &\leqslant \frac{e^{\tau/2}(4.1C_0+4.1c_0+1.6)}{198e^{\tau/2}}\times 21r_3(1+z^{\tau_3+1})(1-\Phi(z))\\
&\quad+\frac{21r_4^{1/2}}{198e^{\tau/2}}(1+z^{\tau_4+1})(1-\Phi(z))\\
&\leqslant 0.44(C_0+c_0)\delta(z)(1-\Phi(z))+r_3(1+z^{\tau_3+1})(1-\Phi(z))\\
&\quad+r_4^{1/2}(1+z^{\tau_4+1})(1-\Phi(z))\\
&\leqslant (0.44C_0+0.44c_0+1)\delta(z)(1-\Phi(z)).
\end{aligned}
\tag{5.85}
$$

Combining (5.82) and (5.85) yields (5.50).

(v) Proof of (5.51). Note that (see, e.g., p. 2010 of Chen and Shao (2004))

$$
|f'(w)|\leqslant\begin{cases}1-\Phi(z) & \text{if } w<0,\\ \left(\dfrac{w\Phi(w)}{\phi(w)}+1\right)\big(1-\Phi(z)\big) & \text{if } 0\leqslant w\leqslant z,\\ 1 & \text{otherwise.}\end{cases}
\tag{5.86}
$$

Observe that

$$
\begin{aligned}
|I_2| &\leqslant \mathbb{E}\{|f'(W)(1-\mathbb{E}\{\hat{K}_1|W\})|\,\mathbf{1}(W\leqslant 0)\}\\
&\quad+\mathbb{E}\{|f'(W)(1-\mathbb{E}\{\hat{K}_1|W\})|\,\mathbf{1}(0<W\leqslant z)\}\\
&\quad+\mathbb{E}\{|f'(W)(1-\mathbb{E}\{\hat{K}_1|W\})|\,\mathbf{1}(W>z)\}\\
&:= I_{21}+I_{22}+I_{23}.
\end{aligned}
$$

For I_{21}, since $-1\leqslant w\Phi(w)/\phi(w)\leqslant 0$ for $w\leqslant 0$, and by (2.8) and (5.86), we have for $z\geqslant 8$

$$
\begin{aligned}
I_{21} &\leqslant (1-\Phi(z))\,\mathbb{E}\{|\mathbb{E}\{\hat{K}_1|W\}-1|\}\\
&\leqslant 2r_1(1-\Phi(z))\\
&\leqslant 0.1r_1(1+z^{\tau_1+2})(1-\Phi(z)).
\end{aligned}
\tag{5.87}
$$

For I_{22}, by (2.8) and (5.86), we have

$$
\begin{aligned}
I_{22} &\leqslant (2\pi)^{1/2}(1-\Phi(z))\,\mathbb{E}\{(We^{W^2/2}+1)|\mathbb{E}\{\hat{K}_1|W\}-1|\,\mathbf{1}(0<W\leqslant z)\}\\
&\leqslant 2.6(1-\Phi(z))I_{24}+5.2r_1(1-\Phi(z)),
\end{aligned}
\tag{5.88}
$$

where

$$
\begin{aligned}
I_{24} &:= \mathbb{E}\{We^{W^2/2}|\mathbb{E}\{\hat{K}_1|W\}-1|\,\mathbf{1}(0<W\leqslant z)\}\\
&= \sum_{j=1}^{[z]}\mathbb{E}\{We^{W^2/2}|\mathbb{E}\{\hat{K}_1|W\}-1|\,\mathbf{1}(j-1<W\leqslant j)\}\\
&\quad +\mathbb{E}\{We^{W^2/2}|\mathbb{E}\{\hat{K}_1|W\}-1|\,\mathbf{1}([z]<W\leqslant z)\}\\
&:= I_{241}+I_{242}.
\end{aligned}
$$

For I_{241}, noting that

$$
w^2/2-jw\leqslant (j-1)^2/2-j(j-1)=-j^2/2+1/2 \quad \text{for } j-1<w\leqslant j,
$$

«««< HEAD we have for $z\geqslant 8$, by (2.8) and (5.22), ======= we have for $z\geqslant 8$, by (2.8) and (5.22), »»»> 622099701ad839dfddec132728eab45bce27be61

$$
\begin{aligned}
I_{241} &\leqslant \sum_{j=1}^{[z]} j\,\mathbb{E}\{e^{W^2/2-jW+jW}|\mathbb{E}\{\hat{K}_1|W\}-1|\,\mathbf{1}(j-1\leqslant W\leqslant j)\}\\
&\leqslant e^{1/2}\sum_{j=1}^{[z]} je^{-j^2/2}\,\mathbb{E}\{|\mathbb{E}\{\hat{K}_1|W\}-1|\Psi_{z_0,j}(W)\}\\
&\leqslant 4e^{1/2}r_1\sum_{j=1}^{\lfloor z\rfloor} j(1+j^{\tau_1})\\
&\leqslant 13.2r_1(1+z^{\tau_1+2}).
\end{aligned}
\tag{5.89}
$$

where in the last inequality we use the similar inequality as that in (5.61). Similarly, for $z\geqslant 8$,

$$
I_{242}\leqslant 4e^{1/2}r_1z(1+z^{\tau_1})\leqslant 1.7r_1(1+z^{\tau_1+2}). \tag{5.90}
$$

By (5.89) and (5.90), we have

$$
I_{24}\leqslant 15r_1(1+z^{\tau_1+2}) \quad \text{for } z\geqslant 8. \tag{5.91}
$$

By (5.88) and (5.91), we have for $z\geqslant 8$,

$$
\begin{aligned}
I_{22} &\leqslant 39r_1(1+z^{\tau_1+2})(1-\Phi(z))+5.2r_1(1-\Phi(z))\\
&\leqslant 40r_1(1+z^{\tau_1+2})(1-\Phi(z)).
\end{aligned}
\tag{5.92}
$$

As for I_{23}, by Lemma 5.3 and (2.8), (5.6) and (5.86) and recalling $z \geqslant 8$, we have

$$\begin{aligned} I_{23} &\leqslant \mathbb{E}\{|\mathbb{E}\{\hat{K}_1|W\} - 1|\,\mathbf{1}(W > z)\} \\ &\leqslant e^{-z^2}\,\mathbb{E}\{|\mathbb{E}\{\hat{K}_1|W\} - 1|\Psi_{z_0,z}(W)\} \\ &\leqslant 4r_1(1+z)(1+z^{\tau_1})e^{-z^2/2} \\ &\leqslant 10.1r_1(1+z)(1+z^{\tau_1})\phi(z) \\ &\leqslant 10.1r_1(1+z)^2(1+z^{\tau_1})(1-\Phi(z)) \\ &\leqslant 25.2r_1(1+z^{\tau_1+2})(1-\Phi(z)). \end{aligned} \tag{5.93}$$

By (5.87), (5.92) and (5.93), we have

$$|I_2| \leqslant 66r_1(1+z^{\tau_1+2})(1-\Phi(z)). \tag{5.94}$$

(vi) Proof of (5.52). It is known that $0 \leqslant f(w) \leqslant 1$ (see p. 2010 of Chen and Shao (2004)). Note that by (5.43) and (5.54),

$$|I_3| \leqslant I_{31} + I_{32} + I_{33},$$

where

$$\begin{aligned} I_{31} &= (1-\Phi(z))\,\mathbb{E}\{|R|\,\mathbf{1}(W \leqslant 0)\}, \\ I_{32} &= \sqrt{2\pi}(1-\Phi(z))\,\mathbb{E}\{e^{W^2/2}|R|\,\mathbf{1}(0 < W \leqslant z)\}, \\ I_{33} &= \mathbb{E}\{|R|\,\mathbf{1}(W > z)\}. \end{aligned}$$

For I_{31}, by (2.7) with $t = 0$, we have

$$I_{31} \leqslant r_0(1+z^{\tau_0+1})(1-\Phi(z)) \quad \text{for } z \geqslant 8.$$

For I_{32}, similar to (5.92), we have

$$I_{32} \leqslant 60r_0(1+z^{\tau_0+1})(1-\Phi(z)).$$

For I_{33}, similar to (5.93), we have

$$I_{33} \leqslant 21r_0(1+z^{\tau_0+1})(1-\Phi(z)).$$

Combining the foregoing inequalities, we have

$$|I_3| \leqslant 82r_0(1+z^{\tau_0+1})(1-\Phi(z)).$$

This proves (5.52). $\square$

Proof of Lemma 5.2. In this proof, we develop a Berry–Esseen bound using the idea in Chen, Röllin and Xia (2020). Let

$$\gamma := \sup_{z\in\mathbb{R}}\big|\mathbb{P}(W \leqslant z) - \Phi(z)\big|,$$

and let $\varepsilon = \gamma/2$. Consider the Stein equation (5.41), and we denote by $f_{z,\varepsilon}$ the solution to (5.41), which is given in (5.43). Let $g_{z,\varepsilon}(w) = wf_{z,\varepsilon}(w)$. By Chen and Shao (2004), we have

$$0 \leqslant f_{z,\varepsilon} \leqslant 1, \quad \|f'_{z,\varepsilon}\| \leqslant 1. \tag{5.95}$$

Note that

$$\gamma \leqslant \sup_{z\in\mathbb{R}}|\mathbb{E}\{h_{z,\varepsilon}(W)\} - Nh_{z,\varepsilon}| + 0.4\varepsilon. \tag{5.96}$$

Now, we bound the first term of (5.96). By (1.2) and (5.41), we have

$$\begin{aligned}
&\mathbb{E}\{h_{z,\varepsilon}(W)\} - Nh_{z,\varepsilon}\\
&\quad= \mathbb{E}\{f'_{z,\varepsilon}(W)\} - \mathbb{E}\{Wf_{z,\varepsilon}(W)\}\\
&\quad= \mathbb{E}\{f'_{z,\varepsilon}(W)\} - \mathbb{E}\{Rf_{z,\varepsilon}(W)\} - \mathbb{E}\bigg\{\int_{-\infty}^{\infty} f'_{z,\varepsilon}(W+u)\hat{K}(u)du\bigg\}\\
&\quad:= J_1 + J_2 + J_3 + J_4 + J_5,
\end{aligned} \tag{5.97}$$

where

$$\begin{aligned}
J_1 &= \mathbb{E}\{f'_{z,\varepsilon}(W)(1-\hat{K}_1)\},\\
J_2 &= -\mathbb{E}\{Rf_{z,\varepsilon}(W)\},\\
J_3 &= -\mathbb{E}\bigg\{\int_{|u|>1}\big(f'_{z,\varepsilon}(W+u) - f'_{z,\varepsilon}(W)\big)\hat{K}(u)du\bigg\},\\
J_4 &= -\mathbb{E}\bigg\{\int_{|u|\leqslant 1}\big(f'_{z,\varepsilon}(W+u) - f'_{z,\varepsilon}(W)\big)K(u)du\bigg\},\\
J_5 &= -\mathbb{E}\bigg\{\int_{|u|\leqslant 1}\big(f'_{z,\varepsilon}(W+u) - f'_{z,\varepsilon}(W)\big)(\hat{K}(u) - K(u))du\bigg\}.
\end{aligned}$$

By (2.7)–(2.11) with $t = 0$ and noting $1 \leqslant \Psi_{\beta,0}(w) \leqslant 2$, we have

$$\begin{aligned}
\mathbb{E}|R| \leqslant 2r_0, \qquad &\mathbb{E}|\mathbb{E}\{\hat{K}_1|W\} - 1| \leqslant 2r_1,\\
\mathbb{E}\hat{K}_{2,0} \leqslant 2r_2, \qquad &\mathbb{E}\hat{K}_{3,0} \leqslant 2r_3, \qquad \mathbb{E}\hat{K}_{4,0} \leqslant 2r_4.
\end{aligned} \tag{5.98}$$

Now, by (5.95) and (5.98), we have

$$|J_1| \leqslant 2r_1, \quad |J_2| \leqslant 2r_0, \quad |J_3| \leqslant 2\,\mathbb{E}\{\hat{K}_{2,0}\} \leqslant 4r_2. \tag{5.99}$$

By Eqs. (2.12), (2.15) and (2.16) of Chen, Röllin and Xia (2020), we have with $a = 0.18$,

$$\begin{aligned}
|J_4| &\leqslant 4r_2 + \frac{4\gamma + 0.8\varepsilon}{\varepsilon}r_2,\\
|J_5| &\leqslant a\gamma + 0.2a\varepsilon + \frac{2r_3}{a} + (2a + 0.4/a)r_4^{1/2} + 5r_4^{1/2}.
\end{aligned} \tag{5.100}$$

Combining (5.96), (5.97), (5.99) and (5.100) gives

$$\gamma \leqslant 0.4\gamma + 2r_0 + 2r_1 + 16.8r_2 + 11.2r_3 + 7.6r_4^{1/2}.$$

Solving the recursive inequality yields the desired result. □

6 PROOF OF THEOREM 3.1

We apply Theorem 2.1 to prove Theorem 3.1. To this end, we first prove a preliminary lemma. Denote by $C, C_1, C_2, \ldots$ absolute constants, which may take different values in different places.

6.1 Preliminary lemmas

Let $\hat{K}_1, \hat{K}_{2,t}, \hat{K}_{3,t}, \hat{K}_{4,t}$ and M_t be as in (2.1)–(2.5) with $\hat{K}(u)$ in (3.1) and $h(t) = \mathbb{E}\{\Psi_{\beta,t}(W)\}$. The following lemma provides the upper bounds of the terms in Condition (A1), whose proof is put in Appendix A.

Lemma 6.1. *Let $m_0 = (a_n^{1/3}/4) \wedge (a_n/16)$. For $0 \leqslant t, \beta \leqslant m_0$, we have*

$$\mathbb{E}\{|\mathbb{E}\{\hat{K}_1|W\} - 1|\Psi_{\beta,t}(W)\} \leqslant C_1(b\kappa n a_n^{-3} + b^{1/2}\kappa^{1/2}n^{1/2}a_n^{-2})(1+t)h(t), \tag{6.1}$$

$$\mathbb{E}\{\hat{K}_{2,t}\Psi_{\beta,t}(W)\} \leqslant C_2 bna_n^{-3}h(t), \tag{6.2}$$

$$\mathbb{E}\{\hat{K}_{3,t}\Psi_{\beta,t}(W)\} \leqslant C_3(\kappa bna_n^{-3} + \kappa^2b^2n^2a_n^{-5})(1+t^2)h(t), \tag{6.3}$$

$$\mathbb{E}\{\hat{K}_{4,t}\Psi_{\beta,t}(W)\} \leqslant C_4(\kappa bna_n^{-4} + \kappa^2b^2n^2a_n^{-6})(1+t^2)h(t), \tag{6.4}$$

and

$$\sup_{0\leqslant t\leqslant m_0} M_t \leqslant C_5 bna_n^{-2}. \tag{6.5}$$

6.2 Proof of Theorem 3.1

Proof of Theorem 3.1. We apply Theorem 2.1 to prove Theorem 3.1. Recalling that $\theta_n = b^{1/2}n^{1/2}a_n^{-1}$, by Lemma 6.1, we have condition (A1) is satisfied with $r_0 = \tau_0 = 0$ and

$$r_1 = C_1\kappa\theta_n(\theta_n+1)a_n^{-1}, \qquad \tau_1 = 1, \tag{6.6}$$

$$r_2 = C_2\theta_n^2 a_n^{-1}, \qquad \tau_2 = 0, \tag{6.7}$$

$$r_3 = C_3(\kappa\theta_n^2+1)^2a_n^{-1}, \qquad \tau_3 = 2,$$

$$r_4 = C_4(\kappa\theta_n^2+1)^2a_n^{-2}, \qquad \tau_4 = 2,$$

$$\rho = C_5\theta_n^2, \tag{6.8}$$

where C_1, C_2, C_3, C_4 and C_5 are as in Lemma 6.1. Recalling the definition of $\delta(t)$ in (2.15), and that $m_0 = (a_n^{1/3}/4) \wedge (a_n/16)$, we have

$$\begin{aligned} \delta(m_0) &\geqslant (r_0 + r_1 + r_2 + r_3 + r_4^{1/2})(1 + m_0^3) \\ &\geqslant C\theta_n\kappa(1 + \theta_n + \kappa\theta_n^3)(a_n^{-1} + a_n^2 \wedge 1) \geqslant C\theta_n\kappa. \end{aligned} \tag{6.9}$$

Combining (6.8) and (6.9) and noting that $\kappa \geqslant 1$, we get the right hand side of (2.14) is less than

$$\begin{aligned} C\Big(\frac{1}{\kappa\theta_n} + \theta_n^2 + 1\Big)\delta(z) &\leqslant 2C\Big(\frac{1}{\theta_n} + \theta_n^2\Big)(r_1 + r_2 + r_3 + r_4^{1/2})(1 + z^3) \\ &\leqslant C'\delta_n(1 + z^3), \end{aligned} \tag{6.10}$$

where $\delta_n = \kappa^2 a_n^{-1}(1+\theta_n^6)$, C' is an absolute constant and we use the fact that $x^2+1/x > 1$ for $x > 0$. On the other hand, by (6.6) and (6.7), we have

$$r_1^{1/3} + r_2^{1/3} \leqslant C\kappa^{1/3} a_n^{-1/3}(1 + \theta_n)^{2/3}. \tag{6.11}$$

Applying Theorem 2.1 and by (6.10) and (6.11), we obtain the desired result. □

7 PROOF OF THEOREM 4.1

In this section, we use the exchangeable pair to construct Stein's identity (1.2). For any $k \geqslant 1$ and k-fold index $\boldsymbol{i} \in \mathbb{N}^k$, we denote by i_j its j-th element. Let $[n]_k := \{\boldsymbol{i} = (i_1, \dots, i_k) \in \mathbb{N}^k : 1 \leqslant i_1 \neq \dots \neq i_k \leqslant n\}$ be a class of k-fold indices. Let $\boldsymbol{I} := (I_1, I_2)$ be chosen uniformly from $[n]_2$ independent of π and $\boldsymbol{X}$, and let $W' = W - X_{I_1,\pi(I_1)} - X_{I_2,\pi(I_2)} + X_{I_1,\pi(I_2)} + X_{I_2,\pi(I_1)}$. Then, it follows that (W, W') is an exchangeable pair. Moreover, we have

$$\begin{aligned} \mathbb{E}\{W - W'|\boldsymbol{X}, \pi\} &= \frac{1}{n(n-1)} \sum_{\boldsymbol{i} \in [n]_2} \mathbb{E}\{X_{i_1,\pi(i_1)} + X_{i_2,\pi(i_2)} - X_{i_1,\pi(i_2)} - X_{i_2,\pi(i_1)}|\boldsymbol{X}, \pi\} \\ &= \frac{2}{n-1}(W - R), \end{aligned}$$

where

$$R = -\frac{1}{n}\sum_{i=1}^n \sum_{j=1}^n X_{i,j}. \tag{7.1}$$

By exchangeability, with $\lambda = 2/(n-1)$ and $\Delta = W - W'$, we have

$$\begin{aligned} 0 &= \mathbb{E}\{(W - W')(f(W) + f(W'))\} = 2\,\mathbb{E}\{\Delta f(W)\} - \mathbb{E}\{\Delta(f(W) - f(W - \Delta))\} \\ &= 2\lambda\,\mathbb{E}\{(W + R)f(W)\} - \mathbb{E}\{\Delta(f(W) - f(W - \Delta))\}. \end{aligned} \tag{7.2}$$

Rearranging (7.2) yields

$$\mathbb{E}\{Wf(W)\} = \mathbb{E}\int_{-\infty}^{\infty} f'(W + u)\hat{K}(u)du + \mathbb{E}\{Rf(W)\},$$

where

$$\begin{aligned}\hat{K}(u) &= \frac{1}{2\lambda}\,\mathbb{E}\{\Delta\big(\mathbf{1}(-\Delta \leqslant u \leqslant 0) - \mathbf{1}(0 < u \leqslant -\Delta)\big)|\boldsymbol{X}, \pi\} \\ &= \frac{1}{4n}\sum_{\boldsymbol{i}\in[n]_2} D_{\boldsymbol{i},\pi(\boldsymbol{i})}\big(\mathbf{1}(-D_{\boldsymbol{i},\pi(\boldsymbol{i})} \leqslant u \leqslant 0) - \mathbf{1}(0 < u \leqslant -D_{\boldsymbol{i},\pi(\boldsymbol{i})})\big),\end{aligned} \tag{7.3}$$

and $D_{\boldsymbol{i},\boldsymbol{j}} = X_{i_1,j_1} + X_{i_2,j_2} - X_{i_1,j_2} - X_{i_2,j_1}$ for any $\boldsymbol{i} = (i_1, i_2)$ and $\boldsymbol{j} = (j_1, j_2)$. Therefore, the condition (1.2) is satisfied.

In what follows, we denote by $C, C_1, C_2, \dots$ absolute constants, which may take different values in different places.

7.1 Preliminary lemma

The following lemma will be useful in the proof of Theorem 4.1, and the proof of this lemma is put in Appendix B. Let $\hat{K}_1, \hat{K}_{2,t}, \hat{K}_{3,t}, \hat{K}_{4,t}$ and M_t be as in (2.1)–(2.5) with $\hat{K}(u)$ defined as in (7.3).

Lemma 7.1. *For $n \geqslant 4$ and $0 \leqslant t, \beta \leqslant \alpha_n^{1/3}/64$, we have*

$$\mathbb{E}\{|R|\Psi_{\beta,t}(W)\} \leqslant C_0 b\alpha_n^{-1}h(t), \tag{7.4}$$

$$\mathbb{E}\{|\hat{K}_1 - 1|\Psi_{\beta,t}(W)\} \leqslant C_1 b(n\alpha_n^{-3} + n^{1/2}\alpha_n^{-2} + n^{-1/2})h(t), \tag{7.5}$$

$$\mathbb{E}\{\hat{K}_{2,t}\Psi_{\beta,t}(W)\} \leqslant C_2 bn\alpha_n^{-3}\,\mathbb{E}\{\Psi_{\beta,t}(W)\}, \tag{7.6}$$

$$\mathbb{E}\{\hat{K}_{3,t}\Psi_{\beta,t}(W)\} \leqslant C_3 b^2(n\alpha_n^{-3} + n^2\alpha_n^{-5})(1+t^2)\,\mathbb{E}\{\Psi_{\beta,t}(W)\}, \tag{7.7}$$

$$\mathbb{E}\{\hat{K}_{4,t}\Psi_{\beta,t}(W)\} \leqslant C_4 b^2(n\alpha_n^{-4} + n^2\alpha_n^{-6})(1+t^2)\,\mathbb{E}\{\Psi_{\beta,t}(W)\} \tag{7.8}$$

and

$$\sup_{0\leqslant t\leqslant m_0} M_t \leqslant C_5 bn\alpha_n^{-2}. \tag{7.9}$$

7.2 Proof of Theorem 4.1

Proof of Theorem 4.1. We apply Theorem 2.1 to prove Theorem 4.1. Recalling that $\theta_n = n^{1/2}\alpha^{-1}$, by Lemma 7.1, we have condition (A1) is satisfied with $m_0 = \alpha_n^{1/3}/64$,

$$\begin{aligned}&r_0 = C_0 b\alpha_n^{-1}, \quad \tau_0 = 0, \quad r_1 = C_1 b((\theta_n^2 + \theta_n)\alpha_n^{-1} + n^{-1/2}), \quad \tau_1 = 0, \quad r_2 = C_2 b\theta_n^2\alpha_n^{-1}, \quad \tau_2 = 0, \\ &r_3 = 2C_3 b^2(\theta_n^2+1)^2\alpha_n^{-1}, \quad \tau_3 = 2, \quad r_4 = 2C_4 b^2(\theta_n^2+1)^2\alpha_n^{-2}, \quad \tau_4 = 2, \quad \rho = C_5 b\theta_n^2,\end{aligned} \tag{7.10}$$

where C_1, C_2, C_3, C_4 and C_5 are as in Lemma 7.1. Recalling the definition of $\delta(t)$ in (2.15), and noting that $m_0 = \alpha_n^{1/3}/64$, we have

$$\delta(m_0) \geqslant r_2(1+m_0^3) \geqslant r_2 m_0^3 \geqslant Cb\theta_n^2. \tag{7.11}$$

Combining (7.10) and (7.11), we have that the right hand side of (2.14) is less than

$$C\Big(\frac{1}{\theta_n^2}+\theta_n^2+1\Big)\delta(z) \leqslant 2C\Big(\frac{1}{\theta_n^2}+\theta_n^2\Big)(r_0+r_1+r_2+r_3+r_4^{1/2})(1+z^3) \leqslant C\delta_n(1+z^3), \tag{7.12}$$

where $\delta_n = (\alpha_n^{-1}+n^{-1/2})(\theta_n^{-2}+\theta_n^6)$, and we used the fact that $x^2+1/x^2>1$ for $x>0$ in the first inequaltiy. On the other hand, by (7.10), we have

$$\begin{aligned} r_0^{1/(\tau_0+1)} &\leqslant Cb\alpha_n^{-1}, \\ r_1^{1/(\tau_1+2)} &\leqslant Cb^{1/2}(\theta_n+1)(1+\theta_n^{-1/2})\alpha_n^{-1/2}, \\ r_2^{1/(\tau_2+3)} &\leqslant Cb^{1/3}\theta_n^{2/3}\alpha_n^{-1/3}. \end{aligned} \tag{7.13}$$

By (7.13),

$$r_0^{1/(\tau_0+1)}+r_1^{1/(\tau_1+2)}+r_2^{1/(\tau_2+3)} \leqslant Cb(1+\theta_n)(1+\theta_n^{-1/2})\alpha_n^{-1/3}. \tag{7.14}$$

Applying Theorem 2.1, and by (7.12) and (7.14), we obtain the desired result. □

A PROOF OF LEMMA 6.1

Throughout this section, we write $h(t)=\mathbb{E}\{\Psi_{\beta,t}(W)\}$ and $T_i=\sum_{j\in B_i}|X_i|$. In what follows, we give two general lemmas, which will be used in the proof of Lemma 6.1. The following lemmas give us inequalities like $\mathbb{E}\{\xi\Psi_{\beta,t}(W)\} \leqslant h(t)\,\mathbb{E}\{f(\xi,X_{\mathcal{J}})\}$ for some special random variable ξ and function f of ξ and $X_{\mathcal{J}}$, where $X_{\mathcal{J}}=\{X_j, j\in\mathcal{J}\}$.

Lemma A.1. *Let $W_i := W-\sum_{k\in B_i}X_j$ and let $\zeta_i=\zeta(X_{A_i})\geqslant 0$ be a function of X_{A_i}. Then, for $0\leqslant t\leqslant m_0$,*

$$\mathbb{E}\big\{\zeta_i e^{3tT_i}\Psi_{\beta,t}(W_i)\big\} \leqslant 81b^{1/4}h(t)\,\mathbb{E}\big\{\zeta_i e^{3a_nT_i/8}\big\}, \tag{A.1}$$

$$\mathbb{E}\big\{\zeta_i T_i^2 e^{3tT_i}\Psi_{\beta,t}(W_i)\big\} \leqslant C\kappa^2 b^{1/4}h(t)\tau^{-1}\,\mathbb{E}\big\{\zeta_i^2 e^{3a_nT_i/4}\big\} + C\kappa^2\tau a_n^{-4}bh(t), \tag{A.2}$$

where $C>0$ is an absolute constant and $\tau>0$ is any positive number.

Lemma A.2. *For each $i\in\mathcal{J}$, let $\xi_i=\xi(X_{A_i})$ be a function of X_{A_i} satisfying that $\mathbb{E}\xi_i=0$. Let $S=\sum_{i\in\mathcal{J}}\xi_i$. For $0\leqslant t,\beta\leqslant m_0$ and any positive number τ, we have*

$$\begin{aligned} \mathbb{E}\big\{S^2\Psi_{\beta,t}(W)\big\} &\leqslant 81b^{1/4}\kappa h(t)\sum_{i\in\mathcal{J}}\mathbb{E}\{\xi_i^2 e^{3a_nT_i/8}\} \\ &\quad + Cb\kappa^2t^2h(t)\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}\setminus N_i}\mathbb{E}|\xi_j|\big(\tau^{-1}\,\mathbb{E}\{\xi_i^2e^{3a_nT_i/4}\}+\tau a_n^{-4}\big). \end{aligned} \tag{A.3}$$

Based on Lemmas A.1 and A.2 we are ready to give the proof of Lemma 6.1 now.

Proof of Lemma 6.1. Recall that $\hat{K}(u), K(u), \hat{K}_i(u)$ and $K_i(u)$ are defined as in (3.1) and (3.2), and recall that

$$\hat{K}_1 = \int_{-\infty}^{\infty} \hat{K}(t)dt = \sum_{i\in\mathcal{J}} X_iY_i, \quad \hat{K}_{2,t} = \sum_{i\in\mathcal{J}} \int_{-\infty}^{\infty} |u|e^{t|u|}\hat{K}_i(u)du. \tag{A.4}$$

Since $\mathbb{E}W^2 = 1$, it follows that $\mathbb{E}\hat{K}_1 = 1$, which implies that $K_1 = \sum_{i\in\mathcal{J}}\int_{t\in\mathbb{R}} K_i(t)dt = 1$. For (6.1), by Jensen's and Hölder's inequalities, we have

$$\begin{aligned}\mathbb{E}\{|\mathbb{E}[\hat{K}_1|W] - K_1|\Psi_{\beta,t}(W)\} &\leqslant \mathbb{E}\{|\hat{K}_1 - K_1|\Psi_{\beta,t}(W)\}\\ &\leqslant h(t)^{1/2}(\mathbb{E}\{|\hat{K}_1 - K_1|^2\Psi_{\beta,t}(W)\})^{1/2}.\end{aligned} \tag{A.5}$$

Applying Lemma A.2 with $\xi_i = X_iY_i - \mathbb{E}\{X_iY_i\}$ and $\tau = b^{1/2}$ yields

$$\mathbb{E}\{|\hat{K}_1 - K_1|^2\Psi_{\beta,t}(W)\} \leqslant G_1 + G_2, \tag{A.6}$$

where

$$\begin{aligned}G_1 &= 81b^{1/4}\kappa h(t)\sum_{i\in\mathcal{J}}\mathbb{E}\{\xi_i^2 e^{a_nT_i/8}\},\\ G_2 &= Cb\kappa^2t^2h(t)\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}\setminus N_i}\mathbb{E}|\xi_j|\big(b^{-1/2}\,\mathbb{E}\{\xi_i^2e^{3a_nT_i/4}\} + b^{1/2}a_n^{-4}\big).\end{aligned}$$

Recalling that $T_i = \sum_{j\in B_i}|X_j|$, we have $|\xi_i| \leqslant T_i^2 + \mathbb{E}T_i^2$, and thus, for $0 < s \leqslant 3/4$,

$$\begin{aligned}\mathbb{E}(\{X_iY_i - \mathbb{E}X_iY_i\}^2e^{sa_nT_i}) &\leqslant a_n^{-4}\,\mathbb{E}(\{a_n^4T_i^4 + \mathbb{E}a_n^4T_i^4\}e^{sa_nT_i})\\ &\leqslant Ca_n^{-4}(\mathbb{E}e^{a_nT_i})^{s+1/4}\\ &\leqslant Cb^{s+1/4}a_n^{-4},\end{aligned} \tag{A.7}$$

where we used the inequality that $y^4 \leqslant Ce^{y/4}$ for $y \geqslant 0$ and some $C > 0$. Moreover,

$$\mathbb{E}|\xi_j| \leqslant 2\,\mathbb{E}T_j^2 \leqslant Ca_n^{-2}\,\mathbb{E}e^{a_nT_j/2} \leqslant Cb^{1/2}a_n^{-2}. \tag{A.8}$$

Substituting (A.7) and (A.8) to (A.6) gives

$$\mathbb{E}\{|\hat{K}_1 - K_1|^2\Psi_{\beta,t}(W)\} \leqslant C(b\kappa na_n^{-4} + b^2\kappa^2n^2a_n^{-6})h(t)(1+t^2),$$

which proves (6.1) together with (A.5).

Next we prove (6.2). Recalling that $\hat{K}_i(u)$ is defined as in (3.1), by (A.20) and applying Lemma A.1 with $\xi_i = |\hat{K}_i(u)|$, we have

$$\begin{aligned}|\,\mathbb{E}\{\hat{K}_i(u)\Psi_{\beta,t}(W)\}| &\leqslant \mathbb{E}\{|\hat{K}_i(u)|e^{tT_i}\Psi_{\beta,t}(W_i)\}\\ &\leqslant 81b^{1/4}h(t)\,\mathbb{E}\{|\hat{K}_i(u)|e^{3a_nT_i/8}\}.\end{aligned}$$

Thus, by (A.4) and recalling that $t \leqslant a_n/16$ and $|X_i| \leqslant T_i, |Y_i| \leqslant T_i$, we have

$$\mathbb{E}\{|\hat{K}_{2,t}|\Psi_{\beta,t}(W)\} \leqslant 81b^{1/4}h(t)\sum_{i\in\mathcal{J}}\mathbb{E}\{|X_i^2Y_i|e^{a_nT_i/2}\}$$

$$
\begin{aligned}
&\leqslant 81 b^{1/4} h(t) \sum_{i \in \mathcal{J}} \mathbb{E}\{T_i^3 e^{a_n T_i/2}\} \\
&\leqslant C b n a_n^{-3} h(t).
\end{aligned}
$$

Next we prove (6.3) and (6.4) together. By definition,

$$
\mathbb{E}\{\hat{K}_{3,t} \Psi_{\beta,t}(W)\} = \int_{|u| \leqslant 1} e^{2t|u|}\, \mathbb{E}\Big\{\big(\hat{K}(u) - K(u)\big)^2 \Psi_{\beta,t}(W)\Big\} du \tag{A.9}
$$

and

$$
\mathbb{E}\{\hat{K}_{4,t} \Psi_{\beta,t}(W)\} = \int_{|u| \leqslant 1} |u| e^{2t|u|}\, \mathbb{E}\Big\{\big(\hat{K}(u) - K(u)\big)^2 \Psi_{\beta,t}(W)\Big\} du. \tag{A.10}
$$

For fixed u, applying Lemma A.2 with $\xi_i = \hat{K}_i(u) - K_i(u)$ and $\tau = b^{1/2} a_n$, we have

$$
\mathbb{E}\big\{\big(\hat{K}(u) - K(u)\big)^2 \Psi_{\beta,t}(W)\big\} = H_1(u) + H_2(u), \tag{A.11}
$$

where

$$
\begin{aligned}
&H_1(u) = 81 b^{1/4} \kappa h(t) \sum_{i \in \mathcal{J}} \mathbb{E}\{(\hat{K}_i(u) - K_i(u))^2 e^{3a_n T_i/8}\}, \\
&H_2(u) = C b \kappa^2 t^2 h(t) \sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{J} \setminus N_i} \mathbb{E}|\hat{K}_j(u) - K_j(u)| \big(b^{-1/2} a_n^{-1}\, \mathbb{E}\{(\hat{K}_i(u) - K_i(u))^2 e^{3a_n T_i/4}\} + b^{1/2} a_n^{-3}\big).
\end{aligned}
$$

For $H_1(u)$, recalling that $|X_i| \leqslant T_i, |Y_i| \leqslant T_i$ and $t \leqslant a_n/16$,

$$
\begin{aligned}
\int_{-\infty}^{\infty} e^{2t|u|}\, \mathbb{E}\{\hat{K}_i(u)^2 e^{3a_n T_i/8}\} du &\leqslant \mathbb{E}\{|X_i^2 Y_i| e^{a_n T_i/2}\} \\
&\leqslant C a_n^{-3}\, \mathbb{E}\{(a_n T_i)^3 e^{a_n T_i/2}\} \\
&\leqslant C b^{3/4} a_n^{-3},
\end{aligned} \tag{A.12}
$$

and similarly,

$$
\int_{-\infty}^{\infty} e^{2t|u|}\, \mathbb{E}\{K_i(u)^2 e^{3a_n T_i/8}\} du \leqslant C b^{3/4} a_n^{-3}. \tag{A.13}
$$

For $H_2(u)$, note that $|\hat{K}_i(u)| \leqslant |X_i|^2$. We have

$$
\mathbb{E}\{|\hat{K}_i(u)|^2 e^{3a_n/4}\} \leqslant \mathbb{E}\{|X_i|^2 e^{3a_n/4}\} \leqslant \mathbb{E}\{T_i^2 e^{3a_n/4}\} \leqslant C b a_n^{-2}
$$

and

$$
\mathbb{E}\{|K_i(u)|^2 e^{3a_n/4}\} \leqslant C b a_n^{-2}.
$$

Then,

$$
H_2(u) \leqslant C b^{3/2} \kappa^2 t^2 h(t) \sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{J} \setminus N_i} \mathbb{E}|\hat{K}_j(u) - K_j(u)|. \tag{A.14}
$$

Similar to (A.12) and (A.13),

$$\int_{-\infty}^{\infty} e^{2t|u|}\,\mathbb{E}|\hat{K}_j(u)-K_i(u)|du \leqslant \mathbb{E}\{(|X_iY_i|+\mathbb{E}|X_iY_i|)e^{2tT_i}\} \leqslant Cb^{1/2}a_n^{-2}. \tag{A.15}$$

Substituting (A.11)–(A.15) to (A.9) gives (6.3). The inequality (6.4) can be shown similarly.

Now, it remains to prove (6.5). By definition,

$$\begin{aligned}\sup_{0\leqslant t\leqslant m_0} M_t &= \int_{|u|\leqslant 1} e^{m_0|u|}|K(u)|du \\ &\leqslant \sum_{i\in\mathcal{J}}\mathbb{E}\Big\{\int_{|u|\leqslant 1}|\hat{K}_i(u)|e^{m_0|u|}du\Big\} \\ &\leqslant 2\sum_{i\in\mathcal{J}}\mathbb{E}\Big\{|X_iY_i|e^{m_0|Y_i|}\Big\} \leqslant 2a_n^{-2}\sum_{i\in\mathcal{J}}\mathbb{E}\big\{|a_n^2T_i^2|e^{m_0T_i}\big\} \\ &\leqslant Ca_n^{-2}b.\end{aligned} \tag{A.16}$$

This completes the proof. □

A.1 Proofs of preliminary lemmas

Proof of Lemma A.1. By Hölder's inequality, for any random variables $U_1, U_2, U_3 \geqslant 0$, we have

$$\mathbb{E}\{U_1U_2U_3\} \leqslant (\mathbb{E}\{U_1U_2^{(1+\varepsilon)/\varepsilon}\})^{\varepsilon/(1+\varepsilon)}(\mathbb{E}\{U_1U_3^{(1+\varepsilon)}\})^{1/(1+\varepsilon)}, \tag{A.17}$$

where $\varepsilon = 16m_0/a_n$. Then

$$0<\varepsilon\leqslant 1, \quad \varepsilon m_0^2 \leqslant 1/3 \quad \text{and} \quad (1+\varepsilon)m_0/\varepsilon \leqslant a_n/8. \tag{A.18}$$

Applying (A.17) with $U_1 = \zeta_i, U_2 = e^{tT_i}$ and $U_3 = \Psi_{\beta,t}(W_i)$, and by (A.18), we have

$$\begin{aligned}&\mathbb{E}\big\{\zeta_i\Psi_{\beta,t}(W_i)e^{3tT_i}\big\} \\ &\quad\leqslant \big(\mathbb{E}\big\{\zeta_ie^{3(1+\varepsilon)tT_i/\varepsilon}\big\}\big)^{\varepsilon/(1+\varepsilon)}\big(\mathbb{E}\big\{\zeta_i\Psi_{\beta,t}(W_i)^{1+\varepsilon}\big\}\big)^{1/(1+\varepsilon)} \\ &\quad= \big(\mathbb{E}\big\{\zeta_ie^{3(1+\varepsilon)tT_i/\varepsilon}\big\}\big)^{\varepsilon/(1+\varepsilon)}\big(\mathbb{E}\zeta_i\big)^{1/(1+\varepsilon)}\big(\mathbb{E}\big\{\Psi_{\beta,t}(W_i)^{1+\varepsilon}\big\}\big)^{1/(1+\varepsilon)} \\ &\quad\leqslant \mathbb{E}\big\{\zeta_ie^{3(1+\varepsilon)tT_i/\varepsilon}\big\}\big(\mathbb{E}\big\{\Psi_{\beta,t}(W_i)^{1+\varepsilon}\big\}\big)^{1/(1+\varepsilon)} \\ &\quad\leqslant \mathbb{E}\big\{\zeta_ie^{3a_nT_i/8}\big\}\,\mathbb{E}\big\{\Psi_{\beta,t}(W_i)^{1+\varepsilon}\big\},\end{aligned} \tag{A.19}$$

where the equality in the third line follows from the fact that W_i is independent of ζ_i and the last inequality follows from (A.18) and $\Psi_{\beta,t}\geqslant 1$. Recalling the definition of $\Psi_{\beta,t}(w)$ in (2.6), we have for any u and v,

$$\Psi_{\beta,t}(u+v) \leqslant \Psi_{\beta,t}(u)e^{t|v|}. \tag{A.20}$$

By (A.20) and Hölder's inequality,

$$\mathbb{E}\big\{\Psi_{\beta,t}(W_i)^{1+\varepsilon}\big\} \leqslant \mathbb{E}\big\{\Psi_{\beta,t}(W)^{1+\varepsilon}e^{(1+\varepsilon)tT_i}\big\} := H_1 \times H_2, \tag{A.21}$$

where

$$H_1 = \mathbb{E}\big\{\Psi_{\beta,t}(W)^{(1+\varepsilon)^2}\big\}, \quad H_2 = \big(\mathbb{E}\big\{e^{(1+\varepsilon)^2 tT_i/\varepsilon}\big\}\big)^{\varepsilon/(1+\varepsilon)}.$$

Recalling the definition of $\Psi_{\beta,t}(w)$, we have

$$\Psi_{\beta,t}(w) \leqslant 2e^{m_0^2} + 1 \leqslant 3e^{m_0^2} \quad \text{for } 0 \leqslant \beta, t \leqslant m_0,$$

which further implies that

$$\Psi_{\beta,t}(w)^{(1+\varepsilon)^2} \leqslant (3e^{m_0^2})^{2\varepsilon+\varepsilon^2}\Psi_{\beta,t}(w). \tag{A.22}$$

By (A.18) and (A.22) we have

$$H_1 \leqslant 27e^{m_0^2(2\varepsilon+\varepsilon^2)}h(t) \leqslant 27e^{3m_0^2\varepsilon}h(t) \leqslant 81h(t). \tag{A.23}$$

For H_2, by (A.18) and by Hölder's inequality again, we have

$$H_2 \leqslant \mathbb{E}e^{a_nT_i/4} \leqslant b^{1/4}. \tag{A.24}$$

Combining (A.19), (A.21), (A.23) and (A.24) yields (A.1).

Now, we prove (A.2). Expanding the square term of the left hand side of (A.2), we have for all $\tau > 0$,

$$\begin{aligned} \mathbb{E}\big\{\zeta_i T_i^2 e^{3tT_i}\Psi_{\beta,t}(W_i)\big\} &= \sum_{j\in B_i}\sum_{k\in B_i}\mathbb{E}\big\{\zeta_i X_j X_k e^{3tT_i}\Psi_{\beta,t}(W_i)\big\} \\ &\leqslant \kappa\sum_{j\in B_i}\mathbb{E}\big\{\zeta_i X_j^2 e^{3tT_i}\Psi_{\beta,t}(W_i)\big\} \\ &\leqslant \frac{\kappa^2}{2\tau}\mathbb{E}\{\zeta_i^2 e^{6tT_i}\Psi_{\beta,t}(W_i)\} + \frac{\kappa\tau}{2}\sum_{j\in B_i}\mathbb{E}\big\{X_j^4\Psi_{\beta,t}(W_i)\big\}. \end{aligned} \tag{A.25}$$

For the first term of the right hand side of (A.25), by (A.1) with replacing ζ_i by ζ_i^2 and $3tT_i$ by $6tT_i$, we obtain

$$\mathbb{E}\{\zeta_i^2 e^{6tT_i}\Psi_{\beta,t}(W_i)\} \leqslant 81b^{1/4}h(t)\,\mathbb{E}\{\zeta_i^2 e^{3a_nT_i/4}\}. \tag{A.26}$$

For the second term of right hand side of (A.25), by (A.20), we have for any $j \in B_i$, with $W_{ij} = W - \sum_{k\in B_i\cup B_j} X_k$,

$$\mathbb{E}\big\{X_j^4\Psi_{\beta,t}(W_i)\big\} \leqslant \mathbb{E}\{X_j^4 e^{tT_j}\Psi_{\beta,t}(W_{ij})\}.$$

Similar to (A.1), we obtain

$$\mathbb{E}\{X_j^4 e^{tT_j}\Psi_{\beta,t}(W_{ij})\} \leqslant 81b^{1/2}h(t)\,\mathbb{E}\{X_j^4 e^{3a_nT_j/8}\}. \tag{A.27}$$

Observing that $X_j \leqslant T_j$, we have the expectation term at the right hand side of (A.27) can be bounded by

$$\mathbb{E}\{X_j^4 e^{tT_j}\} \leqslant \alpha_n^{-4}\,\mathbb{E}\{(a_n T_j)^4 e^{3a_n T_j/8}\} \leqslant C\alpha_n^{-4}\,\mathbb{E}e^{a_n T_j/2} \leqslant Ca_n^{-4} b^{1/2} h(t). \tag{A.28}$$

Substituting (A.26)–(A.28) to (A.25) yields (A.2). $\square$

Proof of Lemma A.2. Expanding the left hand side of (A.3) yields

$$\mathbb{E}\big\{S^2 \Psi_{\beta,t}(W)\big\} := I_1 + I_2, \tag{A.29}$$

where

$$I_1 = \sum_{i\in\mathcal{J}}\sum_{j\in N_i} \mathbb{E}\{\xi_i\xi_j\Psi_{\beta,t}(W)\}, \quad I_2 = \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}\setminus N_i} \mathbb{E}\{\xi_i\xi_j\Psi_{\beta,t}(W)\}.$$

We now give the bounds of I_1 and I_2 separately. Observe that

$$\{(i,j) : i\in\mathcal{J}, j\in N_i\} = \{(i,j) : B_i\cap B_j \neq \varnothing\} = \{(i,j) : j\in\mathcal{J}, i\in N_j\}. \tag{A.30}$$

Recall that $W_i = W - \sum_{j\in B_i} X_j$. For I_1, we have

$$\begin{aligned}
I_1 &\leqslant \frac{1}{2}\sum_{i\in\mathcal{J}, j\in N_i} \mathbb{E}\{(\xi_i^2+\xi_j^2)\Psi_{\beta,t}(W)\} \\
&= \frac{1}{2}\sum_{i\in\mathcal{J}}\sum_{j\in N_i}\mathbb{E}\{\xi_i^2\Psi_{\beta,t}(W)\} + \frac{1}{2}\sum_{i\in\mathcal{J}}\sum_{j\in N_i}\mathbb{E}\{\xi_j^2\Psi_{\beta,t}(W)\} \\
&= \frac{1}{2}\sum_{i\in\mathcal{J}}\sum_{j\in N_i}\mathbb{E}\{\xi_i^2\Psi_{\beta,t}(W)\} + \frac{1}{2}\sum_{j\in\mathcal{J}}\sum_{i\in N_j}\mathbb{E}\{\xi_j^2\Psi_{\beta,t}(W)\} \\
&= \sum_{i\in\mathcal{J}}\sum_{j\in N_i}\mathbb{E}\{\xi_i^2\Psi_{\beta,t}(W)\} \leqslant \kappa\sum_{i\in\mathcal{J}}\mathbb{E}\{\xi_i^2\Psi_{\beta,t}(W_i)e^{tT_i}\},
\end{aligned} \tag{A.31}$$

where we used (A.30), (A.20) and $|N_i| \leqslant \kappa$ in the last line. By Lemma A.1 with $\zeta_i = \xi_i^2$, we obtain

$$\mathbb{E}\big\{\xi_i^2\Psi_{\beta,t}(W_i)e^{tT_i}\big\} \leqslant 81 b^{1/4} h(t)\,\mathbb{E}\big\{\xi_i^2 e^{3a_nT_i/8}\big\}. \tag{A.32}$$

Substituting (A.32) into (A.31), we have

$$I_1 \leqslant 81 b^{1/4}\kappa h(t)\sum_{i\in\mathcal{J}}\mathbb{E}\{\xi_i^2 e^{3a_nT_i/8}\}. \tag{A.33}$$

For $i,j\in\mathcal{J}$, let $W_{ij} = W - \sum_{k\in B_i\cup B_j} X_k$, $V_{ij} = \sum_{k\in B_i\cup B_j} X_k, T_{ij} = \sum_{k\in B_i\cup B_j}|X_k|$. It is easy to see that $|T_{ij}| \leqslant T_i + T_j$. In order bound I_2, for any $i\in\mathcal{J}$ and $j\notin N_i$, by

Taylor's expansion, we have

$$
\begin{aligned}
&\mathbb{E}\{\xi_i\xi_j\Psi_{\beta,t}(W)\}\\
&\quad= \mathbb{E}\{\xi_i\xi_j\Psi_{\beta,t}(W_i)\} + \mathbb{E}\{\xi_i\xi_j V_i\Psi'_{\beta,t}(W_i)\} + \int_0^1 \mathbb{E}\{\xi_i\xi_j V_i^2\Psi''_{\beta,t}(W_i+uV_i)\}du\\
&\quad= \mathbb{E}\{\xi_i\xi_j\Psi_{\beta,t}(W_i)\} + \mathbb{E}\{\xi_i\xi_j V_i\Psi'_{\beta,t}(W_{ij})\}\\
&\qquad+ \int_0^1 \mathbb{E}\{\xi_i\xi_j V_i(V_{ij}-V_i)\Psi''_{\beta,t}(W_{ij}+u(W_i-W_{ij}))\}du\\
&\qquad+ \int_0^1 \mathbb{E}\{\xi_i\xi_j V_i^2\Psi''_{\beta,t}(W_i+uV_i)\}du.
\end{aligned}
$$

If $j\notin N_i$, then ξ_i is independent of (ξ_j, W_i) and ξ_j is independent of (ξ_i, V_i, W_{ij}). Recalling that $\mathbb{E}\xi_i=\mathbb{E}\xi_j=0$ by assumption, we have $\mathbb{E}\{\xi_i\xi_j\Psi_{\beta,t}(W_i)\}=\mathbb{E}\{\xi_i\xi_j V_i\Psi'_{\beta,t}(W_{ij})\}=0$. By (5.26) and by the monotonicity of $\Psi_{\beta,t}(\cdot)$, we have for any $0\leqslant u\leqslant 1$,

$$
\begin{aligned}
&|\mathbb{E}\{\xi_i\xi_j V_i(V_{ij}-V_i)\Psi''_{\beta,t}(W_{ij}+u(W_i-W_{ij}))\}|\\
&\quad\leqslant t^2\,\mathbb{E}\{|\xi_i\xi_j V_i(V_{ij}-V_i)|\Psi_{\beta,t}(W_{ij}+u(W_i-W_{ij}))\}\\
&\quad\leqslant t^2\sum_{l\in B_i}\sum_{m\in B_j\setminus B_i}\mathbb{E}\{|\xi_i\xi_j X_l X_m|(\Psi_{\beta,t}(W_i)+\Psi_{\beta,t}(W_{ij}))\}\\
&\quad\leqslant 2t^2\sum_{l\in B_i}\sum_{m\in B_j\setminus B_i}\mathbb{E}\{|\xi_i\xi_j X_l X_m|\Psi_{\beta,t}(W_{ij})e^{t(T_i+T_j)}\},
\end{aligned}
$$

and similarly,

$$
|\mathbb{E}\{\xi_i\xi_j V_i^2\Psi''_{\beta,t}(W_i+uV_i)\}|\leqslant 2t^2\sum_{l\in B_i}\sum_{m\in B_i}\mathbb{E}\big\{|\xi_i\xi_j X_l X_m|\Psi_{\beta,t}(W_{ij})e^{t(T_i+T_j)}\big\}.
$$

Observe that

$$
(T_i+T_j)^2 e^{t(T_i+T_j)}\leqslant 4(T_i^2e^{2tT_i}+T_j^2e^{2tT_j}).
$$

Hence, it follows that

$$
\begin{aligned}
I_2 &\leqslant 2t^2\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}\setminus N_i}\sum_{l\in B_i}\sum_{m\in B_i\cup B_j}\mathbb{E}\{|\xi_i\xi_j X_l X_m|\Psi_{\beta,t}(W_{ij})e^{t(T_i+T_j)}\}\\
&\leqslant 2t^2\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}\setminus N_i}\mathbb{E}\big\{|\xi_i\xi_j(T_i+T_j)^2|e^{t(T_i+T_j)}\Psi_{\beta,t}(W_{ij})\big\}\\
&\leqslant 8t^2\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}\setminus N_i}\mathbb{E}\big\{|\xi_i\xi_j|(T_i^2e^{2tT_i}+T_j^2e^{2tT_j})\Psi_{\beta,t}(W_{ij})\big\}\\
&= 16t^2\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}\setminus N_i}\mathbb{E}\big\{|\xi_i\xi_j|T_i^2e^{2tT_i}\Psi_{\beta,t}(W_{ij})\big\},
\end{aligned}
\tag{A.34}
$$

where we used (A.30) in the last line. If $j \in \mathcal{J} \setminus N_i$, then ξ_j is independent of (ξ_i, T_i, W_{ij}). Therefore,

$$
\begin{aligned}
I_2 &\leqslant 16t^2 \sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{J} \setminus N_i} \mathbb{E}|\xi_j| \, \mathbb{E}\big\{|\xi_i| T_i^2 e^{2tT_i} \Psi_{\beta,t}(W_{ij})\big\} \\
&\leqslant 16t^2 \sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{J} \setminus N_i} \mathbb{E}|\xi_j| \, \mathbb{E}\big\{|\xi_i| T_i^2 e^{3tT_i} \Psi_{\beta,t}(W_i)\big\} \\
&\leqslant Cb\kappa^2 t^2 h(t) \sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{J} \setminus N_i} \mathbb{E}|\xi_j| \big(\tau^{-1} \, \mathbb{E}\{\xi_i^2 e^{3a_n T_i/4}\} + \tau a_n^{-4}\big).
\end{aligned}
$$

Combining (A.33) and (A.34), we complete the proof. □

B PROOF OF LEMMA 7.1

This section includes three subsections. In Appendix B.1, we prove Lemma 7.1. Before that, we give some preliminary lemmas, whose proofs are given in Appendices B.2 and B.3.

B.1 Proof of Lemma 7.1

For any $\boldsymbol{i} \in [n]_k$, we write i_l be the l-th element of $\boldsymbol{i}$ and write $\pi(\boldsymbol{i}) = (\pi(i_1), \ldots, \pi(i_k))$, Let $A(\boldsymbol{i}) = \{i_1, \ldots, i_k\}$ be the set of elements in $\boldsymbol{i}$. For any $\boldsymbol{i}, \boldsymbol{j} \in [n]_k$ and any matrix $(x_{i,j})_{1 \leqslant i,j \leqslant n}$, let $x_{\boldsymbol{i},\boldsymbol{j}} = (x_{i,j} : i \in A(\boldsymbol{i}), j \in A(\boldsymbol{j}))$. Let $T_{\boldsymbol{i},\boldsymbol{j}} = \sum_{l=1}^{k} |X_{i_l, j_l}|$. For any two positive integers $m \leqslant n$, we write $n_{(m)} := \prod_{i=1}^{m} (n - i + 1)$ as the *descending factorial*. Let $h(t) = \mathbb{E}\{\Psi_{\beta,t}(W)\}$. The following preliminary lemmas are useful in the proof of Lemma 7.1.

Lemma B.1. *Let $0 \leqslant m \leqslant 2$ be an integer and assume that $n \geqslant 4$. Let σ be a uniform permutation on $[n-m]$ which is independent of $\boldsymbol{X}$ and let $S = \sum_{i \in [n-m]} X_{i,\sigma(i)}$. For $k = 1, 2$, and for any $\boldsymbol{i}, \boldsymbol{j} \in [n-m]_k$, let $\zeta_{\boldsymbol{i},\boldsymbol{j}} := \zeta(X_{\boldsymbol{i},\boldsymbol{j}})$ be a positive function of $X_{\boldsymbol{i},\boldsymbol{j}}$. We have*

$$
\mathbb{E}\{\zeta_{\boldsymbol{i},\sigma(\boldsymbol{i})} \Psi_{\beta,t}(S)\} \leqslant 4b^{1/16} \, \mathbb{E}\Psi_{\beta,t}(S) \max_{\boldsymbol{v} \in [n-m]_k} \mathbb{E}\{\zeta_{\boldsymbol{i},\boldsymbol{v}} e^{tT_{\boldsymbol{i},\boldsymbol{v}}}\}, \tag{B.1}
$$

$$
\mathbb{E}\{\zeta_{\sigma^{-1}(\boldsymbol{j}),\boldsymbol{j}} \Psi_{\beta,t}(S)\} \leqslant 4b^{1/16} \, \mathbb{E}\Psi_{\beta,t}(S) \max_{\boldsymbol{v} \in [n-m]_k} \mathbb{E}\{\zeta_{\boldsymbol{i},\boldsymbol{v}} e^{tT_{\boldsymbol{i},\boldsymbol{v}}}\}, \tag{B.2}
$$

$$
\mathbb{E}\{\zeta_{\sigma^{-1}(\boldsymbol{j}),\sigma(\boldsymbol{i})} \Psi_{\beta,t}(S)\} \leqslant 4b^{1/8} \, \mathbb{E}\Psi_{\beta,t}(S) \max_{\boldsymbol{u},\boldsymbol{v} \in [n-m]_k} \mathbb{E}\{\zeta_{\boldsymbol{u},\boldsymbol{v}} e^{t(T_{\boldsymbol{i},\boldsymbol{v}} + T_{\boldsymbol{u},\boldsymbol{j}})}\}. \tag{B.3}
$$

Proof. We only prove (B.1), because (B.2) and (B.3) can be shown similarly. For any $\boldsymbol{i} \in [n-m]_k$, let $S^{(\boldsymbol{i})} = \sum_{i' \notin A(\boldsymbol{i})} X_{i',\sigma(i')}$. By definition, we have

$$
\begin{aligned}
\mathbb{E}\{\zeta_{\boldsymbol{i},\sigma(\boldsymbol{i})} \Psi_{\beta,t}(S)\} &= \frac{1}{(n-m)_k} \sum_{\boldsymbol{j} \in [n-m]_k} \mathbb{E}\{\zeta_{\boldsymbol{i},\boldsymbol{j}} \Psi_{\beta,t}(S) | \sigma(\boldsymbol{i}) = \boldsymbol{j}\} \\
&\leqslant \frac{1}{(n-m)_k} \sum_{\boldsymbol{j} \in [n-m]_k} \mathbb{E}\{\zeta_{\boldsymbol{i},\boldsymbol{j}} e^{tT_{\boldsymbol{i},\boldsymbol{j}}} \Psi_{\beta,t}(S^{(\boldsymbol{i})}) | \sigma(\boldsymbol{i}) = \boldsymbol{j}\}.
\end{aligned} \tag{B.4}
$$

Since $X_{\boldsymbol{i},\boldsymbol{j}}$ is conditionally independent of $S^{(\boldsymbol{i})}$ given the event that $\sigma(\boldsymbol{i}) = \boldsymbol{j}$, and $X_{\boldsymbol{i},\boldsymbol{j}}$ is independent of σ, the conditional expectation term of (B.4) can be rewritten as

$$\mathbb{E}\{\zeta_{\boldsymbol{i},\boldsymbol{j}} e^{tT_{\boldsymbol{i},\boldsymbol{j}}} \Psi_{\beta,t}(S^{(\boldsymbol{i})}) | \sigma(\boldsymbol{i}) = \boldsymbol{j}\} = \mathbb{E}\{\zeta_{\boldsymbol{i},\boldsymbol{j}} e^{tT_{\boldsymbol{i},\boldsymbol{j}}}\} \, \mathbb{E}\{\Psi_{\beta,t}(S^{(\boldsymbol{i})}) | \sigma(\boldsymbol{i}) = \boldsymbol{j}\}. \tag{B.5}$$

For the second term on the RHS of (B.5), note that $\Psi_{\beta,t}(w + x) \leqslant e^{t|x|}\Psi_{\beta,t}(w)$ and $|S - S^{(\boldsymbol{i})}| \leqslant T_{\boldsymbol{i},\boldsymbol{j}}$ given $\sigma(\boldsymbol{i}) = \boldsymbol{j}$. With $\varepsilon = \alpha_n^{-2/3}$, by Hölder's inequality, and noting that $T_{\boldsymbol{i},\boldsymbol{j}}$ is independent of σ, we have

$$\begin{aligned} &\mathbb{E}\{\Psi_{\beta,t}(S^{(\boldsymbol{i})}) | \sigma(\boldsymbol{i}) = \boldsymbol{j}\} \\ &\quad \leqslant \mathbb{E}\{e^{tT_{\boldsymbol{i},\boldsymbol{j}}} \Psi_{\beta,t}(S) | \sigma(\boldsymbol{i}) = \boldsymbol{j}\} \\ &\quad \leqslant \big(\mathbb{E}\big\{e^{(1+\varepsilon)tT_{\boldsymbol{i},\boldsymbol{j}}/\varepsilon}\big\}\big)^{\varepsilon/(1+\varepsilon)} \big(\mathbb{E}\big\{\Psi_{\beta,t}^{1+\varepsilon}(S)\big\} \bigm| \sigma(\boldsymbol{i}) = \boldsymbol{j}\big)^{1/(1+\varepsilon)}. \end{aligned} \tag{B.6}$$

Noting that $0 \leqslant t \leqslant \beta \leqslant \alpha_n^{1/3}/64$, we obtain

$$0 < \varepsilon < 1, \quad (1+\varepsilon)t/\varepsilon \leqslant 2t/\varepsilon \leqslant \alpha_n/32, \quad \varepsilon\beta t \leqslant 0.001. \tag{B.7}$$

For the first term in the RHS of (B.6), by (4.4), we obtain

$$\mathbb{E}\big\{e^{(1+\varepsilon)tT_{\boldsymbol{i},\boldsymbol{j}}/\varepsilon}\big\} \leqslant \mathbb{E}e^{\alpha_n T_{\boldsymbol{i},\boldsymbol{j}}/32} \leqslant \max_{i,j\in[n-m]} \mathbb{E}e^{\alpha_n |X_{i,j}|/16} \leqslant b^{1/16}. \tag{B.8}$$

Noting that $\Psi_{\beta,t}(w) \leqslant 2e^{t\beta} + 1 \leqslant 3e^{t\beta}$, we have

$$\Psi_{\beta,t}(w)^{1+\varepsilon} \leqslant (3e^{t\beta})^{\varepsilon}\Psi_{\beta,t}(w) \leqslant 4\Psi_{\beta,t}(w).$$

Thus,

$$\mathbb{E}\{\Psi_{\beta,t}(S)^{1+\varepsilon} | \sigma(\boldsymbol{i}) = \boldsymbol{j}\} \leqslant 4\,\mathbb{E}\{\Psi_{\beta,t}(S) | \sigma(\boldsymbol{i}) = \boldsymbol{j}\}. \tag{B.9}$$

Combining (B.4)–(B.6), (B.8) and (B.9), we have

$$\mathbb{E}\{\zeta_{\boldsymbol{i},\boldsymbol{j}} e^{tT_{\boldsymbol{i},\boldsymbol{j}}} \Psi_{\beta,t}(S^{(\boldsymbol{i})}) | \sigma(\boldsymbol{i}) = \boldsymbol{j}\} \leqslant 4b^{1/16} \max_{\boldsymbol{v}\in[n-m]_k} \mathbb{E}\{\zeta_{\boldsymbol{i},\boldsymbol{v}} e^{tT_{\boldsymbol{i},\boldsymbol{v}}}\} \, \mathbb{E}\{\Psi_{\beta,t}(S) | \sigma(\boldsymbol{i}) = \boldsymbol{j}\}.$$

Now, taking summation over $\boldsymbol{j} \in [n-m]_k$ yields (B.1). Using a similar argument, we obtain (B.2) and (B.3). ☐

Lemma B.2. *For $\boldsymbol{i}, \boldsymbol{j}$, let $\xi_{\boldsymbol{i},\boldsymbol{j}} := \xi(X_{\boldsymbol{i},\boldsymbol{j}})$ be a function of $X_{\boldsymbol{i},\boldsymbol{j}}$ such that $\mathbb{E}\xi_{\boldsymbol{i},\pi(\boldsymbol{i})} = 0$. For any $\boldsymbol{i} \in [n]_2$ and $\boldsymbol{i}' \in [n]_2^{(\boldsymbol{i})}$, where $[n]_2^{(\boldsymbol{i})} := \{(k,l) \in [n]_2 : k, l \in [n] \setminus A(\boldsymbol{i})\}$, we have*

$$\begin{aligned} &|\,\mathbb{E}\{\xi_{\boldsymbol{i},\pi(\boldsymbol{i})}\xi_{\boldsymbol{i}',\pi(\boldsymbol{i}')}\Psi_{\beta,t}(W)\}| \\ &\quad \leqslant Cb(1+t^2)h(t)n^{-4} \sum_{\boldsymbol{j},\boldsymbol{j}'\in[n]_2} \big(\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}|T_{\boldsymbol{i},\boldsymbol{j}}^2 e^{2t|T_{\boldsymbol{i},\boldsymbol{j}}|}\}\,\mathbb{E}\{|\xi_{\boldsymbol{i}',\boldsymbol{j}'}|\} + \mathbb{E}\{|\xi_{\boldsymbol{i}',\boldsymbol{j}'}|T_{\boldsymbol{i}',\boldsymbol{j}'}^2 e^{2t|T_{\boldsymbol{i}',\boldsymbol{j}'}|}\}\,\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}|\}\big) \\ &\qquad + Cb(1+t^2)h(t)n^{-4} \sum_{\boldsymbol{j},\boldsymbol{j}'\in[n]_2} \big(\alpha_n^{-2} + n^{-1} + \mathbf{1}(E_{\boldsymbol{j},\boldsymbol{j}'})\big)\,\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}|\}\,\mathbb{E}\{|\xi_{\boldsymbol{i}',\boldsymbol{j}'}|\}. \end{aligned}$$

where $E_{\boldsymbol{j},\boldsymbol{j}'} = \{A(\boldsymbol{j}) \cap A(\boldsymbol{j}') \neq \varnothing\}$.

Lemma B.3. *For $\boldsymbol{i}, \boldsymbol{j} \in [n]_2$, let $g_{\boldsymbol{i},\boldsymbol{j}}(u) = D_{\boldsymbol{i},\boldsymbol{j}}\big(\mathbf{1}(-D_{\boldsymbol{i},\boldsymbol{j}} \leqslant u \leqslant 0) - \mathbf{1}(0 < u \leqslant -D_{\boldsymbol{i},\boldsymbol{j}})\big)$ and $\bar{g}_{\boldsymbol{i},\boldsymbol{j}}(u) = g_{\boldsymbol{i},\boldsymbol{j}}(u) - \mathbb{E}g_{\boldsymbol{i},\pi(\boldsymbol{i})}(u)$. We have*

$$\begin{aligned}\bigg|\int_{|u|\leqslant 1} |u|^v e^{2t|u|}\,\mathbb{E}\bigg\{\Big(\sum_{\boldsymbol{i}\in[n]_2} \bar{g}_{\boldsymbol{i},\pi(\boldsymbol{i})}(u)\Big)^2 \Psi_{\beta,t}(W)\bigg\}du\bigg| \\ \leqslant Cb^2(n^2\alpha_n^{-5-v} + n\alpha_n^{-3-v})(1+t^2)h(t), \quad \textit{for } v = 0, 1. \quad \text{(B.10)}\end{aligned}$$

Lemma B.4. *For $\boldsymbol{i}, \boldsymbol{j} \in [n]_2$, let $H_{\boldsymbol{i},\boldsymbol{j}}(u) = D^2_{\boldsymbol{i},\boldsymbol{j}}$ and $\bar{H}_{\boldsymbol{i},\boldsymbol{j}}(u) = H_{\boldsymbol{i},\boldsymbol{j}}(u) - \mathbb{E}H_{\boldsymbol{i},\pi(\boldsymbol{i})}(u)$. We have*

$$\bigg|\mathbb{E}\bigg\{\Big(\sum_{\boldsymbol{i}\in[n]_2} \bar{H}_{\boldsymbol{i},\pi(\boldsymbol{i})}\Big)^2 \Psi_{\beta,t}(W)\bigg\}\bigg| \leqslant Cb^2(n^4\alpha_n^{-6} + n^3\alpha_n^{-4})(1+t^2)h(t). \tag{B.11}$$

The proofs of Lemmas B.2 and B.3 are given in Appendix B.3. The proof of Lemma B.4 is similar to that of Lemma B.3 and thus we omit the details.

Now, we are ready to give the proof of Lemma 7.1.

Proof of Lemma 7.1. We prove (7.4)–(7.9) one by one.

(i). Proof of (7.4). Let $\bar{X}_{i,j} = X_{i,j} - a_{i,j}$. We have

$$\mathbb{E}\{|R|\Psi_{\beta,t}(W)\} \leqslant h^{1/2}(t)(\mathbb{E}\{R^2\Psi_{\beta,t}(W)\})^{1/2}. \tag{B.12}$$

By (4.1) and (7.1), we have

$$\begin{aligned}\mathbb{E}\{R^2\Psi_{\beta,t}(W)\} &= \frac{1}{n^2}\,\mathbb{E}\bigg\{\bigg|\sum_{i=1}^n\sum_{j=1}^n \bar{X}_{i,j}\bigg|^2 \Psi_{\beta,t}(W)\bigg\} \\ &= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\sum_{i'=1}^n\sum_{j'=1}^n \mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W)\}.\end{aligned} \tag{B.13}$$

Now, for fixed $i, j, i', j' \in [n]$,

$$\begin{aligned}&\mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W)\} \\ &= \mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W), \pi(i) = j, \pi(i') = j'\} + \mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W), \pi(i) = j, \pi(i') \neq j'\} \\ &\quad + \mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W), \pi(i) \neq j, \pi(i') = j'\} + \mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W), \pi(i) \neq j, \pi(i') \neq j'\}.\end{aligned}$$

For the first term, with $W^{(i,i')} = W - \sum_{k\in\{i,i'\}} X_{k,\pi(k)}$,

$$\begin{aligned}&|\,\mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W)|\pi(i) = j, \pi(i') = j'\}| \\ &\quad\leqslant \frac{1}{2}\,\mathbb{E}\{(\bar{X}^2_{i,j} + \bar{X}^2_{i',j'})\Psi_{\beta,t}(W)|\pi(i) = j, \pi(i') = j'\} \\ &\quad\leqslant \frac{1}{2}\,\mathbb{E}\{(\bar{X}^2_{i,j} + \bar{X}^2_{i',j'})e^{t(|X_{i,j}|+|X_{i',j'}|)}\Psi_{\beta,t}(W^{(i,i')})|\pi(i) = j, \pi(i') = j'\} \\ &\quad= \frac{1}{2}\,\mathbb{E}\{(\bar{X}^2_{i,j} + \bar{X}^2_{i',j'})e^{t(|X_{i,j}|+|X_{i',j'}|)}\}\,\mathbb{E}\{\Psi_{\beta,t}(W^{(i,i')})|\pi(i) = j, \pi(i') = j'\},\end{aligned}$$

where in the last line we used the fact that $(X_{i,j}, X_{i',j'})$ and $W^{(i,i')}$ are conditionally independent given $\pi(i) = j$ and $\pi(i') = j'$. Recalling that that $t \leqslant \alpha_n^{1/3}/4 \leqslant \alpha_n/4$ and by (4.4), we have

$$\begin{aligned} \mathbb{E}\{\bar{X}_{i,j}^2 e^{t|X_{i,j}|}\} &\leqslant C\alpha_n^{-2}\,\mathbb{E}\{|\alpha_n X_{i,j}|^2 e^{\alpha_n|X_{i,j}|/4}\} \\ &\leqslant C\alpha_n^{-2}\,\mathbb{E}\{e^{\alpha_n|X_{i,j}|/2}\} \leqslant Cb^{1/2}\alpha_n^{-2}. \end{aligned} \tag{B.14}$$

Choosing $\varepsilon = \alpha_n^{-2/3}$ and recalling (4.4) and (B.7), we have

$$\mathbb{E}\{e^{(1+\varepsilon)t(|X_{i,j}|+|X_{i',j'}|)/\varepsilon}\} \leqslant \mathbb{E}\{e^{\alpha_n(|X_{i,j}|+|X_{i',j'}|)/2}\} \leqslant Cb,$$

and

$$\begin{aligned} &\mathbb{E}\{\Psi_{\beta,t}(W^{(i,i')})|\pi(i) = j, \pi(i') = j'\} \\ &\quad\leqslant \mathbb{E}\{e^{t|W-W^{(i,j)}|}\Psi_{\beta,t}(W)|\pi(i) = j, \pi(i') = j'\} \\ &\quad\leqslant (\mathbb{E}\{e^{(1+\varepsilon)t(|X_{i,j}|+|X_{i',j'}|)/\varepsilon}\})^{\frac{\varepsilon}{1+\varepsilon}}(\mathbb{E}\{\Psi_{\beta,t}^{1+\varepsilon}(W)|\pi(i) = j, \pi(i') = j'\})^{\frac{1}{1+\varepsilon}} \\ &\quad\leqslant Cb^{1/2}\,\mathbb{E}\{\Psi_{\beta,t}(W)|\pi(i) = j, \pi(i') = j'\}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} &|\mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W)|\pi(i) = j, \pi(i') = j'\}| \\ &\qquad\qquad\leqslant Cb\alpha_n^{-2}\,\mathbb{E}\{\Psi_{\beta,t}(W)|\pi(i) = j, \pi(i') = j'\}. \end{aligned} \tag{B.15}$$

Moreover, noting that $X_{i',j'}$ is independent of $(X_{i,j}, W)$ given $\pi(i) = j$ and $\pi(i') \neq j'$, we have

$$\begin{aligned} &\mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W)|\pi(i) = j, \pi(i') \neq j'\} \\ &\qquad= \mathbb{E}\{\bar{X}_{i',j'}\}\,\mathbb{E}\{\bar{X}_{i,j}\Psi_{\beta,t}(W)|\pi(i) = j, \pi(i') \neq j'\} = 0. \end{aligned} \tag{B.16}$$

Similarly,

$$\mathbb{E}\{X_{i,j}X_{i',j'}\Psi_{\beta,t}(W)|\pi(i) \neq j, \pi(i') = j'\} = 0. \tag{B.17}$$

Furthermore, if $\pi(i) \neq j'$ and $\pi(i') \neq j'$, we have

$$\begin{aligned} &\mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W)|\pi(i) \neq j, \pi(i') \neq j'\} \\ &\quad= \mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\}\,\mathbb{E}\{\Psi_{\beta,t}(W)|\pi(i) \neq j, \pi(i') \neq j'\} \\ &\quad= \begin{cases} \mathbb{E}\{\bar{X}_{i,j}^2\}\,\mathbb{E}\{\Psi_{\beta,t}(W)|\pi(i) \neq j, \pi(i') \neq j'\} & \text{if } i = i' \text{ and } j = j', \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

By (B.14), we have

$$|\mathbb{E}\{\bar{X}_{i,j}\bar{X}_{i',j'}\Psi_{\beta,t}(W)|\pi(i) \neq j, \pi(i') \neq j'\}|$$

$$\leqslant Cb\alpha_n^{-2}\,\mathbb{E}\{\Psi_{\beta,t}(W)|\pi(i)\neq j,\pi(i')\neq j'\}\,\mathbf{1}((i,j)=(i',j')). \tag{B.18}$$

Substituting (B.15)–(B.18) to (B.13) and using (B.12) yields (7.4).

(ii). Proof of (7.5). Recalling the definition of $\hat{K}_1$ in (2.1) and $\hat{K}(u)$ in (7.3), we have

$$\hat{K}_1 = \frac{1}{4n}\sum_{\boldsymbol{i}\in[n]_2} D^2_{\boldsymbol{i},\pi(\boldsymbol{i})}. \tag{B.19}$$

By (4.2) and (B.19), one can verify (see, e.g., Chen and Fang (2015)) that $|\mathbb{E}\hat{K}_1-1|\leqslant 2/\sqrt{n}$. Thus,

$$\mathbb{E}\{|\hat{K}_1-1|\Psi_{\beta,t}(W)\}\leqslant \mathbb{E}\{|\hat{K}_1-\mathbb{E}\hat{K}_1|\Psi_{\beta,t}(W)\}+\frac{2}{\sqrt{n}}\,\mathbb{E}\{\Psi_{\beta,t}(W)\}. \tag{B.20}$$

For the first term of the R.H.S. of (B.20), recalling that $h(t)=\mathbb{E}\Psi_{\beta,t}(W)$, by Hölder's inequality, we have

$$\mathbb{E}\{|\hat{K}_1-\mathbb{E}\hat{K}_1|\Psi_{\beta,t}(W)\}\leqslant \frac{h^{1/2}(t)}{4n}\biggl(\mathbb{E}\biggl\{\Bigl(\sum_{\boldsymbol{i}\in[n]_2}(D^2_{\boldsymbol{i},\pi(\boldsymbol{i})}-\mathbb{E}D^2_{\boldsymbol{i},\pi(\boldsymbol{i})})\Bigr)^2\Psi_{\beta,t}(W)\biggr\}\biggr)^{1/2}. \tag{B.21}$$

Applying Lemma B.4 to the expectation in the RHS of (B.21), we obtain

$$\mathbb{E}\{|\hat{K}_1-\mathbb{E}\hat{K}_1|\Psi_{\beta,t}(W)\}\leqslant Cb(n\alpha_n^{-3}+n^{1/2}\alpha_n^{-2})h(t). \tag{B.22}$$

Combining (B.20) and (B.22) yields (7.5).

(iii). Proof of (7.6). Recalling $\hat{K}(u)$ as in (7.3), we have

$$\begin{aligned} E\{\hat{K}_2(t)\Psi_{\beta,t}(W)\} &= \int_{-\infty}^{\infty}|u|e^{t|u|}\,\mathbb{E}\{\hat{K}(u)\Psi_{\beta,t}(W)\}du\\ &\leqslant \frac{1}{4n}\sum_{\boldsymbol{i}\in[n]_2}\mathbb{E}\{|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|^3e^{t|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|}\Psi_{\beta,t}(W)\}.\end{aligned} \tag{B.23}$$

Then, applying Lemma B.1 with $k=2$, $m=0$, $\sigma=\pi$, $S=W$, and $\zeta_{\boldsymbol{i},\boldsymbol{j}}(u)=|D_{\boldsymbol{i},\boldsymbol{j}}|^3e^{t|D_{\boldsymbol{i},\boldsymbol{j}}|}$, we have for any $\boldsymbol{i}\in[n]_2$,

$$\begin{aligned} &\mathbb{E}\{|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|^3e^{t|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|}\Psi_{\beta,t}(W)\}\\ &\quad\leqslant Cb^{1/8}h(t)\max_{\boldsymbol{j}\in[n]_2}\mathbb{E}\{|D_{\boldsymbol{i},\boldsymbol{j}}|^3e^{t|D_{\boldsymbol{i},\boldsymbol{j}}|+tT_{\boldsymbol{i},\boldsymbol{j}}}\}\\ &\quad\leqslant Cb^{1/8}h(t)\alpha_n^{-3}\max_{j_1,j_2\in[n]}\sum_{i\in\{i_1,i_2\}}\sum_{j\in\{j_1,j_2\}}\mathbb{E}\{|\alpha_nX_{i,j}|^3e^{6t|X_{i,j}|}\}\\ &\quad\leqslant Cb\alpha_n^{-3}h(t).\end{aligned} \tag{B.24}$$

Therefore, by (B.23) and (B.24), we have

$$\mathbb{E}\{\hat{K}_2(t)\Psi_{\beta,t}(W)\}\leqslant Cbn\alpha_n^{-3}.$$

(iv). Proofs of (7.7) and (7.8). Recall the definitions in (2.3), (2.4) and (7.3), and we have

$$\mathbb{E}\{\hat{K}_{3,t}\Psi_{\beta,t}(W)\} = \int_{|u|\leqslant 1} e^{2t|u|}\,\mathbb{E}\{(\hat{K}(u) - K(u))^2\Psi_{\beta,t}(W)\}du. \tag{B.25}$$

By Lemma B.3 with $v = 0$, we complete the proof of (7.7). By Lemma B.4 with $v = 1$, the inequality (7.8) follows similarly.

(v). Proofs of (7.9). Recalling the definition of $\hat{K}(u)$ in (7.3), by Fubini's theorem we have

$$\sup_{0\leqslant t\leqslant \alpha_n^{1/3}/64} M_t \leqslant \frac{1}{n}\sum_{\boldsymbol{i}\in[n]_2} \mathbb{E}\{e^{\alpha_n|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|/64}|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|^2\} \leqslant Cbn\alpha^{-2} \tag{B.26}$$

where the last inequality follows from the similar argument in (B.24).

B.2 Some useful lemmas

In order to prove Lemma B.2, we need to show some preliminary lemmas. Recall that $\mathcal{S}_n$ is the collection of permutations from $[n] \to [n]$.

Lemma B.5. *For $n \geqslant 4$, $m = 0, 1, 2$, let S and σ be defined as in Lemma B.1. For any $i, j \in [n - m]$, we have*

$$|\,\mathbb{E}\{X_{i,\sigma(i)}\Psi'_{\beta,t}(S)\}| \leqslant C(n^{-1}\alpha_n^{-1} + \alpha_n^{-2})b^{1/2}(1 + t^2)\,\mathbb{E}\Psi_{\beta,t}(S), \tag{B.27}$$
$$|\,\mathbb{E}\{X_{\sigma^{-1}(j),j}\Psi'_{\beta,t}(S)\}| \leqslant C(n^{-1}\alpha_n^{-1} + \alpha_n^{-2})b^{1/2}(1 + t^2)\,\mathbb{E}\Psi_{\beta,t}(S), \tag{B.28}$$
$$|\,\mathbb{E}\{X_{\sigma^{-1}(j),\sigma(i)}\Psi'_{\beta,t}(S)\}| \leqslant C(n^{-1}\alpha_n^{-1} + \alpha_n^{-2})b^{1/2}(1 + t^2)\,\mathbb{E}\Psi_{\beta,t}(S). \tag{B.29}$$

Proof of Lemma B.5. We only prove (B.27), because (B.28) and (B.29) can be shown similarly. Note that

$$\begin{aligned}
\mathbb{E}\{X_{i,\sigma(i)}\Psi'_{\beta,t}(S)\} &= \frac{1}{n-m}\sum_{j=1}^{n-m}\mathbb{E}\{X_{i,j}\Psi'_{\beta,t}(S)|\sigma(i) = j\}\\
&= \frac{1}{n-m}\sum_{j=1}^{n-m}\mathbb{E}\{X_{i,j}\Psi'_{\beta,t}(S^{(i)})|\sigma(i) = j\}\\
&\quad + \frac{1}{n-m}\sum_{j=1}^{n-m}\mathbb{E}\{X_{i,j}(\Psi'_{\beta,t}(S) - \Psi'_{\beta,t}(S^{(i)}))|\sigma(i) = j\}\\
&:= I_1 + I_2.
\end{aligned} \tag{B.30}$$

Denote by $\tau_{i,j}$ the transposition of i and j, and define

$$\sigma_{i,j} = \begin{cases} \sigma & \text{if } \sigma(i) = j,\\ \sigma\circ\tau_{i,\sigma^{-1}(j)} & \text{if } \sigma(i) \neq j. \end{cases}$$

Then $\sigma_{i,j}(i) = j$. For any given distinct $k_1, \ldots, k_{n-m-1} \in [n-m]\setminus\{i\}$ and $l_1, \ldots, l_{n-m-1} \in [n-m] \setminus \{j\}$, denote by A the event that $\{\sigma_{i,j}(k_u) = l_u, u = 1, \ldots, n-m-1\}$. Then,

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A, \sigma(i) = j) + \sum_{u=1}^{n-m-1} \mathbb{P}(A, \sigma(i) = l_u, \sigma(k_u) = j) \\ &= \frac{1}{(n-m)!} + (n-m-1)\frac{1}{(n-m)!} = \frac{1}{(n-m-1)!}.\end{aligned}$$

On the other hand, we have

$$\mathbb{P}(\sigma(k_u) = l_u, u = 1, \ldots, n-m-1 | \sigma(i) = j) = \frac{1}{(n-m-1)!}.$$

This proves that $\mathscr{L}(\sigma_{i,j}) = \mathscr{L}(\sigma|\sigma(i) = j)$. Moreover, with $S^{(i)} = S - X_{i,\sigma(i)}$, $S_{i,j} = \sum_{i'=1}^{n} X_{i',\sigma_{i,j}(i')}$, and let $S_{i,j}^{(i)} = S_{i,j} - X_{i,j}$, it follows that

$$\mathscr{L}(S_{i,j}^{(i)}) = \mathscr{L}(S^{(i)}|\sigma(i) = j).$$

Noting that $X_{i,j}$ is independent of $\Psi_{\beta,t}(S^{(i)})$ conditional on $\sigma(i) = j$, and recalling that $\mathbb{E}X_{i,j} = a_{i,j}$, we have

$$\mathbb{E}\{X_{i,j}\Psi'_{\beta,t}(S^{(i)})|\sigma(i) = j\} = a_{i,j}\, \mathbb{E}\{\Psi'_{\beta,t}(S^{(i)})|\sigma(i) = j\} = a_{i,j}\, \mathbb{E}\{\Psi'_{\beta,t}(S_{i,j}^{(i)})\}.$$

Therefore, recalling that $\sum_{j\in[n]} a_{i,j} = 0$ by assumption (4.1), we have

$$\begin{aligned}I_1 &= \frac{\mathbb{E}\Psi'_{\beta,t}(S)}{n-m} \sum_{j\in[n-m]} a_{i,j} + \frac{1}{n-m} \sum_{j\in[n-m]} a_{i,j}\big(\mathbb{E}\Psi'_{\beta,t}(S_{i,j}^{(i)}) - \mathbb{E}\Psi'_{\beta,t}(S)\big) \\ &= I_{11} + I_{12},\end{aligned} \tag{B.31}$$

where

$$\begin{aligned}I_{11} &= -\frac{\mathbb{E}\Psi'_{\beta,t}(S)}{n-m} \sum_{j\in[n]\setminus[n-m]} a_{i,j}, \\ I_{12} &= \frac{1}{n-m} \sum_{j\in[n-m]} a_{i,j}\big(\mathbb{E}\Psi'_{\beta,t}(S_{i,j}^{(i)}) - \mathbb{E}\Psi'_{\beta,t}(S)\big).\end{aligned}$$

For I_{11}, by (4.4) and Jensen's inequality,

$$\begin{aligned}\max_{i,j}|a_{i,j}| &\leqslant \max_{i,j} \mathbb{E}|X_{i,j}| \leqslant \alpha_n^{-1} \max_{i,j} \mathbb{E}\{|\alpha_n X_{i,j}|\} \\ &\leqslant \alpha_n^{-1} \max_{i,j} \log \mathbb{E}e^{|\alpha_n X_{i,j}|} \leqslant \alpha_n^{-1} \log b.\end{aligned} \tag{B.32}$$

Thus, by (5.27), noting that $0 \leqslant m \leqslant 2$ and $n - m \geqslant n/2$, we have

$$|I_{11}| \leqslant \frac{tm\alpha_n^{-1} \log b}{n-m}\, \mathbb{E}\Psi_{\beta,t}(S) \leqslant Ctn^{-1}\alpha_n^{-1} \log b\, \mathbb{E}\Psi_{\beta,t}(S). \tag{B.33}$$

For I_{12}, note that

$$\begin{aligned}S - S_{i,j}^{(i)} &= (X_{i,\sigma(i)} + X_{\sigma^{-1}(j),j} - X_{\sigma^{-1}(j),\sigma(i)})\,\mathbf{1}(\sigma(i) \neq j) + X_{i,\sigma(i)}\,\mathbf{1}(\sigma(i) = j)\\ &\leqslant |X_{i,\sigma(i)}| + |X_{\sigma^{-1}(j),j}| + |X_{\sigma^{-1}(j),\sigma(i)}|.\end{aligned}$$

Moreover,

$$\begin{aligned}&|\Psi'_{\beta,t}(S_{i,j}^{(i)}) - \Psi'_{\beta,t}(S)|\\ &\quad\leqslant t|S - S_{i,j}^{(i)}|e^{t|S-S_{i,j}^{(i)}|}\Psi_{\beta,t}(S)\\ &\quad\leqslant 3t\Psi_{\beta,t}(S)\Big(|X_{i,\sigma(i)}|e^{3t|X_{i,\sigma(i)}|} + |X_{\sigma^{-1}(j),j}|e^{3t|X_{\sigma^{-1}(j),j}|} + |X_{\sigma^{-1}(j),\sigma(i)}|e^{3t|X_{\sigma^{-1}(j),\sigma(i)}|}\Big).\end{aligned}$$

Applying Lemma B.1 with $k = 1$ and $\zeta_{i,j} = |X_{i,j}|e^{3t|X_{i,j}|}$, and noting that $t \leqslant \alpha_n/64$, we have

$$\begin{aligned}\mathbb{E}|\Psi'_{\beta,t}(S_{i,j}^{(i)}) - \Psi'_{\beta,t}(S)| &\leqslant 36b^{1/8}t\,\mathbb{E}\Psi_{\beta,t}(S)\max_{i,j\in[n]}\mathbb{E}\{|X_{i,j}|e^{4t|X_{i,j}|}\}\\ &\leqslant 36\alpha_n^{-1}b^{1/8}t\,\mathbb{E}\Psi_{\beta,t}(S)\max_{i,j\in[n]}\mathbb{E}\{|\alpha_n X_{i,j}|e^{\alpha_n|X_{i,j}|/16}\}\\ &\leqslant C\alpha_n^{-1}b^{1/8}t\,\mathbb{E}\Psi_{\beta,t}(S)\max_{i,j\in[n]}\mathbb{E}\{e^{\alpha_n|X_{i,j}|/8}\}\\ &\leqslant C\alpha_n^{-1}b^{1/4}t\,\mathbb{E}\Psi_{\beta,t}(S).\end{aligned}\tag{B.34}$$

By (B.32) and (B.34), we obtain

$$|I_{12}| \leqslant C\alpha_n^{-2}b^{1/4}\log bt\,\mathbb{E}\Psi_{\beta,t}(S).\tag{B.35}$$

Combining (B.33) and (B.35) yields

$$I_1 \leqslant Cb^{1/2}(n^{-1}\alpha_n^{-1} + \alpha_n^{-2})t\,\mathbb{E}\Psi_{\beta,t}(S).\tag{B.36}$$

For I_2, observing that

$$\mathbb{E}\{X_{i,j}(\Psi'_{\beta,t}(S) - \Psi'_{\beta,t}(S^{(i)}))|\sigma(i) = j\} \leqslant t^2\,\mathbb{E}\{|X_{i,j}|^2e^{t|X_{i,j}|}\Psi_{\beta,t}(S)\},$$

we have

$$I_2 \leqslant t^2\,\mathbb{E}\{|X_{i,\sigma(i)}|^2e^{t|X_{i,\sigma(i)}|}\Psi_{\beta,t}(S)\}.$$

Applying Lemma B.1 with $k = 1$ and $\zeta_{i,j} = |X_{i,j}|^2e^{t|X_{i,j}|}$, we have

$$\begin{aligned}I_2 &\leqslant Ct^2b^{1/8}h(t)\max_{i,j\in[n]}\mathbb{E}\{|X_{i,j}|^2e^{2t|X_{i,j}|}\}\\ &\leqslant Ct^2b^{1/4}\alpha_n^{-2}\,\mathbb{E}\Psi_{\beta,t}(S).\end{aligned}\tag{B.37}$$

Combining (B.36) and (B.37) yields (B.27). $\square$

Recall that $\tau_{i,j}$ is the transposition of i and j. For $n \geqslant 4$, $m = 0, 1, 2$, and any permutation $\sigma \in \mathcal{S}_{n-m}$, define the transform

$$\mathcal{P}_{\boldsymbol{i},\boldsymbol{j}}\sigma = \begin{cases} \sigma & \text{if } \sigma(\boldsymbol{i}) = \sigma(\boldsymbol{j}), \\ \sigma \circ \tau_{\sigma^{-1}(j_1), i_1} & \text{if } \sigma(i_1) \neq j_1 \text{ and } \sigma(i_2) = \sigma(j_2), \\ \sigma \circ \tau_{\sigma^{-1}(j_2), i_2} & \text{if } \sigma(i_1) = j_1 \text{ and } \sigma(i_2) \neq \sigma(j_2), \\ \sigma \circ \tau_{\sigma^{-1}(j_2), i_1} \circ \tau_{\sigma^{-1}(j_1), i_2} \circ \tau_{i_1, i_2} & \text{if } \sigma(i_1) \neq j_1 \text{ and } \sigma(i_2) \neq \sigma(j_2). \end{cases} \tag{B.38}$$

The transform (B.38) was constructed by Goldstein (2005), and further applied by Chen and Fang (2015) to prove the Berry–Esseen bound for combinatorial central limit theorems. In the following lemmas, we use this transform to calculate the conditional expectation of functions of W given $\pi(i_1) = j_1$ and $\pi(i_2) = j_2$.

Lemma B.6. *Let S and σ be defined as in Lemma B.1. For any $\boldsymbol{i} = (i_1, i_2) \in [n-m]_2$, $\boldsymbol{j} = (j_1, j_2) \in [n-m]_2$ and $1 \leqslant p, q \leqslant 2$, we have*

$$\mathbb{E}\{|X_{\sigma^{-1}(j_q),\sigma(i_p)}|\Psi'_{\beta,t}(S)\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2)\} \leqslant Cbtn^{-1}\alpha_n^{-1}\,\mathbb{E}\Psi_{\beta,t}(S). \tag{B.39}$$

Proof of Lemma B.6. By the law of total expectation, for any $1 \leqslant p, q \leqslant 2$, we have

$$\begin{aligned} &\mathbb{E}\{|X_{\sigma^{-1}(j_q),\sigma(i_p)}|\Psi'_{\beta,t}(S)\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2)\} \\ &= \sum_{u,v\in[n-m]} \mathbb{E}\big\{|X_{u,v}|\Psi'_{\beta,t}(S)\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2) \bigm| \sigma(u) = j_q, \sigma(i_p) = v\big\} \\ &\qquad \times \mathbb{P}(\sigma(u) = j_q, \sigma(i_p) = v), \\ &\leqslant \sum_{u,v\in[n-m]} t\,\mathbb{E}\big\{|X_{u,v}|e^{t|X_{u,j_q}|}\Psi_{\beta,t}(S^{(u)})\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2) \bigm| \sigma(u) = j_q, \sigma(i_p) = v\big\} \\ &\qquad \times \mathbb{P}(\sigma(u) = j_q, \sigma(i_p) = v), \end{aligned} \tag{B.40}$$

where we used (5.27) in the last line. Since $(X_{u,v}, X_{u,j_q})$ is independent of $(S^{(u)}, \pi)$, we have

$$\begin{aligned} &\mathbb{E}\big\{|X_{u,v}|e^{t|X_{u,j_q}|}\Psi_{\beta,t}(S^{(u)})\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2) \bigm| \sigma(u) = j_q, \sigma(i_p) = v\big\} \\ &\qquad \leqslant \mathbb{E}\{|X_{u,v}|e^{t|X_{u,j_q}|}\}\,\mathbb{E}\big\{\Psi_{\beta,t}(S^{(u)})\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2) \bigm| \sigma(u) = j_q, \sigma(i_p) = v\big\} \\ &\qquad \leqslant Cb^{1/8}t\alpha_n^{-1}\,\mathbb{E}\big\{\Psi_{\beta,t}(S^{(u)})\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2) \bigm| \sigma(u) = j_q, \sigma(i_p) = v\big\}. \end{aligned} \tag{B.41}$$

By Hölder's inequality, we have

$$\begin{aligned} &\mathbb{E}\big\{\Psi_{\beta,t}(S^{(u)})\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2) \bigm| \sigma(u) = j_q, \sigma(i_p) = v\big\} \\ &\qquad \leqslant \mathbb{E}\big\{e^{t|X_{u,j_q}|}\Psi_{\beta,t}(S)\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2) \bigm| \sigma(u) = j_q, \sigma(i_p) = v\big\} \\ &\qquad \leqslant (\mathbb{E}\{e^{(1+\varepsilon)|X_{u,j_q}|/\varepsilon}\})^{\varepsilon/(1+\varepsilon)}\,\mathbb{E}\big\{\Psi_{\beta,t}^{1+\varepsilon}(S)\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2) \bigm| \sigma(u) = j_q, \sigma(i_p) = v\big\} \\ &\qquad \leqslant Cb^{1/64}\,\mathbb{E}\big\{\Psi_{\beta,t}(S)\,\mathbf{1}(\sigma(i_1) = j_1 \text{ or } \sigma(i_2) = j_2) \bigm| \sigma(u) = j_q, \sigma(i_p) = v\big\}, \end{aligned} \tag{B.42}$$

where the last inequality follows from (B.8) and (B.9). Substituting (B.41) and (B.42) into (B.40), we have

$$\begin{aligned}
&\mathbb{E}\{|X_{\sigma^{-1}(j_q),\sigma(i_p)}|\Psi'_{\beta,t}(S)\,\mathbf{1}(\sigma(i_1)=j_1 \text{ or } \sigma(i_2)=j_2)\} \\
&\qquad \leqslant Cb^{1/4}t\alpha_n^{-1}\,\mathbb{E}\{\Psi_{\beta,t}(S)\,\mathbf{1}(\sigma(i_1)=j_1 \text{ or } \sigma(i_2)=j_2)\}.
\end{aligned} \tag{B.43}$$

For the expectation term on the right hand side of (B.43),

$$\begin{aligned}
&\mathbb{E}\{\Psi_{\beta,t}(S)\,\mathbf{1}(\sigma(i_1)=j_1 \text{ or } \sigma(i_2)=j_2)\} \\
&\quad = \sum_{v_1,v_2\in[n-m]} \mathbb{E}\{\Psi_{\beta,t}(S)\,\mathbf{1}(\sigma(i_1)=j_1 \text{ or } \sigma(i_2)=j_2)|\sigma(i_1)=v_1,\sigma(i_2)=v_2\} \\
&\qquad\quad \times \mathbb{P}(\sigma(i_1)=v_1,\sigma(i_2)=v_2) \\
&\quad = \sum_{v_1,v_2\in[n-m]} \mathbf{1}(v_1=j_1 \text{ or } v_2=j_2)\,\mathbb{E}\{\Psi_{\beta,t}(S)|\sigma(i_1)=v_1,\sigma(i_2)=v_2\} \\
&\qquad\quad \times \mathbb{P}(\sigma(i_1)=v_1,\sigma(i_2)=v_2).
\end{aligned} \tag{B.44}$$

Let $\boldsymbol{i}=(i_1,i_2)$, $\boldsymbol{v}=(v_1,v_2)$, $\sigma_{\boldsymbol{i},\boldsymbol{v}}=\mathcal{P}_{\boldsymbol{i},\boldsymbol{v}}\sigma$ and $S_{\boldsymbol{i},\boldsymbol{v}}=\sum_{r=1}^{n-m}X_{r,\sigma_{\boldsymbol{i},\boldsymbol{v}}(r)}$. By (3.14) of Chen and Fang (2015) (see also Lemma 4.5 of Chen, Goldstein and Shao (2010)), we have

$$\mathbb{E}\{\Psi_{\beta,t}(S)|\sigma(i_1)=v_1,\sigma(i_2)=v_2\}=\mathbb{E}\{\Psi_{\beta,t}(S_{\boldsymbol{i},\boldsymbol{v}})\}. \tag{B.45}$$

Moreover, by the construction of $S_{\boldsymbol{i},\boldsymbol{v}}$, it follows that

$$\begin{aligned}
|S-S_{\boldsymbol{i},\boldsymbol{v}}| \leqslant{}& |X_{i_1,v_1}|+|X_{i_2,v_2}|+\sum_{i\in\{i_1,i_2\}}\sum_{v\in\{v_1,v_2\}}|X_{\sigma^{-1}(v),\sigma(i)}| \\
&+\sum_{i\in\{i_1,i_2\}}|X_{i,\sigma(i)}|+\sum_{v\in\{v_1,v_2\}}|X_{\sigma^{-1}(v),v}|.
\end{aligned} \tag{B.46}$$

By (A.20) and Hölder's inequality we have

$$\begin{aligned}
\mathbb{E}\{\Psi_{\beta,t}(S_{\boldsymbol{i},\boldsymbol{v}})\} &\leqslant \mathbb{E}\{e^{t|S-S_{\boldsymbol{i},\boldsymbol{v}}|}\Psi_{\beta,t}(S)\} \\
&\leqslant (\mathbb{E}\{e^{(1+\varepsilon)t|S-S_{\boldsymbol{i},\boldsymbol{v}}|/\varepsilon}\})^{\varepsilon/(1+\varepsilon)}\,\mathbb{E}\{\Psi_{\beta,t}(S)^{1+\varepsilon}\}.
\end{aligned} \tag{B.47}$$

By the similar argument to (B.8) and (B.9) again, we obtain

$$\mathbb{E}\{\Psi_{\beta,t}(S_{\boldsymbol{i},\boldsymbol{v}})\} \leqslant Cb^{1/2}\,\mathbb{E}\{\Psi_{\beta,t}(S)\}. \tag{B.48}$$

Combining (B.44), (B.45) and (B.47), we obtain

$$\begin{aligned}
&\mathbb{E}\{\Psi_{\beta,t}(S)\,\mathbf{1}(\sigma(i_1)=j_1 \text{ or } \sigma(i_2)=j_2)\} \\
&\quad \leqslant Cb^{1/2}\,\mathbb{E}\{\Psi_{\beta,t}(S)\}\sum_{v_1,v_2\in[n-m]}\mathbf{1}(v_1=j_1 \text{ or } v_2=j_2)\,\mathbb{P}(\sigma(i_1)=v_1,\sigma(i_2)=v_2) \\
&\quad \leqslant Cb^{1/2}n^{-1}\,\mathbb{E}\{\Psi_{\beta,t}(S)\}.
\end{aligned} \tag{B.49}$$

By (B.43) and (B.49), we complete the proof. □

The following lemma, whose proof is based on Lemmas B.5 and B.6, plays an important role in the proof of Lemma B.2.

Lemma B.7. *Let π, $\boldsymbol{X}$ and W be defined as in Theorem 4.1, and recall that $\mathcal{P}_{\boldsymbol{i},\boldsymbol{j}}$ is defined as in (B.38). For $\boldsymbol{i},\boldsymbol{j} \in [n]_2$, $\boldsymbol{i}' \in [n]_2^{(\boldsymbol{i})}$ and $\boldsymbol{j}' \in [n]_2^{(\boldsymbol{j})}$, let $\mathcal{I} = (\boldsymbol{i},\boldsymbol{i}')$, $\mathcal{J} = (\boldsymbol{j},\boldsymbol{j}')$, and*

$$\pi_{\boldsymbol{i},\boldsymbol{j}} = \mathcal{P}_{\boldsymbol{i},\boldsymbol{j}}\pi, \quad \pi_{\mathcal{I},\mathcal{J}} = \mathcal{P}_{\boldsymbol{i}',\boldsymbol{j}'}\pi_{\boldsymbol{i}',\boldsymbol{j}'}, \quad W_{(\mathcal{I},\mathcal{J})}^{(\mathcal{I})} = \sum_{i' \in [n]\setminus\{i_1,i_2,i_1',i_2'\}} X_{i',\pi_{\mathcal{I},\mathcal{J}}(i')}.$$

Then,

$$|\,\mathbb{E}\{\Psi_{\beta,t}(W_{\mathcal{I},\mathcal{J}}^{(\mathcal{I})})\} - h(t)| \leqslant Cb^2(n^{-1} + \alpha_n^{-2})(1+t^2)h(t). \tag{B.50}$$

Proof. Recall that $h(t) = \mathbb{E}\Psi_{\beta,t}(W)$. To bound the difference between $\mathbb{E}\Psi_{\beta,t}(W_{\mathcal{I},\mathcal{J}})$ and $\mathbb{E}\Psi_{\beta,t}(W)$, we consider the following three steps. In the first step, we construct an auxiliary random variable $S_{\boldsymbol{i}',\boldsymbol{j}'}^{(\boldsymbol{i},\boldsymbol{i}')}$ that is close to W and has the same distribution as $W_{\mathcal{I},\mathcal{J}}^{(\mathcal{I})}$. In the rest, we apply Taylor's expansion to calculate the difference of the expectations.

Step 1. Constructing $S_{\boldsymbol{i},\boldsymbol{j}}^{(\boldsymbol{i},\boldsymbol{i}')}$. Note that π is a random permutation chosen uniformly from $\mathcal{S}_n$, and it follows from Eq. (3.14) of Chen and Fang (2015) (see also Lemma 4.5 of Chen, Goldstein and Shao (2010)) that

$$\mathscr{L}(\pi_{\boldsymbol{i},\boldsymbol{j}}) = \mathscr{L}(\pi|\pi(\boldsymbol{i}) = \boldsymbol{j}). \tag{B.51}$$

Write $W_{\boldsymbol{i},\boldsymbol{j}} = \sum_{i' \in [n]} X_{i',\pi_{\boldsymbol{i},\boldsymbol{j}}(i')}$ and $W_{\boldsymbol{i},\boldsymbol{j}}^{(\boldsymbol{i})} = \sum_{i' \notin A(\boldsymbol{i})} X_{i',\pi_{\boldsymbol{i},\boldsymbol{j}}(i')}$, and it follows from (B.51) that $\mathscr{L}(W_{\boldsymbol{i},\boldsymbol{j}}) = \mathscr{L}(W|\pi(\boldsymbol{i}) = \boldsymbol{j})$ and $\mathscr{L}(W_{\boldsymbol{i},\boldsymbol{j}}^{(\boldsymbol{i})}) = \mathscr{L}(W^{(\boldsymbol{i})}|\pi(\boldsymbol{i}) = \boldsymbol{j})$. To calculate $\mathbb{E}\Psi_{\beta,t}(W_{\boldsymbol{i},\boldsymbol{j}})$, we introduce an auxiliary permutation σ as follows. Let σ be a uniform permutation from $[n] \setminus \{i_1,i_2\}$ to $[n] \setminus \{j_1,j_2\}$, independent of everything else, and let

$$S^{(\boldsymbol{i})} = \sum_{i' \notin A(\boldsymbol{i})} X_{i',\sigma(i')}. \tag{B.52}$$

It also follows from Lemma 4.5 of Chen, Goldstein and Shao (2010) that

$$\mathscr{L}(W_{\boldsymbol{i},\boldsymbol{j}}^{(\boldsymbol{i})}) = \mathscr{L}(S^{(\boldsymbol{i})}). \tag{B.53}$$

Moreover, noting that $\{i_1,i_2\} \cap \{i_1',i_2'\} = \varnothing$, using (B.51) twice implies that

$$\mathscr{L}(\pi_{\mathcal{I},\mathcal{J}}) = \mathscr{L}(\pi|\pi(\mathcal{I}) = \mathcal{J}). \tag{B.54}$$

Recalling (B.38), we define

$$\sigma_{\boldsymbol{i}',\boldsymbol{j}'} = \mathcal{P}_{\boldsymbol{i}',\boldsymbol{j}'}\sigma, \quad S_{\boldsymbol{i}',\boldsymbol{j}'}^{(\boldsymbol{i})} = \sum_{i' \in [n]\setminus\{\boldsymbol{i}\}} X_{i',\sigma_{\boldsymbol{i}',\boldsymbol{j}'}(i')}, \quad S_{\boldsymbol{i}',\boldsymbol{j}'}^{(\boldsymbol{i},\boldsymbol{i}')} = \sum_{i' \in [n]\setminus(\{\boldsymbol{i}\}\cup\{\boldsymbol{i}'\})} X_{i',\sigma_{\boldsymbol{i}',\boldsymbol{j}'}(i')}.$$

Then, it follows by definition that $\mathscr{L}(W^{(\mathcal{I})}_{\mathcal{I},\mathcal{J}}) = \mathscr{L}(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'})$.

Step 2. Bounding $|\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'}) - h(t)|$. We first bound $|\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'}) - \mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})})|$, and the bound of $|\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})}) - h(t)|$ can be obtained similarly. By Taylor's expansion,

$$\Psi_{\beta,t}(w) = \Psi_{\beta,t}(w_0) + (w - w_0)\Psi'_{\beta,t}(w_0) + \frac{1}{2}(w - w_0)^2\, \mathbb{E}\Psi''_{\beta,t}(w_0 + U(w - w_0)), \tag{B.55}$$

where U is a uniform random variable over the interval $[0, 1]$ independent of all others. Applying Taylor's expansion (B.55) with $w_0 = S^{(\boldsymbol{i})}$ and $w = S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'}$, we have

$$\begin{aligned}\mathbb{E}\{\Psi_{\beta,t}(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'})\} = \mathbb{E}\{\Psi_{\beta,t}(S^{(\boldsymbol{i})})\} + \mathbb{E}\big\{\big(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'} - S^{(\boldsymbol{i})}\big)\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\big\} \\ + \mathbb{E}\big\{\big(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'} - S^{(\boldsymbol{i})}\big)^2\Psi''_{\beta,t}\big(S^{(\boldsymbol{i})} + U(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'} - S^{(\boldsymbol{i})})\big)\big\}.\end{aligned} \tag{B.56}$$

Denote by $B_{\boldsymbol{i}',\boldsymbol{j}'}$ the event that $\{\sigma(i_1') = j_1' \text{ or } \sigma(i_2') = j_2'\}$ and denote by $B^c_{\boldsymbol{i},\boldsymbol{j}}$ the complement of $B_{\boldsymbol{i},\boldsymbol{j}}$. For the second term of (B.56), by the construction of $\mathcal{P}_{\boldsymbol{i}',\boldsymbol{j}'}$, we have

$$\begin{aligned}
&|\,\mathbb{E}\{(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'} - S^{(\boldsymbol{i})})\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\}| \\
&\quad\leqslant |\,\mathbb{E}\{X_{\sigma^{-1}(j_1'),\sigma(i_1')}\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\,\mathbf{1}(B_{\boldsymbol{i}',\boldsymbol{j}'})\}| \\
&\qquad + |\,\mathbb{E}\{X_{\sigma^{-1}(j_2'),\sigma(i_2')}\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\,\mathbf{1}(B_{\boldsymbol{i}',\boldsymbol{j}'})\}| \\
&\qquad + |\,\mathbb{E}\{X_{\sigma^{-1}(j_1'),\sigma(i_2')}\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\,\mathbf{1}(B^c_{\boldsymbol{i}',\boldsymbol{j}'})\}| \\
&\qquad + |\,\mathbb{E}\{X_{\sigma^{-1}(j_2'),\sigma(i_1')}\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\,\mathbf{1}(B^c_{\boldsymbol{i}',\boldsymbol{j}'})\}| \\
&\qquad + \sum_{i'\in\{i_1',i_2'\}} |\mathbb{E}\{X_{i',\sigma(i')}\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\}| + \sum_{j'\in\{j_1',j_2'\}} |\mathbb{E}\{X_{\sigma^{-1}(j'),j'}\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\}| \\
&\quad\leqslant \sum_{i'\in\{i_1',i_2'\}}\sum_{j'\in\{j_1',j_2'\}} \mathbb{E}\{|X_{\sigma^{-1}(j'),\sigma(i')}|\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\,\mathbf{1}(B_{\boldsymbol{i}',\boldsymbol{j}'})\} \\
&\qquad + |\,\mathbb{E}\{X_{\sigma^{-1}(j_1'),\sigma(i_2')}\Psi'_{\beta,t}(S^{(\boldsymbol{i})})| + |\,\mathbb{E}\{X_{\sigma^{-1}(j_2'),\sigma(i_1')}\Psi'_{\beta,t}(S^{(\boldsymbol{i})})| \\
&\qquad + \sum_{i'\in\{i_1',i_2'\}} |\mathbb{E}\{X_{i',\sigma(i')}\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\}| + \sum_{j'\in\{j_1',j_2'\}} |\mathbb{E}\{X_{\sigma^{-1}(j'),j'}\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\}|.
\end{aligned}$$

Applying Lemmas B.5 and B.6 with $S = S^{(\boldsymbol{i})}$ under a relabeling of indices, we obtain

$$|\,\mathbb{E}\{(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'} - S^{(\boldsymbol{i})})\Psi'_{\beta,t}(S^{(\boldsymbol{i})})\}| \leqslant C(n^{-1}\alpha_n^{-1} + \alpha_n^{-2})b(1 + t^2)\,\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})}). \tag{B.57}$$

For the third term of (B.56), we have

$$
\begin{aligned}
&\big|\mathbb{E}\big\{(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'}-S^{(\boldsymbol{i})})^2\Psi''_{\beta,t}\big(S^{(\boldsymbol{i})}+U(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'}-S^{(\boldsymbol{i})})\big)\big\}\big|\\
&\quad\leqslant t^2\,\mathbb{E}\big\{(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'}-S^{(\boldsymbol{i})})^2 e^{t|(S^{(\boldsymbol{i},\boldsymbol{i}')}_{\boldsymbol{i}',\boldsymbol{j}'}-S^{(\boldsymbol{i})})|}\Psi_{\beta,t}(S^{(\boldsymbol{i})})\big\}\\
&\quad\leqslant 24t^2\bigg\{\sum_{k\in\{i_1,i_2,i_1',i_2'\}}\ \sum_{l\in\{j_1,j_2,j_1',j_2'\}}\mathbb{E}\big\{|X_{\sigma^{-1}(l),\sigma(k)}|^2e^{24t|X_{\sigma^{-1}(l),\sigma(k)}|}\Psi_{\beta,t}(S^{(\boldsymbol{i})})\big\}\\
&\qquad\qquad+\sum_{i\in\{i_1,i_2,i_1',i_2'\}}\mathbb{E}\big\{X^2_{i,\sigma(i)}e^{24t|X_{i,\sigma(i)}|}\Psi_{\beta,t}(S^{(\boldsymbol{i})})\big\}\\
&\qquad\qquad+\sum_{j\in\{j_1,j_2,j_1',j_2'\}}\mathbb{E}\big\{X^2_{\sigma^{-1}(j),j}e^{24t|X_{i,\sigma(i)}|}\Psi_{\beta,t}(S^{(\boldsymbol{i})})\big\}\bigg\}.
\end{aligned}
\tag{B.58}
$$

Applying Lemma B.1 with $S=S^{(\boldsymbol{i})}$ and $\zeta_{i,j}=|X_{i,j}|^2e^{24t|X_{i,j}|}$ under a relabeling of indices, and recalling that $t\leqslant\alpha_n/64$, we obtain

$$
\begin{aligned}
&\mathbb{E}\big\{|X_{\sigma^{-1}(l),\sigma(k)}|^2e^{24t|X_{\sigma^{-1}(l),\sigma(k)}|}\Psi_{\beta,t}(S^{(\boldsymbol{i})})\big\}\\
&\quad\leqslant 4b^{1/8}h(t)\max_{u,v\in[n]}\mathbb{E}\big\{|X_{u,v}|^2e^{24t|X_{u,v}|+t(|X_{k,v}|+|X_{u,l}|)}\big\}\\
&\quad\leqslant Cb^{1/8}\alpha_n^{-2}\,\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})})\max_{u,v\in[n]}\mathbb{E}\big\{(1+|\alpha_nX_{u,v}|^2)e^{13\alpha_n|X_{u,v}|/32}\big\}\\
&\quad\leqslant Cb^{1/8}\alpha_n^{-2}\,\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})})\max_{u,v\in[n]}\mathbb{E}\big\{e^{\alpha_n|X_{u,v}|/2}\big\}\\
&\quad\leqslant Cb\alpha_n^{-2}\,\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})}),
\end{aligned}
$$

where we used (4.4) in the last line. Similarly, we obtain

$$
\begin{aligned}
\mathbb{E}\big\{X^2_{i,\pi(i)}e^{24t|X_{i,\pi(i)}|}\Psi_{\beta,t}(W)\big\}&\leqslant Cb\alpha_n^{-2}\,\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})}),\\
\mathbb{E}\big\{X^2_{\pi^{-1}(j),j}e^{24t|X_{\pi^{-1}(j),j}|}\Psi_{\beta,t}(W)\big\}&\leqslant Cb\alpha_n^{-2}\,\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})}).
\end{aligned}
$$

Hence, we have

$$
\text{R.H.S. of (B.58)}\leqslant Cb\alpha_n^{-2}(1+t^2)h(t).
$$

Together with (B.56) and (B.57), we have

$$
|\,\mathbb{E}\{\Psi_{\beta,t}(W^{(\mathcal{I})}_{\mathcal{I},\mathcal{J}})\}-\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})})|\leqslant Cb(n^{-1}+\alpha_n^{-2})(1+t^2)\,\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})}). \tag{B.59}
$$

A similar argument yields

$$
|\,\mathbb{E}\Psi_{\beta,t}(S^{(\boldsymbol{i})})-h(t)|\leqslant Cb(n^{-1}+\alpha_n^{-2})(1+t^2)h(t). \tag{B.60}
$$

Combining (B.59) and (B.60) and recalling that $t\leqslant\alpha_n/64$, we obtain

$$
|\,\mathbb{E}\{\Psi_{\beta,t}(W^{(\mathcal{I})}_{\mathcal{I},\mathcal{J}})\}-h(t)|\leqslant Cb^2(n^{-1}+\alpha_n^{-2})(1+t^2)h(t).
$$

This completes the proof. □

B.3 Proofs of Lemmas B.2 and B.3

Now, we apply Lemmas B.1 and B.7 to prove Lemma B.2.

Proof of Lemma B.2. Applying the Taylor expansion (B.55) yields

$$\begin{aligned}
&\mathbb{E}\{\xi_{\boldsymbol{i},\pi(\boldsymbol{i})}\xi_{\boldsymbol{i}',\pi(\boldsymbol{i}')}\Psi_{\beta,t}(W)\}\\
&\quad = \frac{1}{(n)_4}\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\{\xi_{\boldsymbol{i},\pi(\boldsymbol{i})}\xi_{\boldsymbol{i}',\pi(\boldsymbol{i}')}\Psi_{\beta,t}(W)|\pi(\mathcal{I})=\mathcal{J}\}\\
&\quad = Q_1+Q_2+Q_3,
\end{aligned} \tag{B.61}$$

where

$$\begin{aligned}
Q_1 &= \frac{1}{(n)_4}\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}\Psi_{\beta,t}(W^{(\mathcal{I})})|\pi(\mathcal{I})=\mathcal{J}\},\\
Q_2 &= \frac{1}{(n)_4}\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}V_{\mathcal{I},\mathcal{J}}\Psi'_{\beta,t}(W^{(\mathcal{I})})|\pi(\mathcal{I})=\mathcal{J}\},\\
Q_3 &= \frac{1}{(n)_4}\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}V_{\mathcal{I},\mathcal{J}}^2\Psi''_{\beta,t}(W^{(\mathcal{I})}+U(W-W^{(\mathcal{I})}))|\pi(\mathcal{I})=\mathcal{J}\},
\end{aligned}$$

and where $\mathcal{I}=(i_1,i_2,i'_1,i'_2)$, $\mathcal{J}=(j_1,j_2,j'_1,j'_2)$, $W^{(\mathcal{I})}=\sum_{i\in[n]\setminus A(\mathcal{I})}X_{i\pi(i)}$, $V_{\mathcal{I},\mathcal{J}}=X_{i_1,j_1}+X_{i_2,j_2}+X_{i'_1,j'_1}+X_{i'_2,j'_2}$ and U is a uniform random variable on $[0,1]$ independent of $\boldsymbol{X}$ and π.

For Q_1, as $(\xi_{\boldsymbol{i},\boldsymbol{j}},\xi_{\boldsymbol{i}',\boldsymbol{j}'})$ is conditionally independent of $W^{(\mathcal{I})}$ given $\pi(\mathcal{I})=\mathcal{J}$, and as $(\xi_{\boldsymbol{i},\boldsymbol{j}},\xi_{\boldsymbol{i}',\boldsymbol{j}'})$ is also independent of π, we have

$$\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}\Psi_{\beta,t}(W^{(\mathcal{I})})|\pi(\mathcal{I})=\mathcal{J}\}=\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}\}\,\mathbb{E}\{\Psi_{\beta,t}(W^{(\mathcal{I})})|\pi(\mathcal{I})=\mathcal{J}\}. \tag{B.62}$$

Let $W_{\mathcal{I},\mathcal{J}}$ be defined as in Lemma B.7. By (B.54), we have

$$\begin{aligned}
\mathbb{E}\{\Psi_{\beta,t}(W^{(\mathcal{I})})|\pi(\mathcal{I})=\mathcal{J}\} &= \mathbb{E}\{\Psi_{\beta,t}(W_{\mathcal{I},\mathcal{J}}^{(\mathcal{I})})\}\\
&= h(t)+|\mathbb{E}\{\Psi_{\beta,t}(W_{\mathcal{I},\mathcal{J}}^{(\mathcal{I})})\}-h(t)|.
\end{aligned}$$

Taking summations over $\boldsymbol{j}\in[n]_2,\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}$ on both sides of (B.62) and applying

Lemma B.7 gives

$$
\begin{aligned}
|Q_1| &\leqslant \frac{h(t)}{(n)_4}\Big|\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\big(\mathbb{E}\xi_{\boldsymbol{i},\boldsymbol{j}}\,\mathbb{E}\xi_{\boldsymbol{i}',\boldsymbol{j}'}\big)\Big| \\
&\quad + \frac{1}{(n)_4}\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\big(\mathbb{E}|\xi_{\boldsymbol{i},\boldsymbol{j}}|\,\mathbb{E}|\xi_{\boldsymbol{i}',\boldsymbol{j}'}||\mathbb{E}\{\Psi_{\beta,t}(W_{\mathcal{I},\mathcal{J}}^{(\mathcal{I})})\} - h(t)|\big) \\
&\leqslant \frac{h(t)}{(n)_4}\Big|\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\xi_{\boldsymbol{i},\boldsymbol{j}}\,\mathbb{E}\xi_{\boldsymbol{i}',\boldsymbol{j}'}\Big| \\
&\quad + Cb^2n^{-4}(n^{-1}+\alpha_n^{-2})(1+t^2)h(t)\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\big(\mathbb{E}|\xi_{\boldsymbol{i},\boldsymbol{j}}|\,\mathbb{E}|\xi_{\boldsymbol{i}',\boldsymbol{j}'}|\big).
\end{aligned}
\tag{B.63}
$$

For the first term of the right hand side of (B.63), since $\mathbb{E}\xi_{\boldsymbol{i}',\pi(\boldsymbol{i}')} = 0$, it follows that $\sum_{\boldsymbol{j}'\in[n]_2}\mathbb{E}\xi_{\boldsymbol{i}',\boldsymbol{j}'} = 0$. Therefore,

$$
\Big|\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\xi_{\boldsymbol{i}',\boldsymbol{j}'}\Big| = \Big|\sum_{\boldsymbol{j}'\in[n]_2\setminus[n]_2^{(\boldsymbol{j})}}\mathbb{E}\xi_{\boldsymbol{i}',\boldsymbol{j}'}\Big| \leqslant \sum_{\boldsymbol{j}'\in[n]_2}\mathbf{1}(E_{\boldsymbol{j},\boldsymbol{j}'})\,\mathbb{E}|\xi_{\boldsymbol{i}',\boldsymbol{j}'}|, \tag{B.64}
$$

where $E_{\boldsymbol{j},\boldsymbol{j}'} = \{A(\boldsymbol{j})\cap A(\boldsymbol{j}')\neq\varnothing\}$. Therefore, we have

$$
|Q_1| \leqslant Cb(1+t^2)n^{-4}h(t)\sum_{\boldsymbol{j},\boldsymbol{j}'\in[n]_2}\big(\mathbf{1}(E_{\boldsymbol{j},\boldsymbol{j}'}) + \alpha_n^{-2} + n^{-1}\big)\big(\mathbb{E}|\xi_{\boldsymbol{i},\boldsymbol{j}}|\,\mathbb{E}|\xi_{\boldsymbol{i}',\boldsymbol{j}'}|\big). \tag{B.65}
$$

Now we consider Q_2. Since $\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}V_{\mathcal{I},\mathcal{J}}$ is independent of $W^{(\mathcal{I})}$ conditional on $\pi(\mathcal{I}) = \mathcal{J}$, we have

$$
\begin{aligned}
Q_2 &= \frac{1}{(n)_4}\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}V_{\mathcal{I},\mathcal{J}}\}\,\mathbb{E}\{\Psi'_{\beta,t}(W^{(\mathcal{I})})|\pi(\mathcal{I}) = \mathcal{J}\} \\
&= \frac{1}{(n)_4}\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}V_{\mathcal{I},\mathcal{J}}\}\,\mathbb{E}\{\Psi'_{\beta,t}(W_{\mathcal{I},\mathcal{J}}^{(\mathcal{I})})\} \\
&= Q_{2,1} + Q_{2,2} + Q_{2,3},
\end{aligned}
\tag{B.66}
$$

where

$$
\begin{aligned}
Q_{2,1} &= \frac{1}{(n)_4}\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}V_{\boldsymbol{i},\boldsymbol{j}}\}\,\mathbb{E}\{\Psi'_{\beta,t}(W)\}, \\
Q_{2,2} &= \frac{1}{(n)_4}\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}V_{\boldsymbol{i}',\boldsymbol{j}'}\}\,\mathbb{E}\{\Psi'_{\beta,t}(W)\}, \\
Q_{2,3} &= \frac{1}{(n)_4}\sum_{\boldsymbol{j}\in[n]_2}\sum_{\boldsymbol{j}'\in[n]_2^{(\boldsymbol{j})}}\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}V_{\mathcal{I},\mathcal{J}}\}\,\mathbb{E}\{\Psi''_{\beta,t}(W + U(W_{\mathcal{I},\mathcal{J}}^{(\mathcal{I})} - W))(W_{\mathcal{I},\mathcal{J}}^{(\mathcal{I})} - W)\},
\end{aligned}
$$

and $V_{\boldsymbol{i},\boldsymbol{j}} = X_{i_1,j_1} + X_{i_2,j_2}$, $V_{\boldsymbol{i}',\boldsymbol{j}'} = X_{i'_1,j'_1} + X_{i'_2,j'_2}$. Similar to (B.64), we have

$$\begin{aligned} |Q_{2,1}| &\leqslant Cn^{-4}th(t) \sum_{\boldsymbol{j},\boldsymbol{j}' \in [n]_2} \mathbf{1}(E_{\boldsymbol{j},\boldsymbol{j}'}) |\,\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}} V_{\boldsymbol{i},\boldsymbol{j}}\}\,\mathbb{E}\{\xi_{\boldsymbol{i}',\boldsymbol{j}'}\}|, \\ |Q_{2,2}| &\leqslant Cn^{-4}th(t) \sum_{\boldsymbol{j},\boldsymbol{j}' \in [n]_2} \mathbf{1}(E_{\boldsymbol{j},\boldsymbol{j}'}) |\,\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\}\,\mathbb{E}\{\xi_{\boldsymbol{i}',\boldsymbol{j}'} V_{\boldsymbol{i}',\boldsymbol{j}'}\}|. \end{aligned} \tag{B.67}$$

Next, we consider $Q_{2,3}$. By (5.26) and using a similar argument for (B.58),

$$\begin{aligned} &|\mathbb{E}\{\Psi''_{\beta,t}(W + U(W^{(\mathcal{I})}_{\mathcal{I},\mathcal{J}} - W))(W^{(\mathcal{I})}_{\mathcal{I},\mathcal{J}} - W)\}| \\ &\quad \leqslant t^2\,\mathbb{E}\{\Psi_{\beta,t}(W) e^{t(W^{(\mathcal{I})}_{\mathcal{I},\mathcal{J}} - W)} |W^{(\mathcal{I})}_{\mathcal{I},\mathcal{J}} - W|\} \\ &\quad \leqslant Cb\alpha_n^{-1}(1+t^2)h(t). \end{aligned} \tag{B.68}$$

Therefore,

$$|Q_{2,3}| \leqslant Cb(1+t^2)h(t)n^{-4}\alpha_n^{-1}\Big(\sum_{\boldsymbol{j},\boldsymbol{j}' \in [n]_2} |\,\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}} V_{\boldsymbol{i},\boldsymbol{j}}\}\,\mathbb{E}\{\xi_{\boldsymbol{i}',\boldsymbol{j}'}\}| + \sum_{\boldsymbol{j},\boldsymbol{j}' \in [n]_2} |\,\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\}\,\mathbb{E}\{\xi_{\boldsymbol{i}',\boldsymbol{j}'} V_{\boldsymbol{i}',\boldsymbol{j}'}\}|\Big). \tag{B.69}$$

By (B.66), (B.67) and (B.69), we have

$$\begin{aligned} |Q_2| &\leqslant Cb(1+t^2)h(t)n^{-4} \sum_{\boldsymbol{j},\boldsymbol{j}' \in [n]_2} \big(\alpha_n^{-1} + \mathbf{1}(E_{\boldsymbol{j},\boldsymbol{j}'})\big)\big(|\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}} V_{\boldsymbol{i},\boldsymbol{j}}\}\,\mathbb{E}\{\xi_{\boldsymbol{i}',\boldsymbol{j}'}\}| + |\mathbb{E}\{\xi_{\boldsymbol{i}',\boldsymbol{j}'} V_{\boldsymbol{i}',\boldsymbol{j}'}\}\,\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\}|\big) \\ &\leqslant Cb(1+t^2)h(t)n^{-4} \sum_{\boldsymbol{j},\boldsymbol{j}' \in [n]_2} \big(\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}| T^2_{\boldsymbol{i},\boldsymbol{j}}\}\,\mathbb{E}\{|\xi_{\boldsymbol{i}',\boldsymbol{j}'}|\} + \mathbb{E}\{|\xi_{\boldsymbol{i}',\boldsymbol{j}'}| T^2_{\boldsymbol{i}',\boldsymbol{j}'}\}\,\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}|\}\big) \\ &\quad + Cb(1+t^2)h(t)n^{-4} \sum_{\boldsymbol{j},\boldsymbol{j}' \in [n]_2} \big(\alpha_n^{-2} + \mathbf{1}(E_{\boldsymbol{j},\boldsymbol{j}'})\big)\,\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}|\}\,\mathbb{E}\{|\xi_{\boldsymbol{i}',\boldsymbol{j}'}|\}, \end{aligned} \tag{B.70}$$

where the second inequality follows from Cauchy's inequality and the fact that $|V_{\boldsymbol{i},\boldsymbol{j}}| \leqslant T_{\boldsymbol{i},\boldsymbol{j}}$ for any $\boldsymbol{i},\boldsymbol{j}$. Finally, for Q_3, by (5.26) again, and noting that $V_{\mathcal{I},\mathcal{J}} = V_{\boldsymbol{i},\boldsymbol{j}} + V_{\boldsymbol{i}',\boldsymbol{j}'}$, we have

$$\begin{aligned} &|\mathbb{E}\{\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'} V^2_{\mathcal{I},\mathcal{J}} \Psi''_{\beta,t}(W^{(\mathcal{I})} + U(W - W^{(\mathcal{I})}))|\pi(\boldsymbol{i}) = \boldsymbol{j}; \pi(\boldsymbol{i}') = \boldsymbol{j}'\} \\ &\quad \leqslant t^2\,\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}| V^2_{\mathcal{I},\mathcal{J}} \Psi_{\beta,t}(W^{(\mathcal{I})}) e^{t|V_{\mathcal{I},\mathcal{J}}|} |\pi(\boldsymbol{i}) = \boldsymbol{j}; \pi(\boldsymbol{i}') = \boldsymbol{j}'\}| \\ &\quad = t^2\,\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}| V^2_{\mathcal{I},\mathcal{J}} e^{t|V_{\mathcal{I},\mathcal{J}}|}\}\,\mathbb{E}\{\Psi_{\beta,t}(W^{(\mathcal{I})})|\pi(\boldsymbol{i}) = \boldsymbol{j}; \pi(\boldsymbol{i}') = \boldsymbol{j}'\} \\ &\quad = t^2\,\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}| V^2_{\mathcal{I},\mathcal{J}} e^{t|V_{\mathcal{I},\mathcal{J}}|}\}\,\mathbb{E}\{\Psi_{\beta,t}(W^{(\mathcal{I})}_{\mathcal{I},\mathcal{J}})\} \\ &\quad \leqslant 2t^2\,\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}\xi_{\boldsymbol{i}',\boldsymbol{j}'}|(T^2_{\boldsymbol{i},\boldsymbol{j}} e^{2t|T_{\boldsymbol{i},\boldsymbol{j}}|} + T^2_{\boldsymbol{i}',\boldsymbol{j}'} e^{2t|T_{\boldsymbol{i}',\boldsymbol{j}'}|})\}\,\mathbb{E}\{\Psi_{\beta,t}(W) e^{t|W^{(\mathcal{I})}_{\mathcal{I},\mathcal{J}} - W|}\}, \end{aligned} \tag{B.71}$$

where we use $|V_{\boldsymbol{i},\boldsymbol{j}}| \leqslant T_{\boldsymbol{i},\boldsymbol{j}}$ for any $\boldsymbol{i},\boldsymbol{j}$. Then applying Lemma B.1 to the second expectation in the last line of (B.71) and similar to (B.58), we have

$$\mathbb{E}\{\Psi_{\beta,t}(W) e^{t|W^{(\mathcal{I})}_{\mathcal{I},\mathcal{J}} - W|}\} \leqslant Cbh(t). \tag{B.72}$$

By (B.71) and (B.72), we have

$$|Q_3| \leqslant Cbn^{-4}t^2h(t) \sum_{\boldsymbol{j},\boldsymbol{j}'\in[n]_2} \big(\mathbb{E}\{|\xi_{\boldsymbol{i},\boldsymbol{j}}|T_{\boldsymbol{i},\boldsymbol{j}}^2 e^{2t|T_{\boldsymbol{i},\boldsymbol{j}}|}\}\,\mathbb{E}|\xi_{\boldsymbol{i}',\boldsymbol{j}'}| + \mathbb{E}\{|\xi_{\boldsymbol{i}',\boldsymbol{j}'}|T_{\boldsymbol{i}',\boldsymbol{j}'}^2 e^{2t|T_{\boldsymbol{i}',\boldsymbol{j}'}|}\}\,\mathbb{E}|\xi_{\boldsymbol{i},\boldsymbol{j}}|\big). \tag{B.73}$$

Combining (B.61), (B.65), (B.70) and (B.73) yields the desired result. □

We finish our paper by proving Lemma B.3, which is based on Lemmas B.1 and B.2.

Proof of Lemma B.3. First, for $v = 0$, expanding the square term in (B.10) yields

$$\mathbb{E}\Big\{\Big(\sum_{\boldsymbol{i}\in[n]_2} \bar{g}_{\boldsymbol{i},\pi(\boldsymbol{i})}(u)\Big)^2 \Psi_{\beta,t}(W)\Big\} = H_1(u) + H_2(u), \tag{B.74}$$

where

$$H_1(u) = \frac{1}{16n^2} \sum_{\boldsymbol{i}\in[n]_2} \sum_{\boldsymbol{i}'\in[n]_2\setminus[n]_2^{(\boldsymbol{i})}} \mathbb{E}\{\bar{g}_{\boldsymbol{i},\pi(\boldsymbol{i})}(u)\bar{g}_{\boldsymbol{i}',\pi(\boldsymbol{i}')}(u)\Psi_{\beta,t}(W)\},$$

$$H_2(u) = \frac{1}{16n^2} \sum_{\boldsymbol{i}\in[n]_2} \sum_{\boldsymbol{i}'\in[n]_2^{(\boldsymbol{i})}} \mathbb{E}\{\bar{g}_{\boldsymbol{i},\pi(\boldsymbol{i})}(u)\bar{g}_{\boldsymbol{i}',\pi(\boldsymbol{i}')}(u)\Psi_{\beta,t}(W)\}.$$

For $H_1(u)$, by Young's inequality,

$$\begin{aligned} H_1(u) &\leqslant \frac{1}{16n^2} \sum_{\boldsymbol{i}\in[n]_2} \sum_{\boldsymbol{i}'\in[n]_2\setminus[n]_2^{(\boldsymbol{i})}} \mathbb{E}\big\{\big(g_{\boldsymbol{i},\pi(\boldsymbol{i})}^2(u) + g_{\boldsymbol{i}',\pi(\boldsymbol{i}')}^2(u)\big)\Psi_{\beta,t}(W)\big\} \\ &\leqslant Cn^{-1} \sum_{\boldsymbol{i}\in[n]_2} \mathbb{E}\{g_{\boldsymbol{i},\pi(\boldsymbol{i})}^2(u)\Psi_{\beta,t}(W)\}. \end{aligned} \tag{B.75}$$

Taking integration on both sides of (B.75) implies

$$\begin{aligned} &\int_{|u|\leqslant 1} e^{2t|u|}\,\mathbb{E}\{g_{\boldsymbol{i},\pi(\boldsymbol{i})}^2(u)\Psi_{\beta,t}(W)\}du \\ &\quad\leqslant Cn^{-1}\,\mathbb{E}\Big\{\Psi_{\beta,t}(W)\Big(|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|^3 e^{2t|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|} + \mathbb{E}|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|^3 e^{2t|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|}\Big)\Big\}. \end{aligned} \tag{B.76}$$

Applying Lemma B.1 with $k = 2, m = 0, \sigma = \pi$ and $\zeta_{\boldsymbol{i},\boldsymbol{j}} = |D_{\boldsymbol{i},\boldsymbol{j}}|^3 e^{2t|D_{\boldsymbol{i},\boldsymbol{j}}|}$, we have

$$\begin{aligned} &\mathbb{E}\{|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|^3 e^{2t|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|}\Psi_{\beta,t}(W)\} \\ &\quad\leqslant Cb\,\mathbb{E}\{\Psi_{\beta,t}(W)\} \max_{\boldsymbol{v}\in[n]_2} \mathbb{E}\{|D_{\boldsymbol{i},\boldsymbol{v}}|^3 e^{2t|D_{\boldsymbol{i},\boldsymbol{v}}|} e^{t|\sum_{r=1}^2 X_{i_r,v_r}|}\} \\ &\quad\leqslant Cb^{1/16}\alpha_n^{-3}\,\mathbb{E}\{\Psi_{\beta,t}(W)\} \max_{\boldsymbol{v}\in[n]_2} \mathbb{E}\{|\alpha_n D_{\boldsymbol{i},\boldsymbol{v}}|^3 e^{2t|D_{\boldsymbol{i},\boldsymbol{v}}|} e^{t|\sum_{r=1}^2 X_{i_r,v_r}|}\} \\ &\quad\leqslant Cb\alpha_n^{-3}\,\mathbb{E}\{\Psi_{\beta,t}(W)\}, \end{aligned} \tag{B.77}$$

and similarly,

$$\mathbb{E}|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|^3 e^{2t|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|} \leqslant Cb\alpha_n^{-3}. \tag{B.78}$$

By (B.75)–(B.78), we have

$$\int_{|u|\leqslant 1} e^{2t|u|} H_1(u)du \leqslant Cbn\alpha_n^{-3}\,\mathbb{E}\{\Psi_{\beta,t}(W)\}. \tag{B.79}$$

For $H_2(u)$, by Lemma B.2, with $\xi_{\boldsymbol{i},\boldsymbol{j}} = g_{\boldsymbol{i},\boldsymbol{j}}(u)$, we have any $\boldsymbol{i} \in [n]_2, \boldsymbol{i}' \in [n]_2^{(\boldsymbol{i})}$, and $\boldsymbol{j}, \boldsymbol{j}' \in [n]_2$,

$$\begin{aligned}
&\mathbb{E}\{\bar{g}_{\boldsymbol{i},\pi(\boldsymbol{i})}(u)\bar{g}_{\boldsymbol{i}',\pi(\boldsymbol{i}')}(u)\Psi_{\beta,t}(W)\}\\
&\quad\leqslant Cb(1+t^2)h(t)n^{-4}\sum_{\boldsymbol{j},\boldsymbol{j}'\in[n]_2}\mathbb{E}\{|\bar{g}_{\boldsymbol{i},\boldsymbol{j}}(u)|T_{\boldsymbol{i},\boldsymbol{j}}^2 e^{2t|T_{\boldsymbol{i},\boldsymbol{j}}|}\}\,\mathbb{E}\{|\bar{g}_{\boldsymbol{i}',\boldsymbol{j}'}(u)|\}\\
&\qquad + Cb(1+t^2)h(t)n^{-4}\sum_{\boldsymbol{j},\boldsymbol{j}'\in[n]_2}\mathbb{E}\{|\bar{g}_{\boldsymbol{i}',\boldsymbol{j}'}(u)|T_{\boldsymbol{i}',\boldsymbol{j}'}^2 e^{2t|T_{\boldsymbol{i}',\boldsymbol{j}'}|}\}\,\mathbb{E}\{|\bar{g}_{\boldsymbol{i},\boldsymbol{j}}(u)|\}\\
&\qquad + Cb(1+t^2)h(t)n^{-4}\sum_{\boldsymbol{j},\boldsymbol{j}'\in[n]_2}\big(\alpha_n^{-2}+n^{-1}+\mathbf{1}(E_{\boldsymbol{j},\boldsymbol{j}'})\big)\,\mathbb{E}\{|\bar{g}_{\boldsymbol{i},\boldsymbol{j}}(u)|\}\,\mathbb{E}\{|\bar{g}_{\boldsymbol{i}',\boldsymbol{j}'}(u)|\}\\
&\quad := H_{21}(u)+H_{22}(u)+H_{23}(u).
\end{aligned} \tag{B.80}$$

First we consider $H_{21}(u)$. Since $\mathbb{E}\{|\bar{g}_{\boldsymbol{i},\boldsymbol{j}}(u)|\} \leqslant 2\,\mathbb{E}|D_{\boldsymbol{i},\boldsymbol{j}}|$, by Fubini's theorem, we have for any $\boldsymbol{i} \in [n]_2, \boldsymbol{i}' \in [n]_2^{(\boldsymbol{i})}$, and $\boldsymbol{j}, \boldsymbol{j}' \in [n]_2$,

$$\begin{aligned}
&\left|\int_{|u|\leqslant 1} e^{2t|u|}\,\mathbb{E}\{|\bar{g}_{\boldsymbol{i},\boldsymbol{j}}(u)|T_{\boldsymbol{i},\boldsymbol{j}}^2 e^{2t|T_{\boldsymbol{i},\boldsymbol{j}}|}\}\,\mathbb{E}\{|\bar{g}_{\boldsymbol{i}',\boldsymbol{j}'}(u)|\}du\right|\\
&\quad\leqslant 2\left|\int_{|u|\leqslant 1} e^{2t|u|}\,\mathbb{E}\{|\bar{g}_{\boldsymbol{i},\boldsymbol{j}}(u)|T_{\boldsymbol{i},\boldsymbol{j}}^2 e^{2t|T_{\boldsymbol{i},\boldsymbol{j}}|}\}\,\mathbb{E}\{|D_{\boldsymbol{i}',\boldsymbol{j}'}|\}du\right|\\
&\quad\leqslant 2\,\mathbb{E}\Big\{\Big(|D_{\boldsymbol{i},\boldsymbol{j}}|^2 e^{2t|D_{\boldsymbol{i},\boldsymbol{j}}|}+\mathbb{E}\big\{|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|^2 e^{2t|D_{\boldsymbol{i},\pi(\boldsymbol{i})}|}\big\}\Big)T_{\boldsymbol{i},\boldsymbol{j}}^2 e^{2t|T_{\boldsymbol{i},\boldsymbol{j}}|}\Big\}\,\mathbb{E}\{|D_{\boldsymbol{i}',\boldsymbol{j}'}|\},
\end{aligned} \tag{B.81}$$

where use the same notation as in Lemma B.2. Then, by (4.4) and Young's inequality, we have

$$\begin{aligned}
&\mathbb{E}\Big\{|D_{\boldsymbol{i},\boldsymbol{j}}|^2 e^{2t|D_{\boldsymbol{i},\boldsymbol{j}}|}T_{\boldsymbol{i},\boldsymbol{j}}^2 e^{2t|T_{\boldsymbol{i},\boldsymbol{j}}|}\Big\}\,\mathbb{E}\{|D_{\boldsymbol{i}',\boldsymbol{j}'}|\}\\
&\quad\leqslant C\alpha_n^{-5}\,\mathbb{E}\big\{|\alpha_n^{-1}D_{\boldsymbol{i},\boldsymbol{j}}|^2 e^{2t|D_{\boldsymbol{i},\boldsymbol{j}}|}(\alpha_n^{-1}T_{\boldsymbol{i},\boldsymbol{j}})^2 e^{t|T_{\boldsymbol{i},\boldsymbol{j}}|}\big\}\,\mathbb{E}\{\alpha_n^{-1}|D_{\boldsymbol{i}',\boldsymbol{j}'}|\}\\
&\quad\leqslant Cb\alpha_n^{-5}.
\end{aligned} \tag{B.82}$$

Similar to (B.82),

$$\mathbb{E}\Big\{|D_{\boldsymbol{i},\boldsymbol{j}}|^2 e^{2t|D_{\boldsymbol{i},\boldsymbol{j}}|}\Big\}\,\mathbb{E}\Big\{T_{\boldsymbol{i},\boldsymbol{j}}^2 e^{2t|T_{\boldsymbol{i},\boldsymbol{j}}|}\Big\}\,\mathbb{E}\{|D_{\boldsymbol{i}',\boldsymbol{j}'}|\} \leqslant Cb\alpha_n^{-5}. \tag{B.83}$$

Using (B.81)–(B.83), we have

$$\left|\int_{|u|\leqslant 1} e^{2t|u|}\,\mathbb{E}\{|\bar{g}_{\boldsymbol{i},\boldsymbol{j}}(u)|T_{\boldsymbol{i},\boldsymbol{j}}^2 e^{2t|T_{\boldsymbol{i},\boldsymbol{j}}|}\}\,\mathbb{E}\{|\bar{g}_{\boldsymbol{i}',\boldsymbol{j}'}(u)|\}du\right| \leqslant Cb\alpha_n^{-5}. \tag{B.84}$$

Furthermore, we have,

$$\left|\int_{|u|\leqslant 1} e^{2t|u|} H_{21}(u)du\right| \leqslant Cb(1+t^2)\alpha_n^{-5}h(t) \tag{B.85}$$

Moreover, by the same argument, we have

$$\begin{aligned} \left|\int_{|u|\leqslant 1} e^{2t|u|} H_{22}(u)du\right| &\leqslant Cb(1+t^2)\alpha_n^{-5}h(t) \\ \left|\int_{|u|\leqslant 1} e^{2t|u|} H_{23}(u)du\right| &\leqslant Cb(1+t^2)(\alpha_n^{-5}+n\alpha_n^{-3})h(t) \end{aligned} \tag{B.86}$$

By (B.85) and (B.86), we have we have

$$\int_{|u|\leqslant 1} e^{2t|u|} H_2(u)du \leqslant Cb^2(n^2\alpha_n^{-5}+n\alpha_n^{-3})(1+t^2)\,\mathbb{E}\{\Psi_{\beta,t}(W)\}. \tag{B.87}$$

By (B.79) and (B.87), we complete the proof (B.10) for $v=0$. The inequality (B.10) for the case $v=1$ can be shown similarly. □

ACKNOWLEDGEMENTS

Liu Songhao was partially supported by Hong Kong RGC GRF 14302418. Zhuo-Song Zhang was supported by the Singapore Ministry of Education Academic Research Fund Tier 2 grant MOE2018-T2-2-076. Both authors would like to thank Qi-Man Shao and Xiao Fang for their helpful advises.

REFERENCES

P. Baldi and Y. Rinott (1989). On Normal Approximations of Distributions in Terms of Dependency Graphs. *Ann. Probab.* **17**, 1646–1650.

P. Baldi, Y. Rinott and C. Stein (1989). A normal approximation for the number of local maxima of a random function on a graph. In *Probability, Statistics, and Mathematics*, pages 59–81. Elsevier.

S. Chatterjee (2014). A short survey of Stein's method. *Contribution to the Proceedings of the ICM 2014*.

L. Chen, L. Goldstein and Q. Shao (2010). *Normal Approximation by Stein's Method*. Probability and Its Applications. Springer Berlin Heidelberg.

L. H. Y. Chen and Q.-M. Shao (2004). Normal approximation under local dependence. *Ann. Probab.* **32**.

L. H. Y. Chen, X. Fang and Q.-M. Shao (2013). From Stein identities to moderate deviations. *Ann. Probab.* **41**, 262–293.

L. H. Y. Chen, A. Röllin and A. Xia (2020). Palm theory, random measures and Stein couplings. *Available at* arXiv 2004.05026.

L. H. Chen and X. Fang (2015). On the error bound in a combinatorial central limit theorem. *Bernoulli* **21**, 335–359.

H. Cramér (1938). Sur un nouveau théorème-limite de la théorie des probabilités. In *Colloquium on Theory of Probability*, volume 736, pages 5–23.

A. Dembo and Y. Rinott (1996). Some Examples of Normal Approximations by Stein's Method. In D. Aldous and R. Pemantle, editors, *Random Discrete Structures*, The IMA Volumes in Mathematics and Its Applications, pages 25–44. Springer New York.

X. Fang (2019). Wasserstein-2 bounds in normal approximation under local dependence. *Electron. J. Probab.* **24**, 1–14.

A. N. Frolov (2019). On large deviations for combinatorial sums. *Available at* arXiv:1901.04244.

L. Goldstein (2005). Berry-esseen bounds for combinatorial central limit theorems and pattern occurrences, using zero and size biasing. *J. Appl. Probab.* **42**, 661–683.

L. Heinrich (1982). A method for the derivation of limit theorems for sums of m-dependent random variables. *Z. Wahrsch. Verw. Gebiete* **60**, 501–515.

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

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

Z. Hu, J. Robinson and Q. Wang (2007). Cramér-type large deviations for samples from a finite population. *Ann. Statist.* **35**, 673–696.

Y. V. Linnik (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*, volume 2, pages 289–306. Univ of California Press.

V. V. Petrov (1975). *Sums of independent random variables*, volume 82. Springer Science & Business Media.

M. Raic (2003). Normal approximation by Stein's method. In *Proceedings of the 7th Young Statisticians Meeting*, pages 71–97.

M. Raič (2019). A multivariate Berry–Esseen theorem with explicit constants. *Bernoulli* **25**, 2824–2853.

Y. Rinott (1994). On normal approximation rates for certain sums of dependent random variables. *J. Comput. Appl. Math.* **55**, 135–143.

Q.-M. Shao, M. Zhang and Z.-S. Zhang (2021). Cramér-type moderate deviation theorems for nonnormal approximation. *Ann. App. Probab.* **31**, 247—283.

C. Stein (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, Volume 2: Probability Theory*, pages 583–602, Berkeley, Calif. University of California Press.

Z.-S. Zhang (2019). Cramér-type moderate deviation of normal approximation for exchangeable pairs. *Available at* arXiv 1901.09526.