Preprint

DUAL INDUCTION CLT FOR HIGH-DIMENSIONAL m-DEPENDENT DATA

BY HEEJONG BONG [1], ARUN KUMAR KUCHIBHOTLA [1] AND ALESSANDRO RINALDO [1]

[1] *Department of Statistics and Data Science, Carnegie Mellon University, {hbong, arunku, arinaldo}@cmu.edu*

In this work, we provide a $1/\sqrt{n}$-rate finite sample Berry–Esseen bound for m-dependent high-dimensional random vectors over the class of hyper-rectangles. This bound imposes minimal assumptions on the random vectors such as nondegenerate covariances and finite third moments. The proof uses inductive relationships between anti-concentration inequalities and Berry–Esseen bounds, which are inspired by the classical Lindeberg swapping method and the concentration inequality approach for dependent data. Performing a dual induction based on the relationships, we obtain tight Berry–Esseen bounds for dependent samples.

1. Introduction. Recent advances in technology have led to the unprecedented availability of large-scale spatiotemporal data. An important challenge in the analyses of such data is to provide a theoretical guarantee of statistical inferences under temporal dependence. Many existing theoretical studies, such as Liu (2020), relied on parametric or distributional assumptions to give a valid confidence interval, but the validity of the assumptions remains questionable in real-world applications. Central limit theorems (CLTs) provide one of the most general methods of statistical inference based on Gaussian approximation. To account for high dimensional data in real-life applications, there has been a recent surge of work on CLTs with increasing dimensions. Unlike the multivariate Berry–Esseen bounds which can control the difference between the probabilities associated with the average of random vectors and the corresponding Gaussian for arbitrary convex sets, in high dimensions, one needs to restrict the class of sets to a class of "sparse" sets such as hyper-rectangles. Although most of the literature (Norvaiša and Paulauskas, 1991; Chernozhukov, Chetverikov and Kato, 2013, 2017a; Deng and Zhang, 2020; Lopes, Lin and Müller, 2020; Fang and Koike, 2020; Chernozhukov, Chetverikov and Koike, 2020; Kuchibhotla and Rinaldo, 2020; Kuchibhotla, Mukherjee and Banerjee, 2021; Koike, 2021; Lopes, 2022; Chernozhuokov et al., 2022; Chernozhukov et al., 2023) is focused on independent observations, some works have also considered extensions of high-dimensional CLTs to the important case of dependent data including m-dependence, dependency graphs, and physical/functional dependence (Zhang and Wu, 2017; Zhang and Cheng, 2018; Chang, Chen and Wu, 2021; Kojevnikov and Song, 2022) Dependent data CLTs are very crucial for applications in Econometrics and causal inference (under interference).

In this paper, we focus on high-dimensional CLTs over the class of hyper-rectangles. That is, our objective is to bound the Kolmogorov-Smirnov statistic between a summation of samples $X_1, \ldots, X_n \in \mathbb{R}^p$ and its Gaussian approximation Y, denoted by

$$\mu\left(\sum_{i=1}^n X_i, Y\right) \equiv \sup_{r\in\mathbb{R}^p} \left|\mathbb{P}[\sum_{i=1}^n X_i \in A_r] - \mathbb{P}[Y \in A_r]\right|,$$

where $A_r \equiv \{x \in \mathbb{R}^p : x \preceq r\}$ and for two vectors $a, b \in \mathbb{R}^p$, $a \preceq b$ means that $a_k \leq b_k$ for every $k \in [p]$. For independent samples, there has been a flurry of novel results since the seminal work of Chernozhukov, Chetverikov and Kato (2013). A popular approach has been the Lindeberg interpolation, leading to an $n^{-1/6}$ rate (Bentkus et al., 2000; Lopes, 2022).

Keywords and phrases: High-dimensional inference, Central limit theorem, Berry–Esseen bound, m-dependence structure.

Recently, Kuchibhotla and Rinaldo (2020) used a recursion method over the Lindeberg interpolation to establish a high-dimensional Berry–Esseen bound for independent random variables with the rate at most $(\log^3 n \log^8(ep)/n)^{1/2}$ under minimal assumptions: nondegenerate covariances and finite third moments of the observations. We highlight their use of the inductive relationship between anti-concentration inequalities and Berry–Esseen bounds. The importance of anti-concentration inequalities in proving Berry–Esseen bound has been well-acknowledged in the literature. Bentkus et al. (2000) showed explicitly how the dependence of anti-concentration inequalities on p affects that of Berry–Esseen bounds. While the past literature on CLTs for independent observations mainly used bounds for Kolmogorov-Smirnov statistic in terms of anti-concentration inequality, a novel contribution of Kuchibhotla and Rinaldo (2020) was an (implicit) formulation of the backward relationship: induction of anti-concentration inequalities from Berry–Esseen bounds. We extend this idea to include a bound for anti-concentration in terms of the Kolmogorov-Smirnov statistic and demonstrate how the dual induction obtains the tight Berry-Esseen bounds.

In addition, our work extends this approach to m-dependent cases, where $X_i \perp\!\!\!\perp X_j$ if $|i-j|>m$. While previous work has demonstrated this rate for univariate cases (Shergin, 1980; Chen and Shao, 2004), the tightest rate known in high-dimensional cases is $1/n^{1/6}$ for sub-exponential random vectors (Chang, Chen and Wu, 2021). Our paper improves upon this by providing an enhanced Berry–Esseen bound under a more generalized setting called m-ring dependence. We introduce this setting in Section 2 and present our main results in Section 3, which showcase the optimal scaling of $1/\sqrt{n}$ under non-degenerate covariance and finite third moments. To establish these results, we used a combination of the Lindenberg swapping technique and the dual induction argument, overcoming the challenges posed by the m-ring dependence structure; the outline is provided in Section 5. We extensively compare our results to existing works under independence and m-dependence in Section 3.1. Finally, in Section 4, we conclude our paper by summarizing the main findings and outlining avenues for future research.

2. Random vectors with m-ring dependence.

We start with introducing a slight generalization of m-dependence as follows: $X_1,\dots,X_n \in \mathbb{R}^p$ are said to have *m-ring dependence* if $X_i \perp\!\!\!\perp X_j$ for any i and j satisfying $\min\{|i-j|, n-|i-j|\} > m$. If $X_1,\dots,X_n$ is m-dependent, then

$$\min\{|i-j|, n-|i-j|\} > m \quad \Rightarrow \quad |i-j|>m \quad \Rightarrow \quad X_i \perp\!\!\!\perp X_j.$$

This implies that $X_1,\dots,X_n$ is also m-ring dependent. In other words, m-dependence is a special case of the m-ring dependence. Let $Y_1,\dots,Y_n$ be jointly Gaussian random vectors with mean zero and the same second moment as $X_1,\dots,X_n$. i.e.,

$$\mathrm{Var}[(Y_1^\top,\dots,Y_n^\top)^\top] = \mathrm{Var}[(X_1^\top,\dots,X_n^\top)^\top].$$

For subset $I \subseteq \{1,2,\dots,n\}$, define

$$\mathscr{X}_I := \{X_k : k \in I\}, \quad \mathscr{Y}_I := \{Y_k : k \in I\}.$$

and

$$X_I := \sum_{k\in I} X_k, \quad Y_I := \sum_{k\in I} Y_k. \tag{1}$$

To streamline our discussion, we introduce notation for index intervals. For any values of i and j that satisfy $1 \le i < j \le n$, we will denote the index set $\{i,\dots,j\}$ as $[i,j]$. Specifically, when referring to the complete index set of n random vectors, we will use $[1,n]$. To align with the conventions of real intervals, we will employ parentheses to represent open-ended intervals; e.g., $(1,n]$ denotes the set $\{2,\dots,n\}$. For any $I \subset [1,n]$, let $\sigma_{\min,I}$ and $\underline{\sigma}_I$ be

$$\sigma^2_{\min,I} := \min_{k\in[p]} \mathrm{Var}[Y_I^{(k)}], \quad \text{and} \quad \underline{\sigma}_I^2 := \lambda_{\min}(\mathrm{Var}[Y_I|\mathscr{Y}_{I^\complement}]),$$

where the superscript $Y_I^{(k)}$ notates each k-th element of p-dimensional random vector Y_I (as defined in Eq. (1)), and $I^\complement$ is the index set complement. i.e., $I^\complement \equiv [1,n]\setminus I$. We note that the conditional variance is given by the Schur complement of the marginal covariance matrix:

$$\mathrm{Var}[Y_I|\mathscr{Y}_{I^\complement}] = \mathrm{Var}[Y_I] - \mathrm{Cov}[Y_I, \mathrm{vec}(\mathscr{Y}_{I^\complement})]\,\mathrm{Var}[\mathrm{vec}(\mathscr{Y}_{I^\complement})]^{-1}\,\mathrm{Cov}[\mathrm{vec}(\mathscr{Y}_{I^\complement}), Y_I],$$

where $\mathrm{vec}(\mathscr{Y}_{I^\complement})$ indicates the vectorized representation by concatenation. i.e., $\mathrm{vec}(\mathscr{Y}_{I^\complement}) \equiv (Y_k^\top : k \in I^\complement)^\top$. In case $m=0$, i.e., $X_1,\dots,X_n$ are independent, then $\mathrm{Var}[Y_I|\mathscr{Y}_{I^\complement}] = \mathrm{Var}[Y_I]$. Also, for $q \geq 1$, let $L_{q,i}$ and $\nu_{q,i}$ be

$$L_{q,i} := \max_k \mathbb{E}[|X_i^{(k)}|^q] + \max_k \mathbb{E}[|Y_i^{(k)}|^q], \quad \text{and} \quad \nu_{q,i} := \mathbb{E}[\|X_i\|_\infty^q] + \mathbb{E}[\|Y_i\|_\infty^q]. \tag{2}$$

Denote the averages of $L_{q,i}$'s and $\nu_{q,i}$'s by

$$\overline{L}_q = \frac{1}{n}\sum_{i=1}^n L_{q,i} \text{ and } \overline{\nu}_q = \frac{1}{n}\sum_{i=1}^n \nu_{q,i}. \tag{3}$$

For any subset $I \subset [1,n]$, let

$$\overline{L}_{q,I} = \frac{1}{|I|}\sum_{j\in I} L_{q,j} \text{ and } \overline{\nu}_{q,I} = \frac{1}{|I|}\sum_{j\in I} \nu_{q,j}. \tag{4}$$

We note that due to Jensen's inequality, $\nu_{2,i} \leq \nu_{q,i}^{2/q}$ for $i \in [1,n]$ and $\overline{\nu}_2 \leq \overline{\nu}_q^{2/q}$ for any $q \geq 2$.

Notation. In the following argument, $C(\dots)$ is a constant with implicit dependency on the parameters in the parentheses, whose value changes across lines. For absolute constants with no dependency, we omit the parentheses and denote them by C. $\mathbf{1}$ stands for the vector with elements 1 in the appropriate dimension at each line. $\mathbb{I}\{\cdot\}$ denotes the indicator function.

3. High-dimensional Berry–Esseen bound for m-ring dependent random samples with nondegenerate covariance matrices. In this section, we state a high-dimensional Berry–Esseen bound for random vectors with m-ring dependence and nondegenerate covariance matrices. That is, the minimum eigenvalues of sums of consecutive samples are bounded away from 0. We assume the existence of constants $\sigma_{\min}, \underline{\sigma} > 0$ such that for all $I = [i,j]$ or $[j,n]\cup[1,i]$ with some $1 \leq i < j \leq n$,

$$\sigma^2_{\min,I} \geq \sigma^2_{\min}\cdot|I|, \tag{MIN-VAR}$$

$$\underline{\sigma}_I^2 \geq \underline{\sigma}^2 \cdot \max\{|I|-2m, 0\}, \tag{MIN-EV}$$

$$\sigma_{\min} \leq \underline{\sigma}\sqrt{\log(4ep)/2}, \tag{VAR-EV}$$

where $|I|$ is the number of elements in I. The assumption of strongly non-denegerate covariance (i.e., $\lambda_{\min}(\mathrm{Var}[X_i])$ is bounded away from zero) has been commonly used in high-dimensional CLTs under independence (Kuchibhotla and Rinaldo, 2020; Chernozhukov, Chetverikov and Koike, 2020; Fang and Koike, 2021; Lopes, 2022). Building upon these works, Assumptions (MIN-VAR) and (MIN-EV) extend the assumption to m-dependence cases in the same spirit with Assumption (3) in Shergin (1980). It is also worth noting that by subtracting $2m$ from the interval length, Assumption (MIN-EV) allows for the possibility of complete dependence between $X_{[i,j]}$ and the adjacent elements when the interval $[i,j]$ is sufficiently short (e.g., $|[i,j]| \leq 2m$). Assumption (VAR-EV), on the other hand, is not present in previous CLTs. However, removing this assumption does not significantly impact the resulting Berry–Esseen bound; see Remark 3.3 for more details.

THEOREM 3.1. *Suppose that Assumptions* (MIN-VAR), (MIN-EV) *and* (VAR-EV) *hold. Then, if $m=1$ and $q \geq 3$,*

$$\mu\left(X_{[1,n]}, Y_{[1,n]}\right) \leq \frac{C\log(en)}{\sigma_{\min}}\sqrt{\frac{\log(pn)}{n}}\left[\frac{\overline{L}_3}{\underline{\sigma}^2}\log^2(ep) + \left(\frac{\overline{\nu}_q}{\underline{\sigma}^2}\log^2(ep)\right)^{1/(q-2)}\right],$$

for some universal constant $C > 0$.

The steps and details of the proof for a simplified version can be found in Section 5. However, the complete proof necessitates an additional technique that will be explained in Section 5.6. For the comprehensive proof, please refer to Appendix A.3.

The dimension complexity of the convergence (in distribution) induced by the above theorem is worth noting. For convergence to zero of $\mu(X_{[1,n]}, Y_{[1,n]})$, Theorem 3.1 requires

$$\max\left\{\frac{\overline{L}_3^2}{\sigma_{\min}^2\underline{\sigma}^4}\log^5 p, \frac{\overline{\nu}_q^{2/(q-2)}}{\sigma_{\min}^2\underline{\sigma}^{4/(q-2)}}(\log p)^{4/(q-2)}\right\} = o(n).$$

The dimension complexity may vary depending on the problem or random variables of interest due to the dependency on p of $\nu_{q,j} \equiv \mathbb{E}[\|X_j\|_\infty^q] + \mathbb{E}[\|Y_j\|_\infty^q]$, which changes with the tail behavior of X_j. For instance, in the case of sub-Gaussian X_j, the dimension complexity is determined by the first term, and the requirement becomes $\log^5 p = o(n)$.

Under independence, the best-known dimension complexity in the literature for the same setting was $\log^4 p = o(n)$ (see Chernozhukov, Chetverikov and Koike, 2020, Corollary 2.1; hereafter, we cite this paper by CCK20). Considering the negligible distinction between 1-dependence and independence for large n, the difference in the dimension complexity is nontrivial. The key difference lies in the different settings of the two results. Only requiring $q \geq 3$, our result applies even with a finite third moment of X_j, while CCK20 require the finite fourth moment. Assuming higher-order moments to be finite has been shown to improve the resulting Berry–Esseen bounds. For example, Fang and Koike (2020) demonstrated how the Berry–Esseen bound for convex sets improved assuming finite fourth moments compared to Bentkus (2005) based on finite third moments. The following result provides such an improvement of Theorem 3.1 when $q \geq 4$.

THEOREM 3.2. *Suppose that Assumptions* (MIN-VAR), (MIN-EV) *and* (VAR-EV) *hold. Then, if $m=1$ and $q \geq 4$,*

$$\mu\left(X_{[1,n]}, Y_{[1,n]}\right) \leq \frac{C\log(en)}{\sigma_{\min}}\sqrt{\frac{\log(pn)}{n}}\left[\frac{\overline{L}_3}{\underline{\sigma}^2}\log^{3/2}(ep) + \frac{\overline{L}_4^{1/2}}{\underline{\sigma}}\log(ep) + \left(\frac{\overline{\nu}_q}{\underline{\sigma}^2}\right)^{1/(q-2)}\log(ep)\right],$$

for some constant $C > 0$.

The proof follows the same steps as outlined in Section 5, with the addition of a technique that will be described in detail in Section 5.4. The complete proof can be found in Appendix A.4.

Now Theorem 3.2 induces the following dimension complexity:

$$\max\left\{\frac{\overline{L}_3^2}{\sigma_{\min}^2\underline{\sigma}^4}\log^4 p, \frac{\overline{L}_4}{\sigma_{\min}^2\underline{\sigma}^2}\log^3 p, \frac{\overline{\nu} q^{2/(q-2)}}{\sigma_{\min}^2\underline{\sigma}^{4/(q-2)}}\log^3 p\right\} = o(n).$$

In the case of sub-Gaussian X_j, the dimension complexity is $\log^4 p = o(n)$, which matches that of CCK20. In Section 3.1.1, we will explain how the dimension complexities were obtained and compare the theorems under independence and different tail behaviors of X_j.

REMARK 3.3. Assumption (VAR-EV) is not commonly found in the existing literature and may be considered restrictive in many practical applications. However, the Berry-Esseen bounds in Theorems 3.1 and 3.2 still hold even without this assumption. In the absence of the assumption, the factor $C/\sigma_{\min}$ in Theorems 3.1 and 3.2 is replaced by $C/\min\{\sigma_{\min}, \underline{\sigma}\sqrt{\log(ep)}\}$; see Section 5.5 for a detailed explanation of the changes in the proof.

The Berry-Esseen bounds for $m > 1$ can be obtained as corollaries of Theorems 3.1 and 3.2 using an argument from Theorem 2 of Shergin (1980); also, see Theorem 2.6 of Chen and Shao (2004). Let $n' = \lfloor n/m \rfloor$, $X_i' = X_{((i-1)m, im]}$ for $i \in [1, n')$ and $X_{n'}' = X_{((n'-1)m, n]}$. We define Y_i' similarly for $i \in [1, n']$. Then, $X_1', \ldots, X_{n'}'$ are 1-dependent random vectors in $\mathbb{R}^p$. By applying Theorems 3.1 and 3.2 to $\mu(X_{[1,n']}', Y_{[1,n']}') = \mu(X_{[1,n]}, Y_{[1,n]})$, we obtain the following corollary.

COROLLARY 3.4. *Suppose that Assumptions* (MIN-VAR), (MIN-EV)*, and* (VAR-EV) *hold and that* $m > 1$. *If* $q \geq 3$,

$$\mu\left(X_{[1,n]}, Y_{[1,n]}\right) \leq \frac{C\log(en/m)}{\sigma_{\min}}\sqrt{\frac{\log(pn/m)}{n}}\left[(m+1)^2\frac{\overline{L}_3}{\underline{\sigma}^2}\log^2(ep) + \left((m+1)^{q-1}\frac{\overline{\nu}_q}{\underline{\sigma}^2}\log^2(ep)\right)^{1/(q-2)}\right],$$

for some universal constant $C > 0$. *If* $q \geq 4$,

$$\mu\left(X_{[1,n]}, Y_{[1,n]}\right) \leq \frac{C\log(en/m)}{\sigma_{\min}}\sqrt{\frac{\log(pn/m)}{n}} \times\left[(m+1)^2\frac{\overline{L}_3}{\underline{\sigma}^2}\log^{3/2}(ep) + (m+1)^{3/2}\frac{\overline{L}_4^{1/2}}{\underline{\sigma}}\log(ep) + \left((m+1)^{q-1}\frac{\overline{\nu}_q}{\underline{\sigma}^2}\right)^{1/(q-2)}\log(ep)\right],$$

for some universal constant $C > 0$.

3.1. *Comparison with existing literature.* In this section, we compare our main results with existing results in the high-dimensional CLT literature. In Section 3.1.1, we explore the implication of our results under independence and compare it with the nearly optimal result presented by CCK20. The key finding is that both results imply the same dimension complexity for sub-Weibull X_j (including sub-Gaussian and sub-exponential cases), while our result allows for more general conditions such as finite third-order moments and m-dependence. Additionally, in Section 3.1.2, we demonstrate the significant improvement achieved by our work in the Berry-Esseen bound under m-dependence, surpassing previous works.

3.1.1. *Under independence.* Let's consider the case where the random variables X_i are independent. Since independence holds for all pairs of X_i, regardless of the difference between their indices, we can say that $X_i \perp\!\!\!\perp X_j$ when $|i-j|$ is greater than or equal to 1. Thus, according to the definition of 1-dependence, independence is a special case of 1-dependence. As a result, Theorem 3.1 and Theorem 3.2 readily hold for independent X_i. In this subsection, we compare this implication under independence with the work of CCK20. Specifically, we compare their Theorems 2.1 and 2.2 with our Theorem 3.2 since all three theorems assume the existence of finite fourth moments. It's worth noting that under independence, the assumptions (MIN-EV) and (MIN-VAR) of our result can be expressed as follows:

$$\min_{k\in[p]} \mathrm{Var}[Y^{(k)}_{[i,j]}] \geq \sigma^2_{\min} \cdot |[i,j]|,$$

$$\lambda_{\min}(\mathrm{Var}[Y_{[i,j]}]) \geq \underline{\sigma}^2 \cdot \max\{|[i,j]|-2, 0\},\ \forall i,j.$$

These assumptions require the covariance matrix of $X_{[i,j]}$ to be strongly non-degenerate for all pairs (i,j). This assumption is commonly made in high-dimensional CLTs, although CCK20 employed a weaker version assuming that the covariance matrix of the scaled average is well-approximated by a strictly positive-semidefinite matrix. A recent work of Fang et al. (2023) obtained a Berry–Esseen bound of $1/\sqrt{n}$-rate without the strong non-degeneracy assumption. It would be interesting to explore an extension of their work in our setting. In terms of moment assumption, our result presents a significant improvement. In CCK20, Theorems 2.1 and 2.2 require finite fourth moments. However, our Theorem 3.1 only assumes finite third moments, which means it accommodates scenarios where fourth moments may be infinite. For example, this scenario can occur in high-dimensional linear regression problems where $X_i = \xi_i W_i$ for heavy-tailed univariate errors ξ_i and light-tailed covariates W_i. If ξ_i's have a finite third (conditional on W_i) moment but an infinite fourth moment, then $\mathbb{E}\|X_i\|_\infty^4$ can be infinite; see Chernozhukov et al. (2023, Section 4.1) for an application to penalty parameter selection in lasso using bootstrap.

We now shift our focus to the convergence rate of the established Berry-Esseen bounds. Both theorems establish convergence rates of order $1/\sqrt{n}$ with respect to n up to logarithmic factors when p remains fixed. Therefore, a more meaningful comparison lies in the dimension complexities imposed by the theorems. As previously discussed, these dimension complexities are dependent on the tail behavior of X_j. We assume X_j to be i.i.d. and $\mathrm{Var}[X_j^{(k)}] = 1$ for all $k = 1, \ldots, n$. Let's consider the following two scenarios:

1. $|X_j^{(k)}| \leq B$ for all $j = 1, \ldots, n$ and $k = 1, \ldots, p$ almost surely;
2. $\|X_j^{(k)}\|_{\psi_\alpha} \leq B$ and $\frac{1}{n}\sum_{i=1}^n \mathbb{E}[|X_i^{(k)}|^4] \leq B^2$ for all $j = 1, \ldots, n$ and $k = 1, \ldots, p$, where $\|\cdot\|_{\psi_\alpha}$ is the Orlicz norm with respect to $\psi_\alpha(x) \equiv \exp(x^\alpha)$ for $\alpha \leq 2$.

These scenarios correspond to the first two conditions considered by CCK20, which presented the respective Berry–Esseen bounds as Corollary 2.1. In the first scenario,

$$\mu(X_{[1,n]}, Y_{[1,n]}) \leq \begin{cases} \frac{CB}{\sqrt{n}\underline{\sigma}^2} \log^{3/2}(ep)\log(en), & \text{from Corollary 2.1, CCK20,} \\ \frac{CB}{\sqrt{n}\underline{\sigma}^2} \log^{3/2}(ep)\sqrt{\log(pn)}\log(en), & \text{from our Theorem 3.2.} \end{cases}$$

The resulting dimension complexity of CCK20 is $\log^3 p = o(n)$. They demonstrated that this complexity is optimal in Remark 2.1. For our result, the bound is derived using the fact that $L_{3,j} \leq BL_{2,j} = B$, $L_{4,j} \leq B^2 L_{2,j} = B^2$, and $\nu_{q,j} \leq B^q$ for all $q \geq 4$. The resulting dimension complexity of $\log^4 p = o(n)$. Therefore, our Berry-Esseen bound is suboptimal when X_j are bounded random vectors. Next, suppose that X_j are sub-Weibull(α) for $\alpha \leq 2$.

When $\alpha = 2$, X_j are sub-Gaussian, where $\alpha = 1$ means X_j are sub-exponential. In this scenario,

$$\mu(X_{[1,n]}, Y_{[1,n]}) \leq \begin{cases} \frac{CB}{\sqrt{n}\underline{\sigma}^2}\left(\log(en)\log^{3/2}(ep) + \log^{(3/2+1/\alpha)}(ep)\right), & \text{from Corollary 2.1, CCK20,} \\ \frac{CB}{\sqrt{n}\underline{\sigma}^2}\sqrt{\log(pn)}\log(en)\left(\log^{3/2}(ep) + \log^{(3/2+1/\alpha)}(ep)\right), & \text{from our Theorem 3.2.} \end{cases}$$

Corollary 2.1 of CCK20 presented the result only for sub-Gaussian X_j, but the proof readily extends to the other cases with $\alpha < 2$. This leads to the dimension complexity of $\log^{3+2/\alpha}(ep) = o(n)$. For our result, the bound is derived using the fact that $L_{3,j} \leq \sqrt{L_{2,j}L_{4,j}} \leq B$, $L_{4,j} \leq B^2$, and $\nu_{q,j} \leq CB^q\sigma_{\min}^q \log^{q/\alpha}(ep)$ for any $q \geq 4$ (see Corollary 7.4, Zhang and Chen, 2020). The resulting dimension complexity is the same as CCK20.

3.1.2. *Under m-dependence.* In this section, we compare Corollary 3.4 to existing CLT results for m-dependent random variables. First, we revisit Shergin (1980)'s result in univariate case. In Theorem 2 therein, the author showed that if $\mathbb{E}[|X_i|^q] < \infty$ for $q \geq 3$ and $i = 1, \ldots, n$, then

$$\mu\left(X_{[1,n]}, Y_{[1,n]}\right) \leq C(q, \ldots)\left[(m+1)^{q-1}\frac{\sum_{i=1}^n \mathbb{E}[|X_i|^q]}{(\mathbb{E}[X_{[1,n]}^2])^{q/2}}\right]^{1/(q-2)},‘$$

where $C(q, \ldots)$ is a constant with implicit dependency on q and an additional assumption similar to (MIN-VAR). We note that for univariate X_j, Assumptions (MIN-VAR) and (MIN-EV) are equivalent. Under the notation of (MIN-VAR), because $\mathbb{E}[X_{[1,n]}^2] \geq n\sigma_{\min}^2 = n\underline{\sigma}^2$ and $\mathbb{E}[|X_i|^q] = L_{q,i} = \nu_{q,i}$,

$$\begin{aligned} \mu\left(X_{[1,n]}, Y_{[1,n]}\right) &\leq C(q, \ldots)\left[n^{1-q/2}(m+1)^{q-1}\frac{\overline{L}_q}{\sigma_{\min}^q}\right]^{1/(q-2)} \\ &\leq \frac{C(q, \ldots)}{\sigma_{\min}\sqrt{n}}\left[(m+1)^{q-1}\frac{\overline{L}_q}{\underline{\sigma}^2}\right]^{1/(q-2)}. \end{aligned} \tag{5}$$

The dependency on m cannot be improved following the results of Berk (1973). We observe that given p fixed, the upper bound terms in Corollary 3.4 have the equivalent dependencies on n and m with Eq. (5) up to the logarithmic factor.

In the realm of high-dimensional cases, prior works have established Berry–Esseen bounds with various rates. For instance, Zhang and Cheng (2018) derived a bound with an implicit presentation of the rate of their Berry–Esseen bound and its dependence on sample assumptions in their Appendix A.2. More recently, by employing the large-small-block approach, similar to Romano and Wolf (2000) in univariate cases, Chang, Chen and Wu (2021) derived a bound of order $O(m^{2/3}\text{polyn}(\log pn)n^{-1/6})$ specifically for sub-exponential X_j sub-exponential random vectors in their Section 2.1.2. Notably, Kojevnikov and Song (2022) achieved a rate of $O(\log(ep)^{5/4}n^{-1/4})$ without assuming independence, but their result relied on stronger assumptions such as X_j being a martingale difference sequence, more restrictive than m-dependence.

Our contribution, presented in Corollary 3.4, significantly advances the literature by providing a state-of-the-art rate of $O(m^2\text{polyn}(\log(pn))n^{-1/2})$. This result is achieved under minimal assumptions over the high-dimensional CLT literature, making it a valuable addition to the field.

4. Discussion. We derived a $1/\sqrt{n}$ scaling of Berry–Esseen bound for high-dimensional m-dependent random vectors over hyper-rectangles. This result only required finite third moments of the random vectors, among others like sub-Gaussian or sub-exponential conditions. If p were fixed, the rate implied by our result is $(m+1)^2(\log n/m)^{3/2}/\sqrt{n}$. This rate matches that of Shergin (1980) on 1-dimensional m-dependent random variables, and this dependency on m cannot be improved. Our result supports the high-dimensional CLT over hyper-rectangles under m-dependency between samples.

Our advancement in the Gaussian approximation rate of m-dependent samples could benefit the theoretical analyses under physical dependence frameworks. Zhang and Cheng (2018) introduced the m-approximation technique to study the Gaussian approximation of weakly dependent time series under physical dependence. The technique extends the Berry–Esseen bounds for m-dependent samples to weaker temporal dependencies (see Theorem 2.1 and the end of Section 2.2 therein). Similarly, Chang, Chen and Wu (2021) extended the $n^{-1/6}$ rate under m-dependence to samples with physical dependence. The resulting rate in Theorem 3 was better than the best rate of Zhang and Wu (2017).

Another important future direction is extending our technique to samples with generalized graph dependency. Random vectors $X_1, X_2, \dots, X_n \in \mathbb{R}^p$ are said to have dependency structure defined by graph $G = ([n], E)$ if $X_i \perp\!\!\!\perp X_j$ if $(i,j) \in E$. Graph dependency generalizes m-dependence as a special case by taking $E = \{(i,j) : |i-j| \le m\}$. The only CLT result up to our best knowledge has been Chen and Shao (2004) for 1-dimensional samples with graph dependency. Extending their result to high-dimensional samples has a huge potential to advance statistical analyses on network data, which is another data type with increasing availability.

5. Proof techniques and sketch. The proofs of Theorems 3.1 and 3.2 involve intricate layers of advanced techniques, making them challenging to comprehend at a glance. To aid readers' understanding, we initially establish the simplest version of the proof, focusing on $3 \le q$ and 1-dependence, instead of 1-ring dependence. In this particular case, we have $X_1 \perp\!\!\!\perp X_n$. This results in a similar Berry-Esseen bound with Theorem 3.1, but with

$$L_{3,\max} \equiv \max_{i\in[1,n]} L_{3,i} \quad \text{and} \quad \nu_{q,\max} \equiv \max_{i\in[1,n]} \nu_{q,i}$$

in place of $\overline{L}_3$ and $\overline{\nu}_q$, respectively. In the proof, we use the inductive relationship between anti-concentration probabilities and Kolmogorov-Smirnov statistics. Anti-concentration refers to the probability of a random variable to be contained in a small subset (typically an annulus). An anti-concentration probability bound commonly used in the CLT literature, as well as in this work, is of Gaussian random vector. Nazarov (2003) and Chernozhukov, Chetverikov and Kato (2017b) showed an upper bound for the probability that a Gaussian random vector is contained in $A_{r,\delta} \equiv \{x \in \mathbb{R}^p : x \preceq r + \delta\mathbf{1}\} \setminus \{x \in \mathbb{R}^p : x \preceq r - \delta\mathbf{1}\}$ for $r \in \mathbb{R}^p$ and $\delta \in [0, \infty)$.

LEMMA 5.1 (Gaussian anti-concentration inequality; Nazarov, 2003; Chernozhukov, Chetverikov and Kato, 2017b). *For a random vector $Y \sim N(0, \Sigma)$ in $\mathbb{R}^p$, $r \in \mathbb{R}^p$, and $\delta \in [0, \infty)$,*

$$\mathbb{P}[Y \in A_{r,\delta}] \le C\delta\sqrt{\frac{\log(ep)}{\min_{i=1,\dots,p} \Sigma_{ii}}}$$

for an absolute constant $C > 0$.

The anti-concentration probability of our interest is of $X_{[1,i]}$ conditional on the other X_j. We denote the supremum of the probability over r by

$$\kappa_{[1,i]}(\delta) \equiv \sup_{r\in\mathbb{R}^p} \mathbb{P}[X_{[1,i]} \in A_{r,\delta} | \mathscr{X}_{(i,n]}].$$

We also denote the Kolmogorov-Smirnov statistics of our interest by

$$\mu_{[1,i]} \equiv \mu(X_{[1,i]}, Y_{[1,i]}).$$

We start the proof by deriving the inductive relationship from $\kappa_{[1,i]}(\delta)$ to $\mu_{[1,i]}$: for $i \le n$ and $\delta > \sigma_{\min}$,

$$\sqrt{i}\mu_{[1,i]} \le C(\nu_q, \sigma_{\min}, \underline{\sigma})\delta \log(ep) + C(\nu_q, \underline{\sigma}) \frac{\log(en)(\log(ep))^{3/2}}{\delta} \sup_{j:j<i} \sqrt{j}\kappa_{[1,j]}(\delta). \quad (6)$$

To derive this, we use the Lindeberg swapping technique, similar to that used in Kuchibhotla and Rinaldo (2020). However, their standard approach is restricted by the presence of dependency among the random variables X_j. Our key contribution in this aspect lies in addressing the added complexity due to the dependence structure. For the details, please refer to Section 5.1.

In light of Eq. (6), our objective is to provide an upper bound for the anti-concentration probabilities $\kappa_{[1,j]}(\delta)$. When the random variables X_j are independent, $\kappa_{[1,i]}(\delta)$ represents the marginal anti-concentration probability since the condition on $\mathscr{X}_{(i,n]}$ in the definition of $\kappa_{[1,i]}(\delta)$ can be omitted. Consequently, a straightforward upper bound for $\kappa_{[1,i]}(\delta)$ arises from Lemma 5.1: for $i \le n$,

$$\kappa_{[1,i]}(\delta) \le \mu_{[1,i]} + C\delta\sqrt{\frac{\log(ep)}{\sigma_{\min} n}}. \quad (7)$$

Kuchibhotla and Rinaldo (2020) implicitly employed a dual induction approach using Eqs. (6) and (7) to establish the Berry-Esseen bound with the desired rate over n, i.e., $1/\sqrt{n}$ up to the logsarithm. However, Eq. (7) falls short when dealing with 1-dependence, as $X_{[1,i]}$ becomes dependent on $\mathscr{X}_{(i,n]}$ within the definition of $\kappa_{[1,i]}(\delta)$. In the case of univariate dependent X_j, Chen and Shao (2004) derived a non-inductive upper bound for the conditional anti-concentration probability using a telescoping method (refer to Proposition 3.2 therein). Nevertheless, extending this method to high-dimensional cases presents a non-trivial challenge. In our work, we adopt a similar intuition, but instead of aiming for a non-inductive bound, we establish an inductive relationship from $\mu_{[1,i]}$ to $\kappa_{[1,i]}(\delta)$: for $i < n$ and $\delta > \sigma_{\min}$,

$$\sqrt{i}\kappa_{[1,i]}(\delta) \le C\left(\frac{\delta + \nu_1}{\sigma_{\min}}\sqrt{\log(ep)} + \max_{j:j\le i-2} \sqrt{j}\mu_{[1,j]}\right). \quad (8)$$

This main contribution is summarized as Lemma 5.5 in Section 5.2. Finally, we conduct a dual induction using Eqs. (6) and (8) and conclude the proof.

The proofs of Theorems 3.1 and 3.2 share similar steps but incorporate additional techniques. The explanation of these techniques is provided in the last two subsections. In Section 5.4, we present the iterated Lindeberg swapping method, which helps improve the dimension complexity when a finite fourth moment condition is satisfied. Additionally, in Section 5.6, we introduce permutation arguments to enhance the Berry-Esseen bounds by replacing the maximal moments with the average moments. For the complete proofs, please refer to Appendices A.3 and A.4.

5.1. *Induction from κ to μ for $3 \le q$.* Here we prove the inductive relationship for $\mu_{[1,n]}$, but the same proof applies to $\mu_{[1,i]}$ with any i smaller than n. The quantity we want to control concerns expectations of indicator functions, which are not smooth. For this reason, most proofs of CLTs apply a smoothing to replace indicator functions by smooth functions. We use the mixed smoothing proposed by Chernozhukov, Chetverikov and Koike (2020): namely, for $r \in \mathbb{R}^p$ and $\delta, \phi > 0$,

$$\rho^{\varepsilon}_{r,\phi}(x) \equiv \mathbb{E}[f_{r,\phi}(x + \varepsilon Z)],$$

where

$$f_{r,\phi}(x) \equiv \begin{cases} 1, & \text{if } \max\{x^{(i)} - r^{(i)} : i \in [d]\} < 0, \\ 1 - \phi \max\{x^{(i)} - r^{(i)} : i \in [d]\}, & \text{if } 0 \le \max\{x^{(i)} - r^{(i)} : i \in [d]\} < 1/\phi, \\ 0, & \text{if } 1/\phi \le \max\{x^{(i)} - r^{(i)} : i \in [d]\}. \end{cases}$$

This smoothing comes at the cost of smoothing bias terms, according to Lemma 1 of Kuchibhotla and Rinaldo (2020) and Lemma 2.1 of Chernozhukov, Chetverikov and Koike (2020). Here we summarize the two results into the following lemma.

LEMMA 5.2. *Suppose that X is a p-dimensional random vector, and $Y \sim N(0, \Sigma)$ is a p-dimensional Gaussian random vector. Then, for any $\delta > 0$,*

$$\mu(X,Y) \le C \frac{\delta \log(ep) + \sqrt{\log(ep)}/\phi}{\sqrt{\min_{i=1,\ldots,p} \Sigma_{ii}}} + C \sup_{r \in \mathbb{R}^d} \left| \mathbb{E}[\rho^{\delta}_{r,\phi}(X)] - \mathbb{E}[\rho^{\delta}_{r,\phi}(Y)] \right|.$$

Because $\min_{i=1,\ldots,p} \Sigma_{ii} \ge n\sigma^2_{\min}$, Lemma 5.2 implies

$$\mu_{[1,n]} \le \frac{C}{\sqrt{n}} \frac{\delta \log(ep)}{\sigma_{\min}} + \frac{C}{\sqrt{n}} \frac{\sqrt{\log(ep)}}{\phi \sigma_{\min}} + C \sup_{r \in \mathbb{R}^d} \left| \mathbb{E}[\rho^{\delta}_{r,\phi}(X_{[1,n]})] - \mathbb{E}[\rho^{\delta}_{r,\phi}(Y_{[1,n]})] \right|.$$

Lindeberg swapping. The standard Lindeberg swapping approach upper bounds

$$\sup_{r \in \mathbb{R}^d} \left| \mathbb{E}[\rho^{\delta}_{r,\phi}(X_{[1,n]})] - \mathbb{E}[\rho^{\delta}_{r,\phi}(Y_{[1,n]})] \right|$$

by decomposing it into

$$\begin{aligned} &\sup_{r \in \mathbb{R}^d} \left| \mathbb{E}[\rho^{\delta}_{r,\phi}(X_{[1,n]})] - \mathbb{E}[\rho^{\delta}_{r,\phi}(Y_{[1,n]})] \right| \\ &= \sup_{r \in \mathbb{R}^p} \left| \sum_{j=1}^{n} \mathbb{E}\left[\rho^{\delta}_{r,\phi}(W^{\mathsf{C}}_{[j,j]} + X_j) - \rho^{\delta}_{r,\phi}(W^{\mathsf{C}}_{[j,j]} + Y_j) \right] \right|, \end{aligned} \tag{9}$$

where $W^{\mathsf{C}}_{[i,j]} \equiv X_{[1,i)} + Y_{(j,n]}$, and further bounding the each term by third-order remainder terms after Taylor expansions up to order 3 and the second moment matching between X_j and Y_j. In the subsequent discussion, we will use the notation W as a wildcard, representing either X or Y depending on the context.

$$
\begin{aligned}
&\mathbb{E}[\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}} + X_j)] \\
&= \mathbb{E}[\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}})] + \mathbb{E}\left[\left\langle \nabla\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}}), X_j\right\rangle\right] + \mathbb{E}\left[\frac{1}{2}\left\langle \nabla^2\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}}), X_j^{\otimes 2}\right\rangle\right] \\
&\quad + \mathbb{E}\left[\frac{1}{2}\int_0^1 (1-t)^2\left\langle \nabla^3\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}} + tX_j), X_j^{\otimes 3}\right\rangle dt\right] \qquad (10)\\
&= \mathbb{E}[\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}})] + \frac{1}{2}\left\langle \mathbb{E}[\nabla^2\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}})], \mathbb{E}[X_j^{\otimes 2}]\right\rangle \\
&\quad + \frac{1}{2}\int_0^1 (1-t)^2\mathbb{E}\left[\left\langle \nabla^3\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}} + tX_j), X_j^{\otimes 3}\right\rangle\right] dt,
\end{aligned}
$$

so by $\mathbb{E}[X_j^{\otimes 2}] = \mathbb{E}[Y_j^{\otimes 2}]$,

$$
\begin{aligned}
&\mathbb{E}[\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}} + X_j)] - \mathbb{E}[\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}} + Y_j)] \\
&= \frac{1}{2}\int_0^1 (1-t)^2\mathbb{E}\left[\left\langle \nabla^3\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}} + tX_j), X_j^{\otimes 3}\right\rangle\right] dt \\
&\quad - \frac{1}{2}\int_0^1 (1-t)^2\mathbb{E}\left[\left\langle \nabla^3\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}} + tX_j), Y_j^{\otimes 3}\right\rangle\right] dt.
\end{aligned}
$$

However, in the case of 1-dependence, the second equality of Eq. (10) no longer holds due to the dependency between $W_{[j,j]}^{\mathsf{C}}$ and X_j. To address this issue, we introduce Taylor expansions on $\nabla\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}})$ and $\nabla^2\varphi_r^\varepsilon(W_{[j,j]}^{\mathsf{C}})$ to break the dependency before proceeding with the second-order moment matching. This additional step involves meticulous calculations and lengthy specification of remainder terms. We provide the full details in Appendix C.2. As a result,

$$
\sum_{j=1}^n \mathbb{E}\left[\rho_{r,\phi}^\delta(W_{[j,j]}^{\mathsf{C}} + X_j) - \rho_{r,\phi}^\delta(W_{[j,j]}^{\mathsf{C}} + Y_j)\right] = \sum_{j=1}^n \mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,1)} - \mathfrak{R}_{Y_j}^{(3,1)}\right], \qquad (11)
$$

where $\mathfrak{R}_{X_j}^{(3,1)}$ and $\mathfrak{R}_{Y_j}^{(3,1)}$ are remainder terms of the Taylor expansions specified in Appendix C.2.

Remainder lemma. Then, we upper bound the remainder terms using the upper bounds for the differentials of $\rho_{r,\phi}^\varepsilon$. In particular, CCK20 showed in Lemmas 6.1 and 6.2 that

$$
\sup_{w\in\mathbb{R}^p} \sum_{i_1,\ldots,i_\alpha} \sup_{\|z\|_\infty \le \frac{2\delta}{\sqrt{\log(ep)}}} \left|\nabla^{(i_1,\ldots,i_\alpha)}\rho_{r,\phi}^\delta(w+z)\right| \le C\frac{\phi^\gamma(\log(ep))^{(\alpha-\gamma)/2}}{\delta^{\alpha-\gamma}}
$$

for any $\gamma \in [0,1]$. For the event that $W_{[j,j]}^{\mathsf{C}}$ is in the anulus $A_{r,\delta'}$ for some δ' we will specify shortly, we use the above inequality to bound the remainder term. Out of the event, the differential is sufficiently small. Hence, the upper bounds involves with the conditional anti-concentration probability $\mathbb{P}[W_{[j,j]}^{\mathsf{C}} \in A_{r,\delta'} | X_{[1,j)}]$, resulting into the following lemma.

LEMMA 5.3. *Suppose that Assumption (MIN-EV) holds. For W representing either X or Y and $j \in [1,n]$,*

$$
\left|\mathbb{E}\left[\mathfrak{R}_{W_j}^{(3,1)}\right]\right| \le C\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\left[L_{3,\max} + \phi^{q-3}\nu_{q,\max}\right]\min\{1, \kappa_{[1,j-1)}(\delta_{n-j}^o) + \kappa_j^o\},
$$

where $\delta_{n-j}^2 = \delta^2 + \underline{\sigma}^2 \max\{n-j, 0\}$, $\delta_{n-j}^o = 12\delta_{n-j}\sqrt{\log(pn)}$ *and* $\kappa_j^o = \frac{\delta_{n-j}\log(ep)}{\sigma_{\min}\sqrt{\max\{j,1\}}}$, *as long as* $\delta \geq \sigma_{\min}$ *and* $\phi\delta \geq \frac{1}{\log(ep)}$.

Putting all the results back to $\mu_{[1,n]}$, we get

$$
\begin{aligned}
\mu_{[1,n]} &\leq \frac{C}{\sqrt{n}} \frac{\delta \log(ep)}{\sigma_{\min}} + \frac{C}{\sqrt{n}} \frac{\sqrt{\log(ep)}}{\phi\sigma_{\min}} + C\sum_{j=1}^{n} \left| \mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,1)} - \mathfrak{R}_{Y_j}^{(3,1)} \right] \right| \\
&\leq \frac{C}{\sqrt{n}} \frac{\delta \log(ep)}{\sigma_{\min}} + \frac{C}{\sqrt{n}} \frac{\sqrt{\log(ep)}}{\phi\sigma_{\min}} \\
&\quad + C\sum_{j=1}^{n} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max} \right] \min\{1, \kappa_{[1,j-1)}(\delta_{n-j}^o) + \kappa_j^o\}.
\end{aligned}
$$

Partitioning the sum. Let $J_n \equiv n(1 - \frac{\sigma_{\min}^2}{\underline{\sigma}^2 \log^2(4ep)})$. We note that by Assumption (VAR-EV), $J_n \geq \frac{n}{2}$. The choice of J_n stems from comparing 1 and $\kappa_{[1,j)}(\delta_{n-j}^o) + \kappa_j^o$. In detail, note that

$$\kappa_{[1,j)}(\delta_{n-j}^o) + \kappa_j^o \geq 1$$

if $\kappa_j^o = \frac{\delta_{n-j}\log(ep)}{\sigma_{\min}\sqrt{\max\{j,1\}}} = \frac{\sqrt{\delta^2 + \underline{\sigma}^2 \max\{n-j,0\}}\log(ep)}{\sigma_{\min}\sqrt{\max\{j,1\}}} \geq 1$.

For $j < J_n$,

$$
\begin{aligned}
&\sum_{j<J_n} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max} \right] \min\{1, \kappa_{[1,j-1)}(\delta_{n-j}^o) + \kappa_j^o\} \\
&\leq \sum_{j<J_n} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max} \right] \\
&\leq \frac{C}{\sqrt{n}} \frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\sigma_{\min}} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max} \right],
\end{aligned}
$$

because

$$\sum_{j=1}^{\lfloor J_n \rfloor} \frac{1}{\delta_{n-j}^3} \leq \int_{n-\lfloor J_n \rfloor}^{n} \frac{1}{(\delta^2 + t\underline{\sigma}^2)^{3/2}} dt \leq -\frac{2}{\underline{\sigma}^2\delta_n} + \frac{2}{\underline{\sigma}^2\delta_{n-\lfloor J_n \rfloor}} \leq \frac{C}{\underline{\sigma}^3\sqrt{n - J_n}} \leq \frac{C\log(ep)}{\underline{\sigma}^2\sigma_{\min}\sqrt{n}}. \tag{12}$$

For $j \geq J_n$,

$$
\begin{aligned}
&\sum_{j\geq J_n} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max} \right] \min\{1, \kappa_{[1,j-1)}(\delta_{n-j}^o) + \kappa_j^o\} \\
&\leq \sum_{j\geq J_n} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max} \right] \kappa_{[1,j-1)}(\delta_{n-j}^o) \\
&\quad + \sum_{j\geq J_n} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max} \right] \kappa_j^o.
\end{aligned}
$$

The last term is upper bounded by

$$\sum_{j\geq J_n}\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\left[L_{3,\max}+\phi^{q-3}\nu_{q,\max}\right]\kappa_j^o$$
$$\leq\frac{C}{\sqrt{n}}\frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\sigma_{\min}}\log\left(1+\frac{\sqrt{n}\underline{\sigma}}{\delta}\right)\left[L_{3,\max}+\phi^{q-3}\nu_{q,\max}\right].$$

because

$$\sum_{j=\lceil J_n\rceil}^{n}\frac{\kappa_j^o}{\delta_{n-j}^3}=\sum_{j=\lceil J_n\rceil}^{n}\frac{\log(ep)}{\delta_{n-j}^2\sigma_{\min}\sqrt{j}}\leq\frac{C\log(ep)}{\underline{\sigma}^2\sigma_{\min}\sqrt{n}}\log\left(1+\frac{\sqrt{n}\underline{\sigma}}{\delta}\right). \tag{13}$$

In sum, we obtain the following lemma about the relationship between $\mu_{[1,n]}\equiv\mu(X_{[1,n]},Y_{[1,n]})$ and the conditional anti-concentration probability $\kappa_{[1,i]}(\delta)\equiv\sup_{r\in\mathbb{R}^p}\mathbb{P}[X_{[1,i]}\in A_{r,\delta}|\mathscr{X}_{(i,n]}]$.

LEMMA 5.4. *If Assumptions* (MIN-VAR), (MIN-EV) *and* (VAR-EV) *hold, and* $q\geq 3$, *then for any* $\delta\geq\sigma_{\min}$,

$$\mu_{[1,n]}\leq\frac{C}{\sqrt{n}}\frac{\delta\log(ep)}{\sigma_{\min}}+\frac{C}{\sqrt{n}}\frac{\sqrt{\log(ep)}}{\phi\sigma_{\min}}$$
$$+\frac{C}{\sqrt{n}}\frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\sigma_{\min}}\log\left(1+\frac{\sqrt{n}\underline{\sigma}}{\delta}\right)\left[L_{3,\max}+\phi^{q-3}\nu_{q,\max}\right]$$
$$+C\sum_{j\geq J_n}\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\left[L_{3,\max}+\phi^{q-3}\nu_{q,\max}\right]\kappa_{[1,j-3)}(\delta_{n-j}^o),$$

for some absolute constant $C>0$.

5.2. *Induction from μ to κ.* Having obtained the induction from κ to μ in the previous subsection, we now proceed to obtain an induction from μ to κ. This is the step that was implicitly used in the proofs of high-dimensional CLTs for independent observations (e.g., Kuchibhotla and Rinaldo, 2020). However, as mentioned in Section 5, the dependence between $X_{[1,i]}$ and $\mathscr{X}_{(i,n]}$ in $\kappa_{[1,i]}(\delta)\equiv\sup_{r\in\mathbb{R}^p}\mathbb{P}[X_{[1,i]}\in A_{r,\delta}|\mathscr{X}_{(i,n]}]$ makes the step nontrivial. We make a breakthrough using a similar appraoch described in Section 5.1, where we used the Taylor expansion to eliminate the dependency. However, once again, the conditional anti-concentration probability involves a conditional expectation of an indicator function, which lacks smoothness. So we first apply a smoothing technique to the indicator function and leverage the Taylor expansion on the smoothed indicator, subsequently bounding the resulting remainder terms.

Smoothing. For the conditional anti-concentration probability, we use a standard smoothing, rather than the mixed smoothing we used in Section 5.1. For $r\in\mathbb{R}^p$ and $\delta\in[0,\infty)$, let

$$\varphi_{r,\delta}^{\varepsilon}(x)=\mathbb{E}[\mathbb{I}\{x+\varepsilon Z\in A_{r,\delta}\}],$$

where Z is the p-dimensional standard Gaussian random vector. For some $h > 0$,

$$\begin{aligned}\varphi^{\varepsilon}_{r,\delta}(x) - \mathbb{I}\{x \in A_{r,\delta}\} &= \int (\mathbb{I}\{x + \varepsilon z \in A_{r,\delta}\} - \mathbb{I}\{x \in A_{r,\delta}\})\phi(z)dz \\ &= \int_{\|z\|_\infty \le 10\sqrt{\log(ph)}} (\mathbb{I}\{x + \varepsilon z \in A_{r,\delta}\} - \mathbb{I}\{x \in A_{r,\delta}\})\phi(z)dz \\ &\quad + \int_{\|z\|_\infty > 10\sqrt{\log(ph)}} (\mathbb{I}\{x + \varepsilon z \in A_{r,\delta}\} - \mathbb{I}\{x \in A_{r,\delta}\})\phi(z)dz \\ &\ge -\mathbb{I}\{\|x - \partial A_{r,\delta}\|_\infty \le 10\varepsilon\sqrt{\log(ph)}\}\mathbb{I}\{x \in A_{r,\delta}\} \\ &\quad - \mathbb{P}[\|Z\|_\infty > 10\sqrt{\log(ph)}],\end{aligned}$$

where $\partial A_{r,\delta}$ is the boundary of $A_{r,\delta}$, Z is the p-dimensional standard Gaussian random vector and $\phi(z)$ is the density function of Z. Hence,

$$\begin{aligned}\varphi^{\varepsilon}_{r,\delta}(x) &\ge \mathbb{I}\{x \in A_{r,\delta}\} - \mathbb{I}\{\|x - \partial A_{r,\delta}\|_\infty \le 10\varepsilon\sqrt{\log(ph)}\}\mathbb{I}\{x \in A_{r,\delta}\} \\ &\quad - \mathbb{P}[\|Z\|_\infty > 10\sqrt{\log(ph)}] \\ &= \mathbb{I}\{x \in A_{r,\delta-\varepsilon^o}\} - \frac{1}{h^4},\end{aligned} \tag{14}$$

where $\varepsilon^o = 10\varepsilon\sqrt{\log(ph)}$. On the other hand,

$$\begin{aligned}\varphi^{\varepsilon}_{r,\delta}(x) &\le \mathbb{I}\{x \in A_{r,\delta}\} + \mathbb{I}\{\|x - \partial A_{r,\delta}\|_\infty \le 10\varepsilon\sqrt{\log(ph)}\}\mathbb{I}\{x \notin A_{r,\delta}\} \\ &\quad + \mathbb{P}[\|Z\|_\infty > 10\sqrt{\log(ph)}] \\ &= \mathbb{I}\{x \in A_{r,\delta+\varepsilon^o}\} + \frac{1}{h^4}.\end{aligned} \tag{15}$$

As a result, for any $h > 0$,

$$\begin{aligned}&\mathbb{P}[X_{[1,i]} \in A_{r,\delta} | X_{(i,n]}] \\ &\le \mathbb{E}[\varphi^{\varepsilon}_{r,\delta+\varepsilon^o}(X_{[1,i]}) | \mathscr{X}_{(i,n]}] + \frac{1}{h^4}.\end{aligned} \tag{16}$$

Taylor expansion. Applying the Taylor expansion to $\mathbb{E}[\varphi^{\varepsilon}_{r,\delta+\varepsilon^o}(X_{[1,i]}) | \mathscr{X}_{(i,n]}]$,

$$\begin{aligned}&\mathbb{E}[\varphi^{\varepsilon}_{r,\delta+\varepsilon^o}(X_{[1,i]}) | \mathscr{X}_{(i,n]}] \\ &\le \mathbb{E}\left[\varphi^{\varepsilon}_{r,\delta+\varepsilon^o}(X_{[1,i]} - X_{\{2,i-1\}}) | \mathscr{X}_{(i,n]}\right] + \mathbb{E}\left[\mathfrak{R}_{X_{\{2,i-1\}},1} | \mathscr{X}_{(i,n]}\right]\end{aligned}$$

where

$$\mathfrak{R}_{X_{\{2,i-1\}},1} = \int_0^1 \left\langle \nabla \varphi^{\varepsilon}_{r,\delta+\varepsilon^o}(X_{[1,i]} - tX_{\{2,i-1\}}), X_{\{2,i-1\}} \right\rangle dt$$

First, using Eq. (15),

$$\begin{aligned}&\mathbb{E}[\varphi^{\varepsilon}_{r,\delta+\varepsilon^o}(X_{[1,i]} - X_{\{2,i-1\}}) | \mathscr{X}_{(i,n]}] \\ &\le \mathbb{E}[\mathbb{I}\{X_{[1,i]} - X_{\{2,i-1\}} \in A_{r,\delta+2\varepsilon^o}\} + \frac{1}{h^4} | \mathscr{X}_{(i,n]}] \\ &\le \mathbb{E}[\mathbb{P}[X_{[3,i-2]} \in A_{r_1,\delta+2\varepsilon^o} | \mathscr{X}_{(i-1,n]\cup\{1\}}] | \mathscr{X}_{(i,n]}] + \frac{1}{h^4},\end{aligned}$$

where $r_1 = r - X_1 - X_i$ is a Borel measurable function with respect to $\mathscr{X}_{(i-1,n]\cup\{1\}}$. Because $X_{[3,i-2]} \perp\!\!\!\perp \mathscr{X}_{(i-1,n]\cup\{1\}}$,

$$\begin{aligned}&\mathbb{P}[X_{[3,i-2]} \in A_{r_1,\delta+2\varepsilon^o} | \mathscr{X}_{(i-1,n]\cup\{1\}}]\\ &\leq \mathbb{P}[Y_{[3,i-2]} \in A_{r_1,\delta+2\varepsilon^o}] + 2\mu(X_{[3,i-2]}, Y_{[3,i-2]})\\ &\leq C\frac{\delta + 20\varepsilon\sqrt{\log(ph)}}{\sigma_{\min}}\sqrt{\frac{\log(ep)}{i_0 - 2}} + 2\mu_{[3,i-2]},\end{aligned} \tag{17}$$

almost surely due to the Gaussian anti-concentration inequality (Lemma 5.1).

Bounding the remainder. Bounding the remainder term proceeds similarly with the proof of the remainder lemma (Lemma 5.3), resulting into an upper bound with an conditional anti-concentration probability bound ($\kappa_{(2,i-1)}(\epsilon^o)$ in the following lemma). We relegate the bounding details to Appendix B.4.

LEMMA 5.5. *Suppose that Assumptions* (MIN-VAR) *and* (MIN-EV) *hold. For any* $i \in [3, n)$ *and* $\delta > 0$,

$$\begin{aligned}&\kappa_{[1,i]}(\delta)\\ &\leq C\left(\frac{\sqrt{\log(ep)}}{\varepsilon}\left(\nu_{1,2} + \nu_{1,i-1}\right)\kappa_{(2,i-1)}(\varepsilon^o) + \mu_{(2,i-1)}\right)\\ &\quad + \min\left\{1, C\frac{\delta + 2\varepsilon^o}{\sigma_{\min}}\sqrt{\frac{\log(ep)}{i-2}}\right\} + \frac{C}{\sigma_{\min}}(\nu_{1,2} + \nu_{1,i-1})\sqrt{\frac{\log(ep)}{i-2}},\end{aligned}$$

where $\varepsilon^o \equiv 20\varepsilon\sqrt{\log(p(i_o - 2))}$, *as long as* $\varepsilon \geq \sigma_{\min}$.

This resembles the relationship within κ described in Eq. (3.16) of Chen and Shao (2004). In their work on univariate observations, however, the κ in the righthand side was $\kappa_{[1,i]}(\epsilon^o)$ instead of with the reduced index set $(2, i-1)$. Hence, they could plug-in $\delta = \epsilon^o$ and upper bound $\kappa_{(2,i-1)}(\epsilon^o)$ for some suitable ϵ, which we refered to 'a telescoping method' in Section 5. This resulted into a non-inductive upper bound for $\kappa_{[1,i]}(\delta)$ (see Eqs. (3.17) and (3.18) therein). In our setting, the reduced index set $(2, i-1)$ makes the technique ineffective. Alternatively, we proceed to the dual induction levering on the reduced index set.

5.3. *Dual Induction.* Let our induction hypothesis on n be

$$\sqrt{n}\mu_{[1,n]} \leq \widetilde{\mu}_{1,n} L_{3,\max} + \widetilde{\mu}_{2,n}\nu_{q,\max}^{1/(q-2)}, \tag{HYP-BE-1}$$

where

$$\widetilde{\mu}_{1,n} = \mathfrak{C}_1\frac{(\log(ep))^{3/2}\sqrt{\log(pn)}}{\underline{\sigma}^2\sigma_{\min}}\log(en),$$

$$\widetilde{\mu}_{2,n} = \mathfrak{C}_2\frac{\log(ep)\sqrt{\log(pn)}}{\underline{\sigma}^{2/(q-2)}\sigma_{\min}}\log(en)$$

for universal constants $\mathfrak{C}_1$ and $\mathfrak{C}_2$ whose values do not change in this subsection. If $\mathfrak{C}_1, \mathfrak{C}_2 \geq 1$, then (HYP-BE-1), requiring $\mu_{[1,n]} \leq 1$ only, trivially holds for $n \leq 4$.

Now we consider the case of $n > 4$. Suppose that the induction hypotheses hold for all intervals with lengths smaller than n. We first derive an anti-concentration inequality for any such intervals. Without loss of generality, we only consider the intervals $[1, i]$ with $i < n$. We claim that

$$\sqrt{i}\kappa_{[1,i]}(\delta) \le \widetilde{\kappa}_{1,i} L_{3,\max} + \widetilde{\kappa}_{2,i} \nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{3,i} \nu_{2,\max}^{1/2} + \widetilde{\kappa}_4 \delta, \qquad \text{(HYP-AC-1)}$$

where $\widetilde{\kappa}_{1,i} = \mathfrak{C}_{1,\kappa}\widetilde{\mu}_{1,i}$, $\widetilde{\kappa}_{2,i} = \mathfrak{C}_{2,\kappa}\widetilde{\mu}_{2,i}$, $\widetilde{\kappa}_{3,i} = \mathfrak{C}_{3,\kappa}\frac{\log(ep)\sqrt{\log(pi)}}{\sigma_{\min}}$ and $\widetilde{\kappa}_4 = \mathfrak{C}_{4,\kappa}\frac{\sqrt{\log(ep)}}{\sigma_{\min}}$ for some universal constants $\mathfrak{C}_{1,\kappa}$, $\mathfrak{C}_{2,\kappa}$, $\mathfrak{C}_{3,\kappa}$ and $\mathfrak{C}_{4,\kappa}$ whose values do not change in this subsection. If $\mathfrak{C}_{1,\kappa}, \mathfrak{C}_{2,\kappa}, \mathfrak{C}_{3,\kappa}, \mathfrak{C}_{4,\kappa} \ge 1$, then (HYP-BE-1), requiring $\kappa_{[1,n]}(\delta) \le 1$ for all $\delta > 0$ only, trivially holds for $i \le 16$. For $i > 16$, assume that Eq. (HYP-AC-1) holds for all smaller i's. By Lemma 5.5, for any $\varepsilon \ge \sigma_{\min}$ and $\delta > 0$,

$$\begin{aligned}
&\kappa_{[1,i]}(\delta) \\
&\le C\frac{\sqrt{\log(ep)}}{\varepsilon}\nu_{1,\max}\min\{1, \kappa_{(1,i)}(\varepsilon^o)\} \\
&\quad + C\mu_{(1,i)} + C\frac{\delta + 2\varepsilon^o}{\sigma_{\min}}\sqrt{\frac{\log(ep)}{i-2}} + C\frac{\nu_{1,\max}}{\sigma_{\min}}\frac{\log(ep)}{\sqrt{i-2}},
\end{aligned}$$

where $\varepsilon^o = 20\varepsilon\sqrt{\log(p(i-2))}$ and $C > 0$ is an absolute constant. Due to (HYP-AC-1) and (HYP-BE-1) on interval $(1, i) \subsetneq [1, i]$,

$$\begin{aligned}
&\kappa_{[1,i]}(\delta) \\
&\le \frac{C}{\sqrt{i-2}}\frac{\sqrt{\log(ep)}}{\varepsilon}\nu_{1,\max}\left[\widetilde{\kappa}_{1,i-2}L_{3,\max} + \widetilde{\kappa}_{2,i-2}\nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{3,i-2}\nu_{2,\max}^{1/2} + \widetilde{\kappa}_4\delta\right] \\
&\quad + \frac{C}{\sqrt{i-2}}\widetilde{\mu}_{1,i-2}L_{3,\max} + \frac{C}{\sqrt{i-2}}\widetilde{\mu}_{2,i-2}\nu_{q,\max}^{1/(q-2)} \\
&\quad + C\frac{\delta + 2\varepsilon^o}{\sigma_{\min}}\sqrt{\frac{\log(ep)}{i-2}} + C\frac{\nu_{1,\max}}{\sigma_{\min}}\frac{\log(ep)}{\sqrt{i-2}}.
\end{aligned}$$

As a result, we obtain a recursive inequality on $\widetilde{\kappa}$'s that

$$\begin{aligned}
&\sqrt{i}\kappa_{[1,i]}(\delta) \\
&\le \mathfrak{C}'\frac{\sqrt{\log(ep)}}{\varepsilon}\nu_{1,\max}\left[\widetilde{\kappa}_{1,i-2}L_{3,\max} + \widetilde{\kappa}_{2,i-2}\nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{3,i-2}\nu_{2,\max}^{1/2} + \widetilde{\kappa}_4\delta\right] \\
&\quad + \mathfrak{C}'\left[\widetilde{\mu}_{1,i-2}L_{3,\max} + \widetilde{\mu}_{2,i-2}\nu_{q,\max}^{1/(q-2)} + \frac{\delta + 2\varepsilon^o}{\sigma_{\min}}\sqrt{\log(ep)} + \frac{\nu_{1,\max}}{\sigma_{\min}}\log(ep)\right],
\end{aligned} \tag{18}$$

for some universal constant $\mathfrak{C}'$, whose value does not change in this subsection. Plugging in $\varepsilon = 2\mathfrak{C}'\sqrt{\log(ep)}\nu_{2,\max}^{1/2} \ge 2\mathfrak{C}'\sqrt{\log(ep)}\nu_{1,\max} \ge \sigma_{\min}$,

$$\begin{aligned}
&\sqrt{i}\kappa_{[1,i]}(\delta) \\
&\le \frac{1}{2}\left[\widetilde{\kappa}_{1,i-2}L_{3,\max} + \widetilde{\kappa}_{3,i-2}\nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{3,i-2}\nu_{1,\max} + \widetilde{\kappa}_4\delta\right] \\
&\quad + \mathfrak{C}'\left[\widetilde{\mu}_{1,i-2}L_{3,\max} + \widetilde{\mu}_{2,i-2}\nu_{q,\max}^{1/(q-2)}\right] \\
&\quad + \mathfrak{C}'\frac{\log(ep)}{\sigma_{\min}}\nu_{1,\max} + 40\mathfrak{C}'\frac{\log(ep)\sqrt{\log(p(i-2))}}{\sigma_{\min}}\nu_{2,\max}^{1/2} + \mathfrak{C}'\frac{\sqrt{\log(ep)}}{\sigma_{\min}}\delta \\
&\le \widetilde{\kappa}_{1,i}L_{3,\max} + \widetilde{\kappa}_{2,i}\nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{3,i}\nu_{2,\max}^{1/2} + \widetilde{\kappa}_4\delta,
\end{aligned}$$

where $\mathfrak{C}_{1,\kappa} = \mathfrak{C}_{2,\kappa} = \mathfrak{C}_{4,\kappa} = \max\{2\mathfrak{C}', 1\}$ and $\mathfrak{C}_{3,\kappa} = \max\{82\mathfrak{C}', 1\}$.

Now we prove (HYP-BE-1) on n. We first upper bound the last term in Lemma 5.4:

$$C \sum_{j \geq J_n} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max}\right] \kappa_{[1,j-1)}(\delta_{n-j}^o).$$

Applying (HYP-AC-1) to $\kappa_{[1,j-1)}(\delta_{n-j}^o)$ in $\mathfrak{T}_{2,1}$,

$$\begin{aligned}
&C \sum_{j \geq J_n} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max}\right] \kappa_{[1,j-1)}(\delta_{n-j}^o) \\
&\leq C \sum_{j \geq J_n} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3 \sqrt{\max\{j-2, 1\}}} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max}\right] \\
&\quad \times \left[\widetilde{\kappa}_{1,j-2} L_{3,\max} + \widetilde{\kappa}_{2,j-2}\nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{3,j-2}\nu_{2,\max}^{1/2} + \widetilde{\kappa}_4 \delta_{n-j}^o\right]
\end{aligned}$$

We recall that $\widetilde{\kappa}_{1,n-1} = \mathfrak{C}_{1,\kappa}\widetilde{\mu}_{1,n-1}$, $\widetilde{\kappa}_{2,n-1} = \mathfrak{C}_{2,\kappa}\widetilde{\mu}_{2,n-1}$, $\widetilde{\kappa}_{3,n-1} = \mathfrak{C}_{3,\kappa}\frac{\log(ep)\sqrt{\log(p(n-1))}}{\sigma_{\min}}$ and $\widetilde{\kappa}_4 = \mathfrak{C}_{4,\kappa}\frac{\sqrt{\log(ep)}}{\sigma_{\min}}$. Based on Eq. (15) in Kuchibhotla and Rinaldo (2020), saying

$$\begin{aligned}
&\sum_{j=\lceil J_n \rceil}^{n} \frac{1}{\delta_{n-j}^2} \leq \frac{C}{\underline{\sigma}^2} \log\left(1 + \frac{\sqrt{n}\underline{\sigma}}{\delta}\right), \\
&\sum_{j=\lceil J_n \rceil}^{n} \frac{1}{\delta_{n-j}^3} \leq \frac{2}{\delta\underline{\sigma}^2},
\end{aligned} \tag{19}$$

we obtain

$$\begin{aligned}
&C \sum_{j \geq J_n} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max}\right] \kappa_{[1,j-1)}(\delta_{n-j}^o) \\
&\leq \frac{C}{\sqrt{n}} \frac{(\log(ep))^{3/2}}{\underline{\sigma}^2} \left[L_{3,\max} + \phi^{q-3}\nu_{q,\max}\right] \\
&\quad \times \left[\begin{array}{l} (\widetilde{\mu}_{1,n-1}L_{3,\max} + \widetilde{\mu}_{2,n-1}\nu_{q,\max}^{1/(q-2)})\frac{1}{\delta} + \frac{\log(ep)\sqrt{\log(pn)}}{\sigma_{\min}}\frac{\nu_{2,\max}^{1/2}}{\delta} \\ + \frac{\sqrt{\log(ep)\log(pn)}}{\sigma_{\min}} \log\left(1 + \frac{\sqrt{n}\underline{\sigma}}{\delta}\right) \end{array}\right].
\end{aligned}$$

In sum, as long as $\delta \geq \nu_{2,\max}^{1/2}\sqrt{\log(ep)} \geq \sigma_{\min}$ and $\phi > 0$, we have the recursive inequality on $\widetilde{\mu}$'s that

$$\begin{aligned}
&\sqrt{n}\mu_{[1,n]} \\
&\leq \mathfrak{C}'' \frac{(\log(ep))^{3/2}}{\underline{\sigma}^2\delta}\left[L_{3,\max} + \phi^{q-3}\nu_{q,\max}\right]\left[\widetilde{\mu}_{1,n-1}L_{3,\max} + \widetilde{\mu}_{2,n-1}\nu_{q,\max}^{1/(q-2)}\right] \\
&\quad + \mathfrak{C}''\left[\frac{\delta\log(ep)}{\sigma_{\min}} + \frac{\sqrt{\log(ep)}}{\phi\sigma_{\min}}\right] \\
&\quad + \mathfrak{C}''\frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\sigma_{\min}}\log\left(1 + \frac{\sqrt{n}\underline{\sigma}}{\delta}\right)\left[L_{3,\max} + \phi^{q-3}\nu_{q,\max}\right] \\
&\quad + \mathfrak{C}''\frac{(\log(ep))^{3/2}}{\underline{\sigma}^2}\log\left(1 + \frac{\sqrt{n}\underline{\sigma}}{\delta}\right)\left[L_{3,\max} + \phi^{q-3}\nu_{q,\max}\right]\frac{\sqrt{\log(ep)\log(pn)}}{\sigma_{\min}},
\end{aligned}$$

where $\mathfrak{C}''$ is a universal constant whose value does not change in this subsection. Taking $\delta = \frac{\max\{4\mathfrak{C}'',1\}}{\sqrt{\log(ep)}}\left(\frac{L_{3,\max}}{\underline{\sigma}^2}(\log(ep))^2 + \left(\frac{\nu_{q,\max}}{\underline{\sigma}^2}(\log(ep))^2\right)^{\frac{1}{q-2}}\right) \geq \overline{\nu}_2^{1/2}\sqrt{\log(ep)}$ and $\phi = \frac{1}{\delta\sqrt{\log(ep)}}$,

$$\begin{aligned}
&\sqrt{n}\mu_{[1,n]} \\
&\leq \frac{1}{2}\max_{j<n}\widetilde{\mu}_{1,j}L_{3,\max} + \frac{1}{2}\max_{j<n}\widetilde{\mu}_{2,j}\nu_{q,\max}^{1/(q-2)} \\
&\quad + \mathfrak{C}^{(3)}\left(L_{3,\max}\frac{(\log(ep))^2}{\underline{\sigma}^2} + \nu_{q,\max}^{1/(q-2)}\frac{(\log(ep))^{2/(q-2)}}{\underline{\sigma}^{2/(q-2)}}\right)\frac{\sqrt{\log(pn)}}{\sigma_{\min}}\log(en)
\end{aligned}$$

for another universal constant $\mathfrak{C}^{(3)}$, whose value only depends on $\mathfrak{C}''$. Taking $\mathfrak{C}_1 = \mathfrak{C}_2 = \max\{2\mathfrak{C}^{(3)}, 1\}$,

$$\begin{aligned}
\widetilde{\mu}_{1,n} &= \mathfrak{C}_1\frac{(\log(ep))^2\sqrt{\log(pn)}}{\underline{\sigma}^2\sigma_{\min}}\log(en), \\
\widetilde{\mu}_{2,n} &= \mathfrak{C}_2\frac{(\log(ep))^{2/(q-2)}\sqrt{\log(pn)}}{\underline{\sigma}^{2/(q-2)}\sigma_{\min}}\log(en)
\end{aligned}$$

satisfies

$$\sqrt{n}\mu_{[1,n]} \leq \widetilde{\mu}_{1,n}L_{3,\max} + \widetilde{\mu}_{2,n}\nu_{q,\max}^{1/(q-2)},$$

which proves (HYP-BE-1) at n. This proves our theorem.

5.4. *For* $4 \leq q$. In cases with finite fourth moment, we can obtain a better sample complexity by decomposing the third order remainder $\mathfrak{R}_{W_j}^{(3,1)}$. Based on the Taylor expansions up to order 4,

$$\sum_{j=1}^{n}\mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,1)}-\mathfrak{R}_{Y_j}^{(3,1)}\right]$$

$$\begin{aligned}&=\frac{1}{6}\sum_{j=1}^{n}\Big\langle\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(X_{[1,j-1)}+Y_{(j+1,n]})\right],\mathbb{E}[X_j^{\otimes 3}]\Big\rangle\\&+\frac{1}{2}\sum_{j=1}^{n-1}\langle\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(X_{[1,j-1)}+Y_{(j+2,n]})\right],\mathbb{E}[X_j\otimes X_{j+1}\otimes(X_j+X_{j+1})]\rangle\\&+\sum_{j=1}^{n-2}\langle\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(X_{[1,j-1)}+Y_{(j+3,n]})\right],\mathbb{E}\left[X_j\otimes X_{j+1}\otimes X_{j+2}\right]\rangle\\&+\sum_{j=1}^{n}\mathbb{E}\left[\mathfrak{R}_{X_j}^{(4,1)}-\mathfrak{R}_{Y_j}^{(4,1)}\right],\end{aligned}\tag{20}$$

where $\mathfrak{R}_{X_j}^{(4,1)}$ and $\mathfrak{R}_{Y_j}^{(4,1)}$ are remainder terms of the Taylor expansions specified in Appendix C.2. We re-apply the Lindeberg swapping. For brevity, we only look at the first term, but similar arguments apply to the other third order moment terms. We observe that

$$\begin{aligned}&\Big\langle\mathbb{E}[\nabla^3\rho_{r,\phi}^{\varepsilon}(X_{[1,j-1)}+Y_{(j+1,n)})],\mathbb{E}[X_j^{\otimes 3}]\Big\rangle\\&=\Big\langle\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(Y_{[1,j-1)}+Y_{(j+1,n)})\right],\mathbb{E}[X_j^{\otimes 3}]\Big\rangle\\&\quad+\Big\langle\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(X_{[1,j-1)}+Y_{(j+1,n)})-\nabla^3\rho_{r,\phi}^{\varepsilon}(Y_{[1,j-1)}+Y_{(j+1,n)})\right],\mathbb{E}[X_j^{\otimes 3}]\Big\rangle.\end{aligned}\tag{21}$$

For the first term, because $Y_{[1,j-1)}+Y_{(j+1,n)}$ is Gaussian, Lemma 6.2 in Chernozhukov, Chetverikov and Koike (2020) and Assumption (VAR-EV) imply that

$$\begin{aligned}\left|\Big\langle\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(Y_{[1,j-1)}+Y_{(j+1,n)})\right],\mathbb{E}[X_j^{\otimes 3}]\Big\rangle\right|&\le\frac{C}{n^{3/2}}L_{3,j}\frac{(\log(ep))^{3/2}}{\underline{\sigma}^3}\\&\le\frac{C}{n^{3/2}}L_{3,j}\frac{(\log(ep))^{2}}{\underline{\sigma}^2\sigma_{\min}}.\end{aligned}\tag{22}$$

For the second term, we re-apply the Lindeberg swapping: for $j=3,\dots,n-1$,

$$\begin{aligned}&\Big\langle\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(X_{[1,j-1)}+Y_{(j+1,n)})\right]-\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(Y_{[1,j-1)}+Y_{(j+1,n)})\right],\mathbb{E}[X_j^{\otimes 3}]\Big\rangle.\\&=\sum_{k=1}^{j-2}\Big\langle\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(X_{[1,k)}+X_k+Y_{(k,j-1)\cup(j+1,n)})\right]\\&\qquad-\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(X_{[1,k)}+Y_k+Y_{(k,j-1)\cup(j+1,n)})\right],\mathbb{E}[X_j^{\otimes 3}]\Big\rangle.\end{aligned}\tag{23}$$

By the Taylor expansion cnetered at $X_{[1,k)}+Y_{(k,j-1)\cup(j+1,n]}$, the difference can be rewritten as a sum of sixth order remainder terms:

$$\begin{aligned}&\sum_{j=1}^{n}\Big\langle\mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\varepsilon}(X_{[1,j-1)}+Y_{(j+1,n)})-\nabla^3\rho_{r,\phi}^{\varepsilon}(Y_{[1,j-1)}+Y_{(j+1,n)})\right],\mathbb{E}[X_j^{\otimes 3}]\Big\rangle\\&=\sum_{j=3}^{n}\sum_{k=1}^{j-2}\mathbb{E}\left[\mathfrak{R}_{X_j,X_k}^{(6,1)}-\mathfrak{R}_{X_j,Y_k}^{(6,1)}\right],\end{aligned}\tag{24}$$

where $\mathfrak{R}^{(6,1)}_{X_j,W_k}$ is the sixth order remainder terms. The detail of the expansion and specification of the remainder term are given in Appendix C.4. Hence,

$$\sum_{j=1}^{n}\Big\langle \mathbb{E}\left[\nabla^3\rho^{\varepsilon}_{r,\phi}(X_{[1,j-1)}+Y_{(j+1,n]})\right], \mathbb{E}[X_j^{\otimes 3}]\Big\rangle$$

$$\leq \frac{C}{\sqrt{n}}L_{3,\max}\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}+\sum_{j=3}^{n}\sum_{k=1}^{j-2}\mathbb{E}\left[\mathfrak{R}^{(6,1)}_{X_j,X_k}-\mathfrak{R}^{(6,1)}_{X_j,Y_k}\right].$$

Similarly,

$$\sum_{j=1}^{n-1}\big\langle \mathbb{E}\left[\nabla^3\rho^{\varepsilon}_{r,\phi}(X_{[1,j-1)}+Y_{(j+2,n]})\right], \mathbb{E}[X_j\otimes X_{j+1}\otimes(X_j+X_{j+1})]\big\rangle$$

$$\leq \frac{C}{\sqrt{n}}L_{3,\max}\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}+\sum_{j=3}^{n-1}\sum_{k=1}^{j-2}\mathbb{E}\left[\mathfrak{R}^{(6,2)}_{X_j,X_k}-\mathfrak{R}^{(6,2)}_{X_j,Y_k}\right], \text{ and}$$

$$\sum_{j=1}^{n-2}\big\langle \mathbb{E}\left[\nabla^3\rho^{\varepsilon}_{r,\phi}(X_{[1,j-1)}+Y_{(j+3,n]})\right], \mathbb{E}\left[X_j\otimes X_{j+1}\otimes X_{j+2}\right]\big\rangle$$

$$\leq \frac{C}{\sqrt{n}}L_{3,\max}\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}+\sum_{j=3}^{n-2}\sum_{k=1}^{j-2}\mathbb{E}\left[\mathfrak{R}^{(6,3)}_{X_j,X_k}-\mathfrak{R}^{(6,3)}_{X_j,Y_k}\right],$$

where $\mathfrak{R}^{(6,2)}_{X_j,W_k}$ and $\mathfrak{R}^{(6,3)}_{X_j,W_k}$ are similarly derived sixth-order remainder terms. Putting all above terms together, we get

$$\mu_{[1,n]}\leq \frac{C}{\sqrt{n}}\frac{\delta\log(ep)}{\sigma_{\min}}+\frac{C}{\sqrt{n}}\frac{\sqrt{\log(ep)}}{\phi\sigma_{\min}}+\frac{C}{\sqrt{n}}L_{3,\max}\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}$$
$$+\sum_{j=1}^{n}\Big|\mathbb{E}\left[\mathfrak{R}^{(4,1)}_{X_j}-\mathfrak{R}^{(4,1)}_{Y_j}\right]\Big|+\sum_{j=3}^{n}\sum_{k=1}^{j-2}\Big|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X_j,X_k}-\mathfrak{R}^{(6)}_{X_j,Y_k}\right]\Big|,$$

where $\mathfrak{R}^{(6)}_{X_j,W_k}=\frac{1}{6}\mathfrak{R}^{(6,1)}_{X_j,W_k}+\frac{1}{2}\mathfrak{R}^{(6,2)}_{X_j,W_k}+\mathfrak{R}^{(6,3)}_{X_j,W_k}$. A similar remainder lemma with Lemma 5.3 can be given for $\mathfrak{R}^{(4,1)}_{W_j}$ and $\mathfrak{R}^{(6)}_{X_j,W_k}$ (see Lemma A.1), and summing up the upperbounds iteratively over k and j results in a finite fourth moment version of Lemma 5.4 (see Lemma A.2). Finally, the dual induction of this lemma with Lemma 5.5 proves the desired Berry–Esseen bound with finite fourth moments. The proof details are relegated to Appendix A.1.

REMARK 5.6. As we discussed in Section 3.1.1, the bottleneck of our Berry–Esseen bound is often the first term with the third moment (i.e., L_3). A significant improvement by the iterated Lindeberg swapping is reducing the term's order of $\log(ep)$ from $\sqrt{\log(pn)}\log^2(ep)$ to $\sqrt{\log(pn)}\log^{3/2}(ep)$. One may repeat the Lindeberg swapping to further improve the order. For example, the first term of $\mathfrak{R}^{(6)}_{X_j,W_k}$ is

$$\frac{1}{2}\int_0^1(1-t)^2\Big\langle\nabla^6\rho^{\varepsilon}_{r,\phi}(X_{[1,k)}+tW_k+Y_{(k,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k^{\otimes 3}\Big\rangle\, dt.$$

Based on the Taylor expansion up to order 7,

$$\begin{aligned}
&\frac{1}{2}\int_0^1 (1-t)^2\Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + tW_k + Y_{(k,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k^{\otimes 3}\Big\rangle\, dt \\
&= \frac{1}{6}\Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + Y_{(k,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k^{\otimes 3}\Big\rangle \\
&\quad + \frac{1}{6}\int_0^1 (1-t)^3\Big\langle \nabla^7 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + tW_k + Y_{(k,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k^{\otimes 4}\Big\rangle\, dt \\
&= \frac{1}{6}\Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-1)} + Y_{(k+1,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k^{\otimes 3}\Big\rangle \\
&\quad + \frac{1}{6}\int_0^1 \Big\langle \nabla^7 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-1)} + t(X_{k-1}+Y_{k+1}) + Y_{(k+1,j-1)\cup(j+1,n)}), \\
&\qquad\qquad X_j^{\otimes 3}\otimes W_k^{\otimes 3}\otimes (X_{k-1}+Y_{k+1})\Big\rangle\, dt \\
&\quad + \frac{1}{6}\int_0^1 (1-t)^3\Big\langle \nabla^7 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + tW_k + Y_{(k,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k^{\otimes 4}\Big\rangle\, dt.
\end{aligned}$$

Like Eq. (21), one may decompose the first term and re-apply the Lindeberg swapping:

$$\begin{aligned}
&\frac{1}{6}\Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-1)} + Y_{(k+1,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k^{\otimes 3}\Big\rangle \\
&= \frac{1}{6}\Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(Y_{[1,k-1)} + Y_{(k+1,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k^{\otimes 3}\Big\rangle \\
&\quad + \sum_{l=1}^{k-2}\frac{1}{6}\Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,l)} + X_l + Y_{(l,k-1)} + Y_{(k+1,j-1)\cup(j+1,n)}) \\
&\qquad\qquad - \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,l)} + Y_l + Y_{(l,k-1)} + Y_{(k+1,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k^{\otimes 3}\Big\rangle.
\end{aligned}$$

To make a successful improvement, we recommend using piece-wise quadratic $f_{r,\phi}$, instead of the piece-wise linear one defined in Section 5.1. This choice of $f_{r,\phi}$ allows improved remainder lemmas for the sixth, seventh and ninth-order remainder terms. At the end, infinitely repeating the Lindeberg swapping may improve the order asymptotically to $\sqrt{\log(pn)}\log(ep)$.

In this paper, we do not pursue further refining Theorem 3.2. Because Y_j is Gaussian, the dimension complexity cannot be improved from $\log^4(ep)$ due to the last term with the q-th moment (i.e., ν_q). Hence, further Lindeberg swappings do not help match the dimension complexity of CCK20 under bounded X_j.

5.5. *Without Assumption* (VAR-EV). Assumption (VAR-EV) was invoked at specific points in the proofs for the simplified versions. We made use of this assumption in two key instances: first, in our selection of J_n during the "Partitioning the sum" step in Section 5.1, and second, in obtaining the upper bound in Eq. (22).

For the first case, without (VAR-EV), let $J_n \equiv n\left(1-\min\left\{\frac{1}{2}, \frac{\sigma^2_{\min}}{\underline{\sigma}^2\log^2(4ep)}\right\}\right)$, which gaurantees $J_n \geq \frac{n}{2}$ naturally. Then, the rightmost upper bound in Eq. (12) is $\frac{C}{\underline{\sigma}^3\sqrt{n}}$ rather than

$\frac{C\log(ep)}{\underline{\sigma}^2\sigma_{\min}\sqrt{n}}$. Summing the replaced upper bound over j, we obtain

$$\begin{aligned}
&\sum_{j=1}^{n}\left|\mathbb{E}\left[\mathfrak{R}^{(3,1)}_{X_j}-\mathfrak{R}^{(3,1)}_{Y_j}\right]\right| \\
&\leq \frac{C}{\sqrt{n}}\frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\min\{\sigma_{\min},\underline{\sigma}\sqrt{\log(ep)}\}}\log\left(1+\frac{\sqrt{n}\underline{\sigma}}{\delta}\right)\left[L_{3,\max}+\phi^{q-3}\nu_{q,\max}\right] \\
&\quad + C\sum_{j\geq J_n}\frac{(\log(ep))^{3/2}}{\delta^3_{n-j}}\left[L_{3,\max}+\phi^{q-3}\nu_{q,\max}\right]\kappa_{[1,j-3)}(\delta^o_{n-j})
\end{aligned}$$

The resulting induction lemma from κ to μ is (that is, Lemma 5.4 becomes)

$$\begin{aligned}
\mu_{[1,n]} \leq\ & \frac{C}{\sqrt{n}}\frac{\delta\log(ep)}{\sigma_{\min}}+\frac{C}{\sqrt{n}}\frac{\sqrt{\log(ep)}}{\phi\sigma_{\min}} \\
&+\frac{C}{\sqrt{n}}\frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\min\{\sigma_{\min},\underline{\sigma}\sqrt{\log(ep)}\}}\log\left(1+\frac{\sqrt{n}\underline{\sigma}}{\delta}\right)\left[L_{3,\max}+\phi^{q-3}\nu_{q,\max}\right] \\
&+C\sum_{j\geq J_n}\frac{(\log(ep))^{3/2}}{\delta^3_{n-j}}\left[L_{3,\max}+\phi^{q-3}\nu_{q,\max}\right]\kappa_{[1,j-3)}(\delta^o_{n-j}),
\end{aligned}$$

and the same dual induction dervies the desired conclusion.

For Eq. (22), without (VAR-EV), the rightmost upper bound is $\frac{C}{n^{3/2}}L_{3,j}\frac{(\log(ep))^{3/2}}{\underline{\sigma}^3}$ rather than $\frac{C}{n^{3/2}}L_{3,j}\frac{(\log(ep))^{2}}{\underline{\sigma}^2\sigma_{\min}}$. That is,

$$\left|\left\langle\mathbb{E}\left[\nabla^3\rho^{\varepsilon}_{r,\phi}(Y_{[1,j-1)}+Y_{(j+1,n)})\right],\mathbb{E}[X_j^{\otimes 3}]\right\rangle\right|\leq\frac{C}{n^{3/2}}L_{3,j}\frac{(\log(ep))^{3/2}}{\underline{\sigma}^3}.$$

As a result,

$$\begin{aligned}
\mu_{[1,n]}\leq\ & \frac{C}{\sqrt{n}}\frac{\delta\log(ep)}{\sigma_{\min}}+\frac{C}{\sqrt{n}}\frac{\sqrt{\log(ep)}}{\phi\sigma_{\min}}+\frac{C}{\sqrt{n}}L_{3,\max}\frac{(\log(ep))^2}{\underline{\sigma}^2\min\{\sigma_{\min},\underline{\sigma}\sqrt{\log(ep)}\}} \\
&+\sum_{j=1}^{n}\left|\mathbb{E}\left[\mathfrak{R}^{(4,1)}_{X_j}-\mathfrak{R}^{(4,1)}_{Y_j}\right]\right|+\sum_{j=3}^{n}\sum_{k=1}^{j-2}\left|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X_j,X_k}-\mathfrak{R}^{(6)}_{X_j,Y_k}\right]\right|.
\end{aligned}$$

Moreover, the summation of $\mathfrak{R}^{(4,1)}_{X_j}$ and $\mathfrak{R}^{(6)}_{X_j,X_k}$, we use the same J_n to partition them. The dual induction based on the resulting induction lemma dervies the desired conclusion.

5.6. 1-*ring dependence and permutation argument.* We note that the Lindeberg swapping in Eq. (9) is not symmetric with respect to the indices. The asymmetry resulted in a worse rates in Lemmas 5.4 and 5.5, by having the maximal moment terms, $L_{3,\max}$ and $\nu_{q,\max}$. To obtain an improved Berry–Esseen bounds with averaged moment terms, $\overline{L}_3$ and $\overline{\nu}_q$, as in Theorems 3.1 and 3.2, it is desired to relax the asymmetry in the Lindeberg swappings. One such way is to take the average of the upper bounds over permutations of the indices as done in Deng (2020); Deng and Zhang (2020). However, because 1-dependence is specific to the index ordering, the only permutation preserving the dependence structure is the flipping of the indices (i.e., $X_1\mapsto X_n, X_2\mapsto X_{n-1},\ldots,X_n\mapsto X_1$), which is not sufficient for our purpose. We allow more permutations by weakening the dependence structure to 1-ring dependence. By allowing X_1 and X_n dependent on each other, index rotations (i.e.,

$X_1 \mapsto X_{j_o}, X_2 \mapsto X_{j_o+1}, \ldots, X_n \mapsto X_{j_o-1}$ with some $j_o \in [1,n]$) are added to the catalog of available permutations. By averaging the upper bound in Eq. (9) over the permutations, we obtain

$$\begin{aligned} &\left|\mathbb{E}[\rho^{\delta}_{r,\phi}(X_{[1,n]})] - \mathbb{E}[\rho^{\delta}_{r,\phi}(Y_{[1,n]})]\right| \\ &\le \frac{1}{n}\sum_{j_o=1}^{n}\sum_{j=1}^{n-1}\left|\mathbb{E}\left[\rho^{\delta}_{r,\phi}(X_{(j_o,j_o+j)_n} + X_{[j_o+j]_n} + Y_{(j_o+j,j_o+n]_n})\right.\right. \qquad (25) \\ &\qquad\qquad \left.\left. -\rho^{\delta}_{r,\phi}(X_{(j_o,j_o+j)_n} + Y_{[j_o+j]_n} + Y_{(j_o+j,j_o+n]_n})\right]\right|, \end{aligned}$$

where $[j_o+j]_n$ is j_o+j modulo n, and

$$[i,j]_n \equiv \{[i]_n, [i+1]_n, \ldots, [j-1]_n, [j]_n\}.$$

The subscript n notates that the interval is defined modulo n. If the ambient modulo is obvious, we omit the subscript. The other types of intervals, $(i,j]_n, [i,j)_n$ and $(i,j)_n$, are similarly defined. For full notation details, please refer to Appendix A.2. A similar permutation argument also applies to Lemma 5.5; see Lemma A.6. The dual induction on the resulting induction lemmas proves Theorem 3.1 for $3 \le q < 4$. We relegate the proof deatils to Appendix A.3.

For $4 \le q$, there is the second Lindeberg swapping during the decomposition of the third order remainder terms (e.g., Eq. (23)). The same permutation argument as in Eq. (25) provides the following averaged version: for $3 \le j \le n$,

$$\begin{aligned} &\left\langle \mathbb{E}[\nabla^3\rho^{\delta}_{r,\phi}(X_{[1,j-1)} + Y_{(j+1,n]}) - \nabla^3\rho^{\delta}_{r,\phi}(Y_{[1,j-1)} + Y_{(j+1,n]})], \mathbb{E}[X_j^{\otimes 3}]\right\rangle \\ &\le \frac{1}{j-2}\sum_{k_o=1}^{j-2}\sum_{k=1}^{j-2}\left\langle \mathbb{E}[\nabla^3\rho^{\delta}_{r,\phi}(X_{[k_o,k_o+k)_{j-2}} + X_{[k_o+k]_{j-2}} + Y_{(k_o+k,k_o+j-1)_{j-2}\cup(j,n]})\right. \qquad (26) \\ &\qquad \left. - \nabla^3\rho^{\delta}_{r,\phi}(X_{[k_o,k_o+k)_{j-2}} + Y_{[k_o+k]_{j-2}} + Y_{(k_o+k,k_o+j-1)_{j-2}\cup(j,n]})], \mathbb{E}[X_j^{\otimes 3}]\right\rangle. \end{aligned}$$

Then, the dual induction with Lemma A.6 proves Theorem 3.2 for $4 \le q$. We relegate the proof deatils to Appendix A.4.

REFERENCES

BENTKUS, V. (2005). A Lyapunov-type bound in Rd. *Theory of Probability & Its Applications* **49** 311–323.

BENTKUS, V., GÖTZE, F., PAULAUSKAS, V. and RAČKAUSKAS, A. (2000). The accuracy of Gaussian approximation in Banach spaces. In *Limit theorems of probability theory* 25–111. Springer.

BERK, K. N. (1973). A central limit theorem for m-dependent random variables with unbounded m. *The Annals of Probability* 352–354.

CHANG, J., CHEN, X. and WU, M. (2021). Central limit theorems for high dimensional dependent data. *arXiv preprint arXiv:2104.12929*.

CHEN, L. H. and SHAO, Q.-M. (2004). Normal approximation under local dependence. *The Annals of Probability* **32** 1985–2028.

CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *The Annals of Statistics* **41** 2786–2819.

CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2017a). Central limit theorems and bootstrap in high dimensions. *The Annals of Probability* **45** 2309–2352.

CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2017b). Detailed proof of Nazarov's inequality. *arXiv preprint arXiv:1711.10696*.

CHERNOZHUKOV, V., CHETVERIKOV, D. and KOIKE, Y. (2020). Nearly optimal central limit theorem and bootstrap approximations in high dimensions. *arXiv preprint arXiv:2012.09513*.

CHERNOZHUKOV, V., CHETVERIKOV, D., KATO, K. and KOIKE, Y. (2023). High-dimensional data bootstrap. *Annual Review of Statistics and Its Application* **10** 427–449.

CHERNOZHUOKOV, V., CHETVERIKOV, D., KATO, K. and KOIKE, Y. (2022). Improved central limit theorem and bootstrap approximations in high dimensions. *The Annals of Statistics* **50** 2562–2586.

DENG, H. (2020). Slightly conservative bootstrap for maxima of sums. *arXiv preprint arXiv:2007.15877*.

DENG, H. and ZHANG, C.-H. (2020). Beyond Gaussian approximation: Bootstrap for maxima of sums of independent random vectors. *The Annals of Statistics* **48**.

FANG, X. and KOIKE, Y. (2020). Large-dimensional central limit theorem with fourth-moment error bounds on convex sets and balls. *arXiv preprint arXiv:2009.00339*.

FANG, X. and KOIKE, Y. (2021). High-dimensional central limit theorems by Stein's method. *The Annals of Applied Probability* **31** 1660–1686.

FANG, X., KOIKE, Y., LIU, S.-H. and ZHAO, Y.-K. (2023). High-dimensional Central Limit Theorems by Stein's Method in the Degenerate Case. *arXiv preprint arXiv:2305.17365*.

KOIKE, Y. (2021). Notes on the dimension dependence in high-dimensional central limit theorems for hyperrectangles. *Japanese Journal of Statistics and Data Science* **4** 257–297.

KOJEVNIKOV, D. and SONG, K. (2022). A Berry–Esseen bound for vector-valued martingales. *Statistics & Probability Letters* **186** 109448.

KUCHIBHOTLA, A. K., MUKHERJEE, S. and BANERJEE, D. (2021). High-dimensional CLT: Improvements, non-uniform extensions and large deviations. *Bernoulli* **27** 192–217.

KUCHIBHOTLA, A. K. and RINALDO, A. (2020). High-dimensional CLT for Sums of Non-degenerate Random Vectors: $n^{-(1/2)}$-rate. *arXiv preprint arXiv:2009.13673*.

LANG, S. (2012). *Algebra* **211**. Springer Science & Business Media.

LIU, Z. (2020). Matrix-Variate Graphical Models for High-Dimensional Neural Recordings, PhD thesis, Carnegie Mellon University.

LOPES, M. E. (2022). Central limit theorem and bootstrap approximation in high dimensions: Near 1/n rates via implicit smoothing. *The Annals of Statistics* **50** 2492–2513.

LOPES, M., LIN, Z. and MÜLLER, H. (2020). Bootstrapping max statistics in high dimensions: Near-parametric rates under weak variance decay and application to functional and multinomial data. *The annals of statistics* **48**.

NAZAROV, F. (2003). On the Maximal Perimeter of a Convex Set in $\mathbb{R}^n$ with Respect to a Gaussian Measure. In *Geometric aspects of functional analysis* 169–187. Springer.

NORVAIŠA, R. and PAULAUSKAS, V. (1991). Rate of convergence in the central limit theorem for empirical processes. *Journal of Theoretical Probability* **4** 511–534.

ROMANO, J. P. and WOLF, M. (2000). A more general central limit theorem for m-dependent random variables with unbounded m. *Statistics & probability letters* **47** 115–124.

SHERGIN, V. (1980). On the convergence rate in the central limit theorem for m-dependent random variables. *Theory of Probability & Its Applications* **24** 782–796.

ZHANG, H. and CHEN, S. X. (2020). Concentration inequalities for statistical inference. *arXiv preprint arXiv:2011.02258*.

ZHANG, X. and CHENG, G. (2018). Gaussian approximation for high dimensional vector under physical dependence. *Bernoulli*.

ZHANG, D. and WU, W. B. (2017). Gaussian approximation for high dimensional time series. *The Annals of Statistics* **45** 1895–1919.

APPENDIX A: PROOF OF THEOREMS

A.1. Proof details of Section 5.4. We recall from Section 5.4 that

$$\sup_{r\in\mathbb{R}^d}\Big|\mathbb{E}[\rho^{\delta}_{r,\phi}(X_{[1,n]})]-\mathbb{E}[\rho^{\delta}_{r,\phi}(Y_{[1,n]})]\Big| = \sum_{j=1}^{n}\Big|\mathbb{E}\left[\mathfrak{R}^{(3,1)}_{X_j}-\mathfrak{R}^{(3,1)}_{Y_j}\right]\Big|$$

$$\leq \frac{C}{\sqrt{n}}L_{3,\max}\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}+\sum_{j=1}^{n}\Big|\mathbb{E}\left[\mathfrak{R}^{(4,1)}_{X_j}-\mathfrak{R}^{(4,1)}_{Y_j}\right]\Big|+\sum_{j=3}^{n}\sum_{k=1}^{j-2}\Big|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X_j,X_k}-\mathfrak{R}^{(6)}_{X_j,Y_k}\right]\Big|,$$

where $\mathfrak{R}^{(6)}_{X_j,W_k}=\frac{1}{6}\mathfrak{R}^{(6,1)}_{X_j,W_k}+\frac{1}{2}\mathfrak{R}^{(6,2)}_{X_j,W_k}+\mathfrak{R}^{(6,3)}_{X_j,W_k}$.The upper bounds of the remainder terms are given as the following lemma.

LEMMA A.1. *Suppose that Assumption (MIN-EV) holds. For W representing either X or Y and $j,k\in[1,n]$ such that $k\leq j-2$,*

$$\Big|\mathbb{E}\left[\mathfrak{R}^{(4,1)}_{W_j}\right]\Big|\leq C\phi\left[L_{4,\max}\frac{(\log(ep))^{3/2}}{\delta^3_{n-j}}+\nu_{q,\max}\frac{(\log(ep))^{(q-1)/2}}{\delta^{q-1}_{n-j}}\right]$$
$$\times\min\{1,\kappa_{[1,j-4)}(\delta^o_{n-j})+\kappa^o_j\},$$

$$\Big|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X_j,W_k}\right]\Big|\leq C\phi L_{3,\max}\left[L_{3,\max}\frac{(\log(ep))^{3/2}}{\delta^3_{n-k}}+\nu_{q,\max}\frac{(\log(ep))^{(q-1)/2}}{\delta^{q-1}_{n-k}}\right]$$
$$\times\min\{1,\kappa_{[1,k-3)}(\delta^o_{n-k})+\kappa^o_k\}.$$

where $\delta^2_{n-j}\equiv\delta^2+\underline{\sigma}^2\max\{n-j,0\}$, $\delta^o_{n-j}\equiv 12\delta_{n-j}\sqrt{\log(pn)}$ and $\kappa^o_j\equiv\frac{\delta_{n-j}\log(ep)}{\sigma_{\min}\sqrt{\max\{j,1\}}}$, as long as $\delta\geq\sigma_{\min}$ and $\phi\delta\geq\frac{1}{\log(ep)}$.

Back to Eq. (9), we get

$$\Big|\mathbb{E}[\rho^{\delta}_{r,\phi}(X_{[1,n]})]-\mathbb{E}[\rho^{\delta}_{r,\phi}(Y_{[1,n]})]\Big|$$
$$\leq\frac{C}{\sqrt{n}}\overline{L}_3\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}$$
$$+C\phi\sum_{j=1}^{n}\left[L_{4,\max}\frac{(\log(ep))^{3/2}}{\delta^3_{n-j}}+\nu_{q,\max}\frac{(\log(ep))^{(q-1)/2}}{\delta^{q-1}_{n-j}}\right]\min\{1,\kappa_{[1,j-4)}(\delta^o_{n-j})+\kappa^o_j\}$$
$$+C\phi\sum_{j=3}^{n}L_{3,\max}\sum_{k=1}^{j-2}\left[L_{3,\max}\frac{(\log(ep))^{5/2}}{\delta^5_{n-k}}+\nu_{q,\max}\frac{(\log(ep))^{(q+2)/2}}{\delta^{q+2}_{n-k}}\right]$$
$$\times\min\{1,\kappa_{[1,k-3)]}(\delta^o_{n-k})+\kappa^o_k\},$$

Partitioning the sum. Again, we partition the summations at $J_n = n(1 - \frac{\sigma_{\min}^2}{\underline{\sigma}^2 \log^2(4ep)})$. For the first summation, similar calculations with the finite third moment cases lead to

$$\begin{aligned} & C\phi \sum_{j=1}^{n} \left[L_{4,\max} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} + \nu_{q,\max} \frac{(\log(ep))^{(q-1)/2}}{\delta_{n-j}^{q-1}} \right] \min\{1, \kappa_{j_o+(0,j)}(\delta_{n-j}^o) + \kappa_j^o\} \\ & \le \frac{C\phi}{\sqrt{n}} \left[L_{4,\max} \frac{(\log(ep))^{5/2}}{\underline{\sigma}^2 \sigma_{\min}} + \nu_{q,\max} \frac{(\log(ep))^{(q+1)/2}}{\underline{\sigma}^2 \sigma_{\min} \delta^{q-4}} \right] \log\left(1 + \frac{\sqrt{n}\underline{\sigma}}{\delta}\right) \\ & \quad + C\phi \sum_{j \ge J_n} \left[L_{4,\max} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} + \nu_{q,\max} \frac{(\log(ep))^{(q-1)/2}}{\delta_{n-j}^{q-1}} \right] \kappa_{[1,j)}(\delta_{n-j}^o) \end{aligned}$$

by noting Eq. (12) and that

$$\sum_{j=1}^{\lfloor J_n \rfloor} \frac{1}{\delta_{n-j}^{q-1}} \le \sum_{j=1}^{\lfloor J_n \rfloor} \frac{1}{\delta^{q-4} \delta_{n-j}^3} \le \frac{C \log(ep)}{\underline{\sigma}^2 \sigma_{\min} \delta^{q-4} \sqrt{n}}, \tag{27}$$

$$\sum_{j=\lceil J_n \rceil}^{n} \frac{\kappa_{j-1}^o}{\delta_{n-j}^{q-1}} \le \sum_{j=\lceil J_n \rceil}^{n} \frac{\log(ep)}{\delta^{q-4} \delta_{n-j}^2 \sigma_{\min} \sqrt{j-1}} \le \frac{C \log(ep)}{\underline{\sigma}^2 \sigma_{\min} \delta^{q-4} \sqrt{n}} \log\left(1 + \frac{\sqrt{n}\underline{\sigma}}{\delta}\right), \tag{28}$$

for some universal constant $C > 0$. On the other hand, for the second summation,

$$\begin{aligned} & C\phi \sum_{j=3}^{n} L_{3,\max} \sum_{k=1}^{(j-2)\wedge\lfloor J_n \rfloor} \left[L_{3,\max} \frac{(\log(ep))^{5/2}}{\delta_{n-k}^5} + \nu_{q,\max} \frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}} \right] \\ & \le \frac{C\phi}{\sqrt{n}} \sum_{j=3}^{n} L_{3,\max} \left[L_{3,\max} \frac{(\log(ep))^{7/2}}{\underline{\sigma}^2 \sigma_{\min} \delta_{n-\lfloor J_n \rfloor}^2} + \nu_{q,\max} \frac{(\log(ep))^{(q+4)/2}}{\underline{\sigma}^2 \sigma_{\min} \delta_{n-\lfloor J_n \rfloor}^{q-1}} \right] \\ & \le \frac{C\phi}{\sqrt{n}} L_{3,\max} \left[L_{3,\max} \frac{(\log(ep))^{7/2}}{\underline{\sigma}^4 \sigma_{\min}} + \nu_{q,\max} \frac{(\log(ep))^{(q+4)/2}}{\underline{\sigma}^4 \sigma_{\min} \delta^{q-3}} \right], \end{aligned}$$

where the last inequality comes from $\sum_{j=3}^n \frac{1}{\delta_{n-\lfloor J_n \rfloor}^2} \le \sum_{j=3}^n \frac{1}{(n-\lfloor J_n \rfloor)\underline{\sigma}^2} \le \frac{C}{\underline{\sigma}^2}$, and

$$\begin{aligned} & C\phi \sum_{j=\lceil J_n \rceil}^{n-1} L_{3,\max} \sum_{k=\lceil J_n \rceil}^{j-2} \left[L_{3,\max} \frac{(\log(ep))^{5/2}}{\delta_{n-k}^5} + \nu_{q,\max} \frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}} \right] \kappa_k^o \\ & \le \frac{C\phi}{\sqrt{n}} \sum_{j=\lceil J_n \rceil}^{n-1} L_{3,\max} \left[L_{3,\max} \frac{(\log(ep))^{7/2}}{\underline{\sigma}^2 \sigma_{\min} \delta_{n-j}^2} + \nu_{q,\max} \frac{(\log(ep))^{(q+4)/2}}{\underline{\sigma}^2 \sigma_{\min} \delta_{n-j}^{q-1}} \right] \\ & \le \frac{C\phi}{\sqrt{n}} L_{3,\max} \left[L_{3,\max} \frac{(\log(ep))^{7/2}}{\underline{\sigma}^4 \sigma_{\min}} + \nu_{q,\max} \frac{(\log(ep))^{(q+4)/2}}{\underline{\sigma}^4 \sigma_{\min} \delta^{q-3}} \right] \log\left(1 + \frac{\sqrt{n}\underline{\sigma}}{\delta}\right), \end{aligned}$$

where the inequliaties follow Eqs. (19) and (28), respectively. In sum, we obtain a finite fourth moment version of Lemma 5.4:

LEMMA A.2. *If Assumptions* (MIN-VAR), (MIN-EV) *and* (VAR-EV) *hold, and* $q \geq 4$, *then for any* $\delta \geq \sigma_{\min}$,

$$\begin{aligned}
&\mu_{[1,n]} \\
&\leq \frac{C}{\sqrt{n}}\left[\frac{\delta\log(ep)}{\sigma_{\min}} + \frac{\sqrt{\log(ep)}}{\phi\sigma_{\min}} + L_{3,\max}\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}\right] \\
&\quad + \frac{C\phi}{\sqrt{n}}\left[L_{4,\max}\frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\sigma_{\min}} + \nu_{q,\max}\frac{(\log(ep))^{(q+1)/2}}{\underline{\sigma}^2\sigma_{\min}\delta^{q-4}}\right]\log\left(1+\frac{\sqrt{n}\underline{\sigma}}{\delta}\right) \\
&\quad + \frac{C\phi}{\sqrt{n}}L_{3,\max}\left[L_{3,\max}\frac{(\log(ep))^{7/2}}{\underline{\sigma}^4\sigma_{\min}} + \nu_{q,\max}\frac{(\log(ep))^{(q+4)/2}}{\underline{\sigma}^4\sigma_{\min}\delta^{q-3}}\right]\log\left(1+\frac{\sqrt{n}\underline{\sigma}}{\delta}\right) \\
&\quad + C\phi\sum_{j=\lceil J_n\rceil}^{n}\left[L_{4,\max}\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} + \nu_{q,\max}\frac{(\log(ep))^{(q-1)/2}}{\delta_{n-j}^{q-1}}\right]\kappa_{[1,j-4)}(\delta_{n-j}^o) \\
&\quad + C\phi\sum_{j=\lceil J_n\rceil}^{n}L_{3,\max}\sum_{k=\lceil J_n\rceil}^{j-2}\left[L_{3\max}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5} + \nu_{q,\max}\frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}}\right]\kappa_{[1,k-3)}(\delta_{n-k}^o),
\end{aligned}$$

for some absolute constant $C > 0$.

Dual Induction. In this case, our induction hypothesis on n is

$$\sqrt{n}\mu_{[1,n]} \leq \widetilde{\mu}_{1,n}L_{3,\max} + \widetilde{\mu}_{2,n}L_{4,\max}^{1/2} + \widetilde{\mu}_{3,n}\nu_{q,\max}^{1/(q-2)}, \tag{HYP-BE-2}$$

where

$$\begin{aligned}
\widetilde{\mu}_{1,n} &= \mathfrak{C}_1\frac{(\log(ep))^{3/2}\sqrt{\log(pn)}}{\underline{\sigma}^2\sigma_{\min}}\log(en), \\
\widetilde{\mu}_{2,n} &= \mathfrak{C}_2\frac{\log(ep)\sqrt{\log(pn)}}{\underline{\sigma}\sigma_{\min}}\log(en) \\
\widetilde{\mu}_{3,n} &= \mathfrak{C}_3\frac{\log(ep)\sqrt{\log(pn)}}{\underline{\sigma}^{2/(q-2)}\sigma_{\min}}\log(en)
\end{aligned}$$

for universal constants $\mathfrak{C}_1$, $\mathfrak{C}_2$ and $\mathfrak{C}_3$ whose values do not change in this subsection. If $\mathfrak{C}_1, \mathfrak{C}_2, \mathfrak{C}_3 \geq 1$, then (HYP-BE-2), requiring $\mu_{[1,n]} \leq 1$ only, trivially holds for $n \leq 9$.

Now we consider the case of $n > 9$. A similar induction argument with the finite third moment cases proves

$$\sqrt{i}\kappa_{[1,i]}(\delta) \leq \widetilde{\kappa}_{1,i}L_{3,\max} + \widetilde{\kappa}_{2,i}L_{4,\max}^{1/2} + \widetilde{\kappa}_{3,i}\nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{4,i}\nu_{2,\max}^{1/2} + \widetilde{\kappa}_5\delta, \tag{HYP-AC-2}$$

where $\widetilde{\kappa}_{1,i} = \mathfrak{C}_{1,\kappa}\widetilde{\mu}_{1,i}$, $\widetilde{\kappa}_{2,i} = \mathfrak{C}_{2,\kappa}\widetilde{\mu}_{2,i}$, $\widetilde{\kappa}_{3,i} = \mathfrak{C}_{3,\kappa}\widetilde{\mu}_{3,i}$, $\widetilde{\kappa}_{4,i} = \mathfrak{C}_{4,\kappa}\frac{\log(ep)\sqrt{\log(pi)}}{\sigma_{\min}}$ and $\widetilde{\kappa}_5 = \mathfrak{C}_{5,\kappa}\frac{\sqrt{\log(ep)}}{\sigma_{\min}}$ for some universal constants $\mathfrak{C}_{1,\kappa}$, $\mathfrak{C}_{2,\kappa}$, $\mathfrak{C}_{3,\kappa}$, $\mathfrak{C}_{4,\kappa}$ and $\mathfrak{C}_{5,\kappa}$ whose values do not change in this subsection.

Now we prove (HYP-BE-2) on n. We first upper bound the last two terms in Lemma A.2:

$$C\phi \sum_{j=\lceil J_n\rceil}^{n} \left[L_{4,\max} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} + \nu_{q,\max} \frac{(\log(ep))^{(q-1)/2}}{\delta_{n-j}^{q-1}} \right] \kappa_{[1,j-4)}(\delta_{n-j}^o)$$

$$+ C\phi \sum_{j=\lceil J_n\rceil}^{n} L_{3,\max} \sum_{k=\lceil J_n\rceil}^{j-2} \left[L_{3\max} \frac{(\log(ep))^{5/2}}{\delta_{n-k}^5} + \nu_{q,\max} \frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}} \right] \kappa_{[1,k-3)}(\delta_{n-k}^o).$$

Applying (HYP-AC-2) to $\kappa_{[1,j-4)}(\delta_{n-j}^o)$,

$$\begin{aligned}
&C\phi \sum_{j=\lceil J_n\rceil}^{n} \left[L_{4,\max} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} + \nu_{q,\max} \frac{(\log(ep))^{(q-1)/2}}{\delta_{n-j}^{q-1}} \right] \kappa_{[1,j-4)}(\delta_{n-j}^o) \\
&\le C\phi \sum_{j=\lceil J_n\rceil}^{n} \left[L_{4,\max} \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} + \nu_{q,\max} \frac{(\log(ep))^{(q-1)/2}}{\delta_{n-j}^{q-1}} \right] \\
&\quad \times \left[\widetilde{\kappa}_{1,j-5}\, L_{3,\max} + \widetilde{\kappa}_{2,j-5}\, L_{4,\max}^{1/2} + \widetilde{\kappa}_{3,j-5}\, \nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{4,j-5}\, \nu_{2,\max}^{1/2} + \widetilde{\kappa}_5 \delta_{n-j}^o \right] \\
&\le \frac{C\phi}{\sqrt{n}} \left[L_{4,\max} (\log(ep))^{3/2} + \nu_{q,\max} \frac{(\log(ep))^{(q-1)/2}}{\delta^{q-4}} \right] \\
&\quad \times \left[\begin{array}{l} (\widetilde{\kappa}_{1,j-5}\, L_{3,\max} + \widetilde{\kappa}_{2,j-5}\, L_{4,\max}^{1/2} + \widetilde{\kappa}_{3,j-5}\, \nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{4,j-5}\, \nu_{2,\max}^{1/2}) \displaystyle\sum_{j=\lceil J_n\rceil}^{n} \frac{1}{\delta_{n-j}^3} \\ + \widetilde{\kappa}_5 \sqrt{\log(pn)} \displaystyle\sum_{j=\lceil J_n\rceil}^{n} \frac{1}{\delta_{n-j}^2} \end{array} \right] \\
&\le \frac{C\phi}{\sqrt{n}} \left[L_{4,\max} (\log(ep))^{3/2} + \nu_{q,\max} \frac{(\log(ep))^{(q-1)/2}}{\delta^{q-4}} \right] \\
&\quad \times \left[\begin{array}{l} (\widetilde{\mu}_{1,n-1}\, L_{3,\max} + \widetilde{\mu}_{2,n-1}\, L_{4,\max}^{1/2} + \widetilde{\mu}_{3,n-1}\, \nu_{q,\max}^{1/(q-2)}) \dfrac{1}{\delta} + \dfrac{\log(ep)\sqrt{\log(pn)}}{\sigma_{\min}} \dfrac{\nu_{2,\max}^{1/2}}{\delta} \\ + \dfrac{\sqrt{\log(ep)\log(pn)}}{\sigma_{\min}} \log\left(1 + \dfrac{\sqrt{n}\underline{\sigma}}{\delta}\right) \end{array} \right].
\end{aligned}$$

by Eq. (19) and that $\widetilde{\kappa}_{1,n-1} = \mathfrak{C}_{1,\kappa}\widetilde{\mu}_{1,n-1}$, $\widetilde{\kappa}_{2,n-1} = \mathfrak{C}_{2,\kappa}\widetilde{\mu}_{2,n-1}$, $\widetilde{\kappa}_{3,n-1} = \mathfrak{C}_{3,\kappa}\widetilde{\mu}_{3,n-1}$, $\widetilde{\kappa}_{4,n-1} = \mathfrak{C}_{4,\kappa}\frac{\log(ep)\sqrt{\log(p(n-1))}}{\sigma_{\min}}$ and $\widetilde{\kappa}_5 = \mathfrak{C}_{5,\kappa}\frac{\sqrt{\log(ep)}}{\sigma_{\min}}$. Similarly,

$$
\begin{aligned}
&C\phi \sum_{j=\lceil J_n\rceil}^{n} L_{3,\max} \sum_{k=\lceil J_n\rceil}^{j-2} \left[L_{3\max}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5} + \nu_{q,\max}\frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}}\right] \kappa_{[1,k-3)}(\delta^o_{n-k})\\
&\le C\phi \sum_{j=\lceil J_n\rceil}^{n} L_{3,\max} \sum_{k=\lceil J_n\rceil}^{j-2} \left[L_{3\max}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5} + \nu_{q,\max}\frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}}\right]\\
&\quad\times\left[\widetilde{\kappa}_{1,k-4}\, L_{3,\max} + \widetilde{\kappa}_{2,k-4}\, L_{4,\max}^{1/2} + \widetilde{\kappa}_{3,k-4}\, \nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{4,k-4}\, \nu_{2,\max}^{1/2} + \widetilde{\kappa}_5\delta^o_{n-j}\right]\\
&\le C\phi \sum_{j=\lceil J_n\rceil}^{n} L_{3,\max} \left[L_{3\max}\frac{(\log(ep))^{5/2}}{\delta_{n-j}^3} + \nu_{q,\max}\frac{(\log(ep))^{(q+2)/2}}{\delta_{n-j}^{q}}\right]\\
&\quad\times\left[\widetilde{\kappa}_{1,k-4}\, L_{3,\max} + \widetilde{\kappa}_{2,k-4}\, L_{4,\max}^{1/2} + \widetilde{\kappa}_{3,k-4}\, \nu_{q,\max}^{1/(q-2)} + \widetilde{\kappa}_{4,k-4}\, \nu_{2,\max}^{1/2} + \widetilde{\kappa}_5\delta^o_{n-j}\right]\\
&\le C\phi L_{3,\max}\left[L_{3\max}(\log(ep))^{5/2} + \nu_{q,\max}\frac{(\log(ep))^{(q+2)/2}}{\delta^{q-3}}\right]\\
&\quad\times\left[\begin{array}{l}(\widetilde{\mu}_{1,n-1}\, L_{3,\max} + \widetilde{\mu}_{2,n-1}\, L_{4,\max}^{1/2} + \widetilde{\mu}_{3,n-1}\, \nu_{q,\max}^{1/(q-2)})\frac{1}{\delta} + \frac{\log(ep)\sqrt{\log(pn)}}{\sigma_{\min}}\frac{\nu_{2,\max}^{1/2}}{\delta}\\ + \frac{\sqrt{\log(ep)\log(pn)}}{\sigma_{\min}}\log\left(1+\frac{\sqrt{n}\underline{\sigma}}{\delta}\right)\end{array}\right].
\end{aligned}
$$

In sum, as long as $\delta \ge \nu_{2,\max}^{1/2}\sqrt{\log(ep)} \ge \sigma_{\min}$ and $\phi > 0$,

$$
\begin{aligned}
&\sqrt{n}\mu_{[1,n]}\\
&\le \mathfrak{C}^{(4)}\phi\left[\begin{array}{l}L_{4,\max}\frac{(\log(ep))^{3/2}}{\underline{\sigma}^2\delta} + \nu_{q,\max}\frac{(\log(ep))^{(q-1)/2}}{\underline{\sigma}^2\delta^{q-3}}\\ + L_{3,\max}\left[L_{3,\max}\frac{(\log(ep))^{5/2}}{\underline{\sigma}^4\delta} + \nu_{q,\max}\frac{(\log(ep))^{(q+2)/2}}{\underline{\sigma}^4\delta^{q-2}}\right]\end{array}\right]\\
&\quad\times\left[\widetilde{\mu}_{1,n-1}L_{3,\max} + \widetilde{\mu}_{2,n-1}L_{4,\max}^{1/2} + \widetilde{\mu}_{3,n-1}\nu_{q,\max}^{1/(q-2)}\right]\\
&\quad+\mathfrak{C}^{(4)}\left[\frac{\delta\log(ep)}{\sigma_{\min}} + \frac{\sqrt{\log(ep)}}{\phi\sigma_{\min}} + L_{3,\max}\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}\right]\\
&\quad+\mathfrak{C}^{(4)}\phi\left[\begin{array}{l}L_{4,\max}\frac{(\log(ep))^{3/2}}{\underline{\sigma}^2} + \nu_{q,\max}\frac{(\log(ep))^{(q-1)/2}}{\underline{\sigma}^2\delta^{q-4}}\\ + L_{3,\max}\left[L_{3,\max}\frac{(\log(ep))^{5/2}}{\underline{\sigma}^4} + \nu_{q,\max}\frac{(\log(ep))^{(q+2)/2}}{\underline{\sigma}^4\delta^{q-3}}\right]\end{array}\right]\log\left(1+\frac{\sqrt{n}\underline{\sigma}}{\delta}\right)
\end{aligned}
$$

where $\mathfrak{C}^{(4)}$ is a universal constant whose value does not change in this subsection. Taking $\delta = 8\mathfrak{C}^{(4)}\left(\frac{L_{3,\max}}{\underline{\sigma}^2}\sqrt{\log(ep)} + \left(\frac{L_{4,\max}}{\underline{\sigma}^2}\right)^{\frac{1}{2}} + \left(\frac{\nu_{q,\max}}{\underline{\sigma}^2}\right)^{\frac{1}{q-2}}\right)\sqrt{\log(ep)} \ge \overline{\nu}_2^{1/2}\sqrt{\log(ep)}$ and

$\phi = \frac{1}{\delta\sqrt{\log(ep)}}$,

$$
\begin{aligned}
&\sqrt{n}\mu_{[1,n]} \\
&\leq \frac{1}{2}\max_{j<n}\widetilde{\mu}_{1,j}\overline{L}_3 + \frac{1}{2}\max_{j<n}\widetilde{\mu}_{2,j}\overline{L}_4^{1/2} + \frac{1}{2}\max_{j<n}\widetilde{\mu}_{3,j}\overline{\nu}_q^{1/(q-2)} \\
&\quad + \mathfrak{C}^{(5)}\left(\overline{L}_3\frac{(\log(ep))^{3/2}}{\underline{\sigma}^2} + \overline{L}_4^{1/2}\frac{\log(ep)}{\underline{\sigma}} + \overline{\nu}_q^{1/(q-2)}\frac{\log(ep)}{\underline{\sigma}^{2/(q-2)}}\right)\frac{\sqrt{\log(pn)}}{\sigma_{\min}}\log(en)
\end{aligned}
$$

for another universal constant $\mathfrak{C}^{(5)}$, whose value only depends on $\mathfrak{C}''$. Taking $\mathfrak{C}_1 = \mathfrak{C}_2 = \mathfrak{C}_3 = \max\{2\mathfrak{C}^{(5)}, 1\}$,

$$
\begin{aligned}
\widetilde{\mu}_{1,n} &= \mathfrak{C}_1\frac{(\log(ep))^{3/2}\sqrt{\log(pn)}}{\underline{\sigma}^2\sigma_{\min}}\log(en), \\
\widetilde{\mu}_{2,n} &= \mathfrak{C}_2\frac{\log(ep)\sqrt{\log(pn)}}{\underline{\sigma}\sigma_{\min}}\log(en) \\
\widetilde{\mu}_{3,n} &= \mathfrak{C}_3\frac{\log(ep)\sqrt{\log(pn)}}{\underline{\sigma}^{2/(q-2)}\sigma_{\min}}\log(en)
\end{aligned}
$$

satisfies

$$
\sqrt{n}\mu_{[1,n]} \leq \widetilde{\mu}_{1,n}\overline{L}_3 + \widetilde{\mu}_{2,n}\overline{L}_4^{1/2} + \widetilde{\mu}_{3,n}\overline{\nu}_q^{1/(q-2)},
$$

which proves (HYP-BE-2) at n. This proves our theorem.

A.2. Review on the ring $\mathbb{Z}_n$. To facilitate the notations under permutation arguments, we introduce the notion of integer ring $\mathbb{Z}_n$. Let $\mathbb{Z}_n \equiv \mathbb{Z}/n\mathbb{Z}$ be the ring with additive operation $+$ and multiplicative operation $\cdot$ of modulo n. For brevity, we allow a slight notational conflict to denote the elements by $1, \ldots, n$ such that $n(=0)$ and 1 are additive and multiplicative identities, respectively. It means that the next element of n is 1, which is the same as $n+1$ modulo n. When we need to specify, we denote i modulo n by $[i]_n$. We reference textbooks in abstract algebra such as Lang (2012) for detailed properties of the ring structure.

We also define a distance and intervals in $\mathbb{Z}_n$. For $i, j \in \mathbb{Z}_n$, the distance between i and j is defined as

$$
d(i,j) = \min\{|i - j + kn| : k \in \mathbb{Z}\},
$$

where the operations inside the parentheses are on $\mathbb{Z}$, and the closed interval is defined as

$$
[i,j]_n = \{[i]_n, [i+1]_n, \ldots, [j-1]_n, [j]_n\}.
$$

The subscript n notates that the interval is defined over the ring $\mathbb{Z}_n$. If the ambient ring is obvious, we omit the subscript. The other types of intervals, $(i,j]_n, [i,j)_n$ and $(i,j)_n$, are defined similarly.

A.3. Proof of Theorem 3.1. The proof for 1-ring dependent cases with finite third moments is similar with the 1-dependent cases in Section 5. In 1-ring depent cases, we need to address the additional dependence between X_1 and X_n and the average across the permutations in Section 5.6.

Breaking the ring. First, we note that the Berry-Esseen bound under 1-ring dependence can be reduced to the bound under 1-dependence:

$$\begin{aligned}&\sup_{r\in\mathbb{R}^d}\Big|\mathbb{E}[\rho^\delta_{r,\phi}(X_{[1,n]})]-\mathbb{E}[\rho^\delta_{r,\phi}(Y_{[1,n]})]\Big|\\&\le\sup_{r\in\mathbb{R}^d}\Big|\mathbb{E}[\rho^\delta_{r,\phi}(X_{[1,n)})]-\mathbb{E}[\rho^\delta_{r,\phi}(Y_{[1,n)})]\Big|\\&\quad+\sup_{r\in\mathbb{R}^d}\Big|\mathbb{E}[\rho^\delta_{r,\phi}(X_{[1,n]})]-\mathbb{E}[\rho^\delta_{r,\phi}(X_{[1,n)})]\Big|+\sup_{r\in\mathbb{R}^d}\Big|\mathbb{E}[\rho^\delta_{r,\phi}(Y_{[1,n]})]-\mathbb{E}[\rho^\delta_{r,\phi}(Y_{[1,n)})]\Big|.\end{aligned}$$

Note that we removed X_n and Y_n from $X_{[1,n]}$ and $Y_{[1,n]}$, respectively, to break the 1-ring dependence. By the Taylor expansion centered at $X_{[1,n)}$,

$$\begin{aligned}&\rho^\delta_{r,\phi}(X_{[1,n]})-\rho^\delta_{r,\phi}(X_{[1,n)})\\&=\frac{1}{2}\Big\langle\nabla^2\rho^\delta_{r,\phi}(X_{(1,n-1)}),X_n^{\otimes 2}\Big\rangle+\Big\langle\nabla^2\rho^\delta_{r,\phi}(X_{(2,n-1)}),X_n\otimes X_1\Big\rangle\\&\quad+\Big\langle\nabla^2\rho^\delta_{r,\phi}(X_{(1,n-2)}),X_{n-1}\otimes X_n\Big\rangle+\mathfrak{R}^{(3)}_X,\end{aligned}\tag{29}$$

where $\mathfrak{R}^{(3)}_X$ is specified in Appendix C. This is the same for $\rho^\delta_{r,\phi}(Y_{[1,n]})-\rho^\delta_{r,\phi}(Y_{[1,n)})$ but with Y in place of X.

First Lindeberg swapping. We bound $\sup_{r\in\mathbb{R}^d}\Big|\mathbb{E}[\rho^\delta_{r,\phi}(X_{[1,n)})]-\mathbb{E}[\rho^\delta_{r,\phi}(Y_{[1,n)})]\Big|$ by the Lindeberg swapping as in Section 5.1. Here we define

$$W^{\mathsf{C}}_{[i,j]}\equiv X_{[1,i)}+Y_{(j,n)}.$$

Note that unlike Section 5.1, the n-th element is removed. Then,

$$\sum_{j=1}^{n-1}\mathbb{E}\left[\rho^\delta_{r,\phi}(W^{\mathsf{C}}_{[j,j]}+X_j)-\rho^\delta_{r,\phi}(W^{\mathsf{C}}_{[j,j]}+Y_j)\right]=\sum_{j=1}^{n-1}\mathbb{E}\left[\mathfrak{R}^{(3,1)}_{X_j}-\mathfrak{R}^{(3,1)}_{Y_j}\right],\tag{30}$$

where $\mathfrak{R}^{(3,1)}_{X_j}$ and $\mathfrak{R}^{(3,1)}_{Y_j}$ are remainder terms of the Taylor expansions specified in Appendix C.2.

Second moment decomposition and second Lindeberg swapping. To bound the second moment terms in Eq. (29), we re-apply the Lindeberg swapping. For simplicity, we only look at $\Big\langle\nabla^2\rho^\delta_{r,\phi}(X_{(1,n-1)}),X_n^{\otimes 2}\Big\rangle$, but similar arguments work for the other second moment terms. Because $\mathbb{E}[X_n^{\otimes 2}]=\mathbb{E}[Y_n^{\otimes 2}]$,

$$\begin{aligned}&\Big\langle\mathbb{E}\left[\nabla^2\rho^\delta_{r,\phi}(X_{(1,n-1)})\right],\mathbb{E}\left[X_n^{\otimes 2}\right]\Big\rangle-\Big\langle\mathbb{E}\left[\nabla^2\rho^\delta_{r,\phi}(Y_{(1,n-1)})\right],\mathbb{E}\left[Y_n^{\otimes 2}\right]\Big\rangle\\&=\sum_{j=2}^{n-2}\Big\langle\mathbb{E}\left[\nabla^2\rho^\delta_{r,\phi}(X_{(1,j)}+X_j+Y_{(j,n-1)})-\nabla^2\rho^\delta_{r,\phi}(X_{(1,j)}+Y_j+Y_{(j,n-1)})\right],\mathbb{E}\left[X_n^{\otimes 2}\right]\Big\rangle,\end{aligned}$$

By the Taylor expansion up to order 3,

$$\begin{aligned}&\Big\langle\mathbb{E}\left[\nabla^2\rho^\delta_{r,\phi}(X_{(1,n-1)})\right],\mathbb{E}\left[X_n^{\otimes 2}\right]\Big\rangle-\Big\langle\mathbb{E}\left[\nabla^2\rho^\delta_{r,\phi}(Y_{(1,n-1)})\right],\mathbb{E}\left[X_n^{\otimes 2}\right]\Big\rangle\\&=\sum_{j=2}^{n-2}\mathbb{E}\left[\mathfrak{R}^{(3,2,1)}_{X_j}-\mathfrak{R}^{(3,2,1)}_{Y_j}\right],\end{aligned}$$

where $\mathfrak{R}_{X_j}^{(3,2,1)}$ is the third-order remainder, specified in Appendix C.3. Similarly,

$$\left\langle \mathbb{E}\left[\nabla^2 \rho_{r,\phi}^{\delta}(X_{(2,n-1)})\right], \mathbb{E}\left[X_n \otimes X_1\right]\right\rangle - \left\langle \mathbb{E}\left[\nabla^2 \rho_{r,\phi}^{\delta}(Y_{(2,n-1)})\right], \mathbb{E}\left[X_n \otimes X_1\right]\right\rangle$$
$$= \sum_{j=3}^{n-2} \mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,2,2)} - \mathfrak{R}_{Y_j}^{(3,2,2)}\right], \quad \text{and}$$

$$\left\langle \mathbb{E}\left[\nabla^2 \rho_{r,\phi}^{\delta}(X_{(1,n-2)})\right], \mathbb{E}\left[X_n \otimes X_{n-1}\right]\right\rangle - \left\langle \mathbb{E}\left[\nabla^2 \rho_{r,\phi}^{\delta}(Y_{(1,n-2)})\right], \mathbb{E}\left[X_n \otimes X_{n-1}\right]\right\rangle$$
$$= \sum_{j=2}^{n-3} \mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,2,3)} - \mathfrak{R}_{Y_j}^{(3,2,3)}\right],$$

where $\mathfrak{R}_{W_j}^{(3,2,2)}$ and $\mathfrak{R}_{W_j}^{(3,2,3)}$ are simlarly derived third-order remainder terms. Putting all the above terms together,

$$\rho_{r,\phi}^{\delta}(X_{[1,n]}) - \rho_{r,\phi}^{\delta}(X_{[1,n)}) - \rho_{r,\phi}^{\delta}(Y_{[1,n]}) + \rho_{r,\phi}^{\delta}(Y_{[1,n)})$$
$$= \mathfrak{R}_X^{(3)} - \mathfrak{R}_Y^{(3)} + \sum_{j=2}^{n-2} \mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,2)} - \mathfrak{R}_{Y_j}^{(3,2)}\right],$$

where $\mathfrak{R}_{W_j}^{(3,2)} = \frac{1}{2}\mathfrak{R}_{W_j}^{(3,2,1)} + \mathfrak{R}_{W_j}^{(3,2,2)} + \mathfrak{R}_{W_j}^{(3,2,3)}$, and

$$\left|\mathbb{E}[\rho_{r,\phi}^{\delta}(X_{[1,n]})] - \mathbb{E}[\rho_{r,\phi}^{\delta}(Y_{[1,n]})]\right|$$
$$\le \left|\rho_{r,\phi}^{\delta}(X_{[1,n)}) - \rho_{r,\phi}^{\delta}(Y_{[1,n)})\right| + \left|\rho_{r,\phi}^{\delta}(X_{[1,n]}) - \rho_{r,\phi}^{\delta}(X_{[1,n)}) - \rho_{r,\phi}^{\delta}(Y_{[1,n]}) + \rho_{r,\phi}^{\delta}(Y_{[1,n)})\right|$$
$$\le \left|\mathfrak{R}_X^{(3)} - \mathfrak{R}_Y^{(3)}\right| + \sum_{j=1}^{n-1}\left|\mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,1)} - \mathfrak{R}_{Y_j}^{(3,1)}\right]\right| + \sum_{j=2}^{n-2}\left|\mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,2)} - \mathfrak{R}_{Y_j}^{(3,2)}\right]\right|.$$

Remainder lemma. Similar to Section 5.1, the remainder terms $\mathfrak{R}_W^{(3)}$, $\mathfrak{R}_{W_j}^{(3,1)}$ and $\mathfrak{R}_{W_j}^{(3,2)}$ are upper bounded by conditional anti-concentration probability bounds. For $q > 0$, let

$$\widetilde{L}_{q,j} \equiv \sum_{j'=j-3}^{j+3} L_{q,j'} \quad \text{and} \quad \widetilde{L}_{q,[k]_{j-2}} \equiv \sum_{k'=k-3}^{k+3} L_{q,[k']_{j-2}},$$

where $[k']_{j-2}$ is k' modulo $j-2$, and $\widetilde{\nu}_{q,j}$ and $\widetilde{\nu}_{q,[k]_{j-2}}$ are similarly defined.

LEMMA A.3. *Suppose that Assumption (MIN-EV) holds. For W representing either X or Y and $j \in \mathbb{Z}_n$,*

$$\left|\mathbb{E}\left[\mathfrak{R}_W^{(3)}\right]\right| \le C\left[\widetilde{L}_{3,n} + \phi^{q-3}\widetilde{\nu}_{q,n}\right]\frac{(\log(ep))^{3/2}}{\delta^3}\min\{\kappa_{(2,n-1)}(\delta_0^o) + \kappa_n^o, 1\},$$

$$\left|\mathbb{E}\left[\mathfrak{R}_{W_j}^{(3,1)}\right]\right| \le C\left[\widetilde{L}_{3,j} + \phi^{q-3}\widetilde{\nu}_{q,j}\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\min\{\kappa_{[1,j-3)}(\delta_{n-j}^o) + \kappa_j^o, 1\},$$

$$\left|\mathbb{E}\left[\mathfrak{R}_{W_j}^{(3,2)}\right]\right| \le C\left[\widetilde{L}_{3,j} + \widetilde{L}_{3,n} + \phi^{q-3}\left(\widetilde{\nu}_{q,j} + \nu_{q,n}\right)\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}$$
$$\times \min\{\kappa_{(1,j-2)}(\delta_{n-j}^o) + \kappa_j^o, 1\},$$

where $\delta_{n-j}^2 \equiv \delta^2 + \underline{\sigma}^2 \max\{n-j, 0\}$, $\delta_{n-j}^o \equiv 12\delta_{n-j}\sqrt{\log(pn)}$ *and* $\kappa_j^o \equiv \frac{\delta_{n-j}\log(ep)}{\sigma_{\min}\sqrt{\max\{j,1\}}}$, *as long as* $\delta \geq \sigma_{\min}$ *and* $\phi\delta \geq \frac{1}{\log(ep)}$.

Permutation argument. We apply the permutation arguement to Eq. (9) as in Eq. (25).

$$\begin{aligned}&\left|\mathbb{E}[\rho_{r,\phi}^{\delta}(X_{[1,n]})] - \mathbb{E}[\rho_{r,\phi}^{\delta}(Y_{[1,n]})]\right| \\ &\leq \frac{1}{n}\sum_{j_o=1}^{n}\sum_{j=1}^{n-1}\Big|\mathbb{E}\Big[\rho_{r,\phi}^{\delta}(X_{(j_o,j_o+j)} + X_{j_o+j} + Y_{(j_o+j,j_o+n]}) \\ &\qquad\qquad - \rho_{r,\phi}^{\delta}(X_{(j_o,j_o+j)} + Y_{j_o+j} + Y_{(j_o+j,j_o+n]})\Big]\Big|,\end{aligned}$$

where the indices of X and Y are defined in $\mathbb{Z}_n$. Together with the results in Lemma A.3,

$$\begin{aligned}&\left|\mathbb{E}[\rho_{r,\phi}^{\delta}(X_{[1,n]})] - \mathbb{E}[\rho_{r,\phi}^{\delta}(Y_{[1,n]})]\right| \\ &\leq \frac{C}{n}\sum_{j_o=1}^{n}\left[\widetilde{L}_{3,j_o} + \phi^{q-3}\widetilde{\nu}_{q,j_o}\right]\frac{(\log(ep))^{3/2}}{\delta^3}\min\{1, \kappa_{j_o+(2,n-1)}(\delta_0^o) + \kappa_n^o\} \\ &\quad + \frac{C}{n}\sum_{j_o=1}^{n}\sum_{j=2}^{n-2}\left[\widetilde{L}_{3,j_o} + \phi^{q-3}\widetilde{\nu}_{q,j_o}\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\min\{1, \kappa_{j_o+(1,j-2)}(\delta_{n-j}^o) + \kappa_j^o\} \\ &\quad + \frac{C}{n}\sum_{j_o=1}^{n}\sum_{j=1}^{n-1}\left[\widetilde{L}_{3,j_o+j} + \phi^{q-3}\widetilde{\nu}_{q,j_o+j}\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\min\{1, \kappa_{j_o+(1,j-3)}(\delta_{n-j}^o) + \kappa_j^o\},\end{aligned}$$

where $j_o + (2, n-1)$ is the shifted interval of $(2, n-1)$ by j_o in $\mathbb{Z}_n$, namely, $\{[j_o + 3]_n, \ldots, [j_o + n - 2]_n\}$.

Partitioning the sum. We partition the summation over j at $J_n = n(1 - \frac{\sigma_{\min}^2}{\underline{\sigma}^2\log^2(4ep)})$. A notable distinction from Section 5.1 is that we should also take averages over j_o alongside the summations over j. For $j < J_n$,

$$\begin{aligned}&\frac{C}{n}\sum_{j_o=1}^{n}\sum_{j<J_n}\left[\widetilde{L}_{3,j_o+j} + \phi^{q-3}\widetilde{\nu}_{q,j_o+j}\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\min\{1, \kappa_{j_o+(1,j-3)}(\delta_{n-j}^o) + \kappa_j^o\} \\ &\leq \frac{C}{n}\sum_{j_o=1}^{n}\sum_{j<J_n}\left[\widetilde{L}_{3,j_o+j} + \phi^{q-3}\widetilde{\nu}_{q,j_o+j}\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \\ &= C\sum_{j<J_n}\left[\overline{L}_3 + \phi^{q-3}\overline{\nu}_q\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \\ &\leq \frac{C}{\sqrt{n}}\left[\overline{L}_3 + \phi^{q-3}\overline{\nu}_q\right]\frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\sigma_{\min}},\end{aligned}$$

because of Eq. (12). For $j \geq J_n$,

$$\frac{C}{n}\sum_{j_o=1}^{n}\sum_{j\geq J_n}\left[\widetilde{L}_{3,j_o+j}+\phi^{q-3}\widetilde{\nu}_{q,j_o+j}\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\min\{1,\kappa_{j_o+(0,+j)}(\delta_{n-j}^o)+\kappa_j^o\}$$

$$\leq \frac{C}{n}\sum_{j_o=1}^{n}\sum_{j\geq J_n}\left[\widetilde{L}_{3,j_o+j}+\phi^{q-3}\widetilde{\nu}_{q,j_o+j}\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\kappa_{(j_o,j_o+j)}(\delta_{n-j}^o)$$

$$+C\sum_{j\geq J_n}\left[\overline{L}_3+\phi^{q-3}\overline{\nu}_q\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\kappa_{j-1}^o.$$

The last term is upper bounded by

$$C\sum_{j\geq J_n}\left[\overline{L}_3+\phi^{q-3}\overline{\nu}_q\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\kappa_{j-1}^o$$

$$\leq \frac{C}{\sqrt{n}}\left[\overline{L}_3+\phi^{q-3}\overline{\nu}_q\right]\frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\sigma_{\min}}\log\left(1+\frac{\sqrt{n}\overline{\sigma}}{\delta}\right)$$

because of Eq. (13). In sum, we obtain the following induction from $\kappa_I(\delta)$ for $I \subset \mathbb{Z}_n$ to

$$\mu_{[1,n]} = \mu(X_{[1,n]}, Y_{[1,n]}).$$

LEMMA A.4. *If Assumptions* (MIN-VAR), (MIN-EV) *and* (VAR-EV) *hold, then for any* $\delta \geq \sigma_{\min}$,

$$\mu_{[1,n]}$$

$$\leq \frac{C}{\sqrt{n}}\frac{\delta\log(ep)+\sqrt{\log(ep)}/\phi}{\sigma_{\min}}+\frac{C}{\sqrt{n}}\left[\overline{L}_3+\phi^{q-3}\overline{\nu}_q\right]\frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\sigma_{\min}}\log(en)$$

$$+\frac{C}{n}\sum_{j_o=1}^{n}\sum_{j=\lceil J_n\rceil}^{n}\left[\widetilde{L}_{3,j_o}+\phi^{q-3}\widetilde{\nu}_{q,j_o}\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\kappa_{j_o+(3,j-2)}(\delta_{n-j}^o)$$

$$+\frac{C}{n}\sum_{j_o=1}^{n}\sum_{j=\lceil J_n\rceil}^{n-1}\left[\widetilde{L}_{3,j_o+j}+\phi^{q-3}\widetilde{\nu}_{q,j_o+j}\right]\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\kappa_{j_o+(1,j-4)}(\delta_{n-j}^o),$$

for some absolute constant $C > 0$.

Anti-concentration inequality. We note the following monotonic property of κ:

LEMMA A.5. *Suppose that* I_0 *and* I *are intervals in* $\mathbb{Z}_n$ *satisfying* $I' \subset I$ *and that* $\delta' \geq \delta > 0$. *Then,* $\kappa_I(\delta) \leq \kappa_{I'}(\delta')$.

PROOF.

$$\kappa_{[1,i]}(\delta) = \sup_{r\in\mathbb{R}^p}\mathbb{P}[X_I \in A_{r,\delta}|\mathscr{X}_{\mathbb{Z}_n\setminus I}]$$

$$= \sup_{r\in\mathbb{R}^p}\mathbb{E}\left[\mathbb{P}[X_{I'} \in A_{r-X_{I\setminus I'},\delta}|\mathscr{X}_{\mathbb{Z}_n\setminus I'}]\middle|\mathscr{X}_{\mathbb{Z}_n\setminus I}\right]$$

$$\leq \sup_{r\in\mathbb{R}^p}\mathbb{E}\left[\mathbb{P}[X_{[1,i_o]} \in A_{r-X_{I\setminus I'},\delta'}|\mathscr{X}_{\mathbb{Z}_n\setminus I'}]\middle|\mathscr{X}_{\mathbb{Z}_n\setminus I}\right]$$

$$\leq \sup_{r\in\mathbb{R}^p}\mathbb{E}\left[\kappa_{I'}(\delta')\middle|\mathscr{X}_{\mathbb{Z}_n\setminus I}\right] = \kappa_{I'}(\delta').$$

□

Thus, for any $\delta > 0$, $\kappa_{[1,i]}(\delta) \leq \sum_{j=1}^{i_o-1} \kappa_{[j,j+i_o+1]}$ where $i_o = \lfloor \frac{i}{2} \rfloor$. Thus, we obtain the following averaged version of Lemma 5.5.

LEMMA A.6. *Suppose that Assumptions (MIN-VAR) and (MIN-EV) hold. For any $i \in [6, n)$ and $\delta > 0$,*

$$\kappa_{[1,i]}(\delta) \leq \frac{C}{i_o - 1} \sum_{j=1}^{i_o-1} \left(\frac{\sqrt{\log(ep)}}{\varepsilon} \left(\nu_{1,j+1} + \nu_{1,j+i_o}\right) \kappa_{j+(1,i_o)}(\varepsilon^o) + \mu_{j+(1,i_o)} \right) + \min\left\{1, C\frac{\delta + 2\varepsilon^o}{\sigma_{\min}} \sqrt{\frac{\log(ep)}{i_o - 2}}\right\} + \frac{C}{\sigma_{\min}} \overline{\nu}_{1,(1,i)} \sqrt{\frac{\log(ep)}{i_o - 2}},$$

where $i_o \equiv \lfloor \frac{i}{2} \rfloor$ and $\varepsilon^o \equiv 20\varepsilon\sqrt{\log(p(n_o - 2))}$, as long as $\varepsilon \geq \sigma_{\min}$

Dual Induction. Let our induction hypothesis on n be

$$\sqrt{n}\mu_{[1,n]} \leq \widetilde{\mu}_{1,n} \overline{L}_3 + \widetilde{\mu}_{2,n} \overline{\nu}_q^{1/(q-2)}, \tag{HYP-BE-3}$$

where

$$\widetilde{\mu}_{1,n} = \mathfrak{C}_1 \frac{(\log(ep))^{3/2} \sqrt{\log(pn)}}{\underline{\sigma}^2 \sigma_{\min}} \log(en),$$

$$\widetilde{\mu}_{2,n} = \mathfrak{C}_2 \frac{\log(ep) \sqrt{\log(pn)}}{\underline{\sigma}^{2/(q-2)} \sigma_{\min}} \log(en)$$

for universal constants $\mathfrak{C}_1$ and $\mathfrak{C}_2$ whose values do not change in this subsection. If $\mathfrak{C}_1, \mathfrak{C}_2, \geq 1$, then (HYP-BE-3), requiring $\mu_{[1,n]} \leq 2$ only, trivially holds for $n \leq 16$.

Now we consider the case of $n > 16$. Suppose that the induction hypotheses hold for all intervals with lengths smaller than n. We first derive an anti-concentration inequality for any such intervals. Without loss of generality, we only consider the intervals $[1, i]$ with $i < n$. We claim that

$$\sqrt{i}\kappa_{[1,i]}(\delta) \leq \widetilde{\kappa}_{1,i} \overline{L}_{3,[1,i]} + \widetilde{\kappa}_{2,i} \overline{\nu}_{q,[1,i]}^{1/(q-2)} + \widetilde{\kappa}_{3,i} \overline{\nu}_{2,[1,i]}^{1/2} + \widetilde{\kappa}_4 \delta, \tag{HYP-AC-3}$$

where $\widetilde{\kappa}_{1,i} = \mathfrak{C}_{1,\kappa} \widetilde{\mu}_{1,i}$, $\widetilde{\kappa}_{2,i} = \mathfrak{C}_{2,\kappa} \widetilde{\mu}_{2,i}$, $\widetilde{\kappa}_{3,i} = \mathfrak{C}_{3,\kappa} \frac{\log(ep)\sqrt{\log(pi)}}{\sigma_{\min}}$ and $\widetilde{\kappa}_4 = \mathfrak{C}_{4,\kappa} \frac{\sqrt{\log(ep)}}{\sigma_{\min}}$ for some universal constants $\mathfrak{C}_{1,\kappa}$, $\mathfrak{C}_{2,\kappa}$, $\mathfrak{C}_{3,\kappa}$, and $\mathfrak{C}_{4,\kappa}$ whose values do not change over lines. If $\mathfrak{C}_{1,\kappa}, \mathfrak{C}_{2,\kappa}, \mathfrak{C}_{3,\kappa}, \mathfrak{C}_{4,\kappa}, \geq 1$, then (HYP-BE-3), requiring $\kappa_{[1,n]}(\delta) \leq 1$ for all $\delta > 0$ only, trivially holds for $i \leq 16$. For $i > 16$, assume that (HYP-AC-3) holds for all smaller i's. By Lemma A.6, for any $\varepsilon \geq \sigma_{\min}$ and $\delta > 0$,

$$\kappa_{[1,i]}(\delta) \leq \frac{C}{i_o - 1} \sum_{j=1}^{i_o-1} \frac{\sqrt{\log(ep)}}{\varepsilon} (\nu_{1,j+1} + \nu_{1,j+i_o}) \min\{1, \kappa_{j+(1,i_o)}(\varepsilon^o)\} + \frac{C}{i_o - 1} \sum_{j=1}^{i_o-1} \mu_{j+(1,i_o)} + C\frac{\delta + 2\varepsilon^o}{\sigma_{\min}} \sqrt{\frac{\log(ep)}{i_o - 2}} + C\frac{\overline{\nu}_{1,(1,i)}}{\sigma_{\min}} \frac{\log(ep)}{\sqrt{i_o - 2}},$$

where $i_o = \lfloor \frac{i}{2} \rfloor$, $\varepsilon^o = 20\varepsilon\sqrt{\log(p(i_o-2))}$ and $C > 0$ is an absolute constant. Due to (HYP-AC-3) and (HYP-BE-3) on intervals $j + (1, i_o) \subsetneq [1, i]$,

$$
\begin{aligned}
&\kappa_{[1,i]}(\delta) \\
&\le \frac{C}{(i_o-1)^{3/2}} \sum_{j=1}^{i_o-1} \frac{\sqrt{\log(ep)}}{\varepsilon} (\nu_{1,j+1} + \nu_{1,j+i_o}) \\
&\quad \times \Big[\widetilde{\kappa}_{1,i_o-2} \overline{L}_{3,j+(1,i_o)} + \widetilde{\kappa}_{2,i_o-2} \overline{\nu}_{q,j+(1,i_o)}^{1/(q-2)} + \widetilde{\kappa}_{3,i_o-2} \overline{\nu}_{2,j+(1,i_o)}^{1/2} + \widetilde{\kappa}_4 \delta \Big] \\
&\quad + \frac{C}{(i_o-1)^{3/2}} \sum_{j=1}^{i_o-1} \widetilde{\mu}_{1,i_o-2} \overline{L}_{3,j+(1,i_o)} + \frac{C}{(i_o-1)^{3/2}} \sum_{j=1}^{i_o-1} \widetilde{\mu}_{2,i_o-2} \overline{L}_{4,j+(1,i_o)}^{1/2} \\
&\quad + \frac{C}{(i_o-1)^{3/2}} \sum_{j=1}^{i_o-1} \widetilde{\mu}_{3,i_o-2} \overline{\nu}_{q,j+(1,i_o)}^{1/(q-2)} + C \frac{\delta + 2\varepsilon^o}{\sigma_{\min}} \sqrt{\frac{\log(ep)}{i_o-2}} + C \frac{\overline{\nu}_{1,(1,i)}}{\sigma_{\min}} \frac{\log(ep)}{\sqrt{i_o-2}}.
\end{aligned}
$$

To provide an upper bound in terms of $\overline{L}_3$ and $\overline{\nu}_q$, we use the following lemma based on Jensen's inequality.

LEMMA A.7. *Suppose that $j, k \ge \frac{n}{2}$. For any $q_1, q_2, q_3 > 0$ and $\alpha \le 1$,*

$$
\frac{1}{n} \sum_{j_o=1}^{n} L_{q_1,j_o+j} \overline{L}_{q_2,(j_o,j_o+j)}^{\alpha} \le C \overline{L}_{q_1} \overline{L}_{q_2}^{\alpha}
$$

and

$$
\frac{1}{n} \sum_{j_o=1}^{n} \frac{L_{q_1,j_o+j}}{j-1} \sum_{k_o=1}^{j-1} L_{q_2,j_o+[k_o+k]_{j-1}} \overline{L}_{q_3,j_o+(l,k+l)_{j-1}}^{\alpha} \le C \overline{L}_{q_1} \overline{L}_{q_2} \overline{L}_{q_3}^{\alpha},
$$

where $j_o + (k_o, k_o+k)_{j-1}$ is the shifted interval of $(k_o, k_o+k)_{j-1}$ by j_o in $\mathbb{Z}_n$, namely, $\{j_o + [k_o+1]_{j-1}, \ldots, j_o + [k_o+k-1]_{j-1}\}$. The same inequality holds when some L is replaced with ν.

Based on the lemma,

$$
\frac{1}{i_o-1} \sum_{j=1}^{i_o-1} \overline{L}_{3,j+(1,i_o)} \le C \overline{L}_{3,(2,j-1)},
$$

$$
\frac{1}{i_o-1} \sum_{j=1}^{i_o-1} (\nu_{1,j+1} + \nu_{1,j+i_o}) \overline{L}_{3,j+(1,i_o)} \le C \overline{\nu}_{1,(1,i)} \overline{L}_{3,(2,i-1)},
$$

$$
\frac{1}{i_o-1} \sum_{j=1}^{i_o-1} (\nu_{1,j+1} + \nu_{1,j+i_o}) \overline{\nu}_{2,j+(1,i_o)}^{1/2} \le C \overline{\nu}_{1,(1,i)} \overline{\nu}_{2,(2,i-1)}^{1/2}.
$$

Similar inequalities hold with $\overline{\nu}_q^{1/(q-2)}$ in place of $\overline{L}_3$. As a result, we obtain a similar recursive inequality on $\widetilde{\kappa}$'s with Eq. (18) except that $L_{3,\max}$, $\nu_{q,\max}$, $\nu_{1,\max}$ and $\nu_{2,\max}$ are replaced with $\overline{L}_{q,[1,i]}$, $\overline{\nu}_{q,[1,i]}$, $\overline{\nu}_{1,[1,i]}$ and $\overline{\nu}_{2,[1,i]}$, respectively. Solving the recursive inequality, we prove (HYP-AC-3). The proof of (HYP-BE-3) also proceeds similarly using Lemma A.7. In the end, we obtain a similar recursive inequality for $\widetilde{\mu}$'s with Section 5.3 except that $L_{3,\max}$ and $\nu_{q,\max}$ are replaced with $\overline{L}_3$ and $\overline{\nu}_q$, respectively. Solving the recursive inequality, we prove the desired (HYP-BE-3).

A.4. Proof of Theorem 3.2. We recall from Appendix A.3 that

$$\begin{aligned}&\left|\mathbb{E}[\rho^{\delta}_{r,\phi}(X_{[1,n]})]-\mathbb{E}[\rho^{\delta}_{r,\phi}(Y_{[1,n]})]\right|\\&\le\left|\rho^{\delta}_{r,\phi}(X_{[1,n)})-\rho^{\delta}_{r,\phi}(Y_{[1,n)})\right|+\left|\rho^{\delta}_{r,\phi}(X_{[1,n]})-\rho^{\delta}_{r,\phi}(X_{[1,n)})-\rho^{\delta}_{r,\phi}(Y_{[1,n]})+\rho^{\delta}_{r,\phi}(Y_{[1,n)})\right|\\&\le\left|\mathfrak{R}^{(3)}_X-\mathfrak{R}^{(3)}_Y\right|+\sum_{j=1}^{n-1}\left|\mathbb{E}\left[\mathfrak{R}^{(3,1)}_{X_j}-\mathfrak{R}^{(3,1)}_{Y_j}\right]\right|+\sum_{j=2}^{n-2}\left|\mathbb{E}\left[\mathfrak{R}^{(3,2)}_{X_j}-\mathfrak{R}^{(3,2)}_{Y_j}\right]\right|.\end{aligned}$$

In Section 5.4, we improved the rate by decomposing the third order remainder $\mathfrak{R}^{(3,1)}_{W_j}$ when the fourth moments were finite. Namely,

$$\begin{aligned}&\sum_{j=1}^{n-1}\left|\mathbb{E}\left[\mathfrak{R}^{(3,1)}_{X_j}-\mathfrak{R}^{(3,1)}_{Y_j}\right]\right|\\&\le\frac{C}{\sqrt{n}}\overline{L}_3\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}+\sum_{j=1}^{n-1}\left|\mathbb{E}\left[\mathfrak{R}^{(4,1)}_{X_j}-\mathfrak{R}^{(4,1)}_{Y_j}\right]\right|+\sum_{j=3}^{n-1}\sum_{k=1}^{j-2}\left|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X_j,X_k}-\mathfrak{R}^{(6)}_{X_j,Y_k}\right]\right|.\end{aligned}$$

In 1-ring dependence cases, there exists additional third-order remainder terms $\mathfrak{R}^{(3)}_W$ and $\mathfrak{R}^{(3,2)}_{X_j,W_k}$, which came up while breaking the 1-ring dependence in Appendix A.3. Based on Tyalor expansions up to order 4 and the second moment matching between X_j and Y_j, we decompose the additional remainder terms. First,

$$\begin{aligned}&\mathfrak{R}^{(3)}_X\\&=\frac{1}{6}\left\langle\nabla^3\rho^{\delta}_{r,\phi}(X_{(1,n-1)}),X_n^{\otimes 3}\right\rangle+\left\langle\nabla^3\rho^{\delta}_{r,\phi}(X_{(1,n-3)}),X_{n-2}\otimes X_{n-1}\otimes X_n\right\rangle\\&\quad+\left\langle\nabla^3\rho^{\delta}_{r,\phi}(X_{(2,n-2)}),X_{n-1}\otimes X_n\otimes X_1\right\rangle+\left\langle\nabla^3\rho^{\delta}_{r,\phi}(X_{(3,n-1)}),X_n\otimes X_1\otimes X_2\right\rangle\\&\quad+\frac{1}{2}\left\langle\nabla^3\rho^{\delta}_{r,\phi}(X_{(1,n-2)}),X_{n-1}\otimes X_n\otimes(X_{n-1}+X_n)\right\rangle\\&\quad+\frac{1}{2}\left\langle\nabla^3\rho^{\delta}_{r,\phi}(X_{(2,n-1)}),X_n\otimes X_1\otimes(X_1+X_n)\right\rangle\\&\quad+\mathfrak{R}^{(4)}_X,\end{aligned}\tag{31}$$

where $\mathfrak{R}^{(4)}_X$ is specified in Appendix C. This is the same for $\rho^{\delta}_{r,\phi}(Y_{[1,n]})-\rho^{\delta}_{r,\phi}(Y_{[1,n)})$ but with Y in place of X. To bound the third-order moment terms, we re-apply the Lindeberg swapping and the Taylor expansion up to the sixth order as we did to $\left\langle\mathbb{E}[\nabla^3\rho^{\varepsilon}_{r,\phi}(X_{[1,j-1)}+Y_{(j+1,n)})],\mathbb{E}[X_j^{\otimes 3}]\right\rangle$ in Section 5.4. As a result,

$$\left|\mathfrak{R}^{(3)}_X-\mathfrak{R}^{(3)}_Y\right|\le\frac{C}{\sqrt{n}}\overline{L}_3\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}+\left|\mathfrak{R}^{(4)}_X-\mathfrak{R}^{(4)}_Y\right|+\sum_{k=2}^{n-2}\left|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X,X_k}-\mathfrak{R}^{(6)}_{X,Y_k}\right]\right|,$$

where $\mathfrak{R}^{(6)}_{X,W_k}$ is the sixth-order remainder term specified in Appendix C.

For $\sum_{j=2}^{n-2}\left|\mathbb{E}\left[\mathfrak{R}^{(3,2)}_{X_j}-\mathfrak{R}^{(3,2)}_{Y_j}\right]\right|$, by the Taylor expansion up to the fourth order,

$$\sum_{j=2}^{n-2}\mathbb{E}\left[\mathfrak{R}^{(3,2)}_{X_j}-\mathfrak{R}^{(3,2)}_{Y_j}\right]=\sum_{j=2}^{n-2}\mathbb{E}\left[\mathfrak{R}^{(4,2)}_{X_j}-\mathfrak{R}^{(4,2)}_{Y_j}\right],$$

where $\mathfrak{R}^{(4,2)}_{X_j}$ is the fourth-order remainder, specified in Appendix C.3. Putting all the above results together, we get

$$\begin{aligned}&\left|\mathbb{E}[\rho^{\delta}_{r,\phi}(X_{[1,n]})]-\mathbb{E}[\rho^{\delta}_{r,\phi}(Y_{[1,n]})]\right|\\&\le \frac{C}{\sqrt{n}}\overline{L}_3\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}+\left|\mathfrak{R}^{(4)}_X-\mathfrak{R}^{(4)}_Y\right|+\sum_{j=1}^{n-1}\left|\mathbb{E}\left[\mathfrak{R}^{(4,1)}_{X_j}-\mathfrak{R}^{(4,1)}_{Y_j}\right]\right|+\sum_{j=2}^{n-2}\left|\mathbb{E}\left[\mathfrak{R}^{(4,2)}_{X_j}-\mathfrak{R}^{(4,2)}_{Y_j}\right]\right|\\&\quad+\sum_{k=2}^{n-2}\left|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X,X_k}-\mathfrak{R}^{(6)}_{X,Y_k}\right]\right|+\sum_{j=3}^{n-1}\sum_{k=1}^{j-2}\left|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X_j,X_k}-\mathfrak{R}^{(6)}_{X_j,Y_k}\right]\right|.\end{aligned}$$

Remainder lemma. Similar to Section 5.1, the remainder terms are upper bounded by conditional anti-concentration probability bounds. For $q > 0$, let

$$\widetilde{L}_{q,j} \equiv \sum_{j'=j-4}^{j+4} L_{q,j'} \quad\text{and}\quad \widetilde{L}_{q,[k]_{j-2}} \equiv \sum_{k'=k-4}^{k+4} L_{q,[k']_{j-2}},$$

where $[k']_{j-2}$ is k' modulo $j-2$, and $\widetilde{\nu}_{q,j}$ and $\widetilde{\nu}_{q,[k]_{j-2}}$ are similarly defined.

LEMMA A.8. *Suppose that Assumption (MIN-EV) holds. For W representing either X or Y and $j,k\in\mathbb{Z}_n$ such that $k\le j-2$,*

$$\begin{aligned}\left|\mathbb{E}\left[\mathfrak{R}^{(4)}_W\right]\right| &\le C\phi\left[\widetilde{L}_{4,n}\frac{(\log(ep))^{3/2}}{\delta^3}+\widetilde{\nu}_{q,n}\frac{(\log(ep))^{(q-1)/2}}{\delta^{q-1}}\right]\\&\qquad\times\min\{\kappa_{(3,n-2)}(\delta^o_0)+\kappa^o_n,1\},\end{aligned}$$

$$\begin{aligned}\left|\mathbb{E}\left[\mathfrak{R}^{(4,1)}_{W_j}\right]\right| &\le C\phi\left[\widetilde{L}_{4,j}\frac{(\log(ep))^{3/2}}{\delta^3_{n-j}}+\widetilde{\nu}_{q,j}\frac{(\log(ep))^{(q-1)/2}}{\delta^{q-1}_{n-j}}\right]\\&\qquad\times\min\{\kappa_{[1,j-4)}(\delta^o_{n-j})+\kappa^o_j,1\},\end{aligned}$$

$$\begin{aligned}\left|\mathbb{E}\left[\mathfrak{R}^{(4,2)}_{W_j}\right]\right| &\le C\phi\left[\left(\widetilde{L}_{4,j}+\widetilde{L}_{4,n}\right)\frac{(\log(ep))^{3/2}}{\delta^3_{n-j}}+(\widetilde{\nu}_{q,j}+\widetilde{\nu}_{q,n})\frac{(\log(ep))^{(q-1)/2}}{\delta^{q-1}_{n-j}}\right]\\&\qquad\times\min\{\kappa_{(1,j-2)}(\delta^o_{n-j})+\kappa^o_j,1\},\end{aligned}$$

$$\begin{aligned}\left|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X,W_k}\right]\right| &\le C\phi\widetilde{L}_{3,n}\left[\widetilde{L}_{3,[k]_{n-2}}\frac{(\log(ep))^{5/2}}{\delta^5_{n-k}}+\widetilde{\nu}_{q,[k]_{n-2}}\frac{(\log(ep))^{(q+2)/2}}{\delta^{q+2}_{n-k}}\right]\\&\qquad\times\min\{\kappa_{[1,k-3)}(\delta^o_{n-k})+\kappa^o_k,1\}.\end{aligned}$$

$$\begin{aligned}\left|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X_j,W_k}\right]\right| &\le C\phi\widetilde{L}_{3,j}\left[\widetilde{L}_{3,[k]_{j-2}}\frac{(\log(ep))^{5/2}}{\delta^5_{n-k}}+\widetilde{\nu}_{q,[k]_{j-2}}\frac{(\log(ep))^{(q+2)/2}}{\delta^{q+2}_{n-k}}\right]\\&\qquad\times\min\{\kappa_{[1,k-3)}(\delta^o_{n-k})+\kappa^o_k,1\}.\end{aligned}$$

where $\delta^2_{n-j}\equiv\delta^2+\underline{\sigma}^2\max\{n-j,0\}$, $\delta^o_{n-j}\equiv 12\delta_{n-j}\sqrt{\log(pn)}$ and $\kappa^o_j\equiv\frac{\delta_{n-j}\log(ep)}{\sigma_{\min}\sqrt{\max\{j,1\}}}$, as long as $\delta\ge\sigma_{\min}$ and $\phi\delta\ge\frac{1}{\log(ep)}$.

Permutation argument. Back to Eq. (26), for $3 \le j \le n-1$,

$$\begin{aligned}
&\Big\langle \mathbb{E}[\nabla^3 \rho^\delta_{r,\phi}(X_{[1,j-1)} + Y_{(j+1,n]}) - \nabla^3 \rho^\delta_{r,\phi}(Y_{[1,j-1)} + Y_{(j+1,n]})], \mathbb{E}[X_j^{\otimes 3}] \Big\rangle \\
&\le \frac{1}{j-2} \sum_{k_o=1}^{j-2} \sum_{k=1}^{j-2} \Big\langle \mathbb{E}[\nabla^3 \rho^\delta_{r,\phi}(X_{[k_o,k_o+k)_{j-2}} + X_{[k_o+k]_{j-2}} + Y_{(k_o+k,k_o+j-1)_{j-2}\cup(j,n]}) \\
&\qquad - \nabla^3 \rho^\delta_{r,\phi}(X_{[k_o,k_o+k)_{j-2}} + Y_{[k_o+k]_{j-2}} + Y_{(k_o+k,k_o+j-1)_{j-2}\cup(j,n]})], \mathbb{E}[X_j^{\otimes 3}] \Big\rangle \\
&\le C\phi \frac{\widetilde{L}_{3,j}}{j-2} \sum_{k_o=1}^{j-2} \sum_{k=1}^{j-2} \left[\widetilde{L}_{3,[k_o+k']_{j-2}} \frac{(\log(ep))^{3/2}}{\delta^3_{n-k}} + \widetilde{\nu}_{q,[k_o+k']_{j-2}} \frac{(\log(ep))^{(q-1)/2}}{(\delta_{n-k}/2)^{q-1}} \right] \\
&\qquad\qquad \times \min\{\kappa_{(k_o,k_o+k-3)_{j-2}}(\delta^o_{n-k}) + \kappa^o_k, 1\}.
\end{aligned}$$

The permutation arguement also applies to the first Lindeberg swapping in Eq. (25). Together with the results in Lemma A.8,

$$\begin{aligned}
&\left| \mathbb{E}[\rho^\delta_{r,\phi}(X_{[1,n]})] - \mathbb{E}[\rho^\delta_{r,\phi}(Y_{[1,n]})] \right| \\
&\le \frac{1}{n} \sum_{j_o=1}^{n} \sum_{j=1}^{n-1} \mathbb{E}\left[\rho^\delta_{r,\phi}(X_{(j_o,j_o+j)} + X_{j_o+j} + Y_{(j_o+j,j_o+n]}) - \rho^\delta_{r,\phi}(X_{(j_o,j_o+j)} + Y_{j_o+j} + Y_{(j_o+j,j_o+n]}) \right] \\
&\le \frac{C}{\sqrt{n}} \overline{L}_3 \frac{(\log(ep))^2}{\underline{\sigma}^2 \sigma_{\min}} \\
&\quad + \frac{C\phi}{n} \sum_{j_o=1}^{n} \left[\widetilde{L}_{4,j_o} \frac{(\log(ep))^{3/2}}{\delta^3} + \widetilde{\nu}_{q,j_o} \frac{(\log(ep))^{(q-1)/2}}{(\delta/2)^{q-1}} \right] \min\{1, \kappa_{j_o+(3,n-2)}(\delta^o_0) + \kappa^o_n\} \\
&\quad + \frac{C\phi}{n} \sum_{j_o=1}^{n} \sum_{j=2}^{n-2} \left[\widetilde{L}_{4,j_o} \frac{(\log(ep))^{3/2}}{\delta^3_{n-j}} + \widetilde{\nu}_{q,j_o} \frac{(\log(ep))^{(q-1)/2}}{(\delta_{n-j}/2)^{q-1}} \right] \min\{1, \kappa_{j_o+(1,j-2)}(\delta^o_{n-j}) + \kappa^o_j\} \\
&\quad + \frac{C\phi}{n} \sum_{j_o=1}^{n} \sum_{j=1}^{n-1} \left[\widetilde{L}_{4,j_o+j} \frac{(\log(ep))^{3/2}}{\delta^3_{n-j}} + \widetilde{\nu}_{q,j_o+j} \frac{(\log(ep))^{(q-1)/2}}{(\delta_{n-j}/2)^{q-1}} \right] \min\{1, \kappa_{j_o+(1,j-4)}(\delta^o_{n-j}) + \kappa^o_j\} \\
&\quad + \frac{C\phi}{n} \sum_{j_o=1}^{n} \sum_{j=3}^{n} \frac{\widetilde{L}_{3,j_o+j}}{j-2} \sum_{k_o=1}^{j-2} \sum_{k=1}^{j-2} \left[\widetilde{L}_{3,j_o+[k_o+k]_{j-2}} \frac{(\log(ep))^{5/2}}{\delta^5_{n-k}} + \widetilde{\nu}_{q,j_o+[k_o+k]_{j-2}} \frac{(\log(ep))^{(q+2)/2}}{(\delta_{n-k}/2)^{q+2}} \right] \\
&\qquad\qquad \times \min\{1, \kappa_{j_o+(k_o,k_o+k-3)_{j-2}}(\delta^o_{n-k}) + \kappa^o_k\},
\end{aligned}$$

where $j_o + (k_o, k_o+k)_{j-1}$ is the shifted interval of $(k_o, k_o+k)_{j-1}$ by j_o in $\mathbb{Z}_n$, namely, $\{j_o + [k_o+1]_{j-1}, \ldots, j_o + [k_o+k-1]_{j-1}\}$.

Partitioning the sum. Again, we partition the summations at $J_n = n(1 - \frac{\sigma^2_{\min}}{\underline{\sigma}^2 \log^2(4ep)})$. The calculations of the first three summations are similar to those in Appendix A.3. Here we only take a look at the last summation where the summation iterates. For $k < J_n$,

$$\frac{1}{n}\sum_{j_o=1}^{n}\sum_{j=3}^{n}\frac{\widetilde{L}_{3,j_o+j}}{j-2}\sum_{k_o=1}^{j-2}\sum_{k=1}^{(j-2)\wedge\lfloor J_n\rfloor}\left[\widetilde{L}_{3,j_o+[k_o+k]_{j-2}}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5}+\widetilde{\nu}_{q,j_o+[k_o+k]_{j-2}}\frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}}\right]$$

$$\leq\frac{C}{n}\sum_{j_o=1}^{n}\sum_{j=3}^{n}\widetilde{L}_{3,j_o+j}\sum_{k=1}^{(j-2)\wedge\lfloor J_n\rfloor}\left[\overline{L}_3\frac{n(\log(ep))^{5/2}}{(j-2)\delta_{n-k}^5}+\overline{\nu}_q\frac{n(\log(ep))^{(q+2)/2}}{(j-2)\delta_{n-k}^{q+2}}\right]$$

$$\leq\frac{C}{\sqrt{n}}\sum_{j=3}^{n}\overline{L}_3\left[\overline{L}_3\frac{n(\log(ep))^{7/2}}{(j-2)\underline{\sigma}^2\sigma_{\min}\delta_{n-\lfloor J_n\rfloor}^2}+\overline{\nu}_q\frac{n(\log(ep))^{(q+4)/2}}{(j-2)\underline{\sigma}^2\sigma_{\min}\delta_{n-\lfloor J_n\rfloor}^{q-1}}\right]$$

$$\leq\frac{C}{\sqrt{n}}\overline{L}_3\left[\overline{L}_3\frac{(\log(ep))^{7/2}}{\underline{\sigma}^4\sigma_{\min}}+\overline{\nu}_q\frac{(\log(ep))^{(q+4)/2}}{\underline{\sigma}^4\sigma_{\min}\delta^{q-3}}\right]\log(en),$$

where the third inequality comes from Eq. (12) and the last inequality comes from the fact that $\sum_{j=3}^{n}\frac{n}{(j-2)\delta_{n-\lfloor J_n\rfloor}^2}\leq\sum_{j=3}^{n}\frac{n}{(j-2)(n-\lfloor J_n\rfloor)\underline{\sigma}^2}\leq\frac{C}{\underline{\sigma}^2}\log(en)$. For $k > J_n$,

$$\frac{1}{n}\sum_{j_o=1}^{n}\sum_{j=\lceil J_n\rceil}^{n}\frac{\widetilde{L}_{3,j_o+j}}{j-2}\sum_{k_o=1}^{j-2}\sum_{k=\lceil J_n\rceil}^{j-2}\left[\widetilde{L}_{3,j_o+[k_o+k]_{j-2}}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5}+\widetilde{\nu}_{q,j_o+[k_o+k]_{j-2}}\frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}}\right]$$
$$\times\min\{1,\kappa_{j_o+(k_o,k_o+k)_{j-2}}(\delta_{n-k}^o)+\kappa_k^o\}$$

$$\leq\frac{1}{n}\sum_{i=1}^{n}\sum_{j=\lceil J_n\rceil}^{n}\frac{\widetilde{L}_{3,j_o+j}}{j-2}\sum_{k_o=1}^{j-2}\sum_{k=\lceil J_n\rceil}^{j-2}\left[\widetilde{L}_{3,j_o+[k_o+k]_{j-2}}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5}+\widetilde{\nu}_{q,j_o+[k_o+k]_{j-2}}\frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}}\right]$$
$$\times\kappa_{j_o+(k_o,k_o+k)_{j-2}}(\delta_{n-k}^o)$$
$$+\frac{C}{n}\sum_{j_o=1}^{n}\sum_{j=\lceil J_n\rceil}^{n}\widetilde{L}_{3,j_o+j}\sum_{k=\lceil J_n\rceil}^{j-2}\left[\overline{L}_{3,j_o+(0,j)}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5}+\overline{\nu}_{q,j_o+(0,j)}\frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}}\right]\kappa_k^o.$$

Based on Lemma A.7,

$$\frac{1}{n}\sum_{j_o=1}^{n}\sum_{j=\lceil J_n\rceil}^{n}\widetilde{L}_{3,j_o+j}\sum_{k=\lceil J_n\rceil}^{j-1}\left[\overline{L}_{3,(j_o,j_o+j)}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5}+\overline{\nu}_{q,(j_o,j_o+j)}\frac{(\log(ep))^{(q+2)/2}}{\delta_{n-k}^{q+2}}\right]\kappa_k^o$$

$$\leq\frac{C}{n^{3/2}}\sum_{j_o=1}^{n}\sum_{j=\lceil J_n\rceil}^{n}\widetilde{L}_{3,j_o+j}\left[\overline{L}_{3,(j_o,j_o+j)}\frac{(\log(ep))^{7/2}}{\underline{\sigma}^2\sigma_{\min}\delta_{n-j}^2}+\overline{\nu}_{q,(j_o,j_o+j)}\frac{(\log(ep))^{(q+4)/2}}{\underline{\sigma}^2\sigma_{\min}\delta_{n-j}^{q-1}}\right]$$

$$\leq\frac{C}{\sqrt{n}}\sum_{j=\lceil J_n\rceil}^{n}\overline{L}_3\left[\overline{L}_3\frac{(\log(ep))^{7/2}}{\underline{\sigma}^2\sigma_{\min}\delta_{n-j}^2}+\overline{\nu}_q\frac{(\log(ep))^{(q+4)/2}}{\underline{\sigma}^2\sigma_{\min}\delta_{n-j}^{q-1}}\right]$$

$$\leq\frac{C}{\sqrt{n}}\overline{L}_3\left[\overline{L}_3\frac{(\log(ep))^{7/2}}{\underline{\sigma}^4\sigma_{\min}}+\overline{\nu}_q\frac{(\log(ep))^{(q+4)/2}}{\underline{\sigma}^4\sigma_{\min}\delta^{q-3}}\right]\log\left(1+\frac{\sqrt{n}\underline{\sigma}}{\delta}\right),$$

where the first and last inequliaties follow Eqs. (12) and (19), respectively. In sum, we obtain the following induction lemma.

LEMMA A.9. *If Assumptions* (MIN-VAR)*,* (MIN-EV) *and* (VAR-EV) *hold, then for any* $\delta \geq \sigma_{\min}$,

$$\begin{aligned}
&\mu_{[1,n]} \\
&\leq \frac{C}{\sqrt{n}}\frac{\delta \log(ep)}{\sigma_{\min}} + \frac{C}{\sqrt{n}}\frac{\sqrt{\log(ep)}}{\phi\sigma_{\min}} + \frac{C}{\sqrt{n}}\overline{L}_3\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}} \\
&\quad + \frac{C\phi}{\sqrt{n}}\left[\overline{L}_4\frac{(\log(ep))^{5/2}}{\underline{\sigma}^2\sigma_{\min}} + \overline{\nu}_q\frac{(\log(ep))^{(q+1)/2}}{\underline{\sigma}^2\sigma_{\min}(\delta/2)^{q-4}}\right]\log(en) \\
&\quad + \frac{C\phi}{\sqrt{n}}\overline{L}_3\left[\overline{L}_3\frac{(\log(ep))^{7/2}}{\underline{\sigma}^4\sigma_{\min}} + \overline{\nu}_q\frac{(\log(ep))^{(q+4)/2}}{\underline{\sigma}^4\sigma_{\min}(\delta/2)^{q-3}}\right]\log(en) \\
&\quad + \frac{C\phi}{n}\sum_{j_o=1}^{n}\sum_{j=\lceil J_n\rceil}^{n}\left[\widetilde{L}_{4,j_o}\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} + \widetilde{\nu}_{q,j_o}\frac{(\log(ep))^{(q-1)/2}}{(\delta_{n-j}/2)^{q-1}}\right]\kappa_{j_o+(3,j-2)}(\delta_{n-j}^o) \\
&\quad + \frac{C\phi}{n}\sum_{j_o=1}^{n}\sum_{j=\lceil J_n\rceil}^{n-1}\left[\widetilde{L}_{4,j_o+j}\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} + \widetilde{\nu}_{q,j_o+j}\frac{(\log(ep))^{(q-1)/2}}{(\delta_{n-j}/2)^{q-1}}\right]\kappa_{j_o+(1,j-4)}(\delta_{n-j}^o) \\
&\quad + \frac{C\phi}{n}\sum_{j_o=1}^{n}\sum_{j=\lceil J_n\rceil}^{n}\frac{\widetilde{L}_{3,j_o+j}}{j-2} \\
&\qquad \times \sum_{k_o=1}^{j-2}\sum_{k=\lceil J_n\rceil}^{j-2}\left[\widetilde{L}_{3,j_o+[k_ok]_{j-2}}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5} + \widetilde{\nu}_{q,j_o+[k_o+k]_{j-2}}\frac{(\log(ep))^{(q+2)/2}}{(\delta_{n-k}/2)^{q+2}}\right] \\
&\qquad\qquad \times \kappa_{j_o+(k_o,k_o+k-3)_{j-2}}(\delta_{n-k}^o),
\end{aligned}$$

for some absolute constant $C > 0$.

Dual Induction. Based on Lemmas A.6 and A.9, we proceed the dual induction in the same steps as Appendix A.1, but using Lemma A.7 as in Appendix A.3. As a result, we prove for any n,

$$\sqrt{n}\mu_{[1,n]} \leq \widetilde{\mu}_{1,n}\overline{L}_3 + \widetilde{\mu}_{2,n}\overline{L}_4^{1/2} + \widetilde{\mu}_{3,n}\overline{\nu}_q^{1/(q-2)}, \qquad \text{(HYP-BE-4)}$$

where

$$\begin{aligned}
\widetilde{\mu}_{1,n} &= \mathfrak{C}_1\frac{(\log(ep))^{3/2}\sqrt{\log(pn)}}{\underline{\sigma}^2\sigma_{\min}}\log(en), \\
\widetilde{\mu}_{2,n} &= \mathfrak{C}_2\frac{\log(ep)\sqrt{\log(pn)}}{\underline{\sigma}\sigma_{\min}}\log(en) \\
\widetilde{\mu}_{3,n} &= \mathfrak{C}_3\frac{\log(ep)\sqrt{\log(pn)}}{\underline{\sigma}^{2/(q-2)}\sigma_{\min}}\log(en)
\end{aligned}$$

for universal constants $\mathfrak{C}_1$, $\mathfrak{C}_2$ and $\mathfrak{C}_3$ whose values do not change over lines, and for any $i < n$,

$$\sqrt{i}\kappa_{[1,i]}(\delta) \leq \widetilde{\kappa}_{1,i}\overline{L}_{3,[1,i]} + \widetilde{\kappa}_{2,i}\overline{L}_{4,[1,i]}^{1/2} + \widetilde{\kappa}_{3,i}\overline{\nu}_{q,[1,i]}^{1/(q-2)} + \widetilde{\kappa}_{4,i}\overline{\nu}_{2,[1,i]}^{1/2} + \widetilde{\kappa}_5\delta, \qquad \text{(HYP-AC-4)}$$

where $\widetilde{\kappa}_{1,i} = \mathfrak{C}_{1,\kappa}\widetilde{\mu}_{1,i}$, $\widetilde{\kappa}_{2,i} = \mathfrak{C}_{2,\kappa}\widetilde{\mu}_{2,i}$, $\widetilde{\kappa}_{3,i} = \mathfrak{C}_{3,\kappa}\widetilde{\mu}_{3,i}$, $\widetilde{\kappa}_{4,i} = \mathfrak{C}_{4,\kappa}\frac{\log(ep)\sqrt{\log(pi)}}{\sigma_{\min}}$ and $\widetilde{\kappa}_5 = \mathfrak{C}_{5,\kappa}\frac{\sqrt{\log(ep)}}{\sigma_{\min}}$ for some universal constants $\mathfrak{C}_{1,\kappa}$, $\mathfrak{C}_{2,\kappa}$, $\mathfrak{C}_{3,\kappa}$, $\mathfrak{C}_{4,\kappa}$ and $\mathfrak{C}_{5,\kappa}$ whose values do not change over lines. This proves the desired Berry–Esseen bound.

APPENDIX B: PROOFS OF LEMMAS

B.1. Proof of Lemma 5.2. The smoothing lemma is the result of the serial application of the following two lemmas. Lemma B.2 is a corollary of Theorem 2.1 in Chernozhukov, Chetverikov and Koike (2020). We provide a standalone proof in Appendix B.2.

LEMMA B.1 (Lemma 1, Kuchibhotla and Rinaldo, 2020). *Suppose that X is a p-dimensional random vector, and $Y \sim N(0,\Sigma)$ is a p-dimensional Gaussian random vector. Then, for any $\delta > 0$ and a standard Gaussian random vector Z,*

$$\mu(X,Y) \le C\mu(X+\delta Z, Y+\delta Z) + C\frac{\delta\log(ep)}{\sqrt{\min_{i=1,\dots,p}\Sigma_{ii}}}.$$

LEMMA B.2 (Chernozhukov, Chetverikov and Koike, 2020). *Suppose that X is a p-dimensional random vector, and $Y \sim N(0,\Sigma)$ is a p-dimensional Gaussian random vector. Then, for any $\phi > 0$,*

$$\mu(X,Y) \le \sup_{r\in\mathbb{R}^d}|\mathbb{E}[f_{r,\phi}(X)] - \mathbb{E}[f_{r,\phi}(Y)]| + \frac{C}{\phi}\sqrt{\frac{\log(ep)}{\min_{i=1,\dots,p}\Sigma_{ii}}}.$$

For any $\delta > 0$,

$$\mu(X,Y) \le \mu(X+\delta Z, Y+\delta Z) + C\frac{\delta\log(ep)}{\sqrt{\min_{i=1,\dots,p}\Sigma_{ii}}}. \tag{32}$$

Then, for any $\delta > 0$ and $\phi > 0$,

$$\begin{aligned}
&\mu(X+\delta Z, Y+\delta Z)\\
&\le C\sup_{r\in\mathbb{R}^d}|\mathbb{E}[f_{r,\phi}(X+\delta Z)] - \mathbb{E}[f_{r,\phi}(Y+\delta Z)]| + \frac{C}{\phi}\sqrt{\frac{\log(ep)}{\min_{i=1,\dots,p}\Sigma_{ii}}}\\
&\le C\sup_{r\in\mathbb{R}^d}\left|\mathbb{E}[\rho^{\delta}_{r,\phi}(X)] - \mathbb{E}[\rho^{\delta}_{r,\phi}(Y)]\right| + \frac{C}{\phi}\sqrt{\frac{\log(ep)}{\min_{i=1,\dots,p}\Sigma_{ii}}}.
\end{aligned}$$

In sum,

$$\mu(X,Y) \le C\sup_{r\in\mathbb{R}^d}\left|\mathbb{E}[\rho^{\delta}_{r,\phi}(X)] - \mathbb{E}[\rho^{\delta}_{r,\phi}(Y)]\right| + C\frac{\delta\log(ep) + \sqrt{\log(ep)}/\phi}{\sqrt{\min_{i=1,\dots,p}\Sigma_{ii}}}.$$

B.2. Proof of Lemma B.2. By Lemma B.1,

$$\begin{aligned}
\mathbb{P}[X\in A_r] &\le \mathbb{E}[f_{r,\phi}(X)] = \mathbb{E}[f_{r,\phi}(Y)] + \mathbb{E}[f_{r,\phi}(X)] - \mathbb{E}[f_{r,\phi}(Y)]\\
&\le \mathbb{P}[Y\in A_{r+\frac{1}{\phi}\mathbf{1}}] + \mathbb{E}[f_{r,\phi}(X)] - \mathbb{E}[f_{r,\phi}(Y)]\\
&\le \mathbb{P}[Y\in A_r] + \frac{C}{\phi}\sqrt{\frac{\log(ep)}{\min_{i=1,\dots,p}\Sigma_{ii}}},
\end{aligned}$$

and similarly,

$$\begin{aligned}\mathbb{P}[Y\in A_r] &\leq \mathbb{P}[Y\in A_{r-\frac{1}{\phi}\mathbf{1}}]+\frac{C}{\phi}\sqrt{\frac{\log(ep)}{\min_{i=1,\dots,p}\Sigma_{ii}}}\\ &\leq \mathbb{E}[f_{r-\frac{1}{\phi}\mathbf{1},\phi}(Y)]+\frac{C}{\phi}\sqrt{\frac{\log(ep)}{\min_{i=1,\dots,p}\Sigma_{ii}}}\\ &\leq \mathbb{E}[f_{r-\frac{1}{\phi}\mathbf{1},\phi}(X)]+\mathbb{E}[f_{r-\frac{1}{\phi}\mathbf{1},\phi}(Y)]-\mathbb{E}[f_{r-\frac{1}{\phi}\mathbf{1},\phi}(X)]+\frac{C}{\phi}\sqrt{\frac{\log(ep)}{\min_{i=1,\dots,p}\Sigma_{ii}}}\\ &\leq \mathbb{P}[X\in A_r]+\mathbb{E}[f_{r-\frac{1}{\phi}\mathbf{1},\phi}(Y)]-\mathbb{E}[f_{r-\frac{1}{\phi}\mathbf{1},\phi}(X)]+\frac{C}{\phi}\sqrt{\frac{\log(ep)}{\min_{i=1,\dots,p}\Sigma_{ii}}}.\end{aligned}$$

Hence,

$$\sup_{r\in\mathbb{R}^p}|\mathbb{P}[X\in A_r]-\mathbb{P}[Y\in A_r]|\leq \sup_{r\in\mathbb{R}^p}|\mathbb{E}[f_{r,\phi}(X)]-\mathbb{E}[f_{r,\phi}(Y)]|+\frac{C}{\phi}\sqrt{\frac{\log(ep)}{\min_{i=1,\dots,p}\Sigma_{ii}}}.$$

B.3. Proof of Remainder Lemmas. We observe that all remainder terms are in forms of

$$\frac{1}{(\beta-1)!}\int_0^1(1-t)^{\beta-1}\mathbb{E}\left[\left\langle\nabla^\alpha\rho^\delta_{r,\phi}(X_{J_1}+W^{\not\perp}+Y_{J_2}+tW_J),(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right\rangle\right]dt,$$

where $1\leq\beta\leq\alpha$ and J_1,J_2 and J are subsets of $\mathbb{Z}_n$ satisfying $\mathscr{X}_{J_1}$, $\mathscr{Y}_{J_2}$ and $\mathscr{W}_{\{j_1,\dots,j_{\alpha-\beta}\}\cup J}$ are mutually independent. We also note that $\mathscr{X}_{J_1}\cup\mathscr{Y}_{J_2}\cup\{W^{\not\perp}\}$ is independent from $\mathscr{W}_{\{j_1,\dots,j_{\alpha-\beta}\}}$. Here we prove a succinct form of the remainder lemmas:

LEMMA B.3. *Suppose that* $\mathrm{Var}[Y_{J_2}|\mathscr{Y}_{J_2^\complement}]\succeq\underline{\sigma}^2_{J_2}I_p$ *for some* $\underline{\sigma}^2_{J_2}>0$ *and* p*-dimensional identity matrix* I_p*. Then, for any* $\gamma_1,\gamma_2\in[0,1]$ *and* $\eta>0$*, remainder terms in the above form satisfy*

$$\begin{aligned}&\left|\int_0^1(1-t)^{\beta-1}\mathbb{E}\left[\left\langle\nabla^\alpha\rho^\delta_{r,\phi}(X_{J_1}+W^{\not\perp}+Y_{J_2}+tW_J),(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right\rangle\right]dt\right|\\ &\leq C\left[\frac{(\log(ep))^{\alpha/2}}{\delta'^\alpha}\min\{1,\kappa_{J_1}(\delta^o)+\kappa^o\}\right]\\ &\quad\times\left[\begin{aligned}&\frac{\phi^{\gamma_1}\delta'^{\gamma_1}}{(\log(ep))^{\gamma_1/2}}\left\|\mathbb{E}\left[\left|(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right|\right]\right\|_\infty\\ &+\frac{\phi^{\gamma_2}\delta'^{\gamma_2-\eta}}{(\log(ep))^{(\gamma_2-\eta)/2}}\mathbb{E}\left[\prod_{k=1}^{\alpha-\beta}\|W_{j_k}\|_\infty\|W_J\|_\infty^\beta(\|W_J\|_\infty^\eta+\|W^{\not\perp}\|_\infty^\eta)\right]\end{aligned}\right],\end{aligned}$$

where $\left|(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right|$ *is the element-wise absolute operation,* $\delta'\equiv\sqrt{\delta^2+\underline{\sigma}^2_{J_2}}$*,* $\delta^o\equiv 12\delta'\sqrt{\log(p|J_1|)}$*, and* $\kappa^o\equiv\frac{\delta'\log(ep)}{\sigma_{\min}\sqrt{|J_1|}}$ *as long as* $\delta\geq\sigma_{\min}$ *and* $\phi\delta\geq\frac{1}{\log(ep)}$*.*

We prove the above generalized remainder lemma here. Because $\mathrm{Var}[Y_{J_2}|\mathscr{Y}_{J_2^\complement}]\succeq\underline{\sigma}^2_{J_2}I_p$,

$$Y_{J_2}|\mathscr{Y}_{J_2^\complement}\overset{d}{=}Y^o_{J_2}+\underline{\sigma}_{J_2}\cdot Z,\ \text{ almost surely,}$$

where $Y^o_{J_2}$ is the Gaussian random variable with mean $\mathbb{E}[Y_{J_2}|\mathscr{Y}_{J_2^c}]$ and variance $\mathrm{Var}[Y_{J_2}|\mathscr{Y}_{J_2^c}] - \underline{\sigma}^2_{J_2} I_p$. For brevity, let $W^o \equiv X_{J_1} + Y^o_{J_2}$. Because $\delta' = \sqrt{\delta^2 + \underline{\sigma}^2_{J_2}}$, the remainder term is decomposed into

$$\begin{aligned}
&\int_0^1 (1-t)^{\beta-1}\mathbb{E}\left[\left\langle \nabla^\alpha \rho^\delta_{r,\phi}(X_{J_1} + W^{\not J} + Y_{J_2} + tW_J), (\otimes_{k=1}^{\alpha-\beta} W_{j_k}) \otimes W_J^{\otimes\beta}\right\rangle\right] dt \\
&= \int_0^1 (1-t)^{\beta-1}\mathbb{E}\left[\left\langle \nabla^\alpha \rho^{\delta'}_{r,\phi}(W^o + W^{\not J} + tW_J), (\otimes_{k=1}^{\alpha-\beta} W_{j_k}) \otimes W_J^{\otimes\beta}\right\rangle\right] dt \\
&= \int_0^1 (1-t)^{\beta-1}\mathbb{E}\left[\left\langle \nabla^\alpha \rho^{\delta'}_{r,\phi}(W^o + W^{\not J} + tW_J), (\otimes_{k=1}^{\alpha-\beta} W_{j_k}) \otimes W_J^{\otimes\beta}\right\rangle \mathbb{I}_{W_J,1}\right] dt \\
&\quad + \int_0^1 (1-t)^{\beta-1}\mathbb{E}\left[\left\langle \nabla^{\alpha-1} \rho^{\delta'}_{r,\phi}(W^o + W^{\not J} + tW_J), (\otimes_{k=1}^{\alpha-\beta} W_{j_k}) \otimes W_J^{\otimes\beta}\right\rangle \mathbb{I}_{W_J,2}\right] dt \\
&=: \int_0^1 (1-t)^{\beta-1} T_{W_J,1}(t)dt + \int_0^1 (1-t)^{\beta-1} T_{W_J,2}(t)dt,
\end{aligned}$$

where

$$\mathbb{I}_{W_J,1} = \mathbb{I}_{\{\|W_J\|_\infty \text{ and } \|W^{\not J}\|_\infty < \frac{\delta'}{\sqrt{\log(ep)}}\}},$$

$$\mathbb{I}_{W_J,2} = \mathbb{I}_{\{\|W_J\|_\infty \text{ or } \|W^{\not J}\|_\infty \geq \frac{\delta'}{\sqrt{\log(ep)}}\}}.$$

We upper-bound the terms $T_{W_J,1}(t)$ and $T_{W_J,2}(t)$ separately. First,

$$\begin{aligned}
&|T_{W_J,1}(t)| \\
&= \left|\mathbb{E}\left[\left\langle \nabla^\alpha \rho^{\delta'}_{r,\phi}(W^o + W^{\not J} + tW_J), (\otimes_{k=1}^{\alpha-\beta} W_{j_k}) \otimes W_J^{\otimes\beta}\right\rangle \mathbb{I}_{W_J,1}\right]\right| \\
&= \left|\mathbb{E}\left[\sum_{i_1,\ldots,i_\alpha} \nabla^{(i_1,\ldots,i_\alpha)} \rho^{\delta'}_{r,\phi}(W^o + W^{\not J} + tW_J) \cdot \prod_{k\leq\alpha-\beta} W_{j_k}^{(i_k)} \prod_{k>\alpha-\beta} W_J^{(i_k)} \cdot \mathbb{I}_{W_J,1}\right]\right| \\
&\leq \mathbb{E}\left[\sum_{i_1,\ldots,i_\alpha} \sup_{z\in\mathcal{B}}\left|\nabla^{(i_1,\ldots,i_\alpha)} \rho^{\delta'}_{r,\phi}(W^o + z)\right| \cdot \left|\prod_{k\leq\alpha-\beta} W_{j_k}^{(i_k)} \prod_{k>\alpha-\beta} W_J^{(i_k)}\right| \cdot \mathbb{I}_{W_J,1}\right] \\
&\leq \sum_{i_1,\ldots,i_\alpha} \mathbb{E}\left[\sup_{z\in\mathcal{B}}\left|\nabla^{(i_1,\ldots,i_\alpha)} \rho^{\delta'}_{r,\phi}(W^o + z)\right| \cdot \left|\prod_{k\leq\alpha-\beta} W_{j_k}^{(i_k)} \prod_{k>\alpha-\beta} W_J^{(i_k)}\right| \cdot \mathbb{I}_{W_J,1}\right] \\
&= \sum_{i_1,\ldots,i_\alpha} \mathbb{E}\left[\sup_{z\in\mathcal{B}}\left|\nabla^{(i_1,\ldots,i_\alpha)} \rho^{\delta'}_{r,\phi}(W^o + z)\right|\right] \mathbb{E}\left[\left|\prod_{k\leq\alpha-\beta} W_{j_k}^{(i_k)} \prod_{k>\alpha-\beta} W_J^{(i_k)}\right| \cdot \mathbb{I}_{W_J,1}\right] \\
&\leq \mathbb{E}\left[\sum_{i_1,\ldots,i_\alpha} \sup_{z\in\mathcal{B}}\left|\nabla^{(i_1,\ldots,i_\alpha)} \rho^{\delta'}_{r,\phi}(W^o_j + z)\right|\right] \left\|\mathbb{E}\left[\left|(\otimes_{k=1}^{\alpha-\beta} W_{j_k}) \otimes W_J^{\otimes\beta}\right|\right]\right\|_\infty,
\end{aligned} \tag{33}$$

where $\mathcal{B} = \{z : \|z\|_\infty \leq \frac{2\delta'}{\sqrt{\log(ep)}}\}$, and $\left|(\otimes_{k=1}^{\alpha-\beta} W_{j_k}) \otimes W_J^{\otimes\beta}\right|$ is the element-wise absolute operation. The fifth equality follows the independence between W^o and $\mathscr{W}_{\{j_1,\ldots,j_{\alpha-\beta}\}\cup J}$. We decompose the expectation term on the last line by

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{i_1,\ldots,i_\alpha}\sup_{z\in\mathcal{B}}\left|\nabla^{(i_1,\ldots,i_\alpha)}\rho^{\delta'}_{r,\phi}(W^o_j+z)\right|\right] \\
&\le \mathbb{E}\left[\sum_{i_1,\ldots,i_\alpha}\sup_{z\in\mathcal{B}}\left|\nabla^{(i_1,\ldots,i_\alpha)}\rho^{\delta'}_{r,\phi}(W^o_j+z)\right|\cdot\mathbb{I}_{W^o,1}\right] \\
&\quad+\mathbb{E}\left[\sum_{i_1,\ldots,i_\alpha}\sup_{z\in\mathcal{B}}\left|\nabla^{(i_1,\ldots,i_\alpha)}\rho^{\delta'}_{r,\phi}(W^o_j+z)\right|\cdot\mathbb{I}_{W^o,2}\right],
\end{aligned}
\tag{34}
$$

where the value of $h>0$ will be determined later, and

$$
\mathbb{I}_{W^o,1}=\mathbb{I}_{\left\{\|W^o-\partial A_r\|_\infty\le 12\delta'\sqrt{\log(ph)}\right\}},
$$

$$
\mathbb{I}_{W^o,2}=\mathbb{I}_{\left\{\|W^o-\partial A_r\|_\infty> 12\delta'\sqrt{\log(ph)}\right\}}.
$$

For the first term, based on Lemmas 6.1 and 6.2 of CCK20,

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{i_1,\ldots,i_\alpha}\sup_{z\in\mathcal{B}}\left|\nabla^{(i_1,\ldots,i_\alpha)}\rho^{\delta'}_{r,\phi}(W^o+z)\right|\cdot\mathbb{I}_{W^o,1}\right] \\
&\le \sup_{w\in\mathbb{R}^p}\sum_{i_1,\ldots,i_\alpha}\sup_{z\in\mathcal{B}}\left|\nabla^{(i_1,\ldots,i_\alpha)}\rho^{\delta'}_{r,\phi}(w+z)\right|\cdot\mathbb{E}\left[\mathbb{I}_{W^o,1}\right] \\
&\le C\frac{\phi^{\gamma_1}(\log(ep))^{(\alpha-\gamma_1)/2}}{\delta'^{\alpha-\gamma_1}}\mathbb{P}\left[\|X_{J_1}-\partial A_{r-Y^o_{J_2}}\|_\infty\le 12\delta'\sqrt{\log(ph)}\right] \\
&\le C\frac{\phi^{\gamma_1}(\log(ep))^{(\alpha-\gamma_1)/2}}{\delta'^{\alpha-\gamma_1}}\min\left\{1,\kappa_{J_1}\left(12\delta'\sqrt{\log(ph)}\right)\right\},
\end{aligned}
$$

for any $\gamma_1\in[0,1]$. For the last term, based on Lemma 10.5 of Lopes (2022),

$$
\mathbb{E}\left[\sum_{i_1,\ldots,i_\alpha}\sup_{z\in\mathcal{B}}\left|\nabla^{(i_1,\ldots,i_\alpha)}\rho^{\delta'}_{r,\phi}(W^o+z)\right|\cdot\mathbb{I}_{W^o,2}\right]\le C\frac{1}{\delta'^\alpha h}.
$$

Plugging the last two results in Eqs. (33) and (34), the resulting upper bound for $T_{W_J,1}(t)$ is

$$
\begin{aligned}
|T_{W_J,1}(t)|\le C&\left[\frac{\phi^{\gamma_1}(\log(ep))^{(\alpha-\gamma_1)/2}}{\delta'^{\alpha-\gamma_1}}\min\left\{1,\kappa_{J_1}\left(12\delta'\sqrt{\log(ph)}\right)\right\}+\frac{1}{\delta'^\alpha h}\right] \\
&\times\left\|\mathbb{E}\left[\left|(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right|\right]\right\|_\infty,
\end{aligned}
$$

for any $t\in[0,1]$, $\gamma_1\in[0,1]$ and $h>0$. Replacing the minimum with 1 and minimizing over $h>0$, we get

$$
|T_{W_J,1}(t)|\le C\frac{\phi^{\gamma_1}(\log(ep))^{(\alpha-\gamma_1)/2}}{\delta'^{\alpha-\gamma_1}}\left\|\mathbb{E}\left[\left|(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right|\right]\right\|_\infty.
$$

Moreover, plugging-in

$$
h=\frac{\sigma_{\min}}{\phi^{\gamma_1}\delta'^{1+\gamma_1}}\sqrt{\frac{|J_1|}{(\log(ep))^{\alpha-\gamma_1+2}}},
$$

we obtain

$$
|T_{W_J,1}(t)|\le C\frac{\phi^{\gamma_1}(\log(ep))^{(\alpha-\gamma_1)/2}}{\delta'^{\alpha-\gamma_1}}\left\|\mathbb{E}\left[\left|(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right|\right]\right\|_\infty(\kappa_{J_1}(\delta^o)+\kappa^o),
$$

where $\delta^o = 12\delta'\sqrt{\log(p|J_1|)}$ and $\kappa^o = \frac{\delta' \log(ep)}{\sigma_{\min}\sqrt{|J_1|}}$. In sum,

$$\begin{aligned}&|T_{W_J,1}(t)| \\ &\leq C\frac{\phi^{\gamma_1}(\log(ep))^{(\alpha-\gamma_1)/2}}{\delta'^{\alpha-\gamma_1}}\left\|\mathbb{E}\left[\left|(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right|\right]\right\|_\infty \min\{1,\kappa_{J_1}(\delta^o)+\kappa^o\},\end{aligned}$$

for any $t\in[0,1]$, $\gamma_1\in[0,1]$ as long as $\delta\geq\sigma_{\min}$ and $\phi\delta\geq\frac{1}{\log(ep)}$. Now we bound

$$|T_{W_J,2}(t)| = \left|\mathbb{E}\left[\left\langle\nabla^\alpha\rho_{r,\phi}^{\delta'}(W^o+W^{\not{J}}+tW_J),(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right\rangle\mathbb{I}_{W_J,2}\right]\right|.$$

Conditional on $W^{\not{J}}$ and W_J,

$$\begin{aligned}&\mathbb{E}\left[\left\|\nabla^\alpha\rho_{r,\phi}^{\delta'}\left(W^o+W^{\not{J}}+tW_J\right)\right\|_1\middle| W^{\not{J}},W_J\right] \\ &\leq \mathbb{E}\left[\left\|\nabla^\alpha\rho_{r,\phi}^{\delta'}\left(W^o+W^{\not{J}}+tW_J\right)\right\|_1\mathbb{I}_{\{\|W^o-\partial A_{r'}\|_\infty\leq 10\delta'\sqrt{\log(ph)}\}}\middle| W^{\not{J}},W_J\right] \\ &\quad+\mathbb{E}\left[\left\|\nabla^\alpha\rho_{r,\phi}^{\delta'}\left(W^o+W^{\not{J}}+tW_J\right)\right\|_1\mathbb{I}_{\{\|W^o-\partial A_{r'}\|_\infty> 10\delta'\sqrt{\log(ph)}\}}\middle| W^{\not{J}},W_J\right],\end{aligned}$$

where $r' = r-W^{\not{J}}-tW_J$ is deterministic given $W^{\not{J}}$ and W_J. Applying Lemmas 6.1, 6.2 of Chernozhukov, Chetverikov and Koike (2020) and Lemma 10.5 of Lopes (2022) to the two terms, respectively,

$$\begin{aligned}&\mathbb{E}\left[\left\|\nabla^\alpha\rho_{r,\phi}^{\delta'}\left(W^o+W^{\not{J}}+tW_J\right)\right\|_1\middle| W^{\not{J}},W_J\right] \\ &\leq C\frac{\phi^{\gamma_2}(\log(ep))^{(\alpha-\gamma_2)/2}}{\delta'^{\alpha-\gamma_2}}\mathbb{P}\left[\|W^o-\partial A_{r'}\|_\infty\leq 10\delta'\sqrt{\log(ph)}\middle| W^{\not{J}},W_J\right]+C\frac{1}{\delta'^\alpha h} \\ &\leq C\frac{\phi^{\gamma_2}(\log(ep))^{(\alpha-\gamma_2)/2}}{\delta'^{\alpha-\gamma_2}}\min\left\{1,\kappa_{J_1}\left(10\delta'\sqrt{\log(ph)}\right)\right\}+C\frac{1}{\delta'^\alpha h},\end{aligned}$$

almost surely, for any $\gamma_2\in[0,1]$ and $h>0$. Putting the last two results together,

$$\begin{aligned}&|T_{W_J,2}(t)| \\ &\leq C\left(\frac{\phi^{\gamma_2}(\log(ep))^{(\alpha-\gamma_2)/2}}{\delta'^{\alpha-\gamma_2}}\min\left\{1,\kappa_{J_1}\left(10\delta'\sqrt{\log(ph)}\right)\right\}+\frac{1}{\delta'^\alpha h}\right) \\ &\quad\times\mathbb{E}\left[\prod_{k=1}^{\alpha-\beta}\|W_{j_k}\|_\infty\|W_J\|_\infty^\beta\mathbb{I}_{W_J,2}\right],\end{aligned}$$

for any $\gamma_2\in[0,1]$ and $h>0$. Because $\mathbb{I}_{W_J,2}=\mathbb{I}_{\{\|W_J\|_\infty \text{ or } \|W^{\not{J}}\|_\infty\geq\frac{\delta'}{\sqrt{\log(ep)}}\}}\geq\mathbb{I}_{\{\|W_J\|_\infty>\frac{\delta'}{\sqrt{\log(ep)}}\}}+\mathbb{I}_{\{\|W^{\not{J}}\|_\infty>\frac{\delta'}{\sqrt{\log(ep)}}\}}$,

$$\begin{aligned}&\mathbb{E}\left[\prod_{k=1}^{\alpha-\beta}\|W_{j_k}\|_\infty\|W_J\|_\infty^\beta\mathbb{I}_{W_J,2}\right] \\ &\leq\mathbb{E}\left[\prod_{k=1}^{\alpha-\beta}\|W_{j_k}\|_\infty\|W_J\|_\infty^\beta\mathbb{I}_{\{\|W_J\|_\infty>\frac{\delta'}{\sqrt{\log(ep)}}\}}\right]+\mathbb{E}\left[\prod_{k=1}^{\alpha-\beta}\|W_{j_k}\|_\infty\|W_J\|_\infty^\beta\mathbb{I}_{\{\|W^{\not{J}}\|_\infty>\frac{\delta'}{\sqrt{\log(ep)}}\}}\right] \\ &\leq\frac{(\log(ep))^{\eta/2}}{\delta'^\eta}\mathbb{E}\left[\prod_{k=1}^{\alpha-\beta}\|W_{j_k}\|_\infty\|W_J\|_\infty^\beta(\|W_J\|_\infty^\eta+\|W^{\not{J}}\|_\infty^\eta)\right],\end{aligned}$$

for any $\eta > 0$. The resulting upper bound for $T_{W_J,2}(t)$ is

$$
\begin{aligned}
&|T_{W_J,2}(t)| \\
&\leq C\left[\frac{\phi^{\gamma_2}(\log(ep))^{(\alpha-\gamma_2)/2}}{\delta'^{\alpha-\gamma_2}}\min\left\{1,\kappa_{J_1}\left(10\delta'\sqrt{\log(ph)}\right)\right\}+\frac{1}{\delta'^{\alpha}h}\right] \\
&\quad\times\frac{(\log(ep))^{\eta/2}}{\delta'^{\eta}}\mathbb{E}\left[\prod_{k=1}^{\alpha-\beta}\|W_{j_k}\|_\infty\|W_J\|_\infty^\beta(\|W_J\|_\infty^\eta+\|W^{\not\perp}\|_\infty^\eta)\right],
\end{aligned}
$$

for any $\gamma_2 \in [0,1]$, $\eta > 0$ and $h > 0$. By similar choices of h with for $T_{W_J,1}(t)$, we obtain

$$
\begin{aligned}
&|T_{W_J,2}(t)| \\
&\leq C\left[\frac{\phi^{\gamma_2}(\log(ep))^{(\alpha-\gamma_2+\eta)/2}}{\delta'^{\alpha-\gamma_2+\eta}}\min\left\{1,\kappa_{J_1}(\delta^o)+\kappa^o\right\}\right] \\
&\quad\times\mathbb{E}\left[\prod_{k=1}^{\alpha-\beta}\|W_{j_k}\|_\infty\|W_J\|_\infty^\beta(\|W_J\|_\infty^\eta+\|W^{\not\perp}\|_\infty^\eta)\right],
\end{aligned}
$$

for any $t \in [0,1]$ and $\eta > 0$ as long as $\delta \geq \sigma_{\min}$ and $\phi\delta \geq \frac{1}{\log(ep)}$. Putting the upperbounds for $T_{W_J,1}(t)$ and $T_{W_J,2}(t)$ together, we obtain the following upper bound for the entire remainder term:

$$
\begin{aligned}
&\left|\int_0^1(1-t)^{\beta-1}\mathbb{E}\left[\left\langle\nabla^\alpha\rho_{r,\phi}^\delta(X_{J_1}+W^{\not\perp}+Y_{J_2}+tW_J),(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right\rangle\right]dt\right| \\
&\leq C\left[\frac{(\log(ep))^{\alpha/2}}{\delta'^{\alpha}}\min\left\{1,\kappa_{J_1}(\delta^o)+\kappa^o\right\}\right] \\
&\times\left[\begin{array}{l}\dfrac{\phi^{\gamma_1}\delta'^{\gamma_1}}{(\log(ep))^{\gamma_1/2}}\left\|\mathbb{E}\left[\left|(\otimes_{k=1}^{\alpha-\beta}W_{j_k})\otimes W_J^{\otimes\beta}\right|\right]\right\|_\infty \\ +\dfrac{\phi^{\gamma_2}\delta'^{\gamma_2-\eta}}{(\log(ep))^{(\gamma_2-\eta)/2}}\mathbb{E}\left[\displaystyle\prod_{k=1}^{\alpha-\beta}\|W_{j_k}\|_\infty\|W_J\|_\infty^\beta(\|W_J\|_\infty^\eta+\|W^{\not\perp}\|_\infty^\eta)\right]\end{array}\right],
\end{aligned}
$$

for any $\gamma_1, \gamma_2 \in [0,1]$ and $\eta > 0$ as long as $\delta \geq \sigma_{\min}$ and $\phi\delta \geq \frac{1}{\log(ep)}$. This proves Lemma B.3.

Lemma B.3 implies all the remainder theorems in the main text and appendices of this paper. For example, in Lemma 5.3, one of the third-order remainder terms in $\mathfrak{R}_{W_j}^{(3,1)}$ is

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^1(1-t)^2\left\langle\nabla^3\varphi_r^\varepsilon\left(W_{[j,j]}^{\mathsf{C}}+tW_j\right),X_j^{\otimes 3}\right\rangle dt\right] \\
&=\mathbb{E}\left[\int_0^1(1-t)^2\left\langle\nabla^3\varphi_r^\varepsilon\left(X_{[1,j-1)}+W^{\not\perp}+Y_{(j+1,n]}+tW_j\right),X_j^{\otimes 3}\right\rangle dt\right],
\end{aligned}
$$

where $W_{[i,j]}^{\mathsf{C}} \equiv X_{[1,i)} + Y_{(j,n]}$ and $W^{\not\perp} \equiv X_{j-1} + Y_{j+1}$. Because of Assumption (MIN-EV), $\mathrm{Var}[Y_{(j+1,n]}|\mathscr{Y}_{(j+1,n]}] \succeq \underline{\sigma}^2\max\{n-j-3,0\}\cdot I_p$. Let $\underline{\sigma}_j^2 \equiv \underline{\sigma}^2\max\{j,0\}$ and $\delta_j^2 \equiv \sqrt{\delta^2+\underline{\sigma}_j^2}$. Applying Lemma B.3 to this term with $\gamma_1 = 0$ and $\gamma_2 = \eta = q-3$, we obtain for $q \geq 3$,

$$\left|\mathbb{E}\left[\int_0^1 (1-t)^2 \left\langle \nabla^3 \varphi_r^\varepsilon \left(X_{[1,j-1)} + W^{\not{j}} + Y_{(j+1,n]} + tW_j\right), W_j^{\otimes 3}\right\rangle dt\right]\right|$$
$$\leq C \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \min\{1, \kappa_{[1,j-1)}(\delta_{n-j}^o) + \kappa_j^o\}$$
$$\times \left[\left\|\mathbb{E}\left[\left|W_j^{\otimes 3}\right|\right]\right\|_\infty + \phi^{q-3}\mathbb{E}\left[\|W_j\|_\infty^3(\|W_j\|_\infty^{q-3} + \|X_{j-1}+Y_{j+1}\|_\infty^{q-3})\right]\right]$$
$$\leq C \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left[L_{3,j} + (2\phi)^{q-3}(\nu_{q,j-1} + \nu_{q,j} + \nu_{q,j+1})\right] \min\{1, \kappa_{[1,j-1)}(\delta_{n-j}^o) + \kappa_j^o\},$$

where $\delta_{n-j}^o \equiv 12\delta_{n-j}\sqrt{\log(pn)}$ and $\kappa_j^o \equiv \frac{\delta_{n-j}\log(ep)}{\sigma_{\min}\sqrt{\max\{j,1\}}}$ as long as $\delta \geq \sigma_{\min}$ and $\phi\delta \geq \frac{1}{\log(ep)}$. Based on similar applications of Lemma B.3 to the other terms of $\mathfrak{R}_{X_j}^{(3,1)}$, for $q \geq 3$,

$$\left|\mathbb{E}\left[\mathfrak{R}_{W_j}^{(3,1)}\right]\right| \leq C \frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \left(\sum_{j'=j-3}^{j+3} L_{3,j'} + (2\phi)^{q-3} \sum_{j'=j-3}^{j+3} \nu_{q,j'}\right)$$
$$\times \min\{\kappa_{[1,j-3)}(\delta_{n-j}^o) + \kappa_j^o, 1\},$$

as long as $\delta \geq \sigma_{\min}$ and $\phi\delta \geq \frac{1}{\log(ep)}$. Similar arguments and bounds apply to $\mathfrak{R}_W^{(3)}$ and $\mathfrak{R}_{W_j}^{(3,2)}$. For the fourth-order remainder terms, we consider the first term of $\mathfrak{R}_{W_j}^{(4,1)}$:

$$\mathbb{E}\left[\int_0^1 (1-t)^3 \left\langle \nabla^4 \rho_{r,\phi}^\delta (W_{[j,j]}^{\mathsf{C}} + tX_j), X_j^{\otimes 4}\right\rangle dt\right]$$
$$= \mathbb{E}\left[\int_0^1 (1-t)^3 \left\langle \nabla^4 \rho_{r,\phi}^\delta (X_{[1,j-1)} + W^{\not{j}} + Y_{(j+1,n]} + tX_j), X_j^{\otimes 4}\right\rangle dt\right].$$

Applying Lemma B.3 to this term with $\gamma_1 = \gamma_2 = 1$ and $\eta = q-4$, we obtain for $q \geq 4$,

$$\left|\mathbb{E}\left[\int_0^1 (1-t)^3 \left\langle \nabla^4 \varphi_r^\varepsilon \left(X_{[1,j-1)} + W^{\not{j}} + Y_{(j+1,n]} + tW_j\right), W_j^{\otimes 4}\right\rangle dt\right]\right|$$
$$\leq C\phi \min\{1, \kappa_{[1,j-1)}(\delta_{n-j}^o) + \kappa_j^o\}$$
$$\times \left[\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3}\left\|\mathbb{E}\left[\left|W_j^{\otimes 4}\right|\right]\right\|_\infty + \frac{(\log(ep))^{(q-1)/2}}{(\delta_{n-j}/2)^{q-1}}\mathbb{E}\left[\|W_j\|_\infty^4(\|W_j\|_\infty^{q-4} + \|X_{j-1}+Y_{j+1}\|_\infty^{q-4})\right]\right]$$
$$\leq C\phi \left[\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} L_{4,j} + \frac{(\log(ep))^{(q-1)/2}}{(\delta_{n-j}/2)^{q-1}}(\nu_{q,j-1} + \nu_{q,j} + \nu_{q,j+1})\right] \min\{1, \kappa_{[1,j-1)}(\delta_{n-j}^o) + \kappa_j^o\},$$

as long as $\delta \geq \sigma_{\min}$ and $\phi\delta \geq \frac{1}{\log(ep)}$. Based on similar applications of Lemma B.3 to the other terms of $\mathfrak{R}_{X_j}^{(4,1)}$, for $q \geq 4$,

$$\left|\mathbb{E}\left[\mathfrak{R}_{W_j}^{(4,1)}\right]\right| \leq C\phi \left(\frac{(\log(ep))^{3/2}}{\delta_{n-j}^3} \sum_{j'=j-4}^{j+4} L_{4,j'} + \frac{(\log(ep))^{(q-1)/2}}{(\delta_{n-j}/2)^{q-1}} \sum_{j'=j-4}^{j+4} \nu_{q,j'}\right)$$
$$\times \min\{\kappa_{[1,j-4)}(\delta_{n-j}^o) + \kappa_j^o, 1\},$$

as long as $\delta \geq \sigma_{\min}$ and $\phi\delta \geq \frac{1}{\log(ep)}$. Similar arguments and bounds apply to $\mathfrak{R}_W^{(4)}$ and $\mathfrak{R}_{W_j}^{(4,2)}$. Finally, for the sixth-order remainder terms, we consider the first term of $\mathfrak{R}_{X_j,W_k}^{(6,1)}$:

$$\mathbb{E}\left[\int_0^1 (1-t)^2 \left\langle \nabla^6 \rho_{r,\phi}^{\varepsilon}(X_{[1,k)} + tX_k + Y_{(k,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes W_k^{\otimes 3}\right\rangle dt\right]$$
$$= \mathbb{E}\left[\int_0^1 (1-t)^2 \left\langle \nabla^6 \rho_{r,\phi}^{\varepsilon}(X_{[1,k-1)} + W^{\not k} + Y_{(k+1,j-1)\cup(j+1,n)}) + tX_k, \mathbb{E}[X_j^{\otimes 3}] \otimes W_k^{\otimes 3}\right\rangle dt\right],$$

where $W^{\not k} \equiv X_{k-1} + Y_{k+1}$. Applying Lemma B.3 to this term with $\gamma_1 = \gamma_2 = 1$ and $\eta = q-3$, we obtain for $q \geq 4$,

$$\begin{aligned}
&\left|\mathbb{E}\left[\int_0^1 (1-t)^2 \left\langle \nabla^6 \rho_{r,\phi}^{\varepsilon}(X_{[1,k)} + tX_k + Y_{(k,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes W_k^{\otimes 3}\right\rangle dt\right]\right| \\
&\leq C\phi \min\{1, \kappa_{[1,j-1)}(\delta_{n-j}^o) + \kappa_j^o\} \\
&\quad \times \left[\begin{array}{l} \frac{(\log(ep))^{3/2}}{\delta_{n-k}^3} \left\| \mathbb{E}\left[\left|X_j^{\otimes 3}\right|\right] \mathbb{E}\left[|W_k^{\otimes 3}|\right] \right\|_\infty \\ + \frac{(\log(ep))^{(q-1)/2}}{(\delta_{n-k}/2)^{q-1}} \left\| \mathbb{E}\left[\left|X_j^{\otimes 3}\right|\right] \right\|_\infty \mathbb{E}\left[\|W_k\|_\infty^3 (\|W_k\|_\infty^{q-3} + \|X_{k-1} + Y_{k+1}\|_\infty^{q-3})\right] \end{array}\right] \\
&\leq C\phi L_{3,j}\left(L_{3,k}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5} + (\nu_{q,k-1} + \nu_{q,k} + \nu_{q,k+1})\frac{(\log(ep))^{(q+2)/2}}{(\delta_{n-k}/2)^{q+2}}\right) \\
&\quad \times \min\{\kappa_{[1,k-1)}(\delta_{n-k-5}^o) + \kappa_k^o, 1\}.
\end{aligned}$$

Similarly, we get

$$\begin{aligned}
\left|\mathbb{E}\left[\mathfrak{R}_{X_j,W_k}^{(6,1)}\right]\right| \leq C\phi L_{3,j} &\left(\sum_{k'=k-2}^{k+2} L_{3,k'}\frac{(\log(ep))^{5/2}}{\delta_{n-k}^5} + \sum_{k'=k-3}^{k+3} \nu_{q,k'}\frac{(\log(ep))^{(q+2)/2}}{(\delta_{n-k}/2)^{q+2}}\right) \\
&\times \min\{\kappa_{[1,k-3)}(\delta_{n-k}^o) + \kappa_k^o, 1\}.
\end{aligned}$$

Similar arguments and bounds apply to the other sixth-order remainder terms.

B.4. Proof of Anti-concentration Lemma. We use a similar approach with Appendix B.3 to upper bound $\mathbb{E}\left[\mathfrak{R}_{X_{\{2,i-1\}},1} | \mathscr{X}_{(i,n]}\right]$. First, we decompose $\mathbb{E}\left[\mathfrak{R}_{X_{\{2,i-1\}},1} | \mathscr{X}_{(i,n]}\right]$ by

$$\begin{aligned}
&\mathbb{E}\left[\mathfrak{R}_{X_{\{2,i-1\}},1} | \mathscr{X}_{(i,n]}\right] \\
&= \mathbb{E}\left[\int_0^1 \langle \nabla \varphi_{r,\delta+\varepsilon^o}^{\varepsilon}(X_{[1,i]} - tX_{\{2,i-1\}}), X_{\{2,i-1\}}\rangle dt \middle| \mathscr{X}_{(i,n]}\right] \\
&= \mathbb{E}\left[\int_0^1 \langle \nabla \varphi_{r,\delta+\varepsilon^o}^{\varepsilon}(X_{[1,i]} - tX_{\{2,i-1\}}), \mathscr{X}_{\{2,i-1\}}\rangle \mathbb{I}_{\leq}(t) dt \middle| X_{(i,n]}\right] \\
&\quad + \mathbb{E}\left[\int_0^1 \langle \nabla \varphi_{r,\delta+\varepsilon^o}^{\varepsilon}(X_{[1,i]} - tX_{\{2,i-1\}}), X_{\{2,i-1\}}\rangle \mathbb{I}_{>}(t) dt \middle| \mathscr{X}_{(i,n]}\right],
\end{aligned}$$

where $\mathbb{I}_{\leq}(t)=\mathbb{I}\{\|X_{[1,i]}-tX_{\{2,i-1\}}-\partial A_{r,\delta+\varepsilon^o}\|_\infty\leq\varepsilon^o\}$, $\mathbb{I}_{>}(t)=\mathbb{I}\{\|X_{[1,i]}-tX_{\{2,i-1\}}-\partial A_{r,\delta+\varepsilon^o}\|_\infty>\varepsilon^o\}$. Based on Lemma 2.3 of Fang and Koike (2020),

$$\begin{aligned}&\mathbb{E}\left[\int_0^1\langle\nabla\varphi^{\varepsilon}_{r,\delta+\varepsilon^o}(X_{[1,i]}-tX_{\{2,i-1\}}),X_{\{2,i-1\}}\rangle\mathbb{I}_{\leq}(t)dt\,\middle|\,\mathscr{X}_{(i,n]}\right]\\&\leq\mathbb{E}\left[\int_0^1\|\nabla\varphi^{\varepsilon}_{r,\delta+\varepsilon^o}(X_{[1,i]}-tX_{\{2,i-1\}})\|_1\mathbb{I}_{\leq}(t)\|X_{\{2,i-1\}}\|_\infty dt\,\middle|\,\mathscr{X}_{(i,n]}\right]\\&\leq C\frac{\sqrt{\log(ep)}}{\varepsilon}\mathbb{E}\left[\int_0^1\mathbb{I}_{\leq}(t)\|X_{\{2,i-1\}}\|_\infty dt\,\middle|\,\mathscr{X}_{(i,n]}\right]\end{aligned}$$

We use Tonelli's theorem to switch the order between the integration and the expectation conditional on $\mathscr{X}_{(i-2,3)}$:

$$\begin{aligned}&\mathbb{E}\left[\int_0^1\mathbb{I}_{\leq}(t)\|X_{\{2,i-1\}}\|_\infty dt\,\middle|\,\mathscr{X}_{(i,n]}\right]\\&\leq\mathbb{E}\left[\int_0^1\mathbb{E}\left[\mathbb{I}_{\leq}(t)\|X_{\{2,i-1\}}\|_\infty\middle|\,|\mathscr{X}_{(i-2,n+3)}\right]dt\,\middle|\,\mathscr{X}_{(i,n]}\right]\\&=\mathbb{E}\left[\int_0^1\begin{pmatrix}\mathbb{P}[X_{[3,i-2]}\in A_{r_1,\varepsilon^o}|\mathscr{X}_{(i-2,n+3)}]\\+\mathbb{P}[X_{[3,i-2]}\in A_{r_2,\varepsilon^o}|\mathscr{X}_{(i-2,n+3)}]\end{pmatrix}\|X_{\{2,i-1\}}\|_\infty dt\,\middle|\,\mathscr{X}_{(i,n]}\right]\\&\leq\kappa_{[3,i-2]}(\varepsilon^o)\mathbb{E}[\|X_{\{2,i-1\}}\|_\infty|\mathscr{X}_{(i,n]}]=\kappa_{[3,i-2]}(\varepsilon^o)(\nu_1(X_2)+\nu_1(X_{i-1})),\end{aligned}$$

where $r_1=r-(1-t)X_{\{2,i-1\}}-X_1-X_i-(\delta+\varepsilon^o)\mathbf{1}$ and $r_2=r-(1-t)X_{\{2,i-1\}}-X_1-X_i+(\delta+\varepsilon^o)\mathbf{1}$ are Borel measurable functions with respect to $X_{(i-2,n+3)}$. On the other hand, based on Lemma 10.5 of Lopes (2022),

$$\begin{aligned}&\mathbb{E}\left[\int_0^1\langle\nabla\varphi^{\varepsilon}_{r,\delta+\varepsilon^o}(X_{[1,i]}-tX_{\{2,i-1\}}),X_{\{2,i-1\}}\rangle\mathbb{I}_{>}(t)dt\,\middle|\,\mathscr{X}_{(i,n]}\right]\\&\leq\mathbb{E}\left[\int_0^1\|\nabla\varphi^{\varepsilon}_{r,\delta+\varepsilon^o}(X_{[1,i]}-tX_{\{2,i-1\}})\|_1\mathbb{I}_{>}(t)\|X_{\{2,i-1\}}\|_\infty dt\,\middle|\,\mathscr{X}_{(i,n]}\right]\\&=\frac{1}{\varepsilon h^4}\mathbb{E}\left[\int_0^1\mathbb{E}\left[\|X_{\{2,i-1\}}\|_\infty\middle|\,\mathscr{X}_{(i-2,n+3)}\right]dt\,\middle|\,\mathscr{X}_{(i,n]}\right]\\&=\frac{1}{\varepsilon h^4}\mathbb{E}[\|X_{\{2,i-1\}}\|_\infty|\mathscr{X}_{(i,n]}]=\frac{1}{\varepsilon h^4}(\nu_{1,2}+\nu_{1,i-1}).\end{aligned}$$

Putting the above result and Eq. (17) back to Eq. (16), we get for any $h>0$,

$$\begin{aligned}&\mathbb{P}[X_{[1,i]}\in A_{r,\delta}|\mathscr{X}_{(i,n]}]\\&\leq(\nu_{1,2}+\nu_{1,i-1})\left(C\frac{\sqrt{\log(ep)}}{\varepsilon}\max\left\{1,\kappa_{[3,i-2]}(\varepsilon^o)\right\}+\frac{1}{\varepsilon h^4}\right)\\&\quad+C\frac{\delta+2\varepsilon^o}{\sigma_{\min}}\sqrt{\frac{\log(ep)}{i-2}}+2\mu_{[3,i-2]}.\end{aligned}$$

Plugging in $h = \left(\frac{\sigma_{\min}}{\varepsilon}\sqrt{\frac{i-2}{(\log(ep))^2}}\right)^{1/4}$,

$$
\begin{aligned}
&\mathbb{P}[X_{[1,i]} \in A_{r,\delta} | \mathscr{X}_{(i,n]}] \\
&\leq C(\nu_{1,2} + \nu_{1,i-1}) \frac{\sqrt{\log(ep)}}{\varepsilon} \max\left\{1, \kappa_{[3,i-2]}(\varepsilon^o) + \kappa^o_{i-2}\right\} \\
&\quad + C\frac{\delta + 2\varepsilon^o}{\sigma_{\min}}\sqrt{\frac{\log(ep)}{i-2}} + 2\mu_{[3,i-2]},
\end{aligned}
$$

where $\varepsilon^o = 10\varepsilon\sqrt{\log(p\max\{i_0-2,1\})}$ and $\kappa^o_j = \frac{\varepsilon}{\sigma_{\min}}\sqrt{\frac{\log(ep)}{\max\{j,1\}}}$, as long as $\varepsilon \geq \sigma_{\min}$. Because the righthand side is not dependent on r, for any $\varepsilon \geq \sigma_{\min}, \delta > 0$,

$$
\begin{aligned}
&\kappa_{[1,i]}(\delta) \\
&\leq C\left((\nu_{1,2} + \nu_{1,i-1})\frac{\sqrt{\log(ep)}}{\varepsilon}\kappa_{[3,i-2]}(\varepsilon^o) + \mu_{[3,i-2]}\right) \\
&\quad + C\frac{\delta + 2\varepsilon^o}{\sigma_{\min}}\sqrt{\frac{\log(ep)}{i-2}} + C\frac{\overline{\nu}_{1,(1,i)}}{\sigma_{\min}}\frac{\log(ep)}{\sqrt{i-2}}.
\end{aligned}
$$

B.5. Proof of Lemma A.7. Suppose that $j, k \geq \frac{n}{2}$ and $\alpha \leq 1$. Because $x \mapsto x^\alpha$ is concave, by Jensen's inequality, for any $q_1, q_2 > 0$,

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^n L_{q_1,i+j}\overline{L}^\alpha_{q_2,(i,i+j)} &\leq \overline{L}_{q_1}\left(\frac{\frac{1}{n}\sum_{i=1}^n L_{q_1,i+j}\overline{L}_{q_2,(i,i+j)}}{\overline{L}_{q_1}}\right)^\alpha \\
&\leq \overline{L}_{q_1}\left(\frac{\frac{1}{n(j-1)}\sum_{i=1}^n L_{q_1,i}\sum_{i=1}^n L_{q_2,i}}{\overline{L}_{q_1}}\right)^\alpha \\
&\leq \overline{L}_{q_1}\left(\frac{n}{j-1}\overline{L}_{q_2}\right)^\alpha \leq 2^\alpha \overline{L}_{q_1}\overline{L}_{q_2}.
\end{aligned}
$$

Similarly, for any $q_1, q_2, q_3 > 0$,

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \frac{L_{q_1,i+j}}{j-1}\sum_{l=1}^{j-1} L_{q_2,i+[k+l]_{j-1}}\overline{L}^\alpha_{q_3,i+(l,k+l)_{j-1}} \\
&\leq \frac{1}{n}\sum_{i=1}^n \frac{L_{q_1,i+j}}{j-1}\sum_{l=1}^{j-1} L_{q_2,i+[k+l]_{j-1}}\left(\frac{\frac{1}{n}\sum_{i=1}^n \frac{L_{q_1,i+j}}{j-1}\sum_{l=1}^{j-1} L_{q_2,i+[k+l]_{j-1}}\overline{L}_{q_3,i+(l,k+l)_{j-1}}}{\frac{1}{n}\sum_{i=1}^n \frac{L_{q_1,i+j}}{j-1}\sum_{l=1}^{j-1} L_{q_2,i+[k+l]_{j-1}}}\right)^\alpha \\
&\leq \frac{1}{n}\sum_{i=1}^n L_{q_1,i+j}\overline{L}_{q_2,(i,i+j)}\left(\frac{\frac{1}{j-1}\sum_{i=1}^n \frac{L_{q_1,i+j}}{j-1}\sum_{l=1}^{j-1} L_{q_2,i+[k+l]_{j-1}}\overline{L}_{q_3}}{\frac{1}{n}\sum_{i=1}^n \frac{L_{q_1,i+j}}{j-1}\sum_{l=1}^{j-1} L_{q_2,i+[k+l]_{j-1}}}\right)^\alpha \\
&\leq \frac{1}{n}\sum_{i=1}^n L_{q_1,i+j}\overline{L}_{q_2,(i,i+j)}\left(\frac{n}{j-1}\overline{L}_{q_3}\right)^\alpha \\
&\leq \frac{2^\alpha}{n}\sum_{i=1}^n L_{q_1,i+j}\overline{L}_{q_2,(i,i+j)}\overline{L}^\alpha_{q_3} \\
&\leq 2^{\alpha+1}\overline{L}_{q_1}\overline{L}_{q_2}\sum_{l=1}^{j-1} L_{q_2,i+[k+l]_{j-1}}\overline{L}^\alpha_{q_3}.
\end{aligned}
$$

APPENDIX C: DETAILS OF THE TAYLOR EXPANSIONS FOR $m = 1$

C.1. Breaking the ring. We apply the Taylor expansion to $\rho^{\delta}_{r,\phi}(X_{[1,n]})$ centered at $\rho^{\delta}_{r,\phi}(X_{[1,n)})$ as follows:

$$\begin{aligned}
&\rho^{\delta}_{r,\phi}(X_{[1,n]}) - \rho^{\delta}_{r,\phi}(X_{[1,n)}) \\
&= \Big\langle \nabla \rho^{\delta}_{r,\phi}(X_{[1,n)}), X_n \Big\rangle + \frac{1}{2}\Big\langle \nabla^2 \rho^{\delta}_{r,\phi}(X_{[1,n)}), X_n^{\otimes 2} \Big\rangle \\
&\quad + \frac{1}{2}\int_0^1 (1-t)^2 \Big\langle \nabla^3 \rho^{\delta}_{r,\phi}(X_{[1,n)} + tX_n), X_n^{\otimes 3} \Big\rangle \, dt.
\end{aligned}$$

Because $X_{[1,n)}$ and X_n are dependent via X_1 and X_{n-1} due to 1-dependency, we re-apply the Taylor expansion to $\nabla \rho^{\delta}_{r,\phi}(X_{[1,n)})$ and $\nabla^2 \rho^{\delta}_{r,\phi}(X_{[1,n)})$ centered at $X_{(1,n-1)}$:

$$\begin{aligned}
&\Big\langle \nabla \rho^{\delta}_{r,\phi}(X_{[1,n)}), X_n \Big\rangle \\
&= \Big\langle \nabla \rho^{\delta}_{r,\phi}(X_{(1,n-1)}), X_n \Big\rangle + \Big\langle \nabla^2 \rho^{\delta}_{r,\phi}(X_{(1,n-1)}), X_n \otimes (X_1 + X_{n-1}) \Big\rangle \\
&\quad + \int_0^1 (1-t) \Big\langle \nabla^3 \rho^{\delta}_{r,\phi}(X_{(1,n-1)} + t(X_1 + X_{n-1})), X_n \otimes (X_1 + X_{n-1})^{\otimes 2} \Big\rangle \, dt, \quad \text{and}
\end{aligned}$$

$$\begin{aligned}
&\Big\langle \nabla^2 \rho^{\delta}_{r,\phi}(X_{[1,n)}), X_n^{\otimes 2} \Big\rangle \\
&= \Big\langle \nabla^2 \rho^{\delta}_{r,\phi}(X_{(1,n-1)}), X_n^{\otimes 2} \Big\rangle \\
&\quad + \int_0^1 \Big\langle \nabla^3 \rho^{\delta}_{r,\phi}(X_{(1,n-1)} + t(X_1 + X_{n-1})), X_n^{\otimes 2} \otimes (X_1 + X_{n-1}) \Big\rangle \, dt.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\Big\langle \nabla^2 \rho^{\delta}_{r,\phi}(X_{(1,n-1)}), X_n \otimes X_1 \Big\rangle \\
&= \Big\langle \nabla^2 \rho^{\delta}_{r,\phi}(X_{(2,n-1)}), X_n \otimes X_1 \Big\rangle \\
&\quad + \int_0^1 \Big\langle \nabla^3 \rho^{\delta}_{r,\phi}(X_{(2,n-1)} + tX_2), X_n \otimes X_1 \otimes X_2 \Big\rangle \, dt, \quad \text{and}
\end{aligned}$$

$$\begin{aligned}
&\Big\langle \nabla^2 \rho^{\delta}_{r,\phi}(X_{(1,n-1)}), X_n \otimes X_{n-1} \Big\rangle \\
&= \Big\langle \nabla^2 \rho^{\delta}_{r,\phi}(X_{(1,n-2)}), X_n \otimes X_{n-1} \Big\rangle \\
&\quad + \int_0^1 \Big\langle \nabla^3 \rho^{\delta}_{r,\phi}(X_{(1,n-2)} + tX_2), X_n \otimes X_{n-1} \otimes X_{n-2} \Big\rangle \, dt.
\end{aligned}$$

In sum, because $\mathbb{E}[X_j] = 0$,

$$\begin{aligned}
&\mathbb{E}\left[\rho^{\delta}_{r,\phi}(X_{[1,n]}) - \rho^{\delta}_{r,\phi}(X_{[1,n)})\right] \\
&= \frac{1}{2}\Big\langle \mathbb{E}[\nabla^2 \rho^{\delta}_{r,\phi}(X_{(1,n-1)})], \mathbb{E}[X_n^{\otimes 2}] \Big\rangle + \Big\langle \mathbb{E}[\nabla^2 \rho^{\delta}_{r,\phi}(X_{(2,n-1)})], \mathbb{E}[X_n \otimes X_1] \Big\rangle \\
&\quad + \Big\langle \mathbb{E}[\nabla^2 \rho^{\delta}_{r,\phi}(X_{(1,n-2)})], \mathbb{E}[X_{n-1} \otimes X_n] \Big\rangle + \mathbb{E}\left[\mathfrak{R}^{(3)}_X\right],
\end{aligned}$$

where $\mathfrak{R}_X^{(3)}$ is the summation of all the above third-order remainder terms. That is,

$$
\begin{aligned}
&\mathfrak{R}_X^{(3)} \\
&= \frac{1}{2}\int_0^1 (1-t)^2 \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{[1,n)} + tX_n), X_n^{\otimes 3} \Big\rangle \, dt \\
&\quad + \int_0^1 (1-t) \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(1,n-1)} + t(X_1 + X_{n-1})), X_n \otimes (X_1 + X_{n-1})^{\otimes 2} \Big\rangle \, dt \\
&\quad + \frac{1}{2}\int_0^1 \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(1,n-1)} + t(X_1 + X_{n-1})), X_n^{\otimes 2} \otimes (X_1 + X_{n-1}) \Big\rangle \, dt \\
&\quad + \int_0^1 \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(2,n-1)} + tX_2), X_n \otimes X_1 \otimes X_2 \Big\rangle \, dt \\
&\quad + \int_0^1 \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(1,n-2)} + tX_2), X_n \otimes X_{n-1} \otimes X_{n-2} \Big\rangle \, dt
\end{aligned}
$$

To further decompose $\mathfrak{R}_X^{(3)}$, we apply the Taylor expansion up to order 4. For example,

$$
\begin{aligned}
&\frac{1}{2}\int_0^1 (1-t)^2 \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{[1,n)} + tX_n), X_n^{\otimes 3} \Big\rangle \, dt \\
&= \frac{1}{6}\Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{[1,n)}), X_n^{\otimes 3} \Big\rangle + \frac{1}{6}\int_0^1 (1-t)^3 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{[1,n)} + tX_n), X_n^{\otimes 4} \Big\rangle \, dt.
\end{aligned}
$$

Again, to break the dependency between $\nabla^3 \rho_{r,\phi}^{\delta}(X_{[1,n)})$ and $X_n^{\otimes 3}$, we re-apply the Taylor expansion centered at $X_1 + X_{n-1}$.

$$
\begin{aligned}
&\Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{[1,n)}), X_n^{\otimes 3} \Big\rangle \\
&= \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(1,n-1)}), X_n^{\otimes 3} \Big\rangle \\
&\quad + \int_0^1 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(1,n-1)} + t(X_1 + X_{n-1})), X_n^{\otimes 3} \otimes (X_1 + X_{n-1}) \Big\rangle \, dt.
\end{aligned}
$$

Repeating this to the other third-order remainder terms, we get

$$
\begin{aligned}
&\mathfrak{R}_X^{(3)} \\
&= \frac{1}{6}\Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(1,n-1)}), X_n^{\otimes 3} \Big\rangle + \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(1,n-3)}), X_{n-2} \otimes X_{n-1} \otimes X_n \Big\rangle \\
&\quad + \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(2,n-2)}), X_{n-1} \otimes X_n \otimes X_1 \Big\rangle + \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(3,n-1)}), X_n \otimes X_1 \otimes X_2 \Big\rangle \\
&\quad + \frac{1}{2}\Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(1,n-2)}), X_{n-1} \otimes X_n \otimes (X_{n-1} + X_n) \Big\rangle \\
&\quad + \frac{1}{2}\Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(X_{(2,n-1)}), X_n \otimes X_1 \otimes (X_1 + X_n) \Big\rangle \\
&\quad + \mathfrak{R}_X^{(4)},
\end{aligned}
$$

where

$$
\begin{aligned}
&\mathfrak{R}_X^{(4)}\\
&= \frac{1}{6}\int_0^1 (1-t)^3 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{[1,n)} + tX_n), X_n^{\otimes 4}\Big\rangle\, dt\\
&\quad + \frac{1}{2}\int_0^1 (1-t)^2 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(1,n-1)} + t(X_1 + X_{n-1})), X_n \otimes (X_1 + X_{n-1})^{\otimes 3}\Big\rangle\, dt\\
&\quad + \frac{1}{2}\int_0^1 (1-t) \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(1,n-1)} + t(X_1 + X_{n-1})), X_n^{\otimes 2} \otimes (X_1 + X_{n-1})^{\otimes 2}\Big\rangle\, dt\\
&\quad + \frac{1}{6}\int_0^1 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(1,n-1)} + t(X_1 + X_{n-1})), X_n^{\otimes 3} \otimes (X_1 + X_{n-1})\Big\rangle\, dt\\
&\quad + \int_0^1 (1-t) \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(2,n-1)} + tX_2), X_n \otimes X_1 \otimes X_2^{\otimes 2}\Big\rangle\, dt\\
&\quad + \int_0^1 (1-t) \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(1,n-2)} + tX_{n-2}), X_n \otimes X_{n-1} \otimes X_{n-2}^{\otimes 2}\Big\rangle\, dt\\
&\quad + \frac{1}{2}\int_0^1 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(2,n-1)} + tX_2), X_n \otimes X_1^{\otimes 2} \otimes X_2\Big\rangle\, dt\\
&\quad + \frac{1}{2}\int_0^1 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(1,n-2)} + tX_{n-2}), X_n \otimes X_{n-1}^{\otimes 2} \otimes X_{n-2}\Big\rangle\, dt\\
&\quad + \frac{1}{2}\int_0^1 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(2,n-1)} + tX_2), X_n^{\otimes 2} \otimes X_1 \otimes X_2\Big\rangle\, dt\\
&\quad + \frac{1}{2}\int_0^1 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(1,n-2)} + tX_{n-2}), X_n^{\otimes 2} \otimes X_{n-1} \otimes X_{n-2}\Big\rangle\, dt\\
&\quad + \int_0^1 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(3,n-1)} + tX_3), X_n \otimes X_1 \otimes X_2 \otimes X_3\Big\rangle\, dt\\
&\quad + \int_0^1 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(2,n-2)} + t(X_2 + X_{n-2})), X_n \otimes X_{n-1} \otimes X_1 \otimes (X_2 + X_{n-2})\Big\rangle\, dt\\
&\quad + \int_0^1 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(X_{(1,n-3)} + tX_{n-3}), X_n \otimes X_{n-1} \otimes X_{n-2} \otimes X_{n-3}\Big\rangle\, dt.
\end{aligned}
$$

The third-order moments in Eq. (31) are further decomposed as in Appendix C.4. As a result,

$$
\begin{aligned}
&\Big\langle \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\delta}(X_{(1,n-1)})\right], \mathbb{E}[X_n^{\otimes 3}]\Big\rangle\\
&= \Big\langle \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\delta}(Y_{(1,n-1)})\right], \mathbb{E}[X_n^{\otimes 3}]\Big\rangle + \sum_{k=2}^{n-2} \mathbb{E}\left[\mathfrak{R}_{X,X_k}^{(6,1)} - \mathfrak{R}_{X,Y_k}^{(6,1)}\right],
\end{aligned}
$$

$$
\begin{aligned}
&\Big\langle \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\delta}(X_{(1,n-3)})\right], \mathbb{E}[X_{n-2} \otimes X_{n-1} \otimes X_n]\Big\rangle\\
&= \Big\langle \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\delta}(Y_{(1,n-3)})\right], \mathbb{E}[X_{n-2} \otimes X_{n-1} \otimes X_n]\Big\rangle + \sum_{k=2}^{n-4} \mathbb{E}\left[\mathfrak{R}_{X,X_k}^{(6,2)} - \mathfrak{R}_{X,Y_k}^{(6,2)}\right],
\end{aligned}
$$

$$\begin{aligned}&\left\langle \mathbb{E}\left[\nabla^3\rho^\delta_{r,\phi}(X_{(2,n-2)})\right], \mathbb{E}[X_{n-1}\otimes X_n\otimes X_1]\right\rangle\\&=\left\langle \mathbb{E}\left[\nabla^3\rho^\delta_{r,\phi}(Y_{(2,n-2)})\right], \mathbb{E}[X_{n-1}\otimes X_n\otimes X_1]\right\rangle+\sum_{k=3}^{n-3}\mathbb{E}\left[\mathfrak{R}^{(6,3)}_{X,X_k}-\mathfrak{R}^{(6,3)}_{X,Y_k}\right],\end{aligned}$$

$$\begin{aligned}&\left\langle \mathbb{E}\left[\nabla^3\rho^\delta_{r,\phi}(X_{(3,n-1)})\right], \mathbb{E}[X_n\otimes X_1\otimes X_2]\right\rangle\\&=\left\langle \mathbb{E}\left[\nabla^3\rho^\delta_{r,\phi}(Y_{(3,n-1)})\right], \mathbb{E}[X_n\otimes X_1\otimes X_2]\right\rangle+\sum_{k=4}^{n-2}\mathbb{E}\left[\mathfrak{R}^{(6,4)}_{X,X_k}-\mathfrak{R}^{(6,4)}_{X,Y_k}\right],\end{aligned}$$

$$\begin{aligned}&\left\langle \mathbb{E}\left[\nabla^3\rho^\delta_{r,\phi}(X_{(1,n-2)})\right], \mathbb{E}[X_{n-1}\otimes X_n\otimes (X_{n-1}+X_n)]\right\rangle\\&=\left\langle \mathbb{E}\left[\nabla^3\rho^\delta_{r,\phi}(Y_{(1,n-2)})\right], \mathbb{E}[X_{n-1}\otimes X_n\otimes (X_{n-1}+X_n)]\right\rangle+\sum_{k=2}^{n-3}\mathbb{E}\left[\mathfrak{R}^{(6,5)}_{X,X_k}-\mathfrak{R}^{(6,5)}_{X,Y_k}\right], \text{ and}\end{aligned}$$

$$\begin{aligned}&\left\langle \mathbb{E}\left[\nabla^3\rho^\delta_{r,\phi}(X_{(2,n-1)})\right], \mathbb{E}[X_n\otimes X_1\otimes (X_{n-1}+X_n)]\right\rangle\\&=\left\langle \mathbb{E}\left[\nabla^3\rho^\delta_{r,\phi}(Y_{(2,n-1)})\right], \mathbb{E}[X_n\otimes X_1\otimes (X_1+X_n)]\right\rangle+\sum_{k=3}^{n-2}\mathbb{E}\left[\mathfrak{R}^{(6,6)}_{X,X_k}-\mathfrak{R}^{(6,6)}_{X,Y_k}\right].\end{aligned}$$

Because Y is Gaussian, by Lemma 6.2 in Chernozhukov, Chetverikov and Koike (2020) and Assumption (VAR-EV),

$$\begin{aligned}\left|\left\langle \mathbb{E}\left[\nabla^3\rho^\varepsilon_{r,\phi}(Y_{(1,n-1)})\right], \mathbb{E}[X_n^{\otimes 3}]\right\rangle\right| &\le \frac{C}{n^{3/2}}L_{3,n}\frac{(\log(ep))^{3/2}}{\underline{\sigma}^3}\\&\le \frac{C}{n^{3/2}}L_{3,n}\frac{(\log(ep))^{2}}{\underline{\sigma}^2\sigma_{\min}}.\end{aligned}$$

Putting all the terms together,

$$\left|\mathfrak{R}^{(3)}_X-\mathfrak{R}^{(3)}_Y\right|\le\frac{C}{\sqrt{n}}\overline{L}_3\frac{(\log(ep))^2}{\underline{\sigma}^2\sigma_{\min}}+\left|\mathfrak{R}^{(4)}_X-\mathfrak{R}^{(4)}_Y\right|+\sum_{k=2}^{n-2}\left|\mathbb{E}\left[\mathfrak{R}^{(6)}_{X,X_k}-\mathfrak{R}^{(6)}_{X,Y_k}\right]\right|,$$

where $\mathfrak{R}^{(6)}_{X,W_k}=\frac{1}{6}\mathfrak{R}^{(6,1)}_{X,W_k}+\mathfrak{R}^{(6,2)}_{X,W_k}+\mathfrak{R}^{(6,3)}_{X,W_k}+\mathfrak{R}^{(6,4)}_{X,W_k}+\frac{1}{2}\mathfrak{R}^{(6,5)}_{X,W_k}+\frac{1}{2}\mathfrak{R}^{(6,6)}_{X,W_k}$.

C.2. First Lindeberg swapping under 1-dependence. We apply Taylor's expansion to each term with φ^ε_r of the Lindeberg swapping in Eq. (11). We only show the expansions for $j=3,\dots,n-2$ here, but the calculations for $j=1,2,n-1$ and n are similar. We recall the notations

$$W^{\mathsf{C}}_{[i,j]}\equiv X_{[1,i)}+Y_{(j,n]}\quad\text{and}\quad W^{\perp\!\!\!\perp}_{[i,j]}\equiv W^{\mathsf{C}}_{[i-1,j+1]}.$$

First,

$$\begin{aligned}&\varphi^\varepsilon_r(W^{\mathsf{C}}_{[j,j]}+X_j)\\&=\varphi^\varepsilon_r(W^{\mathsf{C}}_{[j,j]})+\left\langle\nabla\varphi^\varepsilon_r(W^{\mathsf{C}}_{[j,j]}),X_j\right\rangle+\frac{1}{2}\left\langle\nabla^2\varphi^\varepsilon_r(W^{\mathsf{C}}_{[j,j]}),X_j^{\otimes 2}\right\rangle\\&\quad+\frac{1}{2}\int_0^1(1-t)^2\left\langle\nabla^3\varphi^\varepsilon_r(W^{\mathsf{C}}_{[j,j]}+tX_j),X_j^{\otimes 3}\right\rangle\,dt.\end{aligned}$$

We further apply Taylor's expansion to the second and third terms:

$$\begin{aligned}
&\left\langle \nabla\varphi_r^{\varepsilon}(W_{[j,j]}^{\mathsf{C}}), X_j \right\rangle \\
&= \left\langle \nabla\varphi_r^{\varepsilon}(W_{[j,j]}^{\perp\!\!\!\perp}), X_j \right\rangle + \left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j,j]}^{\perp\!\!\!\perp}), X_j \otimes (X_{j-1}+Y_{j+1}) \right\rangle \\
&\quad + \int_0^1 (1-t) \left\langle \nabla^3\varphi_r^{\varepsilon}\left(W_{[j,j]}^{\perp\!\!\!\perp} + t(X_{j-1}+Y_{j+1})\right), X_j \otimes (X_{j-1}+Y_{j+1})^{\otimes 2} \right\rangle \, dt,
\end{aligned}$$

$$\begin{aligned}
&\left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j,j]}^{\mathsf{C}}), X_j^{\otimes 2} \right\rangle \\
&= \left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j,j]}^{\perp\!\!\!\perp}), X_j^{\otimes 2} \right\rangle \\
&+ \int_0^1 (1-t) \left\langle \nabla^3\varphi_r^{\varepsilon}\left(W_{[j,j]}^{\perp\!\!\!\perp} + t(X_{j-1}+Y_{j+1})\right), X_j^{\otimes 2} \otimes (X_{j-1}+Y_{j+1}) \right\rangle \, dt.
\end{aligned}$$

Last,

$$\begin{aligned}
&\left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j,j]}^{\perp\!\!\!\perp}), X_j \otimes (X_{j-1}+Y_{j+1}) \right\rangle \\
&= \left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j,j]}^{\perp\!\!\!\perp}), X_j \otimes X_{j-1} \right\rangle + \left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j,j]}^{\perp\!\!\!\perp}), X_j \otimes Y_{j+1} \right\rangle \\
&= \left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j-1,j]}^{\perp\!\!\!\perp}), X_j \otimes X_{j-1} \right\rangle + \left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j,j+1]}^{\perp\!\!\!\perp}), X_j \otimes Y_{j+1} \right\rangle \\
&+ \int_0^1 \left\langle \nabla^3\varphi_r^{\varepsilon}\left(W_{[j-1,j]}^{\perp\!\!\!\perp} + tX_{j-2}\right), X_j \otimes X_{j-1} \otimes X_{j-2} \right\rangle dt \\
&+ \int_0^1 \left\langle \nabla^3\varphi_r^{\varepsilon}\left(W_{[j,j+1]}^{\perp\!\!\!\perp} + tY_{j+2}\right), X_j \otimes Y_{j+1} \otimes Y_{j+2}) \right\rangle dt.
\end{aligned}$$

In sum,

$$\begin{aligned}
&\mathbb{E}\left[\varphi_r^{\varepsilon}(W_{[j,j]}^{\mathsf{C}} + X_j)\right] \\
&= \mathbb{E}\left[\begin{aligned} &\varphi_r^{\varepsilon}(W_{[j,j]}^{\mathsf{C}}) + \left\langle \nabla\varphi_r^{\varepsilon}(W_{[j,j]}^{\perp\!\!\!\perp}), X_j \right\rangle + \frac{1}{2}\left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j,j]}^{\perp\!\!\!\perp}), X_j^{\otimes 2} \right\rangle \\ &+ \left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j-1,j]}^{\perp\!\!\!\perp}), X_j \otimes X_{j-1} \right\rangle + \left\langle \nabla^2\varphi_r^{\varepsilon}(W_{[j,j+1]}^{\perp\!\!\!\perp}), X_j \otimes Y_{j+1} \right\rangle + \mathfrak{R}_{X_j}^{(3,1)} \end{aligned}\right] \\
&= \mathbb{E}\left[\varphi_r^{\varepsilon}(W_{[j,j]}^{\mathsf{C}})\right] + \left\langle \mathbb{E}\left[\nabla\varphi_r^{\varepsilon}(W_{[j,j]}^{\perp\!\!\!\perp})\right], \mathbb{E}\left[X_j\right] \right\rangle + \frac{1}{2}\left\langle \mathbb{E}\left[\nabla^2\varphi_r^{\varepsilon}(W_{[j,j]}^{\perp\!\!\!\perp})\right], \mathbb{E}\left[X_j^{\otimes 2}\right] \right\rangle \\
&\quad + \left\langle \mathbb{E}\left[\nabla^2\varphi_r^{\varepsilon}(W_{[j-1,j]}^{\perp\!\!\!\perp})\right], \mathbb{E}\left[X_j \otimes X_{j-1}\right] \right\rangle + \mathbb{E}[\mathfrak{R}_{X_j}^{(3,1)}],
\end{aligned}$$

where

$$\mathfrak{R}_{X_j}^{(3,1)} = \frac{1}{2}\int_0^1 (1-t)^2 \Big\langle \nabla^3 \varphi_r^{\varepsilon}\left(W_{[j,j]}^{\mathsf{C}} + tX_j\right), X_j^{\otimes 3}\Big\rangle\, dt$$
$$+ \int_0^1 (1-t)\Big\langle \nabla^3 \varphi_r^{\varepsilon}\left(W_{[j,j]}^{\perp\!\!\!\perp} + t(X_{j-1}+Y_{j+1})\right), X_j \otimes (X_{j-1}+Y_{j+1})^{\otimes 2}\Big\rangle\, dt$$
$$+ \int_0^1 (1-t)\Big\langle \nabla^3 \varphi_r^{\varepsilon}\left(W_{[j,j]}^{\perp\!\!\!\perp} + t(X_{j-1}+Y_{j+1})\right), X_j^{\otimes 2} \otimes (X_{j-1}+Y_{j+1})\Big\rangle\, dt$$
$$+ \int_0^1 \Big\langle \nabla^3 \varphi_r^{\varepsilon}\left(W_{[j-1,j]}^{\perp\!\!\!\perp} + tX_{j-2}\right), X_j \otimes X_{j-1} \otimes X_{j-2}\Big\rangle\, dt$$
$$+ \int_0^1 \Big\langle \nabla^3 \varphi_r^{\varepsilon}\left(W_{[j,j+1]}^{\perp\!\!\!\perp} + tY_{j+2}\right), X_j \otimes Y_{j+1} \otimes Y_{j+2})\Big\rangle\, dt.$$

$\mathfrak{R}_{Y_j}^{(3,1)}$ is similarly derived. For $j = 1, 2, n-1$ and n, $\mathfrak{R}_{W_j}^{(3,1)}$ is the same as Eq. (35) but with zero in place of non-existing terms. Summing over $j = 1, \ldots, n$,

$$\sum_{j=1}^{n} \mathbb{E}\left[\rho_{r,\phi}^{\delta}(W_{[j,j]}^{\mathsf{C}} + X_j) - \rho_{r,\phi}^{\delta}(W_{[j,j]}^{\mathsf{C}} + Y_j)\right] = \sum_{j=1}^{n} \mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,1)}\right]$$

To further decompose $\mathfrak{R}_{X_j}^{(3,1)}$, we apply the Taylor expansion up to order 4. For example, where $j = 4, \ldots, n-3$,

$$\frac{1}{2}\int_0^1 (1-t)^2 \Big\langle \nabla^3 \varphi_r^{\varepsilon}\left(W_{[j,j]}^{\mathsf{C}} + tX_j\right), X_j^{\otimes 3}\Big\rangle\, dt$$
$$= \frac{1}{6}\Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(W_{[j,j]}^{\mathsf{C}}), X_j^{\otimes 3}\Big\rangle + \frac{1}{6}\int_0^1 (1-t)^3 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(W_{[j,j]}^{\mathsf{C}} + tX_j), X_j^{\otimes 4}\Big\rangle\, dt.$$

Again, to break the dependency between $\nabla^3 \rho_{r,\phi}^{\delta}(X_{[1,n)})$ and $X_n^{\otimes 3}$, we re-apply the Taylor expansion centered at $X_1 + X_{n-1}$.

$$\Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(W_{[j,j]}^{\mathsf{C}}), X_j^{\otimes 3}\Big\rangle$$
$$= \Big\langle \nabla^3 \rho_{r,\phi}^{\delta}(W_{[j,j]}^{\perp\!\!\!\perp}), X_j^{\otimes 3}\Big\rangle$$
$$+ \int_0^1 \Big\langle \nabla^4 \rho_{r,\phi}^{\delta}(W_{[j,j]}^{\perp\!\!\!\perp} + t(X_{j-1}+Y_{j+1})), X_n^{\otimes 3} \otimes (X_{j-1}+Y_{j+1})\Big\rangle\, dt.$$

Repeating this to the other third-order remainder terms, we get

$$\mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,1)}\right]$$
$$= \frac{1}{6}\Big\langle \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\delta}(X_{[1,j-1)} + Y_{(j+1,n]})\right], \mathbb{E}[X_j^{\otimes 3}]\Big\rangle$$
$$+ \frac{1}{2}\Big\langle \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\delta}(X_{[1,j-2)} + Y_{(j+1,n]})\right], \mathbb{E}[X_{j-1}^{\otimes 2} \otimes X_j] + \mathbb{E}[X_{j-1} \otimes X_j^{\otimes 2}]\Big\rangle$$
$$+ \Big\langle \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\delta}(X_{[1,j-3)} + Y_{(j+1,n]})\right], \mathbb{E}\left[X_{j-2} \otimes X_{j-1} \otimes X_j\right]\Big\rangle$$
$$+ \mathbb{E}\left[\mathfrak{R}_{X_j}^{(4,1)}\right],$$

where

$$
\begin{aligned}
\mathfrak{R}_{X_j}^{(4,1)} = &\frac{1}{6}\int_0^1 (1-t)^3\Big\langle \nabla^4\rho_{r,\phi}^{\delta}(W_{[j,j]}^{\mathsf{C}} + tX_j), X_j^{\otimes 4}\Big\rangle\, dt \\
&+ \frac{1}{2}\int_0^1 (1-t)^2\left\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j,j]}^{\perp\!\!\!\perp} + t(X_{j-1}+Y_{j+1})\right), X_j \otimes (X_{j-1}+Y_{j+1})^{\otimes 3}\right\rangle\, dt \\
&+ \frac{1}{2}\int_0^1 (1-t)\left\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j,j]}^{\perp\!\!\!\perp} + t(X_{j-1}+Y_{j+1})\right), X_j^{\otimes 2} \otimes (X_{j-1}+Y_{j+1})^{\otimes 2}\right\rangle\, dt \\
&+ \frac{1}{6}\int_0^1 \left\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j,j]}^{\perp\!\!\!\perp} + t(X_{j-1}+Y_{j+1})\right), X_j^{\otimes 3} \otimes (X_{j-1}+Y_{j+1})\right\rangle\, dt \\
&+ \int_0^1 (1-t)\Big\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j-1,j]}^{\perp\!\!\!\perp} + tX_{j-2}\right), X_j \otimes X_{j-1} \otimes X_{j-2}^{\otimes 2}\Big\rangle\, dt \\
&+ \int_0^1 (1-t)\Big\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j,j+1]}^{\perp\!\!\!\perp} + tY_{j+2}\right), X_j \otimes Y_{j+1} \otimes Y_{j+2}^{\otimes 2}\Big\rangle\, dt \\
&+ \frac{1}{2}\int_0^1 \Big\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j-1,j]}^{\perp\!\!\!\perp} + tX_{j-2}\right), X_j \otimes X_{j-1}^{\otimes 2} \otimes X_{j-2}\Big\rangle\, dt \\
&+ \frac{1}{2}\int_0^1 \Big\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j,j+1]}^{\perp\!\!\!\perp} + tY_{j+2}\right), X_j \otimes Y_{j+1}^{\otimes 2} \otimes Y_{j+2}\Big\rangle\, dt \\
&+ \int_0^1 \left\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j-1,j+1]}^{\perp\!\!\!\perp} + t(X_{j-2}+Y_{j+2})\right), X_{j-1} \otimes X_j \otimes Y_{j+1} \otimes X_{j-2}\right\rangle\, dt \\
&+ \int_0^1 \left\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j-1,j+1]}^{\perp\!\!\!\perp} + t(X_{j-2}+Y_{j+2})\right), X_{j-1} \otimes X_j \otimes Y_{j+1} \otimes Y_{j+2}\right\rangle\, dt \\
&+ \frac{1}{2}\int_0^1 \Big\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j-1,j]}^{\perp\!\!\!\perp} + tX_{j-2}\right), X_j^{\otimes 2} \otimes X_{j-1} \otimes X_{j-2}\Big\rangle\, dt \\
&+ \frac{1}{2}\int_0^1 \Big\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j,j+1]}^{\perp\!\!\!\perp} + tY_{j+2}\right), X_j^{\otimes 2} \otimes Y_{j+1} \otimes Y_{j+2}\Big\rangle\, dt \\
&+ \int_0^1 \Big\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j-2,j]}^{\perp\!\!\!\perp} + tX_{j-3}\right), X_j \otimes X_{j-1} \otimes X_{j-2} \otimes X_{j-3}\Big\rangle\, dt \\
&+ \int_0^1 \Big\langle \nabla^4\rho_{r,\phi}^{\delta}\left(W_{[j,j+2]}^{\perp\!\!\!\perp} + tY_{j+3}\right), X_j \otimes Y_{j+1} \otimes Y_{j+2} \otimes Y_{j+3}\Big\rangle\, dt.
\end{aligned} \tag{35}
$$

$\mathfrak{R}_{Y_j}^{(4,1)}$ is similarly derived. For $j = 1, 2, 3, n-2, n-1$ and n, $\mathfrak{R}_{W_j}^{(4,1)}$ is the same as Eq. (35) but with zero in place of proper terms.

C.3. Second Lindeberg swapping. For each $j = 2, \ldots, n-2$,

$$
\begin{aligned}
&\left\langle \nabla^2\rho_{r,\phi}^{\delta}\left(X_{(1,j)} + X_j + Y_{(j,n-1)}\right), X_n^{\otimes 2}\right\rangle \\
&= \left\langle \nabla^2\rho_{r,\phi}^{\delta}\left(X_{(1,j)} + Y_{(j,n-1)}\right), X_n^{\otimes 2}\right\rangle + \mathfrak{R}_{X_j}^{(3,2,1)}
\end{aligned}
$$

where $\mathfrak{R}_{X_j}^{(3,2,1)} = \int_0^1 \langle \nabla^3 \rho_{r,\phi}^{\delta}\left(X_{(1,j)} + tX_j + Y_{(j,n-1)}\right), X_n^{\otimes 2} \otimes X_j \rangle \, dt$. Furthermore,

$$
\begin{aligned}
&\mathfrak{R}_{X_j}^{(3,2,1)} \\
&= \left\langle \nabla^3 \rho_{r,\phi}^{\delta}\left(X_{(1,j)} + Y_{(j,n-1)}\right), X_n^{\otimes 2} \otimes X_j \right\rangle \\
&\quad + \int_0^1 (1-t) \left\langle \nabla^4 \rho_{r,\phi}^{\delta}\left(X_{(1,j)} + tX_j + Y_{(j,n-1)}\right), X_n^{\otimes 2} \otimes X_j^{\otimes 2} \right\rangle dt \\
&= \left\langle \nabla^3 \rho_{r,\phi}^{\delta}\left(X_{(1,j-1)} + Y_{(j+1,n-1)}\right), X_n^{\otimes 2} \otimes X_j \right\rangle \\
&\quad + \int_0^1 (1-t) \left\langle \nabla^4 \rho_{r,\phi}^{\delta}\left(X_{(1,j)} + tX_j + Y_{(j,n-1)}\right), X_n^{\otimes 2} \otimes X_j^{\otimes 2} \right\rangle dt \\
&\quad + \int_0^1 \left\langle \nabla^4 \rho_{r,\phi}^{\delta}\left(X_{(1,j-1)} + t(X_{j-1} + Y_{j+1}) + Y_{(j+1,n-1)}\right), X_n^{\otimes 2} \otimes X_j \otimes (X_{j-1} + Y_{j+1}) \right\rangle dt.
\end{aligned}
$$

Because $\mathbb{E}[X_n^{\otimes 2} \otimes X_j] = \mathbb{E}[X_n^{\otimes 2} \otimes Y_j] = 0$,

$$
\begin{aligned}
&\left\langle \mathbb{E}\left[\nabla^2 \rho_{r,\phi}^{\delta}(X_{(1,n-1)})\right], \mathbb{E}\left[X_n^{\otimes 2}\right] \right\rangle - \left\langle \mathbb{E}\left[\nabla^2 \rho_{r,\phi}^{\delta}(Y_{(1,n-1)})\right], \mathbb{E}\left[X_n^{\otimes 2}\right] \right\rangle \\
&= \sum_{j=2}^{n-2} \mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,2,1)} - \mathfrak{R}_{Y_j}^{(3,2,1)}\right] = \sum_{j=2}^{n-2} \mathbb{E}\left[\mathfrak{R}_{X_j}^{(4,2,1)} - \mathfrak{R}_{Y_j}^{(4,2,1)}\right],
\end{aligned}
$$

where $\mathfrak{R}_{X_j}^{(4,2,1)}$ is the fourth-order remainder term above. Similarly,

$$
\begin{aligned}
&\left\langle \mathbb{E}\left[\nabla^2 \rho_{r,\phi}^{\delta}(X_{(2,n-1)})\right], \mathbb{E}\left[X_n \otimes X_1\right] \right\rangle - \left\langle \mathbb{E}\left[\nabla^2 \rho_{r,\phi}^{\delta}(Y_{(2,n-1)})\right], \mathbb{E}\left[X_n \otimes X_1\right] \right\rangle \\
&= \sum_{j=3}^{n-2} \mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,2,2)} - \mathfrak{R}_{Y_j}^{(3,2,2)}\right] = \sum_{j=3}^{n-2} \mathbb{E}\left[\mathfrak{R}_{X_j}^{(4,2,2)} - \mathfrak{R}_{Y_j}^{(4,2,2)}\right], \text{ and}
\end{aligned}
$$

$$
\begin{aligned}
&\left\langle \mathbb{E}\left[\nabla^2 \rho_{r,\phi}^{\delta}(X_{(1,n-2)})\right], \mathbb{E}\left[X_n \otimes X_{n-1}\right] \right\rangle - \left\langle \mathbb{E}\left[\nabla^2 \rho_{r,\phi}^{\delta}(Y_{(1,n-2)})\right], \mathbb{E}\left[X_n \otimes X_{n-1}\right] \right\rangle \\
&= \sum_{j=2}^{n-3} \mathbb{E}\left[\mathfrak{R}_{X_j}^{(3,2,3)} - \mathfrak{R}_{Y_j}^{(3,2,3)}\right] = \sum_{j=2}^{n-3} \mathbb{E}\left[\mathfrak{R}_{X_j}^{(4,2,3)} - \mathfrak{R}_{Y_j}^{(4,2,3)}\right].
\end{aligned}
$$

We define $\mathfrak{R}_{W_j}^{(3,2)} = \frac{1}{2}\mathfrak{R}_{W_j}^{(3,2,1)} + \mathfrak{R}_{W_j}^{(3,2,2)} + \mathfrak{R}_{W_j}^{(3,2,3)}$ and $\mathfrak{R}_{W_j}^{(4,2)} = \frac{1}{2}\mathfrak{R}_{W_j}^{(4,2,1)} + \mathfrak{R}_{W_j}^{(4,2,2)} + \mathfrak{R}_{W_j}^{(4,2,3)}$.

C.4. Third Lindeberg swapping. For $j = 3, \ldots, n-1$,

$$
\begin{aligned}
&\left\langle \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\varepsilon}(X_{[1,j-1)} + Y_{(j+1,n)})\right] - \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\varepsilon}(Y_{[1,j-1)} + Y_{(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}] \right\rangle. \\
&= \sum_{k=1}^{j-2} \Big\langle \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\varepsilon}(X_{[1,k)} + X_k + Y_{(k,j-1)\cup(j+1,n)})\right] \\
&\qquad\qquad - \mathbb{E}\left[\nabla^3 \rho_{r,\phi}^{\varepsilon}(X_{[1,k)} + Y_k + Y_{(k,j-1)\cup(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}] \Big\rangle.
\end{aligned}
$$

We first consider the case with $6 \le j \le n-1$. For $3 \le k \le j-4$, the Taylor expansion centered at $X_{[1,k)} + Y_{(k,j-1)\cup(j+1,n)}$ implies

$$
\begin{aligned}
&\Big\langle \mathbb{E}\left[\nabla^3 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + X_k + Y_{(k,j-1)\cup(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}] \Big\rangle \\
&= \Big\langle \mathbb{E}[\nabla^3 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + Y_{(k,j-1)\cup(j+1,n)})], \mathbb{E}[X_j^{\otimes 3}] \Big\rangle \\
&\quad + \mathbb{E}\left[\Big\langle \nabla^4 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + Y_{(k,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes X_k \Big\rangle\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\Big\langle \nabla^5 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + Y_{(k,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes X_k^{\otimes 2} \Big\rangle\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\int_0^1 (1-t)^2 \Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + tX_k + Y_{(k,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes X_k^{\otimes 3} \Big\rangle \, dt\right].
\end{aligned}
$$

For the inner product terms with dependent factors, we repeat the Taylor expansion:

1.
$$
\begin{aligned}
&\mathbb{E}\left[\Big\langle \nabla^4 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + Y_{(k,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes X_k \Big\rangle\right] \\
&= \mathbb{E}\left[\Big\langle \nabla^5 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-1)} + Y_{(k+1,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes X_k \otimes (X_{k-1} + Y_{k+1}) \Big\rangle\right] \\
&\quad + \mathbb{E}\Big[\int_0^1 (1-t) \Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-1)} + t(X_{k-1} + Y_{k+1}) + Y_{(k+1,j-1)\cup(j+1,n)}), \\
&\qquad\qquad \mathbb{E}[X_j^{\otimes 3}] \otimes X_k \otimes (X_{k-1} + Y_{k+1})^{\otimes 2} \Big\rangle dt\Big],
\end{aligned}
$$

2.
$$
\begin{aligned}
&\mathbb{E}\left[\Big\langle \nabla^5 \rho^{\varepsilon}_{r,\phi}(X_{[1,k)} + Y_{(k,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes X_k^{\otimes 2} \Big\rangle\right] \\
&= \Big\langle \mathbb{E}\left[\nabla^5 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-1)} + Y_{(k+1,j-1)\cup(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}] \otimes \mathbb{E}\left[X_k^{\otimes 2}\right] \Big\rangle \\
&\quad + \mathbb{E}\Big[\int_0^1 \Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-1)} + t(X_{k-1} + Y_{k+1}) + Y_{(k+1,j-1)\cup(j+1,n)}), \\
&\qquad\qquad \mathbb{E}[X_j^{\otimes 3}] \otimes X_k^{\otimes 2} \otimes (X_{k-1} + Y_{k+1}) \Big\rangle \, dt\Big],
\end{aligned}
$$

3.
$$
\begin{aligned}
&\mathbb{E}\left[\Big\langle \nabla^5 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-1)} + Y_{(k+1,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes X_k \otimes (X_{k-1} + Y_{k-1}) \Big\rangle\right] \\
&= \Big\langle \mathbb{E}[\nabla^5 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-2)} + Y_{(k+1,j-1)\cup(j+1,n)})], \mathbb{E}[X_j^{\otimes 3}] \otimes \mathbb{E}[X_k \otimes X_{k-1}] \Big\rangle \\
&\quad + \mathbb{E}\Big[\int_0^1 \Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-2)} + tX_{k-2} + Y_{(k+1,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes X_k \otimes X_{k-1} \otimes X_{k-2} \Big\rangle \, dt\Big] \\
&\quad + \mathbb{E}\Big[\int_0^1 \Big\langle \nabla^6 \rho^{\varepsilon}_{r,\phi}(X_{[1,k-1)} + tY_{k+2} + Y_{(k+2,j-1)\cup(j+1,n)}), \mathbb{E}[X_j^{\otimes 3}] \otimes X_k \otimes Y_{k+1} \otimes Y_{k+2} \Big\rangle \, dt\Big].
\end{aligned}
$$

For $\left\langle \mathbb{E}\left[\nabla^3\rho^\varepsilon_{r,\phi}(X_{[1,k)}+Y_k+Y_{(k,j-1)\cup(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}]\right\rangle$, the calculation is the same but with Y_k in place of X_k. By the second moment matching,

$$\begin{aligned}
&\Big\langle \mathbb{E}\left[\nabla^3\rho^\varepsilon_{r,\phi}(X_{[1,k)}+X_k+Y_{(k,j-1)\cup(j+1,n)})\right] \\
&\qquad\qquad -\mathbb{E}\left[\nabla^3\rho^\varepsilon_{r,\phi}(X_{[1,k)}+Y_k+Y_{(k,j-1)\cup(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}]\Big\rangle \\
&= \Big\langle \mathbb{E}[\nabla^5\rho^\varepsilon_{r,\phi}(X_{[1,k-2)}+Y_{(k+1,j-1)\cup(j+1,n)})], \mathbb{E}[X_j^{\otimes 3}]\otimes\mathbb{E}[X_k\otimes X_{k-1}]\Big\rangle \\
&\quad - \Big\langle \mathbb{E}[\nabla^5\rho^\varepsilon_{r,\phi}(X_{[1,k-1)}+Y_{(k+2,j-1)\cup(j+1,n)})], \mathbb{E}[X_j^{\otimes 3}]\otimes\mathbb{E}[Y_k\otimes Y_{k+1}]\Big\rangle \\
&\quad + \mathbb{E}\left[\mathfrak{R}^{(6,1)}_{X_j,X_k} - \mathfrak{R}^{(6,1)}_{X_j,Y_k}\right],
\end{aligned}$$

where for $W_k = X_k$ or Y_k,

$$\begin{aligned}
&\mathfrak{R}^{(6,1)}_{X_j,W_k} \\
&= \frac{1}{2}\int_0^1 (1-t)^2\Big\langle \nabla^6\rho^\varepsilon_{r,\phi}(X_{[1,k)}+tW_k+Y_{(k,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k^{\otimes 3}\Big\rangle\, dt \\
&\quad + \int_0^1 (1-t)\Big\langle \nabla^6\rho^\varepsilon_{r,\phi}(X_{[1,k-1)}+t(X_{k-1}+Y_{k+1})+Y_{(k+1,j-1)\cup(j+1,n)}), \\
&\qquad\qquad\qquad X_j^{\otimes 3}\otimes W_k\otimes(X_{k-1}+Y_{k+1})^{\otimes 2}\Big\rangle dt \\
&\quad + \frac{1}{2}\int_0^1 \Big\langle \nabla^6\rho^\varepsilon_{r,\phi}(X_{[1,k-1)}+t(X_{k-1}+Y_{k+1})+Y_{(k+1,j-1)\cup(j+1,n)}), \\
&\qquad\qquad\qquad X_j^{\otimes 3}\otimes W_k^{\otimes 2}\otimes(X_{k-1}+Y_{k+1})\Big\rangle\, dt \\
&\quad + \int_0^1 \Big\langle \nabla^6\rho^\varepsilon_{r,\phi}(X_{[1,k-2)}+tX_{k-2}+Y_{(k+1,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k\otimes X_{k-1}\otimes X_{k-2}\Big\rangle\, dt \\
&\quad + \int_0^1 \Big\langle \nabla^6\rho^\varepsilon_{r,\phi}(X_{[1,k-1)}+tY_{k+2}+Y_{(k+2,j-1)\cup(j+1,n)}), X_j^{\otimes 3}\otimes W_k\otimes Y_{k+1}\otimes Y_{k+2}\Big\rangle\, dt.
\end{aligned} \tag{36}$$

For $k=1$,

$$\begin{aligned}
&\Big\langle \mathbb{E}\left[\nabla^3\rho^\varepsilon_{r,\phi}(X_1+Y_{(1,j-1)\cup(j+1,n)})\right] - \mathbb{E}\left[\nabla^3\rho^\varepsilon_{r,\phi}(Y_1+Y_{(1,j-1)\cup(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}]\Big\rangle \\
&= -\Big\langle \mathbb{E}[\nabla^5\rho^\varepsilon_{r,\phi}(Y_{(3,j-1)\cup(j+1,n)})], \mathbb{E}[X_j^{\otimes 3}]\otimes\mathbb{E}[Y_1\otimes Y_2]\Big\rangle \\
&\quad + \mathbb{E}\left[\mathfrak{R}^{(6,1)}_{X_j,X_1} - \mathfrak{R}^{(6,1)}_{X_j,Y_1}\right],
\end{aligned}$$

where $\mathfrak{R}^{(6,1)}_{X_j,W_1}$ is the same as Eq. (36) but with Y_n and Y_{n-1} in place of X_{k-1} and X_{k-2}, respectively. For $k=2$,

$$\Big\langle \mathbb{E}\left[\nabla^3 \rho^{\varepsilon}_{r,\phi}(X_1+X_2+Y_{(2,j-1)\cup(j+1,n)})\right] - \mathbb{E}\left[\nabla^3 \rho^{\varepsilon}_{r,\phi}(X_1+Y_2+Y_{(2,j-1)\cup(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}]\Big\rangle$$
$$= \Big\langle \mathbb{E}[\nabla^5 \rho^{\varepsilon}_{r,\phi}(Y_{(3,j-1)\cup(j+1,n)})], \mathbb{E}[X_j^{\otimes 3}]\otimes \mathbb{E}[X_2\otimes X_1]\Big\rangle$$
$$- \Big\langle \mathbb{E}[\nabla^5 \rho^{\varepsilon}_{r,\phi}(Y_{(4,j-1)\cup(j+1,n)})], \mathbb{E}[X_j^{\otimes 3}]\otimes \mathbb{E}[Y_2\otimes Y_3]\Big\rangle$$
$$+ \mathbb{E}\left[\mathfrak{R}^{(6,1)}_{X_j,X_2} - \mathfrak{R}^{(6,1)}_{X_j,Y_2}\right],$$

where $\mathfrak{R}^{(6,1)}_{X_j,W_2}$ is the same as Eq. (36) but with Y_n in place of X_{k-2}. For $k=j-3$,

$$\Big\langle \mathbb{E}\left[\nabla^3 \rho^{\varepsilon}_{r,\phi}(X_{[1,j-3)}+X_{j-3}+Y_{\{j-2\}\cup(j+1,n)})\right]$$
$$-\mathbb{E}\left[\nabla^3 \rho^{\varepsilon}_{r,\phi}(X_{[1,j-3)}+Y_{j-3}+Y_{\{j-2\}\cup(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}]\Big\rangle$$
$$= \Big\langle \mathbb{E}[\nabla^5 \rho^{\varepsilon}_{r,\phi}(X_{[1,j-5)}+Y_{(j+1,n)})], \mathbb{E}[X_j^{\otimes 3}]\otimes \mathbb{E}[X_{j-3}\otimes X_{j-4}]\Big\rangle$$
$$- \Big\langle \mathbb{E}[\nabla^5 \rho^{\varepsilon}_{r,\phi}(X_{[1,j-4)}+Y_{(j+1,n)})], \mathbb{E}[X_j^{\otimes 3}]\otimes \mathbb{E}[Y_{j-3}\otimes Y_{j-2}]\Big\rangle$$
$$+ \mathbb{E}\left[\mathfrak{R}^{(6,1)}_{X_j,X_{j-3}} - \mathfrak{R}^{(6,1)}_{X_j,Y_{j-3}}\right],$$

where $\mathfrak{R}^{(6,1)}_{X_j,W_2}$ is the same as Eq. (36) but with 0 in place of Y_{k+2}. For $k=j-2$,

$$\Big\langle \mathbb{E}\left[\nabla^3 \rho^{\varepsilon}_{r,\phi}(X_{[1,j-2)}+X_{j-2}+Y_{(j+1,n)})\right] - \mathbb{E}\left[\nabla^3 \rho^{\varepsilon}_{r,\phi}(X_{[1,j-2)}+Y_{j-2}+Y_{(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}]\Big\rangle$$
$$= \Big\langle \mathbb{E}[\nabla^5 \rho^{\varepsilon}_{r,\phi}(X_{[1,j-4)}+Y_{(j+1,n)})], \mathbb{E}[X_j^{\otimes 3}]\otimes \mathbb{E}[X_{j-2}\otimes X_{j-3}]\Big\rangle$$
$$+ \mathbb{E}\left[\mathfrak{R}^{(6,1)}_{X_j,X_{j-2}} - \mathfrak{R}^{(6,1)}_{X_j,Y_{j-2}}\right],$$

where $\mathfrak{R}^{(6,1)}_{X_j,W_2}$ is the same as Eq. (36) but with 0 in place of Y_{k+1} and Y_{k+2}. By the second moment matching,

$$\Big\langle \mathbb{E}\left[\nabla^3 \rho^{\varepsilon}_{r,\phi}(X_{[1,j-1)}+Y_{(j+1,n)}) - \nabla^3 \rho^{\varepsilon}_{r,\phi}(Y_{[1,j-1)}+Y_{(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}]\Big\rangle$$
$$= \sum_{k=1}^{j-2} \mathbb{E}\left[\mathfrak{R}^{(6,1)}_{X_j,X_k} - \mathfrak{R}^{(6,1)}_{X_j,Y_k}\right].$$

This is also the same for $j\in[1,5]$, where some terms in $\mathfrak{R}^{(6,1)}_{X_j,W_k}$ are zero when appropriate. In sum,

$$\sum_{j=1}^{n}\Big\langle \mathbb{E}\left[\nabla^3 \rho^{\delta}_{r,\phi}(X_{[1,j-1)}+Y_{(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}]\Big\rangle$$
$$= \sum_{j=1}^{n}\Big\langle \mathbb{E}\left[\nabla^3 \rho^{\delta}_{r,\phi}(Y_{[1,j-1)}+Y_{(j+1,n)})\right], \mathbb{E}[X_j^{\otimes 3}]\Big\rangle + \sum_{j=3}^{n}\sum_{k=1}^{j-2} \mathbb{E}\left[\mathfrak{R}^{(6,1)}_{X_j,X_k} - \mathfrak{R}^{(6,1)}_{X_j,Y_k}\right].$$

Similarly,

$$\sum_{j=1}^{n}\Big\langle \mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\delta}(X_{(0\vee(j+2-n),j-2)}+Y_{(j+2,n)})\right], \mathbb{E}[X_j^{\otimes 2}\otimes X_{j+1}]+\mathbb{E}[X_j\otimes X_{j+1}^{\otimes 2}]\Big\rangle$$

$$=\sum_{j=1}^{n}\Big\langle \mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\delta}(X_{(0\vee(j+2-n),j-2)}+Y_{(j+2,n)})\right], \mathbb{E}[X_j^{\otimes 2}\otimes X_{j+1}]+\mathbb{E}[X_j\otimes X_{j+1}^{\otimes 2}]\Big\rangle$$

$$+\sum_{j=4}^{n}\sum_{k=1}^{j-3}\mathbb{E}\left[\mathfrak{R}_{X_j,X_k}^{(6,2)}-\mathfrak{R}_{X_j,Y_k}^{(6,2)}\right],\ \text{ and}$$

$$\sum_{j=1}^{n}\Big\langle \mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\delta}(X_{(0\vee(j+3-n),j-3)}+Y_{(j+3,n)})\right], \mathbb{E}\left[X_j\otimes X_{j+1}\otimes X_{j+2}\right]\Big\rangle$$

$$=\sum_{j=1}^{n}\Big\langle \mathbb{E}\left[\nabla^3\rho_{r,\phi}^{\delta}(Y_{(0\vee(j+3-n),j-3)}+Y_{(j+3,n)})\right], \mathbb{E}\left[X_j\otimes X_{j+1}\otimes X_{j+2}\right]\Big\rangle$$

$$+\sum_{j=5}^{n}\sum_{k=1}^{j-4}\mathbb{E}\left[\mathfrak{R}_{X_j,X_k}^{(6,3)}-\mathfrak{R}_{X_j,Y_k}^{(6,3)}\right],$$

where $\mathfrak{R}_{X_j,W_k}^{(6,2)}$ and $\mathfrak{R}_{X_j,W_k}^{(6,3)}$ are similarly derived as Eq. (36).