Journal of Theoretical Probability
https://doi.org/10.1007/s10959-024-01350-4

Another Look at Stein's Method for Studentized Nonlinear Statistics with an Application to U-Statistics

Dennis Leung[1] · Qi-Man Shao[2] · Liqian Zhang[1]

Received: 17 June 2023 / Revised: 25 May 2024 / Accepted: 4 June 2024

Abstract
We take another look at using Stein's method to establish uniform Berry–Esseen bounds for Studentized nonlinear statistics, highlighting variable censoring and an exponential randomized concentration inequality for a sum of censored variables as the essential tools to carry out the arguments involved. As an important application, we prove a uniform Berry–Esseen bound for Studentized U-statistics in a form that exhibits the dependence on the degree of the kernel.

Keywords Stein's method · Studentized nonlinear statistics · Variable censoring · Randomized concentration inequality · U-statistics · Self-normalized limit theory · Uniform Berry–Esseen bound

Mathematics Subject Classification (2020) 62E17

1 Introduction

We revisit the use of Stein's method to prove uniform Berry–Esseen (B–E) bounds for Studentized nonlinear statistics. Let $X_1, \ldots, X_n$ be independent random variables that serve as some raw data, and for each $i = 1, \ldots, n$, let

$$\xi_i \equiv g_{n,i}(X_i) \tag{1.1}$$

✉ Dennis Leung
dennis.leung@unimelb.edu.au

Qi-Man Shao
shaoqm@sustech.edu.cn

Liqian Zhang
liqian@student.unimelb.edu.au

1 School of Mathematics and Statistics, University of Melbourne, Parkville, Australia

2 Department of Statistics and Data Science, SICM, National Center for Applied Mathematics Shenzhen, Southern University of Science and Technology, Shenzhen, China

Published online: 15 June 2024

for a function $g_{n,i}(\cdot)$ that can also depend on i and n, such that

$$\mathbb{E}[\xi_i] = 0 \text{ for all } i \text{ and } \sum_{i=1}^{n} \mathbb{E}[\xi_i^2] = 1. \tag{1.2}$$

A *Studentized nonlinear statistic* is an asymptotically normal statistic that can be represented in the general form

$$T_{SN} \equiv \frac{W_n + D_{1n}}{(1 + D_{2n})^{1/2}}, \tag{1.3}$$

with $W_n \equiv \sum_{i=1}^n \xi_i$, where the "remainder" terms

$$D_{1n} = D_{1n}(X_1, \dots, X_n) \text{ and } D_{2n} = D_{2n}(X_1, \dots, X_n) \tag{1.4}$$

are some functions of the data, with the additional properties that

$$D_{1n}, D_{2n} \longrightarrow 0 \text{ in probability as } n \text{ tends to } \infty, \text{ and } D_{2n} \geq -1 \text{ almost surely.} \tag{1.5}$$

We adopt the convention that if $1 + D_{2n} = 0$, the value of T_{SN} is taken to be 0, $+\infty$ or $-\infty$ depending on the sign of $W_n + D_{1n}$. Such a statistic is a generalization of the classical Student's t-statistic [13], where the denominator $1+D_{2n}$ acts as a data-driven "self-normalizer" for the numerator $W_n + D_{1n}$.

Many statistics used in practice can be seen as examples of (1.3); hence, developing a general Berry–Esseen-type inequality for T_{SN} is relevant to many applications. The first such attempt based on Stein's method can be found in the semi-review article of Shao et al. [9], whose proof critically relies upon an exponential-type randomized concentration inequality first appearing in Shao [8]. However, while their methodology is sound, there are numerous gaps; most notably, Shao et al. [9] overlooked that the original exponential-type randomized concentration inequality of Shao [8] is developed for a sum of independent random variables with mean zero, which is not well suited for their proof wherein the truncated summands generally do not have mean 0. In fact, truncation itself is an insufficient device to carry the arguments involved, as will be explained in this article.

Our contributions are twofold. First, we put the methodology of Shao et al. [9] on solid footing; this, among other things, is accomplished by adopting variable *censoring* instead of truncation, as well as developing a modified randomized concentration inequality for a sum of censored variables, to rectify the gaps in their arguments. We also present a more user-friendly B–E bound for the statistic T_{SN} when the denominator remainder D_{2n} admits a certain standard form. Second, as an application to a prototypical example of Studentized nonlinear statistics, we establish a uniform B–E bound of the rate $1/\sqrt{n}$ for Studentized U-statistics whose dependence on the degree of the kernel is also explicit; all prior works in this vein only treat the simplest case with a kernel of degree 2. This bound is the most optimal known to date and serves to complete the literature in uniform B–E bounds for Studentized U-statistics.

Notation. $\Phi(\cdot)$ is the standard normal distribution function and $\bar{\Phi}(\cdot) = 1 - \Phi(\cdot)$. The indicator function is denoted by $I(\cdot)$. For $p \geq 1$, $\|Y\|_p \equiv (\mathbb{E}[|Y|^p])^{1/p}$ for a random variable Y. For any $a, b \in \mathbb{R}$, $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$. $C, C_1, C_2 \cdots$ .. denotes positive *absolute* constants that may differ in value from place to place, but does not depend on other quantities nor the distributions of the random variables. For two (possibly multivariate) random variables Y_1 and Y_2, "$Y_1 =_d Y_2$" means Y_1 and Y_2 have the same distribution.

2 General Berry–Esseen Bounds for Studentized Nonlinear Statistics

Let $\xi_1, \ldots, \xi_m$ be as in Sect. 1 that satisfy the assumptions in (1.2). For each $i = 1, \ldots, n$, define

$$\xi_{b,i} \equiv \xi_i I(|\xi_i| \leq 1) + I(\xi_i > 1) - I(\xi_i < -1), \tag{2.1}$$

an upper-and-lower *censored* version of ξ_i, and their sum

$$W_b = W_{b,n} \equiv \sum_{i=1}^{n} \xi_{b,i}. \tag{2.2}$$

Moreover, for each $i = 1, \ldots, n$, we define $W_b^{(i)} \equiv W_b - \xi_{b,i}$ and $W_n^{(i)} \equiv W_n - \xi_i$. We also let

$$\beta_2 \equiv \sum_{i=1}^{n} \mathbb{E}[\xi_i^2 I(|\xi_i| > 1)] \text{ and } \beta_3 \equiv \sum_{i=1}^{n} \mathbb{E}[\xi_i^3 I(|\xi_i| \leq 1)].$$

For any $x \in \mathbb{R}$,

$$f_x(w) \equiv \begin{cases} \sqrt{2\pi} e^{w^2/2} \Phi(w) \bar{\Phi}(x) & w \leq x \\ \sqrt{2\pi} e^{w^2/2} \Phi(x) \bar{\Phi}(w) & w > x \end{cases}; \tag{2.3}$$

is the solution to the Stein equation [12]

$$f_x'(w) - w f_x(w) = I(w \leq x) - \Phi(x). \tag{2.4}$$

Our first result is the following uniform Berry–Esseen bound for the Studentized nonlinear statistic in (1.3):

Theorem 2.1 (Uniform B–E bound for Studentized nonlinear statistics) *Let $X_1, \ldots, X_n$ be independent random variables. Consider the Studentized nonlinear statistic T_{SN} in* (1.3)*, constructed with the linear summands in* (1.1) *that satisfy the condition in*

(1.2)*, and the remainder terms in* (1.4) *that satisfy the condition in* (1.5)*. There exists a positive absolute constant* $C > 0$ *such that*

$$\begin{aligned}\sup_{x\in\mathbb{R}}\Big|P(T_{SN}\le x)-\Phi(x)\Big| &\le \sum_{j=1}^{2}P(|D_{jn}|>1/2)\\ &+C\Bigg\{\beta_2+\beta_3+\|\bar{D}_{1n}\|_2+\mathbb{E}\Big[(1+e^{W_b})\bar{D}_{2n}^2\Big]+\sup_{x\ge 0}\Big|x\mathbb{E}[\bar{D}_{2n}f_x(W_b)]\Big|\\ &+\sum_{j=1}^{2}\sum_{i=1}^{n}\Big(\mathbb{E}[\xi_{b,i}^2]\Big\|(1+e^{W_b^{(i)}})(\bar{D}_{jn}-\bar{D}_{jn}^{(i)})\Big\|_1\\ &+\Big\|\xi_{b,i}(1+e^{W_b^{(i)}/2})(\bar{D}_{jn}-\bar{D}_{jn}^{(i)})\Big\|_1\Big)\Bigg\},\end{aligned} \tag{2.5}$$

where for each $j\in\{1,2\}$ *and each* $i\in\{1,\ldots,n\}$*,*

- $D_{jn}^{(i)}\equiv D_{jn}^{(i)}(X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n)$ *is any function in the raw data except* X_i*;*
- $\bar{D}_{jn}$ *is a censored version of* D_{jn} *defined as*

$$\bar{D}_{jn}\equiv D_{jn}I\Big(|D_{jn}|\le\frac{1}{2}\Big)+\frac{1}{2}I\Big(D_{jn}>\frac{1}{2}\Big)-\frac{1}{2}I\Big(D_{jn}<-\frac{1}{2}\Big);$$

- $\bar{D}_{jn}^{(i)}$ *is a censored version of* $D_{jn}^{(i)}$ *defined as*

$$\bar{D}_{jn}^{(i)}\equiv D_{jn}^{(i)}I\Big(|D_{jn}^{(i)}|\le\frac{1}{2}\Big)+\frac{1}{2}I\Big(D_{jn}^{(i)}>\frac{1}{2}\Big)-\frac{1}{2}I\Big(D_{jn}^{(i)}<-\frac{1}{2}\Big).$$

In applications, $D_{1n}^{(i)}$ and $D_{2n}^{(i)}$ are typically taken as “leave-one-out” quantities constructed in almost identical manner as D_{1n} and D_{2n}, respectively, but without any terms involving the datum X_i, for instance, compared D_{1n} and $D_{1n}^{(i)}$ in (3.12) and (3.27) for the case of a U-statistic. The proof of Theorem 2.1 (“Appendix C”) bypasses the gaps in the proof of the original B–E bound for T_{SN} stated in [9, Theorem 3.1]. As a key step in their approach to proving Shao et al. [9, Theorem 3.1] based on Stein’s method, the exponential-type randomized concentration inequality developed in Shao [8, Theorem 2.7] is applied to control a probability of the type

$$P\left(\Delta_1\le\sum_{i=1}^{n}\xi_i I(|\xi_i|\le 1)\le\Delta_2\right),$$

where Δ_1 and Δ_2 are some context-dependent random quantities. Unfortunately, Shao et al. [9] overlooked that Shao [8, Theorem 2.7] was originally developed for a sum of mean-0 random variables, such as W_n, instead of the sum $\sum_{i=1}^{n}\xi_i I(|\xi_i|\le 1)$ figuring in the prior display, whose truncated summands do not have mean 0 in general. The latter needs to be addressed in some way to mend their arguments, which leads to the exponential randomized concentration inequality (Lemma B.1) developed in this work for the sum W_b in (2.2). Here, the censored summands $\xi_{b,i}$'s are considered instead so that the new inequality can still be proved in much the same way as Shao [8, Theorem 2.7]; replacing the truncated $\xi_i I(|\xi_i| \leq 1)$ with the censored $\xi_{b,i}$ is otherwise permissible, because only the boundedness of the summands is essential under the approach.

The B–E bound stated in Theorem 2.1 is in a primitive form. When applied to specific examples of T_{SN}, various terms in (2.5) have to be further estimated to render a more expressive bound. In that respect, the following apparent properties of censoring will become very useful:

Property 2.2 (Properties of variable censoring) *Let Y and Z be any two real value variables. The following facts hold:*

(i) Suppose, for some $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$ with $a \leq b$,

$$\bar{Y} \equiv aI(Y < a) + YI(a \leq Y \leq b) + bI(Y > b)$$

and

$$\bar{Z} \equiv aI(Z < a) + ZI(a \leq Z \leq b) + bI(Z > b).$$

Then it must be that $|\bar{Y} - \bar{Z}| \leq |Y - Z|$.

(ii) If Y is a non-negative random variable, then it must also be true that

$$YI(0 \leq Y \leq b) + bI(Y > b) \leq Y \text{ for any } b \in (0, \infty),$$

i.e., the upper-censored version of Y is always no larger than Y itself.

In applications of Theorem 2.1, that $\bar{D}_{1n}$ and $\bar{D}_{1n}^{(i)}$ are lower-and-upper censored by the same interval $[-1/2, 1/2]$ implies the bound

$$|\bar{D}_{1n} - \bar{D}_{1n}^{(i)}| \leq |D_{1n} - D_{1n}^{(i)}|, \tag{2.6}$$

by virtue of Property 2.2(i), as well as

$$|\bar{D}_1| \leq |D_1| \tag{2.7}$$

by virtue of Property 2.2(ii) because $|\bar{D}_1|$ is essentially the non-negative $|D_1|$ upper-censored at $1/2$. These bounds imply one can form the further norm estimates

$$\|(1 + e^{W_b^{(i)}})(\bar{D}_{1n} - \bar{D}_{1n}^{(i)})\|_1 \leq C\|D_{1n} - D_{1n}^{(i)}\|_2, \tag{2.8}$$

$$\|\xi_{b,i}(1 + e^{W_b^{(i)}/2})(\bar{D}_{1n} - \bar{D}_{1n}^{(i)})\|_1 \leq C\|\xi_i\|_2\|D_{1n} - D_{1n}^{(i)}\|_2 \tag{2.9}$$

and

$$\|\bar{D}_1\|_2 \leq \|D_1\|_2, \tag{2.10}$$

for the terms in (2.5) related to the numerator remainder D_1; see "Appendix D" for the simple arguments leading to these bounds. The right-hand sides of (2.8)–(2.10) are then amenable to direct second moment calculations to render more expressive terms. We also remark that if, instead, the truncated remainder terms

$$D_{jn} I\Big(|D_{jn}| \le \frac{1}{2}\Big) \text{ and } D_{jn}^{(i)} I\Big(|D_{jn}^{(i)}| \le \frac{1}{2}\Big), \text{ for } j = 1, 2, \tag{2.11}$$

are adopted as in Shao et al. [9, Theorem 3.1], a bound analogous to (2.6) does not hold in general; this also attests to censoring as a useful tool for developing nice B–E bounds under the current approach.

In comparison with the terms related to D_1, some of the terms related to D_2 in (2.5), such as

$$\sup_{x \ge 0} |x \mathbb{E}[\bar{D}_{2n} f_x(W_b)]| \text{ and } \mathbb{E}[e^{W_b} \bar{D}_{2n}^2],$$

are more obscure and have to be estimated on a case-by-case basis for specific examples of T_{SN}. However, in certain applications, the denominator remainder can be perceivably manipulated into the form

$$D_{2n} = \max\Big(-1, \quad \Pi_1 + \Pi_2 \Big) \tag{2.12}$$

lower censored at -1, where Π_1 is defined as

$$\Pi_1 \equiv \sum_{i=1}^n \Big(\xi_{b,i}^2 - \mathbb{E}[\xi_{b,i}^2] \Big), \tag{2.13}$$

and $\Pi_2 \equiv \Pi_2(X_1, \dots, X_n)$ is another data-dependent term. For instance, if a non-negative self-normalizer $1 + D_{2n}$ can be written as the intuitive form

$$1 + D_{2n} = \sum_{i=1}^n \xi_i^2 + E$$

for a data-dependent term $E \equiv E(X_1, \dots, X_n)$ of perceivably smaller order, then D_{2n} can be cast into the form (2.12) because $\sum_{i=1}^n (\mathbb{E}[\xi_{b,i}^2] + \mathbb{E}[(\xi_i^2 - 1) I(|\xi_i| > 1)]) = \sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1$ and one can take

$$\Pi_2 = E - \sum_{i=1}^n \mathbb{E}[(\xi_i^2 - 1) I(|\xi_i| > 1)] + \sum_{i=1}^n (\xi_i^2 - 1) I(|\xi_i| > 1).$$

We now present a more refined version of Theorem 2.1 for Studentized nonlinear statistics whose D_{2n} admits the form (2.12) under an absolute third-moment assumption on ξ_i; the proof is included in "Appendix D".

Theorem 2.3 (Uniform B–E bound for Studentized nonlinear statistics with the denominator remainder (2.12) under a third moment assumption) *Suppose all the conditions in Theorem 2.1 are met, and that* $\mathbb{E}[|\xi_i|^3] < \infty$ *for all* $1 \leq i \leq n$*. In addition, assume* D_{2n} *takes the specific form* (2.12) *with* Π_1 *defined in* (2.13) *and* $\Pi_2 \equiv \Pi_2(X_1, \ldots, X_n)$ *being a function in the raw data* $X_1, \ldots, X_n$*. For each* $i = 1, \ldots, n$*, let*

$$\Pi_2^{(i)} \equiv \Pi_2^{(i)}(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$$

be any function in the raw data except X_i*. Then*

$$\begin{aligned} \sup_{x \in \mathbb{R}} \left| P(T_{SN} \leq x) - \Phi(x) \right| \leq C \Bigg\{ & \sum_{i=1}^n \mathbb{E}[|\xi_i|^3] + \|D_{1n}\|_2 + \|\Pi_2\|_2 \\ & + \sum_{i=1}^n \|\xi_i\|_2 \|D_{1n} - D_{1n}^{(i)}\|_2 \\ & + \sum_{i=1}^n \|\xi_i\|_2 \|\Pi_2 - \Pi_2^{(i)}\|_2 \Bigg\}, \end{aligned} \tag{2.14}$$

where $D_{1n}^{(i)} \equiv D_{1n}^{(i)}(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$ *is as in Theorem 2.1.*

The $\| \cdot \|_2$ terms in (2.14) are now amenable to direct second moment calculations. Hence, if one can cast the denominator remainder into the form (2.12), Theorem 2.3 provides a user-friendly framework to establish B–E bounds for such instances of T_{SN}.

3 Uniform Berry–Esseen Bound for Studentized U-Statistics

We will apply Theorem 2.3 to establish a uniform B–E bound of the rate $1/\sqrt{n}$ for Studentized U-statistics of any degree; all prior works in this vein [2, 4, 5, 9, 15] only offer bounds for Studentized U-statistics of degree 2. We refer the reader to Shao et al. [9] and Jing et al. [5] for other examples of applications, including L-statistics and random sums and functions of nonlinear statistics.

Given independent and identically distributed random variables $X_1, \ldots, X_n$ taking value in a measure space $(\mathcal{X}, \Sigma_{\mathcal{X}})$, a U-statistic of degree $m \in \mathbb{N}_{\geq 1}$ takes the form

$$U_n = \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} h(X_{i_1}, \ldots, X_{i_m}),$$

where $h : \mathcal{X}^m \longrightarrow \mathbb{R}$ is a real-valued function symmetric in its m arguments, also known as the kernel of U_n; throughout, we will assume that

$$\mathbb{E}[h(X_1, \ldots, X_m)] = 0, \tag{3.1}$$

as well as

$$2m < n. \tag{3.2}$$

An important related function of $h(\cdot)$ is the *canonical function*

$$g(x) = \mathbb{E}[h(X_1, \ldots, X_{m-1}, x)] = \mathbb{E}[h(X_1, \ldots, X_m)|X_m = x],$$

which determines the first-order asymptotic behavior of the U-statistic. We will only consider *non-degenerate* U-statistics, which are U-statistics with the property that

$$\sigma_g^2 \equiv \mathrm{var}[g(X_1)] > 0.$$

It is well known that when $\mathbb{E}[h^2(X_1, \ldots, X_m)] < \infty$, $\frac{\sqrt{n}U_n}{m\sigma_g}$ converges weakly to the standard normal distribution as n tends to infinity [6, Theorem 4.2.1]; however, the limiting variance σ_g^2 is typically unknown and has to be substituted with a data-driven estimate. By constructing

$$q_i \equiv \frac{1}{\binom{n-1}{m-1}} \sum_{\substack{1 \le i_1 < \cdots < i_{m-1} \le n \\ i_l \ne i \text{ for } l=1,\ldots,m-1}} h(X_i, X_{i_1}, \ldots, X_{i_{m-1}}), \qquad i = 1, \ldots, n,$$

as natural proxies for $g(X_1), \ldots, g(X_n)$, the most common jackknife estimator for σ_g^2 is

$$s_n^2 \equiv \frac{n-1}{(n-m)^2} \sum_{i=1}^{n} (q_i - U_n)^2$$

[1], which gives rise to the *Studentized* U-statistic

$$T_n \equiv \frac{\sqrt{n}U_n}{ms_n}.$$

Without any loss of generality, we will assume that

$$\sigma_g^2 = 1, \tag{3.3}$$

as one can always replace $h(\cdot)$ and $g(\cdot)$, respectively, by $h(\cdot)/\sigma_g$ and $g(\cdot)/\sigma_g$ without changing the definition of T_n. Moreover, for s_n^* defined as

$$s_n^{*2} \equiv \frac{n-1}{(n-m)^2} \sum_{i=1}^{n} q_i^2,$$

we will also consider the statistic

$$T_n^* \equiv \frac{\sqrt{n}U_n}{ms_n^*}. \tag{3.4}$$

For any $x \in \mathbb{R}$, the event-equivalence relationship

$$\{T_n > x\} = \left\{ T_n^* > \frac{x}{\left(1 + \frac{m^2(n-1)x^2}{(n-m)^2}\right)^{1/2}} \right\} \tag{3.5}$$

is known in the literature; see [7, 10] for instance.

We now state a uniform Berry–Esseen bound for T_n and T_n^*. In the sequel, for any $k \in \{1, \ldots, n\}$ and $p \geq 1$, where no ambiguity arises, we may use $\mathbb{E}[\ell]$ and $\|\ell\|_p$ as the respective shorthands for $\mathbb{E}[\ell(X_1, \ldots, X_k)]$ and $\|\ell(X_1, \ldots, X_k)\|_p$, for a given function $\ell : \mathcal{X}^k \longrightarrow \mathbb{R}$ in k arguments. For example, we may use $\mathbb{E}[|h|^3]$ and $\|h\|_3$ to, respectively, denote the third absolute moment and 3-norm of $h(X_1, \ldots, X_m)$ with inserted data, and $\mathbb{E}[g^2] = \|g\|_2^2 = \sigma_g^2 = 1$ under (3.1) and (3.3).

Theorem 3.1 (Berry–Esseen bound for Studentized U-statistics) *Let $X_1, \ldots, X_n$ be independent and identically distributed random variables taking value in a measure space $(\mathcal{X}, \Sigma_{\mathcal{X}})$. Assume (3.1)–(3.3) and*

$$\mathbb{E}[|h|^3] < \infty, \tag{3.6}$$

then the following Berry–Esseen bound holds:

$$\sup_{x \in \mathbb{R}} |P(T_n \leq x) - \Phi(x)| \leq C \frac{\mathbb{E}[|g|^3] + m(\mathbb{E}[h^2] + \|g\|_3 \|h\|_3)}{\sqrt{n}} \tag{3.7}$$

for a positive absolute constant C; (3.7) also holds with T_n replaced by T_n^*.

To the best of our knowledge, this bound is the most optimal to date in the following sense: improving upon the preceding works of [2, 4, 15], for Studentized U-statistics of degree 2, under the same assumptions as Theorem 3.1, Jing et al. [5, Theorem 3.1] state a bound of the form

$$\sup_{x \in \mathbb{R}} |P(T_n \leq x) - \Phi(x)| \leq C \frac{\mathbb{E}[|h(X_1, X_2)|^3]}{\sqrt{n}}$$

for an absolute constant $C > 0$. In comparison, (3.7) is more optimal for $m = 2$ because all the moment quantities

$$\mathbb{E}[|g(X_1)|^3], \quad \mathbb{E}[|h(X_1, X_2)|^2] \text{ and } \|g(X_1)\|_3 \|h(X_1, X_2)\|_3$$

from (3.7) are all no larger than $\mathbb{E}[|h(X_1, X_2)|^3]$, given the standard moment properties for U-statistics; see (3.10).

In addition, we remark that the original B–E bound for Studentized U-statistics of degree 2 in Shao et al. [9, Theorem 4.2 & Remark 4.1] may have been falsely stated. Given (3.1)–(3.3), for an absolute constant $C > 0$, they stated a seemingly better bound (than (3.7)) of the form

$$\sup_{x\in\mathbb{R}} |P(T_n \leq x) - \Phi(x)| \leq C\frac{\|h(X_1, X_2)\|_2 + \mathbb{E}[|g(X_1)|^3]}{\sqrt{n}},$$

under the weaker assumption (than (3.6)) that $\|g(X_1)\|_3 \vee \|h(X_1, X_2)\|_2 < \infty$[1]. Unfortunately, the latter assumption is inadequate under the current approach based on Stein's method. The main issue is that Shao et al. [9] have ignored crucial calculations that require forming estimates of the rate $O(1/n)$ for an expectation of the type

$$\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_2(X_{i_1}, X_{i_2})\bar{h}_2(X_{j_1}, X_{j_2})],$$

where $1 \leq i_1 < i_2 \leq n$ and $1 \leq j_1 < j_2 \leq n$ are two pairs of sample indices, and $\bar{h}_2(\cdot)$ is the second-order canonical function in the Hoeffding's decomposition of U_n for $m = 2$; see (3.9). To do so, we believe one cannot do away with a third moment assumption on the kernel as in (3.6), where the anxious reader can skip ahead to Lemma E.1(iii) and (iv) for a preview of our estimates. Our proof of Theorem 3.1 rectifies such errors; moreover, it generalizes to a kernel of any degree m, for which the enumerative calculations needed are considerably more involved.

We first set the scene for establishing Theorem 3.1, by letting

$$\xi_i = \frac{g(X_i)}{\sqrt{n}} \tag{3.8}$$

and defining

$$\bar{h}_k(x_1 \ldots, x_k) = h_k(x_1 \ldots, x_k) - \sum_{i=1}^{k} g(x_i) \text{ for } k = 1, \ldots, m, \tag{3.9}$$

where

$$h_k(x_1, \ldots, x_k) = \mathbb{E}[h(X_1, \ldots, X_m)|X_1 = x_1, \ldots, X_k = x_k];$$

in particular, $g(x) = h_1(x)$ and $h(x_1, \ldots, x_m) = h_m(x_1, \ldots, x_m)$. An important property of the functions h_k is that

$$\mathbb{E}\big[|h_k|^p\big] \leq \mathbb{E}\big[|h_{k'}|^p\big] \text{ for any } p \geq 1 \text{ and } k \leq k', \tag{3.10}$$

[1] Actually, the bound claimed in Shao et al. [9, Remark 4.1] is $n^{-1/2}(\|h(X_1, X_2)\|_2 + \|g(X_1)\|_3^3)$, but the omission of the exponent 2 for $\|h(X_1, X_2)\|_2$ is itself a typo in that paper.

which is a consequence of Jensen's inequality:

$$\begin{aligned}\mathbb{E}\Big[|h_k(X_1,\dots,X_k)|^p\Big] &= \mathbb{E}\Big[|\mathbb{E}[h(X_1,\dots,X_m)|X_1,\dots,X_k]|^p\Big] \\ &= \mathbb{E}\Big[\Big|\mathbb{E}[h_{k'}(X_1,\dots,X_{k'})|X_1,\dots,X_k]\Big|^p\Big] \\ &\le \mathbb{E}\Big[\mathbb{E}\Big[|h_{k'}(X_1,\dots,X_{k'})|^p \mid X_1,\dots,X_k\Big]\Big] = \mathbb{E}\Big[|h_{k'}(X_1,\dots,X_{k'})|^p\Big].\end{aligned}$$

One can then write the part of (3.4) without the Studentizer s_n^* as

$$\frac{\sqrt{n}U_n}{m} = W_n + D_{1n}, \tag{3.11}$$

where $W_n \equiv \sum_{i=1}^n \xi_i$ and

$$D_{1n} \equiv \binom{n-1}{m-1}^{-1} \sum_{1\le i_1<\cdots<i_m\le n} \frac{\bar{h}_m(X_{i_1}, X_{i_2}, \dots, X_{i_m})}{\sqrt{n}}, \tag{3.12}$$

are considered as the numerator components under the framework of (1.3). To handle s_n^*, we shall first define

$$\Psi_{n,i} = \sum_{\substack{1\le i_1<\cdots<i_{m-1}\le n \\ i_l\ne i \text{ for } l=1,\dots,m-1}} \frac{\bar{h}_m(X_i, X_{i_1}, \dots, X_{i_{m-1}})}{\sqrt{n}}$$

and write

$$\begin{aligned} q_i &= \frac{1}{\binom{n-1}{m-1}} \sum_{\substack{1\le i_1<\cdots<i_{m-1}\le n \\ i_l\ne i \text{ for } l=1,\dots,m-1}} \left[g(X_i) + \sum_{l=1}^{m-1} g(X_{i_l}) + \bar{h}_m(X_i, X_{i_1}, \dots, X_{i_{m-1}}) \right] \\ &= \sqrt{n}\left[\left(\frac{n-m}{n-1}\right)\xi_i + \frac{m-1}{n-1}W_n\right] + \frac{\sqrt{n}}{\binom{n-1}{m-1}}\Psi_{n,i} \end{aligned}$$

for each i. By further letting

$$\Lambda_n^2 = \sum_{i=1}^n \Psi_{n,i}^2 \quad \text{and} \quad V_n^2 = \sum_{i=1}^n \xi_i^2,$$

the sum $\sum_{i=1}^{n} q_i^2$ can be consequently written as

$$\sum_{i=1}^{n} q_i^2 = n\left(\frac{n-m}{n-1}\right)^2 V_n^2 + \left[n^2\left(\frac{m-1}{n-1}\right)^2 + \frac{2n(n-m)(m-1)}{(n-1)^2}\right] W_n^2$$
$$+ \frac{n}{\binom{n-1}{m-1}^2}\Lambda_n^2 + 2n\left(\frac{n-m}{n-1}\right)\binom{n-1}{m-1}^{-1}\sum_{i=1}^{n}\xi_i\Psi_{n,i} + \frac{2n(m-1)}{(n-1)\binom{n-1}{m-1}}\sum_{i=1}^{n} W_n\Psi_{n,i},$$

which implies one can re-express s_n^{*2} as

$$s_n^{*2} = d_n^2(V_n^2 + \delta_{1n} + \delta_{2n}) \quad \text{for} \quad d_n^2 \equiv \frac{n}{n-1} \tag{3.13}$$

for

$$\delta_{1n} = \left[\frac{n(m-1)^2}{(n-m)^2} + \frac{2(m-1)}{(n-m)}\right] W_n^2 + \frac{(n-1)^2}{\binom{n-1}{m-1}^2 (n-m)^2}\Lambda_n^2$$
$$+ \frac{2(n-1)(m-1)}{(n-m)^2\binom{n-1}{m-1}}\sum_{i=1}^{n} W_n\Psi_{n,i} \tag{3.14}$$

and

$$\delta_{2n} \equiv \frac{2(n-1)}{(n-m)}\binom{n-1}{m-1}^{-1}\sum_{i=1}^{n}\xi_i\Psi_{n,i}.$$

We now present the proof of Theorem 3.1.

Proof of Theorem 3.1 It suffices to consider $x \geq 0$ since otherwise one can replace $h(\cdot)$ by $-h(\cdot)$. Defining

$$b_n = \frac{m^2(n-1)}{(n-m)^2} \text{ and } a_{n,x} = a_n(x) = \frac{1}{(1 + b_n x^2)^{1/2}},$$

we first simplify the problem using the bound

$$|\bar{\Phi}(xa_n(x)) - \bar{\Phi}(x)| \leq \min\left(\frac{m^2(n-1)x^3}{\sqrt{2\pi}(n-m)^2}, \frac{2}{\max(2, \sqrt{2\pi}xa_{n,x})}\right)\exp\left(\frac{-x^2 a_{n,x}^2}{2}\right), \tag{3.15}$$

which will be shown by a "bridging argument" borrowed from Jing et al. [5] at the end of this section. Then, by the triangular inequality, (3.5) and (3.15),

$$|P(T_n \leq x) - \Phi(x)|$$
$$= |P(T_n > x) - \bar{\Phi}(x)|$$

$$
\begin{aligned}
&\le |P(T_n^* > xa_n(x)) - \bar{\Phi}(xa_n(x))| + |\bar{\Phi}(xa_n(x)) - \bar{\Phi}(x)| \\
&\le |P(T_n^* > xa_n(x)) - \bar{\Phi}(xa_n(x))| \\
&\quad + \min\left(\frac{m^2(n-1)x^3}{\sqrt{2\pi}(n-m)^2}, \frac{2}{\max(2, \sqrt{2\pi}xa_{n,x})}\right)\exp\left(\frac{-x^2a_{n,x}^2}{2}\right) \\
&\le |P(T_n^* > xa_n(x)) - \bar{\Phi}(xa_n(x))| + C\frac{m^2}{\sqrt{n}},
\end{aligned}
\tag{3.16}
$$

where the last inequality in (3.16) holds as follows: For $0 \le x \le n^{1/6}$, the term

$$
\frac{m^2(n-1)x^3}{\sqrt{2\pi}(n-m)^2} \le \frac{m^2(n-1)\sqrt{n}}{\sqrt{2\pi}(n-m)^2} \le \frac{m^2(n-1)\sqrt{n}}{\sqrt{2\pi}(n-n/2)^2} \le \frac{2\sqrt{2}m^2}{\sqrt{\pi n}}.
$$

For $n^{1/6} < x < \infty$, since $xa_n(x)$ is strictly increasing in $x \in [0, \infty)$, we have that

$$
\begin{aligned}
\exp(-x^2a_{n,x}^2/2) &\le \exp(-n^{1/3}(1 + b_n n^{1/3})^{-1}/2) \\
&\le \exp\left(-\frac{n^{1/3}}{2}\left(1 + \frac{4m^2(n-1)n^{1/3}}{n^2}\right)^{-1}\right) \\
&\underbrace{\le}_{\text{by } (3.2)} \exp\left(-\frac{n^{1/3}}{2(1+(2m)^{4/3})}\right) \le \exp\left(-\frac{n^{1/3}}{8m^{4/3}}\right) \le \frac{Cm^2}{\sqrt{n}}.
\end{aligned}
$$

Since

$$
m = m\mathbb{E}[g^2] \le \mathbb{E}[h^2] \tag{3.17}
$$

by (3.3) and a classical U-statistic moment bound [6, Lemma 1.1.4], in light of (3.16), to prove (3.7) it suffices to show

$$
|P(T_n^* > x) - \bar{\Phi}(x)| \le C\frac{\mathbb{E}[|g|^3] + m(\mathbb{E}[h^2] + \|g\|_3\|h\|_3)}{\sqrt{n}}, \tag{3.18}
$$

as we have claimed to also hold in Theorem 3.1.

Note that since $2|W_n \sum_{i=1}^n \Psi_{n,i}| \le 2\sqrt{n}|W_n|\Lambda_n$ by Cauchy's inequality,

$$
\begin{aligned}
&\frac{2(n-1)(m-1)}{(n-m)^2\binom{n-1}{m-1}}\left|\sum_{i=1}^n W_n\Psi_{n,i}\right| \le 2\left\{\frac{\sqrt{n}(m-1)}{n-m}|W_n|\right\}\left\{\frac{(n-1)}{\binom{n-1}{m-1}(n-m)}\Lambda_n\right\} \\
&\le \frac{n(m-1)^2}{(n-m)^2}W_n^2 + \frac{(n-1)^2}{\binom{n-1}{m-1}^2(n-m)^2}\Lambda_n^2,
\end{aligned}
\tag{3.19}
$$

and hence we can deduce from (3.14) that

$$
\delta_{1n} \ge 0. \tag{3.20}
$$

With (3.11) and (3.13), one can then rewrite T_n^* as

$$T_n^* = \frac{W_n + D_{1n}}{d_n\sqrt{V_n^2 + \delta_{1n} + \delta_{2n}}}.$$

Now, consider the related statistic

$$\tilde{T}_n^* = \frac{W_n + D_{1n}}{\{\max(0, V_{n,b}^2 + \delta_{1n,b} + \delta_{2n,b})\}^{1/2}},$$

with suitably censored components in the denominator defined as

$$V_{n,b}^2 = \sum_{i=1}^n \xi_{b,i}^2, \quad \delta_{1n,b} = \min(\delta_{1n}, n^{-1/2}) \quad \text{and} \quad \delta_{2n,b} = \frac{2(n-1)}{(n-m)}\binom{n-1}{m-1}^{-1}\sum_{i=1}^n \xi_{b,i}\Psi_{n,i},$$

Note that T_n^* and $\tilde{T}_n^*$ can be related by the inclusions of events

$$\{\tilde{T}_n^* \le d_n x\}\backslash\mathcal{E} \subset \{T_n^* \le x\} \subset \{\tilde{T}_n^* \le d_n x\} \cup \mathcal{E},$$

where $\mathcal{E} \equiv \{\max_{1\le i\le n}|\xi_i| > 1\} \cup \{|\delta_{1n}| > n^{-1/2}\}$. The latter fact implies

$$\begin{aligned}
|P(T_n^* \le x) - \Phi(x)| &\le |P(\tilde{T}_n^* \le d_n x) - \Phi(x)| + P(\mathcal{E}) \\
&\le |P(\tilde{T}_n^* \le d_n x) - \Phi(x)| + \sum_{i=1}^n P(|\xi_i| > 1) + P(|\delta_{1n}| > n^{-1/2}) \\
&\le |P(\tilde{T}_n^* \le d_n x) - \Phi(x)| + \beta_2 + \sqrt{n}\mathbb{E}[|\delta_{1n}|] \\
&\le |P(\tilde{T}_n^* \le d_n x) - \Phi(x)| + \frac{\mathbb{E}[|g|^3]}{\sqrt{n}} + C\frac{m\mathbb{E}[h^2]}{\sqrt{n}}, \qquad (3.21)
\end{aligned}$$

with (3.21) coming from $\beta_2 \le \sum_{i=1}^n \mathbb{E}[\xi_i^3] = \mathbb{E}[|g|^3]/\sqrt{n}$, as well as combining (3.19) with (3.14) as:

$$\begin{aligned}
&\mathbb{E}[|\delta_{1n}|] \\
&\quad\le 2\left[\frac{m(m-1)(n-1)}{(n-m)^2}\right]\mathbb{E}[W_n^2] + \frac{2(n-1)^2}{\binom{n-1}{m-1}^2(n-m)^2}\mathbb{E}[\Lambda_n^2] \\
&\quad= 2\left[\frac{m(m-1)(n-1)}{(n-m)^2}\right] \\
&\qquad+ \frac{2(n-1)^2}{\binom{n-1}{m-1}^2(n-m)^2}\mathbb{E}\left[\left(\sum_{2\le i_1<\cdots<i_{m-1}\le n}\bar{h}_m(X_1, X_{i_1}, \ldots, X_{i_{m-1}})\right)^2\right] \\
&\quad\le \left(\frac{8m}{n} + \frac{4(n-1)^2(m-1)^2}{(n-m)^2(n-m+1)m}\right)\mathbb{E}[h^2],
\end{aligned}$$

where the last inequality follows from (3.17) and $2m < n$, as well as a standard U-statistic bound in Lemma E.1(ii).

In light of (3.21), to prove (3.18), it suffices to bound $|P(\tilde{T}_n^* \le d_n x) - \Phi(x)|$. To this end, we first define

$$\check{T}_n^{**} = \frac{W_n + D_{1n}}{\{\max(0, V_{n,b}^2 + \delta_{2n,b})\}^{1/2}}$$

and

$$\hat{T}_n^{**} = \frac{W_n + D_{1n}}{\{\max(0, V_{n,b}^2 + n^{-1/2} + \delta_{2n,b})\}^{1/2}},$$

which, by (3.20), have the property

$$P(\check{T}_n^{**} \le d_n x) \le P(\tilde{T}_n^* \le d_n x) \le P(\hat{T}_n^{**} \le d_n x) \tag{3.22}$$

Hence, to establish a bound for $|P(\tilde{T}_n^* \le d_n x) - \Phi(x)|$, our strategy is to prove the same bound for $|P(\check{T}_n^{**} \le d_n x) - \Phi(d_n x)|$ and $|P(\hat{T}_n^{**} \le d_n x) - \Phi(d_n x)|$, as well as using the bound

$$|\Phi(d_n x) - \Phi(x)| = \phi(x')(d_n x - x) \le C(d_n - 1) \le C n^{-1/2}, \tag{3.23}$$

coming from the mean value theorem, where $x' \in (x, d_n x)$ and $x\phi(x')$ is a bounded function in $x \in [0, \infty)$. To simplify notation, we will put $\check{T}_n^{**}$ and $\hat{T}_n^{**}$ under one umbrella and define their common placeholder

$$T_n^{**} = \frac{W_n + D_{1n}}{(1 + D_{2n})^{1/2}}, \tag{3.24}$$

where

$$D_{2n} \equiv \max(-1, V_{n,b}^2 - 1 + (n^{-1/2}|0) + \delta_{2n,b}) \tag{3.25}$$

and for $a, b \in \mathbb{R}$, $(a|b)$ represents either a or b; so T_n^{**} is either $\hat{T}_n^{**}$ or $\check{T}_n^{**}$.

Now, we cast (3.25) into the form (2.12) by taking $\Pi_1 = V_{n,b}^2 - \sum_{i=1}^n \mathbb{E}[\xi_{b,i}^2]$ and

$$\Pi_2 = \delta_{2n,b} + (n^{-1/2}|0) - \sum_{i=1}^n \mathbb{E}[(\xi_i^2 - 1)I(|\xi_i| > 1)] \tag{3.26}$$

In order to apply Theorem 2.3 to bound $|P(T_n^{**} \le d_n x) - \Phi(d_n x)|$, we will let $D_{1n}^{(i)}$ and $\Pi_2^{(i)}$, respectively, to be the "leave-one-out" versions of D_{1n} and Π_2 in (3.12) and

(3.26) that omit all the terms involving X_i, i.e,

$$D_{1n}^{(i)} \equiv \binom{n-1}{m-1}^{-1} \sum_{\substack{1\leq i_1<\cdots<i_m\leq n \\ i_l\neq i \text{ for } l=1,\ldots,m}} \frac{\bar{h}_m(X_{i_1}, X_{i_2}, \ldots, X_{i_m})}{\sqrt{n}} \tag{3.27}$$

and

$$\Pi_2^{(i)} \equiv \delta_{2n,b}^{(i)} + (n^{-1/2}|0) - \sum_{\substack{j=1 \\ j\neq i}}^{n} \mathbb{E}[(\xi_j^2 - 1)I(|\xi_j| > 1)] \tag{3.28}$$

for

$$\delta_{2n,b}^{(i)} \equiv \frac{2(n-1)}{\sqrt{n}(n-m)} \binom{n-1}{m-1}^{-1} \sum_{\substack{j=1 \\ j\neq i}}^{n} \xi_{b,j} \sum_{\substack{1\leq i_1<\cdots<i_{m-1}\leq n \\ i_l\neq j,i \text{ for } l=1,\ldots,m-1}} \bar{h}_m(X_j, X_{i_1}, \ldots, X_{i_{m-1}}).$$

We also need the following bounds:

Lemma 3.2 (Moment bounds related to D_{1n} in (3.12)) *Let D_{1n} and $D_{1n}^{(i)}$ be defined as in (3.12) and (3.27). Under the assumptions of Theorem 3.1, the following hold:*

$$\|D_{1n}\|_2 \leq \frac{(m-1)\|h\|_2}{\sqrt{m(n-m+1)}}, \tag{3.29}$$

and

$$\|D_{1n} - D_{1n}^{(i)}\|_2 \leq \frac{\sqrt{2}(m-1)\|h\|_2}{\sqrt{nm(n-m+1)}} \tag{3.30}$$

Proof of Lemma 3.2 This is known in the literature. Refer to Chen et al. [3, Lemma 10.1] for a proof. □

Lemma 3.3 (Moment bounds related to Π_2 in (3.26)) *Consider Π_2 and $\Pi_2^{(i)}$ defined in (3.26) and (3.28). Under the assumptions of Theorem 3.1, the following bounds hold:*

(i)

$$\|\Pi_2\|_2 \leq C\frac{\|g\|_3^3 + m\|g\|_3\|h\|_3}{\sqrt{n}},$$

and

(ii)

$$\|\Pi_2 - \Pi_2^{(i)}\|_2 \leq C\frac{m\|g\|_3\|h\|_3 + m^{1.5}\sqrt{\|h\|_2}}{n}$$

The proof of Lemma 3.3 is deferred to Appendix E. One can then apply Theorem 2.3, along with Lemmas 3.2 and 3.3 as well as (3.17), to give the bound

$$|P(T_n^{**} \le d_n x) - \Phi(d_n x)| \le C \frac{\mathbb{E}[|g|^3] + m(\|g\|_3 \|h\|_3 + \|h\|_2^{3/2})}{\sqrt{n}} \tag{3.31}$$

where we have used the fact that $\sigma_g^2 = 1$ in (3.3) and $\sigma_g \le \|h\|_2$ by virtue of (3.10). From (3.31), one can establish (3.18) with (3.21)–(3.24) and that $\|h\|_2^{3/2} \le \mathbb{E}[h^2]$.

It remains to finish the proof for (3.15): First, it can be seen that

$$0 < a_{n,x} \le 1. \tag{3.32}$$

Because of (3.32), we have

$$|xa_{n,x} - x| = \left| \frac{(a_{n,x}^2 - 1)x}{a_{n,x} + 1} \right| = \left| \left(\frac{b_n}{1 + b_n x^2} \right) \left(\frac{x^3}{a_{n,x} + 1} \right) \right| \le b_n x^3 = \frac{m^2(n-1)x^3}{(n-m)^2},$$

which implies, by the mean value theorem, that

$$|\Phi(xa_{n,x}) - \Phi(x)| \le \phi(xa_{n,x}) \frac{m^2(n-1)x^3}{(n-m)^2} = \frac{m^2(n-1)x^3}{\sqrt{2\pi}(n-m)^2} \exp\left(\frac{-x^2 a_{n,x}^2}{2} \right).$$

At the same time, we also have, by the well-known normal tail bound and (3.32),

$$|\Phi(xa_{n,x}) - \Phi(x)| \le \bar{\Phi}(xa_{n,x}) + \bar{\Phi}(x) \le \frac{2}{\max(2, \sqrt{2\pi} x a_{n,x})} \exp\left(\frac{-x^2 a_{n,x}^2}{2} \right).$$

□

Appendix A. Technical Lemmas

The first two lemmas below concern properties of the $\xi_{b,i}$'s and their sum W_b.

Lemma A.1 (Bound on expectation of $\xi_{b,i}$) *Let* $\xi_{b,i} = \xi_i I(|\xi_i| \le 1) + 1I(\xi_i > 1) - 1I(\xi_i < -1)$ *with* $\mathbb{E}[\xi_i] = 0$. *Then*

$$\left|\mathbb{E}[\xi_{b,i}]\right| \le \mathbb{E}[\xi_i^2 I(|\xi_i| > 1)] \le \mathbb{E}[\xi_i^2]$$

Proof of Lemma A.1

$$\begin{aligned}\left|\mathbb{E}[\xi_{b,i}]\right| &= |\mathbb{E}[(\xi_i - 1)I(\xi_i > 1) + (\xi_i + 1)I(\xi_i < -1)]| \\ &\le \mathbb{E}[(|\xi_i| - 1)I(|\xi_i| > 1)] \le \mathbb{E}[|\xi_i| I(|\xi_i| > 1)] \le \mathbb{E}[|\xi_i|^2 I(|\xi_i| > 1)] \le \mathbb{E}[\xi_i^2].\end{aligned}$$

□

Lemma A.2 (Bennett's inequality for a sum of censored random variables) *Let $\xi_1, \ldots, \xi_n$ be independent random variables with $\mathbb{E}[\xi_i] = 0$ for all $i = 1, \ldots, n$ and $\sum_{i=1}^n \mathbb{E}[\xi_i^2] \leq 1$, and define $\xi_{b,i} = \xi_i I(|\xi_i| \leq 1) + 1I(\xi_i > 1) - 1I(\xi_i < -1)$. For any $t > 0$ and $W_b = \sum_{i=1}^n \xi_{b,i}$, we have*

$$\mathbb{E}[e^{tW_b}] \leq \exp\left(e^{2t}/4 - 1/4 + t/2\right)$$

Proof of Lemma A.2 Note that, by Lemma A.1,

$$\mathbb{E}[e^{tW_b}] = \mathbb{E}[e^{t(W_b - \mathbb{E}[W_b])}]e^{t\mathbb{E}[W_b]} \leq \mathbb{E}[e^{t\sum_{i=1}^n(\xi_{b,i} - \mathbb{E}[\xi_{b,i}])}]e^t.$$

Moreover, by the standard Bennett's inequality [3, Lemma 8.1],

$$\mathbb{E}[e^{t\sum_{i=1}^n(\xi_{b,i} - \mathbb{E}[\xi_{b,i}])}] \leq \exp\left(4^{-1}(e^{2t} - 1 - 2t)\right).$$

□

The next lemmas concern properties of the solution to the Stein equation, f_x in (2.3). It is customary to *define* its derivative at x as $f_x'(x) \equiv xf_x(x) + \bar{\Phi}(x)$ so the Stein equation (2.4) is valid for all w. Moreover, we define

$$g_x(w) = (wf_x(w))' = f_x(w) + wf_x'(w). \tag{A.1}$$

Precisely,

$$f_x'(w) = \begin{cases} \left(\sqrt{2\pi}we^{w^2/2}\Phi(w) + 1\right)\bar{\Phi}(x) & \text{for} \quad w \leq x \\ \left(\sqrt{2\pi}we^{w^2/2}\bar{\Phi}(w) - 1\right)\Phi(x) & \text{for} \quad w > x \end{cases}; \tag{A.2}$$

$$g_x(w) = \begin{cases} \sqrt{2\pi}\bar{\Phi}(x)\left((1+w^2)e^{w^2/2}\Phi(w) + \frac{w}{\sqrt{2\pi}}\right) & \text{for} \quad w \leq x \\ \sqrt{2\pi}\Phi(x)\left((1+w^2)e^{w^2/2}\bar{\Phi}(w) - \frac{w}{\sqrt{2\pi}}\right) & \text{for} \quad w > x \end{cases}. \tag{A.3}$$

Lemma A.3 (Uniform bounds for f_x) *For f_x and f_x', the following bounds are true:*

$$|f_x'(w)| \leq 1, \qquad 0 < f_x(w) \leq 0.63 \quad \textit{and} \quad 0 \leq g_x(w) \quad \textit{for all} \quad w, x \in \mathbb{R}.$$

Moreover, for any $x \in [0, 1]$, $g_x(w) \leq 2.3$ for all $w \in \mathbb{R}$.

Lemma A.4 (Nonuniform bounds for f_x when $x \geq 1$) *For $x \geq 1$, the following are true:*

$$f_x(w) \leq \begin{cases} 1.7e^{-x} & \text{for} \quad w \leq x - 1 \\ 1/x & \text{for} \quad x - 1 < w \leq x \\ 1/w & \text{for} \quad x < w \end{cases} \tag{A.4}$$

and

$$|f_x'(w)| \le \begin{cases} e^{1/2-x} & \text{for} \quad w \le x-1 \\ 1 & \text{for} \quad x-1 < w \le x \\ (1+x^2)^{-1} & \text{for} \quad w > x \end{cases}. \tag{A.5}$$

Moreover, $g_x(w) \ge 0$ *for all* $w \in \mathbb{R}$,

$$g_x(w) \le \begin{cases} 1.6\,\bar{\Phi}(x) & \text{for} \quad w \le 0 \\ 1/w & \text{for} \quad w > x \end{cases}, \tag{A.6}$$

and $g_x(w)$ *is increasing for* $0 \le w \le x$ *with*

$$g_x(x-1) \le xe^{1/2-x} \quad \textit{and} \quad g_x(x) \le x+2.$$

We remark that the nonuniform bounds in Lemma A.4 refine the ones previously collected in Shao et al. [9, Lemma 5.3]; as an aside, a property analogous to (A.5) has been incorrectly stated in Shao et al. [9] without the absolute signs $|\cdot|$ around $f_x'(w)$. The proofs below repeatedly use the well-known inequality [3, p.16 & 38]

$$\frac{we^{-w^2/2}}{(1+w^2)\sqrt{2\pi}} \le \bar{\Phi}(w) \le \min\left(\frac{1}{2}, \frac{1}{w\sqrt{2\pi}}\right) e^{-w^2/2} \text{ for } w > 0. \tag{A.7}$$

Proof of Lemma A.3 The bounds for f_x and f_x', and that $g_x(w) \ge 0$, are well known; see Chen et al. [3, Lemma 2.3]. We will show that g_x in (A.3) is less than 2.3 when $x \in [0, 1]$. Using (A.7), for $w > x$, we have

$$\begin{aligned} g_x(w) &\le \sqrt{2\pi}\,\Phi(x)\left((1+w^2)e^{w^2/2}\bar{\Phi}(w) - \frac{w}{\sqrt{2\pi}}\right) \\ &\le \sqrt{2\pi}\,\Phi(x)\left(\frac{1}{2} + \frac{w}{\sqrt{2\pi}} - \frac{w}{\sqrt{2\pi}}\right) \le \frac{\sqrt{2\pi}\,\Phi(x)}{2} \le 2. \end{aligned}$$

For $0 \le w \le x$,

$$\begin{aligned} g_x(w) &= \sqrt{2\pi}\,\bar{\Phi}(x)\left((1+w^2)e^{w^2/2}\Phi(w) + \frac{w}{\sqrt{2\pi}}\right) \\ &\le \sqrt{2\pi}\,\bar{\Phi}(x)\left((1+x^2)e^{x^2/2}\Phi(x) + \frac{x}{\sqrt{2\pi}}\right) \\ &\le \left\{\left(\frac{\sqrt{2\pi}}{2} + x\right)\Phi(x) + \frac{e^{-x^2/2}}{\sqrt{2\pi}}\right\} \vee \left(\sqrt{2\pi}\,\bar{\Phi}(0)\cdot\Phi(0)\right) \\ &\le \left\{\left(\frac{\sqrt{2\pi}}{2} + 1\right)\Phi(1) + \frac{1}{\sqrt{2\pi}}\right\} \vee 0.63 \le 2.3. \end{aligned}$$

For $w < 0$,

$$\sqrt{2\pi}\bar{\Phi}(x)\left((1+w^2)e^{w^2/2}\Phi(w)+\frac{w}{\sqrt{2\pi}}\right) \leq \sqrt{2\pi}\bar{\Phi}(x)\left(\frac{1}{2}+\frac{|w|}{\sqrt{2\pi}}-\frac{|w|}{\sqrt{2\pi}}\right) \leq 1.26.$$

□

Proof of Lemma A.4 **Proof of** (A.4) **by investigating** (2.3): When $w \leq 0$, by (A.7), $x^2 \geq 2x-1$, and the symmetry of $\phi(\cdot)$, we have that

$$f_x(w) \leq e^{w^2/2}\Phi(w)\frac{e^{-x^2/2}}{x} \leq \frac{e^{-x^2/2}}{2x} \leq \frac{e^{-x+1/2}}{2} \leq 0.9e^{-x}.$$

When $0 < w \leq x-1$, by (A.7), we have

$$f_x(w) \leq e^{(x-1)^2/2}\Phi(w)\frac{e^{-x^2/2}}{x} = \Phi(w)e^{-x+1/2} \leq 1.7e^{-x}.$$

When $x-1 < w \leq x$, by (A.7), we have

$$f_x(w) \leq \frac{e^{(w^2-x^2)/2}\Phi(w)}{x} \leq \frac{1}{x}.$$

When $w > x$, by (A.7), we have

$$f_x(w) \leq \frac{\Phi(x)}{w} \leq \frac{1}{w}.$$

Proof of (A.5) **by investigating** (A.2): When $w \leq 0$, by the symmetry of $\phi(\cdot)$, (A.7) and $x^2 \geq 2x-1$, we have

$$0 = 0\cdot\bar{\Phi}(x) \leq f_x'(w) \leq \left(\frac{1}{1+w^2}\right)\frac{e^{-x+1/2}}{\sqrt{2\pi}} \leq 0.4e^{1/2-x}.$$

When $0 < w \leq x-1$, by (A.7) and $x^2 \geq 2x-1$,

$$0 \leq f_x'(w) \leq \left(\sqrt{2\pi}(x-1)e^{\frac{(x-1)^2}{2}}+1\right)\frac{e^{-x^2/2}}{x\sqrt{2\pi}} \leq \left(\frac{x-1}{x}+\frac{1}{x\sqrt{2\pi}}\right)e^{1/2-x} \leq e^{1/2-x},$$

as $\left(\frac{x-1}{x}+\frac{1}{x\sqrt{2\pi}}\right)$ is increasing as a function in x on $[1,\infty)$. When $x-1 < w \leq x$, by (A.7) we have

$$0 \leq f_x'(w) = \Phi(w)\underbrace{\sqrt{2\pi}we^{w^2/2}\bar{\Phi}(x)}_{\leq 1}+\bar{\Phi}(x) \leq 1.$$

When $w > x$, since $\sqrt{2\pi} w e^{w^2/2}\bar{\Phi}(w) \le 1$ by (A.7), hence $f_x'(w) \le 0$. Moreover, by applying (A.7) again, we have

$$\frac{-1}{x^2+1} \le \left(\frac{w^2}{w^2+1} - 1\right)\Phi(x) \le f_x'(w) \le 0.$$

Proof of (A.6) **by investigating** (A.3): When $w < 0$, by the symmetry of ϕ and (A.7),

$$0 = \sqrt{2\pi}\bar{\Phi}(x)\cdot 0 \le g_x(w) \le \left(\min\left(\frac{1+w^2}{|w|}, \frac{(1+w^2)\sqrt{2\pi}}{2}\right) + w\right)\bar{\Phi}(x) \le 1.6\bar{\Phi}(x),$$

where the last inequality uses the facts that $\frac{(1+w^2)\sqrt{2\pi}}{2} + w \le 1.6$ for $w \in [-1, 0]$ and that $\frac{1+w^2}{|w|} + w = 1/|w|^2 \le 1$ for $w < -1$. When $w > x$, by (A.7),

$$0 \le \sqrt{2\pi}\Phi(x)\cdot 0 \le g_x(w) \le \Phi(x)\left(\frac{1+w^2}{w} - w\right) = \frac{\Phi(x)}{w} \le 1/w.$$

When $0 \le w \le x$, it is easy to see that $g_x(w)$ is non-negative and increasing in w. Moreover, from (A.7) and $x^2 \ge 2x - 1$,

$$\begin{aligned}
g_x(x-1) &= \sqrt{2\pi}\bar{\Phi}(x)\left((2+x^2-2x)e^{x^2/2-x+1/2}\Phi(x-1) + \frac{x-1}{\sqrt{2\pi}}\right)\\
&\le \frac{(2+x^2-2x)}{x}e^{1/2-x}\Phi(x-1) + \frac{x-1}{x\sqrt{2\pi}}e^{-x^2/2}\\
&\le \frac{(4+2x^2-4x)}{2x}e^{1/2-x} + \frac{x-1}{2x}e^{1/2-x}\\
&\le \left(x - \frac{3}{2} + \frac{3}{2x}\right)e^{1/2-x} \le xe^{1/2-x}.
\end{aligned}$$

Lastly, by (A.7), it is easy to see that

$$\begin{aligned}
g_x(x) &= \sqrt{2\pi}\bar{\Phi}(x)\left((1+x^2)e^{x^2/2}\Phi(x) + \frac{x}{\sqrt{2\pi}}\right)\\
&\le \frac{1+x^2}{x}\Phi(x) + \frac{e^{-x^2/2}}{\sqrt{2\pi}} \le \left(\frac{1}{x} + x\right) + \frac{1}{2} \le x + 2
\end{aligned}$$

□

Lemma A.5 (Bound on expectation of $f_x'(W_b^{(i)} + t)$) *Let $x \ge 1$, $t \in \mathbb{R}$, and $W_b^{(i)}$ be as defined in Sect. 1 under the assumptions (1.2). Then there exists an absolute constant $C > 0$ such that*

$$\left|\mathbb{E}[f_x'(W_b^{(i)} + t)]\right| \le C(e^{-x} + e^{-x+t}).$$

Proof of Lemma A.5 From (A.5) in Lemma A.4, we have

$$
\begin{aligned}
|\mathbb{E}[f_x'(W_b^{(i)}+t)]| &\le e^{1/2-x} + \mathbb{E}[I(W_b^{(i)}+t > x-1)] \\
&\le e^{1/2-x} + e^{1-x+t}\mathbb{E}[e^{W_b^{(i)}}]
\end{aligned}
$$

and then apply the Bennett inequality in Lemma A.2. □

Appendix B. Exponential Randomized Concentration Inequality for a Sum of Censored Variables

Lemma B.1 (Exponential randomized concentration inequality for a sum of censored random variables) *Let $\xi_1, \ldots, \xi_n$ be independent random variables with mean zero and finite second moments, and for each $i = 1, \ldots, n$, define*

$$
\xi_{b,i} = \xi_i I(|\xi_i| \le 1) + 1I(\xi_i > 1) - 1I(\xi_i < -1),
$$

an upper-and-lower censored version of ξ_i; moreover, let $W = \sum_{i=1}^n \xi_i$ and $W_b = \sum_{i=1}^n \xi_{b,i}$ be their corresponding sums, and Δ_1 and Δ_2 be two random variables on the same probability space. Assume there exists $c_1 > c_2 > 0$ and $\delta \in (0, 1/2)$ such that

$$
\sum_{i=1}^n \mathbb{E}[\xi_i^2] \le c_1
$$

and

$$
\sum_{i=1}^n \mathbb{E}[|\xi_i| \min(\delta, |\xi_i|/2)] \ge c_2.
$$

Then for any $\lambda \ge 0$, it is true that

$$
\begin{aligned}
&\mathbb{E}[e^{\lambda W_b} I(\Delta_1 \le W_b \le \Delta_2)] \\
&\le \left(\mathbb{E}\left[e^{2\lambda W_b}\right]\right)^{1/2} \exp\left(-\frac{c_2^2}{16 c_1 \delta^2}\right) \\
&\quad + \frac{2e^{\lambda\delta}}{c_2}\left\{\sum_{i=1}^n \mathbb{E}[|\xi_{b,i}| e^{\lambda W_b^{(i)}} (|\Delta_1 - \Delta_1^{(i)}| + |\Delta_2 - \Delta_2^{(i)}|)]\right. \\
&\quad + \mathbb{E}[|W_b| e^{\lambda W_b} (|\Delta_2 - \Delta_1| + 2\delta)] \\
&\quad \left. + \sum_{i=1}^n \left|\mathbb{E}[\xi_{b,i}]\right| \mathbb{E}[e^{\lambda W_b^{(i)}} (|\Delta_2^{(i)} - \Delta_1^{(i)}| + 2\delta)]\right\},
\end{aligned}
$$

where $\Delta_1^{(i)}$ and $\Delta_2^{(i)}$ are any random variables on the same probability space such that ξ_i and $(\Delta_1^{(i)}, \Delta_2^{(i)}, W^{(i)}, W_b^{(i)})$ are independent, where $W^{(i)} = W - \xi_i$ and $W_b^{(i)} = W_b - \xi_{b,i}$.

In particular, by defining $\beta_2 \equiv \sum_{i=1}^n \mathbb{E}[\xi_i^2 I(|\xi_i| > 1)]$ and $\beta_3 \equiv \sum_{i=1}^n \mathbb{E}[\xi_i^3 I(|\xi_i| \leq 1)]$, if $\sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1$ and $\beta_2 + \beta_3 \leq 1/2$, one can take

$$\delta = \frac{\beta_2 + \beta_3}{4}, \quad c_1 = 1 \text{ and } \quad c_2 = \frac{1}{4} \tag{B.1}$$

to satisfy the conditions of the inequality.

Proof of Lemma B.1 It suffices to show the lemma under the assumptions that

$$\Delta_1 \leq \Delta_2 \text{ and } \Delta_1^{(i)} \leq \Delta_2^{(i)}. \tag{B.2}$$

If (B.2) is not true, we can let $\Delta_1^* = \min(\Delta_1, \Delta_2)$, $\Delta_2^* = \max(\Delta_1, \Delta_2)$, $\Delta_1^{*(i)} = \min(\Delta_1^{(i)}, \Delta_2^{(i)})$, $\Delta_2^{*(i)} = \max(\Delta_1^{(i)}, \Delta_2^{(i)})$. Then the assumptions in (B.2) can be seen to be not forgoing any generality by noting that $|\Delta_2^* - \Delta_1^*| = |\Delta_2 - \Delta_1|$ (also $|\Delta_2^{*(i)} - \Delta_1^{*(i)}| = |\Delta_2^{(i)} - \Delta_1^{(i)}|$),

$$\mathbb{E}[e^{\lambda W_b} I(\Delta_1 \leq W_b \leq \Delta_2)] \leq \mathbb{E}[e^{\lambda W_b} I(\Delta_1^* \leq W_b \leq \Delta_2^*)]$$

and

$$|\Delta_1^* - \Delta_1^{*(i)}| + |\Delta_2^* - \Delta_2^{*(i)}| \leq |\Delta_1 - \Delta_1^{(i)}| + |\Delta_2 - \Delta_2^{(i)}|, \tag{B.3}$$

where (B.3) is true by the following fact: If we have real numbers $x_1 \leq x_2$ and $y_1 \leq y_2$, it must be that

$$|x_1 - y_1| + |x_2 - y_2| \leq |x_1 - y_2| + |x_2 - y_1|. \tag{B.4}$$

Without loss of generality, one can assume $x_1 \leq y_1$ and simply prove (B.4) by case considerations:

(i) If $x_1 \leq x_2 \leq y_1 \leq y_2$, then

$$\begin{aligned} |x_1 - y_1| + |x_2 - y_2| &= y_1 - x_1 + y_2 - x_2 \\ &= y_2 - x_1 + y_1 - x_2 = |x_1 - y_2| + |x_2 - y_1|. \end{aligned}$$

(ii) If $x_1 \leq y_1 \leq x_2 \leq y_2$, , then

$$\begin{aligned} |x_1 - y_1| + |x_2 - y_2| &= y_1 - x_1 + y_2 - x_2 \\ &\leq y_2 - x_1 \leq |x_1 - y_2| + |x_2 - y_1|. \end{aligned}$$

(iii) If $x_1 \leq y_1 \leq y_2 \leq x_2$, , then

$$|x_1 - y_1| + |x_2 - y_2| = \underbrace{y_1 - x_1}_{\leq y_2 - x_1} + \underbrace{x_2 - y_2}_{\leq x_2 - y_1} \leq |x_1 - y_2| + |x_2 - y_1|.$$

More generally, a fact like (B.4) can be proved by the rearrangement inequality [11, p.78], but the details are omitted here.

Under the working assumptions in (B.2), for $a < b$, we define the function

$$f_{a,b}(w) = \begin{cases} 0 & \text{for } w \leq a - \delta \\ e^{\lambda w}(w - a + \delta) & \text{for } a - \delta < w \leq b + \delta \\ e^{\lambda w}(b - a + 2\delta) & \text{for } w > b + \delta \end{cases},$$

which has the property

$$|f_{a,b}(w) - f_{a_1,b_1}(w)| \leq e^{\lambda w}(|a - a_1| + |b - b_1|) \text{ for all } w, \quad a < b \text{ and } a_1 < b_1, \tag{B.5}$$

as well as

$$f'_{a,b}(w) \geq 0 \text{ almost surely.}$$

Moreover, we have

$$I_1 + I_2 = \mathbb{E}[W_b f_{\Delta_1,\Delta_2}(W_b)] - \sum_{i=1}^{n} \mathbb{E}[\xi_{b,i}]\mathbb{E}[f_{\Delta_1^{(i)},\Delta_2^{(i)}}(W_b^{(i)}))] \tag{B.6}$$

where

$$I_1 \equiv \sum_{i=1}^{n} \mathbb{E}\left[\xi_{b,i}\left(f_{\Delta_1,\Delta_2}(W_b) - f_{\Delta_1,\Delta_2}(W_b^{(i)})\right)\right] \text{ and}$$
$$I_2 \equiv \sum_{i=1}^{n} \mathbb{E}\left[\xi_{b,i}\left(f_{\Delta_1,\Delta_2}(W_b^{(i)}) - f_{\Delta_1^{(i)},\Delta_2^{(i)}}(W_b^{(i)})\right)\right].$$

Given the property in (B.5), we have

$$|I_2| \leq \sum_{i=1}^{n} \mathbb{E}\Big[|\xi_{b,i}|e^{\lambda W_b^{(i)}}\Big(|\Delta_1 - \Delta_1^{(i)}| + |\Delta_2 - \Delta_2^{(i)}|\Big)\Big]. \tag{B.7}$$

Now we estimate I_1, by first rewriting it as

$$I_1 = \sum_{i=1}^n \mathbb{E}\left[\xi_{b,i}\left(f_{\Delta_1,\Delta_2}(W_b) - f_{\Delta_1,\Delta_2}(W_b^{(i)})\right)\right]$$
$$= \sum_{i=1}^n \mathbb{E}\left[\xi_{b,i}\int_{-\xi_{b,i}}^0 f'_{\Delta_1,\Delta_2}(W_b+t)\mathrm{d}t\right] = \sum_{i=1}^n \mathbb{E}\left[\int_{-\infty}^{\infty} f'_{\Delta_1,\Delta_2}(W_b+t)\hat{K}_i(t)\mathrm{d}t\right],$$

where

$$\hat{K}_i(t) \equiv \xi_{b,i}\{I(-\xi_{b,i} \le t \le 0) - I(0 < t \le -\xi_{b,i})\}.$$

Note that $\xi_{b,i}$ and $I(-\xi_{b,i} \le t \le 0) - I(0 < t \le -\xi_{b,i})$ have the same sign, and it is also true that $0 \le \tilde{K}_i(t) \le \hat{K}_i(t)$ where

$$\tilde{K}_i(t) = \xi_{b,i}\{I(-\xi_{b,i}/2 \le t \le 0) - I(0 < t \le -\xi_{b,i}/2)\}$$

By the fact that $f'_{\Delta_1,\Delta_2}(w) \ge e^{\lambda w} \ge 0$ for all $w \in (\Delta_1 - \delta, \Delta_2 + \delta)$, one can lower bound I_1 as

$$I_1 \ge \sum_{i=1}^n \mathbb{E}\left[\int_{-\infty}^{\infty} f'_{\Delta_1,\Delta_2}(W_b+t)\tilde{K}_i(t)\mathrm{d}t\right]$$
$$\ge \sum_{i=1}^n \mathbb{E}\left[\int_{|t|\le\delta} I(\Delta_1 \le W_b \le \Delta_2) f'_{\Delta_1,\Delta_2}(W_b+t)\tilde{K}_i(t)\mathrm{d}t\right]$$
$$\ge \sum_{i=1}^n \mathbb{E}\left[I(\Delta_1 \le W_b \le \Delta_2)e^{\lambda(W_b-\delta)}|\xi_{b,i}|\min(\delta, |\xi_{b,i}|/2)\right]$$
$$= \mathbb{E}\left[I(\Delta_1 \le W_b \le \Delta_2)e^{\lambda(W_b-\delta)}\left(\sum_{i=1}^n \eta_i\right)\right],$$

where $\eta_i = |\xi_i|\min(\delta, |\xi_i|/2)$, noting that given $\delta < 1/2$, $\min(\delta, |\xi_i|/2) = \min(\delta, |\xi_{b,i}|/2)$ due to the censoring definition of $\xi_{b,i}$. Hence, continuing, we can further lower bound I_1 as

$$I_1 \ge (c_2/2)\mathbb{E}\left[e^{\lambda(W_b-\delta)}I(\Delta_1 \le W_b \le \Delta_2)I\left(\sum_{i=1}^n \eta_i \ge c_2/2\right)\right]$$
$$\ge \frac{c_2}{2e^{\lambda\delta}}\left\{\mathbb{E}\left[e^{\lambda W_b}I(\Delta_1 \le W_b \le \Delta_2)\right] - \mathbb{E}\left[e^{\lambda W_b}I\left(\sum_{i=1}^n \eta_i < c_2/2\right)\right]\right\}$$
$$\ge \frac{c_2}{2e^{\lambda\delta}}\left\{\mathbb{E}\left[e^{\lambda W_b}I(\Delta_1 \le W_b \le \Delta_2)\right] - \sqrt{\mathbb{E}\left[e^{2\lambda W_b}\right]P\left(\sum_{i=1}^n \eta_i < c_2/2\right)}\right\}$$

$$\geq \frac{c_2}{2e^{\lambda\delta}} \left\{ \mathbb{E}[e^{\lambda W_b} I(\Delta_1 \leq W_b \leq \Delta_2)] - \left(\mathbb{E}\left[e^{2\lambda W_b}\right]\right)^{1/2} \exp\left(-\frac{c_2^2}{16c_1\delta^2}\right)\right\}, \tag{B.8}$$

where the last inequality comes from the sub-Gaussian lower tail bound for sum of non-negative random variables [14, Theorem 2.19],

$$P\left(\sum_{i=1}^n \eta_i < c_2/2\right) \leq \exp\left(-\frac{(c_2/2)^2}{2\sum_{i=1}^n \mathbb{E}[\eta_i^2]}\right) \leq \exp\left(-\frac{c_2^2}{8c_1\delta^2}\right).$$

Clearly, since $|f_{\Delta_1,\Delta_2}(w)| \leq e^{\lambda w}(\Delta_2 - \Delta_1 + 2\delta)$, we have, from (B.6),

$$\begin{aligned} I_1 + I_2 \leq\ & \mathbb{E}[|W_b| e^{\lambda W_b}(|\Delta_2 - \Delta_1| + 2\delta)] \\ & + \sum_{i=1}^n \left|\mathbb{E}[\xi_{b,i}]\right| \mathbb{E}[e^{\lambda W_b^{(i)}}(|\Delta_2^{(i)} - \Delta_1^{(i)}| + 2\delta)] \end{aligned} \tag{B.9}$$

Combining (B.7), (B.8) and (B.9), the proof is done.

If $\sum_{i=1}^n \mathbb{E}[\xi_i^2] = 1$ and $\beta_2 + \beta_3 \leq 1/2$, one can apparently take $c_1 = 1$. The parameter choices of c_2 and δ in (B.1) can be justified as follows: Using the fact that [3, p.259]

$$\min(x, y) \geq y - \frac{y^2}{4x} \text{ for } x > 0 \text{ and } y \geq 0,$$

by taking $\delta = (\beta_2 + \beta_3)/4$, we have

$$\begin{aligned} \sum_{i=1}^n \mathbb{E}[|\xi_i| \min(\delta, |\xi_i|/2)] &\geq \sum_{i=1}^n \mathbb{E}[|\xi_i| I(|\xi_i| \leq 1) \min(\delta, |\xi_i|/2)] \\ &\geq \sum_{i=1}^n \left[\frac{\mathbb{E}[\xi_i^2 I(|\xi_i| \leq 1)]}{2} - \frac{\mathbb{E}[|\xi_i|^3 I(|\xi_i| \leq 1)]}{16\delta}\right] = \frac{1-\beta_2}{2} - \frac{\beta_3}{16\delta} \\ &= \frac{1}{2} - \frac{8\delta\beta_2 + \beta_3}{16\delta} \underbrace{\geq}_{\delta \leq 1/8} \frac{1}{2} - \frac{\beta_2 + \beta_3}{16\delta} = \frac{1}{4}. \end{aligned}$$

□

Appendix C. Proof of Theorem 2.1

This section presents the proof of Theorem 2.1. The approach is similar to that of Shao et al. [9, Theorem 3.1], but there are quite a number of differences stemming from correcting the numerous gaps in the latter. It suffices to consider $x \geq 0$, or else we can consider $-T_{SN}$ instead[2]. Moreover, without loss of generality, we can assume

$$\beta_2 + \beta_3 < 1/2; \tag{C.1}$$

otherwise, it must be true that $|P(T_{SN} \le x) - \Phi(x)| \le 2(\beta_2 + \beta_3)$. Since

$$1 + s/2 - s^2/2 \le (1+s)^{1/2} \le 1 + s/2 \text{ for all } s \ge -1,$$

we have the two inclusions

$$\{T_{SN} > x\} \subset \{W_n + D_{1n} - xD_{2n}/2 > x\} \cup \{x + x(D_{2n} - D_{2n}^2)/2 < W_n + D_{1n} \le x + xD_{2n}/2\}$$

and

$$\{T_{SN} > x\} \supset \{W_n + D_{1n} - xD_{2n}/2 > x\}.$$

Hence, it suffices to establish the bounds

$$\begin{aligned} P(x + x(D_{2n} - D_{2n}^2)/2 \le W_n + D_{1n} \le x + xD_{2n}/2) \le \sum_{j=1}^{2} P(|D_{jn}| > 1/2) \\ + C\Big\{\beta_2 + \beta_3 + \mathbb{E}\Big[(1 + e^{W_b})\bar{D}_{2n}^2\Big] \\ + \sum_{j=1}^{2}\sum_{i=1}^{n} \|\xi_{b,i} e^{W_b^{(i)}/2}(\bar{D}_{jn} - \bar{D}_{jn}^{(i)})\|_1\Big\} \end{aligned} \tag{C.2}$$

and

$$\begin{aligned} |P(W_n + D_{1n} - xD_{2n}/2 \le x) - \Phi(x)| \le \sum_{j=1}^{2} P(|D_{jn}| > 1/2) \\ + C\Big\{\beta_2 + \beta_3 + \|\bar{D}_{1n}\|_2 + \mathbb{E}\Big[(1 + e^{W_b})\bar{D}_{2n}^2\Big] \\ + \Big|x\mathbb{E}[\bar{D}_{2n} f_x(W_b)]\Big| \\ + \sum_{j=1}^{2}\sum_{i=1}^{n}\Big(\mathbb{E}[\xi_{b,i}^2]\Big\|(1 + e^{W_b^{(i)}})(\bar{D}_{jn} - \bar{D}_{jn}^{(i)})\Big\|_1 \\ + \Big\|\xi_{b,i}(1 + e^{W_b^{(i)}/2})(\bar{D}_{jn} - \bar{D}_{jn}^{(i)})\Big\|_1\Big)\Big\} \end{aligned} \tag{C.3}$$

[2] For a given $x < 0$, if one can uniformly bound $|P(T_{SN} < x + \epsilon) - \Phi(x + \epsilon)|$ for all $\epsilon \in (x, 0)$, one can then similarly bound $|P(T_{SN} \le x) - \Phi(x)|$ by taking limits on both sides as $\epsilon \longrightarrow 0$.

separately. Before starting to prove them, we introduce the following notation:

$$\bar{\Delta}_{1n,x} = \frac{x(\bar{D}_{2n} - \bar{D}_{2n}^2)}{2} - \bar{D}_{1n} \text{ and } \bar{\Delta}_{2n,x} = \frac{x\bar{D}_{2n}}{2} - \bar{D}_{1n}.$$

C.1 Proof of (C.2)

We further introduce

$$\bar{\Delta}^{(i)}_{1n,x} = \frac{x(\bar{D}^{(i)}_{2n} - (\bar{D}^{(i)}_{2n})^2)}{2} - \bar{D}^{(i)}_{1n} \text{ and } \bar{\Delta}^{(i)}_{2n,x} = \frac{x\bar{D}^{(i)}_{2n}}{2} - \bar{D}^{(i)}_{1n}.$$

Noting that

$$\begin{aligned} &P\Big(x + x(D_{2n} - D_{2n}^2)/2 \le W_n + D_{1n} \le x + xD_{2n}/2\Big) \\ &\le P_0 + \sum_{j=1}^{2} P(|D_{jn}| > 1/2) + \beta_2, \end{aligned} \tag{C.4}$$

where

$$P_0 = P(x + \bar{\Delta}_{1n,x} \le W_b \le x + \bar{\Delta}_{2n,x}),$$

it suffices to bound P_0. Since $\bar{D}_{2n} - \bar{D}_{2n}^2 \ge -3/4$ and hence $\frac{1}{2}(x + \bar{\Delta}_{1n,x}) \ge \frac{1}{2}(\frac{5x}{8} - \frac{1}{2}) > \frac{x}{4} - \frac{1}{4}$, in light of (C.1), applying Lemma B.1 with the parameters in (B.1) and $\lambda = 1/2$ implies that

$$\begin{aligned} e^{x/4-1/4}P_0 &\le \mathbb{E}[e^{W_b/2} I(x + \bar{\Delta}_{1n,x} \le W_b \le x + \bar{\Delta}_{2n,x})] \\ &\le \left(\mathbb{E}\left[e^{W_b}\right]\right)^{1/2} \exp\left(-\frac{1}{16(\beta_2+\beta_3)^2}\right) \\ &\quad + 8e^{(\beta_2+\beta_3)/8}\Bigg\{\sum_{i=1}^{n} \mathbb{E}\Big[|\xi_{b,i}|e^{W_b^{(i)}/2}\Big(|\bar{\Delta}_{1n,x} - \bar{\Delta}^{(i)}_{1n,x}| + |\bar{\Delta}_{2n,x} - \bar{\Delta}^{(i)}_{2n,x}|\Big)\Big] \\ &\quad + \mathbb{E}\left[|W_b|e^{W_b/2}\left(|\bar{\Delta}_{2n,x} - \bar{\Delta}_{1n,x}| + \frac{\beta_2+\beta_3}{2}\right)\right] \\ &\quad + \sum_{i=1}^{n} \Big|\mathbb{E}[\xi_{b,i}]\Big| \mathbb{E}\left[e^{W_b^{(i)}/2}\left(|\bar{\Delta}^{(i)}_{2n,x} - \bar{\Delta}^{(i)}_{1n,x}| + \frac{\beta_2+\beta_3}{2}\right)\right]\Bigg\} \end{aligned} \tag{C.5}$$

We will bound different terms on the right-hand side of (C.5). First,

$$\mathbb{E}[e^{W_b}] \le \exp(e^2/4 + 1/4) \text{ by Lemma A.2} \tag{C.6}$$

and

$$\exp\left(\frac{-1}{16(\beta_2+\beta_3)^2}\right) \le C(\beta_2+\beta_3). \tag{C.7}$$

Since $\bar{D}_{2n}^2 - (\bar{D}_{2n}^{(i)})^2 = (\bar{D}_{2n} - \bar{D}_{2n}^{(i)})(\bar{D}_{2n} + \bar{D}_{2n}^{(i)})$,

$$\begin{aligned}&\mathbb{E}[|\xi_{b,i}|e^{W_b^{(i)}/2}(|\bar{\Delta}_{1n,x} - \bar{\Delta}_{1n,x}^{(i)}| + |\bar{\Delta}_{2n,x} - \bar{\Delta}_{2n,x}^{(i)}|)]\\ &\quad \le C\mathbb{E}[|\xi_{b,i}|e^{W_b^{(i)}/2}(|\bar{D}_{1n} - \bar{D}_{1n}^{(i)}| + x|\bar{D}_{2n} - \bar{D}_{2n}^{(i)}|)].\end{aligned} \tag{C.8}$$

Moreover, since $\frac{|W_b|}{2} \le e^{|W_b|/2} \le e^{W_b/2} + e^{-W_b/2}$, by Lemma A.2,

$$\begin{aligned}&\mathbb{E}\left[|W_b|e^{W_b/2}\left(|\bar{\Delta}_{2n,x} - \bar{\Delta}_{1n,x}| + \frac{\beta_2+\beta_3}{2}\right)\right]\\ &\quad \le C_1 x\mathbb{E}[(1+e^{W_b})\bar{D}_{2n}^2] + C_2(\beta_2+\beta_3).\end{aligned} \tag{C.9}$$

Lastly, by Lemma A.1, Bennett's inequality (Lemma A.2) and (C.1), we have

$$\begin{aligned}&\sum_{i=1}^n \left|\mathbb{E}[\xi_{b,i}]\right| \mathbb{E}\left[e^{W_b^{(i)}/2}\left(|\bar{\Delta}_{2n,x}^{(i)} - \bar{\Delta}_{1n,x}^{(i)}| + \frac{\beta_2+\beta_3}{2}\right)\right]\\ &\quad \le C\sum_{i=1}^n \left|\mathbb{E}[\xi_{b,i}]\right| \underbrace{\left(x\mathbb{E}[e^{W_b^{(i)}/2}(\bar{D}_{2n}^{(i)})^2] + \beta_2 + \beta_3\right)}_{\le\ C(1+x)} \le C(1+x)\beta_2.\end{aligned} \tag{C.10}$$

Collecting (C.4)–(C.10), we get (C.2).

C.2 Proof of (C.3)

For this part, as a proof device, we let $X_1^*, \ldots, X_n^*$ be independent copies of $X_1, \ldots, X_n$ and in analogy to (1.4), we introduce

$$\begin{aligned}D_{1n,i^*} &= D_{1n}(X_1, \ldots, X_{i-1}, X_i^*, X_{i+1}, \ldots, X_n) \text{ and}\\ D_{2n,i^*} &= D_{2n}(X_1, \ldots, X_{i-1}, X_i^*, X_{i+1}, \ldots, X_n),\\ \bar{D}_{1n,i^*} &= D_{1n,i^*} I\left(|D_{1n,i^*}| \le \frac{1}{2}\right) + \frac{1}{2} I\left(D_{1n,i^*} > \frac{1}{2}\right) - \frac{1}{2} I\left(D_{1n,i^*} < -\frac{1}{2}\right) \text{ and}\\ \bar{D}_{2n,i^*} &= D_{2n,i^*} I\left(|D_{2n,i^*}| \le \frac{1}{2}\right) + \frac{1}{2} I\left(D_{2n,i^*} > \frac{1}{2}\right) - \frac{1}{2} I\left(D_{2n,i^*} < -\frac{1}{2}\right),\end{aligned}$$

as well as

$$\bar{\Delta}_{2n,x,i^*} = \frac{x\bar{D}_{2n,i^*}}{2} - \bar{D}_{1n,i^*},$$

which are correspondingly versions of D_{1n}, D_{2n}, $\bar{D}_{1n}$, $\bar{D}_{2n}$ and $\bar{\Delta}_{2n,x}$ with X_i^* replacing X_i as input. For any pair $1 \leq i, i' \leq n$ and $j \in \{1, 2\}$, we also define

$$D_{jn,i^*}^{(i')} \equiv \begin{cases} D^{(i')}(X_1, \ldots, X_{i-1}, X_i^*, X_{i+1}, \ldots, X_{i'-1}, X_{i'+1}, \ldots, X_n) & \text{if } i < i' \\ D^{(i')}(X_1, \ldots, X_{i'-1}, X_{i'+1}, \ldots, X_{i-1}, X_i^*, X_{i+1}, \ldots, X_n) & \text{if } i > i' \\ D^{(i')}(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n) & \text{if } i = i' \end{cases},$$

i.e., $D_{jn,i^*}^{(i')}$ is a version of the "leave-one-out" $D_{jn}^{(i')}$ with X_i^* replacing X_i as input, and its censored version

$$\bar{D}_{jn,i^*}^{(i')} \equiv D_{jn,i^*}^{(i')} I\left(|D_{jn,i^*}^{(i')}| \leq \frac{1}{2}\right) + \frac{1}{2} I\left(D_{jn,i^*}^{(i')} > \frac{1}{2}\right) - \frac{1}{2} I\left(D_{jn,i^*}^{(i')} < -\frac{1}{2}\right).$$

It suffices to bound $|P(W_b - \bar{\Delta}_{2n,x} \leq x) - \Phi(x)|$ since

$$|P(W_n - \Delta_{2n,x} \leq x) - \Phi(x)| \leq |P(W_b - \bar{\Delta}_{2n,x} \leq x) - \Phi(x)| + \beta_2 + \sum_{j=1}^{2} P(|D_{jn}| > 1/2). \quad \text{(C.11)}$$

First, define the *K function*

$$k_{b,i}(t) = \mathbb{E}[\xi_{b,i}\{I(0 \leq t \leq \xi_{b,i}) - I(\xi_{b,i} \leq t < 0)\}],$$

which has the properties

$$\int_{-\infty}^{\infty} k_{b,i}(t)\mathrm{d}t = \int_{-1}^{1} k_{b,i}(t)\mathrm{d}t = \mathbb{E}[\xi_{b,i}^2] = \|\xi_{b,i}\|_2^2 \quad \text{and}$$
$$\int_{-\infty}^{\infty} |t| k_{b,i}(t)\mathrm{d}t = \int_{-1}^{1} |t| k_{b,i}(t)\mathrm{d}t = \frac{\mathbb{E}[|\xi_{b,i}|^3]}{2} = \frac{\|\xi_{b,i}\|_3^3}{2}. \quad \text{(C.12)}$$

Since

$$\mathbb{E}\Big[\int_{-1}^{1} f_x'(W_b^{(i)} - \bar{\Delta}_{2n,x,i^*} + t) k_{b,i}(t)\mathrm{d}t\Big] = \mathbb{E}\Big[\xi_{b,i}\{f_x(W_b - \bar{\Delta}_{2n,x,i^*}) - f_x(W_b^{(i)} - \bar{\Delta}_{2n,x,i^*})\}\Big]$$

by independence and the fundamental theorem of calculus, from the Stein equation (2.4), one can then write

$$\begin{aligned} &P(W_b - \bar{\Delta}_{2n,x} \leq x) - \Phi(x) \\ &\quad = \mathbb{E}[f_x'(W_b - \bar{\Delta}_{2n,x})] - \mathbb{E}[W_b f_x(W_b - \bar{\Delta}_{2n,x})] \\ &\qquad + \mathbb{E}\Big[\bar{\Delta}_{2n,x}\Big(f_x(W_b - \bar{\Delta}_{2n,x}) - f_x(W_b)\Big)\Big] + \mathbb{E}[\bar{\Delta}_{2n,x} f_x(W_b)] \end{aligned}$$

$$
\begin{aligned}
&= \underbrace{\sum_{i=1}^{n} \mathbb{E}\Big[\int_{-1}^{1} \{f_x'(W_b - \bar{\Delta}_{2n,x}) - f_x'(W_b^{(i)} - \bar{\Delta}_{2n,x,i^*} + t)\} k_{b,i}(t) \mathrm{d}t\Big]}_{R_1} \\
&\quad + \underbrace{\sum_{i=1}^{n} \mathbb{E}[(\xi_i^2 - 1) I(|\xi_i| > 1)] \mathbb{E}[f_x'(W_b - \bar{\Delta}_{2n,x})] - \sum_{i=1}^{n} \mathbb{E}[\xi_{b,i} f_x(W_b^{(i)} - \bar{\Delta}_{2n,x,i^*})] + \mathbb{E}[\bar{\Delta}_{2n,x} f_x(W_b)]}_{R_2} \\
&\quad + \underbrace{\Big\{ - \sum_{i=1}^{n} \mathbb{E}\Big[\xi_{b,i} \Big\{ f_x(W_b - \bar{\Delta}_{2n,x}) - f_x(W_b - \bar{\Delta}_{2n,x,i^*}) \Big\}\Big]\Big\}}_{R_3} \\
&\quad + \underbrace{\mathbb{E}\Big[\bar{\Delta}_{2n,x} \int_0^{-\bar{\Delta}_{2n,x}} f_x'(W_b + t) \mathrm{d}t\Big]}_{R_4} \\
&= R_1 + R_2 + R_3 + R_4.
\end{aligned}
$$

To finish the proof, we will establish the following bounds for R_1, R_2, R_3, R_4:

$$
\begin{aligned}
|R_1| &\le C\Big\{\beta_2 + \beta_3 + \sum_{j=1}^{2}\sum_{i=1}^{n}\Big(\mathbb{E}[\xi_{b,i}^2]\Big\|(1 + e^{W_b^{(i)}})(\bar{D}_{jn} - \bar{D}_{jn}^{(i)})\Big\|_1 \\
&\qquad + \Big\|\xi_{b,i} e^{W_b^{(i)}/2}(\bar{D}_{jn} - \bar{D}_{jn}^{(i)})\Big\|_1\Big)\Big\} \qquad \text{(C.13)}\\
|R_2| &\le 1.63\beta_2 + 0.63\|\bar{D}_{1n}\|_2 + \Big|\frac{x}{2}\mathbb{E}[\bar{D}_{2n} f_x(W_b)]\Big|, \qquad \text{(C.14)}\\
|R_3| &\le C\sum_{j=1}^{2}\sum_{i=1}^{n}\|\xi_{b,i}(1 + e^{W_b^{(i)}/2})(\bar{D}_{jn} - \bar{D}_{jn}^{(i)})\|_1, \qquad \text{(C.15)}\\
|R_4| &\le C\Big(\|\bar{D}_{1n}\|_2 + \mathbb{E}[(1 + e^{W_b})\bar{D}_{2n}^2]\Big). \qquad \text{(C.16)}
\end{aligned}
$$

Then (C.13)–(C.16) together with (C.11) conclude (C.3).

C.2.1 Bound for R_1

Let $g_x(w) = (wf_x(w))'$ as defined in (A.1). By the Stein equation (2.4) and defining $\eta_1 = t - \bar{\Delta}_{2n,x,i^*}$ and $\eta_2 = \xi_{b,i} - \bar{\Delta}_{2n,x}$, we can write

$$R_1 = R_{11} + R_{12},$$

where

$$R_{11} = \sum_{i=1}^{n}\int_{-1}^{1}\mathbb{E}\Big[\int_{t-\bar{\Delta}_{2n,x,i^*}}^{\xi_{b,i}-\bar{\Delta}_{2n,x}} g_x(W_b^{(i)} + u)du\Big]k_{b,i}(t)\mathrm{d}t$$

$$= \underbrace{\sum_{i=1}^{n} \int_{-1}^{1} \mathbb{E}\Big[\int g_x(W_b^{(i)} + u) I(\eta_1 \le u \le \eta_2) du \Big] k_{b,i}(t) \mathrm{d}t}_{R_{11.1}}$$
$$- \underbrace{\sum_{i=1}^{n} \int_{-1}^{1} \mathbb{E}\Big[\int g_x(W_b^{(i)} + u) I(\eta_2 \le u \le \eta_1) du \Big] k_{b,i}(t) \mathrm{d}t}_{R_{11.2}}$$

and

$$R_{12} = \sum_{i=1}^{n} \int_{-1}^{1} \{P(W_b - \bar{\Delta}_{2n,x} \le x) - P(W_b^{(i)} - \bar{\Delta}_{2n,x,i^*} + t \le x)\} k_{b,i}(t) \mathrm{d}t.$$

For $0 \le x < 1$, since $|g_x| \le 2.3$ (Lemma A.3), using the properties in (C.12), we have

$$\begin{aligned} |R_{11}| &\le C \sum_{i=1}^{n} \int_{-1}^{1} \Big(|t| + \|\xi_{b,i}\|_1 + \sum_{j=1}^{2} \|\bar{D}_{jn} - \bar{D}_{jn,i^*}\|_1 \Big) k_{b,i}(t) \mathrm{d}t \\ &\le C \left(\sum_{i=1}^{n} \|\xi_{b,i}\|_3^3 + \sum_{i=1}^{n} \|\xi_{b,i}\|_2^2 \|\xi_{b,i}\|_1 + \sum_{j=1}^{2} \sum_{i=1}^{n} \|\xi_{b,i}\|_2^2 \|\bar{D}_{jn} - \bar{D}_{jn,i^*}\|_1 \right) \\ &\le C \left(\beta_2 + \beta_3 + \sum_{j=1}^{2} \sum_{i=1}^{n} \|\xi_{b,i}\|_2^2 \|\bar{D}_{jn} - \bar{D}_{jn,i^*}\|_1 \right) \quad \text{for } 0 \le x < 1, \end{aligned} \tag{C.17}$$

where we have used $\|\xi_{b,i}\|_1 \le \|\xi_{b,i}\|_2 \le \|\xi_{b,i}\|_3$ and

$$\begin{aligned} \|\xi_{b,i}\|_3^3 &= \mathbb{E}[|\xi_i^3| I(|\xi_i| \le 1)] + \mathbb{E}[I(|\xi_i| > 1)] \\ &\le \mathbb{E}[|\xi_i^3| I(|\xi_i| \le 1)] + \mathbb{E}[\xi_i^2 I(|\xi_i| > 1)] \end{aligned} \tag{C.18}$$

in the last inequality.

For $x \ge 1$, we first bound the integrand of $R_{11.1}$. Using the identity

$$\begin{aligned} 1 &= I(W_b^{(i)} + u \le x - 1) + I(x - 1 < W_b^{(i)} + u, u \le 3x/4) \\ &\quad + I(x - 1 < W_b^{(i)} + u, u > 3x/4) \\ &\le I(W_b^{(i)} + u \le x - 1) + I(x - 1 < W_b^{(i)} + u, W_b^{(i)} + 1 > x/4) \\ &\quad + (x - 1 < W_b^{(i)} + u, u > 3x/4) \end{aligned}$$

and the bounds for $g_x(\cdot)$ in Lemma A.4, in light of $|\bar{\Delta}_{2n,x}| \le \frac{x|\bar{D}_{2n}|}{2} + |\bar{D}_{1n}| \le \frac{1}{2} + \frac{x}{4}$ and $1.6\bar{\Phi}(x) \le xe^{1/2-x}$,

$$\begin{aligned}
&\left|\mathbb{E}\Big[\int g_x(W_b^{(i)}+u)I(\eta_1 \le u \le \eta_2)du\Big]\right| \\
&\le xe^{1/2-x}\|\eta_2-\eta_1\|_1 + (x+2)\Big\{\|I(W_b^{(i)}+1 > x/4)(\eta_2-\eta_1)\|_1 \\
&\quad + \|I(\eta_2 > 3x/4)(\eta_2-\eta_1)\|_1\Big\} \\
&\le xe^{1/2-x}\|\eta_2-\eta_1\|_1 + \frac{x+2}{e^{x/4-1}}\|e^{W_b^{(i)}}(\eta_2-\eta_1)\|_1 + \frac{x+2}{e^{3x/4}}\|e^{\xi_{b,i}-\bar{\Delta}_{2n,x}}(\eta_2-\eta_1)\|_1 \\
&\le \left(xe^{1/2-x} + \frac{e^{3/2}(x+2)}{e^{x/2}}\right)\|\eta_2-\eta_1\|_1 + \frac{x+2}{e^{x/4-1}}\|e^{W_b^{(i)}}(\eta_2-\eta_1)\|_1 \\
&\le \frac{C(x+2)}{e^{x/4}}\Big\{|t| + \|\bar{\Delta}_{2n,x,i^*} - \bar{\Delta}_{2n,x} + \xi_{b,i}\|_1 + \|e^{W_b^{(i)}}(\bar{\Delta}_{2n,x,i^*} - \bar{\Delta}_{2n,x} + \xi_{b,i})\|_1\Big\}
\end{aligned}$$

where we have used the Bennett's inequality (Lemma A.2) via $\|e^{W_b^{(i)}}t\|_1 \le C|t|$. Continuing,

$$\begin{aligned}
&\left|\mathbb{E}\Big[\int g_x(W_b^{(i)}+u)I(\eta_1 \le u \le \eta_2)du\Big]\right| \\
&\le \frac{C(x+2)}{e^{x/4}}\Big\{|t| + \|x(\bar{D}_{2n,i^*} - \bar{D}_{2n}) - (\bar{D}_{1n,i^*} - \bar{D}_{1n}) + \xi_{b,i}\|_1 \\
&\quad + \|e^{W_b^{(i)}}[x(\bar{D}_{2n,i^*} - \bar{D}_{2n}) - (\bar{D}_{1n,i^*} - \bar{D}_{1n}) + \xi_{b,i}]\|_1\Big\} \\
&\le C\Big\{|t| + (1+\|e^{W_b^{(i)}}\|_2)\|\xi_{b,i}\|_2 + \sum_{j=1}^{2}\|(1+e^{W_b^{(i)}})(\bar{D}_{jn,i^*} - \bar{D}_{jn})\|_1\Big\} \\
&\le C\Big\{|t| + \|\xi_{b,i}\|_2 + \sum_{j=1}^{2}\|(1+e^{W_b^{(i)}})(\bar{D}_{jn,i^*} - \bar{D}_{jn})\|_1\Big\},
\end{aligned} \tag{C.19}$$

where the last inequality uses Bennett's inequality (Lemma A.2 giving $\|e^{W_b^{(i)}}\|_2 \le C$). By a completely analogous argument, we also have the bound

$$\begin{aligned}
&\left|\mathbb{E}\Big[\int g_x(W_b^{(i)}+u)I(\eta_2 \le u \le \eta_1)du\Big]\right| \\
&\le C\Big\{|t| + \|\xi_{b,i}\|_2 + \sum_{j=1}^{2}\|(1+e^{W_b^{(i)}})(\bar{D}_{jn,i^*} - \bar{D}_{jn})\|_1\Big\}.
\end{aligned} \tag{C.20}$$

for the integrand of $R_{11.2}$, for $x \geq 1$. Combining (C.19) and (C.20), as well as the integral and moment properties in (C.12) and (C.18), via integrating over t, we have

$$
\begin{aligned}
|R_{11}| &\leq C\left\{\beta_2+\beta_3+\sum_{i=1}^{n}\|\xi_{b,i}\|_2^2\left(\|\xi_{b,i}\|_2+\sum_{j=1}^{2}\|(1+e^{W_b^{(i)}})(\bar{D}_{jn,i^*}-\bar{D}_{jn})\|_1\right)\right\}\\
&\leq C\left\{\beta_2+\beta_3+\sum_{j=1}^{2}\sum_{i=1}^{n}\|\xi_{b,i}\|_2^2\left\|(1+e^{W_b^{(i)}})(\bar{D}_{jn,i^*}-\bar{D}_{jn})\right\|_1\right\} \text{ for } x\geq 1,
\end{aligned}
\tag{C.21}
$$

where the last inequality also uses $\|\xi_{b,i}\|_2^3 \leq \|\xi_{b,i}\|_3^3$ and (C.18). Combining (C.21) with the bound for $x \in [0, 1)$ in (C.17), we get, for all $x \geq 0$,

$$
\begin{aligned}
|R_{11}| &\leq C\left\{\beta_2+\beta_3+\sum_{j=1}^{2}\sum_{i=1}^{n}\mathbb{E}[\xi_{b,i}^2]\left\|(1+e^{W_b^{(i)}})(\bar{D}_{jn}-\bar{D}_{jn,i^*})\right\|_1\right\}\\
&= C\left\{\beta_2+\beta_3+\sum_{j=1}^{2}\sum_{i=1}^{n}\mathbb{E}[\xi_{b,i}^2]\left\|(1+e^{W_b^{(i)}})(\bar{D}_{jn}-\bar{D}_{jn}^{(i)}+\bar{D}_{jn}^{(i)}-\bar{D}_{jn,i^*})\right\|_1\right\}\\
&\leq C\left\{\beta_2+\beta_3+\sum_{j=1}^{2}\sum_{i=1}^{n}\mathbb{E}[\xi_{b,i}^2]\left\|(1+e^{W_b^{(i)}})(\bar{D}_{jn}-\bar{D}_{jn}^{(i)})\right\|_1\right\}
\end{aligned}
\tag{C.22}
$$

where in the last inequality, we have used the fact that $(W_b^{(i)}, \bar{D}_{jn} - \bar{D}_{jn}^{(i)}) =_d (W_b^{(i)}, \bar{D}_{jn,i^*} - \bar{D}_{jn}^{(i)})$.

For R_{12}, its integrand for a given i is bounded by

$$
\begin{aligned}
&P(x+\bar{\Delta}_{2n,x} \leq W_b \leq x-t+\bar{\Delta}_{2n,x,i^*}+\xi_{b,i})\\
&\quad +P(x-t+\bar{\Delta}_{2n,x,i^*}+\xi_{b,i} \leq W_b \leq x+\bar{\Delta}_{2n,x})
\end{aligned}
\tag{C.23}
$$

Since

$$
(x+\bar{\Delta}_{2n,x}) \wedge (x-t+\bar{\Delta}_{2n,x,i^*}+\xi_{b,i}) \geq (3x)/4-5/2 \quad \text{for} \quad |t|\leq 1,
$$

and $\mathbb{E}[e^{W_b}] \leq C$ by Bennett's inequality (Lemma A.2), by defining

$$
\bar{\Delta}_{2n,x,i^*}^{(i')} \equiv \frac{x\bar{D}_{2n,i^*}^{(i')}}{2} - \bar{D}_{1n,i^*}^{(i')} \text{ for } 1 \leq i' \leq n,
$$

we can apply the randomized concentration inequality (Lemma B.1) with the parameters in (B.1) and $\lambda = 1/2$ to bound (C.23) by

$$
\begin{aligned}
&Ce^{-3x/8}\Big\{\beta_2+\beta_3 \\
&\quad+\sum_{i'=1}^{n}\mathbb{E}\Big[|\xi_{b,i'}|e^{W_b^{(i')}/2}\Big(|\bar{\Delta}_{2n,x}-\bar{\Delta}_{2n,x}^{(i')}|+|\bar{\Delta}_{2n,x,i^*}-\bar{\Delta}_{2n,x,i^*}^{(i')}|+I(i'=i)|\xi_{b,i}|\Big)\Big] \\
&\quad+\mathbb{E}\Big[\underbrace{|W_b|e^{W_b/2}}_{\leq 2(1+e^{W_b})}\Big(|\bar{\Delta}_{2n,x}-\bar{\Delta}_{2n,x,i^*}|+|\xi_{b,i}|+|t|+\beta_2+\beta_3\Big)\Big] \\
&\quad+\sum_{i'=1}^{n}\Big|\mathbb{E}[\xi_{b,i'}]\Big|\mathbb{E}\Big[e^{W_b^{(i')}/2}\underbrace{\Big(|t|+|\xi_{b,i}|I(i'\neq i)+|\bar{\Delta}_{2n,x}^{(i')}-\bar{\Delta}_{2n,x,i^*}^{(i')}|+\beta_2+\beta_3\Big)}_{\leq C(1+x)}\Big]\Big\} \\
&\leq C\Big\{\beta_2+\beta_3+\mathbb{E}[|\xi_{b,i}|^2e^{W_b^{(i)}/2}] \\
&\quad+\sum_{j=1}^{2}\sum_{i'=1}^{n}\mathbb{E}\Big[|\xi_{b,i'}|e^{W_b^{(i')}/2}\Big(|\bar{D}_{jn}-\bar{D}_{jn}^{(i')}|+|\bar{D}_{jn,i^*}-\bar{D}_{jn,i^*}^{(i')}|\Big)\Big] \\
&\quad+\mathbb{E}\Big[(1+e^{W_b})\Big(\sum_{j=1}^{2}|\bar{D}_{jn}-\bar{D}_{jn,i^*}|+|\xi_{b,i}|+|t|+\beta_2+\beta_3\Big)\Big]+\sum_{i'=1}^{n}\Big|\mathbb{E}[\xi_{b,i'}]\Big|\mathbb{E}\Big[e^{W_b^{(i')}/2}\Big]\Big\} \\
&\leq C\Big\{\beta_2+\beta_3+\mathbb{E}[|\xi_{b,i}|^2]+\sum_{j=1}^{2}\sum_{i'=1}^{n}\mathbb{E}\Big[|\xi_{b,i'}|e^{W_b^{(i')}/2}\Big(|\bar{D}_{jn}-\bar{D}_{jn}^{(i')}|+|\bar{D}_{jn,i^*}-\bar{D}_{jn,i^*}^{(i')}|\Big)\Big] \\
&\quad+\sum_{j=1}^{2}\|(1+e^{W_b})(\bar{D}_{jn}-\bar{D}_{jn,i^*})\|_1+\|\xi_{b,i}\|_2+|t|\Big\};
\end{aligned}
\tag{C.24}
$$

in (C.24), we have used that $\sum_{i'=1}^{n}|\mathbb{E}[\xi_{b,i'}]|\leq\beta_2$ by Lemma A.1 and

$$
\max(\|e^{W_b}\|_2,\|e^{W_b}\|_1,\mathbb{E}[e^{W_b^{(i')}/2}],\mathbb{E}[e^{W_b^{(i)}/2}])\leq C
$$

by Bennett's inequality (Lemma A.2). Since (C.24) bounds (C.23) which bounds the integrand of R_{12}, on integration with respect to t which has the properties in (C.12), we get

$$
\begin{aligned}
|R_{12}|\leq C\Big\{&\beta_2+\beta_3+\sum_{j=1}^{2}\Big[\sum_{i=1}^{n}\mathbb{E}[\xi_{b,i}^2]\Big\|(1+e^{W_b^{(i)}})(\bar{D}_{jn}-\bar{D}_{jn,i^*})\Big\|_1 \\
&+\sum_{i=1}^{n}\mathbb{E}[\xi_{b,i}^2]\sum_{i'=1}^{n}\mathbb{E}\Big[|\xi_{b,i'}|e^{W_b^{(i')}/2}\big(|\bar{D}_{jn}-\bar{D}_{jn}^{(i')}|+|\bar{D}_{jn,i^*}-\bar{D}_{jn,i^*}^{(i')}|\big)\Big]\Big]\Big\}
\end{aligned}
\tag{C.25}
$$

where we have used $\sum_{i=1}^{n}\|\xi_{b,i}\|_2^4\leq\sum_{i=1}^{n}\|\xi_{b,i}\|_2\|\xi_{b,i}\|_2^2\leq\sum_{i=1}^{n}\mathbb{E}[|\xi_{b,i}|^3]\leq\beta_2+\beta_3$ by (C.18). From (C.25), by defining

$$
W_b^{(i,i')}\equiv\begin{cases}W_b-\xi_{b,i}-\xi_{b,i'} & \text{if } i'\neq i\\ W_b-\xi_{b,i} & \text{if } i'=i\end{cases},
$$

with $e^{W_b^{(i')}/2} \le e^{1/2} e^{W_b^{(i,i')}/2}$, we further get

$$\begin{aligned}
|R_{12}| &\le C\Bigg\{\beta_2 + \beta_3 + \sum_{j=1}^{2}\Bigg[\sum_{i=1}^{n}\mathbb{E}[\xi_{b,i}^2]\|(1+e^{W_b^{(i)}})(\bar{D}_{jn} - \bar{D}_{jn}^{(i)} + \bar{D}_{jn}^{(i)} - \bar{D}_{jn,i^*})\|_1 \\
&\quad + \sum_{i=1}^{n}\mathbb{E}[\xi_{b,i}^2]\sum_{i'=1}^{n}\mathbb{E}\Big[|\xi_{b,i'}|e^{W_b^{(i,i')}/2}\big(|\bar{D}_{jn} - \bar{D}_{jn}^{(i')}| + |\bar{D}_{jn,i^*} - \bar{D}_{jn,i^*}^{(i')}|\big)\Big]\Bigg]\Bigg\} \\
&\le C\Bigg\{\beta_2 + \beta_3 + \sum_{j=1}^{2}\Bigg[\sum_{i=1}^{n}\mathbb{E}[\xi_{b,i}^2]\|(1+e^{W_b^{(i)}})(\bar{D}_{jn} - \bar{D}_{jn}^{(i)})\|_1 \\
&\quad + \sum_{i=1}^{n}\mathbb{E}[\xi_{b,i}^2]\sum_{i'=1}^{n}\mathbb{E}\Big[|\xi_{b,i'}|e^{W_b^{(i,i')}/2}|\bar{D}_{jn} - \bar{D}_{jn}^{(i')}|\Big]\Bigg]\Bigg\}, \qquad \text{(C.26)}
\end{aligned}$$

where we have used that

$$(e^{W_b^{(i)}}, \bar{D}_{jn} - \bar{D}_{jn}^{(i)}) =_d (e^{W_b^{(i)}}, \bar{D}_{jn,i^*} - \bar{D}_{jn}^{(i)}) \text{ and}$$
$$(|\xi_{b,i'}|e^{W_b^{(i,i')}/2}, \bar{D}_{jn} - \bar{D}_{jn}^{(i')}) =_d (|\xi_{b,i'}|e^{W_b^{(i,i')}/2}, \bar{D}_{jn,i^*} - \bar{D}_{jn,i^*}^{(i')})$$

to arrive at (C.26). Lastly, (C.26) can be further simplified as

$$\begin{aligned}
|R_{12}| \le C\Bigg\{\beta_2 + \beta_3 + \sum_{j=1}^{2}\sum_{i=1}^{n}\Bigg(&\mathbb{E}[\xi_{b,i}^2]\Big\|(1+e^{W_b^{(i)}})(\bar{D}_{jn} - \bar{D}_{jn}^{(i)})\Big\|_1 \\
&+\mathbb{E}\Big[|\xi_{b,i}|e^{W_b^{(i)}/2}|\bar{D}_{jn} - \bar{D}_{jn}^{(i)}|\Big]\Bigg)\Bigg\} \qquad \text{(C.27)}
\end{aligned}$$

using $e^{W_b^{(i,i')}/2} \le e^{(W_b^{(i')}+1)/2}$ and $\sum_{i=1}^{n}\mathbb{E}[\xi_{b,i}^2] \le \sum_{i=1}^{n}\mathbb{E}[\xi_i^2] = 1$ by (1.2). Combining (C.22) and (C.27) gives (C.13).

C.2.2 Bound for R_2

Since $|f_x'| \le 1$ by Lemma A.3,

$$\begin{aligned}
&|\sum_{i=1}^{n}\mathbb{E}[(\xi_i^2 - 1)I(|\xi_i| > 1)]\mathbb{E}[f_x'(W_b - \bar{\Delta}_{2n,x})]| \\
&\quad \le \sum_{i=1}^{n}\mathbb{E}[\xi_i^2 I(|\xi| > 1)] \le \beta_2. \qquad \text{(C.28)}
\end{aligned}$$

Moreover, by independence, Lemma A.1 and that $|f_x| \le 0.63$ from Lemma A.3,

$$\left|\sum_{i=1}^n \mathbb{E}[\xi_{b,i} f(W_b^{(i)} - \bar{\Delta}_{2n,x,i^*})]\right| = \left|\sum_{i=1}^n \mathbb{E}[\xi_{b,i}]\mathbb{E}[f(W_b^{(i)} - \bar{\Delta}_{2n,x,i^*})]\right|$$
$$\le 0.63 \sum_{i=1}^n |\mathbb{E}[\xi_{b,i}]| \le 0.63 \sum_{i=1}^n \mathbb{E}[\xi_i^2 I(|\xi_i| > 1)] = 0.63\beta_2.$$

Lastly, by $|f_x| \le 0.63$ and the definition of $\bar{\Delta}_{2n,x}$,

$$|\mathbb{E}[\bar{\Delta}_{2n,x} f_x(W_b)]| \le 0.63\|\bar{D}_{1n}\|_2 + \left|\frac{x}{2}\mathbb{E}[\bar{D}_{2n} f_x(W_b)]\right|$$

Hence, we established (C.14).

C.2.3 Bound for R_3

By mean value theorem, given $|f_x'| \le 1$ (Lemma A.3),

$$|f_x(W_b - \bar{\Delta}_{2n,x}) - f_x(W_b - \bar{\Delta}_{2n,x,i^*})| \le C|\bar{\Delta}_{2n,x} - \bar{\Delta}_{2n,x,i^*}|$$
$$\le C(|\bar{D}_{1n} - \bar{D}_{1n,i^*}| + x|\bar{D}_{2n} - \bar{D}_{2n,i^*}|).$$

Hence,

$$|R_3| \le C \sum_{j=1}^2 \sum_{i=1}^n \|\xi_{b,i}(\bar{D}_{jn} - \bar{D}_{jn,i^*})\|_1$$
$$= C \sum_{j=1}^2 \sum_{i=1}^n \|\xi_{b,i}(\bar{D}_{jn} - \bar{D}_{jn}^{(i)} + \bar{D}_{jn}^{(i)} - \bar{D}_{jn,i^*})\|_1 \text{ for } 0 \le x \le 1. \quad \text{(C.29)}$$

For $x > 1$, given $|\bar{\Delta}_{2n,x}| \vee |\bar{\Delta}_{2n,x,i^*}| \le \frac{1}{2} + \frac{x}{4}$, by (A.5) in Lemma A.4 and $|f_x'| \le 1$ (Lemma A.3),

$$|f_x(W_b - \bar{\Delta}_{2n,x}) - f_x(W_b - \bar{\Delta}_{2n,x,i^*})|$$
$$\le |f_x(W_b - \bar{\Delta}_{2n,x}) - f_x(W_b - \bar{\Delta}_{2n,x,i^*})|\Big[I(W_b \le 3x/4 - 3/2) + I(W_b > 3x/4 - 3/2)\Big]$$
$$\le C\Big(e^{1/2-x} + I(W_b > 3x/4 - 3/2)\Big)\Big(|\bar{D}_{1n} - \bar{D}_{1n,i^*}| + x|\bar{D}_{2n} - \bar{D}_{2n,i^*}|\Big)$$
$$\le C\Big(e^{-x} + e^{-3x/8}e^{W_b/2}\Big)\Big(|\bar{D}_{1n} - \bar{D}_{1n,i^*}| + x|\bar{D}_{2n} - \bar{D}_{2n,i^*}|\Big)$$
$$\le C\Big(e^{-x} + e^{-3x/8}e^{W_b^{(i)}/2}\Big)\Big(|\bar{D}_{1n} - \bar{D}_{1n,i^*}| + x|\bar{D}_{2n} - \bar{D}_{2n,i^*}|\Big),$$

where we have used $e^{W_b/2} \le e^{1/2} e^{W_b^{(i)}/2}$ in the last inequality. Hence,

$$|R_3| \le C \sum_{j=1}^{2} \sum_{i=1}^{n} \|\xi_{b,i}(1 + e^{W_b^{(i)}/2})(\bar{D}_{jn} - \bar{D}_{jn}^{(i)} + \bar{D}_{jn}^{(i)} - \bar{D}_{jn,i^*})\|_1 \text{ for } x > 1 \tag{C.30}$$

Because $(\xi_{b,i}, W_b^{(i)}, \bar{D}_{jn} - \bar{D}_{jn}^{(i)}) =_d (\xi_{b,i}, W_b^{(i)}, \bar{D}_{jn,i^*} - \bar{D}_{jn}^{(i)})$, (C.29) and (C.30) establish (C.15).

C.2.4 Bound for R_4

Using that $|f_x'| \le 1$ in Lemma A.3, for $0 \le x \le 1$,

$$\mathbb{E}\Big[\bar{\Delta}_{2n,x} \int_0^{-\bar{\Delta}_{2n,x}} f_x'(W_b + t)\mathrm{d}t\Big] \le C\bar{\Delta}_{2n,x}^2 \le C(\|\bar{D}_{1n}\|_2^2 + \|\bar{D}_{2n}\|_2^2) \le C(\|\bar{D}_{1n}\|_2 + \|\bar{D}_{2n}\|_2^2).$$

For $x > 1$, using (A.5) in Lemma A.4 and that $|f_x'| \le 1$ in Lemma A.3, given $|\bar{\Delta}_{2n,x}| \le \frac{1}{2} + \frac{x}{4}$

$$\begin{aligned}
&\mathbb{E}\Big[\bar{\Delta}_{2n,x} \int_0^{-\bar{\Delta}_{2n,x}} f_x'(W_b + t)\mathrm{d}t\Big] \\
&\le e^{1/2-x}\mathbb{E}[\bar{\Delta}_{2n,x}^2] + \mathbb{E}[I(W_b \ge 3x/4 - 3/2)\bar{\Delta}_{2n,x}^2] \\
&\le C(e^{-x}\mathbb{E}[\bar{\Delta}_{2n,x}^2] + e^{-3x/4}\mathbb{E}[e^{W_b}\bar{\Delta}_{2n,x}^2]) \\
&\le C\Big\{2e^{-x}\Big(\|\bar{D}_{1n}\|_2^2 + \frac{x^2}{4}\|\bar{D}_{2n}\|_2^2\Big) + 2e^{-3x/4}\mathbb{E}\Big[e^{W_b}\Big(\bar{D}_{1n}^2 + \frac{x^2}{4}\bar{D}_{2n}^2\Big)\Big]\Big\} \\
&\le C(\|\bar{D}_{1n}\|_2 + \mathbb{E}[(1 + e^{W_b})\bar{D}_{2n}^2]),
\end{aligned}$$

where we have used $\mathbb{E}[e^{W_b}|\bar{D}_{1n}|^2] \le \mathbb{E}[e^{W_b}|\bar{D}_{1n}|] \le \|e^{W_b}\|_2\|\bar{D}_{1n}\|_2 \le C\|\bar{D}_{1n}\|_2$ by Lemma A.2 and $\|\bar{D}_{1n}\|_2^2 \le \|\bar{D}_{1n}\|_2$. This establishes (C.16).

Appendix D. Proof of Theorem 2.3

We first verify (2.8)–(2.10), which will also be used in the proof of Theorem 2.3; (2.10) is immediate from (2.7). We can prove (2.8) with Hölder's inequality as

$$\begin{aligned}
\|(1 + e^{W_b^{(i)}})(\bar{D}_{1n} - \bar{D}_{1n}^{(i)})\|_1 &\le \|1 + e^{W_b^{(i)}}\|_2\|\bar{D}_{1n} - \bar{D}_{1n}^{(i)}\|_2 \\
&\le \Big(1 + \exp(e^4/8 - 1/8 + 1/2)\Big)\Big\|D_{1n} - D_{1n}^{(i)}\Big\|_2,
\end{aligned}$$

where we have also used Bennett's inequality (Lemmas A.2) and (2.6) at the end. Similarly, (2.9) can be proved as

$$\begin{aligned}
&\|\xi_{b,i}(1+e^{W_b^{(i)}/2})(\bar{D}_{1n}-\bar{D}_{1n}^{(i)})\|_1\\
&\quad\le \|\xi_{b,i}(1+e^{W_b^{(i)}/2})\|_2\|\bar{D}_{1n}-\bar{D}_{1n}^{(i)}\|_2\\
&\quad= \|\xi_{b,i}\|_2\|1+e^{W_b^{(i)}/2}\|_2\|\bar{D}_{1n}-\bar{D}_{1n}^{(i)}\|_2\\
&\quad\le \Big(1+\exp(e^2/8-1/8+1/4)\Big)\|\xi_i\|_2\Big\|D_{1n}-D_{1n}^{(i)}\Big\|_2,
\end{aligned}$$

where we have also used the independence of $e^{W_b^{(i)}}$ and $\xi_{b,i}$.

Our next task is to bound the other terms in the general bound of Theorem 2.1. Let

$$\bar{\Pi}_k = \Pi_k I(|\Pi_k| \le 1) + I(\Pi_k > 1) - I(\Pi_k < -1) \text{ for } k = 1, 2.$$

Since $|D_{2n}| \le |\Pi_1| + |\Pi_2|$, and $|\bar{D}_{2n}|$ is precisely $|D_{2n}|$ as a non-negative random variable upper-censored at $1/2$, it must be that $|\bar{D}_{2n}| \le |\bar{\Pi}_1| + |\bar{\Pi}_2|$, which further implies

$$\bar{D}_{2n}^2 \le 2(\bar{\Pi}_1^2 + \bar{\Pi}_2^2). \tag{D.1}$$

From (D.1) and $\bar{\Pi}_2^2 \le |\bar{\Pi}_2|$, we can get

$$\mathbb{E}[\bar{D}_{2n}^2] \le 2(\|\Pi_1\|_2^2 + \|\Pi_2\|_2) \tag{D.2}$$

On the other hand, define

$$D_{2n}^{(i)} = \max\Bigg(-1, \sum_{1\le i'\le n, i'\ne i}(\xi_{b,i'}^2 - \mathbb{E}[\xi_{b,i'}^2]) + \Pi_2^{(i)}\Bigg).$$

By Property 2.2(i), one can then write

$$\begin{aligned}
\|(1+e^{W_b^{(i)}})(\bar{D}_{2n}-\bar{D}_{2n}^{(i)})\|_1 &\le \|(1+e^{W_b^{(i)}})(\xi_{b,i}^2-\mathbb{E}[\xi_{b,i}^2])\|_1 + \|(1+e^{W_b^{(i)}})(\Pi_2-\Pi_2^{(i)})\|_1\\
&\le \|1+e^{W_b^{(i)}}\|_3\|\xi_{b,i}^2-\mathbb{E}[\xi_{b,i}^2]\|_{3/2} + \|1+e^{W_b^{(i)}}\|_2\|\Pi_2-\Pi_2^{(i)}\|_2\\
&\le C\Big((\mathbb{E}[|\xi_i|^3])^{2/3} + \|\Pi_2-\Pi_2^{(i)}\|_2\Big)
\end{aligned} \tag{D.3}$$

and

$$\begin{aligned}
&\|\xi_{b,i}(1+e^{W_b^{(i)}/2})(\bar{D}_{2n}-\bar{D}_{2n}^{(i)})\|_1\\
&\quad\le \|\xi_{b,i}(1+e^{W_b^{(i)}/2})(\xi_{b,i}^2-\mathbb{E}[\xi_{b,i}^2])\|_1 + \|\xi_{b,i}(1+e^{W_b^{(i)}/2})(\Pi_2-\Pi_2^{(i)})\|_1\\
&\quad\le \|\xi_{b,i}\|_3\|1+e^{W_b^{(i)}}\|_3\|\xi_{b,i}^2-\mathbb{E}[\xi_{b,i}^2]\|_{3/2} + \|\xi_{b,i}\|_2\|1+e^{W_b^{(i)}}\|_2\|\Pi_2-\Pi_2^{(i)}\|_2\\
&\quad\le C\Big(\mathbb{E}[|\xi_i|^3] + \|\xi_i\|_2\|\Pi_2-\Pi_2^{(i)}\|_2\Big),
\end{aligned} \tag{D.4}$$

where we have applied Bennett's inequality (Lemma A.2) to both (D.3) and (D.4) at the end. To complete the proof, it suffices to show the bounds

$$\mathbb{E}[e^{W_b}\bar{D}_{2n}^2] \leq C\left\{\sum_{i=1}^n \|\xi_{b,i}\|_3^3 + \|\Pi_2\|_2\right\} \tag{D.5}$$

and

$$\sup_{x\geq 0}|x\mathbb{E}[\bar{D}_{2n} f_x(W_b)]| \leq C\Big(\|\Pi_1\|_2^2 + \sum_{i=1}^n \|\xi_{b,i}\|_3^3 + \|\Pi_2\|_2\Big), \tag{D.6}$$

because Theorem 2.3 is then just a corollary of Theorem 2.1 by collecting (2.8)–(2.10), (D.2)–(D.6), as well as the simple facts

$$\beta_2 + \beta_3 \leq \sum_{i=1}^n \mathbb{E}[|\xi_i|^3], \qquad \mathbb{E}[|\xi_{b,i}|^2] \leq \|\xi_{b,i}\|_2 \leq \|\xi_i\|_2 \leq \|\xi_i\|_3,$$

$$P(|D_{1n}| > 1/2) \leq 2\|D_{1n}\|_2, \quad \|\Pi_1\|_2^2 \leq \sum_{i=1}^n \mathbb{E}[\xi_{b,i}^4] \leq \sum_{i=1}^n \mathbb{E}[|\xi_{b,i}|^3] \leq \sum_{i=1}^n \mathbb{E}[|\xi_i|^3],$$

and

$$\begin{aligned} P(|D_{2n}| > 1/2) &\leq P(|\Pi_1| + |\Pi_2| > 1/2) \\ &\leq P(|\Pi_1| > 1/4) + P(|\Pi_2| > 1/4) \\ &\leq C(\|\Pi_1\|_2^2 + \|\Pi_2\|_2). \end{aligned}$$

D.1 Proof of (D.5).

First, letting $W_b^{(i,j)} \equiv W_b - \xi_{b,i} - \xi_{b,j}$ for $1 \leq i \neq j \leq n$, we have

$$\begin{aligned} \mathbb{E}[\Pi_1^2 e^{W_b}] &= \sum_{i=1}^n \mathbb{E}[(\xi_{b,i}^2 - \mathbb{E}[\xi_{b,i}^2])^2 e^{\xi_{b,i}}]\mathbb{E}[e^{W_b^{(i)}}] \\ &\quad + \sum_{1\leq i\neq j\leq n} \mathbb{E}[(\xi_{b,i}^2 - \mathbb{E}[\xi_{b,i}^2])e^{\xi_{b,i}}]\mathbb{E}[(\xi_{b,j}^2 - \mathbb{E}[\xi_{b,j}^2])e^{\xi_{b,j}}]\mathbb{E}[e^{W_b^{(i,j)}}] \\ &= \sum_{i=1}^n \mathbb{E}[(\xi_{b,i}^2 - \mathbb{E}[\xi_{b,i}^2])^2 e^{\xi_{b,i}}]\mathbb{E}[e^{W_b^{(i)}}] \\ &\quad + \sum_{1\leq i\neq j\leq n} \mathbb{E}[(\xi_{b,i}^2 - \mathbb{E}[\xi_{b,i}^2])(e^{\xi_{b,i}} - 1)]\mathbb{E}[(\xi_{b,j}^2 - \mathbb{E}[\xi_{b,j}^2])(e^{\xi_{b,j}} - 1)]\mathbb{E}[e^{W_b^{(i,j)}}] \\ &\leq C\left(\sum_{i=1}^n \mathbb{E}[\xi_{b,i}^4] + \sum_{1\leq i\neq j\leq n} \mathbb{E}\Big[|\xi_{b,i}^2 - \mathbb{E}[\xi_{b,i}^2]||\xi_{b,i}|\Big]\mathbb{E}\Big[|\xi_{b,j}^2 - \mathbb{E}[\xi_{b,j}^2]||\xi_{b,j}|\Big]\mathbb{E}\Big[e^{W_b^{(i,j)}}\Big]\right) \end{aligned}$$

$$\leq C\left\{\sum_{i=1}^{n}\|\xi_{b,i}\|_3^3+\sum_{1\leq i\neq j\leq n}\|\xi_{b,i}\|_3^3\|\xi_{b,j}\|_2^2\right\}\leq C\sum_{i=1}^{n}\|\xi_{b,i}\|_3^3 \tag{D.7}$$

by Lemma A.2 that $|e^s-1|\leq |s|(e^a-1)/a$ for $s\leq a$ and $a>0$,

$$\begin{aligned}&\mathbb{E}[|\xi_{b,i}^2-\mathbb{E}[\xi_{b,i}^2]||\xi_{b,i}|]\\&\leq\left\{(\|\xi_{b,i}^2-\mathbb{E}[\xi_{b,i}^2]\|_{3/2}\|\xi_{b,i}\|_3)\wedge\mathbb{E}[|\xi_{b,i}^2-\mathbb{E}[\xi_{b,i}^2]|]\right\}\\&\leq 2\left\{\|\xi_{b,i}\|_3^3\wedge\|\xi_{b,i}\|_2^2\right\}\text{ for any } i=1,\dots,n,\end{aligned}$$

and $\sum_{j=1}^{n}\|\xi_{b,i}\|_3^3\|\xi_{b,j}\|_2^2\leq\|\xi_{b,i}\|_3^3$. Second, by Lemma A.2,

$$\mathbb{E}[\bar{\Pi}_2^2e^{W_b}]\leq\mathbb{E}[\bar{\Pi}_2^4]^{1/2}(\mathbb{E}[e^{2W_b}])^{1/2}\leq C\mathbb{E}[\Pi_2^2]^{1/2}=C\|\Pi_2\|_2 \tag{D.8}$$

Combining (D.1), (D.7) and (D.8) gives (D.5).

D.2 Proof of (D.6).

Since $\sup_{x\geq 0}|xf_x(w)|\leq C$ (which uses (A.4) in Lemma A.4 and that $|f_x|\leq 0.63$ in Lemma A.3),

$$\begin{aligned}\sup_{x\geq 0}|x\mathbb{E}[(D_{2n}-\bar{D}_{2n})f_x(W_b)]|&\leq\sup_{x\geq 0}x\mathbb{E}[(|D_{2n}|-1/2)|f_x(W_b)|I(|D_{2n}|>1/2)]\\&\leq C\mathbb{E}[|D_{2n}|I(|D_{2n}|>1/2)]\\&\leq C\Big(\mathbb{E}[|\Pi_1|I(|D_{2n}|>1/2)]+\mathbb{E}[|\Pi_2|]\Big)\\&\leq C\Big(\mathbb{E}\Big[|\Pi_1|\Big\{I(|\Pi_1|>1/4)+I(|\Pi_2|>1/4)\Big\}\Big]+\mathbb{E}[|\Pi_2|]\Big)\\&\leq C\Big(\mathbb{E}[4\Pi_1^2+2|\Pi_1||\Pi_2|^{1/2}]+\mathbb{E}[|\Pi_2|]\Big)\\&\leq C\Big(\mathbb{E}[5\Pi_1^2+|\Pi_2|]+\mathbb{E}[|\Pi_2|]\Big)\\&\leq C\Big(\|\Pi_1\|_2^2+\|\Pi_2\|_2\Big),\end{aligned}$$

where we have used that $I(|\Pi_1|>1/4)\leq 4|\Pi_1|$, $I(|\Pi_2|>1/4)\leq 2|\Pi_2|^{1/2}$ and $2|\Pi_1||\Pi_2|^{1/2}\leq|\Pi_1|^2+|\Pi_2|$. Noting that

$$x\mathbb{E}[\bar{D}_{2n}f_x(W_b)]=x\mathbb{E}[(\bar{D}_{2n}-D_{2n})f_x(W_b)]+x\mathbb{E}[D_{2n}f_x(W_b)],$$

the above implies

$$\sup_{x\geq 0}\Big|x\mathbb{E}[\bar{D}_{2n}f_x(W_b)]\Big|\leq C\Big(\|\Pi_1\|_2^2+\|\Pi_2\|_2\Big)+\sup_{x\geq 0}\Big|x\mathbb{E}[D_{2n}f_x(W_b)]\Big|, \tag{D.9}$$

so for the rest of this section we focus on bounding $\sup_{x\geq 0}\left|x\mathbb{E}[D_{2n}f_x(W_b)]\right|$. From the form of D_{2n} in (2.12), by defining $\Pi = \Pi_1 + \Pi_2$, we have

$$x\mathbb{E}[D_{2n}f_x(W_b)] = \mathbb{E}[x\Pi f_x(W_b)] - \mathbb{E}[xf_x(W_b)I(\Pi < -1)(1+\Pi)],$$

so it suffices to establish

$$\begin{aligned}&\left|\mathbb{E}[x\Pi f_x(W_b)]\right| \vee \left|\mathbb{E}[xf_x(W_b)I(\Pi < -1)(1+\Pi)]\right|\\ &\leq C\left(\sum_{i=1}^n \mathbb{E}[|\xi_{b,i}|^3] + \|\Pi_2\|_2\right) \text{ for all } x \geq 0.\end{aligned} \tag{D.10}$$

We first bound $\left|\mathbb{E}[xf_x(W_b)I(\Pi < -1)(1+\Pi)]\right|$. Since

$$\begin{aligned}\mathbb{E}[xf_x(W_b)I(\Pi < -1)(1+\Pi)] = &\ \mathbb{E}[xf_x(W_b)I(\Pi < -1)]\\ &+\mathbb{E}[xf_x(W_b)\Pi I(\Pi < -1)],\end{aligned} \tag{D.11}$$

we will bound the two terms on the right hand side separately. As $xf_x(w)$ is bounded for all $x \geq 0$ (Lemma A.3 and (A.4) in Lemma A.4), we have

$$\begin{aligned}\left|\mathbb{E}[xf_x(W_b)I(\Pi < -1)]\right| &\leq \mathbb{E}\Big[|xf_x(W_b)|I(\Pi < -1)\Big]\\ &\leq C\sum_{j=1}^2 P(\Pi_j < -1/2) \leq C\Big(\|\Pi_1\|_2^2 + \|\Pi_2\|_2\Big)\end{aligned}$$

and

$$\begin{aligned}\left|\mathbb{E}[xf_x(W_b)\Pi I(\Pi < -1)]\right| &\leq C\mathbb{E}[|\Pi|I(\Pi < -1)]\\ &\leq C\Big(\mathbb{E}[|\Pi_1|I(\Pi < -1)] + \|\Pi_2\|_2\Big)\\ &\leq C\left(\|\Pi_1\|_2\sqrt{\sum_{j=1}^2 P(\Pi_j < -1/2)} + \|\Pi_2\|_2\right)\\ &\leq C\Big(\|\Pi_1\|_2\sqrt{\|\Pi_1\|_2^2 + \|\Pi_2\|_2} + \|\Pi_2\|_2\Big)\\ &\leq C\Big(\|\Pi_1\|_2^2 + \|\Pi_1\|_2\sqrt{\|\Pi_2\|_2} + \|\Pi_2\|_2\Big)\\ &\leq C\Big(\|\Pi_1\|_2^2 + \|\Pi_2\|_2\Big),\end{aligned}$$

where the second last inequality uses $\sqrt{\|\Pi_1\|_2^2 + \|\Pi_2\|_2} \le \|\Pi_1\|_2 + \sqrt{\|\Pi_2\|_2}$ and the last inequality uses that $2|ab| \le a^2 + b^2$ for any $a, b \in \mathbb{R}$. So the part of (D.10) regarding $\left|\mathbb{E}[x f_x(W_b) I(\Pi < -1)(1 + \Pi)]\right|$ is proved because $\|\Pi_1\|_2^2 = \sum_{i=1}^n (\mathbb{E}[\xi_{b,i}^4] - (\mathbb{E}[\xi_{b,i}^2])^2) \le \sum_{i=1}^n \mathbb{E}[|\xi_{b,i}|^3]$.

Next we bound $\left|\mathbb{E}[x\Pi f_x(W_b)]\right|$, and we will control the two terms on the right-hand side of

$$|\mathbb{E}[x\Pi f_x(W_b)]| \le x|\mathbb{E}[\Pi_1 f_x(W_b)]| + x|\mathbb{E}[\Pi_2 f_x(W_b)]|. \tag{D.12}$$

For the first term $x|\mathbb{E}[\Pi_1 f_x(W_b)]|$, we write

$$\begin{aligned}
\left|\mathbb{E}[\Pi_1 f_x(W_b)]\right| &= \left|\sum_{i=1}^n \mathbb{E}\Big[(\xi_{b,i}^2 - \mathbb{E}[\xi_{b,i}^2])(f_x(W_b) - f_x(W_b^{(i)}))\Big]\right| \\
&= \left|\sum_{i=1}^n \mathbb{E}\Big[(\xi_{b,i}^2 - \mathbb{E}[\xi_{b,i}^2]) \int_0^{\xi_{b,i}} \mathbb{E}[f_x'(W_b^{(i)} + t)]\mathrm{d}t\Big]\right| \\
&\le \sum_{i=1}^n \mathbb{E}\Big[(\xi_{b,i}^2 + \mathbb{E}[\xi_{b,i}^2]) \int_0^{|\xi_{b,i}|} |\mathbb{E}[f_x'(W_b^{(i)} + t)]|\mathrm{d}t\Big],
\end{aligned} \tag{D.13}$$

where the second equality uses the independence of $W_b^{(i)}$ and $\xi_{b,i}$. From (D.13) and Lemma A.5, for any $x \ge 1$, we have that

$$\begin{aligned}
\left|\mathbb{E}[\Pi_1 f_x(W_b)]\right| &\le C\sum_{i=1}^n \mathbb{E}\Big[(\xi_{b,i}^2 + \mathbb{E}[\xi_{b,i}^2]) \int_0^{|\xi_{b,i}|} (e^{-x} + e^{-x+t})\mathrm{d}t\Big] \\
&\le C\sum_{i=1}^n \mathbb{E}\Big[(\xi_{b,i}^2 + \mathbb{E}[\xi_{b,i}^2]) \int_0^{|\xi_{b,i}|} (e^{-x} + e^{-x+1})\mathrm{d}t\Big] \text{ (as } |\xi_{b,i}| \le 1) \\
&\le Ce^{-x}\sum_{i=1}^n \Big(\mathbb{E}[|\xi_{b,i}|^3] + \mathbb{E}[|\xi_{b,i}|^2]\mathbb{E}[|\xi_{b,i}|]\Big) \\
&\le Ce^{-x}\sum_{i=1}^n \mathbb{E}[|\xi_{b,i}|^3],
\end{aligned}$$

which implies

$$\sup_{x\ge 1} x\left|\mathbb{E}[\Pi_1 f_x(W_b)]\right| \le C\sum_{i=1}^n \mathbb{E}[|\xi_{b,i}|^3]. \tag{D.14}$$

Moreover, for $0 \le x < 1$, since $|f_x'| \le 1$ (Lemma A.3), from (D.13) we get

$$\sup_{0\le x<1} x\left|\mathbb{E}[\Pi_1 f_x(W_b)]\right| \le \sum_{i=1}^n \left(\mathbb{E}[|\xi_{b,i}|^3] + \mathbb{E}[|\xi_{b,i}|^2]\mathbb{E}[|\xi_{b,i}|]\right)$$
$$\le 2\sum_{i=1}^n \mathbb{E}[|\xi_{b,i}|^3]. \tag{D.15}$$

For the term $x|\mathbb{E}[\Pi_2 f_x(W_b)]|$, given that $\sup_{x\ge 0}|xf_x(w)| \le C$ for all w (explained at the beginning of Sect. D.2), we have

$$\sup_{x\ge 0} x|\mathbb{E}[\Pi_2 f_x(W_b)]| \le \sup_{x\ge 0}\mathbb{E}[|\Pi_2||xf_x(W_b)|] \le C\|\Pi_2\|_1 \le C\|\Pi_2\|_2, \tag{D.16}$$

Combining (D.12) and (D.14)–(D.16) proves the part of (D.10) regarding $|\mathbb{E}[x\Pi f_x(W_b)]|$.

Appendix E. Proof of Lemma 3.3

In this section, we adopt the following notation: For any natural numbers $k' \le k$, we denote $[k' : k] \equiv \{k', \ldots, k\}$ and $[k] \equiv \{1, \ldots, k\}$. Moreover, for any natural number $k \ge 1$, we let

$$\bar{h}_{k,\{i_1,\ldots,i_k\}} \equiv \bar{h}_k(X_{i_1}, \ldots, X_{i_k})$$

with respect to the function $\bar{h}_k(\cdot)$ in (3.9). To prove Lemma 3.3, we need the following technical lemmas proven, respectively, in Appendices F.1 and F.2.

Lemma E.1 (Useful kernel bounds) *Under assumptions (3.1)–(3.3),*

(i) For any $k \in [m]$,

$$\mathbb{E}[\bar{h}_k^2] \le \mathbb{E}[h_k^2] \le \frac{k}{m}\mathbb{E}[h^2]$$

(ii) For any $i \in [n]$,

$$\mathbb{E}\left[\left(\sum_{\substack{1\le i_1<\cdots<i_{m-1}\le n\\ i_l\ne i \text{ for } l\in[m-1]}} \bar{h}_m(X_i, X_{i_1}, \ldots, X_{i_{m-1}})\right)^2\right]$$
$$\le \frac{2(m-1)^2}{n(n-m+1)}\binom{n-1}{m-1}\binom{n}{m}\mathbb{E}[h^2];$$

(iii) For each $i \in [n]$, consider $\xi_{b,i}$ defined in (2.1) with ξ_i defined in (3.8). Given $k_1, k_2 \in [m]$, for any $1 \le i_1 < \cdots < i_{k_1} \le n$ and $1 \le j_1 < \cdots < j_{k_2} \le n$, we

have

$$\left|\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{k_1,\{i_1,\dots,i_{k_1}\}}\bar{h}_{k_2,\{j_1,\dots,j_{k_2}\}}]\right| \leq \frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \frac{2d\|h\|_2}{n}$$

where

$$d = |(\{i_1,\dots,i_{k_1}\} \cap \{j_1,\dots,j_{k_2}\})\backslash\{1,2\}|,$$

the number of elements in the intersection of $\{i_1,\dots,i_{k_1}\}$ *and* $\{j_1,\dots,j_{k_2}\}$ *that are not* 1 *or* 2*.*

(iv) If, in addition to all the conditions in (iii)*, it is true that* $1 \notin \{j_1,\dots,j_{k_2}\}$ *and* $2 \notin \{i_1,\dots,i_{k_1}\}$*, then we have the bound*

$$\left|\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{k_1,\{i_1,\dots,i_{k_1}\}}\bar{h}_{k_2,\{j_1,\dots,j_{k_2}\}}]\right| \leq \frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \frac{2d\|h\|_2}{n^{3/2}}$$

Lemma E.2 (Counting identities and bounds) *Let* m, n *be non-negative integers such that* $m \leq n$*.*

(i) Suppose n_1 *and* n_2 *are non-negative integers such that* $n_1 + n_2 = n$*. Then*

$$\sum_{k=0}^{m}\binom{n_1}{k}\binom{n_2}{m-k} = \binom{n}{m}.$$

(ii) Suppose k *is a non-negative integer such that* $k \leq m$*. Then*

$$\binom{n}{k}\binom{n-k}{m-k} = \binom{n}{m}\binom{m}{k}.$$

(iii) For positive integers a, b, e *such that* $b + e \leq a$*, we have*

$$\binom{a}{b} - \binom{a-e}{b} \leq \binom{a}{b}\frac{be}{a-b+1}.$$

In addition to the lemmas above, we will make use of the following enumerative equalities, whenever the binomial coefficients involved are well defined:

$$\binom{n-2}{m-1} = \binom{n-1}{m-1}\frac{n-m}{n-1}, \tag{E.1}$$

$$\binom{n-2}{m-2} = \binom{n-1}{m-1}\frac{m-1}{n-1}, \tag{E.2}$$

$$\binom{n-3}{m-2} = \binom{n-1}{m-1}\frac{(m-1)(n-m)}{(n-1)(n-2)} \tag{E.3}$$

$$\binom{n-3}{m-3} = \binom{n-1}{m-1} \frac{(m-1)(m-2)}{(n-1)(n-2)} \quad \text{, and} \tag{E.4}$$

$$\binom{n-4}{m-4} = \binom{n-1}{m-1} \frac{(m-1)(m-2)(m-3)}{(n-1)(n-2)(n-3)}. \tag{E.5}$$

E.1 Proof of Lemma 3.3(*i*)

We shall further let

$$\Pi_{21} \equiv (n^{-1/2}|0) - \sum_{i=1}^{n} \mathbb{E}[(\xi_i^2 - 1)I(|\xi_i| > 1)] \text{ and}$$

$$\Pi_{22} \equiv \delta_{2n,b} = \frac{2(n-1)}{(n-m)} \binom{n-1}{m-1}^{-1} \sum_{i=1}^{n} \xi_{b,i} \Psi_{n,i}, \tag{E.6}$$

so $\Pi_2 = \Pi_{21} + \Pi_{22}$. It suffices to show these bounds for Π_{21} and Π_{22} in (E.6):

$$\|\Pi_{21}\|_2^2 \leq C \left(\frac{\|g\|_3^6}{n} + \frac{1}{n} \right) \leq C \frac{\|g\|_3^6}{n}. \tag{E.7}$$

$$\|\Pi_{22}\|_2^2 \leq C \frac{m^2 \|g\|_3^2 \|h\|_3^2}{n} \tag{E.8}$$

From there, since $\|\Pi_2\|_2 \leq \|\Pi_{21}\|_2 + \|\Pi_{22}\|_2$, Lemma 3.3($i$) is proved.

E.1.1 Proof of (E.7)

We first note that

$$\sum_{i=1}^{n} \mathbb{E}\Big[(\xi_i^2 - 1)I(|\xi_i| > 1)\Big] \leq \sum_{i=1}^{n} \mathbb{E}\Big[\xi_i^2 I(|\xi_i| > 1)\Big] \leq \sum_{i=1}^{n} \mathbb{E}[|\xi_i|^3] = \mathbb{E}[|g|^3]/\sqrt{n},$$

which gives $(\sum_{i=1}^{n} \mathbb{E}[(\xi_i^2 - 1)I(|\xi_i| > 1)])^2 \leq (\mathbb{E}[|g|^3])^2/n$, and hence (E.7).

E.1.2 Proof of (E.8)

It is trivial for $m = 1$ since $\Psi_{n,i} = 0$. For $m \geq 2$, first write

$$\Pi_{22}^2 = \frac{4(n-1)^2}{(n-m)^2 n} \binom{n-1}{m-1}^{-2} \left(\sum_{i=1}^{n} \xi_{b,i} \sum_{\substack{1 \leq i_1 < \cdots < i_{m-1} \leq n \\ i_l \neq i \text{ for } l \in [m-1]}} \bar{h}_m(X_i, X_{i_1}, \ldots, X_{i_{m-1}}) \right)^2,$$

which implies immediately from $2m < n$ in (3.2) that

$$\mathbb{E}\left[\Pi_{22}^2\right] \leq \frac{16}{n}\binom{n-1}{m-1}^{-2}\mathbb{E}\left[\left(\sum_{i=1}^n \xi_{b,i}\sum_{\substack{1\leq i_1<\cdots<i_{m-1}\leq n\\ i_l\neq i \text{ for } l\in[m-1]}}\bar{h}_m(X_i, X_{i_1}, \ldots, X_{i_{m-1}})\right)^2\right]. \tag{E.9}$$

Upon expanding the above expectation,

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{i=1}^n \xi_{b,i}\sum_{\substack{1\leq i_1<\cdots<i_{m-1}\leq n\\ i_l\neq i \text{ for } l\in[m-1]}}\bar{h}_{m,\{i,i_1,\ldots,i_{m-1}\}}\right)^2\right]\\
&= \sum_{i=1}^n \mathbb{E}\left[\left(\xi_{b,i}\sum_{\substack{1\leq i_1<\cdots<i_{m-1}\leq n\\ i_l\neq i \text{ for } l\in[m-1]}}\bar{h}_{m,\{i,i_1,\ldots,i_{m-1}\}}\right)^2\right]\\
&\quad + \sum_{1\leq i\neq j\leq n}\mathbb{E}\left[\left(\xi_{b,i}\sum_{\substack{1\leq i_1<\cdots<i_{m-1}\leq n\\ i_l\neq i \text{ for } l\in[m-1]}}\bar{h}_{m,\{i,i_1,\ldots,i_{m-1}\}}\right)\right.\\
&\quad \times \left.\left(\xi_{b,j}\sum_{\substack{1\leq j_1<\cdots<j_{m-1}\leq n\\ j_l\neq j \text{ for } l\in[m-1]}}\bar{h}_{m,\{j,j_1,\ldots,j_{m-1}\}}\right)\right]\\
&= n\mathbb{E}\left[\left(\xi_{b,1}\sum_{\substack{1\leq i_1<\cdots<i_{m-1}\leq n\\ i_l\neq 1 \text{ for } l=1,\ldots,m-1}}\bar{h}_{m,\{1,i_1,\ldots,i_{m-1}\}}\right)^2\right] + \qquad (\text{E.10})\\
&n(n-1)\mathbb{E}\left[\left(\xi_{b,1}\sum_{\substack{1\leq i_1<\cdots<i_{m-1}\leq n\\ i_l\neq 1 \text{ for } l\in[m-1]}}\bar{h}_{m,\{1,i_1,\ldots,i_{m-1}\}}\right)\right.\\
&\left.\left(\xi_{b,2}\sum_{\substack{1\leq j_1<\cdots<j_{m-1}\leq n\\ j_l\neq 2 \text{ for } l\in[m-1]}}\bar{h}_{m,\{2,j_1,\ldots,j_{m-1}\}}\right)\right]. \qquad (\text{E.11})
\end{aligned}$$

We need to control the two expectations in (E.10) and (E.11). We first bound the expectation in (E.10). With the definition in (3.9) and that

$$\begin{aligned}
&\mathbb{E}[\bar{h}_{m,\{1,i_1,\ldots,i_{m-1}\}}\bar{h}_{m,\{1,j_1,\ldots,j_{m-1}\}}]\\
&\quad = \mathbb{E}[\bar{h}^2_{1,\{1\}}] = 0 \text{ if } |\{i_1,\ldots,i_{m-1}\}\cap\{j_1,\ldots,j_{m-1}\}| = 0,
\end{aligned}$$

we can write

$$
\begin{aligned}
&\mathbb{E}\left[\left(\xi_{b,1}\sum_{\substack{1\le i_1<\cdots<i_{m-1}\le n\\ i_l\ne 1 \text{ for } l\in[m-1]}}\bar{h}_{m,\{1,i_1,\ldots,i_{m-1}\}}\right)^2\right]\\
&=\mathbb{E}\left[\xi_{b,1}^2\sum_{k=0}^{m-1}\left(\sum_{\substack{1\le i_1<\cdots<i_{m-1}\le n\\ 1\le j_1<\cdots<j_{m-1}\le n\\ i_l,j_l\ne 1 \text{ for } l\in[m-1]\\ |\{i_1,\ldots,i_{m-1}\}\cap\{j_1,\ldots,j_{m-1}\}|=k}}\bar{h}_{m,\{1,i_1,\ldots,i_{m-1}\}}\bar{h}_{m,\{1,j_1,\ldots,j_{m-1}\}}\right)\right]\\
&=\sum_{k=1}^{m-1}\binom{n-1}{k}\binom{n-k-1}{m-k-1}\binom{n-m}{m-k-1}\mathbb{E}\left[\xi_{b,1}^2\bar{h}_{k+1}^2(X_1\ldots,X_{k+1})\right]\\
&\le\sum_{k=1}^{m-1}\binom{n-1}{k}\binom{n-k-1}{m-k-1}\binom{n-m}{m-k-1}\frac{k+1}{m}\mathbb{E}[h^2],
\end{aligned}
\tag{E.12}
$$

where the last inequality comes from Lemma E.1(i) and that $\xi_{b,1}^2\le 1$. Continuing from (E.12), we can get

$$
\begin{aligned}
&\mathbb{E}\left[\left(\xi_{b,1}\sum_{\substack{1\le i_1<\cdots<i_{m-1}\le n\\ i_l\ne 1 \text{ for } l\in[m-1]}}\bar{h}_{m,\{1,i_1,\ldots,i_{m-1}\}}\right)^2\right]\\
&\le\sum_{k=1}^{m-1}\binom{n-1}{k}\binom{n-k-1}{m-k-1}\binom{n-m}{m-k-1}\frac{k+1}{m}\mathbb{E}[h^2]\\
&=\frac{1}{m}\binom{n-1}{m-1}\sum_{k=1}^{m-1}\binom{m-1}{k}\binom{n-m}{m-k-1}(k+1)\mathbb{E}[h^2] \text{ by Lemma E.2}(ii)\\
&=\frac{m-1}{m}\binom{n-1}{m-1}\sum_{k=1}^{m-1}\binom{m-2}{k-1}\frac{k+1}{k}\binom{n-m}{m-1-k}\mathbb{E}[h^2]\\
&\le 2\binom{n-1}{m-1}\sum_{k=0}^{m-2}\binom{m-2}{k}\binom{n-m}{m-2-k}\mathbb{E}[h^2]\\
&=2\binom{n-1}{m-1}\binom{n-2}{m-2}\mathbb{E}[h^2] \text{ by Lemma E.2}(i)\\
&=2\frac{m-1}{n-1}\binom{n-1}{m-1}^2\mathbb{E}[h^2]
\end{aligned}
\tag{E.13}
$$

Now we bound the expectation in (E.11). First first expand it as

$$\mathbb{E}\Bigg[\Bigg(\xi_{b,1}\sum_{\substack{1\le i_1<\cdots<i_{m-1}\le n\\ i_l\neq 1 \text{ for } l\in[m-1]}}\bar{h}_{m,\{1,i_1,\ldots,i_{m-1}\}}\Bigg)\Bigg(\xi_{b,2}\sum_{\substack{1\le j_1<\cdots<j_{m-1}\le n\\ j_l\neq 2 \text{ for } l\in[m-1]}}\bar{h}_{m,\{2,j_1,\ldots,j_{m-1}\}}\Bigg)\Bigg]$$

$$= \binom{n-2}{m-1}\binom{n-2-(m-1)}{m-1}\Bigg(\underbrace{\mathbb{E}\left[\xi_{b,1}\bar{h}_{m,\{1,\ldots,m\}}\right]}_{=\mathbb{E}[\mathbb{E}[\xi_{b,1}\bar{h}_{m,\{1,\ldots,m\}}|X_1]]=\mathbb{E}[\xi_{b,1}\bar{h}_1(X_1)]=0}\Bigg)^2$$

$$+ 2\times\binom{n-2}{m-2}\binom{n-2-(m-2)}{m-1}\underbrace{\mathbb{E}\Big[\xi_{b,1}\xi_{b,2}\bar{h}_{m,\{1,2,\ldots,m\}}\bar{h}_{m,\{2,m+1,\ldots,2m-1\}}\Big]}_{=\mathbb{E}[\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{2,\{1,2\}}\bar{h}_{1,\{2\}}|X_1,X_2]]=0 \text{ since } \bar{h}_{1,\{2\}}=0}$$

$$+ 2\times\underbrace{\sum_{\substack{1\le i_1<\cdots<i_{m-2}\le n\\ 1\le j_1<\cdots<j_{m-1}\le n\\ i_l,j_v\neq 1,2, \text{ for } l\in[m-2],v\in[m-1]\\ |\{i_1,\ldots,i_{m-2}\}\cap\{j_1,\ldots,j_{m-1}\}|\ge 1}}\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{m,\{1,2,i_1,\ldots,i_{m-2}\}}\bar{h}_{m,\{2,j_1,\ldots,j_{m-1}\}}]}_{\equiv EA}$$

$$+ \underbrace{\sum_{\substack{1\le i_1<\cdots<i_{m-1}\le n\\ 1\le j_1<\cdots<j_{m-1}\le n\\ i_l,j_l\neq 1,2, \text{ for } l\in[m-1]\\ |\{i_1,\ldots,i_{m-1}\}\cap\{j_1,\ldots,j_{m-1}\}|\ge 1}}\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{m,\{1,i_1,\ldots,i_{m-1}\}}\bar{h}_{m,\{2,j_1,\ldots,j_{m-1}\}}]}_{\equiv EB}$$

$$+ \underbrace{\sum_{\substack{1\le i_1<\cdots<i_{m-2}\le n\\ 1\le j_1<\cdots<j_{m-2}\le n\\ i_l,j_l\neq 1,2, \text{ for } l\in[m-2]}}\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{m,\{1,2,i_1,\ldots,i_{m-2}\}}\bar{h}_{m,\{1,2,j_1,\ldots,j_{m-2}\}}]}_{\equiv EC}, \quad \text{(E.14)}$$

and will then bound each of EA, EB, and EC.

We start with EA, and it suffices to assume $m \geq 3$, otherwise one cannot expect the two sets $\{i_1,\ldots,i_{m-2}\}$ and $\{j_1,\ldots,j_{m-1}\}$ indexing a given summand

$$\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{m,\{1,2,i_1,\ldots,i_{m-2}\}}\bar{h}_{m,\{2,j_1,\ldots,j_{m-1}\}}]$$

of EA to intersect for at least one element. Using the fact that the data $X_1,\ldots,X_n$ are i.i.d., if the two index sets have $k \in [m-2]$ common elements not in the set $\{1,2\}$, one can write the summand as

$$\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{m,\{1,2,i_1,\ldots,i_{m-2}\}}\bar{h}_{m,\{2,j_1,\ldots,j_{m-1}\}}]$$
$$= \mathbb{E}[\xi_{b,1}\xi_{b,2}\,\bar{h}_m(X_1,X_2,\ldots,X_m)\,\bar{h}_m(X_2,X_3,\ldots,X_{k+2},X_{m+1},\ldots,X_{2m-1-k})].$$

From this, we can alternatively write

$$EA = \sum_{k=1}^{m-2} \binom{n-2}{k} \binom{n-2-k}{m-2-k} \binom{n-m}{m-1-k}$$
$$\mathbb{E}\Big[\xi_{b,1}\xi_{b,2}\ \bar{h}_{m,[1:m]}\ \bar{h}_{m,[2:(k+2)]\cup[(m+1):(2m-k-1)]}\Big];$$

from this, we can then form the bound

$$\begin{aligned}
|EA| &\leq \sum_{k=1}^{m-2} \binom{n-2}{k} \binom{n-2-k}{m-2-k} \binom{n-m}{m-1-k} \\
&\quad \Big|\mathbb{E}\Big[\xi_{b,1}\xi_{b,2}\ \bar{h}_{m,[1:m]}\ \bar{h}_{m,[2:(k+2)]\cup[(m+1):(2m-k-1)]}\Big]\Big| \\
&= \binom{n-2}{m-2} \sum_{k=1}^{m-2} \binom{m-2}{k} \binom{n-m}{m-1-k} \\
&\quad \Big|\mathbb{E}\Big[\xi_{b,1}\xi_{b,2}\ \bar{h}_{m,[1:m]}\ \bar{h}_{m,[2:(k+2)]\cup[(m+1):(2m-k-1)]}\Big]\Big| \\
&\qquad\qquad \text{by Lemma E.2}(ii) \\
&\leq \binom{n-2}{m-2} \sum_{k=1}^{m-2} \binom{m-2}{k} \binom{n-m}{m-1-k} \left\{ \frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \frac{2k\|h\|_2}{n} \right\} \\
&\qquad\qquad \text{by Lemma E.1}(iii) \\
&= \binom{n-2}{m-2} \left\{ \left[\binom{n-2}{m-1} - \binom{n-m}{m-1} \right] \frac{9.5\|g\|_3^2\|h\|_3^2}{n} + (m-2)\binom{n-3}{m-2} \frac{2\|h\|_2}{n} \right\},
\end{aligned}$$

where the last line comes from the equalities

$$\begin{aligned}
\sum_{k=1}^{m-2} \binom{m-2}{k} \binom{n-m}{m-1-k} &= \sum_{k=0}^{m-2} \binom{m-2}{k} \binom{n-m}{m-1-k} - \binom{n-m}{m-1} \\
&= \binom{n-2}{m-1} - \binom{n-m}{m-1} \text{ by Lemma E.2}(i)
\end{aligned}$$

and

$$\begin{aligned}
\sum_{k=1}^{m-2} k \binom{m-2}{k} \binom{n-m}{m-1-k} &= (m-2) \sum_{k=1}^{m-2} \binom{m-3}{k-1} \binom{n-m}{m-1-k} \\
&= (m-2) \sum_{k=0}^{m-3} \binom{m-3}{k} \binom{n-m}{m-2-k} \\
&= (m-2) \binom{n-3}{m-2} \text{ coming from Lemma E.2}(i)
\end{aligned}$$

Continuing, we get

$$
\begin{aligned}
|EA| &\le \binom{n-2}{m-2}\left\{\binom{n-2}{m-1}\frac{9.5(m-2)(m-1)\|g\|_3^2\|h\|_3^2}{(n-m)n} + (m-2)\binom{n-3}{m-2}\frac{2\|h\|_2}{n}\right\} \\
&\qquad \text{by Lemma E.2}(iii) \\
&= \binom{n-1}{m-1}^2\left\{\frac{9.5(m-2)(m-1)^2\|g\|_3^2\|h\|_3^2}{(n-1)^2 n} + \frac{2(m-1)^2(m-2)(n-m)\|h\|_2}{n(n-1)^2(n-2)}\right\} \\
&\qquad \text{by (E.1), (E.2) and (E.3)} \\
&\le C\binom{n-1}{m-1}^2\frac{m^3\|g\|_3^2\|h\|_3^2}{n^3},
\end{aligned} \tag{E.15}
$$

where the last line uses $2m < n$, and $1 = \sigma_g \le \|h\|_2 \le \|h\|_3$.

Now we bound EB. Analogously to EA, we first write

$$
\begin{aligned}
|EB| &\le \sum_{k=1}^{m-1}\binom{n-2}{k}\binom{n-2-k}{m-1-k}\binom{n-m-1}{m-1-k} \\
&\quad \Big|\mathbb{E}[\xi_{b,1}\xi_{b,2}\,\bar{h}_m(X_1, X_3, \ldots, X_{m+1})\,\bar{h}_m(X_2, \underbrace{X_3, \ldots, X_{k+2}}_{k \text{ shared}}, X_{m+2}, \ldots, X_{2m-k})]\Big| \\
&= \binom{n-2}{m-1}\sum_{k=1}^{m-1}\binom{m-1}{k}\binom{n-m-1}{m-1-k} \\
&\quad \Big|\mathbb{E}[\xi_{b,1}\xi_{b,2}\,\bar{h}_m(X_1, X_3, \ldots, X_{m+1})\,\bar{h}_m(X_2, \underbrace{X_3, \ldots, X_{k+2}}_{k \text{ shared}}, X_{m+2}, \ldots, X_{2m-k})]\Big| \\
&\qquad \text{by Lemma E.2}(ii) \\
&\le \binom{n-2}{m-1}\sum_{k=1}^{m-1}\binom{m-1}{k}\binom{n-m-1}{m-1-k}\left(\frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \frac{2k\|h\|_2}{n^{3/2}}\right) \text{ by Lemma E.1}(iv) \\
&= \binom{n-2}{m-1}\left\{\left[\binom{n-2}{m-1} - \binom{n-m-1}{m-1}\right]\frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \binom{n-3}{m-2}\frac{2(m-1)\|h\|_2}{n^{3/2}}\right\},
\end{aligned}
$$

where in the last equality, we have used

$$
\begin{aligned}
\sum_{k=1}^{m-1}\binom{m-1}{k}\binom{n-m-1}{m-1-k} &= \sum_{k=0}^{m-1}\binom{m-1}{k}\binom{n-m-1}{m-1-k} - \binom{n-m-1}{m-1} \\
&= \binom{n-2}{m-1} - \binom{n-m-1}{m-1} \text{ by Lemma E.2}(i)
\end{aligned}
$$

and

$$
\begin{aligned}
\sum_{k=1}^{m-1}\binom{m-1}{k}\binom{n-m-1}{m-1-k}k &= (m-1)\sum_{k=1}^{m-1}\binom{m-2}{k-1}\binom{n-m-1}{m-1-k} \\
&= (m-1)\sum_{k=0}^{m-2}\binom{m-2}{k}\binom{n-m-1}{m-2-k}
\end{aligned}
$$

$$= (m-1)\binom{n-3}{m-2} \text{ by Lemma E.2}(i)$$

Continuing, we get

$$\begin{aligned}
|EB| &\leq \binom{n-2}{m-1}\left\{\binom{n-2}{m-1}\frac{(m-1)^2}{n-m}\frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \binom{n-3}{m-2}\frac{2(m-1)\|h\|_2}{n^{3/2}}\right\} \\
&\qquad \text{by Lemma E.2}(iii) \\
&= \binom{n-1}{m-1}^2\left\{\frac{9.5(m-1)^2(n-m)\|g\|_3^2\|h\|_3^2}{(n-1)^2 n} + \frac{2(m-1)^2(n-m)^2\|h\|_2}{(n-1)^2(n-2)n^{3/2}}\right\} \\
&\qquad \text{by (E.1) and (E.3)} \\
&\leq C\binom{n-1}{m-1}^2\frac{m^2\|g\|_3^2\|h\|_3^2}{n^2},
\end{aligned} \tag{E.16}$$

where the last line uses $2m < n$, and $1 = \sigma_g \leq \|h\|_2 \leq \|h\|_3$.

Lastly, for EC, in an analogous manner as EA and EB, we first write it as

$$\begin{aligned}
EC = \sum_{k=0}^{m-2}&\binom{n-2}{k}\binom{n-2-k}{m-2-k}\binom{n-m}{m-2-k} \\
&\mathbb{E}[\xi_{b,1}\xi_{b,2}\,\bar{h}_m(X_1, X_2, \ldots, X_m) \\
&\bar{h}_m(X_1, X_2, \underbrace{X_3, \ldots, X_{k+2}}_{k \text{ shared, empty if } k=0}, X_{m+1}, \ldots, X_{2m-k-2})].
\end{aligned}$$

Then we can bound

$$\begin{aligned}
|EC| &\leq \sum_{k=0}^{m-2}\binom{n-2}{k}\binom{n-2-k}{m-2-k}\binom{n-m}{m-2-k} \\
&\quad \left|\mathbb{E}[\xi_{b,1}\xi_{b,2}\,\bar{h}_m(X_1, X_2, \ldots, X_m)\,\bar{h}_m(X_1, \ldots, X_{k+2}, X_{m+1}, \ldots, X_{2m-k-2}]\right| \\
&\leq \binom{n-2}{m-2}\sum_{k=0}^{m-2}\binom{m-2}{k}\binom{n-m}{m-2-k} \\
&\quad \left|\mathbb{E}[\xi_{b,1}\xi_{b,2}\,\bar{h}_m(X_1, X_2, \ldots, X_m)\,\bar{h}_m(X_1, \ldots, X_{k+2}, X_{m+1}, \ldots, X_{2m-k-2}]\right| \\
&\qquad \text{by Lemma E.2}(ii) \\
&\leq \binom{n-2}{m-2}\sum_{k=0}^{m-2}\binom{m-2}{k}\binom{n-m}{m-2-k}\left(\frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \frac{2k\|h\|_2}{n}\right) \text{ by Lemma E.1}(iii) \\
&= \binom{n-2}{m-2}\left\{\binom{n-2}{m-2}\frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \binom{n-3}{m-3}\frac{2(m-2)\|h\|_2}{n}\right\},
\end{aligned}$$

where the last equality comes from

$$\sum_{k=0}^{m-2}\binom{m-2}{k}\binom{n-m}{m-2-k}=\binom{n-2}{m-2}\text{ by Lemma E.2}(i)$$

and for $m \geq 3$,

$$\begin{aligned}\sum_{k=0}^{m-2}\binom{m-2}{k}\binom{n-m}{m-2-k}k&=\sum_{k=1}^{m-2}\binom{m-2}{k}\binom{n-m}{m-2-k}k\\&=(m-2)\sum_{k=1}^{m-2}\binom{m-3}{k-1}\binom{n-m}{m-2-k}\\&=(m-2)\sum_{k=0}^{m-3}\binom{m-3}{k}\binom{n-m}{m-3-k}\\&=(m-2)\binom{n-3}{m-3}\text{ by Lemma E.2}(i).\end{aligned}$$

Continuing, we get by (E.2) and (E.4),

$$\begin{aligned}|EC|&\leq\binom{n-2}{m-2}\left\{\binom{n-2}{m-2}\frac{9.5\|g\|_3^2\|h\|_3^2}{n}+\binom{n-3}{m-3}\frac{2(m-2)\|h\|_2}{n}\right\}\\&=\binom{n-1}{m-1}^2\left\{\frac{9.5(m-1)^2\|g\|_3^2\|h\|_3^2}{n(n-1)^2}+\frac{2(m-1)^2(m-2)^2\|h\|_2}{n(n-1)^2(n-2)}\right\}\\&\leq C\binom{n-1}{m-1}^2\left\{\frac{m^2\|g\|_3^2\|h\|_3^2}{n^3}+\frac{m^4\|h\|_2}{n^4}\right\}\end{aligned}\tag{E.17}$$

Substituting (E.15), (E.16), and (E.17) into (E.14), we get that

$$\begin{aligned}&\left|\mathbb{E}\left[\left(\xi_{b,1}\sum_{\substack{1\leq i_1<\cdots<i_{m-1}\leq n\\ i_l\neq 1\text{ for }l\in[m-1]}}\bar{h}_{m,\{1,i_1,\ldots,i_{m-1}\}}\right)\left(\xi_{b,2}\sum_{\substack{1\leq j_1<\cdots<j_{m-1}\leq n\\ j_l\neq 2\text{ for }l\in[m-1]}}\bar{h}_{m,\{2,j_1,\ldots,j_{m-1}\}}\right)\right]\right|\\&\leq C\binom{n-1}{m-1}^2\frac{m^2\|g\|_3^2\|h\|_3^2}{n^2},\end{aligned}\tag{E.18}$$

where we have used that $2m < n$ and $1 = \|g\|_2 \leq \|h\|_2 \leq \|h\|_3$. Finally, collecting (E.9), (E.10), (E.11), (E.13) and (E.18), we obtain (E.8).

E.2 Proof of Lemma 3.3(*ii*)

Note that

$$\delta_{2n,b}-\delta_{2n,b}^{(i)}=A+B,$$

where

$$A = \frac{2(n-1)}{\sqrt{n}(n-m)} \binom{n-1}{m-1}^{-1} \xi_{b,i} \sum_{\substack{1 \le i_1 < \cdots < i_{m-1} \le n \\ i_l \ne i \text{ for } l \in [m-1]}} \bar{h}_m(X_i, X_{i_1}, \ldots, X_{i_{m-1}})$$

and

$$B = \frac{2(n-1)}{\sqrt{n}(n-m)} \binom{n-1}{m-1}^{-1} \sum_{\substack{1 \le j \le n \\ j \ne i}} \Bigg(\xi_{b,j} \sum_{\substack{1 \le i_1 < \cdots < i_{m-2} \le n \\ i_l \ne j, i \text{ for } l = 1, \ldots, m-2}} \bar{h}_m(X_j, X_i, X_{i_1}, \ldots, X_{i_{m-2}}) \Bigg).$$

From (3.26) and (3.28), we first write

$$\begin{aligned} \|\Pi_2 - \Pi_2^{(i)}\|_2 &\le \mathbb{E}[(\xi_i^2 - 1) I(|\xi_i| > 1)] + \|\delta_{2n,b} - \delta_{2n,b}^{(i)}\|_2 \\ &\le \frac{\mathbb{E}[g^2]}{n} + \|A\|_2 + \|B\|_2, \end{aligned} \tag{E.19}$$

by Lemma A.1, where

$$A = \frac{2(n-1)}{\sqrt{n}(n-m)} \binom{n-1}{m-1}^{-1} \xi_{b,i} \sum_{\substack{1 \le i_1 < \cdots < i_{m-1} \le n \\ i_l \ne i \text{ for } l \in [m-1]}} \bar{h}_m(X_i, X_{i_1}, \ldots, X_{i_{m-1}})$$

and

$$B = \frac{2(n-1)}{\sqrt{n}(n-m)} \binom{n-1}{m-1}^{-1} \sum_{\substack{1 \le j \le n \\ j \ne i}} \Bigg(\xi_{b,j} \sum_{\substack{1 \le i_1 < \cdots < i_{m-2} \le n \\ i_l \ne j, i \text{ for } l = 1, \ldots, m-2}} \bar{h}_m(X_j, X_i, X_{i_1}, \ldots, X_{i_{m-2}}) \Bigg).$$

So we will bound $\|A\|_2$ and $\|B\|_2$, which is trivial for $m = 1$ as $\bar{h}_1(\cdot) = 0$. For $m \ge 2$, by Lemma E.1(ii),

$$\begin{aligned} \mathbb{E}[A^2] &\le \frac{4(n-1)^2}{n(n-m)^2} \binom{n-1}{m-1}^{-2} \mathbb{E}\Bigg[\Bigg(\sum_{\substack{1 \le i_1 < \cdots < i_{m-1} \le n \\ i_l \ne i \text{ for } l \in [m-1]}} \bar{h}_m(X_i, X_{i_1}, \ldots, X_{i_{m-1}}) \Bigg)^2 \Bigg] \\ &\le \frac{8(n-1)^2 (m-1)^2 \mathbb{E}[h^2]}{(n-m)^2 n (n-m+1) m} \le C \frac{m \mathbb{E}[h^2]}{n^2} \end{aligned} \tag{E.20}$$

Moreover, for B, we first expand its second moment as

$$\mathbb{E}[B^2]$$

$$= \frac{4(n-1)^2}{n(n-m)^2} \binom{n-1}{m-1}^{-2} \mathbb{E}\Bigg[\Bigg(\sum_{\substack{1 \le j \le n \\ j \ne i}} \Bigg(\xi_{b,j} \sum_{\substack{1 \le i_1 < \cdots < i_{m-2} \le n \\ i_l \ne j, i \text{ for } l = 1, \ldots, m-2}} \bar{h}_{m,\{j,i,i_1,\ldots,i_{m-2}\}} \Bigg) \Bigg)^2 \Bigg]$$

$$
\begin{aligned}
&= \frac{4(n-1)^2}{n(n-m)^2}\binom{n-1}{m-1}^{-2} \\
&\quad \times \Bigg\{(n-1) \underbrace{\sum_{\substack{1\le i_1<\cdots<i_{m-2}\le m-2 \\ 1\le j_1<\cdots<j_{m-2}\le m-2 \\ i_l, j_l\neq 1,2 \text{ for } l\in[m-2]}} \mathbb{E}[\xi_{b,1}^2 \bar{h}_{m,\{1,2,i_1,\dots,i_{m-2}\}} \bar{h}_{m,\{1,2,j_1,\dots,j_{m-2}\}}]}_{\equiv ED} \\
&\quad + (n-1)(n-2) \underbrace{\sum_{\substack{1\le i_1<\cdots<i_{m-2}\le n \\ 1\le j_1<\cdots<j_{m-2}\le n \\ i_l\neq 1,3 \text{ for } l\in[m-2] \\ j_l\neq 2,3 \text{ for } l\in[m-2]}} \mathbb{E}[\xi_{b,1}\xi_{b,2} \bar{h}_{m,\{1,3,i_1,\dots,i_{m-2}\}} \bar{h}_{m,\{2,3,j_1,\dots,j_{m-2}\}}]}_{\equiv EE} \Bigg\}.
\end{aligned}
\tag{E.21}
$$

To bound ED, we first note that, by $|\xi_{b,1}| \le 1$, Hölder's inequality and Lemma E.1(i), each of its summand can be bounded as

$$
\left| \mathbb{E}[\xi_{b,1}^2 \bar{h}_{m,\{1,2,i_1,\dots,i_{m-2}\}} \bar{h}_{m,\{1,2,j_1,\dots,j_{m-2}\}}] \right| \le \mathbb{E}[h^2] \tag{E.22}
$$

Then, by considering the number of elements $k \in [m-2]$ shared by the sets $\{i_1, \dots, i_{m-2}\}$ and $\{j_1, \dots, j_{m-2}\}$ indexing each such summand, we have the bound

$$
\begin{aligned}
|ED| &\le \sum_{k=0}^{m-2} \binom{n-2}{k}\binom{n-2-k}{m-2-k}\binom{n-m}{m-2-k}\mathbb{E}[h^2] \\
&= \binom{n-2}{m-2}\sum_{k=0}^{m-2}\binom{m-2}{k}\binom{n-m}{m-2-k}\mathbb{E}[h^2] \text{ by Lemma E.2}(ii) \\
&= \binom{n-2}{m-2}^2 \mathbb{E}[h^2] \text{ by Lemma E.2}(i) \\
&= \binom{n-1}{m-1}^2 \left(\frac{m-1}{n-1}\right)^2 \mathbb{E}[h^2] \text{ by (E.2).}
\end{aligned}
\tag{E.23}
$$

To bound EE, we first break it down as

$$
\begin{aligned}
EE &= \\
&\underbrace{\sum_{\substack{1\le i_1<\cdots<i_{m-2}\le n \\ 1\le j_1<\cdots<j_{m-2}\le n \\ i_l\neq 1,2,3 \text{ for } l\in[m-2] \\ j_l\neq 1,2,3 \text{ for } l\in[m-2]}} \mathbb{E}[\xi_{b,1}\xi_{b,2} \bar{h}_{m,\{1,3,i_1,\dots,i_{m-2}\}} \bar{h}_{m,\{2,3,j_1,\dots,j_{m-2}\}}}_{\equiv EE_1}
\end{aligned}
$$

$$+ \underbrace{\sum_{\substack{1\le i_1<\cdots<i_{m-3}\le n \\ 1\le j_1<\cdots<j_{m-2}\le n \\ i_l\neq 1,2,3 \text{ for } l\in[m-3] \\ j_l\neq 1,2,3 \text{ for } l\in[m-2]}} \mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{m,\{1,2,3,i_1,\ldots,i_{m-3}\}}\bar{h}_{m,\{2,3,j_1,\ldots,j_{m-2}\}}]}_{\equiv EE_2}$$

$$+ \underbrace{\sum_{\substack{1\le i_1<\cdots<i_{m-2}\le n \\ 1\le j_1<\cdots<j_{m-3}\le n \\ i_l\neq 1,2,3 \text{ for } l\in[m-2] \\ j_l\neq 1,2,3 \text{ for } l\in[m-3]}} \mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{m,\{1,3,i_1,\ldots,i_{m-2}\}}\bar{h}_{m,\{1,2,3,j_1,\ldots,j_{m-3}\}}]}_{\equiv EE_3}$$

$$+ \underbrace{\sum_{\substack{1\le i_1<\cdots<i_{m-3}\le n \\ 1\le j_1<\cdots<j_{m-3}\le n \\ i_l\neq 1,2,3 \text{ for } l\in[m-3] \\ j_l\neq 1,2,3 \text{ for } l\in[m-3]}} \mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_{m,\{1,2,3,i_1,\ldots,i_{m-3}\}}\bar{h}_{m,\{1,2,3,j_1,\ldots,j_{m-3}\}}}_{\equiv EE_4} . \tag{E.24}$$

Using Lemma E.1(iv), one can then bound EE_1 as

$$\begin{aligned}
&|EE_1| \\
&\le \sum_{k=0}^{m-2} \binom{n-3}{k}\binom{n-3-k}{m-2-k}\binom{n-1-m}{m-2-k}\left(\frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \frac{2d\|h\|_2}{n^{3/2}}\right) \\
&\le \sum_{k=0}^{m-2} \binom{n-3}{k}\binom{n-3-k}{m-2-k}\binom{n-1-m}{m-2-k}\left(\frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \frac{2\|h\|_2^2}{n}\right) \\
&\qquad\qquad \text{by (3.17) and } d \le m \le n \\
&\le 11.5\binom{n-3}{m-2}\sum_{k=0}^{m-2}\binom{m-2}{k}\binom{n-1-m}{m-2-k}\frac{\|g\|_3^2\|h\|_3^2}{n} \\
&\qquad\qquad \text{by Lemma E.2}(ii) \text{ and } \|h\|_2 \le \|h\|_3 \\
&= 11.5\binom{n-3}{m-2}^2\frac{\|g\|_3^2\|h\|_3^2}{n} \text{ by Lemma E.2}(i) \\
&= 11.5\binom{n-1}{m-1}^2\frac{(m-1)^2(n-m)^2\|g\|_3^2\|h\|_3^2}{n(n-1)^2(n-2)^2} \text{ by (E.3).}
\end{aligned} \tag{E.25}$$

For EE_2 and EE_3, using Lemma E.1(iii), one can bound them similarly as

$$\begin{aligned}
&\max(|EE_2|, |EE_3|) \\
&\le \sum_{k=0}^{m-3}\binom{n-3}{k}\binom{n-3-k}{m-3-k}\binom{n-m}{m-2-k}\left(\frac{9.5\|g\|_3^2\|h\|_3^2}{n} + \frac{2(2+k)\|h\|_2}{n}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \binom{n-3}{m-3}\sum_{k=0}^{m-3}\binom{m-3}{k}\binom{n-m}{m-2-k}\left(\frac{9.5\|g\|_3^2\|h\|_3^2}{n}+\frac{2(2+k)\|h\|_2}{n}\right) \text{ by Lemma E.2}(ii)\\
&= \binom{n-3}{m-3}\binom{n-3}{m-2}\frac{9.5\|g\|_3^2\|h\|_3^2+4\|h\|_2}{n}\\
&\quad + \binom{n-3}{m-3}\sum_{k=1}^{m-3}\binom{m-4}{k-1}\binom{n-m}{m-2-k}\frac{2(m-3)\|h\|_2}{n} \text{ by Lemma E.2}(i)\\
&= \binom{n-3}{m-3}\left\{\binom{n-3}{m-2}\frac{9.5\|g\|_3^2\|h\|_3^2+4\|h\|_2}{n}+\sum_{k=0}^{m-4}\binom{m-4}{k}\binom{n-m}{m-3-k}\frac{2(m-3)\|h\|_2}{n}\right\}\\
&= \binom{n-3}{m-3}\left\{\binom{n-3}{m-2}\frac{9.5\|g\|_3^2\|h\|_3^2+4\|h\|_2}{n}+\binom{n-4}{m-3}\frac{2(m-3)\|h\|_2}{n}\right\}\text{by Lemma E.2}(i)\\
&= \binom{n-3}{m-3}\left\{\binom{n-3}{m-2}\frac{9.5\|g\|_3^2\|h\|_3^2+4\|h\|_2}{n}+\binom{n-3}{m-3}\frac{2(m-3)(n-m)\|h\|_2}{(n-3)n}\right\}\\
&= \binom{n-1}{m-1}^2\left\{\frac{(m-1)^2(m-2)(n-m)(9.5\|g\|_3^2\|h\|_3^2+4\|h\|_2)}{(n-1)^2(n-2)^2 n}\right.\\
&\quad \left.+\frac{2(m-1)^2(m-2)^2(m-3)(n-m)\|h\|_2}{(n-1)^2(n-2)^2(n-3)n}\right\} \text{ by (E.3) and (E.4)}\\
&\le C\binom{n-1}{m-1}^2\left\{\frac{m^3\|g\|_3^2\|h\|_3^2}{n^4}+\frac{m^5\|h\|_2}{n^5}\right\}\text{by } 1\le\|g\|_3 \text{ and } \|h\|_2\le\|h\|_3. \qquad \text{(E.26)}
\end{aligned}
$$

Lastly, for EE_4, using Lemma E.1(iii), one can bound it as

$$
\begin{aligned}
&|EE_4|\\
&\quad \le \sum_{k=0}^{m-3}\binom{n-3}{k}\binom{n-3-k}{m-3-k}\binom{n-m}{m-3-k}\left(\frac{9.5\|g\|_3^2\|h\|_3^2}{n}+\frac{2(3+k)\|h\|_2}{n}\right)\\
&\quad = \binom{n-3}{m-3}\sum_{k=0}^{m-3}\binom{m-3}{k}\binom{n-m}{m-3-k}\\
&\quad \left(\frac{9.5\|g\|_3^2\|h\|_3^2}{n}+\frac{2(3+k)\|h\|_2}{n}\right) \text{ by Lemma E.2}(ii)\\
&\quad = \binom{n-3}{m-3}\left\{\binom{n-3}{m-3}\frac{9.5\|g\|_3^2\|h\|_3^2+6\|h\|_2}{n}\right.\\
&\qquad \left.+\frac{2(m-3)\|h\|_2}{n}\sum_{k=1}^{m-3}\binom{m-4}{k-1}\binom{n-m}{m-4-(k-1)}\right\}\\
&\quad = \binom{n-3}{m-3}\left\{\binom{n-3}{m-3}\frac{9.5\|g\|_3^2\|h\|_3^2+6\|h\|_2}{n}\right.\\
&\qquad \left.+\frac{2(m-3)\|h\|_2}{n}\binom{n-4}{m-4}\right\} \text{ by Lemma E.2}(i)\\
&\quad \le C\binom{n-1}{m-1}^2\left\{\frac{m^4\|g\|_3^2\|h\|_3^2}{n^5}+\frac{m^6\|h\|_2}{n^6}\right\}\text{by (E.4), (E.5),}
\end{aligned}
$$

$$1 \leq \|g\|_3 \text{ and } \|h\|_2 \leq \|h\|_3. \tag{E.27}$$

Combining (E.24), (E.25), (E.26), (E.27), and $2m < n$, we get that

$$|EE| \leq C\binom{n-1}{m-1}^2 \left\{ \|g\|_3^2 \|h\|_3^2 \left(\frac{m^2}{n^3} + \frac{m^3}{n^4} + \frac{m^4}{n^5} \right) + \|h\|_2 \left(\frac{m^5}{n^5} \right) \right\}. \tag{E.28}$$

Combining (E.21), (E.23), and (E.28), we get

$$\begin{aligned} \mathbb{E}[B^2] &\leq C\left\{ \frac{m^2}{n^2}\mathbb{E}[h^2] + \left[\|g\|_3^2 \|h\|_3^2 \left(\frac{m^2}{n^2} + \frac{m^3}{n^3} + \frac{m^4}{n^4} \right) + \|h\|_2 \left(\frac{m^5}{n^4} \right) \right] \right\} \\ &\leq C\left\{ \frac{m^2 \|g\|_3^2 \|h\|_3^2}{n^2} + \frac{m^5 \|h\|_2}{n^4} \right\}, \end{aligned} \tag{E.29}$$

where we have used $2m < n$, as well as $\|h\|_2 \leq \|h\|_3$ and $1 = \|g\|_2 \leq \|g\|_3$ in the last line. Combining (E.19), (E.20), and (E.29) gives Lemma 3.3(ii).

Appendix F. Proof of Lemmas E.1 and E.2

F.1 Proof of Lemma E.1

The proof for (i) and (ii) can be found in Chen et al. [3, Ch.10, Appendix]. We will focus on proving (iii) and (iv). For any subset $\{i_1, \ldots, i_k\} \subset [n]$, we will denote

$$X_{\{i_1,\ldots,i_k\}} = \{X_{i_1}, \ldots, X_{i_k}\}.$$

To simplify the notation, we also denote

$$I = \{i_1, \ldots, i_{k_1}\} \text{ and } J = \{j_1, \ldots, j_{k_2}\},$$

as well as

$$h_I = h_{k_1}(X_{i_1}, \ldots, X_{i_{k_1}}) \text{ and } \bar{h}_I = \bar{h}_{k_1,\{i_1,\ldots,i_{k_1}\}}$$

and

$$h_J = h_{k_2}(X_{j_1}, \ldots, X_{j_{k_2}}) \text{ and } \bar{h}_I = \bar{h}_{k_2,\{j_1,\ldots,i_{j_2}\}}.$$

First, it suffices to assume both

$$k_1, k_2 \geq 2$$

because if any of k_1 and k_2 is equal to 1, then one of $\bar{h}_{k_1,\{i_1,\ldots,i_{k_1}\}}$ and $\bar{h}_{k_2,\{j_1,\ldots,j_{k_2}\}}$ must be equal to zero by the definition in (3.9), so the bound is trivial. Moreover, one can further assume without loss of generality that the index sets I and J are such that

$$I\backslash\{1,2\} = J\backslash\{1,2\} = [3:(d+2)] \text{ if } d > 0, \tag{F.1}$$

in which case it must be true that $|I\backslash\{1,2\}| = |J\backslash\{1,2\}| = d$. This is because for any I and J, we have

$$\begin{aligned}
&\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_I\bar{h}_J] \\
&\quad= \mathbb{E}\Big[\mathbb{E}[\xi_{b,1}\xi_{b,2}\bar{h}_I\bar{h}_J \mid X_{\{1,2\}\cup(I\cap J)}]\Big] \\
&\quad= \mathbb{E}\Big[\xi_{b,1}\xi_{b,2}\mathbb{E}[\bar{h}_I\bar{h}_J \mid X_{\{1,2\}\cup(I\cap J)}]\Big] \\
&\quad= \mathbb{E}\Big[\xi_{b,1}\xi_{b,2}\mathbb{E}[\bar{h}_I \mid X_{\{1,2\}\cup(I\cap J)}]\mathbb{E}[\bar{h}_J \mid X_{\{1,2\}\cup(I\cap J)}]\Big] \\
&\qquad\qquad \text{because } I\backslash\Big(\{1,2\}\cup(I\cap J)\Big) \text{ and } J\backslash\Big(\{1,2\}\cup(I\cap J)\Big) \text{ are disjoint} \\
&\quad= \mathbb{E}\Big[\xi_{b,1}\xi_{b,2}\mathbb{E}[\bar{h}_I \mid X_{(I\cap\{1,2\})\cup(I\cap J)}]\mathbb{E}[\bar{h}_J \mid X_{(J\cap\{1,2\})\cup(I\cap J)}]\Big] \\
&\quad= \mathbb{E}\Big[\xi_{b,1}\xi_{b,2}\bar{h}_{(I\cap\{1,2\})\cup(I\cap J)}\bar{h}_{(J\cap\{1,2\})\cup(I\cap J)}\Big].
\end{aligned}$$

Since

$$\Big((I\cap\{1,2\})\cup(I\cap J)\Big)\backslash\{1,2\} = (I\cap J)\backslash\{1,2\} = \Big((J\cap\{1,2\})\cup(I\cap J)\Big)\backslash\{1,2\}$$

and

$$|(I\cap J)\backslash\{1,2\}| = d \text{ by assumption,}$$

by the i.i.d.'ness of the data $X_1, \ldots, X_n$ it suffices to assume (F.1).

By the definition in (3.9), we perform the expansion

$$\begin{aligned}
&\mathbb{E}[\xi_{b,1}\xi_{b,2}\,\bar{h}_I\,\bar{h}_J] \\
&\quad= \mathbb{E}\Big[\xi_{b,1}\xi_{b,2}\Big(h_I - \sum_{i\in I\cap\{1,2\}} g(X_i) - \sum_{i\in I\backslash\{1,2\}} g(X_i)\Big) \\
&\quad\Big(h_J - \sum_{j\in J\cap\{1,2\}} g(X_j) - \sum_{j\in J\backslash\{1,2\}} g(X_j)\Big)\Big] \\
&\quad= \underbrace{\mathbb{E}[\xi_{b,1}\xi_{b,2}\,h_I\,h_J]}_{\equiv HH} \\
&\qquad \underbrace{- \sum_{i\in I\cap\{1,2\}} \mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_i)\,h_J]}_{\equiv GH_1} \underbrace{- \sum_{j\in J\cap\{1,2\}} \mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_j)\,h_I]}_{\equiv GH_2} \\
&\qquad \underbrace{- \sum_{i\in I\backslash\{1,2\}} \mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_i)\,h_J]}_{\equiv GH_3} \underbrace{- \sum_{j\in J\backslash\{1,2\}} \mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_j)\,h_I]}_{\equiv GH_4}
\end{aligned}$$

$$+\underbrace{\sum_{i\in I\cap\{1,2\}}\sum_{j\in J\cap\{1,2\}}\mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_i)g(X_j)]}_{\equiv GG_1}+\underbrace{\sum_{i\in I\setminus\{1,2\}}\sum_{j\in J\setminus\{1,2\}}\mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_i)g(X_j))]}_{\equiv GG_2},$$

recognizing that the last batch of expansion terms

$$\sum_{i\in I\cap\{1,2\}}\sum_{j\in J\setminus\{1,2\}}\underbrace{\mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_i)g(X_j)]}_{=\mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_i)]\mathbb{E}[g(X_j)]=0}+\sum_{i\in I\setminus\{1,2\}}\sum_{j\in J\cap\{1,2\}}\underbrace{\mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_i)g(X_j)]}_{\mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_j)]\mathbb{E}[g(X_i)]=0}$$

vanish. The remaining terms in each row of the expansion above are bounded as follows:

F.1.1 Bound on *HH*:

$$\begin{aligned}|HH| &= \left|\mathbb{E}[\xi_{b,1}\xi_{b,2}\,h_I\,h_J]\right| \le \left\|\xi_{b,1}\xi_{b,2}\right\|_3\left\|h_I\,h_J\right\|_{3/2}\\ &= \left(\mathbb{E}[|\xi_{b,1}|^3]\mathbb{E}[|\xi_{b,2}|^3]\right)^{1/3}\left(\mathbb{E}\left[\left|h_I\right|^{3/2}\left|h_J\right|^{3/2}\right]\right)^{2/3}\\ &\le \left(\mathbb{E}[|\xi_{b,1}|^3]\mathbb{E}[|\xi_{b,2}|^3]\right)^{1/3}\left(\||h_I|^{3/2}\|_2\||h_J|^{3/2}\|_2\right)^{2/3} \quad \text{by Cauchy's inequality}\\ &\le n^{-1}\|g\|_3^2\|h\|_3^2, \end{aligned} \tag{F.2}$$

where the last line comes from (3.10) with $|I|\vee|J|\le m$.

F.1.2 Bound on $GH_1 + GH_2$:

$$\begin{aligned}&|GH_1+GH_2|\\ &\le \sum_{i\in I\cap\{1,2\}}\|\xi_{b,1}\xi_{b,2}g(X_i)\|_{3/2}\|h_J\|_3+\sum_{j\in J\cap\{1,2\}}\|\xi_{b,1}\xi_{b,2}g(X_j)\|_{3/2}\|h_I\|_3\\ &= |I\cap\{1,2\}|\cdot\|\xi_{b,1}\xi_{b,2}g(X_1)\|_{3/2}\|h_J\|_3+|J\cap\{1,2\}|\cdot\|\xi_{b,1}\xi_{b,2}g(X_1)\|_{3/2}\|h_I\|_3\\ &\le 4\|\xi_{b,1}\xi_{b,2}g(X_1)\|_{3/2}\|h\|_3 \quad \text{by (3.10)}\\ &= 4\|\xi_{b,1}g(X_1)\|_{3/2}\|\xi_{b,2}\|_{3/2}\|h\|_3 \quad \text{by independence}\\ &\le 4\left(\mathbb{E}[n^{-3/4}|g(X_1)|^3]\right)^{2/3}\left(\mathbb{E}[n^{-3/4}|g(X_2)|^{3/2}]\right)^{2/3}\|h\|_3\\ &= 4n^{-1}\|g\|_3^2\|g\|_{3/2}\|h\|_3\\ &\le 4n^{-1}\|g\|_3^2\|h\|_3, \end{aligned} \tag{F.3}$$

where the last inequality is true because $\|g\|_{3/2}\le\|g\|_2=\sigma_g=1$.

F.1.3 General bound on $GH_3 + GH_4$:

$$
\begin{aligned}
&|GH_3 + GH_4| \\
&\quad \le \sum_{i \in I \backslash \{1,2\}} \|\xi_{b,1}\xi_{b,2}g(X_i)\|_2 \|h_J\|_2 + \sum_{j \in J \backslash \{1,2\}} \|\xi_{b,1}\xi_{b,2}g(X_j)\|_2 \|h_I\|_2 \\
&\quad = |I \backslash \{1,2\}| \cdot \|\xi_{b,1}\xi_{b,2}g(X_3)\|_2 \|h_J\|_2 + |J \backslash \{1,2\}| \cdot \|\xi_{b,1}\xi_{b,2}g(X_3)\|_2 \|h_I\|_2 \\
&\quad \le 2d \|\xi_{b,1}\xi_{b,2}g(X_3)\|_2 \|h\|_2 \text{ by (3.10) and (F.1)} \\
&\quad \le 2d \|\xi_1\|_2 \|\xi_2\|_2 \|g(X_3)\|_2 \|h\|_2 \text{ by independence} \\
&\quad = 2dn^{-1} \|h\|_2 \text{ by (3.3).}
\end{aligned} \tag{F.4}
$$

F.1.4 Special bound on $GH_3 + GH_4$ under $1 \notin J$ and $2 \notin I$:

$$
\begin{aligned}
&|GH_3 + GH_4| \\
&\quad = \left| \sum_{i \in I \backslash \{1,2\}} \mathbb{E}[\xi_{b,1}]\mathbb{E}[\xi_{b,2}g(X_i)\, h_J] + \sum_{j \in J \backslash \{1,2\}} \mathbb{E}[\xi_{b,2}]\mathbb{E}[\xi_{b,1}g(X_j)\, h_I] \right| \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{by } 1 \notin J \text{ and } 2 \notin I \\
&\quad \le \sum_{i \in I \backslash \{1,2\}} \left|\mathbb{E}[\xi_{b,1}]\right| \|\xi_{b,2}g(X_i)\|_2 \|h\|_2 \\
&\qquad + \sum_{j \in J \backslash \{1,2\}} \left|\mathbb{E}[\xi_{b,2}]\right| \|\xi_{b,1}g(X_j)\|_2 \|h\|_2 \text{ by (3.10)} \\
&\quad \le 2d \cdot \left|\mathbb{E}[\xi_{b,1}]\right| \|\xi_{b,1}g(X_3)\|_2 \|h\|_2 \text{ by (F.1)} \\
&\quad \le 2d\mathbb{E}[\xi_1^2]\, \|\xi_1\|_2\, \|g(X_3)\|_2 \|h\|_2 \text{ by Lemma A.1 and independence} \\
&= 2dn^{-3/2} \|h\|_2 \text{ by } \sigma_g^2 = 1 \text{ in (3.3).}
\end{aligned} \tag{F.5}
$$

F.1.5 Bound on $GG_1 + GG_2$

$$
\begin{aligned}
&|GG_1 + GG_2| \\
&\quad \le 2\Big(\mathbb{E}[|\xi_{b,1}g^2(X_1)|] \cdot |\mathbb{E}[\xi_{b,2}]| + \mathbb{E}[|\xi_{b,1}g(X_1)|] \cdot \mathbb{E}[|\xi_{b,2}g(X_2)|]\Big) \\
&\qquad + \left| \sum_{i \in I \backslash \{1,2\}} \sum_{j \in J \backslash \{1,2\}} \mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_i)g(X_j))] \right| \\
&\quad = 2\Big(\mathbb{E}[|\xi_{b,1}g^2(X_1)|] \cdot |\mathbb{E}[\xi_{b,2}]| + \mathbb{E}[|\xi_{b,1}g(X_1)|] \cdot \mathbb{E}[|\xi_{b,2}g(X_2)|]\Big) \\
&\qquad + d \cdot |\mathbb{E}[\xi_{b,1}]| \cdot |\mathbb{E}[\xi_{b,2}]| \cdot \mathbb{E}[g^2(X_3)],
\end{aligned}
$$

where the last equality uses that

$$
\begin{aligned}
\mathbb{E}[\xi_{b,1}\xi_{b,2}g(X_i)g(X_j))] &= \mathbb{E}[\xi_{b,1}\xi_{b,2}]\mathbb{E}[g(X_i)]\mathbb{E}[g(X_j))] \\
&= 0 \text{ if } i \ne j \text{ and } i, j \notin \{1, 2\},
\end{aligned}
$$

as well as the working assumption in (F.1). Continuing, we get

$$
\begin{aligned}
&|GG_1 + GG_2| \\
&\le 2\Big(\mathbb{E}[g^2(X_1)] \cdot |\mathbb{E}[\xi_{b,2}]| + n^{-1}\mathbb{E}[g^2(X_1)] \cdot \mathbb{E}[g^2(X_2)]\Big) \\
&\quad + d \cdot |\mathbb{E}[\xi_{b,1}]| \cdot |\mathbb{E}[\xi_{b,2}]| \cdot \mathbb{E}[g^2(X_3)] \\
&\le 2(n^{-1} + n^{-1}) + dn^{-2} \text{ by Lemma A.1 and } \mathbb{E}[g(X_1^2)] = 1 \text{ in (3.3)} \\
&\le 4n^{-1} + \frac{d}{2m} n^{-1} \text{ by } 2m < n \\
&\le 4.5n^{-1} \text{ by } d \le m. \qquad (F.6)
\end{aligned}
$$

F.1.6 Summary

Recall $1 = \sigma_g \le \|g\|_3 \le \|h\|_3$. Combining (F.2), (F.3), (F.4), (F.6) gives Lemma E.1(iii), and combining (F.2), (F.3), (F.5), (F.6) gives Lemma E.1(iv).

F.2 Proof of Lemma E.2

Statement (i) is the Vandermonde's identity, which counts the number of ways to choose m balls from n_1 red balls and n_2 green balls, by summing over $k \in [0 : m]$ the number of ways to choose k red balls and $m - k$ green balls. Statement (ii) counts the number of ways to choose m balls out of a bag of n balls and paint k of the m chosen balls as red, in two different ways. Statement (iii) comes from

$$
\begin{aligned}
\binom{a}{b} - \binom{a-e}{b} &= \binom{a}{b}\left(1 - \frac{(a-e)\dots(a-e-b+1)}{a\cdots(a-b+1)}\right) \\
&= \binom{a}{b}\left(1 - \prod_{j=a-b+1}^{a}\left(1 - \frac{e}{j}\right)\right) \\
&\le \binom{a}{b} \sum_{j=a-b+1}^{a} \frac{e}{j} \\
&\le \binom{a}{b} \frac{be}{a-b+1}.
\end{aligned}
$$

Author Contributions D. L. and Q. S. conceived the research. D.L. wrote the manuscript and proved the main results. L.Z. improved Theorem 3.1.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Data Availability We do not analyze or generate any datasets, because our work proceeds within a theoretical and mathematical approach.

Declarations

Conflict of interest This research is partially supported by National Nature Science Foundation of China NSFC 12031005, Shenzhen Outstanding Talents Training Fund, China, and the Professor Maurice H. Belz Scholarship from the University of Melbourne.

References

1. Arvesen, J.N.: Jackknifing U-statistics. Ann. Math. Stat. **40**(6), 2076–2100 (1969)
2. Callaert, H., Veraverbeke, N.: The order of the normal approximation for a studentized U-statistic. Ann. Stat. **9**(1), 194–200 (1981)
3. Chen, L.H., Goldstein, L., Shao, Q.-M.: Normal Approximation by Stein's Method, vol. 2. Springer, New York (2011)
4. Helmers, R.: The Berry–Esseen bound for studentized U-statistics. Can. J. Stat. **13**(1), 79–82 (1985)
5. Jing, B.-Y., Wang, Q., Zhao, L.: The Berry–Esséen bound for studentized statistics. Ann. Probab. **28**(1), 511–535 (2000)
6. Korolyuk, V.S., Borovskich, Y.V.: Theory of U-Statistics, vol. 273. Springer, New York (2013)
7. Lai, T.L., Shao, Q.-M., Wang, Q.: Cramér type moderate deviations for Studentized U-statistics. ESAIM Probab. Stat. **15**, 168–179 (2011)
8. Shao, Q.-M.: Stein's method, self-normalized limit theory and applications. In: Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) (In 4 Volumes) Vol. I: Plenary Lectures and Ceremonies Vols. II–IV: Invited Lectures, pp. 2325–2350. World Scientific (2010)
9. Shao, Q.-M., Zhang, K., Zhou, W.-X.: Stein's method for nonlinear statistics: a brief survey and recent progress. J. Stat. Plan. Inference **168**, 68–89 (2016)
10. Shao, Q.-M., Zhou, W.-X.: Cramér type moderate deviation theorems for self-normalized processes. Bernoulli **22**(4), 2029–2079 (2016)
11. Steele, J.M.: The Cauchy–Schwarz Master Class. AMS/MAA Problem Books Series. Mathematical Association of America, Washington, DC; Cambridge University Press, Cambridge. An introduction to the art of mathematical inequalities (2004)
12. Stein, C.: 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, vol. 6, pp. 583–603. University of California Press (1972)
13. Student. "The probable error of a mean. *Biometrika*, 1–25 (1908)
14. Victor, H., la Peña, D., Lai, T.L., Shao, Q.-M.: Self-Normalized Processes: Limit Theory and Statistical Applications, vol. 204. Springer, New York (2009)
15. Zhao, L.: The rate of the normal approximation for a studentized U-statistic. Sci. Explor. **3**(2), 45–52 (1983)